A sequential algorithm to detect diffusion switching along intracellular particle trajectories

Vincent Briane, Myriam Vimond, Cesar Augusto Valades Cruz, Antoine Salomon, Christian Wunder, Charles Kervrann

A Sequential Algorithm to Detect Diffusion Switching along Intracellular Particle Trajectories

V. Briane [1,2], M. Vimond [2], C.A. Valades Cruz [3], A. Salomon [1], C. Wunder [3], C. Kervrann [1,*]

[1] Inria, Centre de Rennes Bretagne Atlantique, Serpico Project-Team, Rennes, 35042, France,

[2] CREST (Ensai, Université Bretagne Loire), Bruz, 35170, France

[3] Institut Curie, PLS Research University, Cellular and Chemical Biology, U1143 INSERM / UMR 3666 CNRS, 26 Rue d'Ulm, 75248 Paris Cedex 05, France.

Abstract

Recent advances in molecular biology and fluorescence microscopy imaging have made possible the inference of the dynamics of single molecules in living cells. When we observe a long trajectory (more than 100 points), it is possible that the particle switches mode of motion over time. Then, an issue is to estimate the temporal change-points that is the times at which a change of dynamics occurs.

We propose a non-parametric procedure based on test statistics [Briane et al., 2018] computed on local windows along the trajectory to detect the change-points. This algorithm controls the number of false change-point detections in the case where the trajectory is fully Brownian. A Monte Carlo study is proposed to demonstrate the performances of the method and also to compare the procedure to two competitive algorithms. At the end, we illustrate the efficacy of the method on real data in 2D and 3D, depicting the motion of mRNA complexes –called mRNP– in neuronal dendrites, Galectin-3 endocytosis and trafficking within the cell.

A user-friendly Matlab package containing examples and the code of the simulations used in the paper is available at `http://serpico.rennes.inria.fr/doku.php?id=software:cpanalysis:index`.

1 Introduction

High-resolution fluorescence imaging have made possible the observation of the dynamic behaviour of single molecules. The study of these new data with state-of-the-art statistical and image analysis techniques will allow to decipher the dynamic coordination and organization of interacting molecules, which determine the different functions of the cell, see [Kervrann et al., 2016] and references therein. Indeed living cells are dynamic structures and their constituent particles are constantly moving within and between cellular compartments, domains, and microdomains. Tracking algorithms [Chenouard et al., 2014, Maroulas et al., 2015, Roudot et al., 2017] allow to reconstruct frame by frame the trajectory of a particle over a fixed time period. This trajectory is a time series describing

the behavior of the particle which can change over time. In this context, change-point detection (see [Truong et al., 2018] for a review) is an importance task since change-points are often indicative of a new interaction of the particle with another component of the cell.

In biophysics, the dynamics of these trajectories are usually classified into three groups : subdiffusion, superdiffusion and Brownian motion. The definition of these dynamics is related to the criterion of the mean square displacement, a function of time widely used in biophysics and in cellular imaging to quantify motion. If the MSD function is linear, the trajectory is Brownian [Qian et al., 1991]. If the MSD is sublinear (respectively superlinear) the trajectory is a subdiffusion (respectively a superdiffusion), see [Bressloff, 2014, Chapter 7] and [Metzler and Klafter, 2000]. The biological interpretation of subdiffusion is that the particle is confined in a domain or evolves in an open but crowded area [Berry and Chaté, 2014, Bressloff and Newby, 2013, Section 3]. Superdiffusion occurs when the particle is transported actively via molecular motors along the microtubules [Bressloff and Newby, 2013, Section 4]. Finally, when the particle evolves freely inside the cytosol, it undergoes Brownian motion [Bressloff and Newby, 2013, Section 2].

There exist several examples of proteins whose dynamic mode switches over time. For instance, transmembrane proteins such as AMPA receptors oscillate between subdiffusion and Brownian motion [Hoze et al., 2012]. As another example, a virus invading a cell switches motion between superdiffusion along microtubules and Brownian motion in the cytosol [Lagache et al., 2009]. Here, we focus on the family of Galectins which is known to be involved in important physiological processes such as immune response, cellular development and cancer progression. Galectins act through binding to specific carbohydrates on intracellular and extracellular proteins and lipids. We study trafficking of one member of this lectin family, Galectin-3 (Gal-3). Extracellular Gal-3 is able to bind to the plasma membrane and gets internalized via the formation of vesicular transport carriers [Lakshminarayan et al., 2014]. Those endocytic transport carriers are able to fuse with cellular structures, first with early endosomes, and later on with recycling endosomes or late endosomes. New transport carriers are formed via a fission process, when Gal-3 leaves one of those cellular structures. Movement of Gal-3 containing transport vesicles between those compartments is facilitated with the help of motor proteins. Therefore trafficking of Galectins within the cellular environment can be seen as a constant switching between different types of motion. Estimating the change points and identifying the different modes of motions between these change points is essential to characterize the trafficking behaviour of the Gal-3 trajectories.

Our aim is to estimate the times at which the particle changes from one type of diffusion (superdiffusion, subdiffusion or Brownian motion) to another type of diffusion. This problem can be formulated as a multiple change-point detection problem and has been an active domain of research in biophysics recently. Some approaches use a sliding window to detect a dynamic change in the trajectory. [Simson et al., 1995] and [Meilhac et al., 2006] propose a procedure which computes locally the largest displacement from the starting point of each segment or a confinement index based on the MSD ; below a critical threshold, a short segment will be labeled as a confinement zone. The algorithm of [Bouzigues and Dahan, 2007] is based on the computation of a speed correlation index on sliding windows in order to detect directed motion from Brownian motion. [Arcizet et al., 2008] fit locally the MSD to the power law in order to segment the trajectory into subtrajectories driven by either Brownian motion or directed Brownian motion. The parameters of these three last procedures are heuristically set. Other methods are based on feature parameter classification using a supervised support vector classification [Helmuth et al., 2007] or back-propagation neural network [Dosset et al., 2016]. Furthermore, some approaches assume a parametric model of motion. [Yin et al., 2017] use log-likelihood ratio tests to divide the trajectory into segments that have different diffusion coefficients and/or velocities ; implicitly each segment is assumed to be driven by a

Brownian motion with constant drift (equivalently velocity). [Monnier et al., 2015] propose a Bayesian hidden Markov model. In their model, the particle switches between Brownian motion and Brownian motion with a constant drift. [Türkcan and Masson, 2013] propose an adaptation of their Bayesian decision tree method in order to discriminate Brownian motion and a parametric confined motion by selecting a model which minimizes the BIC criterion ; they use a sliding window approach, computing the BIC on subtrajectories. Among these procedures, only [Helmuth et al., 2007] and [Dosset et al., 2016] consider the three modes of diffusions to segment the trajectory.

In this paper, we derive an algorithm based on the statistical test procedure of [Briane et al., 2018]. Indeed, [Briane et al., 2018] defined a non-parametric three-decision test to distinguish the three types of diffusion and showed that their procedure is consistent under parametric alternatives. Now, the main principle of the proposed algorithm takes its inspiration from the sequential procedure of [Cao and Wu, 2015] and can be summarized as follows. Using a sliding window, we detect a candidate change-point as a point having a high difference between two test statistics which are computed in a local neighbourhood of the point. Then we define clusters of candidate change-points. For each cluster, our change-point estimate is chosen as the change-point candidate having the highest difference between the two local test statistics.

The present paper is organized as follows. In Section 2, we exhibit the inference model and recall the test procedure of [Briane et al., 2018]. In Section 3, we present our sequential procedure for detecting change-points along a trajectory. In Section 4, we assess the performance of the procedure on Monte Carlo simulations and on real 2D and 3D data, depicting the motion of mRNA complexes –called mRNP– in neuronal dendrites, and Galectin-3 endocytosis and movement within the cell.

2 Approach

In this section, we first expose the change-point model. Then, we consider our problem as a statistical test. Finally we present briefly the classification procedure of [Briane et al., 2018] that is used in our procedure.

2.1 Change-Point Model

We observe the successive positions of a single particle in a d-dimensional space ($d = 2$ or $d = 3$) at time $t_0, t_1, \ldots, t_{n-1}$. We suppose that the lag time between two consecutive observations is a constant Δ. The observed trajectory of the particle is,

$$\mathbb{X}_n = \left(X_{t_0}, X_{t_1}, \ldots, X_{t_{n-1}}\right), \tag{1}$$

where $X_{t_k} \in \mathbb{R}^d$ is the position of the particle at time $t_k = t_0 + k\Delta$, $k = 0, \ldots, n-1$. In the sequel, we propose a simple change-point model. More complex models related to the models studied in [Briane et al., 2018] are given in Supplementary Materials (Sec.1). Here we assume that the discrete trajectory is generated by a d-dimensional ($d = 2$ or $d = 3$) diffusion process (X_t) strong solution of the stochastic differential equation (SDE) :

$$dX_t = \mu(X_t, t)dt + \sigma dB_t, \qquad t \in [t_0, t_{n-1}], \tag{2}$$

where B_t denotes a $d-$dimensional Brownian motion ; the unknown parameters of the model are the diffusion coefficient $\sigma > 0$ and the drift term $\mu : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^d$.

Furthermore, we assume that there exists a sequence of N change-points on $[t_0, t_{n-1}]$, namely $t_0 = \tau_0 < \tau_1 < \ldots \tau_N < \tau_{N+1} = t_{n-1}$ such that,

$$\mu(x, t) = \mu_j(x) \text{ for } t \in [\tau_j, \tau_{j+1}), \tag{3}$$

The vector of change-points $(\tau_j)_{j=1\dots N}$ and the number of change-points N are unknown. We also assume that for each τ_j there exists $0 \leq j^\star \leq n$ such that $\tau_j = t_{j^*}$. In other words, the change of motion occurs precisely at a sampling time. The parameters (μ_j, σ) are also unknown. Moreover, we assume that μ_j and μ_{j+1} are associated to different types of diffusion. Let us give two examples. First, suppose the particle evolves freely in the cytosol during $[\tau_j, \tau_{j+1})$ and then is transported actively by molecular motors during $[\tau_{j+1}, \tau_{j+2})$. This situation can be modelled by a switch between Brownian motion and Brownian motion with drift (an example of superdiffusion). Then, in terms of drift, we will have $\mu_j(x) = \mathbf{0}_d$ (Brownian motion) and $\mu_{j+1}(x) = \mathbf{v}$ (Brownian with drift) with $\mathbf{v} \in \mathbb{R}^d$ the velocity of the molecular motor. Secondly, we can imagine that the particle evolves freely in the cytosol during $[\tau_j, \tau_{j+1})$ and then is confined in a domain during $[\tau_{j+1}, \tau_{j+2})$. This situation can be modelled by a switch between Brownian motion and the Ornstein-Uhlenbeck process (an example of subdiffusion). In this case, we have $\mu_{j+1}(x) = -\lambda(x - \theta)$ where $\lambda > 0$ models the restoring force toward the center of the confinement domain θ. These types of switches are studied in Section 4. We note that the aforementioned switches are just examples and that the method presented in the sequel can handle other types of switches.

In what follows, we present a sequential procedure to estimate both the number of change-points N and the vector of change-points $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)$. We emphasize that our algorithm do not rely on any parametric assumptions about the particle motion. In the next section, we present the sequential procedure as a statistical test.

2.2 Test Hypotheses for Change-Point Problem

We follow the idea of the multiple testing procedure of [Cao and Wu, 2015] which is relevant for our problem. In our settings, the global null hypothesis is that (X_t) is Brownian on $[t_0, t_{n-1}]$:

$$H_0 : \mathbb{X}_n \text{ is generated from } (\sigma B_t)_{t_0 \leq t \leq t_{n-1}}. \tag{4}$$

The global alternative hypothesis is that there exist N change-points τ such that sub-trajectories on $[\tau_j, \tau_{j+1})$ and $[\tau_{j+1}, \tau_{j+2})$ are coming from different types of diffusion (Brownian, subdiffusion or superdiffusion) defined by their drift μ_j and μ_{j+1}.

Remark 2.1 *The case where the whole trajectory is subdiffusive or superdiffusive belongs to the alternative hypothesis. In this case there is no change-point ($\boldsymbol{\tau} = \emptyset$).*

[Cao and Wu, 2015] observe the outcome of K statistical hypotheses tests H_{0j} versus H_{1j}, $j = 1 \dots K$. The hypothesis H_{0j} corresponds to the observation of noise at location j while H_{1j} matches with the observation of a true signal. The authors assume that there exist clusters of noise H_{0j} and clusters of true signal H_{1j}. The objective is to detect these clusters. For each location j, they compute two local averages of the p-values associated to the tests before and after location j. If these two local averages are not in the same range of values, j is a potential change-point.

In our context, H_{0j} states that the subtrajectory of chosen size k on $[t_j, t_{j+k}]$ is a Brownian motion. For each time t_j, we adapt the idea of [Cao and Wu, 2015] by considering two test statistics computed on the subtrajectories of size k $(X_{t_{j-k}}, X_{t_{j-k-1}}, \dots, X_{t_j})$ (before t_j) and $(X_{t_j}, X_{t_{j+1}}, \dots, X_{t_{j+k}})$ (after t_j). These test statistics are initially developed in [Briane et al., 2018] and allow to classify the sub-trajectories into three groups of diffusion namely Brownian motion, subdiffusion and superdiffusion. Then, the heuristic is to carry multiple testing on these subtrajectories and define clusters of subtrajectories driven by the same type of diffusion. This way, we obtain a procedure, which controls the type I error of the global null hypothesis (4) at level α. In other words, the procedure will not detect falsely a change point when the trajectory is fully Brownian with probability $1 - \alpha$.

2.3 Trajectory Classification with a Three-Decision Test

If there is no change-point ($\boldsymbol{\tau} = \emptyset$), [Briane et al., 2018] propose a three-decision test procedure in order to decide if the trajectory $\mathbb{X}_n$ is subdiffusion, superdiffusion or Brownian motion. The null hypothesis of their test is H_0 ($\mathbb{X}_n$ is Brownian). The two alternative hypothesises are H_1 ($\mathbb{X}_n$ is subdiffusive) and H_2 ($\mathbb{X}_n$ is superdiffusive). [Briane et al., 2018] use the following test statistic to carry their statistical test :

$$T_n = \frac{\max_{i=1,\dots,n-1} \|X_{t_i} - X_{t_0}\|_2}{(t_{n-1} - t_0)^{1/2} \hat{\sigma}_n(t_0 : t_{n-1})} \tag{5}$$

where

$$\hat{\sigma}_n^2(t_0 : t_{n-1}) = \left((1/(2(n-1)\Delta)) \sum_{j=1}^{n-1} \|X_{t_j} - X_{t_{j-1}}\|_2^2 \right)^{1/2}$$

is a consistent estimate of the diffusion coefficient σ under the null hypothesis. The statistic of the maximum is scaled to have a standardized measure. Consequently, under the hypothesis that $\mathbb{X}_n$ comes from a Brownian motion of diffusion coefficient σ, the distribution of T_n does not depend on σ nor Δ but only on the trajectory size n. Then we can define $q_n(\alpha)$ the quantile of order α of the distribution of T_n under the null hypothesis (which depends only on n).

[Briane et al., 2018] interpret T_n as follows. If T_n is low, it indicates that the process stays close to its initial position during the period $[t_0, t_{n-1}]$ then its is likely that it is a subdiffusion. On contrary if T_n is large, the process goes far away from its starting point, as a superdiffusion does with high probability. Consequently, [Briane et al., 2018] declare that the trajectory $\mathbb{X}_n$ is subdiffusive (H_1) if $T_n < q_n(\alpha/2)$, superdiffusive (H_2) if $T_n > q_n(1 - \alpha/2)$ and Brownian (H_0) otherwise. The type I error of this three-decision test is controlled at level α, and we can approximate the quantiles $q_n(\alpha)$ by Monte Carlo estimates.

[Briane et al., 2018] proved the consistency of the test under different parametric models of subdiffusion and superdiffusion, including models, which are not solution of the SDE (2). Then our algorithm, which relies on (5), can in principle be used even if the motion is not governed by the SDE (2), see Supplementary Materials (Sec.1).

3 Methods

In this section, we present our procedure for detecting change-points along a trajectory. We first present the different steps of the procedure. Then, we focus on the tuning of the parameters of the algorithm.

3.1 A Sequential Procedure for the Change-point Estimation

Our procedure comprises three main steps : detecting the potential change-points ; gathering these potential change-points into clusters ; and estimating the change-point in each cluster assuming a cluster contains a single change-point. The input parameter of our method is the size k of the local window. The choice of k and the other tuning parameters is discussed in Section 3.2 and Section 4.

Step 1 : Detecting the Candidate Change-Points Let $1 \leq k \leq n/2$ be fixed. For all index $i \in \{k, k+1, \dots n-k\}$, we consider two subtrajectories of size k starting at X_{t_i} : the backward trajectory $\mathbb{X}_i^- = (X_{t_i}, X_{t_{i-1}}, \dots X_{t_{i-k}})$ and the forward trajectory $\mathbb{X}_i^+ = (X_{t_i}, X_{t_{i+1}}, \dots X_{t_{i+k}})$. Then we carry the test H_{0i} ($\mathbb{X}_i^-$ is Brownian) versus H_{1i} ($\mathbb{X}_i^-$ is subdiffusive) or H_{2i} ($\mathbb{X}_i^-$ is superdiffusive).

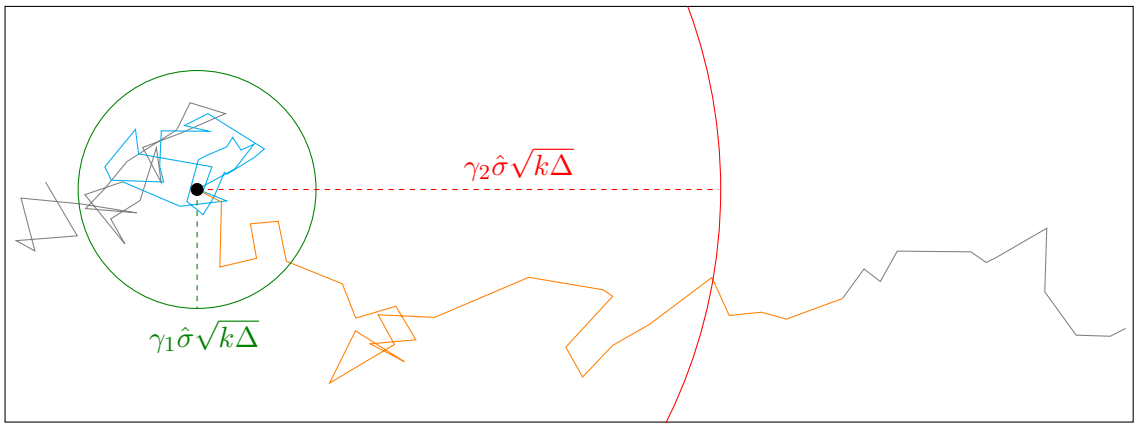


FIGURE 1 – Illustration of the detection step of the candidate change-points with $k = 30$ and $n = 90$. The two-dimensional trajectory is Brownian for $i \in \{0, \dots 60\}$, after the process is a Brownian motion with drift. The change-point $\tau_1 = 60$ is represented by the black dot. The blue (respectively orange) part corresponds to the subtrajectory $\mathbb{X}_{\tau_1}^-$ respectively $\mathbb{X}_{\tau_1}^+$. Comparing the test statistics (6) to (γ_1, γ_2) is equivalent to determine which area, delimited by the red and green circles centred on X_{τ_1}, the particle reaches. The blue subtrajectory stays inside the inner circle : it is classified as subdiffusive. The orange subtrajectory goes outside the outer circle : it is classified as superdiffusive. Then this point is detected as a potential change-point.

Symmetrically, we carry the test for $\mathbb{X}_i^+$. If the outcome of the test is similar for the backward trajectory $\mathbb{X}_i^-$ and the forward trajectory $\mathbb{X}_i^+$ then t_i is unlikely to be a change-point ; hence t_i is not declared as a candidate change-point. Then we carried $2(n-2k+1)$ tests since we test two sets of hypothesises at each time t_i. In this context of multiple testing, we can not use the thresholds $(q_n(\alpha/2), q_n(1-\alpha/2))$ for defining the test as in [Briane et al., 2018] (see also Sec. 2.3) ; in fact such a choice does not control the type I error rate of the global null hypothesis (4) at level α. Instead, we define two thresholds $\gamma_1 < \gamma_2$ to manage the multiple testing issues and to control the type I error rate of the null hypothesis (4) at level α (see Supplementary materials (Sec. 2)).

Now we explain in a pragmatic way how to proceed. We compute the test statistic (5) for the backward and forward trajectory as,

$$B_i = \frac{\max\limits_{j=1,\dots,k} \left\| X_{t_{i-j}} - X_{t_i} \right\|}{(t_i - t_{i-k})^{1/2} \hat{\sigma}(t_i : t_{i-k})}, \quad A_i = \frac{\max\limits_{j=1,\dots,k} \left\| X_{t_{i+j}} - X_{t_i} \right\|}{(t_{i+k} - t_i)^{1/2} \hat{\sigma}(t_i : t_{i+k})}. \tag{6}$$

where $\hat{\sigma}(t_i : t_{i+k})$ (respectively $\hat{\sigma}(t_i : t_{i-k})$) denotes the estimate of the diffusion coefficient from the forward trajectory $\mathbb{X}_i^+$ (respectively the backward trajectory $\mathbb{X}_i^-$). The denomination B_i (respectively A_i) is for *"Before time t_i"* (respectively *"After time t_i"*). Then, we compare the values of A_i and B_i through the statistic Q_i,

$$Q_i = \begin{cases} 0 & \text{if } (A_i, B_i) \in [0, \gamma_1)^2, \\ 0 & \text{if } (A_i, B_i) \in (\gamma_2, +\infty)^2 \\ 0 & \text{if } (A_i, B_i) \in [\gamma_1, \gamma_2]^2 \\ 1 & \text{otherwise.} \end{cases} \tag{7}$$

If $Q_i = 0$ it means that the statistics B_i and A_i belong to the same range of values defined by γ_1 and γ_2. Then, both $\mathbb{X}_i^+$ and $\mathbb{X}_i^-$ are from the same type of diffusion : it is unlikely that t_i is

0 0 0 0 0 0 1 0 0 0 1 1 0 0 **0 0 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1** **1** **1 0 0 0** 0 1 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0.

FIGURE 2 – Illustration of the cluster using the sequence of Q_j computed on the trajectory presented in Figure 3 (b) with $k = 30$, $c = k/2$ and $p = 0.75$. More specifically, the sequence of Q_j correspond to the portion of the trajectory of index $i \in \{140, 200\}$. Hence the detected cluster matches with the second true change-point $\tau_2 = 175$. The bold numbers define a cluster. The true change-point (the boxed number) is contained in this cluster.

a change-point. On the contrary, if $Q_i = 1$ the subtrajectories $\mathbb{X}_i^+$ and $\mathbb{X}_i^-$ are not from the same type of diffusion and t_i is a potential change-point. The detection step is illustrated on a simulated trajectory in Figure 1.

Step 2 : Gathering the Candidate Change-Point into Clusters Originally [Cao and Wu, 2015] propose to define a cluster as $r = k/2$ successive indexes i such that $Q_i \neq 0$. Due to the high level of randomness of the stochastic processes modelling the trajectory, we observe rarely such clusters of size $k/2$ in our case. With this settings, the procedure would fail to detect change-points. A first investigation was to optimize the minimal size r. From an experimental point of view, such optimization would turn out tricky and can lead to overdetection or underdetection depending on the situation. Therefore we propose an other way to build clusters. Even if it is hard to observe successive potential change-points, we argue that a subset of indexes where the concentration of potential change-points is high (even if there are not connected) is likely to contain a true change-point. Then, we define a cluster of potential change-points as a subset of successive indexes $\mathcal{M} = \{i, \dots, i+l\}$ such that, for all subset of size c of successive indexes of $\mathcal{M}$, the proportion of candidate change-points exceeds a proportion p,

$$\sum_{j=m}^{m+c-1} Q_j \geq pc, \quad \forall m = i, \dots, i + l - c + 1. \tag{8}$$

In particular, a cluster needs to have a minimal size of c and if $p < 1$ some points of the clusters may not be candidate change-points (that is $Q_i = 0$). [Cao and Wu, 2015] define a cluster as connected components of candidate change-points whose length is larger than $k/2$. Moreover it is logically necessary that the proportion of candidate change-points is larger than proportion of the other points, which is lead us to choose $p > 1/2$. Note that $p = 1$ is equivalent to build clusters as presented in [Cao and Wu, 2015]. As illustrated on Figure 2, the choice $p = 1$ fails to include the true change-point (the boxed number) into a cluster. In our context, $c = k/2$ and $p = 0.75$ are used, and allow good performances of the algorithm on both simulated data and real data.

Step 3 : Estimating the Change-Point in each Cluster Let $\mathcal{M}_1, \dots \mathcal{M}_{\hat{N}}$ be the clusters defined at the last step. We estimate the change-point of cluster $\mathcal{M}_j$ by the index i for which the difference between B_i and A_i is the highest,

$$\hat{\tau}_j = t_{r_j}, \quad r_j = \max_{i \in \mathcal{M}_j} |B_i - A_i|. \tag{9}$$

Finally, we can summarize the method as follows :

TABLE 1 – Cut-off values (γ_1, γ_2) of Procedure 1 for different trajectory sizes n and window sizes k for dimensions $d = 2, 3$. The cut-off values are estimated with the Monte Carlo Algorithm 1 (Supplementary Materials) using $V = 10\,001$ replications and the default parameters of Procedure 1 $c = k/2$, $p = 0.75$ and with $\alpha = 0.05$.

		$d=2$		$d=3$	
n	k	γ_1	γ_2	γ_1	γ_2
150	20	0.74	3.12	0.96	3.46
150	30	0.79	3.09	1.01	3.37
150	40	0.81	3.05	1.03	3.35
300	20	0.71	3.29	0.91	3.60
300	30	0.74	3.28	0.95	3.59
300	40	0.75	3.27	0.96	3.59

Procedure 1

1. *For a chosen window size k compute B_i and A_i in (6) for $i = k, \dots, n-k$.*
2. *For prespecified cut-off values $\gamma_1 < \gamma_2$ compute Q_i from (7).*
3. *Decompose $\{k, \dots, n-k\} = W_0 \cup W_1$ where $i \in W_0$ if $Q_i = 0$ and $i \in W_1$ if $Q_i = 1$.*
4. *Gather the potential change-points, that is points t_i such that $Q_i = 1$, into clusters $\mathcal{M}_1, \dots, \mathcal{M}_{\hat{N}}$ satisfying Equation (8).*
5. *For each $\mathcal{M}_j$ let $r_j = \max_{i \in \mathcal{M}_j} |B_i - A_i|$ then $\hat{\tau}_j = t_{r_j}$.*

3.2 Tuning parameters

The parameters of Procedure 1 are the size of the window k, the parameters defining the clusters c and p and finally the cut-off-values (γ_1, γ_2). From our experiments, we recommend to set $c = k/2$ and $p = 0.75$. The cut-off values (γ_1, γ_2) are automatically computed and depend on the other parameters of the procedure that is (k, c, p), the trajectory size n and the dimension d. More specifically, as already mentioned in Section 3, (γ_1, γ_2) are chosen in order to control the type I error under the global null hypothesis (4) at level $\alpha \in (0, 1)$. In other words, we choose (γ_1, γ_2) such that, when the trajectory is fully Brownian (without any change-point), we have probability α to detect falsely a change-point. We note that we use the standard value $\alpha = 0.05$ and do not consider α as a parameter of our procedure. Values of (γ_1, γ_2) for different trajectory sizes n and dimensions $d = 2, 3$ are given in Table 1. There are computed with the Monte Carlo algorithm 1, see Supplementary Materials. The way to choose the cut-off values (γ_1, γ_2) is carefully explained and mathematically justified in the Supplementary Materials (Sec. 2). At the end, the window size k is the only input parameter to be set by the user. The influence of parameter k is discussed in Section 4.

4 Results

In this section, we first conduct a Monte Carlo study of the procedure in order to evaluate and compare the performance of the method according to two scenarios, see Table 2 and Figure 3, in the the two-dimensional case. Then, we analyse two-dimensional data depicting long range transport of mRNAs [Monnier et al., 2015]. In the Supplementary Materials (Sec. 2), we compare our method

TABLE 2 – Simulation scenarios for the Monte Carlo study.

Times	Scenario 1	Scenario 2
$[1, 100]$	Brownian	Brownian
$[101, 175]$	Brownian with drift	Ornstein-Uhlenbeck
$[176, 300]$	Brownian	Brownian

to two parametric competing procedures proposed respectively by [Türkcan and Masson, 2013] and [Monnier et al., 2015].

4.1 Simulation Settings

First, we expose the parametric models illustrating subdiffusion and superdiffusion that we use in our simulation. Subdiffusion is modelled by the Ornstein-Uhlenbeck process :

$$dX_t^i = -\lambda(X_t^i - \theta_i)dt + \sigma dB_t^i, \qquad i = 1, 2, \tag{10}$$

where $\lambda > 0$ models the restoring force toward the equilibrium point $\theta = (\theta_1, \theta_2)$; $\sigma > 0$ is the diffusion coefficient. For modelling superdiffusion we use the Brownian motion with drift solution of the SDE :

$$dX_t^i = (v/\sqrt{2})dt + \sigma dB_t^i, \qquad i = 1, 2, \tag{11}$$

where $\sigma > 0$ is the diffusion coefficient and $v > 0$. Then the constant drift $\mathbf{v} = (v, v)/\sqrt{2}$ verifies $\|\mathbf{v}\| = v$.

We simulate trajectories of size $n = 300$ with two change-points occurring at $\tau_1 = 100$ and $\tau_2 = 175$. We study two different scenarios, see Table 2. We set $\sigma = 1$ for the diffusion coefficient of all the processes and $\Delta = 1$ for the step of time. For the Ornstein-Uhlenbeck process (10), we define the equilibrium point as $\theta = X_{\tau_1}$ where X_{τ_1} is the position of the particle at τ_1.

For each scenario, we compute the performances of our procedure for different values of the parameters v (for the Brownian motion with drift) and λ (for the Ornstein-Uhlenbeck process). We run our algorithm with different window sizes k for each situation. We assess the performances of our algorithm with respect to two criteria : the number of change-points detected and the location of these change-points. The change-point location is assessed only on the trajectories for which we detect the right number of change-points that is $N = 2$. We compute the average and standard deviation of the locations. We analyse the results of the simulation on the different scenarios in the next paragraphs.

4.2 Performance of the Procedure

Scenario 1 Table 3 gives us the results associated to Scenario 1 (see Table 2). We can see clearly that, as $\|v\|$ increases, the performance of the method increases with respect to both criteria. For a given window size k, we observe that, on the one hand, the proportion of trajectories for which we detect the right number of change-point ($\hat{N} - N = 0$) tends to 1 as v increases ; on the other hand, given $\hat{N} - N = 0$, the bias and the variance of the estimated change-point decrease to 0 as v increases. We also notice that for the window size $k = 20$, the performance of the algorithm is lower than for $k = 30$ and 40 except when $\|v\| = 2$. As the size of the window is too low, it is hard for the algorithm to detect a Brownian motion with drift with a low drift norm. In particular, when $\|v\| = 0.6$, it does not detect any change-point in most cases ; we note that $\hat{N} - N = -2$ for 42.2% of the trajectories. However, when the drift norm is high, a low window size performs as well as the

TABLE 3 – Performance of the Procedure 1 for Scenario 1 (see Table 2) for different window sizes k and different values of the drift v. The computations are based on 1 001 simulated trajectories from Scenario 1. We compute the proportions of trajectories with $\hat{N} - N = -2$, $\hat{N} - N = \pm 1$, $\hat{N} - N = 0$ and $\hat{N} - N \geq 2$. The column τ_1 (respectively τ_2) gives the empirical average of the first (respectively second) detected change point on 300 trajectories among which we detect the right number of change points ($\hat{N} - N = 0$). The number in brackets is the empirical standard deviation of the estimate of τ_1 and τ_2 computed on these 300 trajectories.

		$\hat{N} - N$					τ_1		τ_2	
v	k	-2	-1	0	1	≥ 2				
0.6	20	42.2	14.1	34.7	5.7	3.3	126.3	(23.7)	153.7	(23.6)
0.6	30	20.9	16.9	55.9	5.5	0.8	115.0	(17.8)	162.8	(18.4)
0.6	40	11.8	18.6	67.6	1.8	0.2	109.4	(15.5)	168.2	(15.1)
0.8	20	6.5	12.9	54.4	17.3	8.9	117.4	(16.7)	157.5	(18.5)
0.8	30	1.6	6.5	84.1	6.3	1.5	107.3	(11.6)	170.1	(14.1)
0.8	40	0.3	4.1	93.2	2.3	0.1	104.7	(9.7)	172.4	(10.0)
1	20	0.2	3.7	63.2	21.6	11.3	108.3	(12.1)	168.2	(13.2)
1	30	0.0	1.9	93.8	3.5	0.8	102.9	(5.9)	173.9	(6.5)
1	40	0.0	0.1	97.7	1.9	0.3	103.2	(6.6)	174.5	(7.3)
2	20	0.0	0.0	96.6	2.4	1.0	101.4	(2.1)	176.0	(2.3)
2	30	0.0	0.3	96.8	2.5	0.4	101.2	(3.4)	175.9	(2.6)
2	40	0.0	0.1	99.2	0.5	0.2	101.5	(2.8)	175.8	(2.9)

larger ones (see the case $\|v\| = 2$). It performs even better if the change-points τ_1, τ_2 are closer. In this case, a large window tends to mix up the two change-points and consequently find only one. We can summarize this as follows : a large window size enables to detect well the change-points associated with a small drift v if the change-points are significantly separated while a small window is able to distinguish two close change-points if the drift v is large enough.

Scenario 2 Table 4 gives us the results associated to Scenario 2 (see Table 2). As in Scenario 1, for a window size $k = 20$ the performance of the algorithm increases as λ increases. However, it does not behave the same way if the window size is 30 or 40. For $k = 30$, the performance increases from $\lambda = 1$ to $\lambda = 2$ but remains the same for larger values of λ. For the window size $k = 40$, the proportion of trajectories with the correct number of detected change-points dramatically drops from 83.6% with $\lambda = 1$ to 54.1% for $\lambda = 4$. At the same time, the proportion of trajectories with $\hat{N} - N = -1$ increases. It means that when λ becomes too high the algorithm mixes up the two change-points and finds only one. As λ is high (clear subdiffusion), we detect a potential change-point very early in the trajectory : as soon as few points of the forward subtrajectory $\mathbb{X}_i^+$ enter in the subdiffusion regime ($t \geq \tau_1$) we classify it as subdiffusive. For example, if λ is big enough we can suppose that the subtrajectory of size k $\mathbb{X}_i^+ = (X_{t_i}, \dots X_{\tau_1}, X_{\tau_1+1}, X_{\tau_1+2})$ will be classified as subdiffusive with only three points in the subdiffusive regime. Then, we get a long sequence of potential change-points. But as k is large, the forward subtrajectory has already reached the second change-point τ_2. Consequently, it begins to detect potential change-points corresponding to the second change-point τ_2. As there is a single cluster of potential change-points, the algorithm only detects one change-point instead of the two expected. From our simulations, we observe that the change-point detected is either close to τ_1 or τ_2 : it estimated correctly one change-point out of the two real change-points. The idea is that, in a way, a large λ (a very clear subdiffusion) makes the two change-points get closer artificially. Then, a large window can not separate them.

TABLE 4 – Performance of the Procedure 1 for Scenario 2 (see Table 2) for different window sizes k and different values of parameter λ. We use the same simulation protocol as for Table 3.

		$\hat{N}-N$								
λ	k	-2	-1	0	1	≥ 2	τ_1		τ_2	
1	20	18.1	44.1	31.8	5.3	0.7	109.9	(20.7)	167.8	(17.9)
1	30	0.8	16.3	78.3	4.2	0.4	104.9	(8.7)	169.9	(9.3)
1	40	0.0	13.2	83.6	3.0	0.2	105.6	(10.8)	170.4	(11.6)
2	20	3.1	22.6	68.1	5.5	0.7	106.4	(8.5)	170.2	(8.1)
2	30	0.1	6.5	89.4	3.4	0.6	107.5	(8.7)	169.1	(8.2)
2	40	0.0	21.2	77.0	1.6	0.2	108.0	(12.7)	169.1	(12.8)
3	20	1.1	17.1	74.8	5.9	1.1	106.3	(5.6)	170.1	(7.9)
3	30	0.0	5.7	90.2	3.2	0.9	108.7	(8.8)	167.6	(8.7)
3	40	0.1	32.1	64.8	2.5	0.5	109.3	(12.9)	166.4	(13.5)
4	20	0.6	12.2	79.4	6.8	1.0	107.2	(6.5)	169.9	(8.4)
4	30	0.0	6.5	89.7	3.1	0.7	109.6	(9.6)	166.5	(9.2)
4	40	0.0	44.3	54.1	1.4	0.2	111.5	(13.3)	166.0	(13.3)

TABLE 5 – Proportions of trajectories (among the trajectories with $\hat{N}=N$) for which subtrajectories are correctly labelled, in scenario 1 and 2. The change points are detected and estimated with Procedure 1. The subtrajectories are labelled using the three-decision test of [Briane et al., 2018] at level 5%. Columns 2 and 3 (respectively 4 and 5) correspond to scenario 1 (respectively scenario 2). For example, in scenario 1 with $v=0.6$, when we use a window of size $k=20$, 73.7% of the trajectories for which we detect $N=2$ change points are labelled as Brownian on $[t_0,\hat{\tau}_1]$, superdiffusive on $[\hat{\tau}_1,\hat{\tau}_2]$ and again Brownian on $[\hat{\tau}_2,t_n]$.

Scenario 1			Scenario 2	
k	v	% right label	λ	% right label
20	0.6	73.7	1	83.0
30	0.6	82.3	1	89.3
40	0.6	86.0	1	85.0
20	0.8	74.7	2	90.7
30	0.8	87.7	2	89.0
40	0.8	88.7	2	88.7
20	1.0	82.0	3	89.3
30	1.0	86.3	3	87.3
40	1.0	88.7	3	86.7
20	2.0	89.7	4	88.7
30	2.0	90.0	4	85.0
40	2.0	90.3	4	84.7

4.3 Post Analysis

Once the change-points are estimated, we can label the type of diffusion on each subtrajectory $(X_{\hat{\tau}_j},\dots,X_{\hat{\tau}_{j+1}})$ delimited by the estimated change-points $\hat{\tau}_j$. For example, we can use the three-decision test of [Briane et al., 2018] to label each subtrajectory (see Sec. 2.2). For a given scenario, value of parameter (λ or v) and window size k, we assessed the labelling of the subtrajectories when the label is given by the three-decision test of [Briane et al., 2018]. Results and details of the evaluation are given in Table 5.

4.4 Real Data

We demonstrate the interest of our algorithm on two different sets of data in two and three dimensions.

2D case : long-range transport of mRNAs. We use the same data [1] as [Monnier et al., 2015] depicting long-range transport of mRNAs in complex with mRNA-binding proteins (mRNPs). In live neuronal cultures, endogenous β-actin mRNP particles alternate between Brownian motion and active transport. In case of active transport (superdiffusion), the particle is driven by molecular motors along microtubule tracks in the neuronal dendrites. The microscopic sequence was obtained using mRNA fluorescence labelling techniques. More specifically, in the experiment of [Monnier et al., 2015], the MS2 bacteriophage capsid protein was tagged with a GFP (Green Fluorescence Protein). As the MS2 bacteriophage capsid protein binds to β-actin mRNP, it allows to track mRNP.

The time resolution of the sequence is $\Delta = 0.1s$. The space resolution is not given but when the Brownian motion with drift is chosen, [Monnier et al., 2015] find a drift parameter with order of magnitude of $1\mu m.s^{-1}$. We set the parameter $K = 2$ for the method of [Monnier et al., 2015], see Supplementary Materials (Sec. 2). In this case, the model 3 (two Brownian motion with different diffusion coefficients) is selected by the method. Then, from our point of view, there is no change of dynamic. We note that we run 100 times the algorithm and did not get the same outcome each time. It is due to the fact that the inference is based on a Monte-Carlo Markov chains (MCMC) algorithm for computing the *a posteriori* estimates. Consequently, the selected model was not the same every times (92 times model 3, 7 times model 4, 1 time model 5). Then, the MCMC algorithm can show some problems of stability giving some contrary outcomes from one run to another.

In Figure 4, we show our results for two window sizes $k = 10$ and $k = 15$. We do not detect any change-point for larger windows. With both window sizes, we detect approximately the same portion of the trajectory as superdiffusive. With the window size $k = 15$, we also detect a subdiffusive part in the trajectory. With $k = 15$, we note that Procedure 1 detects four change- points $\boldsymbol{\tau} = (26, 62, 75, 283)$ while the three-decision test of [Briane et al., 2018] classifies the consecutive subtrajectories $(X_1, \ldots, X_{26})$ and $(X_{26}, \ldots, X_{62})$ both as Brownian. Then we can wonder if the first detected change-point $\hat{\tau}_1 = 26$ is really a true change-point. Furthermore, as we also do not detect the change-point with $k = 10$, we can consider the change-point $\hat{\tau}_1 = 26$ as a false detection. We explain quickly why this issue do no question the validity of Procedure 1. We emphasize that Procedure 1 detects change-points thanks to a local analysis of the trajectory due to the sliding window k. On the other hand, the post-labelling method is not based on a local approach as it classifies the whole subtrajectories defined by the detected change-points. Then, it is completely possible that the results of Procedure 1 and the post-labelling method do not agree.

3D case : Galectin-3 proteins in HeLa cells. We study the movements of Galectin-3 proteins when entering the cell via endocytosis [Lakshminarayan et al., 2014], followed by active transport via endocytic vesicles within the cytosol. Fluorescence images (199 time points, pixel size in x, y 104 nm, distance in z 369 nm) were acquired using a lattice light sheet microscope (LLSM, 3i, Denver, USA) [Chen et al., 2014]. Image volumes of Galactosyltransferase-EGFP (Golgi apparatus for structural orientation) and labeled Gal-3 were recorded every 4.55 seconds, using 20 ms exposure time. 3D data sets were deskewed to account for the 32.8° angle of the detection objective. Subsequent deconvolution was performed using the Richardson-Lucy algorithm. For the segmentation

1. http ://hmm-bayes.org/about/

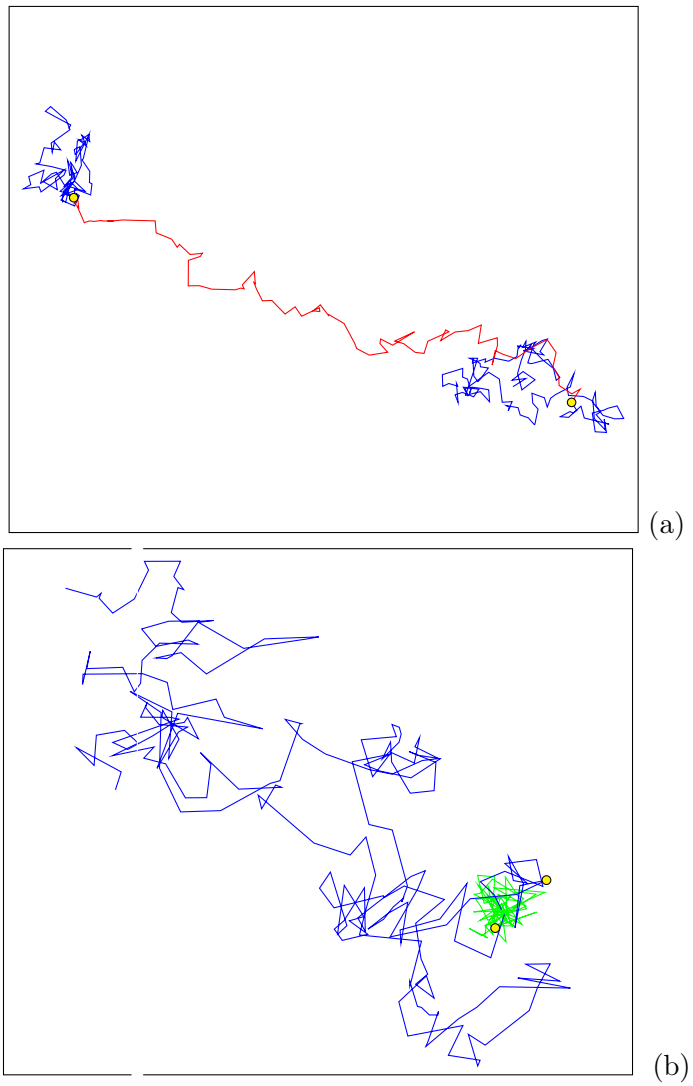
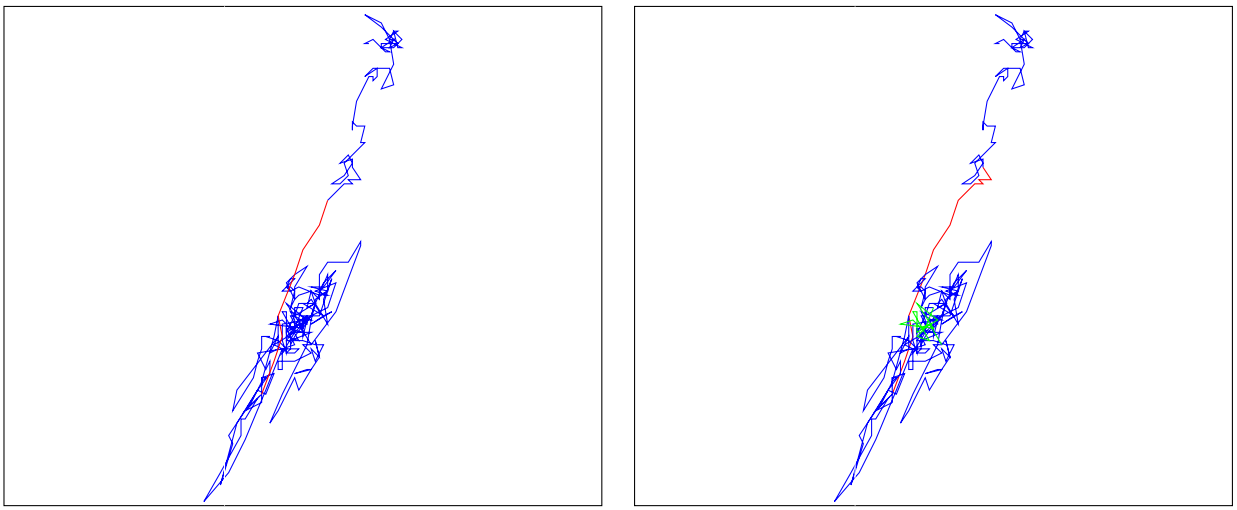


FIGURE 3 – Simulated trajectories. Figure (a), trajectory from Scenario 1 with $v = 0.8$. We detect $\hat{N} = 2$ change-points $\hat{\tau}_1 = 89$ and $\hat{\tau}_2 = 172$ with a window of size 30. Figure (b) trajectory from Scenario 2 with $\lambda = 1$. We detect $\hat{N} = 2$ change-points $\hat{\tau}_1 = 87$ and $\hat{\tau}_2 = 165$ with a window of size 30. The locations of the change-points $X_{\hat{\tau}_1}$ and $X_{\hat{\tau}_2}$ are shown as yellow dots on the trajectories.

algorithm, Golgi and cell masks were defined based on GalT-EGFP images and auto fluorescence signal. The cell contour for each time-point and z-plane were calculated using Matlab Image Processing Toolbox. A similar analysis was performed to estimate the Golgi contour.

The trajectories were obtained from the sequence of images thanks to the Icy tracker used with the default parameters [Chenouard et al., 2013], available on the Icy software (`http:www.icy.org`). Then we used a preprocessing step selecting the trajectories with at least 25 distinct positions and that stop at the same position less than $K = \lfloor n/10 \rfloor$ times (with n the length of the trajectory). We end up with 408 trajectories with mean length 89. Then we run our algorithm with $k = 15$. The computational time of our algorithm is $5.25s$ using four cores working in parallel on a Mac Book Pro version 10.10.1 equipped with 2.8 GHz Intel Core i7, 16 Gb of RAM. In Figure 5, we represent the segmented trajectories in one HeLa cell.



$k = 10$ $k = 15$

FIGURE 4 – Change-point detection on a two-dimensional trajectory depicting neuronal mRNPs. The blue parts correspond to Brownian portions of the trajectory, red part to superdiffusive portions, green part to the subdiffusive portion. The detected change-points are $\boldsymbol{\tau} = (67, 75)$ for $k = 10$ alternating between Brownian, superdiffusion and Brownian. The detected change-points are $\boldsymbol{\tau} = (26, 62, 75, 283)$ for $k = 15$ alternating between Brownian motion, superdiffusion, Brownian motion and finally subdiffusion.

We observed subdiffusive motion of Gal-3 (Figure 5, green), which is characteristic for movement of molecules within a confined compartment, like very early endocytic uptake structures or different endosomes. Trafficking between those compartments is often facilitated through molecular motor proteins, characterized by a fast motion, which we identified as superdiffusive trajectories (in red). Another group of Gal-3 tracks is characterized by Brownian motion within the plasma membrane (blue), indicating an early stage of endocytosis, which is characterized by a slow motion of carbohydrate trapped Gal-3. This event is just preceding an endocytic event of internalization from the plasma membrane. The overall picture of trajectories can be described as confined in small regions and Brownian motions of particles (blue and green), which are interconnected by active transport events (red).

5 Discussion

We proposed a non-parametric algorithm to detect the change-points along a particle trajectory. These change-points are defined as the times at which the particle switches between three modes of motion, namely Brownian motion, subdiffusion and superdiffusion. These types of processes are extensively used in the biophysics literature [Bressloff, 2014] [Berry and Chaté, 2014]. When the trajectory is fully Brownian (our null hypothesis H_0), we control the probability to detect a false change-point at level α. Our algorithm is user-friendly as there is only one parameter to tune, namely the window size k. Future work will involve the development of a multi-scale approach to automatically select the optimal window size k.

We compared our method to [Türkcan and Masson, 2013] and [Monnier et al., 2015]. First, none of these methods propose to distinguish the three types of diffusion. Secondly, we show more reliable

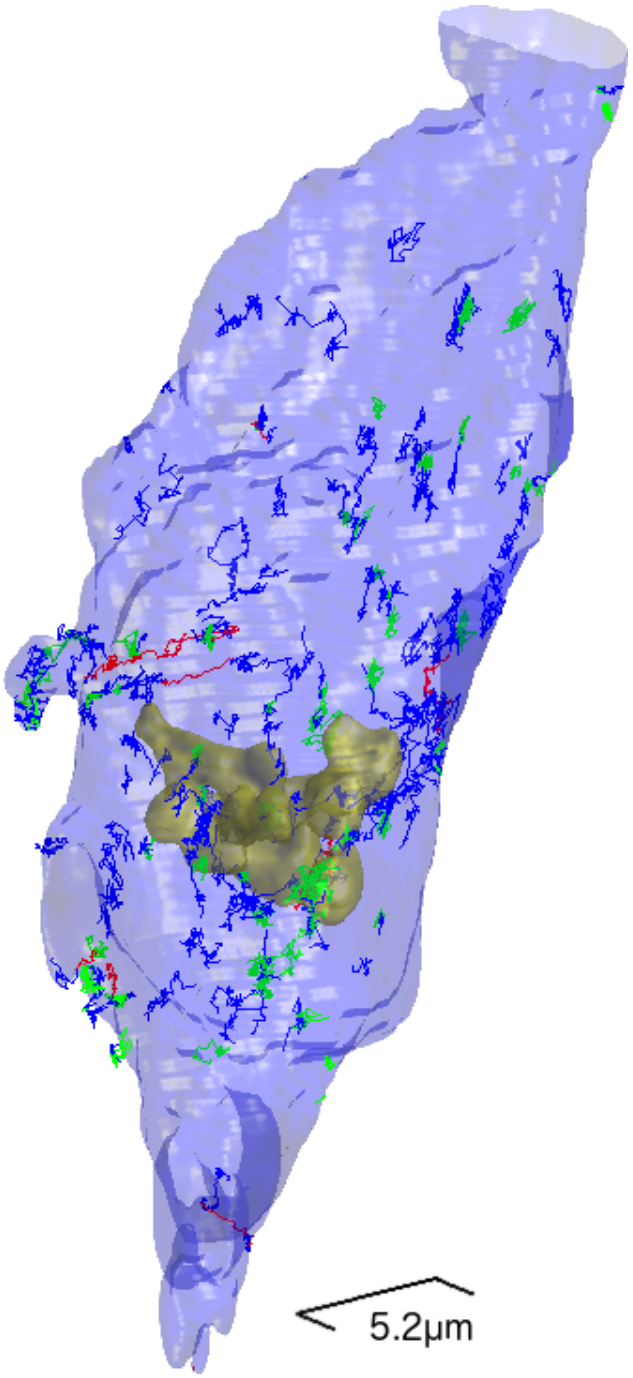


FIGURE 5 – Change point detection over a set of three-dimensional trajectories of Galectin-3 proteins observed in HeLa cells. The blue parts correspond to Brownian portions of the trajectory, red parts to superdiffusive portions, green parts to the subdiffusive portion. In light blue we plot the cell membrane and in yellow the Golgi apparatus for structural orientation.

results than both of the methods on our simulations. Thirdly, our method is much faster than the two others which is an advantage when dealing with a large numbers of trajectories. We also considered real data depicting neuronal mRNPs (mRNAs in complex with mRNA-binding), and another very complex biological example, Gal-3 trafficking from the plasma membrane to different cellular compartments. The analysis of multiple Gal-3 trajectories demonstrates nicely that there is not one typical signature. Biological trafficking events are very multifaceted. The presented algorithm is capable of identifying and characterizing the multistep biological movement, switching several times between subdiffusive, superdiffusive and Brownian motion.

Acknowledgement

The authors greatly acknowledge Ludovic Leconte from the Cell and Tissue Imaging facility (PICT-IBiSA), Institut Curie, member of the National Research Infrastructure France-BioImaging (ANR10-INBS-04).

Author contributions

V. Briane : Conceptualization, Methodology, Modeling, Software, Formal analysis, Investigation, Validation, Visualization, Writing-original draft ; M. Vimond : Formal analysis, Supervision, Investigation, Methodology, Writing-review and editing. C.A. Valades Cruz : Microscopy, Image processing ; A. Salomon : Image processing, Visualization ; C. Wunder : Sample preparation, Microscopy ; C. Kervrann : Formal analysis, Supervision, Investigation, Methodology, Writing-review and editing.

Funding

Funding was provided by Inria Rennes and CREST-Ensai- Université Bretagne Loire. This work was also supported by the French National Research Agency (France-BioImaging infrastructure–ANR-10-INBS-04, DALLISH–ANR-16-CE23-0005). The lattice light sheet microscope was financed by : Labex DCBiol, Labex CellTissPhyBio, HFSP (RGP0029/2014) and ANR-11-LBX-0038.

Références

[Arcizet et al., 2008] Arcizet, D., Meier, B., Sackmann, E., Rädler, J. O., and Heinrich, D. (2008). Temporal analysis of active and passive transport in living cells. *Phys. Review Lett.*, 101(24) :248103.

[Berry and Chaté, 2014] Berry, H. and Chaté, H. (2014). Anomalous diffusion due to hindering by mobile obstacles undergoing brownian motion or ornstein-ulhenbeck processes. *Phys. Review E*, 89(2) :022708.

[Bouzigues and Dahan, 2007] Bouzigues, C. and Dahan, M. (2007). Transient directed motions of gabaa receptors in growth cones detected by a speed correlation index. *Biophysical J.*, 92(2) :654–660.

[Bressloff and Newby, 2013] Bressloff, P. and Newby, J. (2013). Stochastic models of intracellular transport. *Reviews of Modern Physics*, 85(1) :135.

[Bressloff, 2014] Bressloff, P. C. (2014). *Stochastic Processes in Cell Biology*, volume 41. Springer.

[Briane et al., 2018] Briane, V., Kervrann, C., and Vimond, M. (2018). Statistical analysis of particle trajectories in living cells. *Phys. Review E*, 97(6) :062121.

[Cao and Wu, 2015] Cao, H. and Wu, W. B. (2015). Changepoint estimation : another look at multiple testing problems. *Biometrika*, page asv031.

[Chen et al., 2014] Chen, B., Legant, W., Wang, Q., Shao, L., Milkie, D., Davidson, M., Janetopoulos, C., Wu, X., Hammer, J. r., Liu, Z., English, B., Mimori-Kiyosue, Y., Romero, D., Ritter, A., Lippincott-Schwartz, J., Fritz-Laylin, L., Mullins, R., Mitchell, D.M. snd Bembenek, J., Reymann, A., Bohme, R., Grill, S., Wang, J., Seydoux, G., Tulu, U., Kiehart, D., and Betzig, E. (2014). Lattice light-sheet microscopy : imaging molecules to embryos at high spatiotemporal resolution. *Science*, 346 :1257998.

[Chenouard et al., 2013] Chenouard, N., Bloch, I., and Olivo-Marin, J.-C. (2013). Multiple hypothesis tracking for cluttered biological image sequences. *IEEE Trans. Patt. Anal. Mach. Intell.*, 35(11) :2736–27–50.

[Chenouard et al., 2014] Chenouard, N., Smal, I., De Chaumont, F., Maška, M., Sbalzarini, I. F., Gong, Y., Cardinale, J., Carthel, C., Coraluppi, S., Winter, M., et al. (2014). Objective comparison of particle tracking methods. *Nature Meth.*, 11(3) :281.

[Dosset et al., 2016] Dosset, P., Rassam, P., Fernandez, L., Espenel, C., Rubinstein, E., Margeat, E., and Milhiet, P.-E. (2016). Automatic detection of diffusion modes within biological membranes using back-propagation neural network. *BMC Bioinformatics*, 17(1) :197.

[Helmuth et al., 2007] Helmuth, J. A., Burckhardt, C. J., Koumoutsakos, P., Greber, U. F., and Sbalzarini, I. F. (2007). A novel supervised trajectory segmentation algorithm identifies distinct types of human adenovirus motion in host cells. *J. Struct. Biol.*, 159(3) :347–358.

[Hoze et al., 2012] Hoze, N., Nair, D., Hosy, E., Sieben, C., Manley, S., Herrmann, A., Sibarita, J.-B., Choquet, D., and Holcman, D. (2012). Heterogeneity of ampa receptor trafficking and molecular interactions revealed by superresolution analysis of live cell imaging. *Proceedings Natl. Academy Sci.*, 109(42) :17052–17057.

[Kervrann et al., 2016] Kervrann, C., Sorzano, C. Ó. S., Acton, S. T., Olivo-Marin, J.-C., and Unser, M. (2016). A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy. *IEEE J. Select. Topics Signal Process.*, 10(1) :6–30.

[Lagache et al., 2009] Lagache, T., Dauty, E., and Holcman, D. (2009). Quantitative analysis of virus and plasmid trafficking in cells. *Phys. Review E*, 79(1) :011921.

[Lakshminarayan et al., 2014] Lakshminarayan, R., Wunder, C., Becken, U., Howes, M., Benzing, C., Arumugam, S., Sales, S., Ariotti, N., Chambon, V., C., L., Loew, D., Shevchenko, A., Gaus, K., R.G., P., and Johannes, L. (2014). Galectin-3 drives glycosphingolipid-dependent biogenesis of clathrin-independent carriers. *Nature Cell Biol.*, 16(6) :595–606.

[Maroulas et al., 2015] Maroulas, V., Nebenführ, A., et al. (2015). Tracking rapid intracellular movements : A bayesian random set approach. *The Annals of Applied Statistics*, 9(2) :926–949.

[Meilhac et al., 2006] Meilhac, N., Le Guyader, L., Salome, L., and Destainville, N. (2006). Detection of confinement and jumps in single-molecule membrane trajectories. *Phys. Review E*, 73(1) :011915.

[Metzler and Klafter, 2000] Metzler, R. and Klafter, J. (2000). The random walk's guide to anomalous diffusion : a fractional dynamics approach. *Physics Reports*, 339(1) :1–77.

[Monnier et al., 2015] Monnier, N., Barry, Z., Park, H. Y., Su, K.-C., Katz, Z., English, B. P., Dey, A., Pan, K., Cheeseman, I. M., Singer, R. H., et al. (2015). Inferring transient particle transport dynamics in live cells. *Nature Meth.*, 12(9) :838–840.

[Qian et al., 1991] Qian, H., Sheetz, M. P., and Elson, E. L. (1991). Single particle tracking. analysis of diffusion and flow in two-dimensional systems. *Biophysical J.*, 60(4) :910.

[Roudot et al., 2017] Roudot, P., Ding, L., Jaqaman, K., Kervrann, C., and Danuser, G. (2017). Piecewise-stationary motion modeling and iterative smoothing to track heterogeneous particle motions in dense environments. *IEEE Trans. Image Process.*, 26(11) :5395–5410.

[Simson et al., 1995] Simson, R., Sheets, E. D., and Jacobson, K. (1995). Detection of temporary lateral confinement of membrane proteins using single-particle tracking analysis. *Biophysical J.*, 69(3) :989–993.

[Truong et al., 2018] Truong, C., Oudre, L., and Vayatis, N. (2018). A review of change point detection methods. *arXiv preprint arXiv :1801.00718*.

[Türkcan and Masson, 2013] Türkcan, S. and Masson, J.-B. (2013). Bayesian decision tree for the classification of the mode of motion in single-molecule trajectories. *PLoS One*, 8(12) :e82799.

[Yin et al., 2017] Yin, S., Song, N., and Yang, H. (2017). Detection of velocity and diffusion coefficient change points in single-particle trajectories. *Biophysical J.*, 115(2) :217–229.

Supplementary Materials: A Sequential Algorithm to Detect Diffusion Switching along Intracellular Particle Trajectories

1 A Change-Point Model

In this section, we describe a more general change-point model that includes the model presented in Section 2.1 of the paper. As our sequential algorithm is based on the test procedure of Briane et al. (2018), we consider a similar diffusion process as in Briane et al. (2018) adapted for the change-point problem :

$$dX_t = \mu(X_t, t)dt + \sigma(t)dB_t^{\mathfrak{h}(t)}, \qquad t \in [t_0, t_{n-1}], \tag{1.1}$$

where $B^{\mathfrak{h}(t)}$ denotes a $d-$dimensional fractional Brownian motion of Hurst parameter $\mathfrak{h}(t)$; the unknown parameters of the model are the Hurst parameter function $\mathfrak{h} : \mathbb{R}^+ \to (0, 1)$, the diffusion coefficient function $\sigma : \mathbb{R}^+ \to (0, \infty)$ and the drift term $\mu : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^d$. Compared to the model presented in Section 2.1, we add the Hurst parameter $\mathfrak{h}(t)$: when $\mathfrak{h} \neq 1/2$ the SDE (1.1) is driven by fractional Brownian motion $B^{\mathfrak{h}(t)}$ which has correlated increments. Also, in model (1.1), we assume that the diffusion coefficient σ can vary over time while this parameter is constant in the model of the paper.

As before we suppose that the parameters defining the diffusion (1.1) are piecewise constant over time. Then, we assume that there exists a sequence of N change-points on $[t_0, t_{n-1}]$, namely $t_0 = \tau_0 < \tau_1 < \ldots \tau_N < \tau_{N+1} = t_{n-1}$ such that,

$$\mu(x, t) = \mu_j(x), \ \mathfrak{h}(t) = \mathfrak{h}_j, \ \sigma(t) = \sigma_j \text{ for } t \in [\tau_j, \tau_{j+1}). \tag{1.2}$$

The unknown parameters of the model are the vector of change-points $(\tau_j)_{j=1\ldots N}$, the number of change-points N, and the parameters $(\mathfrak{h}_j, \mu_j, \sigma_j)$ of the diffusion process restricted on subinterval $[\tau_j, \tau_{j+1})$. The drift term μ_j is assumed to satisfy the usual Lipschitz and linear growth conditions in order that the SDE (1.1) admits a strong solution

on $[\tau_j, \tau_{j+1})$ (see (Nualart & Ouknine 2002) for $0 < \mathfrak{h} \leq 1/2$ and (Mishura 2008) for $1/2 < \mathfrak{h} < 1$). We extend by continuity the solution on each subinterval to get a solution on $[t_0, t_{n-1}]$.

Again we assume that for each τ_j there exists $0 \leq j^\star \leq n$ such that $\tau_j = t_{j^*}$ (the change of motion occurs precisely at a sampling time). In analogy with the model of the paper, we assume that $(\mathfrak{h}_j, \mu_j)$ and $(\mathfrak{h}_{j+1}, \mu_{j+1})$ are associated to different types of diffusion (Brownian motion, subdiffusion or superdiffusion). We note that the parameter σ_j does not influence the type of diffusion.

Finally we have to mention that the test procedure of Briane et al. (2018) has also been validated in the case of continuous time random walk (CTRW) characterized by a subdiffusive behaviour. CTRW are not defined through stochastic differential equations. Then, our sequential algorithm can deal with an even greater variety of change-point models than the two presented here (for instance transition of motion including CTRW).

2 Choosing the cut-off values (γ_1, γ_2)

Ideally, γ_1 and γ_2 are choosen such that we control the type I error at level $0 < \alpha < 1$; in other words such that we control the probability to detect falsely a change-point when the trajectory is fully Brownian. Then, controlling the type I error at level $0 < \alpha < 1$ is equivalent to have :

$$P_{H_0}\left(\exists i \in \{k, \ldots, n^\star\}, \quad \sum_{j=i}^{i+c-1} Q_j \geq pc\right) \leq \alpha, \tag{2.1}$$

where $n^\star = n - k - c + 1$ and P_{H_0} is the probability under H_0 that is under the hypothesis that the trajectory is fully Brownian. In fact, the left hand side of Equation (2.1) is the probability to build one cluster of minimal size c under H_0, when clusters are defined thanks to Eq. (3.3) (see step 2 of the procedure in Section 3.1 of the paper). Then, controlling the probability in (2.1) at level α under H_0 is equivalent to control the probability to detect falsely a change-point under H_0 at level α (definition of the type I error).

Proposition 1. *Let define* $d_i = \min(B_i, A_i)$ *and* $D_i = \max(B_i, A_i)$ *where* A_i *and* B_i *are the test statistics (3.1), for* $i = k, \ldots, n^\star$. *We define* $\gamma_1^\star$ *and* $\gamma_2^\star$ *as:*

$$\begin{aligned} P_{H_0}\left(\min_{i=k,\ldots,n^\star} d_{i(\lceil pc/2 \rceil)} < \gamma_1^\star\right) &= \frac{\alpha}{2}, \\ P_{H_0}\left(\max_{i=k,\ldots,n^\star} D_{i(c-\lceil pc/2 \rceil)} > \gamma_2^\star\right) &= \frac{\alpha}{2}, \end{aligned} \tag{2.2}$$

where,

- $d_{i(\lceil pc/2 \rceil)}$ *is the* $\lceil pc/2 \rceil$ *smallest element of* $(d_i, \ldots, d_{i+c-1})$,
- $D_{i(c-\lceil pc/2 \rceil)}$ *the* $c-\lceil pc/2 \rceil$ *smallest element of* $(D_i, \ldots, D_{i+c-1})$ *(equivalently the* $\lceil pc/2 \rceil$ *greatest element).*

In other words, $\gamma_1^\star$ *is the quantile of order* $\alpha/2$ *of the random variable* $\min_{i=k,\ldots,n^\star} d_{i(\lceil pc/2 \rceil)}$ *and* $\gamma_2^\star$ *is the quantile of order* $1-\alpha/2$ *of the random variable* $\max_{i=k,\ldots,n^\star} D_{i(c-\lceil pc/2 \rceil)}$. *With the choice of cut-off values* $\gamma_1^\star$ *and* $\gamma_2^\star$ *Procedure 1 with parameters* (k, c, p) *controls the type I error* (2.1) *at level* α.

Proposition 1 provides a choice for γ_1 and γ_2 in order to controll the level of the procedure. These thresholds $\gamma_1^\star$ and $\gamma_2^\star$ can be approximated by Monte Carlo estimate, see Algorithm 1. A proof of Proposition 1 is available in Section 2.1. Nethertheless this choice is too conservative as it is shown in Table 1. We can see that in the proof of Proposition 1 (Section 2.1) since the bound in Equation (2.9) is loose.

We investigate another choice by Monte Carlo experiments, and we recommend to use the heuristic cut-off values $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ defined as follow :

$$\begin{aligned} P_{H_0}\Big(\min_{i=k,\ldots,n^\star} d_{i(\lceil pc \rceil)} < \tilde{\gamma}_1\Big) &= \frac{\alpha}{2}, \\ P_{H_0}\Big(\max_{i=k,\ldots,n^\star} D_{i(c-\lceil pc \rceil)} > \tilde{\gamma}_2\Big) &= \frac{\alpha}{2}. \end{aligned} \tag{2.3}$$

Notice that we replace $pc/2$ in Equation (2.2) by pc. Then, it is straightforward to show that $\gamma_1^\star \leq \tilde{\gamma}_1$ and $\gamma_2^\star \geq \tilde{\gamma}_2$. Table 2 gives Monte Carlo approximations of $(\gamma_1^\star, \gamma_2^\star)$ and $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ as illustration. Procedure 1 with $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ is less conservative : it is more sensitive to the presence of subdiffusion or superdiffusion. We can also see that the estimated error rates are closed to the expected value $\alpha = 5\%$, see Table 1.

Remark 2.1. *We note that the cut-off values defined by* (2.2) *or* (2.3) *potentially depend on the diffusion coefficient* σ *(and on the step of time* Δ*). In fact, the null hypothesis* H_0 *depends on parameter* σ *(and* Δ*). However, the test statistics (3.1) do not depend on* (σ, Δ) *under* H_0*. Consequently, the cut-off values defined by* (2.2) *or* (2.3) *neither depend on* (σ, Δ).

2.1 Proof of Proposition 1

Proof. We suppose that the trajectory $\mathbb{X}_n$ is generated under the null hypothesis (3.2) that is the trajectory is fully Brownian. For simplicity, we note P the probability under H_0

Table 1: The estimated type I error rates (in percentage) for different cut-off values (γ_1, γ_2) when $\alpha = 5\%$. The type I error rate is computed using Monte Carlo experiments over 100 001 Brownian trajectories. We use the default parameters for Procedure 1 that is $c = k/2$ and $p = 0.75$. The estimations are accurate at $\pm 0.14\%$.

		Probability of Type I error	
n	k	with $(\gamma_1^\star, \gamma_2^\star)$	with $(\tilde{\gamma}_1, \tilde{\gamma}_2)$
150	20	0.60	5.21
150	30	0.65	4.81
150	40	0.94	4.56
300	20	0.47	5.04
300	30	0.59	4.89
300	40	0.82	4.83

Table 2: Two choices of cut-off values for Procedure 1 for different trajectory sizes n and window sizes k for dimension $d = 2$. The cut-off values are estimated with the Monte Carlo Algorithm 1 using $V = 10\,001$ replications and the default parameters of Procedure 1 $c = k/2$, $p = 0.75$ and with $\alpha = 0.05$.

n	k	$\gamma_1^\star$	$\gamma_2^\star$	$\tilde{\gamma}_1$	$\tilde{\gamma}_2$
150	20	0.61	3.38	0.74	3.12
150	30	0.65	3.35	0.79	3.09
150	40	0.68	3.28	0.81	3.05
300	20	0.58	3.55	0.71	3.29
300	30	0.62	3.55	0.74	3.28
300	40	0.64	3.52	0.75	3.27

(noted P_{H_0} previously).We want to show that under H_0, Procedure 1 with thresholds γ_1 and γ_2 defined in Proposition 1, controls the probability of the type I error at level α:

$$P\left(\exists i \in \{k, \dots, n^\star\}, \quad \sum_{j=i}^{i+c-1} Q_j \geq pc\right) \leq \alpha \tag{2.4}$$

where $n^\star = n - k - c + 1$.
We express the event $\{Q_i = 1\}$ as:

$$\begin{aligned}\{Q_i = 1\} = &\{\gamma_1 \leq B_i \leq \gamma_2, A_i < \gamma_1\} \cup \{\gamma_1 \leq B_i \leq \gamma_2, A_i > \gamma_2\} \\ &\cup \{B_i < \gamma_1, \gamma_1 \leq A_i \leq \gamma_2\} \cup \{B_i > \gamma_2, \gamma_1 \leq A_i \leq \gamma_2\} \\ &\cup \{B_i < \gamma_1, A_i > \gamma_2\} \cup \{B_i > \gamma_2, A_i < \gamma_1\}\end{aligned} \tag{2.5}$$

Then we deduce the following :

$$\begin{aligned}\{Q_i = 1\} &\subset \{B_i < \gamma_1\} \cup \{A_i < \gamma_1\} \cup \{B_i > \gamma_2\} \cup \{A_i > \gamma_2\} \\ &= \{\min(B_i, A_i) < \gamma_1\} \cup \{\max(B_i, A_i) > \gamma_2\}\end{aligned} \tag{2.6}$$

In the sequel, we note $d_i = \min(B_i, A_i)$ and $D_i = \max(B_i, A_i)$. Then we have:

$$P(Q_i = 1) \leq P(\{d_i < \gamma_1\} \cup \{D_i > \gamma_2\}), \quad i = k, \dots, n^\star. \tag{2.7}$$

This implies the following:

$$P\Big(\sum_{j=i}^{i+c-1} Q_j \geq pc\Big) \leq P\Big(\sum_{j=i}^{i+c-1} \mathbf{1}(\{d_j < \gamma_1\} \cup \{D_j > \gamma_2\}) \geq pc\Big) \tag{2.8}$$

Now, we can bound the right-hand side of Equation (2.8):

$$\begin{aligned}&P\Big(\sum_{j=i}^{i+c-1} \mathbf{1}(\{d_j < \gamma_1\} \cup \{D_j > \gamma_2\}) \geq pc\Big) \\ \leq &P\Big(\sum_{j=i}^{i+c-1} \mathbf{1}(\{d_j < \gamma_1\}) + \mathbf{1}(\{D_j > \gamma_2\}) \geq pc\Big) \\ \leq &P\Big(\sum_{j=i}^{i+c-1} \mathbf{1}(\{d_j < \gamma_1\}) \geq pc/2\Big) + P\Big(\sum_{j=i}^{i+c-1} \mathbf{1}(\{D_j > \gamma_2\}) \geq pc/2\Big)\end{aligned} \tag{2.9}$$

Then, we can express the right-hand side of Equation (2.9) as:

$$\begin{aligned} &P\Big(\sum_{j=i}^{i+c-1} \mathbf{1}(\{d_j < \gamma_1\}) \geq pc/2\Big) + P\Big(\sum_{j=i}^{i+c-1} \mathbf{1}(\{D_j > \gamma_2\}) \geq pc/2\Big) \\ =&P(d_{i(\lceil pc/2 \rceil)} < \gamma_1) + P(D_{i(c-\lceil pc/2 \rceil)} > \gamma_2) \end{aligned} \tag{2.10}$$

Finally we have:

$$\begin{aligned} &P(\exists i \in \{k, \dots, n^\star\}, \quad \sum_{j=i}^{i+c-1} Q_j \geq pc) \\ =&P\Big(\bigcup_{i=k}^{n^\star}\Big\{\sum_{j=i}^{i+c-1} Q_j \geq pc\Big\}\Big) \\ \leq&P\Big(\bigcup_{i=k}^{n^\star}\Big\{\sum_{j=i}^{i+c-1} \mathbf{1}(\{d_j < \gamma_1\}) \geq pc/2\Big\}\Big) + P\Big(\bigcup_{i=k}^{n^\star}\Big\{\sum_{j=i}^{i+c-1} \mathbf{1}(\{D_j > \gamma_2\}) \geq pc/2\Big\}\Big) \\ =&P\Big(\bigcup_{i=k}^{n^\star}\{d_{i(\lceil pc/2 \rceil)} < \gamma_1\}\Big) + P\Big(\bigcup_{i=k}^{n^\star}\{D_{i(c-\lceil pc/2 \rceil)} > \gamma_2\}\Big) \\ =&P\Big(\min_{i=k,\dots,n^\star} d_{i(\lceil pc/2 \rceil)} < \gamma_1\Big) + P\Big(\max_{i=k,\dots,n^\star} D_{i(c-\lceil pc/2 \rceil)} > \gamma_2\Big) \\ =&\frac{\alpha}{2} + \frac{\alpha}{2} = \alpha \end{aligned} \tag{2.11}$$

We go from line 2 to line 3 using Equations (2.8) and (2.9). We go from line 3 to line 47 using Equation (2.10). Finally, we go from line 5 to 6 using the thresholds γ_1 and γ_2 of Proposition 1. It finishes the proof. □

2.2 Monte Carlo Algorithm for Computing (γ_1, γ_2)

Input: n, k, c, p, α, V
// the length n of the trajectory
// the window size k
// the cluster parameters (c, p)
// the level $\alpha \in (0, 1)$
// the number V of Monte Carlo experiments
Result: $\hat{\gamma}_1(n, k, c, p, \alpha)$ $\hat{\gamma}_2(n, k, c, p, \alpha)$
for *i=1* **to** V **do**
  Generate $\mathbb{X}_n^i$ of size n from the null hypothesis (3.2) (see the paper) with $\sigma = 1$ and $\Delta = 1$;
  // Compute the statistics (4.1) (see the paper) along $\mathbb{X}_n^i$
  for *j=k* **to** *n-k* **do**
    Compute (B_j^i, A_j^i) from (4.1);
    Set $d_j^i = \min(B_j^i, A_j^i)$;
    Set $D_j^i = \max(B_j^i, A_j^i)$;
  end
  for *r=k* **to** *n-k-c+1* **do**
    Compute s_r^i the $\lceil pc \rceil$ smallest element of $(d_r^i, \ldots, d_{r+c-1}^i)$;
    Compute S_r^i the $c - \lceil pc \rceil$ smallest element of $(d_r^i, \ldots, d_{r+c-1}^i)$;
  end
  Compute $m_i = \min_r(S_r^i)$ and $M_i = \max_r(s_r^i)$;
end
Let $(\tilde{m}_1, \ldots, \tilde{m}_V)$ the sorted m_is and $(\tilde{M}_1, \ldots, \tilde{M}_V)$ the sorted M_is;
Set $\hat{\gamma}_1(n, k, \alpha) = \tilde{m}_{\lfloor (\alpha/2)V \rfloor}$ and $\hat{\gamma}_2(n, k, \alpha) = \tilde{M}_{\lfloor (1-\alpha/2)V \rfloor}$;

Algorithm 1: Estimation of the cut-off values $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ by Monte Carlo simulations. For estimating $(\gamma_1^\star, \gamma_2^\star)$, one should turn pc into $pc/2$ in this algorithm.

3 Comparisons with Competitive Methods

We compare our method to two other methods. The method of Türkcan & Masson (2013) detects change-points between Brownian motion and parametric models of subdiffusion including the Ornstein-Uhlenbeck process. The method of Monnier et al. (2015) detects change-points between Brownian motion and Brownian motion with drift which is an example of superdiffusion. We note that none of these methods deal with the three types of diffusion (Brownian motion, subdiffusion and superdiffusion) as we do. In this section, we present the two competitive methods and compare their performances to Procedure 1 on simulations. At the end of the section, we give a particular emphasis on the speed and stability of the different methods.

3.1 The Method of Türkcan & Masson (2013) First the method of Türkcan & Masson (2013) is a parametric method. The two parametric models under concern are the Brownian motion and the Ornstein-Uhlenbeck process (called diffusion in a harmonic potential in (Türkcan & Masson 2013)). Türkcan & Masson (2013) select the model that minimizes the BIC criterion. For detecting change-points, the BIC criterion is computed on a sliding window along the trajectory. When the BIC indicates a switch of model and that the new model is confirmed in the next r steps of times, a change is assumed to occur.

We reproduce the simulation described in Türkcan & Masson (2013). We simulate $V = 100$ trajectories of size $n = 500$. First the trajectory undergoes an Ornstein-Uhlenbeck process and at time $\tau_1 = 250$ it switches to a Brownian motion. The two processes share the same diffusion coefficient $\sigma = 0.4472$. The specific parameter of the Ornstein-Uhlenbeck process (4.1) is $\lambda = 7.3870$. The step of time is $\Delta = 0.05$. Results of the two methods are given in Table 3. We can see that our method show better results in both the number $\hat{N}$ of detected change-points and in the location of the change-points. We also emphasize that we do not set $r = 3$ as in (Türkcan & Masson 2013) but we set $r = 51$ which corresponds to the size of the window. With $r = 3$, the method of Türkcan & Masson (2013) detects more than 4 change-points in 91% of the trajectories. Actually, the method is able to detect the change-point, if a collection of about $V = 50$ trajectories showing the same number of change-points at the same location is available. Accordingly, it provides good results in average. However, such a situation is not realistic in practical imaging. In our scenarios, our non-parametric method outperforms the parametric method of Türkcan & Masson (2013).

3.2 The Method of Monnier et al. (2015) Monnier et al. (2015) use two parametric models to fit the displacements of the particle: the Brownian motion and the Brownian

Table 3: Comparison of Procedure 1 and the method of Türkcan & Masson (2013) on the simulation of Türkcan & Masson (2013). We recall that the true change-point is $\tau_1 = 250$.

Method	$\hat{N} - N$				τ_1	
	-1	0	1	≥ 2		
Procedure 1	19	77	3	1	240.5	(29.4)
Method of Türkcan & Masson (2013)	27	59	14	0	176.3	(53.7)

motion with drift. We note that the Brownian motion can be seen as a Brownian motion with a null drift. The two models are actually a unique parametric model with parameters $v = (v_1, v_2)$ and σ (with $v = (0, 0)$ for the Brownian case). Then, Monnier et al. (2015) use hidden Markov models to fit the displacements of the particle over time. The hidden states are defined as a set of a drift parameter and diffusion coefficient $S_k = v_k, \sigma_k$. They estimate both the number of states K, the parameters (v_k, σ_k) and the successive (hidden) states along the trajectories. They also add a constrained $v = 0$ for modelling Brownian motion. Model selection is used with a Bayesian criterion to select the best model. If we assume that $K \leq 2$ and also consider the constrained models with $v = 0$, the competing models are:

Model 1 a single state which is the Brownian with parameter σ_1,

Model 2 a single state which is the Brownian with drift with parameters (v_1,σ_1),

Model 3 two states which are two Brownian with parameter σ_1 and σ_2,

Model 4 two states which are one Brownian and one Brownian with drift with respective parameters σ_1 and (v_2, σ_2),

Model 5 two states which are two Brownian with drift with parameters (v_1, σ_1) and (v_2, σ_2).

In our experiment, we run the method of Monnier et al. (2015) on 100 simulated trajectories from Scenario 1. We assume $K \leq 2$ that is the competing models are the five models aforementioned. Results are given in Table 4. When $v = 0.6$ or 0.8, Monnier et al. (2015) detect no change-point ($\hat{N} - N = -2$) for a large majority of the trajectories. In this case, the selected model can either have one state (models 1 and 2) or have two states but from the same type of diffusion (models 3 and 5). Actually, when $v = 0.6, 0.8$, the preferred model is Brownian only (model 1) for most of the trajectories (see Table 5). Then the drift is too low to select a model involving Brownian with drift. As expected, the performance of

Table 4: Performance of the algorithm of Monnier et al. (2015) for Scenario 1 (see Table 2 of the paper). The computations are based on 100 simulated trajectories from Scenario 1. We compute the proportions of trajectories with $\hat{N} - N = -2$, $\hat{N} - N = \pm 1$, $\hat{N} - N = 0$ and $\hat{N} - N \geq 2$. The column τ_1 (respectively τ_2) gives the empirical average of the first (respectively second) detected change point on the trajectories among which we detect the right number of change points ($\hat{N} - N = 0$). The number in brackets is the empirical standard deviation of the estimates of τ_1 and τ_2. We note that the empirical average and standard deviation estimate of τ_1 and τ_2 are not computed over the same number of trajectories for the different values of the drift v (causing the null standard deviation line 1).

	$\hat{N} - N$								
v	-2	-1	0	1	≥ 2	τ_1		τ_2	
0.6	99	0	1	0	0	93.0	(0.0)	177.0	(0.0)
0.8	82	0	15	1	2	96.0	(7.9)	173.4	(3.9)
1.0	23	0	68	7	2	99.9	(3.9)	174.9	(4.3)
2.0	0	0	96	1	3	100.0	(1.4)	175.0	(1.2)

the method of Monnier et al. (2015) improves as v increases. The method detects the right number of change-points for 96% of the trajectories when $v = 2$. When the method detects at least one change-point, it means that the selected model is the model 5. Even when the right model is chosen, it can over-detect the number of change-points (that is $\hat{N} - N \geq 1$). We have 9% of overdetection when $v = 1$. When the method detects the right number of change-points ($\hat{N} - N = 0$), the location of the change-points are very close to the true locations. For instance when $v = 2$, the average location of the first detected change-point is 100 (which is exactly τ_1) and its standard deviation is 1.4. Finally, our non-parametric method detects better the change-points when the drift is low ($v \leq 0.8$). The quality of detections are similar when the drift is high enough ($v = 2$). For $v = 1$, we have a larger proportion of trajectories detected with the right number of change-points with our method except when we use a window size of $n = 20$ (63.2% with our method *versus* 68% with the method of Monnier et al. (2015)). The locations of the detected change-points among the trajectories with $\hat{N} - N = 0$ are slightly more accurate with the method of Monnier et al. (2015).

3.3 Algorithmic Considerations Finally, we compare the speed and stability of the different methods. The method of Monnier et al. (2015) is time consuming because of the estimation of the *a posteriori* distribution by the Metropolis-Hastings algorithm. Assuming

Table 5: Selected models with the method of Monnier et al. (2015) on 100 simulated trajectories from Scenario 1. BR (respectively BRD) stands for Brownian (respectively Brownian with drift). For instance, when $v = 0.6$, the method of Monnier et al. (2015) states that the best fit is Brownian motion only for 97 trajectories; Brownian motion with drift only for 2 trajectories; a mix of Brownian and Brownian motion with drift for 1 trajectory.

v	BR	BRD	BR-BR	BR-BRD	BRD-BRD
0.6	97	2	0	1	0
0.8	74	8	0	18	0
1	18	5	0	77	0
2	0	0	0	100	0

$K \leq 2$, it took $115s$ in average to deal with one trajectory of the simulation presented in Table 4 (300 points) with four cores working in parallel on a Mac Book Pro version 10.10.1 equipped with 2.8 GHz Intel Core i7, 16 Gb of RAM. In comparison, our method takes less than $0.05s$ to process a trajectory without working in parallel. Both our procedure and the method of Türkcan & Masson (2013) compute quantities on local windows (in our case the statistics (3.1), the BIC of different models for (Türkcan & Masson 2013)). From this aspect, the complexity of these two algorithms is equivalent. However, Türkcan & Masson (2013) needs to estimate the MAP (maximum *a posteriori*) to compute the BIC. They choose a complex likelihood to model the spatial heterogeneity of the motion. Therefore, they use quasi-Newtonian optimization to find the MAP which is the most time consuming step of their procedure. It took in average $11s$ to process a trajectory of the simulation presented in Table 4 (500 points) against less than $0.05s$ for Procedure 1. In term of stability, different runs of the method of Monnier et al. (2015) on the same trajectory can give different results (see Section 4.4 of the paper).This is due to a bad convergence of the Metropolis-Hastings algorithm. In rare cases, the optimization step of Türkcan & Masson (2013) can fail. Procedure 1 does not suffer any of these problems as it does not involve any parameter inference.

References

Briane, V., Kervrann, C. & Vimond, M. (2018), 'Statistical analysis of particle trajectories in living cells', *Physical Review E* **97**(6), 062121.

Mishura, Y. (2008), *Stochastic calculus for fractional Brownian motion and related processes*, Vol. 1929, Springer Science & Business Media.

Monnier, N., Barry, Z., Park, H. Y., Su, K.-C., Katz, Z., English, B. P., Dey, A., Pan, K., Cheeseman, I. M., Singer, R. H. et al. (2015), 'Inferring transient particle transport dynamics in live cells', *Nature methods* **12**(9), 838–840.

Nualart, D. & Ouknine, Y. (2002), 'Regularization of differential equations by fractional noise', *Stochastic Processes and their Applications* **102**(1), 103–116.

Türkcan, S. & Masson, J.-B. (2013), 'Bayesian decision tree for the classification of the mode of motion in single-molecule trajectories', *PLoS One* **8**(12), e82799.