Dense mapping of intracellular diffusion and drift from single-particle tracking data analysis

Antoine Salomon, Cesar Augusto Valades Cruz, Ludovic Leconte, Charles Kervrann

DENSE MAPPING OF INTRACELLULAR DIFFUSION AND DRIFT FROM SINGLE-PARTICLE TRACKING DATA ANALYSIS

Antoine Salomon[1], Cesar Augusto Valades-Cruz[1,2], Ludovic Leconte[2], Charles Kervrann[1]

1. Inria, Centre Rennes-Bretagne Atlantique, SERPICO Project Team, 35042 Rennes, France
2. Institut Curie, PSL Research University, Space-Time imaging of Endomembranes Dynamics Team, UMR144 CNRS, 75248, Paris, France

ABSTRACT

It is of primary interest for biologists to be able to visualize the dynamics of proteins within the cell. In this paper, we propose a new mapping method to robustly estimate dynamics in the entire cell from particle tracks. To obtain satisfying diffusion and drift maps, we use a spatiotemporal kernel estimator. Trajectory classification data is used as input and allows to automatically label particle movements into three classes: confined motion (or subdiffusion), Brownian motion, and directed motion (or superdiffusion). We then use this information to calculate diffusion coefficient and drift maps separately on each class of motion.

Index Terms— Diffusion, drift, mapping, classification, single-particle tracking

1. INTRODUCTION

Diffusion is defined as the ensemble movement of elementary particles (e.g. atoms, ions, molecules,...) from a region of higher concentration to a region of lower concentration. A basic example of diffusion would be the spread of dye in static water. The motion of each particle of dye is a consequence of collisions with fast-moving molecules in fluids, which produces Brownian motion. This results in a spreading gradient at the macroscopic scale. This notion is of particular importance in biology, since it can be used to characterize the stochastic part of particle motion in the intracellular medium. On the other hand, drift is defined as the deterministic active component of particle motion, like for instance active transport of vesicles propelled by molecular motors. Several modelings have been proposed for diffusion modeling. In particular, Einstein and Langevin defined general models [1] [2].

In this paper, we focus on the Langevin equation, which takes into account both drift and diffusion as follows:

$$d\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{x})dt + \sigma(\boldsymbol{x})d\boldsymbol{w} \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^d, d = 2, 3$ is the spatial coordinates of a particle, $\boldsymbol{b} \in \mathbb{R}^d$ the drift vector, $\sigma \in \mathbb{R}^+$ the diffusion coefficient, and $\boldsymbol{w} \in \mathbb{R}^d$ standard Gaussian white noise. This equation can be used to draw a map of both the diffusion and the drift in a cell given the spatial and temporal coordinates of moving particles. To obtain the trajectories in Total internal reflection fluorescence microscopy (TIRF) [3] [4] or in 3D Lattice Light Sheet Microscopy (LLSM) [5], we can use different techniques. Different trackers can be used to map the videos into tracking data, such as the Multiple Hypothesis Tracker [6] [7] or U-track [8] [9]. Given particle tracks, Hozé et al. [10] proposed to exploit tracking data and divide the image domain into square blocks. In each block of the regular grid, diffusion and drift are estimated, providing results in the form of images for the diffusion and vector field matrices for the drift. This particular method was dedicated to Single-Particle Tracking (SPT) - PALM which is able to provide a huge amount of tracks. However, it can give disappointing or inexploitable outputs if the number of trajectories points is too low, with a low spatial resolution depending on the size of the square blocks.

We propose several new features to improve this estimation method including in the case of conventional microscopy. First, we use motion classification to automatically divide and classify trajectories into sub-trajectories of three motion classes: subdiffusion, Brownian motion, and superdiffusion motion. This information serves to calculate diffusion coefficient and drift maps separately on each motion class with the most suitable model (Section 2). To obtain satisfying results, i.e. to avoid the blurring effect due to calculation of both diffusion and drift in cell coordinates where no data is available, we replace the scanning movement of an averaging window by a trajectory-following movement combined with spatial and temporal Gaussian weights. Each diffusion point and each drift vector are then calculated at coordinates corresponding to the coordinates given by the particle tracking, which gives us more detailed results in the form of point clouds and vector fields.

The remainder of the paper is organized as follows: in the next section, we present the trajectory labeling framework. In Section 3, we present the estimators of diffusion and drift. In Section 4, we demonstrate the potential of our estimation approach on artificial and real fluorescence microscopy data.

2. PARTICLE MOTION CLASSIFICATION

The motion of a particle in the cytoplasm can be roughly classified into three classes. In general, particle motion is a stochastic process, resulting of all shocks between the particles and the environment. In what follows, we describe conventional situations. First, free motion (or Brownian motion-) corresponds to pure diffusion (no obstacle and no collision), with no drift. The Langevin equation can be re-written in a simplified form as:

$$dx = \sigma(x)dw. \tag{2}$$

Second, the particle can be bound to a static point. In this case (subdiffusion) stochastic motion is restricted to a small area around the binding point. The Ornstein-Uhlenbeck process has been proposed to model this motion, and can be seen as a particular type of Langevin stochastic process:

$$dx = \theta(x_0 - x)dt + \sigma(x)dw \tag{3}$$

where θ designates a constant corresponding the strength of the binding and x_0 the binding point. Third, superdiffusion corresponds to a mostly deterministic motion. The particle moves in a given direction and motion is not restricted to a small area. This case typically happens when the particle is being actively transported within the cell by molecular motors.

In [11], the authors proposed a statistical method, which given a trajectory as an input, is able to assign a motion class to it. The same particle being able to change the way it moves following its interaction with its cellular environment, the same authors proposed in [12] a statistical method to detect switches between two types of motion along the trajectory. This method is able to divide a trajectory into classified subtrajectories. In what follows, we use the two aforementioned methods to classify particle tracks. We end up with the set $\mathcal{X} \subset \mathbb{R}^d, d = 2, 3$ of all points from all trajectories, divided by trajectories, subtrajectories, and motion class.

3. DENSE MAPPING

If we define $\mathcal{X}$ as the set of all points from all trajectories, we can define $\mathcal{X}_{free}$, $\mathcal{X}_{sub}$ and $\mathcal{X}_{sup}$ respectively as the sets of points classified as Brownian, subdiffusive and superdiffusive. We want to calculate a drift vector and a diffusion coefficient on each point of $\mathcal{X}$. Inside a trajectory, we respectively define the first and second order differences of x_{t_i} as:

$$\Delta^1 x_{t_j} = x_{t_{j+1}} - x_{t_j}, \tag{4}$$

$$\Delta^2 x_{t_j} = x_{t_{j+1}} - 2x_{t_j} + x_{t_{j-1}}. \tag{5}$$

For each point x_{t_i}, we calculate the drift vector and the diffusion coefficient using one of the following formulas, depending on its motion type.
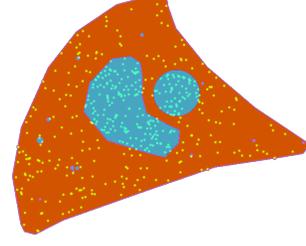
Fig. 1. Example of FluoSIM crossbow cell simulation, used as groundtruth. Each green point represents a moving particle. The cell is labeled in orange.

Brownian motion   We assume that the drift component is null in the case of Brownian motion, i.e. $b(x) = 0$. The diffusion coefficient for the point $x_i \in \mathcal{X}_{free}$ is given by:

$$\hat{\sigma}(x_i) = \frac{\sum\limits_{x \in \mathcal{X}_{free}} w(x, x_i)\|\Delta^2 x\|^2}{2d\Delta t} \tag{6}$$

where $w(x, x_i)$ is the weight assigned to the point x relatively to the point of calculation x_i, $d = 2, 3$ the dimensionality of the data, and Δt the time resolution, i.e. the lag between two frames of the original video.

Subdiffusion   In the case of subdiffusive motion, the drift and diffusion estimators for the point $x_i \in \mathcal{X}_{sub}$ are respectively given by:

$$\hat{b}(x_i) = \frac{\sum\limits_{x \in \mathcal{X}_{sub}} w(x, x_i)\Delta^1 x}{\Delta t}, \tag{7}$$

$$\hat{\sigma}(x_i) = \frac{\sum\limits_{x \in \mathcal{X}_{sub}} w(x, x_i)\|\Delta^1 x - \hat{b}(x_i)\Delta t\|^2}{d\Delta t}. \tag{8}$$

Superdiffusion   In the case of superdiffusive motion, the drift and diffusion estimators for the point $x_i \in \mathcal{X}_{sub}$ are respectively given by the following formulas:

$$\hat{b}(x_i) = \frac{\sum\limits_{x \in \mathcal{X}_{sup}} w(x, x_i)\Delta^1 x}{\Delta t}, \tag{9}$$

$$\hat{\sigma}(x_i) = \frac{\sum\limits_{x \in \mathcal{X}_{sup}} w(x, x_i)\|\Delta^2 x\|^2}{2d\Delta t}. \tag{10}$$

The weight assigned to each point x relatively to point of calculation x_i is the product of a spatial Gaussian function and a temporal Gaussian function. In [10], $w(x, x_i)$ is a constant. In the case where x_i is classified as Brownian motion (the formula is similar for the other motion types), we have:

(a) Diffusion point cloud map ($\mu m^2/s$)   (b) Diffusion point cloud histogram   (c) Drift map
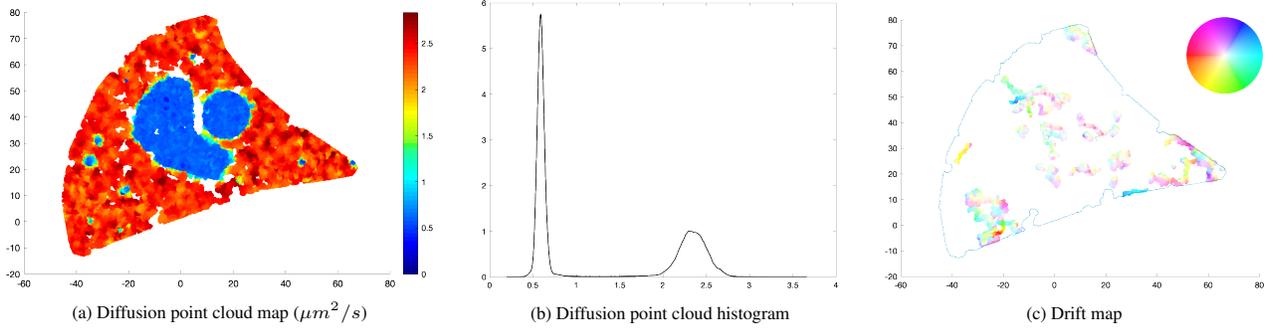
Fig. 2. Results obtained with our method on simulated data. The colormap represents the orientation and the amplitude of vectors.

$$w(\boldsymbol{x}, \boldsymbol{x}_i) = \frac{exp[-2((\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|}{r_S})^2 + (\frac{|t_{\boldsymbol{x}}-t_{\boldsymbol{x}_i}|}{r_T})^2)]}{\sum\limits_{\boldsymbol{x} \in \mathcal{X}_{free}} exp[-2((\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|}{r_S})^2 + (\frac{|t_{\boldsymbol{x}}-t_{\boldsymbol{x}_i}|}{r_T})^2)]} \quad (11)$$

where the radius r_S of the spatial Gaussian, is set up as the kth ($k \in \mathbb{N}$ is a parameter chosen by the user) smallest distance in the following set:

$$dist_{s,i} = \{\|\boldsymbol{x} - \boldsymbol{x}_i\|, \boldsymbol{x} \in \mathcal{X}\} \quad (12)$$

and r_T, the radius of the temporal Gaussian, is set up as:

$$r_T = \max(|t_{\boldsymbol{x}} - t_{\boldsymbol{x}_i}|) \text{ such that } \boldsymbol{x} \in \{\boldsymbol{x}', \|\boldsymbol{x}' - \boldsymbol{x}_i\| \leq r_S\}. \quad (13)$$

In the end, all the calculated estimators can be put together on a graph to form a pseudo-map of drift and diffusion, in the form of a point cloud and a vector field.

4. EXPERIMENTAL RESULTS

To assess the reliability of our method, we conducted several experiments on simulated and real fluorescence microscopy datasets.

Evaluation on simulated datasets   First, we used the FluoSIM software [13] to generate simulated fluorescence microscopy tracking data, by drawing a representation of a 2D crossbow-shaped cell (Fig. 1). The diffusion coefficient was set to 2.4 $\mu m^2/s$ in the orange region, and was set to 0.6 $\mu m^2/s$ in two large regions in the center of the cell, and twelve additional small confinement regions represented in blue. FluoSIM does not feature any simulation of superdiffusive movements, thus only free diffusion and subdiffusion are present here. The sequence shown in Fig. 1 has a length of 1000 frames, a time step of 0.02 s, a spatial resolution of 0.16 μm and includes 450 moving particles. The tracking data was directly extracted from FluoSIM: the tracks are true trajectories with no noise (we did not use an intermediary tracking software on a simulated video sequence).

Our method produced the outputs shown on Fig. 2, while Fig. 3 illustrates the output obtained with our implementation of the method [10]. We can notice a certain amount of false positive results concerning drift. A few trajectories were wrongly classified as superdiffusive, or subdiffusive outside the small confinement regions, thus a drift component was calculated. Nevertheless, our method provided accurate and precise reconstruction of diffusion in comparison to simulation ground truth, both in terms of spatial localizations and parameter values, as we displayed on the diffusion map and its associated histogram (Fig. 2). On the other side, it is hard to distinguish the small confinement regions on the diffusion map generated by the method [10] because of the low spatial resolution (Fig. 3 illustrates the best compromise achieved). The drift map shown in Fig. 3(c) is noisy.
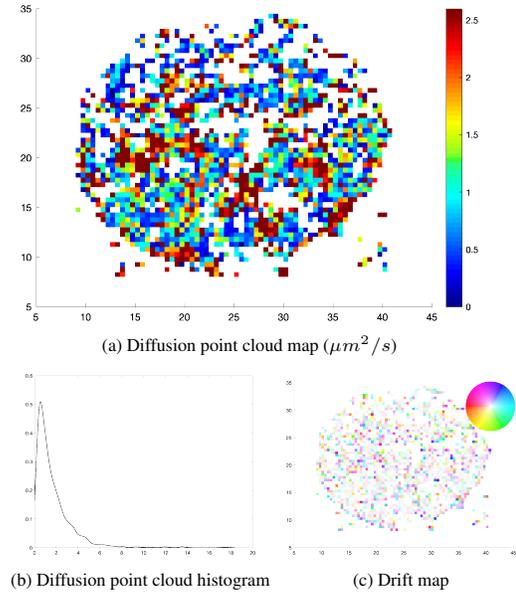
Experiments on real Data (2D)   The method has been assessed on real fluorescence microscopy data, especially in 2D PALM/STORM SPT fluorescence microscopy and in 3D LLSM fluorescence microscopy. Due to the absence of ground truth, we can only give a qualitative assessment of the results, based on what is expected to happen in the specimen.

On the outputs presented in Fig. 4, we can clearly distinguish three modes on the histogram, corresponding to small confinement regions with low diffusion (in blue on the diffusion map), large regions of high diffusion (in red) and the rest of the cell with average diffusion (in blue/green). The diffusion map (and interpolation map) gives clear spatial information about the sample. On the other hand, the outputs obtained with method [10] (Fig. 5) do not allow to distinguish the small confinement regions, as already shown in Fig. 3.
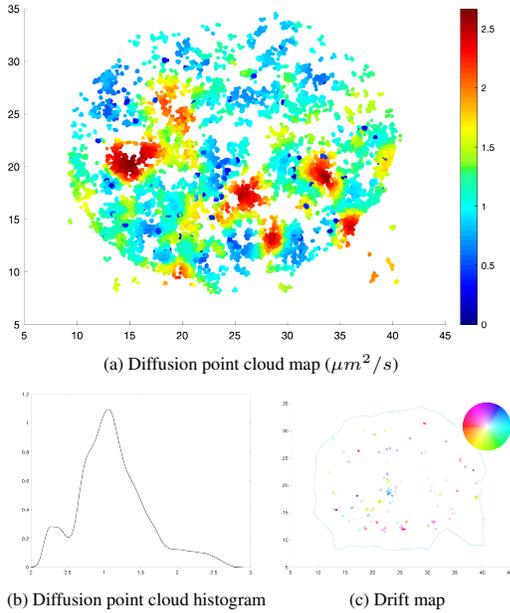
Experiments on real Data (3D)   Our method has also been used to analyse 3D LLSM fluorescence microscopy data and in 3D LLSM fluorescence microscopy. Figure 6 depicts Galectin-3 particle trajectories moving inside a SUM159 cell. The horizontal, vertical, and temporal resolutions are of 104 nm, 325 nm, and 2.25 s respectively. We can clearly distinguish between two regions (corresponding to two modes on the histogram) of low and high diffusion in the cell : the flat side of the specimen has higher diffusion than the rest of cell.

(a) Diffusion point cloud map ($\mu m^2/s$)

(b) Diffusion point cloud histogram

(c) Drift map
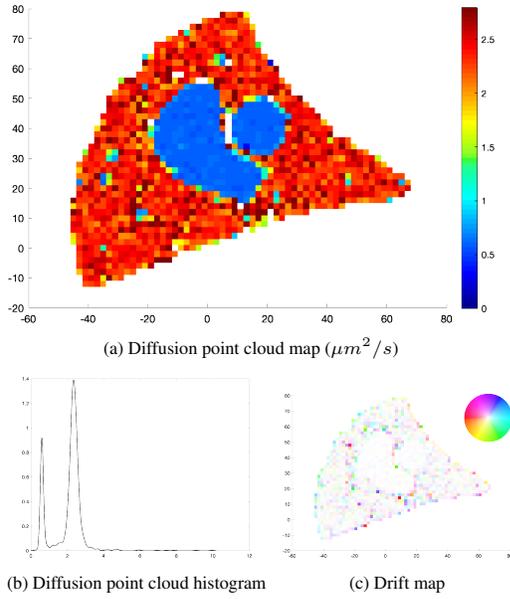
Fig. 3. Results obtained with method [10] on simulated data.

(a) Diffusion point cloud map ($\mu m^2/s$)

(b) Diffusion point cloud histogram

(c) Drift map

Fig. 5. Results obtained with method [10] on 2D PALM/STORM SPT data.

(a) Diffusion point cloud map ($\mu m^2/s$)

(b) Diffusion point cloud histogram

(c) Drift map

Fig. 4. Results obtained with our method on 2D PALM/STORM SPT data.

(a) Diffusion point cloud ($\mu m^2/s$)

(b) Diffusion point cloud histogram

(c) Drift vector field

Fig. 6. Results obtained with our method on 3D LLSM data.

5. CONCLUSION

We have proposed a new mapping method to robustly estimate diffusion and drift from particle tracks. Using a spatiotemporal kernel estimator guided by particle motion classification, this computational approach is more robust than previous methods and provides high-resolution estimations maps, and then is helpful to detect subtle biological events in image sequences.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] A. Einstein, "On the motion of small particles suspended in liquids at restrequired by the molecular-kinetic theory of heat," *Annalen Der Physik*, vol. 17, pp. 549–560, 1905.

[2] P. Langevin, "Sur la théorie du mouvement brownien," *Comptes-rendus de l'Académie des sciences*, vol. 146, pp. 530–533, 1908.

[3] E.J. Ambrose, "A surface contact microscope for the study of cell movements," *Nature*, vol. 178, pp. 1194, November 1956.

[4] D. Axelrod, "Cell-substrate contacts illuminated by total internal reflection fluorescence," *The Journal of Cell Biology*, vol. 89(1), pp. 141–145, April 1981.

[5] B.C. Chen et al., "Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution," *Science*, vol. 346(6208), October 2014.

[6] S.S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 346(6208), pp. 5–18, January 2004.

[7] N. Chenouard et al., "Multiple hypothesis tracking for cluttered biological image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2736–3750, November 2013.

[8] K. Jaqaman et al., "Robust single-particle tracking in live-cell time-lapse sequences," *Nature Methods*, vol. 5, pp. 695–702, August 2008.

[9] P. Roudot et al., "Piecewise-stationary motion modeling and iterative smoothing to track heterogeneous particle motions in dense environments," *IEEE Transactions on Image Processing*, vol. 26, no. 11, pp. 5395–5410, November 2017.

[10] N. Hozé et al., "Heterogeneity of ampa receptor trafficking and molecular interactions revealed by superresolution analysis of live cell imaging," *Proceedings of the National Academy of Sciences*, vol. 109, pp. 17052–17057, October 2012.

[11] V. Briane et al., "A statistical analysis of particle trajectories in living cells," *Physical Review E*, vol. 97, June 2018.

[12] V. Briane et al., "A sequential algorithm to detect diffusion switching along intracellular particle trajectories," *Bioinformatics*, June 2019.

[13] M. Lagardère et al., "Fluosim: simulator of single molecule dynamics for fluorescence live-cell and super-resolution imaging of membrane proteins," February 2020.