Statistical methods for large ensemble of super-resolution stochastic single particle trajectories

N. Hoze [1] and D. Holcman [2]

Abstract

Following recent progresses in super-resolution microscopy obtained in the last decade, massive amount of redundant single stochastic trajectories are now available for statistical analysis. Flows of trajectories of molecules or proteins are sampling the cell membrane or its interior at a very high time and space resolution. Several statistical analysis were developed to extract information contained in these data, such as the biophysical parameters of the underlying stochastic motion to reveal the cellular organization. These trajectories can further reveal hidden subcellular organization. We present here the statistical analysis of these trajectories based on the classical Langevin equation, which serves as a model of trajectories. Parametric and non-parametric estimators are constructed by discretizing the stochastic equations and they allow recovering tethering forces, diffusion tensor or membrane organization from measured trajectories, that differ from physical ones by a localization noise. Modeling, data analysis and automatic detection algorithms serve extracting novel biophysical features such as potential wells and other sub-structures, such as rings at an unprecedented spatiotemporal resolution. It is also possible to reconstruct the surface membrane of a biological cell from the statistics of projected random trajectories.

Introduction

Recent advances in super-resolution microscopy [54, 80, 29, 71] associated with single particle tracking (SPT) [73, 75] have allowed routine collections of tens of thousands of short and long trajectories. These large datasets offer new challenges in analysis, interpretation, search and reconstruction of hidden features, especially in the context of cell biology probed at the molecular resolution. These trajectories explore part of a cell, either in dimension 2 [26, 62] or 3 at the shortest possible resolution, and their paths are critically influenced by local organization, as emphasized in neuronal cells [76, 68]. Thus, a large research effort in that direction has been based on the hypothesis that it would be possible to reconstruct some of the biological nano- to micro- structures from these trajectories. It should also be possible to unravel their underlying physical processes.

In the past recent years, these challenges of cell membrane reconstruction, structural identification and membrane organization have been addressed following the development of novel statistical methods based on stochastic models. These models are the Langevin equation, its Smoluchowski limit and associated models that account for additional localization point identification noise. We present here the development of these statistical approaches

[1]Institut für Integrative Biologie, ETH, Universitätstrasse 16, 8092 Zürich, Switzerland.
[2]DAMTP, University of Cambridge University, CB30DS, UK and Applied Mathematics and Computational Biology, Ecole Normale Supérieure, 46 rue d'Ulm 75005 Paris, France.

that are based on stochastic models, the deconvolution procedure and the numerical simulations used to extract biophysical parameters from single particle trajectories data. Several empirical estimators have been proposed to recover the local diffusion coefficient, vector field and even organized patterns in the drift, such as potential wells.

For studying long trajectories, the most common procedure consists in estimating the mean-square displacement (MSD) of the particle [19]. The MSD is proportional to time and allows in theory estimating the diffusion coefficient. It has been widely used in early applications of long but not necessarily redundant single particle trajectories in a biological context [64, 74]. However, the MSD applied to long trajectories suffers from several issues. First, it is not precise in part because the measured points are correlated [87]. Second, it cannot be used to compute any physical diffusion coefficient when trajectories consists of switching episodes for example alternating between free and confined diffusion, where other algorithm procedures have been developed [27]. At low spatiotemporal resolution of the observed trajectories, the MSD behaves sublinearly with time, a process known as anomalous diffusion [84, 56, 58], which is due in part to the averaging of the different phases of the particle motion. To overcome these issues and distinguish between different types of motion, a stochastic model of the acquired data is needed.

This manuscript is organized as follow: we first describe single stochastic particle trajectories recorded in the context of cell biology in two dimensional membrane and in the cell cytoplasm. In the second part, we introduce Langevin stochastic processes used as a model to interpret physical trajectories. These equations are a key step for the construction of parametric and non-parametric empirical estimators to extract forces, diffusion tensor or membrane organization from measured trajectories. The third part explains how suborganization such as potential wells are reconstructed from trajectories. The fourth section shows how the error in the position localization affects the first and second moments computed from SPT trajectoriess. In section five, we present a different type of SPT trajectories analysis used to recover tethering forces applied on particle belonging to a polymer chain (model of a DNA chromatin locus). In the last section, we show how to reconstruct a surface from the projections of many SPT trajectories. This presentation is intended for statisticians, applied mathematicians, theoretical physicists and more general anybody interested in novel methods based on stochastic processes, recently used to extract information from super-resolution data below the diffraction limit.

1 Presentation of single particle tracking trajectory data in cellular biology

Single particle trajectories can be short and long often acquired from digital images [32, 54]. Before explaining how they are used to investigate the local structure of the environment they sample, we shall present the trajectories in the context of statistical analysis of these data. Indeed, the goal of building a statistical ensemble from these data is to observe local physical properties of the particles, such as velocity, diffusion, confinement or attracting forces reflecting the interactions of the particles with their environment. It might even now

be possible to use modeling to construct from diffusion coefficient (or tensor) the confinement or local density of obstacles reflecting the presence of biological objects of different sizes.

In the context of cell biology single particle trajectories are acquired on tagged molecules or proteins that are not immobile but are usually moving inside the cell either on the two dimensional surface (membrane) or inside the three dimensional cytosol. In all cases, they can be sampled at a nanometer scale of molecular motion. It is possible to observe proteins of interest because they have been genetically or directly associated with a fluorescent molecule that can be tracked under the microscope. The luminescent property of the dye defines their stability and thus the time they will remain bright before they bleach. Collecting these data requires to define a spatial precision, which is usually limited by the diffraction limit [69] and the sampling time acquisition.

Biophysical and molecular processes live at the nanometer scale, that was unaccessible due to the limitation of the spatiotemporal resolution, and the image of a infinitely small particle could not be resolved below hundreds of nanometers (Abbe's diffraction limit [52]). Detecting objects under the diffraction limit is now routinely achieved using different microscopy techniques such as PALM [7], STED [31] or STORM [69]: these methods rely on stochastic activation of isolated particles (molecules of interest) for a short time. Not all particles emit light at the same time, thus it becomes possible to distinguish their position. For all microscopy approaches, the localization method consists in fitting a Gaussian to the illumination profile of any localized point. The standard deviation of the fitted Gaussian, which determines the precision of localization, is reduced when the emitting particle is of small size, the number of photons received during the acquisition is large and the particle motion is limited.

SPT allowed observing moving particles at the micrometer scale, to investigate cell membrane organization [73, 49, 64, 18, 74, 32, 25], but also the cell nucleus dynamics and mRNA production [22, 81]. Due to the constant improvement of the instrumentation, the spatial resolution is continuously decreasing, reaching today values of approximately 20 nm, while the acquisition time step is usually in the range of 10 to 50 ms to capture the shortest events in live tissues. Once points are acquired, the second step is to reconstruct trajectories. This is done by tracking algorithms that allow connecting the dots of acquired points [14, 75].

We now present trajectories of different moving molecules obtained from sptPALM and other techniques (Fig. 1): First, proteins trafficking on the surface of neuronal cells (transmembrane proteins) [26, 62, 53] can now be recorded at a spatial resolution of approximately 30 nm and a time resolution of 20 ms in sub-surface compartments, such as the pre-synaptic membrane [76, 68]. More recently, sptPALM was used to detect the dynamical organization of membrane proteins in plant cells [36, 47], or events of DNA binding by transcription factors in mammalian nucleus [24] and by CRISPR-Cas9 [46]. The physical interpretation of the trajectories depends on the quality of the upstream data acquisition. Although super-resolution image acquisition and particle tracking are crucial to guarantee high quality data, we leave these topics outside as there have been already discussed in [15, 52].
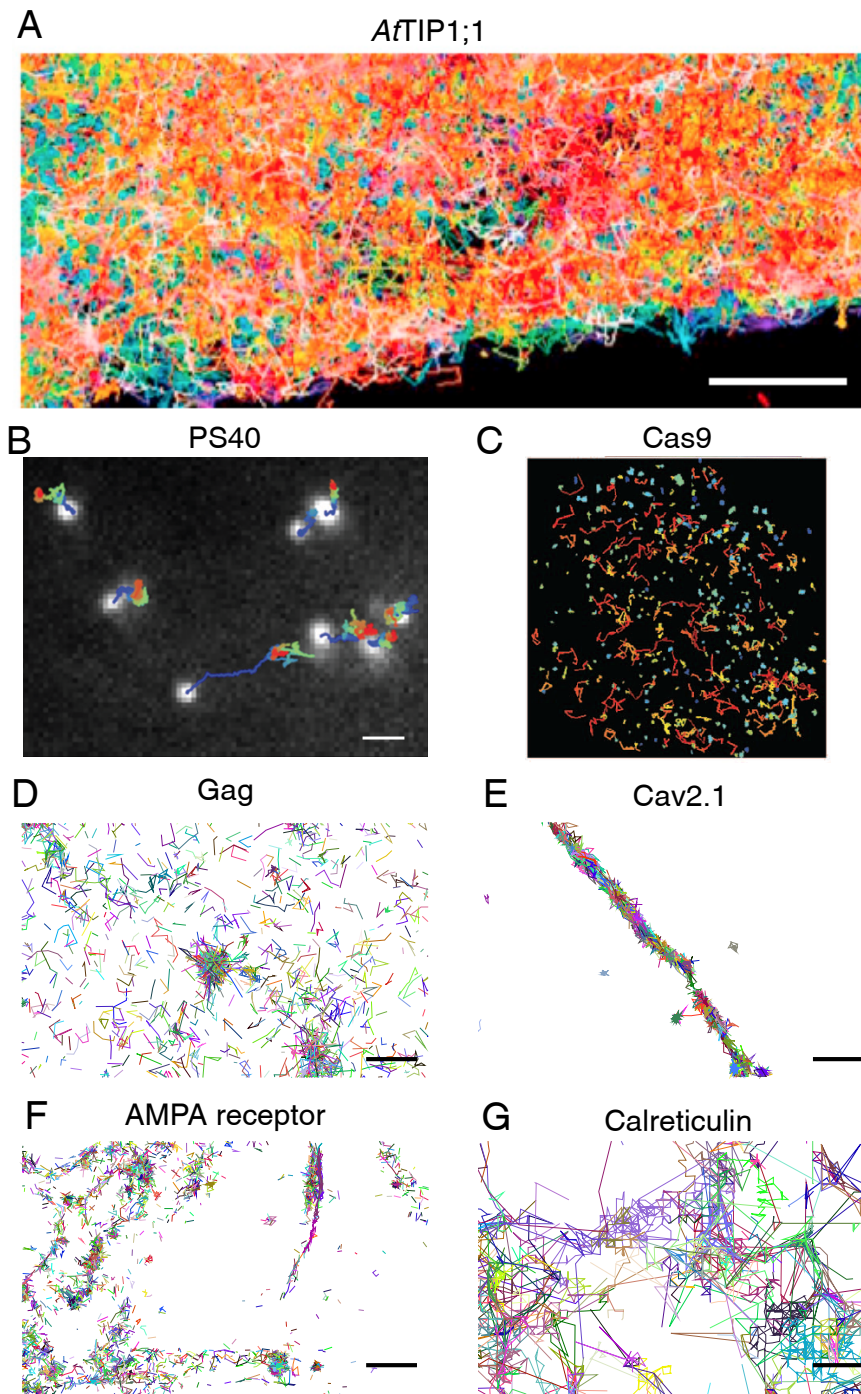
3

Figure 1: **Examples of superresolution trajectories.** (A) Superresolution trajectories of the tonoplast marker AtTIP1;1 in root epidermal cells (adapted from [36]). (B) Interaction of the nanoparticle PS40 with clathrin-coated pits on the plasma membrane of COS-7 cells (adapted from [51]). (C) Dynamics of Cas9 molecules transfected with short interspersed nuclear elements (SINE) of the B2 type in 3T3 cells (adapted from [46]). (D) HIV structural protein Gag at the surface of cellular plasma membrane. (E) Calcium channel Cav2.1 in the presynaptic density. (F) AMPA receptors in hippocampal neurons. (G) Calcium storage protein Calreticulin in the endoplasmic reticulum. Scale bars: A : 5 μm, B, D–G: 1 μm.

2 Stochastic equation as a model of motion

2.1 Langevin and Smoluchowski equations

Langevin's equation [19, 8, 50] describe a stochastic particle driven by a random force ξ and a field of force (e.g., electrostatic, mechanical, etc...). With an external force $F(x, t)$ [13], it is written as

$$m\frac{d^2x}{dt^2} + \Gamma\frac{dx}{dt} - F(x, t) = \dot{\eta}, \tag{1}$$

where m is the mass of the particle and $\Gamma = 6\pi a\rho$ is the friction coefficient of a diffusing particle, ρ the viscosity and η the δ-correlated Gaussian white noise [19, 8, 50]. The force can derive from a potential, $F(x, t) = -U'(x)$ and in that case, with notation $\varepsilon = k_B T$, which represents the energy (k_B is the Boltzmann's constant and T the temperature), the Langevin's equation is transformed into the two dimensional stochastic system

$$\frac{dx}{dt} = v \tag{2}$$

$$\frac{dv}{dt} = -\gamma v - U'(x) + \sqrt{2\varepsilon\gamma}\,\dot{\eta}, \tag{3}$$

where $\gamma = \Gamma/m$ is the dynamical friction coefficient per unit mass.

Langevin's equation is used to describe trajectories where inertia or acceleration matters. For example at very short timescales, when a molecule unbinds from a binding site, it escapes the potential well and the inertia term allows the particles to move away from the attractor and thus prevents immediate rebinding that could plague numerical simulations. In the large friction limit $\gamma \to \infty$ the trajectories $x(t)$ of the Langevin equation 1 converges in probability to these of the Smoluchowski equation [77, 5]

$$\gamma\frac{dx}{dt} + U'(x) = \sqrt{2\varepsilon\gamma}\,\dot{\eta}. \tag{4}$$

Equation 4 is written equivalently as

$$\gamma dx + U'(x)dt = \sqrt{2\varepsilon\gamma}d\eta. \tag{5}$$

2.2 Physical model equations

For a timescale much longer than elementary molecular events, the position of tracked particles is described by the overdamped limit of the Saffman-Delbrück-Langevin model [70, 50]. The model is described in the previous subsection and assumes that the diffusion of a protein or a particle embedded in a membrane surface is generated by a diffusion coefficient D and a field of force $F(X, t)$, according to the overdamped Langevin equation

$$\dot{X} = \frac{F(X(t), t)}{\gamma} + \sqrt{2D}\,\dot{W}, \tag{6}$$

5

where W is a Gaussian white noise and γ is the dynamical viscosity [77]. The source of the driving noise is the thermal agitation of the ambient lipid and membrane molecules. However, due to the acquisition timescale of empirical recorded trajectories, which is too low to follow the thermal fluctuations, rapid events are not resolved in the data, and at this coarser spatiotemporal scale, the motion is described by an effective stochastic equation [37, 40]

$$\dot{X} = a(X)dt + \sqrt{2}B(X)\dot{W}, \tag{7}$$

where $a(X)$ is the drift field and $B(X)$ the diffusion matrix (see Fig. 2). The effective diffusion tensor is given by $D(X) = \frac{1}{2}B(X)B^T(X)$ ($.^T$ denotes the transposition) [78, 77]. The observed effective diffusion tensor is not necessarily isotropic and can be state-dependent, whereas the friction coefficient γ in 6 remains constant and the microscopic diffusion coefficient (or tensor) may remain isotropic.
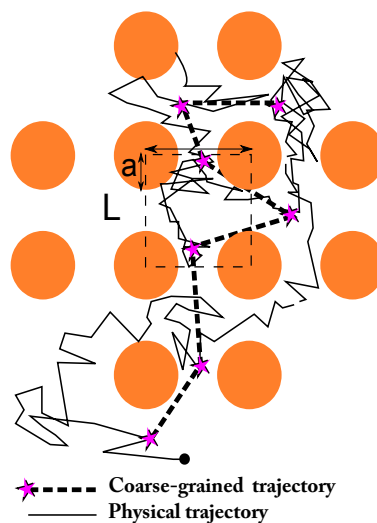

Figure 2: **Scheme of a coarse-grained trajectories from an effective stochastic equation**. A physical trajectory is generated by the overdamped Langevin equation (Eq. 6) while the coarse-grained trajectory is sampled at fixed time interval Δt (indicated by purple stars). The presented obstacles are not directly visible.

2.3 Empirical estimation of the drift and diffusion tensor of a general stochastic process

In this section, we present the construction of empirical estimators that serve to recover physical properties from parametric [37, 41, 87] and non-parametric statistics [4]. Retrieving statistical parameters of a diffusion process from one-dimensional time series statistics have been studied using Bayesian inference in [30, 65]. The approach we present here is based on the direct asymptotic of stochastic equations to construct direct empirical estimators for recovering drift and diffusion tensor. There are derived from discretizing the stochastic equation 7 [44, 78, 79, 33, 34].

6

We note that the models and the analysis presented above assume that processes are stationary, so that the statistical properties of trajectories do not change over time. In practice, this assumption is satisfied when trajectories are acquired for less than a minute, where only few slow changes may occur on the surface of a neuron for example. Time-lapse analysis [37] with a delay of 15 minutes between acquisitions has indeed revealed slow changes over time of biological significance.

The coarse-grained model 7 is recovered from the conditional moments of the trajectory increments $\Delta X = X(t + \Delta t) - X(t)$ (see [77, 23, 82]),

$$a(x) = \lim_{\Delta t \to 0} \frac{\mathbb{E}[\Delta X(t) \mid X(t) = x]}{\Delta t}, \quad D(x) = \lim_{\Delta t \to 0} \frac{\mathbb{E}[\Delta X(t)^T \Delta X(t) \mid X(t) = x]}{2\Delta t}. \tag{8}$$

Here the notation $\mathbb{E}[\cdot \mid X(t) = x]$ means averaging over all trajectories that are at point x at time t. Indeed, the coefficients of eq. 6 can be statistically estimated at each point of the membrane (or rather in the focal plane) from an infinitely large sample of its trajectories in the neighborhood of the point X at time t. In practice, the expectations in 8 are estimated by finite sample averages and Δt is the time-resolution of the recording of the trajectories, as described in [37]. Formulas 8 are approximated in [37] at the time step $\Delta t = 50$ ms, where 200 points falling in any bin is usually enough for the estimation.

To estimate the local drift and diffusion coefficients, the trajectories are first grouped within a small neighbourhood. The field of observation is partitioned into square bins $S(x_k, r)$ of side r and centre x_k (Fig. 3A), and the local drift and diffusion are estimated for each of the square. Considering a sample of N_t trajectories $\{x^i(t_1), \ldots, x^i(t_{N_s})\}$, where t_j are the sampling times, the discretization of equation 8 for the drift $a(x_k) = (a_x(x_k), a_y(x_k))$ at position x_k is

$$
\begin{aligned}
a_x(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_t} \sum_{i=0, \tilde{x}_i^j \in S(x_k, r)}^{N_s - 1} \left(\frac{x_{i+1}^j - x_i^j}{\Delta t} \right) \\
\\
a_y(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_t} \sum_{i=0, \tilde{x}_i^j \in S(x_k, r)}^{N_s - 1} \left(\frac{y_{i+1}^j - y_i^j}{\Delta t} \right),
\end{aligned}
\tag{9}
$$

where N_k is the number of points of trajectory that fall in the square $S(x_k, r)$. Similarly, the components of the effective diffusion tensor $D(x_k)$ are approximated by the empirical sums

$$
\begin{aligned}
D_{xx}(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_t} \sum_{i=0, x_i \in S(x_k, r)}^{N_s - 1} \frac{(x_{i+1}^j - x_i^j)^2}{2\Delta t} \\
D_{yy}(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_t} \sum_{i=0, x_i \in S(x_k, r)}^{N_s - 1} \frac{(y_{i+1}^j - y_i^j)^2}{2\Delta t} \\
D_{xy}(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_t} \sum_{i=0, x_i \in S(x_k, r)}^{N_s - 1} \frac{(x_{i+1}^j - x_i^j)(y_{i+1}^j - y_i^j)}{2\Delta t}.
\end{aligned}
\tag{10}
$$

7

The moment estimation 9 and 10 requires a large number of trajectories passing through each point of the surface, which fits precisely to the massive data generated by the sptPALM technique on biological samples. Indeed, the exact inversion formula eq8 demands in theory an infinite number of trajectories passing through any point x of interest. In practice, the recovery of the drift and diffusion tensor is obtained after a region is subdivided by a square grid of radius r (of the order of 50 to 100 nm). Estimations of the error term in the drift compared to the diffusion coefficient suggest a minimum of 200 points per bin (Supp. Information in [37]).

We illustrate this method on trajectories of vesicular stomatitis virus G protein (VSVG) (Fig. 3B). The trajectories are separated on squares of size $r = 100$ nm. The empirical estimator 9 is applied to the trajectories to compute the drift from the trajectories (Fig. 3C). The average diffusion coefficient $D = \frac{1}{2}(D_{xx} + D_{yy})$ is estimated using eqs. 10 (Fig. 3D).
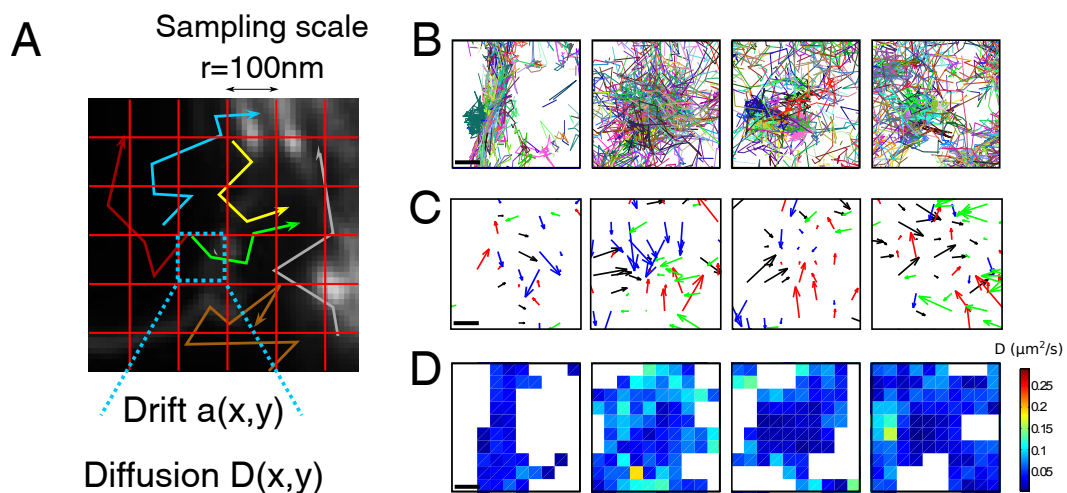

Figure 3: (A) Discretization of the image into pixels of size r, that is used preliminarily to the estimation of the local drift and diffusion (eqs. 9-10). (B-D) Application of the estimators to VSVG superresolution trajectories. (B) Four samples of sptPALM trajectories (n=30,000) of VSVG proteins. (C) Drift field estimation in the four squares. The arrows are colored depending on the drift orientation. Square side is $r = 100$nm. (D) Map of the diffusion coefficient. Scale bars: 200 nm.

3 Interpretation of two-dimensional regions of high trajectories density as potential wells

The drift field $a(\boldsymbol{x})$ in equation 7 represents a force that acts on the diffusing particle, regardless of the existence or not of a potential well [38]. In the case where $\boldsymbol{D}(\boldsymbol{x})$ is locally constant and the coarse-grained drift field $\boldsymbol{b}(\boldsymbol{x})$ is a gradient of a potential

$$a(\boldsymbol{x}) = -\nabla U(\boldsymbol{x}), \tag{11}$$

then the density of particles represents locally the Boltzmann density $e^{-U(\boldsymbol{x})/D}$ [34]. The force field can form potential wells, generically approximated locally as a paraboloid with an elliptic base. It remains a difficult question to extract automatically the axis, the center and the boundary of the elliptic base of the well. Once they are known, within the analytical representation $U(\boldsymbol{x}) = A\left((\frac{x}{r_x})^2 + (\frac{y}{r_y})^2\right) + O(x,y)^2$, the constants A, r_x, r_y are three parameters to be determined.

We present methods to detect empirical potential wells reconstructed from data. In particular, several parameters influence the detections such as the pixel resolution and the minimum number of sampled points of trajectory in each pixel. Another interesting feature is the time evolution of these structures.

3.1 Vector fields, attractors and characteristic index

A determinist vector field is most of the time a local non-zero flow. However, there are rare occasions where it vanishes at isolated points, and each region near these points is of particular interest. The situation is however more complicated when the vector field has random components, especially near critical points, because the noise dominates. Yet, to determine whether a vector field of the extracted drift component contains structures such as basin of attractions of point attractors (which is probably the simplest detectable structure of a generic dynamical system), a robust characterization of the field is needed to differentiate random distributions from those that may contain a physical significance. Due to the time and space resolution limitation, in the context of SPT, small attractors, associated with short-range interactions less than few tens of nano-meters, below the diffraction limit, cannot be resolved.

To characterise a two-dimensional attractor, generated by a gradient field, the first step it to recover the elliptic paraboloid, which is second order Taylor approximation of the field at the non degenerated critical point. The elliptic domain is the basin of attraction, while its depth is a measure of the force field strength. For a circular potential well of depth A and radius r, centered at a point (x_0, y_0), the field is given by

$$U(x,y) = \begin{cases} A((\frac{x-x_0}{r})^2 + (\frac{y-y_0}{r})^2 - 1) & \text{for } (x-x_0)^2 + (y-y_0)^2 < r^2 \\ \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

The associated field of force with a critical point at the center is an ubiquitous pattern of interaction (see Fig. 4A-B). The nature of potential wells in empirical data is still unclear [33] as they can either be due to a direct interaction of a receptor, proteins or a molecule with a partner such as a scaffolding protein or with a coherent ensemble of organized molecules.

To recover the strength A of a potential well from a spatial drift distribution (eq. 9), a minimization procedure was introduced [37] based on evaluating the error between the field of force and the closed gradient:

$$Err(A) = \int_\Omega \| - \nabla U(X) - a(X)\|^2 \frac{dS_X}{|\Omega|} \tag{13}$$

where Ω is the basin of attraction. In practice, using a sampling set of points $X_1, .., X_N$, after discretizing the error 13, we obtain

$$Err_N(A) \;=\; \frac{1}{N}\sum_{i=1}^{N}\| -\nabla U(X_i) - a(X_i)\|^2 = \sum_{i=1}^{N}(a_x^i + 2A\frac{x_i}{r^2})^2 + (a_y^i + 2A\frac{y_i}{r^2})^2, \quad (14)$$

where $a(X_i) = (a_x^i, a_y^i)$ is the empirical vector field on the square centered on X_i, and we set $(x_0, y_0) = (0,0)$ for the sake of simplicity. The harmonic potential that minimizes the distance to the field a is attained for the value A^* given by

$$A^* = -\frac{r^2}{2}\frac{\sum_{i=1}^{N} a_x^i x_i + a_y^i y_i}{\sum_{i=1}^{N} x_i^2 + y_i^2} \quad (15)$$

and the minimal error is

$$Err_N(rA^*) = \sum_{i=1}^{N} a_x^{i\,2} + a_y^{i\,2} - \frac{(\sum_{i=1}^{N} a_x^i x_i + a_y^i y_i)^2}{\sum_i x_i^2 + y_i^2}. \quad (16)$$

The distance between a vector field and one derived from a harmonic potential well in its basin of attraction is characterized by the index

$$S(a) = \begin{cases} 0 \text{ for } r = 0 \\[2ex] \dfrac{Err(A^*)}{\sum_{i=1}^{N}\|a(X_i)\|^2} = 1 - \dfrac{(\sum_{i=1}^{N} a_x^i x_i + a_y^i y_i)^2}{(\sum_{i=1}^{N} x_i^2 + y_i^2)(\sum_{i=1}^{N} a_x^{i\,2} + a_y^{i\,2})} & \text{otherwise.} \end{cases} \quad (17)$$

The index $S(a)$ varies in the interval $[0,1]$. For a pure harmonic potential $S(a) = 0$, while a drift made of independent Brownian direction (no potential) is characterized by $\lim_{N\to\infty} S_N = 1$. This minimization procedure is illustrated in Fig. 4C-E, where the drift is computed from simulated trajectories for potential wells with various strengths. The estimation of the well strength depends on the number of trajectories. The index can be used for elliptic and not only circular wells.

To conclude, empirical estimators are constructed from the first and second moment formula 8. These estimators have been applied to extract features from SPT trajectories [34, 40, 37]. Converging arrows have been interpreted as potential wells, but their biophysical nature remains unclear [33]. They could be due to membrane curvature, direct interactions with organized scaffolding proteins, local receptor interactions with converging cytoskeleton or a combination of all of them. The energies of the wells for membrane receptors or channels are distributed between 1.5 to 8 $k_B T$, although it is slightly reduced for voltage-dependent calcium channels than for AMPA receptors [63, 34]. In the next section, we will show how the localization noise impacts the recovery of the first moments and the extraction of the potential wells. Finally, potentials wells can be organized in ring domains as reported in [38].
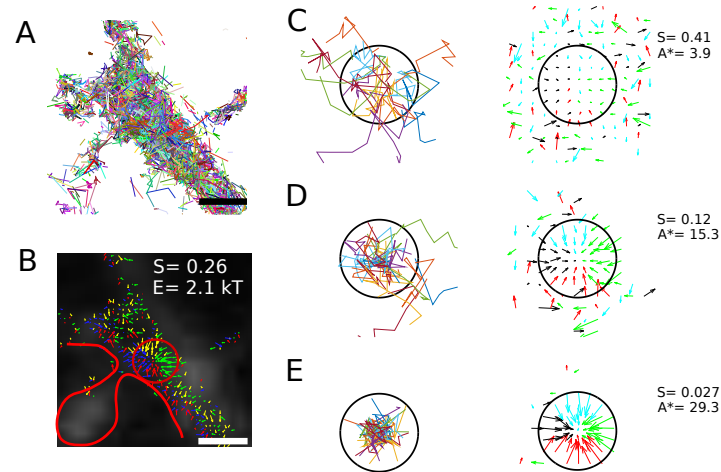
Figure 4: **Potential well strength estimation.** (A) Trajectories of AMPA receptors in a high density region, at the base of a dendritic spine: characterized by a potential well (B), as indicated by converging a vector field (adapted from [34]). Scale bar: 1μm. (C,D,E) Analysis and estimation of potential well strength on simulated trajectories. The process stochastic is $dX_t = -AX_t + \sqrt{2D}dB_t$, inside a disk of radius 1 (black), and pure diffusion $dX_t = \sqrt{2D}dB_t$ outside. Twenty trajectories, each containing ten points are shown (left) and the drift map is estimated over 400 trajectories, using formula 9. Parameters are $D = 1\mu m^2/s$, $\Delta t = 0.05s$, and (C): $A = 2$, (D): $A = 10$, (E): $A = 20$. The index S and the estimated A are indicated for each set of trajectories. Here $R = 1$ (no units), bins are 0.3, so there are about 36 bins per potential well.

11

4 Model of a stochastic process perturbed by error localisation noise and parameter estimations

The statistical methods now assume that SPTs are already reconstructed from an ensemble of points, estimated from tracking algorithms [14]. There are several sources of the noise: the acquired points are corrupted by errors approximated as Gaussian. This simple model accounts for localization errors when the emitters (dye particles) have a constant intensity during the trajectory acquisition and in addition, the noise source is identical for all trajectories. This noise is associated to localization errors introduced by assigning a random position of the detected particle within the Gaussian localization point spread function. This reconstruction procedure is now a classical for tracking and implemented in freely available softwares [14].

However, we do consider here errors originating from an upstream tracking algorithm that mistakenly combines spots from two different molecules. This error can potentially bias the reconstruction of trajectories, but as different noisy time series are combined when trajectories visit the same spatial landmarks in the cell, this error may actually be smoothed out by the averaging procedure based on formula 8.

Finally, we emphasize that statistical issues inherent to approximation for transition densities and time series estimation have been studied in the context of classical statistics, but we emphasize here complications due to small time approximation in the moment estimators [65, 1, 30, 48]. We shall present in this section some singular behaviors when the Gaussian measurement noise corrupts the signal. Thus the approach we present here takes the point of view of analytical approximation and statistical physics.

To reconstruct a diffusion process from trajectories that contains additional errors introduced by tracking algorithms or uncertainty in the detection process [85], a model of these errors is necessary. Indeed, the exact localization of the trajectory points is affected, thus modifying the estimation of the biophysical parameters such as the diffusion coefficient. The different noise terms not only bias the parameter estimates, they also induce a time correlation between the observed points of the trajectories. In this section, we present the possible noise sources and the models developed to take these fluctuation noises into account [6, 87].

A model of the localization noise introduced in measuring the position of a particle is obtained by fitting a Gaussian on the diffraction spot. The localization uncertainty accounts for the width of the Gaussian standard deviation σ, the signal intensity, the intrinsic fluctuation of individual photons and the background noise [85]. In a first approximation, the observed position y_n at time $t_n = n\Delta t$ is then obtained from the true position $x_n = x(t_n)$ by adding a Gaussian variable ξ_n

$$y_n = x_n + \sigma\xi_n. \tag{18}$$

Another source of errors originates from the finite camera exposure, which is due to the fact that collected photons are coming from a moving stochastic particle. The choice of the detected point results from the integration of the particle positions between two acquisition time points. This additional noise, called the motion blur [6], dynamic errors [72] or dynamic

12

localization uncertainty [59] depends on the number of acquired photons and on the displacement of the particle between two time intervals [59]. The motion blur can influence the point spread function resulting in breaking symmetry and non Gaussian noise description. In this section, we consider only a Gaussian localization noise, which is the same for all particles at any acquisition time step. This assumption is discussed in [12], where a time-dependent localization noise is chosen.

The camera shutter state function $\varsigma(t)$ characterizes the opening of the camera during a time interval Δt. The camera shutter state affects the observed position x_{true} of the particle at time $t_n = n\Delta t$ and leads to the choice of the position

$$x_n = \int_{t_{n-1}}^{t_n} x_{true}(t)\varsigma(t - t_{n-1})dt + \sigma\xi_n, \tag{19}$$

where

$$\frac{1}{\Delta t}\int_0^{\Delta t} \varsigma(t)dt = 1. \tag{20}$$

The motion blur coefficient is then given by

$$R = \frac{1}{\Delta t}\int_0^{\Delta t} S(t)(1 - S(t))dt, \tag{21}$$

where $S(t) = \int_0^t \varsigma(v)dv$. Recovering x_{true} from the sequence x_n and estimating the physical diffusion coefficient involves computing various statistical moments such as the mean square displacement. However, this estimation needs to account for the various sources of noise [6, 87, 88]. Indeed, the localization error and finite camera exposure affect the MSD [72, 61, 28], and the motion blur coefficient appears as a correction term in the diffusion coefficient estimation. For example, the finite exposure time critically affects the estimate of the diffusion coefficient for particles undergoing hop diffusion or confined motion [67] and the measured coefficient D decreases with the exposure time. The motion blur has higher impact on fast moving particles, due to integration over longer areas. We present now some empirical estimators.

4.1   Local estimators applied to noisy trajectories

Estimating drift and diffusion with a localization noise from the empirical estimators 9 and 10 supposes that the physical stochastic process [41] is

$$dx = a(x)dt + \sqrt{2D(x)}dw. \tag{22}$$

This section relies on the short time approximation

$$x_{n+1} = x_n + a(x_n)\Delta t + \sqrt{2D(x_n)}\Delta w. \tag{23}$$

Euler's numerical scheme of stochastic equation 23 and approximations of the transition probability density function are commonly used in the context of statistical inference of
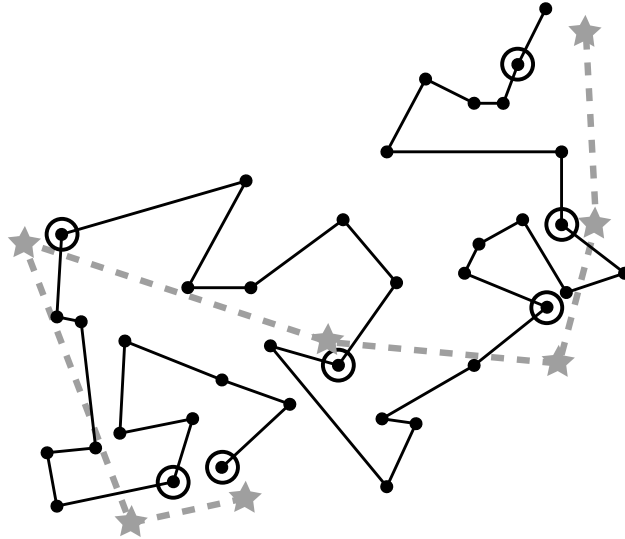
Figure 5: **Observed trajectory (grey dashed line) sampled from a physical process (black line) and perturbed by a localization noise**. At constant time interval Δt, the physical trajectory is sub-sampled (black circles). Localization noise perturbs the localization and the observed points (grey stars) are positioned in a neighborhood of physical ones.

a diffusion process, time series analysis and financial econometrics literature [65, 1, 30]. Here, we discuss only an asymptotic approach to estimate the first and second moments of a stochastic process and the two-dimensional vector field underlying the drift term. The asymptotic approach provide an estimation of the estimators at order $o(\Delta t)$.

We start when the observed points are at positions x_n for the physical process, generated by relation 18. The deconvolution of the trajectories consists in removing the Gaussian localization noise from the measured trajectories acquired at a time resolution Δt. An observed trajectory is a sequence of points y_n sampled from a physical process x_n and corrupted with a Brownian noise $\sigma \xi_n$, where $\sigma > 0$ and ξ_n are i.i.d Gaussian variable of variance 1 (relation 18). We neglect the motion blur coefficient ($R = 0$). To compute the first moment, each component in relation 23 are expanded each as follow:

$$y_{n+1} - y_n = a(y_n - \sigma \xi_n)\Delta t + b(y_n - \sigma \xi_n)\Delta w_n + \sigma(\xi_{n+1} - \xi_n). \tag{24}$$

Using a Taylor expansion

$$a(y - \sigma \xi) = \sum_{k=0}^{n} \frac{(-\sigma)^k}{k!} \frac{\partial a^k(y)}{\partial \boldsymbol{x}^k} \xi^k + O(\sigma^{n+1}), \tag{25}$$

$$b(y - \sigma \xi) = \sum_{k=0}^{n} \frac{(-\sigma)^k}{k!} \frac{\partial b^k(y)}{\partial \boldsymbol{x}^k} \xi^k + O(\sigma^{n+1}). \tag{26}$$

14

at second order (in dimension one),

$$
\begin{aligned}
y_{n+1} - y_n &= \left(a(y_n) - \sigma\xi_n a'(y_n) + \frac{\sigma^2}{2}\xi_n^2 a''(y_n) \right)\Delta t \\
&+ \left(b(y_n) - \sigma\xi_n b'(y_n) + \frac{\sigma^2}{2}\xi_n^2 b''(y_n) \right)\Delta w_n + \sigma(\xi_{n+1} - \xi_n).
\end{aligned}
$$

the expectation is computed with $\mathbb{E}_{\xi_n}(\xi_n^2) = 1$

$$
\mathbb{E}_{w,\xi}(y_{n+1} - y_n \,|\, y_n = y) = \left(a(y) + \frac{\sigma^2}{2}a''(y) + o(\sigma^2) \right) + o(\Delta t), \tag{27}
$$

We conclude that at order two, a correction has to be added to the drift, but when σ is small this contribution is negligible. In particular, this result shows that at the first order the additive noise does not influence the recovery of the vector field and local potential wells. The recovered energy is thus not affected by additive noise. The diffusion coefficient is computed from the second moment for $\Delta t \ll 1$,

$$
\begin{aligned}
\frac{\mathbb{E}_{w,\xi}((y_{n+1} - y_n)^2 \,|\, y_n = y)}{2\Delta t} &= \frac{\sigma^2}{\Delta t} + \frac{1}{2}b^2(y) + \sigma^2 a'(y) + \frac{1}{2}\sigma^2\left(b'^2(y) + \frac{b(y)b''(y)}{2} \right) \\
&+ o(\Delta t) + O(\sigma^2). \tag{28}
\end{aligned}
$$

The correlation term with the drift comes from the product $-\sigma\xi_n a'(y_n)(\sigma(\xi_{n+1} - \xi_n))$. Due to the independent property of the random variables, the product $\mathbb{E}(\xi_n\xi_{n+1}) = 0$, but $\mathbb{E}(\xi_n^2) = 1$. We are thus left with new term $\sigma^2 a'(y)$ in the effective diffusion coefficient, in agreement with the numerical simulations.

Another method to compute the statistical moment uses the transition probability $p_n(y|x) = Pr\{y_{n+1} = y|y_n = x\}$ of the observed point at position y and at time $(n+1)\Delta t$ given that at time $n\Delta t$ it is observed at position x, is computed from the pdf of the physical process by a convolution with the Gaussian distribution of the localization noise. In one dimension, when the diffusion tensor $b(X_n) = \sqrt{2D}$ is uniform in space

$$
p_n(y|x) = p(X_{n+1} + Z_{n+1} = y|X_n + Z_n = x), \tag{29}
$$

where the two processes X_n and Z_n are independent. In $\mathbb{R}$, we have

$$
p_n(y|x) = \int_{\mathbb{R}}\int_{\mathbb{R}} p(X_{n+1} = y_1|X_n = x_1) \frac{e^{-\frac{(x-x_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \frac{e^{-\frac{(y-y_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx_1 dy_1. \tag{30}
$$

For $\Delta t \ll 1$, we use the approximation of the $X_{n+1} - X_n \sim \mathcal{N}(a(X_n)\Delta t, \sqrt{2D\Delta t})$. The transition probability is thus approximated by $p(X_{n+1} = y_1|X_n = x_1) = \dfrac{e^{-\frac{(y_1 - x_1 - a(x_1)\Delta t)^2}{4D\Delta t}}}{\sqrt{4\pi D\Delta t}}$,

which gives that $p(y|x) = \int_{\mathbb{R}} \dfrac{e^{-\frac{(x-x_1)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \dfrac{e^{-\frac{(y-x_1-a(x_1)\Delta t)^2}{2(\sigma^2+2D\Delta t)}}}{\sqrt{2\pi(\sigma^2+2D\Delta t)}} dx_1$. To obtain an explicit expression of this convolution, we use the change of variable $x_1 = x + \sigma\eta$ where $\sigma \ll 1$,

$$p_n(y|x) = \int_{\mathbb{R}} \frac{e^{-\frac{\eta^2}{2}}}{\sqrt{2\pi}} \frac{e^{-\dfrac{(y-x-a(x)\Delta t - \sigma\eta(1+a'(x)\Delta t))^2}{2(\sigma^2+2D\Delta t)}}}{\sqrt{2\pi(\sigma^2+2D\Delta t)}} d\eta.$$

This integral can be regarded as the convolution of two gaussian functions over the real line, and

$$p_n(Y_{n+1}=y|Y_n=x) = \frac{e^{-\dfrac{(y-x-a(x)\Delta t)^2}{2\sigma^2_{\Delta t}(x)}}}{\sigma_{\Delta t}(x)\sqrt{2\pi}}, \tag{31}$$

where

$$\sigma^2_{\Delta t}(x) = 2\sigma^2(1+a'(x)\Delta t) + 2D\Delta t + O(\Delta t)^2. \tag{32}$$

We conclude that the transition probability of the observed process Y_n is Gaussian and $Y_{n+1} - Y_n \sim \mathcal{N}(a(Y_n)\Delta t, \sigma_1(Y_n))$. The observed motion is thus defined by the discrete scheme:

$$\tilde{Y}_{\Delta t}(t+\Delta t) = \tilde{Y}_{\Delta t}(t) + a_{obs}(\tilde{Y}_{\Delta t})\Delta t + \frac{\sigma_{obs,\Delta t}(\tilde{Y}_{\Delta t})}{\sqrt{\Delta t}}\Delta W_t, \tag{33}$$

where $\Delta W_t = W(t+\Delta t) - W(t)$ and W is a Brownian motion of variance 1 and

$$a_{obs}(x) = a(x) \tag{34}$$
$$\sigma_{obs,\Delta t}(x) = \sigma_{\Delta t}(x) = \sqrt{2\sigma^2(1+a'(x)\Delta t) + 2D\Delta t + O(\Delta t)^2}. \tag{35}$$

This approach allows defining the continuous process $\tilde{Y}_{\Delta t}$ from the approximation at the scale Δt, it is solution of the stochastic equation

$$d\tilde{Y}_{\Delta t}(s) = a(\tilde{Y}_{\Delta t})ds + \frac{\sigma_{\Delta s}(\tilde{Y}_{\Delta t})}{\sqrt{\Delta t}}dW_s. \tag{36}$$

The drift of the observed process at a time resolution Δt is the same (at first order in σ) as the physical one, while the diffusion tensor is changed and given by formula 32.

An estimator for the drift at a time resolution Δ of the observed process is constructed from

$$a_{\Delta t}(x) = \mathbb{E}\left[\frac{Y_{n+1}-Y_n}{\Delta t}\Big|Y_n=x\right] = \frac{1}{\Delta t}\int_{\mathbb{R}}(y-x)p(Y_{n+1}=y|Y_n=x)dy = a(x)+o(1). \tag{37}$$

The result 27 is recovered: adding a Gaussian noise on the physical process does not alter the physical deterministic drift at first order in σ, while the diffusion coefficient is

$$D_{\Delta t}(x) = \frac{\sigma^2}{\Delta t} + D + \sigma^2 a'(x) + \frac{a^2(x)}{2}\Delta t + o(\Delta t). \tag{38}$$

16

4.2 Generalization in dimension m

For a m-dimensional stochastic process

$$d\boldsymbol{X} = \boldsymbol{A}(\boldsymbol{X})dt + \sqrt{2D}d\boldsymbol{w}, \tag{39}$$

where $\boldsymbol{A}$ is a vector field and $\boldsymbol{w}$ the classical $m-$dimensional centered Brownian motion of variance 1. The diffusion tensor is assumed to be a constant D. The observed motion is

$$\boldsymbol{Y}_n = \boldsymbol{X}_n + \sigma\boldsymbol{\eta_n}, \tag{40}$$

where $\boldsymbol{\eta_n}$ are i.i.d m-dimensional standard gaussian variable and σ is a small parameter. Then, the estimator for the drift is [41]

$$
\begin{aligned}
\boldsymbol{a}_{\Delta t}(\boldsymbol{x}) &= \mathbb{E}\left[\frac{\boldsymbol{Y}_{n+1} - \boldsymbol{Y}_n}{\Delta t}|\boldsymbol{Y}_n = \boldsymbol{x}\right] \\
&= \boldsymbol{A}(\boldsymbol{x}) + o(\Delta t) + O(\sigma^2),
\end{aligned} \tag{41}
$$

and the estimated diffusion coefficient relates to the physical parameter by

$$
\begin{aligned}
D_{\Delta t}(\boldsymbol{x}) &= E\left[\frac{\|\boldsymbol{Y}_{n+1} - \boldsymbol{Y}_n\|^2}{2m\Delta t}|\boldsymbol{Y}_n = \boldsymbol{x}\right] \\
&= D + \frac{\sigma^2}{\Delta t} + \frac{\sigma^2}{m}div(\boldsymbol{A}) + O(\Delta t),
\end{aligned} \tag{42}
$$

where $div(\boldsymbol{A})$ is the divergence of the drift vector. These formulas show how spatial variations of the drift affect the measured diffusion tensor.

4.3 Other estimators

For a stochastic process containing a drift component, it is not possible to extract the physical diffusion coefficient directly by combining the first and the second moment estimators, which is in contrast with the pure diffusion case (see [6, 87]). We now present an estimator where the Gaussian instrumental noise can be eliminated. Using the difference $\Delta Y_n = Y_{n+1} - Y_n$, we can rewrite

$$
\begin{aligned}
\Delta Y_n &= a(X_n)\Delta t + \sigma(X_n)\Delta W_n + \sigma(\eta_{n+1} - \eta_n) \\
\Delta Y_{n-1} &= a(X_{n-1})\Delta t + \sigma(X_{n-1})\Delta W_{n-1} + \sigma(\eta_n - \eta_{n-1}),
\end{aligned}
$$

where ΔW_n and ΔW_{n-1} are two independent increments of Brownian motion. The expectation is

$$E\left[\frac{(Y_{n+1} - Y_n)(Y_n - Y_{n-1})}{\Delta t}\right] = -\frac{\sigma^2}{\Delta t}E(\eta_n^2) = -\frac{\sigma^2}{\Delta t} + o(1). \tag{43}$$

Using relation 38, we obtain that

$$E\left[\frac{(Y_{n+1} - Y_n)^2}{2\Delta t}|Y_n = x\right] + E\left[\frac{(Y_{n+1} - Y_n)(Y_n - Y_{n-1})}{\Delta t}\right] = D + \sigma^2 a'(x) + o(1).$$
(44)

In this estimator, the instrumental noise is averaged out. There are no direct procedures to get rid of the derivative of the drift term, which can be extracted from the first order moment. However, computing a derivative from noisy data should be done carefully as it introduces singularities and irregularities.

4.4 Recovering the diffusion tensor in dimension 1

Using formula 38, we estimated the diffusion coefficient $\tilde{D}$ in Figs. 6A and 6B. The signal to noise ratio (SNR) is defined as $\frac{D}{\frac{\sigma^2}{\Delta t}}$. A high SNR can be due either to a large sampling rate Δt or to a low positional noise. In our numerical application, we first vary the SNR by fixing the amplitude of the noise σ and by varying the increment Δt (Figs. 6A and 6C) then we vary the parameters the other way around (Figs. 6A and 6D). We also estimated the diffusion coefficient for an OU process (Figs. 6B and 6D). These numerical estimations show that the estimator for the diffusion coefficient is biased for a high SNR for an OU process when the positional noise σ is fixed and the time step Δt increases (Fig. 6C). This counterintuitive result is due to the approximation 23 for the physical motion $X_{n+1} - X_n \sim \mathcal{N}(a(X_n)\Delta t, \sqrt{2D\Delta t})$, which is only applicable at small time steps Δt. However this approximation is perfectly valid for a Brownian motion, as shown in Fig. 6A.

4.5 Empirical estimators associated to an Ornstein-Uhlenbeck process

For an Ornstein-Uhlenbeck (OU) process, $dX = -\lambda(X-\mu)dt+\sqrt{2D}dW$, the pdf is $p(y,t|x,0) = \frac{1}{\sqrt{2\pi D\frac{(1-e^{-2\lambda t})}{\lambda}}}\exp(-\frac{(y-\mu-(x-\mu)e^{-\lambda t})^2}{\frac{2D}{\lambda}(1-e^{-2\lambda t})})$. For the observed motion Y_n given by eq. 30 the pdf is

$$p(Y_{n+1} = y|Y_n = x) \quad \times = \frac{e^{-\frac{(y - \mu - (x - \mu)e^{-\lambda\Delta t})^2}{2(\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t}))}}}{\sqrt{2\pi(\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t}))}}.$$
(45)

and the observed drift and diffusion coefficient at timescale Δt are [41]

$$a_{\Delta t}^{OU}(x) = \frac{1}{\Delta t}\int_{\mathbb{R}}(y - x)p(Y_{n+1} = y|Y_n = x)dy = -(x - \mu)\frac{1 - e^{-\lambda\Delta t}}{\Delta t},$$
(46)

$$D_{\Delta t}^{OU}(x) = \frac{1}{2\Delta t}\int_{\mathbb{R}}(y - x)^2 p(Y_{n+1} = y|Y_n = x)dy$$

$$= \frac{1}{2\Delta t}\left(\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t})\right) + (\mu - x)^2\frac{(1 - e^{-\lambda\Delta t})^2}{2\Delta t}.$$
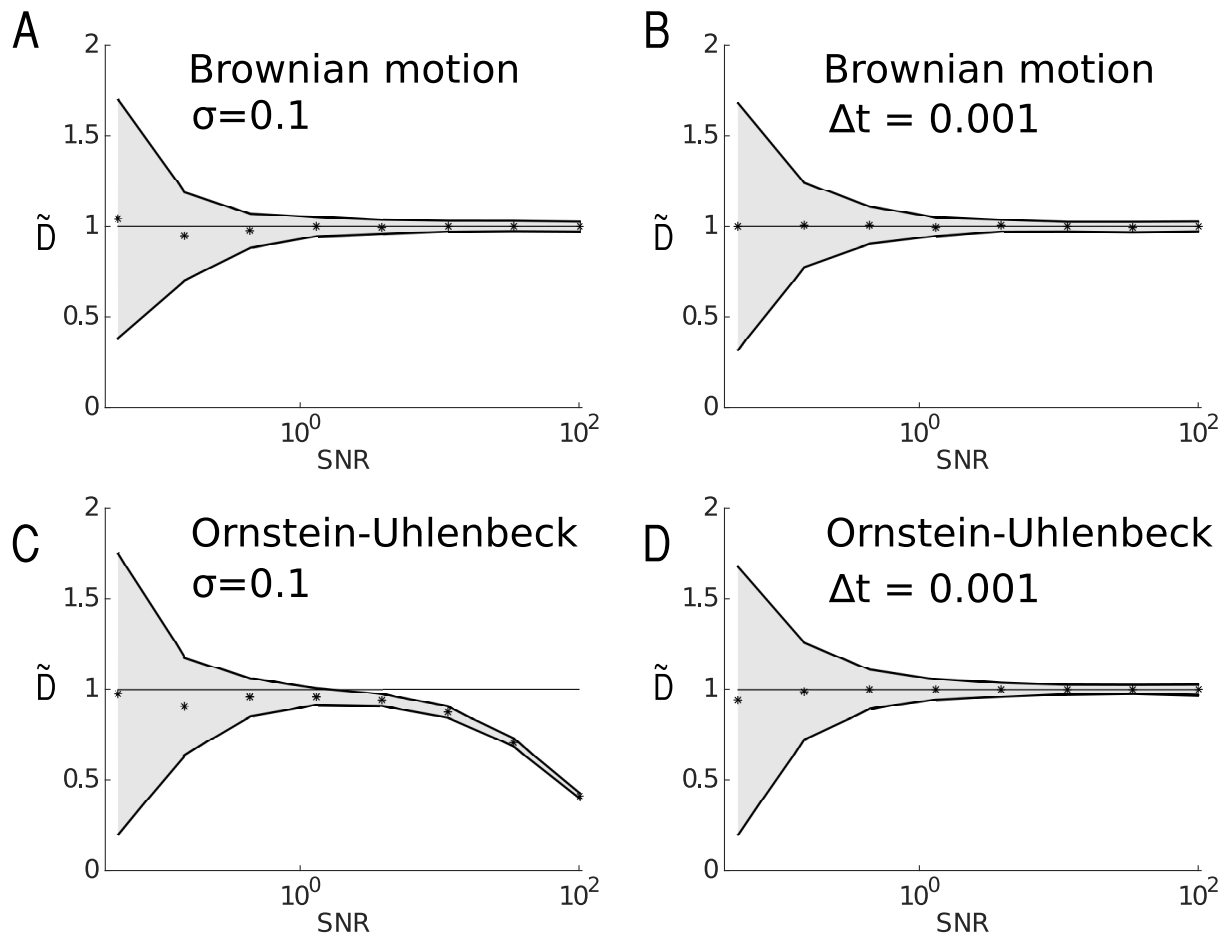
Figure 6: Diffusion coefficients are estimated for a Brownian motion (A and B) and Ornstein-Uhlenbeck process (C and D) and for various values of the signal-noise ratio (SNR) represented in a log-10 scale. Trajectories were simulated using Euler's scheme, and sub-sampled so that the observed trajectories contain 10,000 points, and position noise σ was subsequently added to each point of the trajectories. The diffusion coefficient $D_{\Delta t}$ is estimated using formula 10. Black dots represent $\tilde{D} = D_{\Delta t} - \frac{\sigma^2}{\Delta t}$ and the continuous lines bounding the grey area represent $\tilde{D} \pm s.d.$. The diffusion coefficient is $D = 1$ and for the OU process the drift is $a(x) = -2x$. Variations of the SNR, defined as $\frac{D}{\frac{\sigma^2}{\Delta t}}$, are obtained for fixed position noise (A,C) or fixed sampling time (B,D). In (A) and (C), the positional noise is fixed at $\sigma = 0.1$, while the sampling rate Δt is varying. In (B) and (D), the sampling time is fixed to $\Delta t = 0.001$ and the position noise σ is varying.

19

In Fig. 7, we estimate the local drift and diffusion coefficient for an OU-process and compare the local estimators for the drift 37 with relation 46 (Fig. 7B). For the diffusion tensor, we compare relations 38 and 47 (Fig. 7C). Using the first order approximation for short time step Δt, we compare estimators 37 (resp. 38) that gives improved results compared to 46 (resp. 47), showing the improvement of the $O(\Delta t)$. Note that other classical statistical methods have been developed to recover drift and diffusion tensor [65, 1, 30] based on Bayesian inference or direct empirical sums.
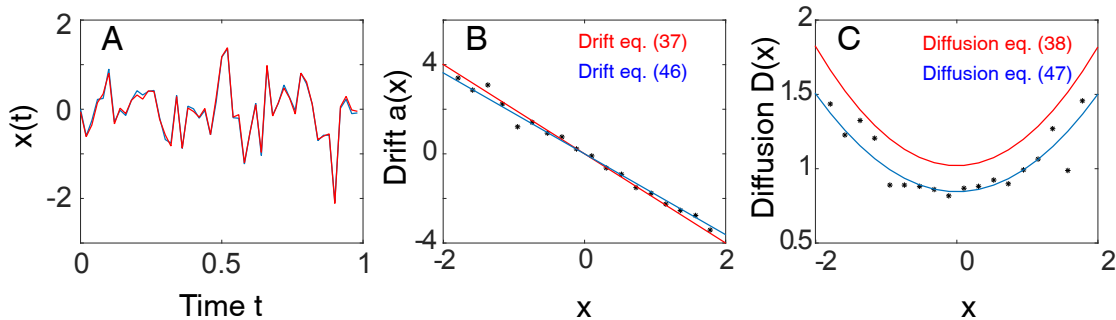

Figure 7: **Recovering an Ornstein-Uhlenbeck process at second order in Δt** (A) Trajectory of a one dimensional OU-process, generated using Euler's scheme (blue curve) and the observed trajectory (red curve) is obtained by subsampling at $\Delta t = 0.1$ and with an additional position noise of standard deviation $\sigma = 0.05$ (SNR=40). The other parameters are fixed to $D = 1$, $\lambda = 2$, $\mu = 0$. The observed trajectories contain 10,000 points. (B) Estimation of the local drift using eq. 9 (black dots), and comparison with the analytical formulas 37 (red) and 46 (blue). (C) Estimation of the local diffusion coefficient using eq. 10 (black dots) and comparison with the analytical formulas 38 (red) and 47 (blue), confirming the effect of the $O(\Delta t)$ term.

4.6 Maximum likelihood estimators for an Ornstein-Uhlenbeck process using the approximated and exact transition probabilities

We now describe briefly the maximum likelihood estimators to recover an OU process from the approximated transition probability. For a OU-process sampled at short time step Δt (eq. 23) with the approximated scheme $X_{n+1} = X_n - \lambda(X_n - \mu)\Delta t + \sqrt{2D}\Delta w$, the transition probability of the observed motion is

$$p(Y_{n+1} = y | Y_n = x) = \frac{e^{-\frac{(y - x + \lambda(x - \mu)\Delta t)^2}{2\sigma_{\Delta t}^2}}}{\sigma_{\Delta t}\sqrt{2\pi}}, \tag{47}$$

where the variance is

$$\sigma_{\Delta t}^2 = 2\sigma^2 + (2D - 2\sigma^2\lambda)\Delta t + O(\Delta t^2).$$

20

The log-likelihood is defined as

$$
\begin{aligned}
\ell(y_1, \ldots, y_{N+1}|\boldsymbol{\theta}) &= \sum_{n=1}^{N} \ln p(y_{n+1}|y_n; \boldsymbol{\theta}) \\
&= -\sum_{n=1}^{N} \ln \sigma_{\Delta t}(y_n, \boldsymbol{\theta}) - \frac{1}{2} \sum_{n=1}^{N} \frac{(y_{n+1} - y_n - a(y_n; \boldsymbol{\theta})\Delta t)^2}{\sigma_{\Delta t}^2(y_n)}.
\end{aligned}
\tag{48}
$$

The likelihood function is applied here on a partially reconstructed OU-process, where the corrupted noise amplitude σ has already been accounted for in the variance term $\sigma_{\Delta t}$. Thus the discretization of the nonlinear SDEs for small time step Δt does not rely on any addition assumption about the initial non-Markovian process of the observed measurements (generated from a discretize OU-process with an added Gaussian noise). The likelihood function 48 expresses the joint distribution of the observations as a product of transition densities, which is in fact valid for Markovian processes (i.e. a time series recorded without measurement error). But the observed measurement noise sequence is not a Markovian process. The exact likelihood would in principle requires using the entire series $p(y_{n+1}|y_n; y_{n-1}..1; ..; y1; \boldsymbol{\theta})$. To overcome this difficulty, the effect of the noise is taken into account directly in the term $\sigma_{\Delta t}$. This procedure approximates a non-Markovian by a Markovian process. The variance term is recovered, however the error in the drift is not accounted for (see relation 27). In general, for an OU-process with constant measurement noise, the exact computation is tractable and detailed are given in ([30] p. 386). Yet, as we shall see the maximum likelihood estimator allows recovering the parameters of interest.

The parameters $\theta_1, \ldots, \theta_m$ and D are computed as maximizers of the likelihood function and thus by differentiating ℓ with respect to the variables $\theta_1, \ldots, \theta_m, D$. The conditions $\frac{\partial \ell}{\partial D} = 0$ and $\frac{\partial \ell}{\partial \theta_i} = 0$ can be rewritten as

$$
\ell(y_1, \ldots, y_{N+1}|\lambda, D) = -N \ln \sigma_{\Delta t} - \frac{1}{2} \sum_{n=1}^{N} \frac{(y_{n+1} - y_n + \lambda(y_n - \mu)\Delta t)^2}{\sigma_{\Delta t}^2}.
$$

Conditions $\frac{\partial \ell}{\partial D} = 0$, $\frac{\partial \ell}{\partial \lambda} = 0$ and $\frac{\partial \ell}{\partial \mu} = 0$ lead to

$$
\sum_{n=1}^{N} y_n(y_{n+1} - y_n + \lambda(y_n - \mu)\Delta t)) = 0,
$$

and thus the empirical estimator $\tilde{\lambda}$ for the parameter λ is

$$
\tilde{\lambda} = -\frac{1}{\Delta t} \frac{\sum_{n=1}^{N} y_n(y_{n+1} - y_n)}{\sum_{n=1}^{N} y_n(y_n - \tilde{\mu})}.
\tag{49}
$$

Similarly, using $\frac{\partial \ell}{\partial \mu} = 0$, we obtain the condition

$$
\tilde{\mu} = \frac{1}{N\tilde{\lambda}\Delta t}(y_{N+1} - y_1) - \frac{1}{N} \sum_{n=1}^{N} y_n.
\tag{50}
$$

21

By combining 49 and 50 we obtain

$$\tilde{\mu} = \frac{\sum_{n=1}^{N} y_n \sum_{n=1}^{N} y_n(y_{n+1} - y_n) - \sum_{n=1}^{N} y_n^2(y_{N+1} - y_1)}{N \sum_{n=1}^{N} y_n(y_{n+1} - y_n) - \sum_{n=1}^{N} y_n(y_{N+1} - y_1)}. \tag{51}$$

Finally, the empirical estimator for the diffusion coefficient is

$$\tilde{D} = \sigma^2(\tilde{\lambda} - \frac{1}{\Delta t}) + \frac{1}{2N\Delta t} \sum_{n=1}^{N} (y_{n+1} - y_n + \tilde{\lambda}(y_n - \tilde{\mu})\Delta t)^2. \tag{52}$$

Estimators can also be built using the exact transition probability of the OU process

$$p_n(Y_{n+1} = y | Y_n = x) \quad = \quad \frac{e^{-\frac{(y - \mu - (x - \mu)e^{-\lambda\Delta t})^2}{2(\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t}))}}}{\sqrt{2\pi(\sigma^2(1 + e^{-2\lambda\Delta t}) + \frac{D}{\lambda}(1 - e^{-2\lambda\Delta t}))}}. \tag{53}$$

and are given by [41]

$$\tilde{\mu} = \frac{1}{N} \left(\sum_{i=2}^{N} y_i \right) + \frac{1}{1 - e^{-\tilde{\lambda}\Delta t}} \frac{y_{N+1} - y_1 e^{-\tilde{\lambda}\Delta t}}{N}, \tag{54}$$

$$\tilde{D} = \frac{\tilde{\lambda}}{1 - e^{-2\tilde{\lambda}\Delta t}} \left(\frac{1}{N} \sum_{i=1}^{N} (y_{i+1} - \tilde{\mu} - (y_i - \tilde{\mu})e^{-\tilde{\lambda}\Delta t})^2 \right) - \sigma^2 \tilde{\lambda} \frac{1 + e^{-2\tilde{\lambda}\Delta t}}{1 - e^{-2\tilde{\lambda}\Delta t}}. \tag{55}$$

The two maximum-likelihood estimators for the OU process are compared using numerical simulations, where the time-step is $\Delta t = 0.1$ and the parameters λ, μ and D are estimated for $n = 500$ observations. The parameters are estimated for various values of the observation noise σ. The results are summarized in Fig. 8: the parametric estimators based on the exact transition probability of the OU process, gives better estimates than for the one using the approximated pdf. To estimate the likelihood function, other approaches are based on Kalman filtering [30].

4.7 Parameter estimation for motion with varying dynamics

For non-stationary processes, a time correlation between points of SPT trajectories renders the analysis quite difficult: indeed when particles can alternate between free and confined motion, or experience rapid changes in the interacting forces [12, 88], the process is not stationary. Similarly, potential wells can appear or disappear [40] at a time scale of minutes. To extract information from non stationary SPT trajectories, the Hierarchical Dirichlet Process Switching Linear Dynamical System (HDS-SLDS) method [20, 21] detects abrupt changes in the forces governing a single trajectory [11] (see also [9, 12]). The method extends hidden Markov models (HMM) [83], in which the different states are associated with a linear dynamical process. The underlying discrete physical model of the motion is
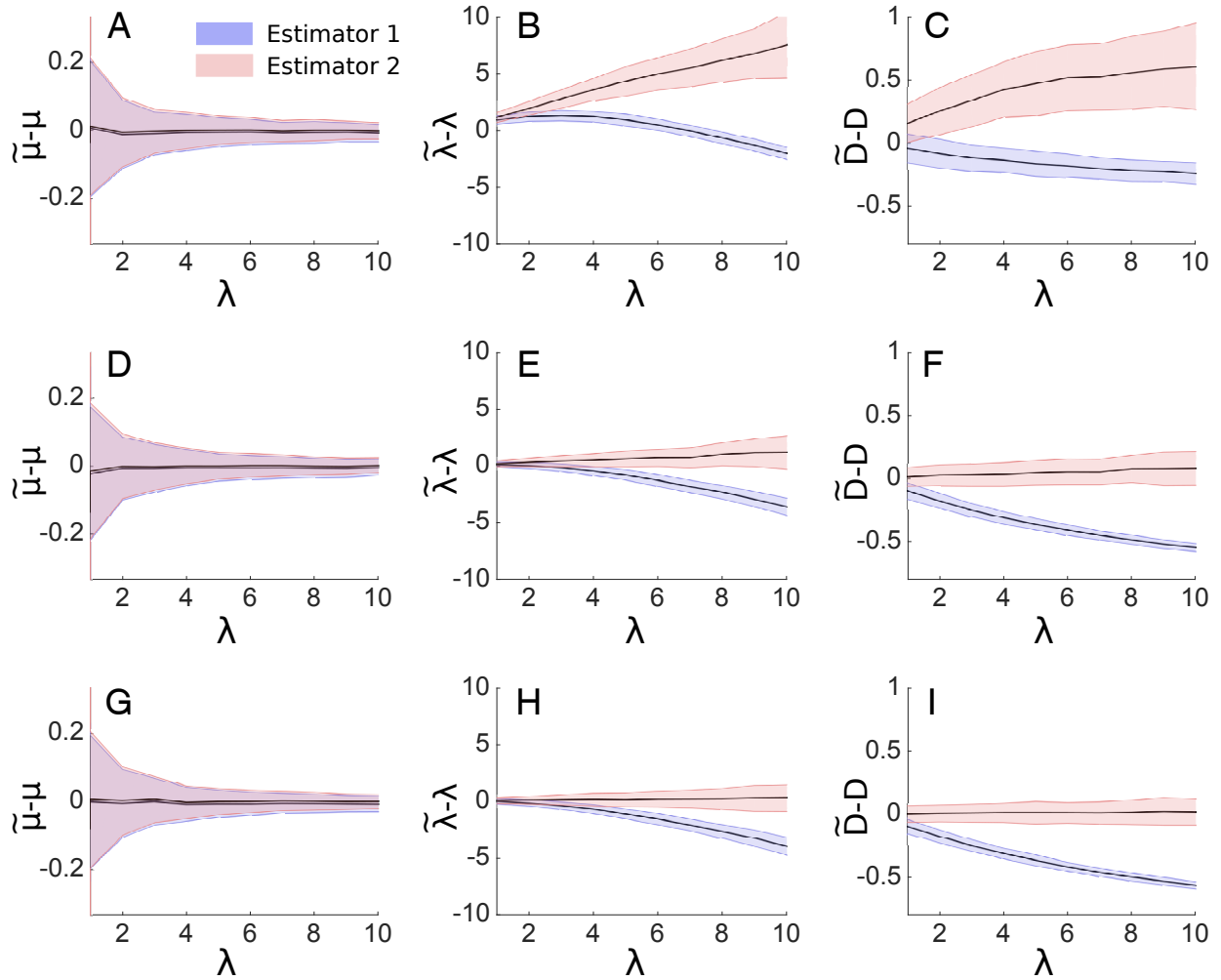
Figure 8: The maximum-likelihood estimators presented in 4.6 (blue : approximated pdf, red : exact pdf) for an OU process with $D = 1$, $\mu = 1$, and various values of λ are compared. From left to right, the plotted estimations are $\tilde{\mu} - \mu$ (A,D,G), $\tilde{\lambda} - \lambda$ (B,E,H), and $\tilde{D} - D$ (C,F,I). The observation time-step is $\Delta t = 0.1$ and from top to bottom, $\sigma = \sqrt{0.1}$ ((A,B,C), SNR=1), $\sigma = \sqrt{0.01}$ ((D,E,F), SNR=10), $\sigma = \sqrt{0.001}$ ((G,H,I), SNR=100). The black line is the mean of the estimation for trajectories of 500 points, and the colored part indicates $\pm$ the standard deviation (adapted from [41])

$$x_{n+1} = \mu + F(x_n) + \Sigma\eta_n; \eta_n \sim \mathcal{N}(0,1), \tag{56}$$

where x_n is the position at the discrete time step t_n, μ is a constant vector, F is a space-dependent force matrix, and η_n a random variable to account for fluctuations in the particle motion. The observed position is the sum of the physical position plus a normal random variable with variance σ (localization noise)

$$y_{n+1} = x_{n+1} + \sigma\vec{\epsilon}_{n+1}; \vec{\epsilon}_{n+1} \sim \mathcal{N}(0,1). \tag{57}$$

When the motion alternates between different states, characterized by the parameters $\theta = (\mu, F, \Sigma, \sigma)$, the procedure consists in decomposing a trajectory in different segments using a non-parametric Bayesian technique. The motion parameters in each sub-trajectory is estimated using a maximum likelihood estimation (MLE) method. The transition between each state is detected as large jumps in consecutive points of the observed trajectory.

In the framework of the HMM, transitions between the different states are modeled as Markovian probabilities. Finally, to classify noise terms generated by thermal motion, external forces or the localization noise, a Kalman filter can be used to estimate the physical position of the particle. This method was used to detect various regimes in chromatin motion in yeast cells and to estimate parameters in each cell stage [11].

4.8 Comparison with classical single particle trajectories analysis

The ensemble of estimators described above are based on the exact expression computed from the statistical moments and pdf of the stochastic processes. We now recall the classical mean-square displacement (MSD) estimators. The diffusion coefficient is computed from single particle trajectories using the MSD, which consists of averaging the square displacement over single trajectories: in one dimension, for a sequence of points $(x_1, \ldots, x_{N+1})$ sampled at constant time intervals Δt, the MSD is given by [64]

$$\rho_n = \frac{1}{N-n+1} \sum_{k=1}^{N-n+1} (x_{k+n} - x_k)^2, \text{ for } 1 \leq n \leq N. \tag{58}$$

For a Brownian particle with a diffusion coefficient D, the MSD is given by

$$\langle \rho_n \rangle = 2Dn\Delta t. \tag{59}$$

This relation between the MSD and the diffusion coefficient is however not a good estimator of the diffusion coefficient for a general stochastic process that contains a drift or a localization noise, but also for a Brownian motion [6, 87, 88, 16]. For example, the accumulation of the noise by adding points along the trajectory affects the extracted value of D. The typical MSD curve differs from a line because at large time lags, the number of points used in the averaging is very low and the squared displacements are not independent from each other, resulting in cumulating errors [64]. Moreover, this procedure does not preserve the spatial heterogeneity of regions explored by the random particle.

The finite exposure time (see relation 19) affects the estimate of the diffusion coefficient for particles undergoing hop diffusion or confined motion [67]. The measured diffusion coefficient D decreases as the exposure time increases. The motion blur has even more effect on fast moving particles, as the particle observed position is the result of integration over longer areas. Including motion blur in the MSD leads to the following estimate for MSD [88]

$$\langle \rho_n \rangle = 2Dn\Delta t + 2(\sigma^2 - 2RD\Delta t). \tag{60}$$

In the case of a motion blur and position noise, optimal estimates of the diffusion coefficient uses an optimal number of points in the MSD formula [59]. Localization error and the finite camera exposure affect the MSD curve and thus perturb the estimation of the diffusion coefficient [72, 61, 28]. In the absence of localization error, the optimal estimate of the diffusion coefficient uses the first two points of the MSD [74, 59]. To conclude, fitting the MSD curve is difficult because the sequence ρ_n are highly correlated. The estimation of the diffusion coefficient becomes imprecise as the number of points included in the fit increases [59]. The MSD remains commonly used [64] to estimate the diffusion coefficient despite localization and dynamical noise that alter significantly the MSD slope at short timescales, which might be erroneously attributed to subdiffusion processes [6].

5 Statistical analysis of a trajectories from a single locus located on a polymer

Long SPT trajectories can be used for extracting biophysical parameters and mechanical forces. Indeed, a dye molecule located on a DNA is often used as a probe to explore the local cellular nucleus environment. We present here several estimators used to reconstruct external and internal forces applied on the DNA molecule and chromatin from the statistics of long locus trajectories. In the stochastic model, the locus motion is driven by local diffusion plus forces between monomers of a model polymer [45, 89, 86]. Monomer motions are usually highly correlated leading to anomalous diffusion behavior [57, 42].

To distinguish external forces applied on a single monomer from intrinsic forces acting on monomers, the first step is to use a polymer model such as Rouse. A stochastic particle is modeled by the Smoluchowski's limit of the Langevin equation: the velocity $\mathbf{v}$ is proportional to a force $\mathbf{f}$ plus an additional white noise, summarized as

$$\mathbf{v} = \frac{\mathbf{f}}{\gamma} + \sqrt{2D}\frac{\mathbf{dw}}{\mathbf{dt}}, \tag{61}$$

where γ is the friction coefficient, D the diffusion coefficient and w is the normalized Wiener process. By averaging over the ensemble of velocity realizations, the first moment or the force field is recovered [37]. However, for a polymer chain, there are internal forces between monomers. In recent statistical methods, it was possible to separate from measured single (monomer) trajectories, internal polymer forces from the external ones acting on a single localized monomer.

25

For an external force which is the gradient of the potential $U_{\text{ext}}(\boldsymbol{R})$ and applied to a Rouse polymer, the physical interactions are described by the energy term

$$\phi(\boldsymbol{R}) = \frac{\kappa}{2} \sum_{j=2}^{N} (\boldsymbol{R}_j - \boldsymbol{R}_{j-1})^2 + U_{\text{ext}}(\boldsymbol{R}), \tag{62}$$

where $\boldsymbol{R} = (\boldsymbol{R}_1, \boldsymbol{R}_2, ..., \boldsymbol{R}_N)$ is the ensemble of monomer positions, connected by a spring of strength $\kappa = d k_B T / b^2$. b is the standard-deviation of the distance between adjacent monomers [17] and d the dimensionality (dim 2 or 3). In the Smoluchowski's limit of the Langevin equation [77], the dynamics of monomer $\boldsymbol{R}_n$ is described by

$$\frac{d\boldsymbol{R}_j}{dt} = -D \nabla_{\boldsymbol{R}_j} \phi(\boldsymbol{R}) + \sqrt{2D} \frac{d\boldsymbol{w}_j}{dt}, \tag{63}$$

for $j = 1, \ldots, N$ and each $\boldsymbol{w}_j$ is an independent d-dimensional white noise with mean zero and variance 1, D is the monomer diffusion coefficient. We will describe specifically the field of force $\nabla_{\boldsymbol{R}_j} U_{\text{ext}}(\boldsymbol{R})$ in the next subsection.

A reference monomer $\boldsymbol{R}_c$ is chosen representing a tagged locus and the first statistical moment of the monomer $\boldsymbol{R}_c$ (averaged over all realizations) is recovered from data from an inversion formula, generalizing eq. 8:

$$\lim_{\Delta t \to 0} \mathbb{E}\{\frac{\boldsymbol{R}_c(t + \Delta t) - \boldsymbol{R}_c(t)}{\Delta t} | \boldsymbol{R}_c = \boldsymbol{x}\} = -D \int_\Omega d\boldsymbol{R}_1 .. \int_\Omega d\boldsymbol{R}_N (\nabla_{\boldsymbol{R}_c} \phi) P(\boldsymbol{R} | \boldsymbol{R}_c = \boldsymbol{x}), \tag{64}$$

where $\mathbb{E}\{.|\boldsymbol{R}_c = \boldsymbol{x}\}$ denotes averaging under the condition that the tagged monomer is at position $\boldsymbol{R}_c = \boldsymbol{x}$. Formula 64 is generic and does not depend on the external forces acting on the polymer.

When a polymer chain evolves in a domain Ω, with reflecting condition on the boundary $\partial \Omega$, the conditional probability $P(\boldsymbol{R} | \boldsymbol{R}_c = \boldsymbol{x})$ is computed from the equilibrium probability distribution function $P(\boldsymbol{R}_1, \boldsymbol{R}_2, ..., \boldsymbol{R}_N)$, which satisfies the Fokker-Planck equation (FPE) in the phase-space $\Omega \times ..\Omega \subset \mathbb{R}^{3N}$,

$$0 = \Delta P(\boldsymbol{R}) + \nabla \cdot (\nabla \phi \, P(\boldsymbol{R})), \tag{65}$$

with boundary condition

$$\phi \frac{\partial P}{\partial \boldsymbol{n}_i} + P \frac{\partial \phi}{\partial \boldsymbol{n}_i} = 0 \text{ for } \boldsymbol{R}_i \in \partial \Omega \quad \text{for } i = 1..N,$$

where $\boldsymbol{n}_i$ is the normal vector to the boundary $\partial \Omega$ at position $\boldsymbol{R}_i$. A permanent force located at position $\boldsymbol{\mu}$ can be approximated by an harmonic well. A force is applied to monomer n located at position $\boldsymbol{R}_n$ and is computed from the gradient of the harmonic potential

$$U_{\text{ext}}(\boldsymbol{R}_n) = \frac{1}{2} k (\boldsymbol{\mu} - \boldsymbol{R}_n)^2, \tag{66}$$

where k is the force constant. Any monomer n is affected by the applied force, but differently from the tagged monomer c and we shall assume that $n < c$. The potential well affects the

dynamics of the entire polymer and specifically the observed locus c. For the mean velocity of monomer c, we get the following relation

$$\lim_{\Delta t \to 0} \mathbb{E}\{\frac{\boldsymbol{R}_c(t + \Delta t) - \boldsymbol{R}_c(t)}{\Delta t} | \boldsymbol{R}_c(t) = \boldsymbol{x}\} = -Dk_{cn}\boldsymbol{x},$$

$$k_{cn} = \frac{k\kappa}{\kappa + (c - n)k}. \tag{67}$$

Expression 67 links the average velocity over empirical trajectories of the observed monomer c to a permanent force applied on monomer n. The coefficient k_{cn} depends on the harmonic well strength k, the inter-monomer spring constant κ and is inversely proportional to the distance $|n - c|$ between monomers n and c along the chain.

To extract the empirical effective spring coefficient k_c from a trajectory given by $\boldsymbol{R}_c(h\Delta t)$ ($h = 1 \ldots N_p$), where N_p is the number of points, the differential quotient in eq. 67 is first computed. This step allows extracting the mean position of the locus. Once the steady state is reached, the time average of the locus position is computed from

$$\langle \boldsymbol{R}_c \rangle \approx \frac{1}{N_p} \sum_{h=1}^{N_p} \boldsymbol{R}_c(h\Delta t).$$

When the polymer interacts with a single interacting potential, the average position $\langle \boldsymbol{R}_c \rangle$ estimates the location where the force is applied. An upper bound for the number of points N_p is of order $(\tau_r/\Delta t)$, where $\tau_r = \frac{|c-n|^\beta}{D\kappa\pi^\beta}$ is the relaxation time for a portion of the chain of a β-polymer between monomers c and n [2].

When the diffusion coefficient D is known (which could be extracted using the second moment estimator 10) and the inter-monomer spring constant κ are known, it is possible to estimate the force from the constant k_{cn} in eq. 67, using the linearity of the force with respect to the position of the locus (see eq. 66). The step size $(\boldsymbol{R}(\Delta t(h+1)) - \boldsymbol{R}(\Delta th))$ is proportional to the locus position $(\boldsymbol{R}(\Delta t(h+1)) - \langle \boldsymbol{R}_c \rangle)$. In the isotropic case, the apparent force constant k_{cn} acting on monomer c is computed from the trajectories of $\boldsymbol{R}_c(t)$ by the estimator

$$\tilde{k}_c \approx \frac{1}{d(N_p - 1)} \sum_{i=1}^{d} \sum_{h=1}^{N_p-1} \frac{R_c^i((h+1)\Delta t) - R_c^i(h\Delta t)}{D\Delta t(R_c^i(h\Delta t) - \langle R_c^i \rangle)}, \tag{68}$$

(d is the dimension and N_p is the number of points). This approach has been validated using stochastic simulations and live cell imaging data [3]. For a strong anchoring ($k \gg \kappa$), we can approximate $k_c \approx \kappa|c - n|^{-1}$. In that case, the empirical effective spring constant can be used to estimate the distance to the interacting monomer.

For a long enough sampled trajectory and a force derived from a stationary potential well, the effective spring coefficient can be recovered directly from the empirical estimator 68. However, with a small sampling step Δt, refined chromatin behavior can be recovered accurately as well as the diffusion coefficient. However, the length of a trajectory is often limited due to photobleaching [66]. In principle, the length of a trajectory may be shorter than the time scale to equilibrium . In that case, formula 68 can still be applied to recover the parameter k_c.

6   Reconstruction of the surfaces of cell from the statistics of observed projected random trajectories [39]

A large samples of short trajectories of diffusing single particles acquired on the surface of a cell are projected onto the plane of a confocal microscope. How these projected data can be used to reconstruct the shape of the membrane surface and physical properties of the receptor motion? A general method for the reconstruction of a two-dimensional surface from the statistics of planar projections of many independent trajectories of a diffusion process on the surface introduced in [39] starts with the finding the original stochastic dynamics from the projections. The reconstruction of the geometrical properties and drift field, which represents physical interactions, falls into a class of inverse problems that have been investigated in the context of extraction of shape from shading [90].

However, due to the irregularity of stochastic trajectories, the reconstruction method requires in practice a large set of data points. The reconstruction scheme is illustrated in Fig. 9. The driftless case can be formulated as follows. The planar projection $\boldsymbol{Y}$ of a
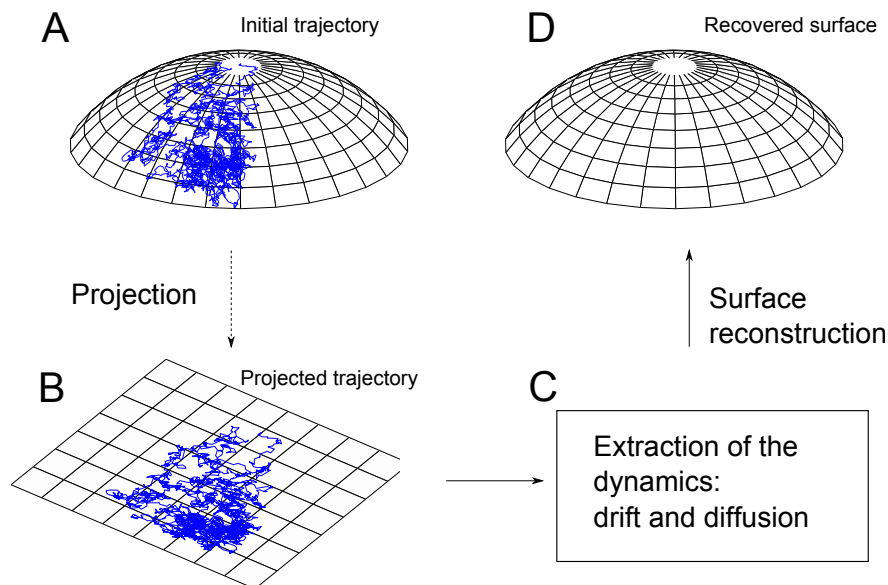

Figure 9: **Reconstruction of a surface from the dynamics of a projected trajectory**. (A) A trajectory of a diffusion process on the unknown surface Σ. (B) A planar projection of the trajectory. (C) Statistical estimates of the drift and the diffusion coefficients of the projected stochastic dynamics. (D) Reconstruction of Σ from the estimated coefficients.

diffusion process $\boldsymbol{X}$, defined on a smooth manifold Σ by the Itô system

$$d\boldsymbol{X} = \sqrt{2}\boldsymbol{B}(\boldsymbol{X})\,d\boldsymbol{W}, \tag{69}$$

where $\boldsymbol{W}$ is Brownian motion on Σ and $\boldsymbol{B}(\boldsymbol{X})$ is the noise matrix, satisfies the equation

$$d\boldsymbol{Y} = \boldsymbol{a}(\boldsymbol{Y}) + \sqrt{2}\tilde{\boldsymbol{B}}(\boldsymbol{Y})\,d\tilde{\boldsymbol{W}}, \tag{70}$$

28

where $\boldsymbol{a}(\boldsymbol{Y})$ and $\tilde{\boldsymbol{B}}(\boldsymbol{Y})$ can be expressed in terms of the derivatives of Σ on the projection plane (e.g., a fixed tangent plane $\prod$). Because Σ is unknown, $\boldsymbol{a}(\boldsymbol{Y})$ and $\tilde{\boldsymbol{B}}(\boldsymbol{Y})$ are the effective drift field and diffusion matrix observed in the plane. Thus the first task is to obtain an analytic expression for these functions. Once explicit expressions for the effective drift field and diffusion tensor in terms of the local surface properties such as the mean curvature are found, the next task is the reconstruction of the surface, which requires solving nonlinear partial differential equations. When the dynamics on the surface contains a drift field, the drift field of the projected dynamics is split into a part that is due to the surface curvature and another due to the original drift.

We present here the main step of the reconstruction for a Brownian motion $\boldsymbol{X}(t)$ on a surface Σ and its planar projection. We assume that Σ has the explicit representation $z = f(x, y)$, where $f(x, y)$ is a sufficiently smooth function defined in a planar domain $\mathcal{D}$ in the (x, y) plane. We can assume that the (x, y) plane is tangent to Σ at the origin $\boldsymbol{0}$, that is, $\boldsymbol{0} \in \Sigma$ and the normal to Σ at $\boldsymbol{0}$ is the z-axis. We fix an orthonormal frame $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$, where $\boldsymbol{k}$ is the unit vector in the direction of the z-axis. Thus

$$\boldsymbol{X}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j} + z(t)\boldsymbol{k}, \quad z(t) = f(x(t), y(t)). \tag{71}$$

For a driftless diffusion $\boldsymbol{X}(t)$ with a symmetric diffusion tensor $\sigma^{i,j}(\boldsymbol{X})$ on Σ with a tangent plane $\prod$ at $\boldsymbol{X}(t) \in \Sigma$, we define the orthonormal frame

$$\boldsymbol{t_x} = \frac{1}{\sqrt{1 + f_x^2}} \begin{pmatrix} 1 \\ 0 \\ f_x \end{pmatrix}, \quad \boldsymbol{t_y} = \frac{1}{\sqrt{(1 + f_x^2)(1 + f_x^2 + f_y^2)}} \begin{pmatrix} f_x f_y \\ -(1 + f_x^2) \\ -f_y \end{pmatrix}$$

$$\boldsymbol{n} = \frac{1}{\sqrt{1 + f_x^2 + f_y^2}} \begin{pmatrix} -f_x \\ -f_y \\ 1 \end{pmatrix}. \tag{72}$$

The orthogonal projection $\tilde{\boldsymbol{X}}(t + \Delta t) = \tilde{x}(t + \Delta t)\boldsymbol{i} + \tilde{y}(t + \Delta t)\boldsymbol{j} + \tilde{z}(t + \Delta t)\boldsymbol{k}$ on $\prod$ of the Brownian motion $\boldsymbol{X}(t + \Delta t) \in \Sigma$ is given in terms of 72 by

$$\tilde{\boldsymbol{X}}(t + \Delta t) = \boldsymbol{X}(t) + \Delta u_1 \boldsymbol{i} + \Delta u_2 \boldsymbol{j} + \Delta u_3 \boldsymbol{k}, \tag{73}$$

where

$$\Delta u_1 = \frac{1}{\sqrt{1 + f_x^2}} \left(\Delta s_1 + \frac{f_x f_y}{\sqrt{1 + f_x^2 + f_y^2}} \Delta s_2 \right),$$

$$\Delta u_2 = -\frac{\sqrt{1 + f_x^2}}{\sqrt{1 + f_x^2 + f_y^2}} \Delta s_2$$

$$\Delta u_3 = \frac{1}{\sqrt{1 + f_x^2}} \left(f_x \Delta s_1 - \frac{f_y}{\sqrt{1 + f_x^2 + f_y^2}} \Delta s_2 \right)$$

and

$$\Delta s_1 = B_{1,1}(\boldsymbol{X}(t)) \Delta w_1 + B_{1,2}(\boldsymbol{X}(t)) \Delta w_2, \quad \Delta s_2 = B_{2,1}(\boldsymbol{X}(t)) \Delta w_1 + B_{2,2}(\boldsymbol{X}(t)) \Delta w_2, \tag{74}$$

29

with $w_1(t), w_2(t)$ independent standard Brownian motions in $\mathbb{R}$. The diffusion tensor $\boldsymbol{\sigma}(\boldsymbol{X})$ is given in terms of the matrix $\boldsymbol{B}(\boldsymbol{X})$ as

$$\boldsymbol{\sigma}(\boldsymbol{X}) = \frac{1}{2}\boldsymbol{B}(\boldsymbol{X})\boldsymbol{B}^T(\boldsymbol{X}). \tag{75}$$

To regain $\boldsymbol{X}(t+\Delta t)$, the point $\tilde{\boldsymbol{X}}(t+\Delta t)$ is projected orthogonally back from $\prod$ to Σ to give

$$\boldsymbol{X}(t+\Delta t) = \tilde{\boldsymbol{X}}(t+\Delta t) + \theta(\Delta t)\boldsymbol{n}(\boldsymbol{X}(t)), \tag{76}$$

where the parameter $\theta(\Delta t) \in \mathbb{R}$ is the solution of the equation

$$z(t+\Delta t) = f(x(t+\Delta t), y(t+\Delta t)) \tag{77}$$

with

$$\begin{cases} x(t+\Delta t) = x(t) + \dfrac{1}{\sqrt{1+f_x^2}}\left(\Delta s_1 + \dfrac{f_x f_y}{\sqrt{1+f_x^2+f_y^2}}\Delta s_2\right) - \dfrac{\theta(\Delta t)f_x}{\sqrt{1+f_x^2+f_y^2}} \\[3mm] y(t+\Delta t) = y(t) - \dfrac{\sqrt{1+f_x^2}}{\sqrt{1+f_x^2+f_y^2}}\Delta s_2 - \dfrac{\theta(\Delta t)f_y}{\sqrt{1+f_x^2+f_y^2}} \\[3mm] z(t+\Delta t) = z(t) + \dfrac{1}{\sqrt{1+f_x^2}}\left(f_x\Delta s_1 - \dfrac{f_y}{\sqrt{1+f_x^2+f_y^2}}\Delta s_2\right) + \dfrac{\theta(\Delta t)}{\sqrt{1+f_x^2+f_y^2}} \end{cases} , \tag{78}$$

where $\theta(\Delta t)$ is such that [39]

$$\theta(\boldsymbol{X}) = \lim_{\Delta t \to 0}\frac{\theta(\Delta t)}{\Delta t} = \frac{f_{xx}}{2(1+f_x^2)}\left[\frac{\sigma_{1,1}(\boldsymbol{X})}{\sqrt{1+f_x^2+f_y^2}} + \frac{\sigma_{2,2}(\boldsymbol{X})f_x^2 f_y^2}{(1+f_x^2+f_y^2)^{3/2}} + \frac{2\sigma_{1,2}(\boldsymbol{X})f_x f_y}{1+f_x^2+f_y^2}\right]$$
$$- \frac{f_{xy}}{1+f_x^2+f_y^2}\left[\sigma_{1,2}(\boldsymbol{X}) + \frac{\sigma_{2,2}(\boldsymbol{X})f_x f_y}{\sqrt{1+f_x^2+f_y^2}}\right] + \frac{f_{yy}}{2}\frac{\sigma_{2,2}(\boldsymbol{X})(1+f_x^2)}{(1+f_x^2+f_y^2)^{3/2}}. \tag{79}$$

In the isotropic case $\boldsymbol{B}(\boldsymbol{X}) = \boldsymbol{I}$, we obtain

$$\theta(\boldsymbol{X}) = \frac{1}{2(1+f_x^2+f_y^2)^{3/2}}\left[(1+f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1+f_x^2)f_{yy}\right], \tag{80}$$

which is the mean curvature at point $\boldsymbol{X}$. Setting

$$\boldsymbol{C}(\boldsymbol{x}) = \begin{pmatrix} \dfrac{1}{\sqrt{1+f_x^2(\boldsymbol{x})}} & \dfrac{f_x(\boldsymbol{x})f_y(\boldsymbol{x})}{\sqrt{1+f_x^2(\boldsymbol{x})}\sqrt{1+f_x^2(\boldsymbol{x})+f_y^2(\boldsymbol{x})}} \\[4mm] 0 & \dfrac{-\sqrt{1+f_x^2(\boldsymbol{x})}}{\sqrt{1+f_x^2(\boldsymbol{x})+f_y^2(\boldsymbol{x})}} \end{pmatrix}, \tag{81}$$

30

we define

$$\tilde{\boldsymbol{B}}(\boldsymbol{X}) = \boldsymbol{C}(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{X}), \quad \tilde{\boldsymbol{\sigma}}(\boldsymbol{X}) = \frac{1}{2}\tilde{\boldsymbol{B}}(\boldsymbol{X})\tilde{\boldsymbol{B}}^{T}(\boldsymbol{x}). \tag{82}$$

Thus the system of stochastic equations in $\mathcal{D}$ that describes the driftless diffusion $z(t) = f(x(t), y(t))$ on Σ, where the projected process on the (x, y) plane, $\boldsymbol{x}(t)$, is

$$d\boldsymbol{x} = \boldsymbol{a}_{\text{geometric}}(\boldsymbol{x})\,dt + \tilde{\boldsymbol{B}}(\boldsymbol{x}, f(\boldsymbol{x}))\,d\boldsymbol{w}, \tag{83}$$

where

$$\boldsymbol{a}_{\text{geometric}}(\boldsymbol{x}) = -\frac{\nabla f(\boldsymbol{x})}{\sqrt{1 + |\nabla f(\boldsymbol{x})|^2}}\,\theta(\boldsymbol{x}, f(\boldsymbol{x})), \tag{84}$$

and $\boldsymbol{w}(t) = (w_1(t), w_2(t))^T$. If $\boldsymbol{B}(\boldsymbol{X}) = \sqrt{2D(\boldsymbol{X})}\boldsymbol{I}$, then

$$\theta(\boldsymbol{X}) = \frac{D(\boldsymbol{X})}{2(1 + f_x^2 + f_y^2)^{3/2}}\left[(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}\right] \tag{85}$$

and 83 becomes

$$
\begin{aligned}
dx_t = &- D(\boldsymbol{X})f_x \left(\frac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{2(1 + |\nabla f|^2)^2}\right) dt \\
&+ \frac{\sqrt{2D(\boldsymbol{X})}}{\sqrt{1 + f_x^2}}\left[dw_1 + \frac{f_x f_y}{\sqrt{1 + |\nabla f|^2}}\,dw_2\right]
\end{aligned}
\tag{86}
$$

$$dy_t = -D(\boldsymbol{X})f_y\left(\frac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{2(1 + f_x^2 + f_y^2)^2}\right) dt - \frac{\sqrt{2D(\boldsymbol{X})(1 + f_x^2)}}{\sqrt{1 + |\nabla f|^2}}\,dw_2.$$

The planar projection of the Brownian motion (or driftless diffusion) has a geometric drift vector and an effective diffusion tensor that are observed. The surface Σ and the diffusion tensor can be recovered from the observed projections.

When the diffusion process $\boldsymbol{X}(t)$ on Σ has a drift vector $\boldsymbol{\beta}(\boldsymbol{X}) = \beta_x \boldsymbol{i} + \beta_y \boldsymbol{j} + \beta_z \boldsymbol{k}$, such that $\beta_z(\boldsymbol{X}) = \beta_x(\boldsymbol{X})f_x(\boldsymbol{x}) + \beta_y(\boldsymbol{X})f_y(\boldsymbol{x})$, the stochastic equations of the projected diffusion contain an additional drift vector

$$d\boldsymbol{x} = \boldsymbol{a}(\boldsymbol{x}, f(\boldsymbol{x}))\,dt + \tilde{\boldsymbol{B}}(\boldsymbol{x}, f(\boldsymbol{x}))\,d\boldsymbol{w}, \tag{87}$$

where

$$\boldsymbol{a}(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x}) + \boldsymbol{a}_{\text{geometric}}(\boldsymbol{x}), \tag{88}$$

with $\boldsymbol{b}(\boldsymbol{x}) = (\beta_x(\boldsymbol{X}), \beta_y(\boldsymbol{X}))$.

31

6.1 Reconstruction procedure

For an isotropic constant diffusion, the projection of Brownian motion from Σ to the (x, y) plane and recover the function $z = f(x, y)$ and the diffusion coefficient D from the effective diffusion tensor

$$\tilde{\sigma}_{1,1} = D\frac{1 + f_y^2}{1 + f_x^2 + f_y^2}, \quad \tilde{\sigma}_{1,2} = -D\frac{f_x f_y}{1 + f_x^2 + f_y^2}, \quad \tilde{\sigma}_{2,2} = D\frac{1 + f_x^2}{1 + f_x^2 + f_y^2} \tag{89}$$

given in 82. The system 89 of three algebraic equations for the unknowns D, f_x, and f_y can be solved explicitly and then Σ is recovered by integrating the gradient $\nabla f(x, y)$ and reduces to solve the single equation new Eikonal equation

$$\frac{\tilde{\sigma}_{2,2}}{\tilde{\sigma}_{1,1}} - 1 = f_x^2 - \frac{\tilde{\sigma}_{2,2}}{4\tilde{\sigma}_{1,1}}\left(-\frac{\tilde{\sigma}_{1,1}}{\tilde{\sigma}_{1,2}}f_x \pm \sqrt{\left(\frac{\tilde{\sigma}_{1,1}}{\tilde{\sigma}_{1,2}}\right)^2 f_x^2 - 4}\right)^2. \tag{90}$$

Before recovering the drift, because

$$\boldsymbol{a}_{\text{geometric}}(\boldsymbol{x}) = -D\left(\begin{array}{c} f_x\left[\dfrac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{2(1 + f_x^2 + f_y^2)^2}\right] \\ f_y\left[\dfrac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{2(1 + f_x^2 + f_y^2)^2}\right] \end{array}\right) \tag{91}$$

equation 91 implies that $\boldsymbol{a}_{\text{geometric}}(\boldsymbol{x}) \parallel \nabla f(\boldsymbol{x})$ so the level curves of $f(x, y)$ can be determined by integrating the geometric drift at any point in the planar projection of Σ. To recover isotropic Brownian motion with constant diffusion coefficient on Σ from its planar projection, we determine the components $\boldsymbol{b}(\boldsymbol{x}) = (b_x(\boldsymbol{x}), b_y(\boldsymbol{x}))^T$ of the drift by solving the system 88, which can be written as

$$b_x(\boldsymbol{x}) - Df_x(\boldsymbol{x})\frac{[1 + f_y^2(\boldsymbol{x})]f_{xx}(\boldsymbol{x}) - 2f_x(\boldsymbol{x})f_y(\boldsymbol{x})f_{xy}(\boldsymbol{x}) + [1 + f_x^2(\boldsymbol{x})]f_{yy}(\boldsymbol{x})}{2[1 + f_x^2(\boldsymbol{x}) + f_y^2(\boldsymbol{x})]^2} = a_1(\boldsymbol{x})$$

$$\tag{92}$$

$$b_y(\boldsymbol{x}) - Df_y(\boldsymbol{x})\frac{[1 + f_y^2(\boldsymbol{x})]f_{xx}(\boldsymbol{x}) - 2f_x(\boldsymbol{x})f_y(\boldsymbol{x})f_{xy}(\boldsymbol{x}) + [1 + f_x^2(\boldsymbol{x})]f_{yy}(\boldsymbol{x})}{2[1 + f_x^2(\boldsymbol{x}) + f_y^2(\boldsymbol{x})]^2} = a_2(\boldsymbol{x}),$$

where now the functions $a_1(\boldsymbol{x})$ and $a_2(\boldsymbol{x})$ are the observed drift components of the projected diffusion. However, numerical reconstructions are not easy and require many points.

7 Conclusion and open questions

This manuscript presented a statistical approach to single particle trajectories, where estimators are constructed to extract first and second order moment properties of the underlying physical processes while removing various sources of noise. The large amount of SPT trajectories allows recovering drift and diffusion tensor at each spatial point at a given resolution.
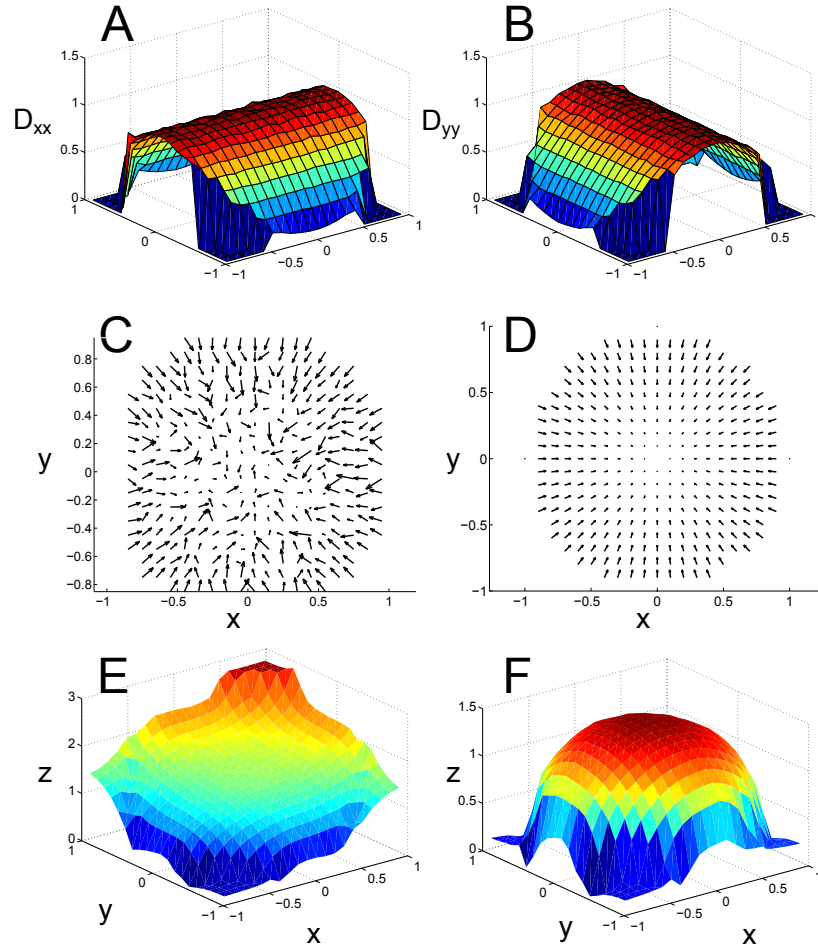
Figure 10: **Projection of spherical Brownian motion.** (**A**) Diffusion coefficient D_{xx} reconstructed using equation 10. (**B**) Diffusion coefficient D_{yy}. (**C**) The drift field $\boldsymbol{a}(\boldsymbol{x}_k) = (a_x(\boldsymbol{x}_k), a_y(\boldsymbol{x}_k))^T$, reconstructed by formula 9. (**D**) The theoretical drift field of the projected Brownian motion on a sphere, obtained from equation 86. (**E**) Reconstruction of the sphere assuming $f_x \geq 0$ and $f_y \geq 0$. (**F**) Reconstruction of the sphere. The parameter values are $D = 1$, $dt = 0.00005$, number of trajectories $N_t = 9,000$, number of points per trajectory $N_s = 1,000$, $\Delta X = 0.1$.

In contrast, computing the MSD along trajectories leads to accumulation of errors due to localization and instrumental noise and also correlates points in the MSD estimation.

Several improvements are expected in the future. For example, diffusion in confined domains has been widely studied in corrugated channels, in domains with obstacles, but the effect of confinement on the observed diffusion coefficient remains difficult to address in particular at the boundary with obstacles. This question is related to the finite time step which is often too large to resolve the molecular size and thus the physical diffusion. Extracting the radius of confinement [49] is very tempting, but in practice, confinement can also be due to an interacting potential. It is a key step to clarify the contribution of both effects and thus the first and second moments should be accurately computed and interpreted. Stochastic simulations and analysis reveal that a random walk has an apparent drift near a wall, which is proportional to $1/\sqrt{\Delta t}$, where Δt is the sampling rate. This relation provides a criteria to differentiate hard-wall confinement from a deterministic drift, for which the estimator is independent of the sampling rate [41]. This criteria could be implemented in confined domains of size larger than the layer $\sqrt{D\Delta t}$. To conclude, a multiscale analysis with various Δt should be used to resolve the presence of a boundary of a subdomain versus the confinement due to a potential well.

In general, how to integrate positional noise and motion blur on the estimate of diffusivity parameters in a confined domain is still an open question. Several studies [37, 55] have now confirmed that regions where trajectories are very confined are due to potential wells. However, extracting the exact properties of the wells remains difficult due to estimation errors of the boundary and the energy constant [63]. Another difficulty in interpreting the SPT trajectories in cell biology is due to the small size of the confinement regions in the range of 100 nm, which is of the order of the localization noise. Resolving hard-wall reflection from interaction potential would be a challenge for statistical methods, but is it really a biophysical issue? Probably hard-walls are a simplification of three dimensional scaffolding proteins at the interface between two- and three- dimensions, that should be addressed using the geometry of the molecules. Measuring and observing particle interactions [10] has recently been applied for G-protein coupled receptors forming dimers or oligomers [43]. But investigating more complicated motions and extracting parameters beside diffusion or interactions (modeled by an Ornstein-Uhlenbeck process) is usually more involved especially for processes driven by molecular aggregation-dissociation. In that case, there is no long-range interaction and thus no potential wells are expected. A statistical analysis is clearly missing.

Most estimators rely on the stationary assumption. Even when a particle motion alternates between free and confined motion [88], it is always possible to decompose the trajectories into sub-time windows, where the motion is clearly identified. This approach of time segmentation was applied to identify transient potential wells that appear and disappear [40], revealing that these structures are generated by dynamical processes, yet to be identified. Indeed, the analysis was performed for sufficiently small windows of time (of the order of seconds) and during that time, potential wells are stationary, but not at a time scale of minutes, where they can evolve [37]. It is always interesting to detect the different time scales involved in subcellular processes. For example, detecting events such as clustering of proteins during the formation of vesicle (clathrin-coated pits [51]) should certainly benefit

34

from a statistical analysis to identify the exact mechanism of clustering.

The method for recovering a stochastic dynamics and the surface using the projection of short stochastic trajectories on a plane is not converging very quickly, because stochastic trajectories are not smooth enough and it is thus difficult to obtain an accurate reconstruction of the surface. A large number of points are necessary to recover surfaces. In dimension two, the sampling number of points is N^2 (compared to N in dimension one). Using simulations of 9,000 trajectories, each composed of 1,000 points on a sphere and projected on the (x, y) plane, the effective diffusion tensor and geometric drift can be evaluated by formulas 9, 10 (Fig. 10A-C). The surface is reconstructed by integrating the gradient, as described in section 6.1 (Fig. 10F). Then the drift vector $\boldsymbol{b}(\boldsymbol{x}) = (b_x(\boldsymbol{x}), b_y(\boldsymbol{x}))^T$ of the projected diffusion is found by subtracting the geometric drift from the measured drift $\boldsymbol{a}(\boldsymbol{x}) = (a_1(\boldsymbol{x}), a_2(\boldsymbol{x}))^T$ (see (92)). This procedure is however not very practical.

Finally, three dimensional SPT trajectories will be soon available [60, 35], and an approximation theory should probably be used to reconstruct the dynamics of the underlying process and the membrane surface or the three dimensional cellular domain, where trajectories evolve. Statistical analysis should be used to improve the spatial resolution based on redundant trajectories, especially when the resolution in the z-direction is much less than in x and y. Probably, the sampling time Δt will not change. Multiscale spatio-temporal data should further improve the reconstruction of cell surfaces.

References

[1] Aït-Sahalia Y, Jacod J, 2014. High-frequency financial econometrics. *Princeton University Press.*

[2] Amitai A, Holcman D, 2013. Polymer model with long-range interactions: Analysis and applications to the chromatin structure.*Phys. Rev. E*, 88(5), p.052604.

[3] Amitai A, Toulouze M, Dubrana K, Holcman D, 2015. Analysis of single locus trajectories for extracting in vivo chromatin tethering interactions. *PLoS Comput Biol*, 11(8), p.e1004433.

[4] Ashley TT, Andersson SB, 2015. Method for simultaneous localization and parameter estimation in particle tracking experiments. *Phys. Rev. E* 92, 052707.

[5] Belopolskaya Ya I, Dalecky Yu L, 2014. *Stochastic equations and differential geometry,* Springer, Dordrecht.

[6] Berglund AJ, 2010. Statistics of camera-based single-particle tracking. *Phys Rev E* 82: 011917.

[7] Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, Davidson MW, Lippincott-Schwartz J, Hess HF, 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science.* 313:1642–1645.

[8] Brush SG, 1976. *The Kind of Motion We Call Heat: A History of the Kinetic Theory of Gases in the Nineteenth Century, Book 1: Physics and the Atomists. Book 2: Statistical Physics and Irreversible Processes.*

[9] Calderon CP, Harris NC, Kiang CH, Cox DD, 2009. Quantifying multiscale noise sources in single-molecule experiments via pathwise statistical inference procedures *J. Phys. Chem. B* 113.

[10] Calderon CP, Weiss LE, Moerner WE, 2015. Robust hypothesis tests for detecting statistical evidence of 2D and 3D interactions in single-molecule measurements, arXiv:1312.6742.

[11] Calderon CP, Bloom K, 2015. Inferring latent states and refining force estimates via hierarchical Dirichlet process modeling in single particle tracking experiments. *PloS one*, 10(9), e0137633.

[12] Calderon CP, 2016. Motion blur filtering: A statistical approach for extracting confinement forces and diffusivity from a single blurred trajectory. *Phys. Rev. E*, 93(5), p.053303.

[13] Chandrasekhar S, 1943. Stochastic problems in physics and astrophysics. *Rev. Mod. Phys.* 15:1-89.

[14] de Chaumont F, et al., 2012. Icy: an open bioimage informatics platform for extended reproducible research, *Nature Methods* 9, 690–696.

[15] Chenouard N, Smal I, de Chaumont F, Maška M, Sbalzarini IF, Gong Y, et al. 2014. Objective comparison of particle tracking methods. *Nature Methods*, 11(3), 281–289.

[16] Destainville N, Salomé L, 2006. Quantification and correction of systematic errors due to detector time-averaging in single-molecule tracking experiments. *Biophys. J.*, 90(2), L17-L19.

[17] M. Doi and S. F. Edwards, 1986. The Theory of Polymer Dynamics (Clarendon Press, Oxford, UK).

[18] Edidin M. 1992. Patches, posts and fences: proteins and plasma membrane domains. *Trends in cell biology*, 2(12), 376-380.

[19] Einstein A, 1956. *Investigations on the Theory of the Brownian Movement.* translated and reprinted by Dover, New York .

[20] Fox E, Sudderth E, Jordan M, Willsky A, 2010. Bayesian Nonparametric Methods for Learning Markov Switching Processes. *IEEE Signal Process Mag* 27: 43–54.

[21] Fox E, Sudderth E, Jordan M, Willsky A, 2011. Bayesian Nonparametric Inference of Switching Dynamic Linear Models. *IEEE Trans Signal Process* 59: 1569–1585.

[22] Fusco D, Accornero N, Lavoie B, Shenoy SM, Blanchard JM, Singer RH, Bertrand E, 2003. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Current Biology* 13:161167.

[23] Friedrich R, Peinke J, 1997. Description of a turbulent cascade by a Fokker-Planck equation. *Phys. Rev. Lett.* 78:863-866.

[24] Gebhardt JC, Suter DM, Roy R, Zhao ZW, Chapman AR, Basu S, Maniatis T, Xie XS, 2013. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nat. Methods* 10, 421–426.

[25] Ghosh RN, Webb WW, 1994. Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. *Biophys. J.* 66, 13011318.

[26] Giannone G, Hosy E, Levet F, Constals A, Schulze K, Sobolevsky AI, Rosconi MP, Gouaux E, Tampe R, Choquet D, Cognet L, 2010. Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. Biophys. J. 99:1303–1310.

[27] Gorelashvili M, Emmert M, Hodeck K, Heinrich D, Amoeboid migration mode adaption in quasi-3D spatial density gradients of varying lattice geometry, *New J. Phys.* 16, 075012 (2014)

[28] Goulian M, Simon SM, 2000. Tracking single proteins within cells. *Biophys. J.*, 79:2188–2198.

[29] Gunzenhäuser J, Wyss R, Manley S, 2014. A quantitative approach to evaluate the impact of fluorescent labeling on membrane-bound HIV-Gag assembly by titration of unlabeled proteins. *PLoS One.* 9(12): e115095.

[30] Hamilton JD, 1994. Time series analysis. p. 385. Princeton: Princeton university press.

[31] Hell SW, Wichmann J, 1994. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters*, 19:780–782.

[32] Hess ST, Gould TJ, Gudheti MV, Maas SA, Mills KD, Zimmerberg J, 2007. Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. *Proc. Natl. Acad. Sci. USA* 104:17370–17375.

[33] Holcman D, 2013. Unraveling novel features hidden in superresolution microscopy data. *Commun Integr Biol.* **6** (3):e23893.

[34] Holcman D, Hoze N, Schuss Z, 2015. Analysis and interpretation of superresolution single-particle trajectories. *Biophys. J.*, 109:1761-1771.

[35] Holden SJ, Pengo T, Meibom KL, Fernandez Fernandez C, Collier J, Manley S, 2014. High throughput 3D super-resolution microscopy reveals Caulobacter crescentus in vivo Z-ring organization. *Proc Natl Acad Sci U S A.* 111:4566–4571.

[36] Hosy E, Martinière A, Choquet D, Maurel C, Luu DT, 2015. Super-resolved and dynamic imaging of membrane proteins in plant cells reveal contrasting kinetic profiles and multiple confinement mechanisms. *ISTIC*, 8(2).

[37] Hoze N, Nair D, Hosy E, Sieben C, Manley S, Herrmann A, Sibarita JB, Choquet D, Holcman D, 2012. Heterogeneity of receptor trafficking and molecular interactions revealed by superresolution analysis of live cell imaging, *Proc. Natl. Acad. Sci. USA* 109: 17052–17057.

[38] Hoze N, Holcman D, 2013. Potential wells for AMPA receptors organized in ring nanodomains, ArXiv q-bio arXiv:1309.3436.

[39] Hoze N, Schuss Z, Holcman D, 2013. Reconstruction of surface and stochastic dynamics from a planar projection of trajectories, *SIAM Journal on Imaging Sciences* 6:2430–2449.

[40] Hoze N, Holcman D, 2014. Residence times of receptors in dendritic spines analyzed by stochastic simulations in empirical domains. *Biophys. J.* 107:3008-17.

[41] Hoze N, Holcman D, 2015. Recovering a stochastic process from super-resolution noisy ensembles of single-particle trajectories. *Phys. Rev. E* 92 (5), 052109.

[42] Kantor Y, Kardar M, 2007. Anomalous diffusion with absorbing boundary. *Phys. Rev. E* 76:061121.

[43] Jonas KC, Huhtaniemi I, Hanyaloglu AC, 2015. Single-molecule resolution of G protein-coupled receptor (GPCR) complexes. *Methods in Cell Biology.*

[44] Karlin S, Taylor HM, *A Second Course on Stochastic Processes* (Academic Press, 1981).

[45] Kepten E, Bronshtein I, Garini Y, 2011. Ergodicity convergence test suggests telomere motion obeys fractional dynamics. *Phys. Rev. E* **83**, 041919.

[46] Knight SC, Xie L, Deng W, et al. 2015. Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*, 350: 823–826.

[47] Komis G, Šamajová O, Ovečka M, Šamaj J, 2015. Super-resolution microscopy in plant cell imaging. *Trends in plant science*, 20:834-843.

[48] Kupferman R, Pavliotis GA, Stuart AM, 2004. Ito versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. *Phys. Rev. E*, 70:036120.

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

[50] Langevin P, 1908. Sur la théorie du mouvement brownien. *CR Acad. Sci. Paris*, 146(530-533), 530.

[51] Li Y, Shang L, Nienhaus GU, 2016. Super-resolution imaging-based single particle tracking reveals dynamics of nanoparticle internalization by live cells. *Nanoscale.*

[52] Liu Z, Lavis LD, Betzig E, 2015. Imaging live-cell dynamics and structure at the single-molecule level. *Molecular cell*, 58:644–659.

[53] Lu HE, MacGillavry HD, Frost NA, Blanpied TA, 2014. Multiple spatial and kinetic subpopulations of CaMKII in spines and dendrites as resolved by single-molecule tracking PALM. *The Journal of Neuroscience*, 34:7600–7610.

[54] Manley S, Gillette JM, Patterson GH, Shroff H, Hess HF, Betzig E, Lippincott-Schwartz J, 2008. High-density mapping of single molecule trajectories with photoactivated localization microscopy. Nat. Methods, 5:155–157.

[55] Masson JB, Dionne P, Salvatico C, Renner M, Specht CG, Triller A, Dahan M, 2014. Mapping the energy and diffusion landscapes of membrane proteins at the cell surface using high-density single-molecule imaging and Bayesian inference: Application to the multiscale dynamics of glycine receptors in the neuronal membrane. Biophys. J. 106:74–83.

[56] Meroz Y, Sokolov IM, 2015. A toolbox for determining subdiffusive mechanisms, *Physics Reports* 573:1–30.

[57] Metzler R, Klafter J, 2000. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics reports.* 339:1-77.

39

[58] Metzler R, Jeon JH, Cherstvy A, Barkai E, 2014. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.* 16:24128-24164.

[59] Michalet X, 2010. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Phys. Rev. E*, 82(4), 041914.

[60] Min J, Holden SJ, Carlini L, Unser M, Manley S, Ye JC, 2014. 3D high-density localization microscopy using hybrid astigmatic/ biplane imaging and sparse image reconstruction. *Biomed Opt Express.* 5: 3935–3948.

[61] Montiel D, Cang H, Yang H, 2006. Quantitative characterization of changes in dynamical behavior for single-particle tracking studies. *The Journal of Physical Chemistry B*, 110:19763–19770.

[62] Nair D, Hosy E, Petersen JD, Constals A, Giannone G, Choquet D, Sibarita JB, 2013. Super-resolution imaging reveals that AMPA receptors inside synapses are dynamically organized in nanodomains regulated by PSD95. *J Neurosci.*, 33:13204-13224.

[63] Parutto P, Hoze N, Holcman D, (preprint).

[64] Qian H, Sheetz MP, Elson EL, 1991. Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. *Biophys. J.*, 60:910.

[65] Rao BP, 1999. Statistical inference for diffusion type processes. Arnold Publishers, London.

[66] Rines DR, Thomann D, Dorn JF, Goodwin P, Sorger PK, 2011. Live cell imaging of yeast. *Cold Spring Harb Protoc.*

[67] Ritchie K, Shan XY, Kondo J, Iwasawa K, Fujiwara T, Kusumi A, 2005. Detection of Non-Brownian Diffusion in the Cell Membrane in Single Molecule Tracking. *Biophys. J.*, 88:2266-2277.

[68] Rossier O, Giannone G, 2015. The journey of integrins and partners in a complex interactions landscape studied by super-resolution microscopy and single protein tracking. *Exp Cell Res.* 343: 28–34.

[69] Rust MJ, Bates M, Zhuang X, 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods*, 3:793-796.

[70] Saffman PG, Delbrück M, 1975. Brownian motion in biological membranes, *Proc. Nat. Acad. Sci., USA*, 72:3111–3113.

[71] Sage D, Kirshner H, Pengo T, Stuurman N, Min J, Manley S, Unser M, 2015. Quantitative evaluation of software packages for single-molecule localization microscopy. *Nat Methods*, 12:717–724.

[72] Savin T, Doyle PS, 2005. Static and Dynamic Errors in Particle Tracking Microrheology. *Biophys. J.*, 88:623-638.

[73] Saxton MJ, 1995. Single-particle tracking: effects of corrals. Biophys. J. 69:389–398.

[74] Saxton MJ, 1997. Single-particle tracking: the distribution of diffusion coefficients. *Biophys. J.*, 72:1744.

[75] Saxton MJ, 2008. Single-particle tracking: connecting the dots. *Nat. Methods*, 5:671-672.

[76] Schneider R, Hosy E, Kohl J, Klueva J, Choquet D, Thomas U, Voigt A, Heine M, 2015. Mobility of calcium channels in the presynaptic membrane. *Neuron*, 86:672-679.

[77] Schuss Z, 2010. *Theory and Applications of Stochastic Processes: An Analytical Approach.* Applied Mathematical Sciences vol.170, Springer NY.

[78] Schuss Z, 2011. *Nonlinear Filtering and Optimal Phase Tracking.* Applied Mathematical Sciences vol. 180, Springer NY.

[79] Schuss Z, 2013. *Brownian Dynamics at Boundaries and Interfaces In Physics, Chemistry, and Biology*, Applied Mathematical Sciences, Volume 186 , Springer NY.

[80] Sengupta P, Van Engelenburg S, Lippincott-Schwartz J, 2012. Visualizing cell structure and function with point-localization superresolution imaging, *Developmental cell*,6: 1092–102.

[81] Shav-Tal Y, Darzacq X, Shenoy SM, Fusco D, Janicki SM, Spector DL, Singer RH. 2004. Dynamics of single mRNPs in nuclei of living cells. *Science* 304:17971800.

[82] Siegert S, Friedrich R, Peinke J, 1998. Analysis of data sets of stochastic systems, *Phys. Lett. A* 243:275–280.

[83] Teh YW, Jordan MI, Beal MJ, Blei DM, 2006. Hierarchical Dirichlet Processes. *J. Am. Stat. Assoc.*, 101: 1566–1581.

[84] Tejedor V, Benichou O, Voituriez R, Jungmann R, Simmel F, Selhuber-Unkel C, Oddershede LB, Metzler R, 2010. Quantitative analysis of single particle trajectories: mean maximal excursion method, *Biophys. J.*, 98: 1364–1372.

[85] Thompson RE, Larson DR, Webb WW, 2002. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.*, 82:2775-83.

[86] Tokuda N, Terada TP, Sasai M, 2012. Dynamical modeling of three-dimensional genome organization in interphase budding yeast. *Biophys. J.* 102:296–304.

[87] Vestergaard CL, Blainey P, Flyvbjerg H, 2014. Optimal estimation of diffusion coefficients from single-particle trajectories. *Phys. Rev. E*, 89:022726.

[88] Vestergaard et al., 2015. Estimation of motility parameters from trajectory data. *The european physical journal special topics.*

[89] Wong H, Marie-Nelly H, Herbert S, Carrivain P, Blanc H, Koszul R, Fabre E, Zimmer C, 2012. A predictive computational model of the dynamic 3D interphase yeast nucleus. *Curr. Biol.*, 22:1881-1890.

[90] Zhang R, Tsai PS, Cryer JE Shah M, 1999. Shape-from-shading: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21:690-706.