High-throughput super-resolution single particle trajectory analysis reconstructs organelle dynamics and membrane re-organization

P. Parutto[1*], J. Heck[2], M. Lu[3], C. Kaminski[3], M. Heine[2] E. Avezov[4], and D. Holcman[15 *]

September 2, 2021

Abstract

Super-resolution imaging can generate thousands of single-particle trajectories. These data can potentially reconstruct subcellular organization and dynamics, as well as measure disease-linked changes. However, computational methods that can derive quantitative information from such massive datasets are currently lacking. Here we present data analysis and algorithms that are broadly applicable to reveal local binding and trafficking interactions and organization of dynamic subcellular sites. We applied this analysis to the endoplasmic reticulum and neuronal membrane. The method is based on spatio-temporal time window segmentation that explores data at multiple levels and detects the architecture and boundaries of high density regions in areas that are hundreds of nanometers. By statistical analysis of a large number of datapoints, the present method allows measurements of nano-region stability. By connecting highly dense regions, we reconstructed the network topology of the ER, as well as molecular flow redistribution, and the local space explored by trajectories. Segmenting trajectories

[*1] Group of Data Modeling and Computational Biology, IBENS, Ecole Normale Supérieure,75005 Paris, France.[*]new address: UK Dementia Research Institute at the University of Cambridge and Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0AH, UK. [2] Research Group Functional Neurobiology at the Institute of Developmental Biology and Neurobiology, Johannes Gutenberg University Mainz, Mainz, Germany [3] Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge CB3 0AS, UK and [4] UK Dementia Research Institute at the University of Cambridge and Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0AH, UK. [5] DAMPT, University Of Cambridge, DAMPT and Churchill College CB30DS, United Kingdom.

at appropriate scales extracts confined trajectories, allowing quantification of dynamic interactions between lysosomes and the ER. A final step of the method reveals the motion of trajectories relative to the ensemble, allowing reconstruction of dynamics in normal ER and the atlastin-null mutant. Our approach allows users to track previously inaccessible large scale dynamics at high resolution from massive datasets. The algorithm is available as an ImageJ plugin that can be applied by users to large datasets of overlapping trajectories.

1 Introduction

Sub-cellular compartments are focused sites where large numbers of molecules dynamically interact to support cellular function [1]. The trajectories of ions and proteins as they move between the cytoplasm, plasma membrane and organelles are critical to cellular function [2, 3]. These dynamics occur at the Endoplasmic Reticulum (ER) [4, 5, 6], the mitochondrial network, endosomes and lysosomes and microtubules, and are impacted by the local properties of these different environments [7]. Several experimental paradigms can measure these constitutive molecular motions at subcellular sites, including the reciprocal fluorescence recovery after photobleaching (FRAP) [8, 9] which locally depletes fluorescence and measures time scales and fraction of recovery. Analysis of these data can reveal trafficking at a population level. In contrast, photoactivation [10] consists of activating molecules in a local region of the cell and reveals their spread over a transient time-frame. Combined with diffusion modeling and stochastic simulations, various biophysical properties can be measured, including diffusion coefficients and the fraction and time scale of recovery [11]. These methods provide information on the dynamic function of organelles, but are insufficient to identify and reconstruct high density regions. These approaches also cannot examine phase separation stability or the local spaces is explored by molecules at a nanoscale resolution. Statistical analysis [12, 13] of a large ensemble of super-resolution single-particle trajectories SPTs (Fig. 1A-B), has the potential to reveal local molecular interactions. Molecules are not uniformly distributed inside a cell but instead form heterogeneous aggregates, possibly in phase-separated nanodomains, characterized by high density regions (HDRs). Such regions are characterized by reduced velocity movement of molecules and confinement. These local areas can be enriched with calcium channels at neuronal synapses, store-operated calcium entry receptors such as STIM1 on spine apparatus, and can also be found at ER nodes [14, 15, 16, 17]. Interestingly, these ubiquitous structures are transient, yet

persist with a time scale that is longer than that associated with molecular trafficking. In short, many interactions that are critical to the cell are regulated and controlled by sub-cellular mechanisms that currently cannot be easily captured by quantitative analysis.

To determine the underlying physical properties of molecular trafficking, various computational modeling methods have been developed to analyze SPTs. These models include those based on classical free [18] and confined [19, 20, 9, 21] diffusion, active deterministic motion or a mixture of deterministic and stochastic models [22]. Based on this theoretical framework, analysis of SPTs has revealed the dynamics of local chromatin organization in the nucleus [23, 24, 25], synaptic receptor trafficking at neuronal synapses [26] and SVG virus assembly [12]. A significant recent advance is the analysis of massive numbers of overlapping SPTs. This statistical analysis can reveal the properties of molecular trajectories. However analyzing these data can potentially also allow quantification of membrane dynamics [27] and may give insight into organelle organization.

Here we have developed a method based on hybrid algorithms, and automated analysis pipeline for SPTs (Fig. 1C), based on the Langevin [28] equation of motion to provide statistical analysis of these data. This method estimates biophysical characteristics and is capable of reconstructing nanodomain sizes and boundaries using the classical physical model of a potential well [29, 28], well known since Kramer's work in 1940. This method allowed us to characterize calcium channel organisation in the membranes of hippocampal neurons. We also present an algorithm that reconstructs a network from SPTs based on the clustering of low-velocity trajectory fragments, and use it to reconstruct ER network topology as well as estimate the time scale of lysosome trafficking and ER network interactions. Finally, these methods allowed us to extract the motion of trajectories relative to the reconstructed network, thereby revealing the redistribution of trajectories inside the ER of normal and atlastin-null mutant cells [30]. The diversity of datasets used here demonstrate the broad applicability of these methods to cell biological processes. The methods are available to users as elementary ImageJ plugins.

2 Results

2.1 An algorithm that reconstructs nanodomains in high density regions

In order to extract sub-cellular regions where a high density of molecular interactions are occurring, we develop a novel computational approach and automated pipeline Fig. 1C, based on stochastic equations, multiscale analysis, optimal estimators and maximum likelihood statistics. In contrast with methods to extract density of points [31, 32] or to compute the Maximum Likelihood Estimator (MLE) [33, 34], the present approach combines density of points with local dynamics associated with the elementary displacement $\Delta X = \boldsymbol{X}(t + \Delta t) - \boldsymbol{X}(t)$, where $\boldsymbol{X}(t)$ is the position of the particle at time t [13, 14] to extract the field of forces. We could thereby assess the nature, organization and stability of a large amount of HDRs, by collecting statistics that can reveal hidden cellular organization. HDRs could previously only be characterized by manual curation based on extracting parameters of potential well, a concept in classical physics [35, 28] that describes the stability of dynamic systems, such as the motion of a bead attached to a spring. In contrast, using combined optimization methods, the present method allows users to automatically extract local diffusion coefficients, energetics, local field potential, and most importantly, local boundaries.

The method relies on an interesting observation that nanodomains of high density tend to have an elliptic shape. Thus the first step to detect them is to recover their center and boundaries (see method and eq. 12). Non-automated classification algorithms have used the density of points [33] or displacements (ΔX) separately. But these procedures often lead to parameter estimations that are not completely satisfactory due to a shallow minimum of the associated error function (see Method). This shallow minimum leads to large variability and possible mistakes in the estimation of most of the nanodomain parameters such as the boundary and the energy of nanodomains.

To overcome these difficulties, we developed a hybrid algorithm (Fig. 1C), described in the Methods section, which combines two independent procedures starting with a principal component analysis (PCA) to recover the elliptic boundary and followed by a MLE to compute the effective diffusion coefficient and drift properties.

To test this hybrid algorithm, we first constructed a ground truth dataset. This dataset consists of trajectories (Fig. S1 Aa-b) generated by the stochastic equation 2, where the force defining the nanodomain is defined by equa-

tion 12 and characterized by converging arrows (Fig. S1 Ac). We use a time discretization $\Delta t = 20$ and 50 ms and generated 10^3 trajectories of size 20 points. The new algorithm comprises three steps :

1. Automatic determination of bins with the highest-density of points (Fig. S1 Ba).

2. For each such bin, we iterate over square regions of increasing width w_k around the bin center. For each iteration k, we apply a PCA to estimate the semi-axis a_k, b_k, the center c_k and a score $\mathcal{L}$ (eq. 24) based on the points falling in the square of size w_k (Fig. S1 Bb-c). We iterate until we reach a threshold size defined by the user, as the maximum size of a well (Fig. S1Bc).

3. In the termination step, we select the optimal value of the iteration providing the optimal parameter (see Method paragraph 6.3).

The reconstruction is illustrated in Fig. S1C for three wells (see also Table S1 and S2 where we used two time steps $\Delta t = 20$ and 50 ms to test the reconstruction of circular and elliptic boundaries). Finally, we showed that this hybrid algorithm performs much better than two previous algorithms in estimating all parameters, especially those at low energy at 4kT (Table S1). Other algorithms are based on the local density of points [33], or on computing the drift term [14] only.
At this stage, we have thus validated the hybrid algorithm by using a ground-truth data sets to identify and automatically reconstruct nanodomains. Our novel hybrid algorithm out-performed existing models for all parameter estimations and in terms of defining the exact shape of the nanodomain.

2.2 Analysis of SPTs for the endogeneous voltage-gated calcium channels reveals their organization and weaker stability in nanodomains

We applied the hybrid algorithm to study the dynamics of calcium voltage channels (CaV2.1) tracking on the surface of neuronal cells for two over-expressed splice variants CaV2.1$_{\Delta 47}$ and CaV2.1$_{+47}$, previously shown to shape synaptic short-term plasticity [14]. We also examined endogenously tagged CaV2.1 channels. Using a large number of redundant SPTs (see Methods), we were able to automatically detect the nanodomains defined as HDRs. We plotted the trajectories, density, diffusion maps and the drift field associated with these HDRs (Fig. 2A-C). The algorithm allowed us to identify the geometrical characteristics of nanodomains approximated as

ellipses with semi-axis lengths $a = 143 \pm 51$ nm and $b = 104 \pm 33$ nm for Cav2.1$_{\Delta 47}$. These parameters are similar for Cav2.1$_{+47}$ (Fig. 2D), while it was smaller for endogenous Cav2.1 with $a = 100 \pm 47$ nm and $b = 73 \pm 31$ nm. The diffusion coefficient was $D = 0.091 \pm 0.052$ $\mu m^2/s$ for the $\Delta 47$ variant, $D = 0.087 \pm 0.045$ $\mu m^2/s$ for the +47 variant and smaller for endogenous CaV2.1 with $D = 0.069 \pm 0.051$ $\mu m^2/s$.
Interestingly although the associated energy was ~ 3.9 kT for the two variants and 3.3 kT for endogenous CaV 2.1, we found that the associated residence times of a receptor in a nanodomain was around $174 - 181$ ms and 94 ms respectively (Table S3). We also tested two methods to estimate the diffusion coefficient, either using a static grid [33] or a sliding disk convolved with a cosine function (Fig. S1) and used this approach in subsequent analyses. To conclude, these estimations of residence time were larger (almost double) compared to the ones we reported in [14] and thus our new approach quantifies channel dynamics in the neuronal membrane more precisely. The approach also captures differences in variant forms of channels, revealing a significant reduction in interactions for the endogenous fraction.

2.3 Time-lapse analysis reveals the stability of CaV nanodomains over time

To investigate the stability of nanodomains across time, we used a time-lapse analysis (Fig. S3) with sliding windows of 20 s and no overlap. This analysis allows determination of the lifetime of a nanodomain, which is given by the number of successive windows where it is detected (Fig. S4 A). For example, the trajectories obtained during a 250 s experiment are split into 13 20 s windows ($0 - 20$ s, $20 - 40$ s, ..., $240 - 260$ s). We searched for the presence of potential wells in each window (Fig S4A-B). To follow a well across successive time windows, we consider that two wells identified at time t_k and t_{k+1} are the same if the distance between their centers is less than 250 nm. The ensemble of consecutive times $(t_q, ...t_r)$ where a well is first detected at time t_q and disappears at time t_{r+1} is used to define the stability duration $\tau = t_r - t_q$. This analysis allows us to follow the size of the small and large axis of the wells and the associated energy over time (Fig. S4C-D). Finally, we found that 55% of the wells were present for more than 20 s (one time window) and that their average duration is ~ 56 s for both $\Delta 47$ and +47 variants (Fig. S4E and Table S4). However, nanodomain stability is reduced to $T = 46$ s for endogenous CaV2.1. All of these durations are longer than the ~ 30 s that we previously reported [14], indicating that the present algorithm advances our ability to capture the dynamics of these high

density enigmatic subcellular domains.

2.4 The hybrid algorithm reveals that ER node are nanodomains defined by an attracting field of force

To further explore the range of applicability of the present hybrid algorithm, we analysed SPTs recorded from the ER of HEK293T, COS-7 WT as well as atlastin knockout COS-7 dATL cells. In such mutant, the morphology of peripheral ER tubules is altered but it is unclear how the ER flow is affected. Since nodes have been previously characterized as HDRs [17], we asked here whether these nanodomains could further be defined as potential wells. Applying the hybrid algorithm reveals several potential wells (Fig. 3A-B red ellipses) precisely located at nodes forming high density regions. We further estimated the diffusion and (Fig. 3C) drift maps, observing converging arrows patterns in these regions (Fig. 3D and Fig. S4-A), a classical feature of potential wells [22]. Interestingly, these HDRs were characterized by ellipses with a large semi-axis $a = 219 \pm 71$ nm and $b = 155 \pm 56$ nm for HEK293T (see also Fig. 3E and table S6 for COS-7 WT and COS-7 dATL). The hybrid algorithm further reveals an average energy of $E = 3.3 \pm 0.9$ kT and a mean residence time of $\tau = 101 \pm 64$ ms. Interestingly, although the elliptic parameters are not much different in the case of COS-7 WT and COS-7 dATL, a difference can be observed in the dynamical parameters characterizing the transport of the material across the ER network. To conclude, the present hybrid algorithm reveals that ER nodes concentrate trafficking of luminal molecules by a spring-force type mechanism, the origin of which should be further explored.

3 Organelle network reconstruction from a large number of SPTs

We next wanted to check whether our approach can be used to define the structure of the ER and be generalized to other organelles. For these goals, we developed a novel method and algorithm to reconstruct the network from SPTs.

3.1 Graph Reconstruction Algorithm (GRA) to unravel the ER network

Although SPTs can be used to explore the ER network architecture [17], we still lacked a method for automated reconstruction of the ER. We therefore

developed the algorithm to see if the dynamic architecture of this complex organelle can be recovered from SPTs.

The starting point is to colour trajectory displacements depending on their instantaneous velocity. This reveals a dynamical segregation of the ER into nodes and tubules (Fig. 4A-C). Based on this segregation, we developed the Graph Reconstruction Algorithm (GRA) to recover the ER structure from SPTs. The GRA consists of three steps: 1-identifying the HDRs formed by low-velocity (blue regions) trajectory fragments forming the nodes of the graph (Fig. 4D-G), 2- determining the high-velocity trajectory fragments connecting the previously detected nodes (Fig. 4H-K) and 3-constructing the associated graph (Fig. 4L), as described in Materials and Methods. The first step of the method relies on the density-based clustering algorithm (dbscan) [36], which requires that the distribution of trajectories be quite heterogeneous so that small regions of high density are well-separated. In an additional step, a direction can be attributed to the links (i.e. tubules) based on the percentage of trajectories going in the same direction between two nodes: when the percentage is around 50%, the node is bidirectional, however if the percentage is much lower, the node is defined as unidirectional. This procedure allows us reconstruct a two-dimensional graph for the organelle network that can be used to study further statistical properties.

3.2 Graph Reconstruction Algorithm of the ER network of COS-7 dATL cells reveals aberrant organization and trafficking

Next we examined whether disruptions to organelle structure and dynamics can be captured using our algorithm. Atlastin-1 is a GTPase that mediates homotypic membrane fusion in the ER. We analyzed SPTs recorded from an ER luminal probe in atlastin knockout (dATL) COS-7 cells. We present colour-coded trajectories that follow their instantaneous velocity (Fig. 5A-B), the density and diffusion maps (Fig. 5C-D), as well as the histogram of the apparent diffusion coefficients (Fig. 5E), revealing a mean of $D_{app} = 1.56 \pm 0.83\ \mu m^2/s$.

We applied the GRA and obtained a graph reconstruction of the ER (Fig. 5F) where the HDRs (red) are connected by blue segments. This approach allows quantification of the distribution of distances between nodes with a mean $d_{nodes} = 0.83 \pm 0.27\ \mu m$ (Fig. 5G) and the node projected area $S_{nodes} = 0.15 \pm 0.12\ \mu m^2$ (Fig. 5H). Note that the RGA could miss some non-explored ER regions or regions that are sub-sampled. To conclude,

the GRA allows automated reconstruction of organelle networks based on SPT exploration once there is enough heterogeneity in the distribution of datapoints.

3.3 Graph Reconstruction Algorithm and SPTs segmentation reveal the duration of ER-lysosome interactions

To demonstrate the broad applicability of the algorithms, we analyzed an ensemble of trajectories from lysosomes (Fig. 9.6A). These trajectories are characterized by a distribution of instantaneous velocities in the range $[0 - 3.5]$ $\mu m/s$ (Fig. 9.6B). However low and high velocities are not segregated, but are instead found in similar regions (Fig. 9.6A left and right). The distribution $f(v)$ of velocities (Fig. 9.6B), can be fitted by a sum of two exponentials

$$f(v) = A\exp(-\frac{v}{v_0}) + B\exp(-\frac{v}{v_1}), \tag{1}$$

where a best fit approximation gives $v_0 = 0.06$ $\mu m/s$ (95% confidence interval $[0.057, 0.072]$) and $v_1 = 0.6$ $\mu m/s$ (95% confidence interval $[0.383, 1.322]$) (coefficient of determination $R^2 = 0.95$), with $A = 0.20$ and $B = 0.013$. This fit suggests that the distribution of lysosomes is largely driven by low velocity components. The rare appearance of high velocity components suggests a possible switch between slow and fast motions. Finally, note that 86% of displacements are associated with a velocity of less than 0.5 $\mu m/s$ and 12.9% are in the range of [0.5-1.5] $\mu m/s$.

To further study how lysosomes move in the cytoplasm, we computed relevant density and diffusion maps (Fig. 9.6C-D) and found that the motion had a diffusion component (Fig. 9.6E), with an average apparent diffusion coefficient of $D_{\text{app}} = 0.062 \pm 0.040$ $\mu\text{m}^2/\text{s}$. Interestingly, regions of low diffusion coefficients colocalized with regions of high density in the density map [37, 27] (Fig. 9.6A-C).

We then isolated regions of high density using a method based on the density of points (Methods subsection 6.8.2), revealing an ensemble of $n = 95$ HDR sub-domains, approximated by ellipses (magenta in Fig. 9.6F) of semi-axis lengths $a = 516 \pm 196$ nm (large) and $b = 278 \pm 143$ nm (small) (Fig. 9.6G). By considering the displacements connecting different regions, we reconstructed (Method subsection 6.8.2) a network explored by the lysosomes, Fig 9.6H, where high density regions (red circles) are connected by direct lines (yellow). Interestingly, the histogram of average velocities between these regions is not symmetric (Fig. 9.6I) with a mean velocity $v = 1.03 \pm 0.32$

μm/s, which clearly deviates from diffusion, as computed from the Rayleigh distribution. This deviation suggests that the transitions between these regions are driven by an active motion. Moreover, the overlay between ER (white) and the lysosome reconstructed network (Fig. 9.6G) suggests that the lysosome trajectories follow the topology of the ER network [7]. To conclude, our analysis reveals that lysosomes travel along a network that strongly colocalizes with the ER. However, high and low velocities occur in similar regions. Since lysosomes move along microtubules, this present statistical analysis suggests that the ER-microtubule network shapes lysosome trafficking.

4 Trajectory re-synchronization approach reveals single local molecular dynamic exploration inside an ensemble

In the previous result sections, we reconstructed the networks hidden inside SPTs data. We shall now introduce a last step in our method which re-synchronizes trajectories that fall inside the same subcellular are, but were acquired at random times. This approach allows us to study the dynamics of trajectories with respect to the ensemble of trajectories that visit the same spatial region. The approach also enables determination of the local spatio-temporal properties that trajectories explore at the single unit level.

4.1 Trajectory segmentation reveals ER-lysosomes interaction time scale

To study the possible interactions between lysosomes and the ER (Fig. 9.7A) reconstructed network in Fig. 9.6, we focused on the confined portion found along individual trajectories (see method section 6.8.3 and Fig. 9.7A). We hypothesise that the lysosome motion can switch between a directed and confined motion (Fig. 9.7B). We first show that the lysosome can indeed switch between directed and confined motion (Fig. 9.7C).
To recover the size of the confinement areas, we fitted ellipses over these regions and obtained average semi-axes lengths (Fig. 9.7B) of $a = 232 \pm 77$ nm (large) and $b = 94 \pm 47$ nm (small). Furthermore, this approach allowed us to estimate the confinement strength λ by considering that the confined motion could be generated by a spring force, modeled by an Ornstein-Uhlenbeck process [28]. We found that the spring force is $\lambda = 0.123 \pm 0.025\ s^{-1}$ (Fig. 9.7C), associated with an average local diffusion coefficient of $D = 0.032 \pm 0.002$

μm^2/s (Fig. 9.7D) for a total of $n = 818$ confinement regions. Finally, the distribution of times in confined regions could be well-approximated by a single exponential with a time constant $\tau = 5.35$ s (Fig. 9.7E). The average residence time of lysosomes in these regions is $\bar{\tau} = 30 \pm 12$ s which can be interpreted as a time where lysosomes could interact with the ER. To conclude, the present algorithm reveals that as lysosomes travel along a network that strongly colocalizes with the ER, the velocity can switch from large to small displacements and the trajectories can become restricted into regions of size $\sim$ 200nm, on a timescale of 5 s. This could correspond to a change in directionality of movement or a direct interaction with the ER. Our approach can therefore extract changes in lysosome dynamics that may reflect functional interactions from complex data.

4.2 Trajectory re-synchronization approach shows how a single trajectory explores single nodes

After an ER network is reconstructed from SPTs by the GRA (Fig. 9.8A), the node-tubule topology emerges. Thus it becomes possible to study how trajectories locally explore the network by synchronizing them upon exit from a chosen node (Fig.9.8B). Interestingly, we found that the mean instantaneous velocity at exit is $v_e = 30.2 \pm 10.2$ $\mu m/s$ and keeps decreasing during the next 200 to 300 ms (Fig. 9.8C). Escape occurs in equal directions (Fig. 9.8D), as shown in four examples where we followed their dispersal. To characterize this dispersion, we plotted the dispersion index (section S5), revealing two phases (Fig. 9.8B), one below 100ms, showing a rapid dispersion, followed by a second phase with less expansion. These two phases can be interpreted as : for the first one, trajectories escape from a well, then in the second phase, trajectories tend to be re-captured for a certain time in nodes, thus preventing a fast exploration of the network.

4.3 Trajectory re-synchronization reveals novel local dynamics within dATL ER-tubules

We next analyzed SPTs recorded in the ER of COS-7 dATL cells which lack the ER membrane-shaping protein atlastin (double knockout of the ATL-1/2 genes [30]) and exhibit a disrupted peripheral ER morphology with elongated tubules [30]. These SPTs revealed changes in the local space exploration of trajectories (Fig. S7A). Under normal conditions, trajectories are mostly located in nodes [17], while here trajectories predominantly explore long tubules (fig. S7B) with lengths of 5.4 ± 2.44 μm, much longer than the

$\sim$1 μm found for the tubules of WT cells. Inside these tubules, we found that trajectories exhibit a "stuttering" behavior around different positions that lasts for seconds. To characterize this behavior, we estimated several parameters such as the duration that a trajectory spent around a given position $\tau_{ll} = 79 \pm 76$ ms (Fig. S7D), the transition time between different positions $\tau_t = 27 \pm 15$ ms (Fig. S7E), the length of a transition step $\Delta_{ll} = 0.53 \pm 0.45$ μm(Fig. S7F) and finally the standard deviation around the stable positions $SD = 0.14 \pm 0.07$ μm (Fig. S7G).

To conclude, following the reconstruction of networks using our algorithm we were able to re-position and re-synchronize SPTs. Using analyses of the ER-lysosome, ER in normal COS-7 cells, and ER in COS-7 dATL cells, the algorithm revealed trajectories explored by the local geometrical space and the associated time scales.

5 Discussion and concluding remarks

We present here a general method and the associated algorithms that can automatically characterize nanodomains where trajectories accumulate. Our approach generate graph representations organelle networks from SPTs. Automatically finding nanodomains is useful to extract large statistics (size, energy of potential wells, mean residence time of particles) and compare their properties. Further, by quantifying the trajectories inside and outside HDRs, we could recover membrane organization, as well as determine the local redistribution of dynamic organelles and proteins. By reconstructing a graph of ER or lysosome networks from SPTs we can recover molecular flow at the nanoscale level. We found here that HDRs could be characterized as attractive nanoregion (potential wells), and this generic representation suggests a universal mechanism of molecular stabilization that probably requires further investigation. Interestingly, these structures can be transiently remodeled in time as revealed by the present time-lapse analysis.

5.1 Universality of high density nanoregions characterized as potential wells

High density nanoregions are now associated with potential wells for several receptors and channels such as CaV [14], AMPAR [13], Glycine receptors [38] or GAGs [39, 40]. Interestingly, nodes of the ER can also be characterized as potential wells, which may reflect retention of luminal flow or to allow protein maturation. This representation suggests a generic membrane organization

to retain particles (receptors, channels, proteins,etc...) in a field of force with long-range interactions.

Interestingly, the geometry of these regions and their energy profile are independent of the experimental conditions, further confirming again their stability. Note that the physical nature of potential well remains unclear [41]. The present method could be applied to analyse molecular crowding and the dynamics of nanodomains, thus clarifying processes relevant for phase separation at synapses [42]. With the development of new labelling methods, improved fluorophores and the ability to tag endogenous populations of molecules via CRISPR/Cas, it will soon be possible to investigate phase separation at a population level and with SPT, to track endogenous dynamics, offering novel opportunities for the present approach.

5.2 Trafficking in networks

The present method allowed us to reconstruct the ER network. The reconstruction algorithm reveals that lysosome trajectories follow ER network. This reconstructed network is further segregated into nodes and links, but low and high velocities are now much more mixed compared to the reconstruction obtained from luminal proteins. It is possible that lysosomes follow the cytoskeleton network which is correlated with the ER [7].

In addition, the distribution of lysosome velocity follows a double exponential (displacement histogram in Fig. 9.6) with fast ($\sim$0.6 $\mu m/s$) and slow ($\sim$0.06 $\mu m/s$) components. However, a more detailed analysis revealed that these velocities can be further subdivided into: 1) confined motion (Fig. 9.7) characterized by a residence time of $\sim$ 5 s. 2) deterministic motion between HDRs (Fig.9.7 F-G), characterized by a distribution of velocity with an average of 1.03 μm/s. It would be interesting to better characterize the switch between confined and rectilinear motion. Regions of deterministic velocities and those where diffusion can be found are often not well separated, suggesting that lysosomes can use various modes of transport, independently of the subregions where there are located. We found however some regions characterized by a high density of trajectories, with a low velocity, suggesting that there are possible trapping mechanisms to retain lysosomes in specific subregions of the ER, possibly at exit sites [43]. This mode of motion is quite different from the internal motion inside the ER lumen or on its membrane: in the first case, the node-tubule topology is associated with a diffusion-drift dynamics, while in the second case, the motion on the ER membrane is likely diffusion-based [17]. To conclude, the present analysis reveals that the confined time is around 5s, suggesting that during this time, lysosomes

may be trapped in interaction with the ER.

Future work should reveal interaction times between Lysosomes and the ER. By applying our algorithm to different cells and organelles, we have shown that information on the boundaries and dynamics of subcellular interactions can be revealed from large SPT datasets. The automated algorithms presented here can be applied by users to analyse hundreds of thousands of trajectories and to study nanodomains with almost no human intervention and are available as an elementary ImageJ plugin.

6 Material and Methods

The method section is organized as several subsections describing the physical model to interpret SPTs, the associated data estimators, the potential well description, several hybrid algorithms to analyse automatically the regions of high densities and finally a method to reconstruct a graph for organelle networks.

6.1 Diffusion model, velocity, vector fields and empirical estimators

In the Smoluchowski's limit of the Langevin equation [28], the position $\boldsymbol{X}(t)$ of a molecule is described by

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

where $F(\boldsymbol{X})$ is a field of force, W is a white noise, γ is the friction coefficient [44] and D is the diffusion coefficient. At a coarser spatio-temporal scale, the motion can be coarse-grained as a stochastic process [13, 37]

$$\dot{X} = a(\boldsymbol{X}) + \sqrt{2}B(\boldsymbol{X})\dot{W}, \tag{3}$$

where $a(\boldsymbol{X})$ is the drift field and $B(\boldsymbol{X})$ the diffusion matrix. The effective diffusion tensor is given by $D(\boldsymbol{X}) = \frac{1}{2}B(\boldsymbol{X})B^T(\boldsymbol{X})$ ($.^T$ denotes the transposition) [45, 44]. The drift of the stochastic model from eq. 3 can be recovered from SPTs acquired at any infinitesimal time step Δt by estimating the conditional moments of the trajectory displacements $\Delta\boldsymbol{X} = \boldsymbol{X}(t+\Delta t) - \boldsymbol{X}(t)$ [44, 46, 47, 37, 22]

$$a(x) = \lim_{\Delta t \to 0} \frac{\mathbb{E}[\Delta\boldsymbol{X}(t) \,|\, \boldsymbol{X}(t) = x]}{\Delta t}, \tag{4}$$

$$D(x) = \lim_{\Delta t \to 0} \frac{\mathbb{E}[\Delta\boldsymbol{X}(t)^T \Delta\boldsymbol{X}(t) \,|\, \boldsymbol{X}(t) = x]}{2\Delta t}. \tag{5}$$

The notation $\mathbb{E}[\cdot \,|\, \boldsymbol{X}(t) = x]$ represents averaging over all trajectories that are passing at point x at time t. To estimate the local drift $a(\boldsymbol{X})$ and diffusion coefficients $D(\boldsymbol{X})$ at each point $\boldsymbol{X}$ of the membrane and at a fixed time resolution Δt, we use a similar procedure to the one for the estimation of the density (section 6.4) based on a square grid.
The local estimators to recover the vector field and diffusion tensor [33]

consist in grouping points of trajectories within a lattice of square bins $S(x_k, \Delta x)$ centered at x_k and of width Δx. For an ensemble of N two-dimensional trajectories $\{\boldsymbol{X}_i(t_j) = (x_i^{(1)}(t_j), x_i^{(2)}(t_j)), i = 1..N, j = 1..M_i\}$ with M_i the number of points in trajectory $\boldsymbol{X}_i$ and successive points are recorded with an acquisition time $t_{j+1} - t_j = \Delta t$. The discretization of eq. 4 for the drift $a(x_k) = (a^{(1)}(x_k), a^{(2)}(x_k))$ in a bin centered at position x_k is

$$a^{(u)}(x_k) \approx \frac{1}{N_k} \sum_{i=1}^{N} \sum_{j=0, x_i(t_j) \in S(x_k, \Delta x)}^{M_i - 1} \left(\frac{x_i^{(u)}(t_{j+1}) - x_i^{(u)}(t_j)}{\Delta t} \right), \quad (6)$$

where $u = 1..2$ and N_k is the number of points from any trajectory falling in the square $S(x_k, r)$. Similarly, the components of the effective diffusion tensor $D(x_k)$ are approximated by the empirical sums

$$D^{(u,v)}(x_k) \approx \frac{1}{N_k} \sum_{i=1}^{N} \sum_{j=0, \boldsymbol{X}_i(t_j) \in S(x_k, \Delta x)}^{M_i - 1} \frac{[x_i^{(u)}(t_{j+1}) - x_i^{(u)}(t_j)][x_i^{(v)}(t_{j+1}) - x_i^{(v)}(t_j)]}{2\Delta t}. \quad (7)$$

The centers of the bin and their size Δx are free parameters that are optimized during the estimation procedure.

6.1.1 Variant estimation of the diffusion coefficient and drift

To increase the accuracy of the diffusion and drift maps, we weighted the points in the moving windows with a cosine function (would also be possible to use wavelets). In that case, the new estimator for the drift field is now

$$a^{(u)}(x_k) \approx \frac{\sum_{i=1}^{N} \sum_{j=0, x_j(t_j) \in \mathcal{D}(x_k, r)}^{M_i - 1} \left(\frac{\left(x_i^{(u)}(t_{j+1}) - x_i^{(u)}(t_j) \right)}{\Delta t} w_{i,j}(x_k, r) \right)}{\sum_{i=1}^{N} \sum_{j=0, \boldsymbol{x}_i(t_j) \in \mathcal{D}(x_k, r)}^{N_s - 1} w_{i,j}(x_k, r)} \quad (8)$$

with N_k the number of points of the trajectories falling in the disk $\mathcal{D}(x_k, r)$ of radius r and centered at x_k. The weight of a displacement starting at $X_i(t_j)$ with respect to the disk $\mathcal{D}(x_k, \Delta x)$ is given by

$$w_{i,j}(x_k, r) = \cos \left(\frac{\pi}{2} \frac{||\boldsymbol{X}_i(t_j) - x_k||}{r} \right), \quad (9)$$

with $||.||$ the Euclidean norm. In that case, we can choose a refined grid $S(x_k, (\Delta x)')$ with bin size $(\Delta x)' = \Delta x / l_{sc}$, where l_{sc} is a scaling factor. The role of the cosine weights w is to decrease continuously the influence of the points falling near the boundary.
Similarly, the generalized formula for the effective diffusion tensor $D(x_k)$ are given by the weighted sums

$$D^{(u,v)}(x_k) \approx \frac{\sum_{j=1}^{N} \sum_{i=0, \boldsymbol{X}_i(t_j) \in \mathcal{D}(x_k,r)}^{M_i-1} \frac{\left((x_i^{(u)}(t_{j+1}) - x_i^{(u)}(t_j))(x_i^{(v)}(t_{j+1}) - x_i^{(v)}(t_j))\right) w_{i,j}(x_k,r)}{2\Delta t}}{\sum_{i=1}^{N} \sum_{j=0, \boldsymbol{X}_i(t_j) \in \mathcal{D}(x_k,r)}^{M_i-1} w_{i,j}(x_k,r)},$$

where the weights w are given by eq. 9.

6.1.2 Local point density estimation

The local density of points ρ can be determined using a procedure similar to the drift or diffusion estimation by the image plane into a square bin $S(x_k, \Delta x)$. We then compute for each square of S centered at x_k

$$\rho_{\Delta x}(x_k) = \frac{N_k}{(\Delta x)^2}, \tag{10}$$

where N_k is the number of trajectory points falling into the bin centered at x_k. In practice, it usually helps to smooth this density estimation by applying a local average using a small $d \times d$ kernel with $d \sim 1, 3, 5$.

6.2 Estimating potential well parameters

In this subsection, we present the estimators for the two main parameters of potential wells: the extent of their boundary and their associated energy [13, 38, 33]. We recall that the diffusion coefficient inside a well is considered to be constant and is computed from the second moment of the displacement for all points falling inside the boundary of the well (see eq. 4). The potential well model allows to estimate the residence time using the classical escape time formula [44, 13] for a circular well

$$\tau_e \approx \frac{Dr^2}{4A^2} e^{\frac{A}{D}}, \tag{11}$$

with r the radius of the well, A its attraction coefficient and D its diffusion coefficient. In the case of an elliptic well, we obtain an approximate circular boundary using the harmonic mean of the semi-axes $r = \sqrt{ab}$, where a and b are the large and the small-axes lengths respectively. This approximation holds for $a \approx b$.

6.2.1 Parabolic potential well representation

To extract the energy of well, we consider the basin of attraction of a truncated elliptic parabola with the associated energy function

$$U(\boldsymbol{X}) = \begin{cases} A\left[\left(\dfrac{x^{(1)} - \boldsymbol{\mu}^{(1)}}{a}\right)^2 + \left(\dfrac{x^{(2)} - \boldsymbol{\mu}^{(2)})}{b}\right)^2 - 1\right], & \boldsymbol{X} \in \mathcal{B} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where $\boldsymbol{X} = [x^{(1)}, x^{(2)}]$, $\boldsymbol{\mu} = [\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}]$ is the center of the well, a, b are the elliptic semi-axes lengths and the elliptic boundary is defined by

$$\mathcal{B} = \{\boldsymbol{X}\text{such that } A\left[\left(\frac{x^{(1)} - \boldsymbol{\mu}^{(1)}}{a}\right)^2 + \left(\frac{x^{(2)} - \boldsymbol{\mu}^{(2)})}{b}\right)^2 - 1\right] = 0\}. \tag{13}$$

6.2.2 Recovering the center $\boldsymbol{\mu}$

The center of the nanodomain $\mathcal{B}$ is estimated as the center of mass of the cloud of points falling inside the HDR. We use the empirical averaging formula

$$\boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{X}_i, \tag{14}$$

where N is the total number of points such that $X_i \in \mathcal{B}$.

6.2.3 Covariance matrix Σ

We use the sample estimator of the Covariance matrix defined for a cloud of N two-dimensional points $\boldsymbol{X}_i = (x_i^{(1)}, x_i^{(2)})$ as

$$\sigma^{(u,v)} = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i^{(u)} - \mu^{(u)}\right)\left(x_i^{(v)} - \mu^{(v)}\right). \tag{15}$$

6.2.4 Confidence ellipse estimation $\varepsilon = (\mu, a, b, \varphi)$

We define the boundary of the well as the $X\%$ confidence ellipse of the associated Gaussian density distribution of center $\boldsymbol{\mu}$ and covariance Σ. Using the singular value decomposition method, we decompose the covariance matrix Σ as

$$\Sigma = UDV \tag{16}$$

where U, V are unitary matrices and D is diagonal. The values in D represent the variance in each dimension along the principal components. The values of D follow a Chi-Square distribution with $n = 2$ degrees of freedom. Therefore the semi-axes lengths a, b can be obtained at the $x\%$ from D as

$$a = \sqrt{\psi_x D^{(1,1)}}, \qquad b = \sqrt{\psi_x D^{(2,2)}}, \tag{17}$$

with ψ_x is given by $P(v < \psi_x) = x$ for a Chi-Square distribution with two degrees of freedom (for example $\psi_{99} = 9.210$, $\psi_{95} = 5.991$ and $\psi_{90} = 4.605$). Finally, the orientation φ of the ellipse is defined by the angle

$$\varphi = \text{atan}\left(\frac{U^{(2,1)}}{U^{(1,1)}}\right). \tag{18}$$

6.2.5 Maximum likelihood estimators (MLE) based on an Ornstein-Uhlenbeck model

Using the potential well representation from eq. 2 in the stochastic model presented in eq.12 leads to a truncated Ornstein-Uhlenbeck process, centered at μ, with an attraction coefficient λ and diffusion coefficient $\sigma = \sqrt{2D}$. The probability density function $p(\boldsymbol{X}(t_{j+1})|\boldsymbol{X}(t_j))$ for $j = 1..(M-1)$ of observing two successive positions of the same trajectory $\boldsymbol{X}(t_j)$ and $\boldsymbol{X}(t_{j+1})$, separated by a time step $t_{j+1} - t_j = \Delta t$ is given by

$$\boldsymbol{X}(t_{j+1})|X(t_j) \sim \mathcal{N}(m(X(t_j)), s), \tag{19}$$

with

$$m(X(t_j)) = X(t_j)e^{-\lambda\Delta t} + \mu(1 - e^{-\lambda\Delta t}) \tag{20}$$

and

$$s = \frac{\sigma^2(1 - e^{-2\lambda\Delta t})}{2\lambda}, \tag{21}$$

which we rewrite as

$$\begin{aligned} m(X(t_j)) &= \mu\lambda\beta + X(t_j)(1 - \lambda\beta) \\ s &= \sigma^2(\beta - \tfrac{1}{2}\lambda\beta^2), \end{aligned} \tag{22}$$

and

$$\beta = \frac{1 - e^{-\lambda\Delta t}}{\lambda}. \tag{23}$$

The log-likelihood function of observing the successive pairs $(\boldsymbol{X}_i(t_j), \boldsymbol{X}_i(t_{j+1}))$, $i = 1 \dots N$, possibly from various trajectories, is given by

$$\begin{aligned}
\mathcal{L}(\mu, \lambda, \sigma | X_1, .., X_n) &= \sum_{i=1}^{N} \log(p(X_i(t_{j+1}), X_i(t_j)) \qquad (24)\\
&= \sum_{i=1}^{N} \left[\log\left(\frac{1}{\sqrt{2\pi s}}\right) - \frac{(X_i(t_{j+1}) - m(X_i(t_j)))^2}{2s} \right] \\
&= -\frac{1}{2} \sum_{i=1}^{N} \left[\log(2\pi s) + \frac{(X_i(t_{j+1}) - m(X_i(t_j)))^2}{s} \right].
\end{aligned}$$

The corresponding maximum likelihood estimators for λ and D, $\tilde{\lambda}$ and $\tilde{D}$ are obtained by solving the system of equations

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \mu} &= 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= 0 \qquad (25)\\
\frac{\partial \mathcal{L}}{\partial D} &= 0,
\end{aligned}$$

from which we obtain the empirical estimators for the drift coefficient

$$\tilde{\lambda}_N = -\frac{1}{\Delta t} \log \left(\frac{\left(\sum_{i=1}^{N} X_i(t_{j+1}) X_i(t_j) \right) - \left(\frac{1}{N} \sum_{i=1}^{N} X_i(t_j) \right) \left(\sum_{i=1}^{N} X_i(t_{j+1}) \right)}{\left(\sum_{i=1}^{N} X_i(t_j)^2 \right) - \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^{N} X_i(t_j) \right)^2} \right), \qquad (26)$$

and the diffusion coefficient:

$$\tilde{D}_N = \frac{\lambda}{N(1 - e^{-2\lambda \Delta t})} \sum_{i=1}^{N} [X_i(t_{j+1}) - m(X_i(t_j))]^2. \qquad (27)$$

Note that the parameter λ in eq. 27 can be computed from the estimator $\tilde{\lambda}_N$.

6.3 Hybrid density-drift algorithm

In this sub-section, we present two variants of an algorithm to detect the main characteristics of a potential well from some observed trajectories: the center μ, the semi-axes lengths $a \geq b$, the orientation φ, the field strength A, the diffusion coefficient D and the energy E.

6.3.1 Fixed Spatial Scale Hybrid Density-Drift (FSHDD) Algorithm

Initiation Search for high-density regions: the image is partitioned by a grid $G_{\Delta x}$ with square bins of size Δx. from which we compute the density map $\rho_{\Delta x}(x)$ (eq. (10)). We then select the bins from $\rho_{\Delta x}(x)$ with the highest $d\%$ density as possible regions containing a potential well.

Iterations For each region obtained in the initiation step, we apply an iterative procedure that is going to consider increasingly larger square neighborhoods around this region. For each iteration $k = 1..K$, we keep only the trajectories contained inside the square $\Gamma_{k,\Delta x}$ of size $[(2k+1)\times(2k+1)](\Delta x)^2$ and centered at point μ_{k-1}. The point μ_k is the center of mass of points falling in the square Γ_k (eq. 14) and μ_0 is the center of the initial high-density bin. The elliptic semi-axes a_k, b_k are computed as the $x\%$ confidence ellipse (eq. 17) from the covariance matrix C_k (eq. (15)) and the angle φ_k is the orientation of the ellipse (eq. 18). These parameters define the elliptic boundary of the well at iteration k

$$\varepsilon_k(\Delta x) = (\mu_k, a_k, b_k, \varphi_k). \tag{28}$$

The field coefficient A_k and diffusion coefficient D_k are computed from eq. 26 and eq. 27 respectively, for the trajectories contained inside ε_k. Specifically, we obtain A_k from

$$\lambda_k^{(1)} = \frac{2A_k}{a_k^2} \text{ and } \lambda_k^{(2)} = \frac{2A_k}{b_k^2}. \tag{29}$$

We repeat this procedure K times, with $K = \lceil \frac{M_s}{\Delta x} \rceil$ for the spatial parameter Δx and the maximum region size M_s, defined by the user.

Termination This step consists in selecting the best iteration among K: we evaluate for each iteration $k > 1$ the likelihood score $\mathcal{L}_k = \mathcal{L}(\mu_k, \lambda_k, \sigma_k | X_1, ..X_p \in \varepsilon_k - \varepsilon_{k-1})$ defined by eq. (24) but computed for sub-trajectories falling in the ring formed by the ellipses ε_{k-1} and ε_k. The best iteration k^* is selected as the first local maximum of the curve $\mathcal{L}_k$.

6.3.2 Multiscale Hybrid algorithm (MSHA)

We generalize the hybrid density-drift (FSHDD) algorithm defined above for a fixed spatial scale, by now varying the grid size Δx, in the range

$\Delta x_1 < \Delta x_2 < .. < \Delta x_N$ selected by the user. The purpose of this MHSA is to select the optimal size Δx_{i^*} that minimizes the error

$$k^*_{\Delta x_i^*} = \min_{i=1..N} \mathcal{L}(\mu_{k^*}, \lambda_{k^*}, \sigma_{k^*} | X_1, ..X_p \in \varepsilon_{k^*}(\Delta x_i)), \tag{30}$$

where $k^*_{\Delta x_i^*}$ is the iteration that minimizes $\mathcal{L}$ across all the spatial scales Δx_i, $i = 1..N$.

6.4 Density-Based Algorithm

The Density-Based Algorithm (DBA) uses the density of points estimated around the local density maximum of a high-density region. The algorithm uses the idea of level set of the Gaussian density distribution of points inside a potential well. We define the level set Γ_α with respect to a local maximum M^* as the ensemble of all trajectory points falling in bins with a density greater than αM^*:

$$\Gamma_\alpha = \{\boldsymbol{X}_i \,\text{such that}\, \rho_e(\boldsymbol{x}) > \alpha M^*\}, \tag{31}$$

where ρ_e is the empirical point density, estimated over the bins of a square grid (eq. 10) and $\alpha \in [0, 1]$ is a density threshold. For the points $\boldsymbol{X}_i = (x_i^{(1)}, x_i^{(2)})$ located in Γ_α, the center $\boldsymbol{\mu}$ of the distribution is approximated by the empirical estimators based on eq. 14 but restricted to the points in Γ_α:

$$\hat{\mu}_\alpha^{(u)} = \frac{1}{N_p} \sum_{\{k=1, \boldsymbol{X}_k \in \Gamma_\alpha\}}^{N_p} x_k^{(u)}, \tag{32}$$

with N_p the number of points in the ensemble Γ_α and $u = 1..2$. Similarly, we extend the estimator for the covariance matrix Σ (eq. 15) to

$$\hat{\sigma}_\alpha^{(u,v)} = \frac{1}{N_p - 1} \sum_{\{k=1, \boldsymbol{X}_k \in \Gamma_\alpha\}}^{N_p} (x_k^{(u)} - \hat{\mu}_\alpha^{(u)})(x_k^{(v)} - \hat{\mu}_\alpha^{(v)}). \tag{33}$$

We now define the density-based algorithm:

Initiation Search for high-density regions: the image is partitioned by a grid $G_{\Delta x}$ with square bins of size Δx from which we compute the density map $\rho_{\Delta x}(x)$ (eq. (10)). We then select the bins from $\rho_{\Delta x}(x)$ with the highest $d\%$ density as possible regions containing a potential well.

Iterations For each selected high-density bin, we initiated the well center $\tilde{\boldsymbol{\mu}}_0$ at the center of the bin. We then construct a refined grid centered at $\tilde{\boldsymbol{\mu}}_0$ and with bin size $\Delta x/p$, where p is chosen by the user (usually around 2 to 4). In this grid, we compute the centers $\boldsymbol{\mu}_{0,\alpha_k}$, for different values of α: $\alpha_1 < \ldots < \alpha_N$, selected by the user. The refined center $\boldsymbol{\mu}_0$ is obtained as the center of mass of the centers $\boldsymbol{\mu}_{0,\alpha_k}$ for $k = 1..N$.

We then apply an iterative procedure that considers increasingly larger concentric annulus of center $\boldsymbol{\mu}_0$, width Δ_r and radius r_k for $k = 1..K$. Where the number of iterations K is determined based on a minimal $r_1 = r_{min}$ and maximal $r_K = r_{max}$ distances defined by the user. For each iteration k, we compute the confidence ellipse $\varepsilon_k = (\boldsymbol{\mu}_0, a_k, b_k, \varphi_k)$ (see sub-section 6.2.4) obtained from the covariance matrix Σ_k (eq. 33) computed only for the points that fall in the annulus of radius r_k. We then search for the iteration r^* that maximizes the ratio $C_v(r_k) = \sqrt{a_k/b_k}$ and use it to define the refined distance to the center

$$r_e(\boldsymbol{X}) = \sqrt{\left(x^{(1)} - \mu_0^{(1)}\right)^2 + \kappa\left(x^{(2)} - \mu_0^{(2)}\right)^2}, \tag{34}$$

where $\kappa = C_v(r^*)$, that transforms an ellipse into a circle with the same center.

Finally, we compute the density of points $N_e(r_k)$ falling in the annulus of radius r_k based on the refined distance measure r_e.

Termination We select the first iteration k^* such that $N_e(r_k^*) > N_e(r_{k^*-1})$: it is the first iteration where the derivative of the density with respect to the distance to the center, stops decreasing. This criteria is more stable on empirical data than searching for the minimum of the density (see figs. 3&4 panel B3 of [33]). The elliptic boundary of the well ε^* is centered at $\boldsymbol{\mu}_0$, has semi-axis lengths a^*, b^* given by $a^* = r_{k^*}$ and $b^* = \kappa r_{k^*}$ and orientation φ_{k^*}. We then use the ML estimator (eq. 27) to estimate the constant diffusion coefficient D inside ε^*. Finally, to compute A^* we use the diagonal form of the covariance matrix estimated (eq. 33) for all the points falling in ε^*:

$$\Sigma = \frac{D}{A}\begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}. \tag{35}$$

and estimate

$$A^* = \frac{D}{2}\left(\frac{(a^*)^2}{\sigma_{11}} + \frac{(b^*)^2}{\sigma_{22}}\right). \tag{36}$$

6.5 Drift-based algorithm

The drift based algorithm consists in using an error function in the space of the vector field to estimate the characteristics of a well.

Initiation Search for high-density regions: the image is partitioned by a grid $G_{\Delta x}$ with square bins of size Δx from which we compute the density map $\rho_{\Delta x}(x)$ (eq. (10)). We then select the bins from $\rho_{\Delta x}(x)$ with the highest $d\%$ density as possible regions containing a potential well.

Iterations For each region selected in the initiation, we apply the following iterative procedure for $k = 1..K$:

(a) We select only the sub-trajectories falling into a square $S_k(\boldsymbol{\mu}_k, \Delta x)$ centered at $\boldsymbol{\mu}_{k-1}$ and of size $(2k+1)\Delta x \times (2k+1)\Delta x$. We take $\boldsymbol{\mu}_0$ to be the center of the high-density bin.

(b) We estimate the elliptic well boundary $\varepsilon_k = [\boldsymbol{\mu}_k, a_k, b_k, \varphi_k]$ as the $X\%$ confidence ellipse (see sub-section 6.2.4) from the cloud of points in $S_k(c_k, \Delta)$. Where X is a parameter selected by the user (usually $90, 95$ or 99).

(c) We then compute a new grid $G_{\Delta x}(\boldsymbol{\mu}_k)$ centered at $\boldsymbol{\mu}_k$, that we use to estimate the local drift map (eq. 6) $a_k(\boldsymbol{X}) = [a_k^{(1)}(\boldsymbol{X}), a_k^{(2)}(\boldsymbol{X})]$ and estimate the attraction coefficient A_k of the well using the least-square regression formula

$$A_k = \frac{1}{2} \frac{\sum_{i=1}^{M} \frac{a_k^{(1)}(\boldsymbol{X}_i)x_i^{(1)}}{a^2} + \frac{a_k^{(2)}(\boldsymbol{X}_i)x_i^{(2)}}{b^2}}{\sum_{i=1}^{M} \frac{(x_i^{(1)})^2}{a^4} + \frac{(x_i^{(2)})^2}{b^4}}, \tag{37}$$

where $\boldsymbol{X}_i = [x_i^{(1)}, x_i^{(2)}]$ $(i = 1 \dots M)$ are the centers of the M bins from $G_{\Delta x}(\boldsymbol{\mu}_k)$ that are contained inside the ellipse ε_k.

(d) Finally, we estimate the quality of the well (parabolic index) based on the residual least square error:

$$S_k(a_k, A_k) = 1 - \frac{\left(\sum_{i=1}^{M} \frac{a_k^{(1)}(\boldsymbol{X}_i)x_i^{(1)}}{a^2} + \frac{a_k^{(2)}(\boldsymbol{X}_i)x_i^{(2)}}{b^2}\right)^2}{\left(\sum_{i=1}^{M} \frac{(x_i^{(1)})^2}{a^4} + \frac{(x_i^{(2)})^2}{b^4}\right)\left(\sum_{i=1}^{M} ||a_k(\boldsymbol{X}_i)||^2\right)}. \tag{38}$$

The index $S_k \in [0,1]$ is defined such that $S_k \to 0$ for a drift field generated by a parabolic potential well and $S_k \to 1$ for a random drift vector field as observed for diffusive motion.

The number of iterations is given by $N = \lfloor w_{max}/\Delta x \rfloor$ where w_{max} is the maximum size of the region to consider and is given by the user.

Termination We select the iteration k^* that minimizes the parabolic index S: $k^* = \arg\min_{k=1\ldots K} S_k(a_k, A_k)$. We estimate the diffusion coefficient inside the well using the local estimator (eq. 4) for all the displacements inside the ellipse ε_{k^*}.

6.6 Sliding window analysis to study the stability of the wells over time

To determine the stability of the potential wells, we use a non-overlapping sliding window of 20s [14], to recover the ellipse characteristics, as shown on different examples in Supplementary Fig. 3A,C. When a well disappears in a given time window, but reappears latter, we kept the well for the entire period.

6.7 Reconstructing a graph for a network explored by SPTs

We describe here an algorithm to reconstruct a graph of a network explored by SPTs. This algorithm is based on the heterogeneous distribution of points caused by trajectories spending more time inside nodes than in tubules.

6.7.1 Velocity based graph reconstruction algorithm (Vebgral) of a network explored by SPTs

The Vebgral algorithmic procedure to detect nodes (junctions) and inter-junction links (tubules) uses the large amount of recorded SPTs. The present algorithm first generates an ensemble of points from slow trajectory segments based on a maximum displacement length threshold v_L and then uses the `dbscan` algorithm [36] to cluster these points based on their local density. The algorithm requires specification of two ensembles of parameters:

1. An ensemble of distances $R = \{R_u, u = 1..U\}$ (in μm) defining the neighborhood radius around points.

2. An ensemble of counts $N = \{N_v, v = 1..V\}$ defining the numbers of points required in the neighborhood to form a cluster [36].

A pair of these two parameters (R_u, N_v) for any u, v define a local density $\frac{N_v}{R_u}$ (points/μm^2) inside each cluster. The values of R and N depend on the local number of recorded trajectories and can vary inside the image. For each dataset, these values can be determined such that the computed clusters overlap with the structure of the organelle formed by the trajectories.

We now present the steps of the algorithm:

1. We form the ensemble of points belonging to low-velocity trajectory fragments $S = \{X_i(t_j), |\frac{||\Delta X_i(t_j)||}{\Delta t} \geq v_L\}$.

2. We apply the `dbscan` procedure with parameters (R_U, N_1) to obtain a first ensemble of K clusters $c_1, \ldots, c_K$ from the points in S.

3. We then refine these initial clusters by searching for clusters possessing more than N_{max} points. For any cluster c_k possessing more than N_{max} points:

   (a) We iteratively re-apply the `dbscan` algorithm with the more stringent parameter pair (R_u, N_v) for $u = 2..U$ and $v = (V-1)..1$ and replace the initial cluster with the resulting sub-cluster(s). We continue iterating over the generated sub-clusters until they all possess less than N_{max} points.

4. We then approximate the boundary of each cluster either by its maximum volume ellipsoid or its convex hull polygon and assign each point discarded in step 1 to the cluster in which they fall if possible.

5. Finally, we merge any overlapping pair of clusters by computing the boundary of the combined ensemble of points (either elliptic or the convex hull) and we iterate this procedure until no more clusters overlap.

This first step allows to find the K nodes of the network. In the second step, we define tubules by constructing a connectivity matrix C of size $K \times K$ where $c_{i,j}$ $(1 \leq i, j \leq K)$ is the number of trajectory displacements starting in node i and arriving in node j. Specifically, we increment $c_{i,j}$ for each data point $X_k(t_l)$ $(1 \leq k \leq N_t,\ 0 \leq l < M_k - 1)$ in the following cases:

1. $X_k(t_l)$ is located in node i and $X_k(t_{l+1})$ in node j

2. $X_k(t_l)$ is located in node i, $X_k(t_{l+1})$ does not belong to any node and $X_k(t_{l+2})$ is located in node j (in this case $0 \leq l < M_k - 2$).

6.8 Lysosome analysis

6.8.1 Trajectories analysis for lysosome SPTs

To study the dynamics of lysosomes, we plotted the distribution of instantaneous velocities, computed from each trajectory displacement $\boldsymbol{X}(t+\Delta t) - \boldsymbol{X}(t)$ by

$$v = \frac{\boldsymbol{X}(t+\Delta t) - \boldsymbol{X}(t)}{\Delta t}, \tag{39}$$

where $\Delta t = 1.5$s. We approximate the distribution of instantaneous velocities using a two exponential model obtained by fitting $f(v) = Ae^{-\frac{v}{v_0}} + Be^{-\frac{v}{v_1}}$ to the distribution using MATLAB's `fit` toolbox. The density and diffusion maps were computed using the estimators 4 described above.

6.8.2 Local high-density region analysis: ellipse approximation of the boundary

High-density regions are extracted from trajectories as follows: we construct a density map (eq. 10) based on a square grid with bin size $\Delta x = 480$ nm. From this density map, we select only the 5% highest density bins and in case multiple such bins appear within a distance of two squares of each other, only the one with the highest value was kept. For each selected bin of center c, we computed a refined density map of size 5×5 squares, centered at c and with bin size $\Delta x' = 200$ nm. From this local map, we collected trajectory points falling into the bins that have a density $> 80\%$ of the maximal bin value [33] and use them to estimate the elliptic boundary of the region from the 95% confidence ellipse (see sub-section 6.2.4). Finally, when a pair of ellipses overlap, we replaced them by the ellipse computed over their combined ensemble of points and iterated this procedure until no more overlaps could be found.

6.8.3 Transient confinement detection

To detect transient confinement periods along individual trajectories, we used the following procedure: for each point Xt_j of a trajectory, we considered the ensemble of its successors $e_{t_j,n} = \{X(t_j), \ldots, X(t_{j+n})\}$, where initially $n = N_{nh}$ is set by the user. We then computed the center of mass $\mu_{t_j,n}$ and checked that all the points $X(t_k) \in e_{t_j,n}$ have a distance to the center of mass $||X(t_j) - \mu_{t_k,n}|| < R_{nh}$, for a chosen distance threshold R_{nh} ($||.||$ is the Euclidean norm). We then iterate the procedure, considering increasingly larger ensembles of successors $n = N_{nh}, N_{nh}+1, ...N_{nh}+K$ until

either reaching the end of the recorded trajectory or when the next point $\boldsymbol{X}(t_{j+n+1})$ do not fall into the neighborhood of the center of mass $\mu_{t_j,n}$. The confinement duration is then computed by considering the difference $t_{j+n} - t_j$ in time between the two endpoints of the ensemble. Finally, the spring constant λ and diffusion coefficient D of the confinement is obtained by applying the Ornstein-Ulenbeck maximum likelihood estimators [48, 33], where the OU-process is given by

$$dX = -\lambda(X - \mu)dt + \sqrt{2D}dW. \tag{40}$$

6.9 ImageJ plugin

The present method and algorithms are implemented into a imageJ plugin called "TrajectoryAnalysis". The plugin allows to reconstruct the various maps (trajectories, density, drift, diffusion), detect potential wells and reconstruct the graph associated to trajectories. It allows to extract various statistics such as the distribution of diffusion coefficients or the energy and the size of potential wells.

7 Author contributions

PP and DH conceived and designed the research plan, computational methods and algorithms and wrote the manuscript. PP wrote the algorithms. M.H., E.A. J.H. and MH designed the experiments and collected super-resolution microscopy data.

8 Acknowledgement

D. H. 's research has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 882673), Plan Cancer-INSERM Projet 19CS145-00 and ANR NEUC-0001.

9 Figures

9.1 Fig.1 High-throughput SPT analysis pipeline

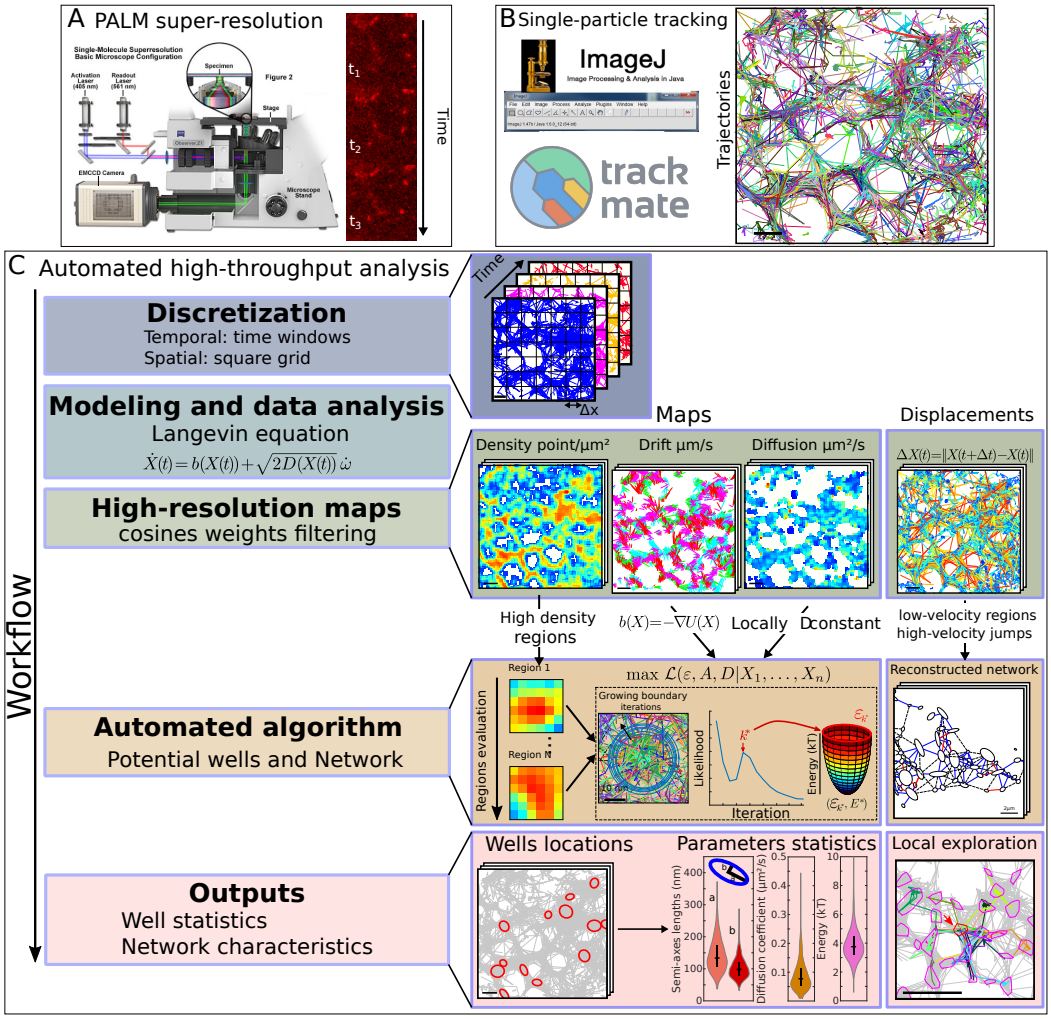


Figure 1:

legend fig. 1

High-throughput automated trajectory analysis workflow. A. Acquisition device and raw data of a single-particle experiment. **B.** Raw data from A are

transformed into trajectories using classical softwares such as Trackmate available as an ImageJ plugin. **C.** Schematic description of the high-throughput analysis implemented here: trajectories are first discretized spatially using a square grid and a time-windows generate a segmentation in time. Thet are interpreted using the Langevin equation, allowing us to generate high-resolution maps of the local trajectory motion. High-density / low-velocity regions of the maps are generated by automated machine-learning type algorithms to detect potential wells and to reconstruct network. Outputs consist in location maps of wells, their statistics and reconstructed network allowing to analyse how trajectories locally exploration the nanophysiology scale.

9.2 Fig.2 Large statistics of CaV nanodomains located on dendrite automatically detected

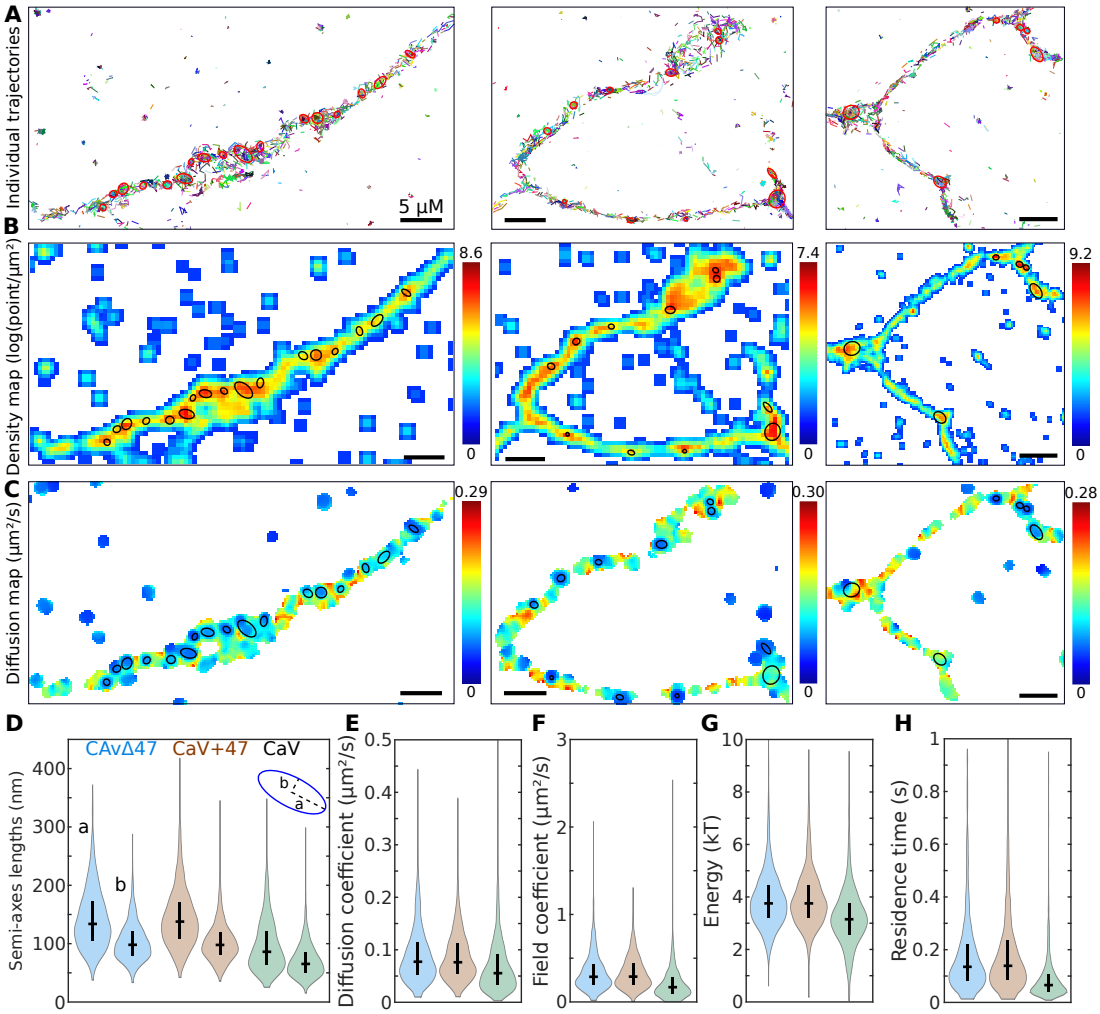


Figure 2:

legend fig. 2

Super-resolution SPTs automated analysis reveals CaV nanodomain organization. A. Examples of trajectories displaying high-density regions (red ellipses). **B.** Associated density map, presenting the local point density (in log(points/μm^2)) and computed over a grid with bin sizes of 50 nm and locally averaged over a 3x3

Gaussian kernel. **C.** Associated diffusion map, displaying the local diffusion coefficient ($\mu m^2/s$) computed over a grid with bin sizes of 80 nm with at least 15 displacements and estimated with the cosine-round method (Material and Method) with a radius of 100 nm.**D-H.** Population characteristics for three experimental conditions: tracking CaV2.1-47, CaV2.1+47 or CaV2.1. Nanodomain represented as potential wells are recovered using the hybrid algorithm. D: Distribution of semi-axes lengths of elliptic well boundaries, E: Diffusion coefficients inside wells, F: Field coefficients A, G: Energies of the well (kT), H: Residence time distribution of a trajectory inside a well (H). Mean and variance values are shown in Table S3.

9.3 Fig.3 HDR present in organelle such as the ER network

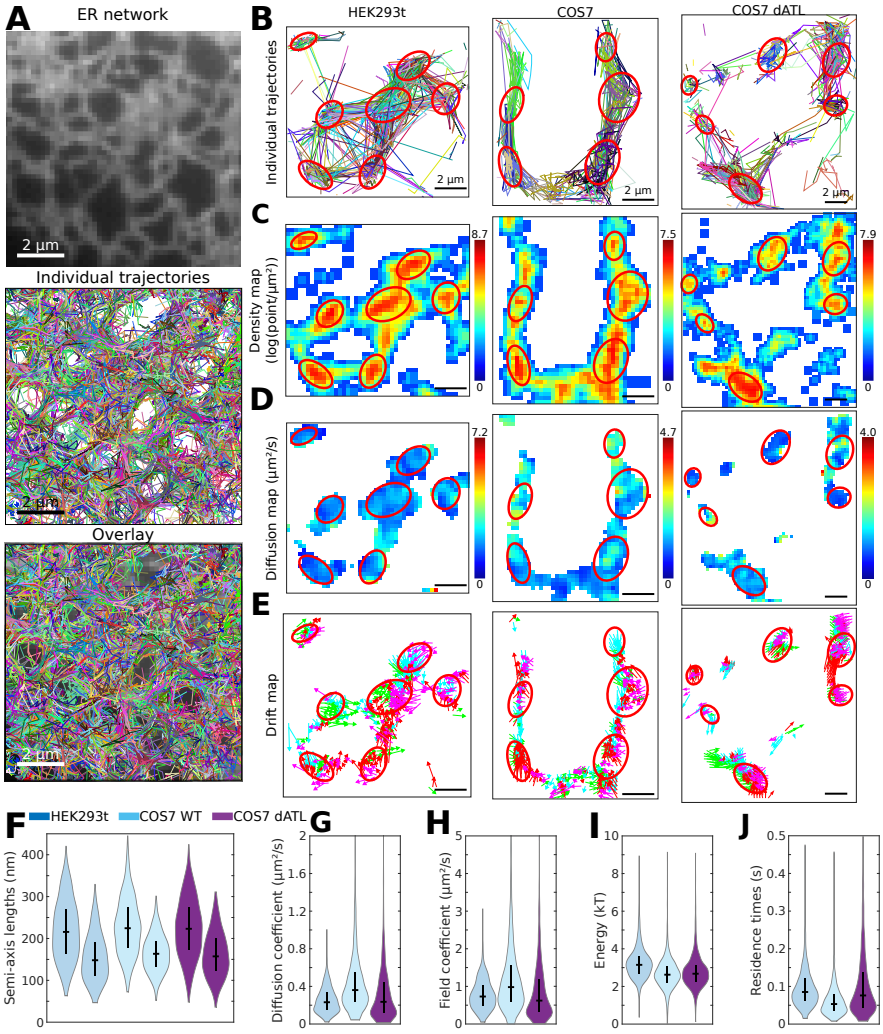


Figure 3:

legend fig. 3

ER nodes revealed by super-resolution SPTs are characterized as potential wells. A. Super-resolution ER network. Single trajectories recorded in the lumen and coded by a given color (second) and overlay (Third) **B.** Single trajectories coded by a given color and recorded inside the ER for three different cells

(HEK293t, COS7, COS7-dATL). HDRs are associated with ER node (red ellipse). **B-C-D.** Density, Diffusion and Drift maps for the regions shown in **A.**. Arrows in the drift maps are colored according to the direction: West (purple), East (green), South (blue) and North (red). **E-J.** Statistics of the wells reconstructed from the hybrid algorithm (Material and Method): semi-axis (longest a and shortest b), diffusion coefficients D, Field coefficient A, Energy of the well E and the estimated residence time.

9.4 Fig.4 Algorithm to automatically reconstruct a network

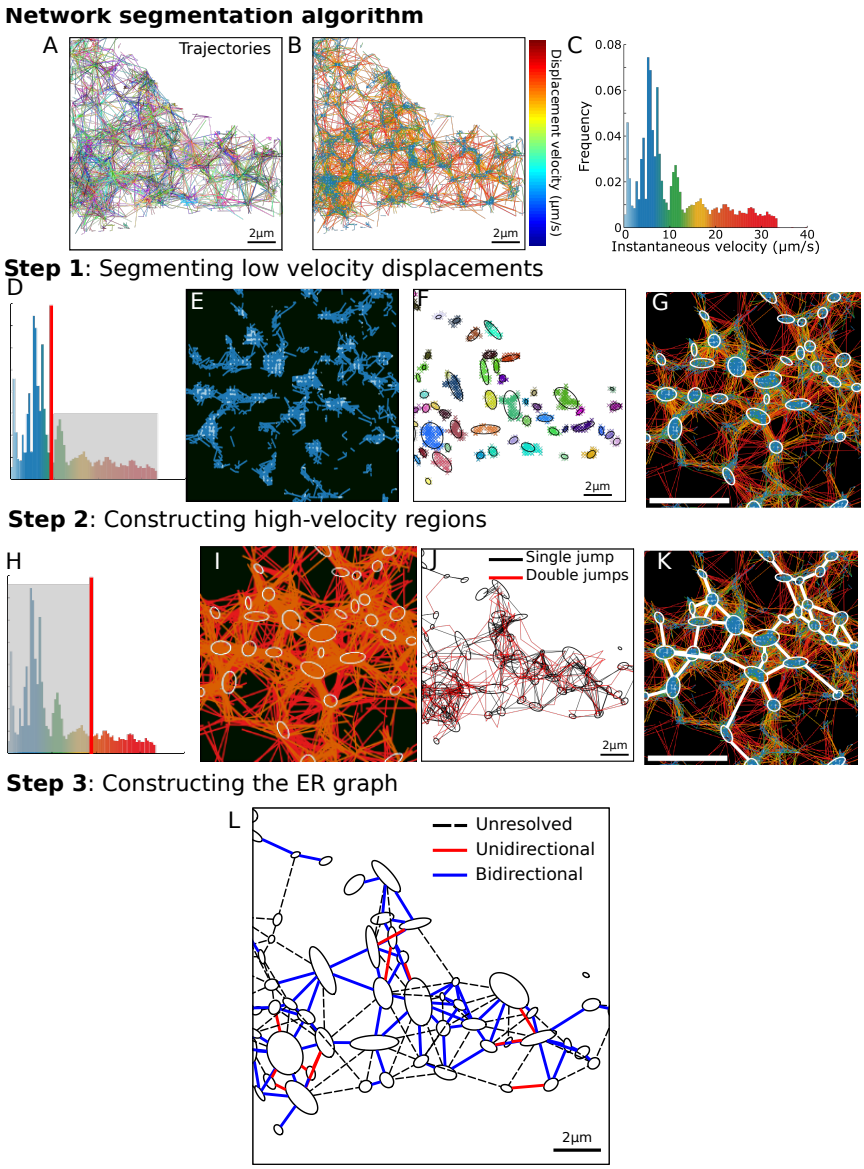


Figure 4:

legend fig. 4

Automatically network segmentation revealed by super-resolution SPTs. A-C. SPTs (A), density map (B) and color coded distribution according to amplitude of instantaneous velocity (C). Step 1: detecting regions of low velocities. **D.** dbscan procedure to cluster (E) and merge them into ellipses (F) and embedding into the network (G). Step 2: characterization of high velocity regions from the histogram (H), segmented between ellipses (step 1) showing one or two fast jumps (J) and finally reconstructed in the ensemble of trajectories (K). Step 3: abstract graph reconstruction of an ER showing nodes and tubules, that could be bi- or unidirectional. Step 4: Synchronizing trajectories reveal a switching dynamics occurring inside tubules.

9.5 Fig.5 ER network in dATL cells

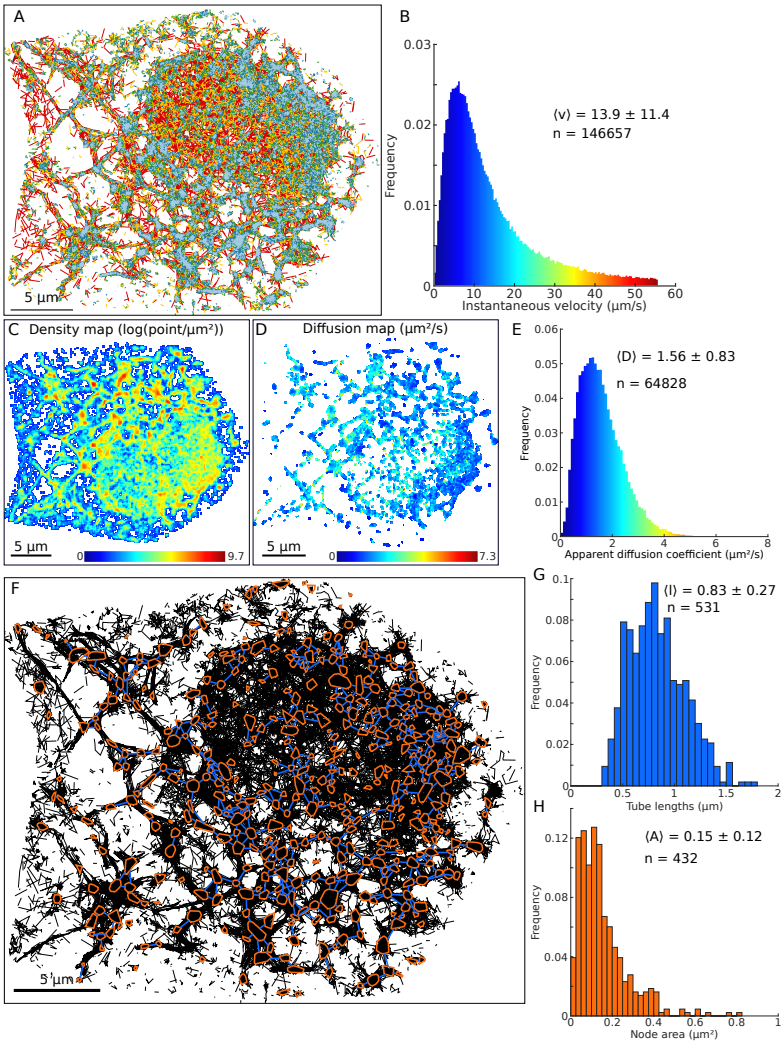


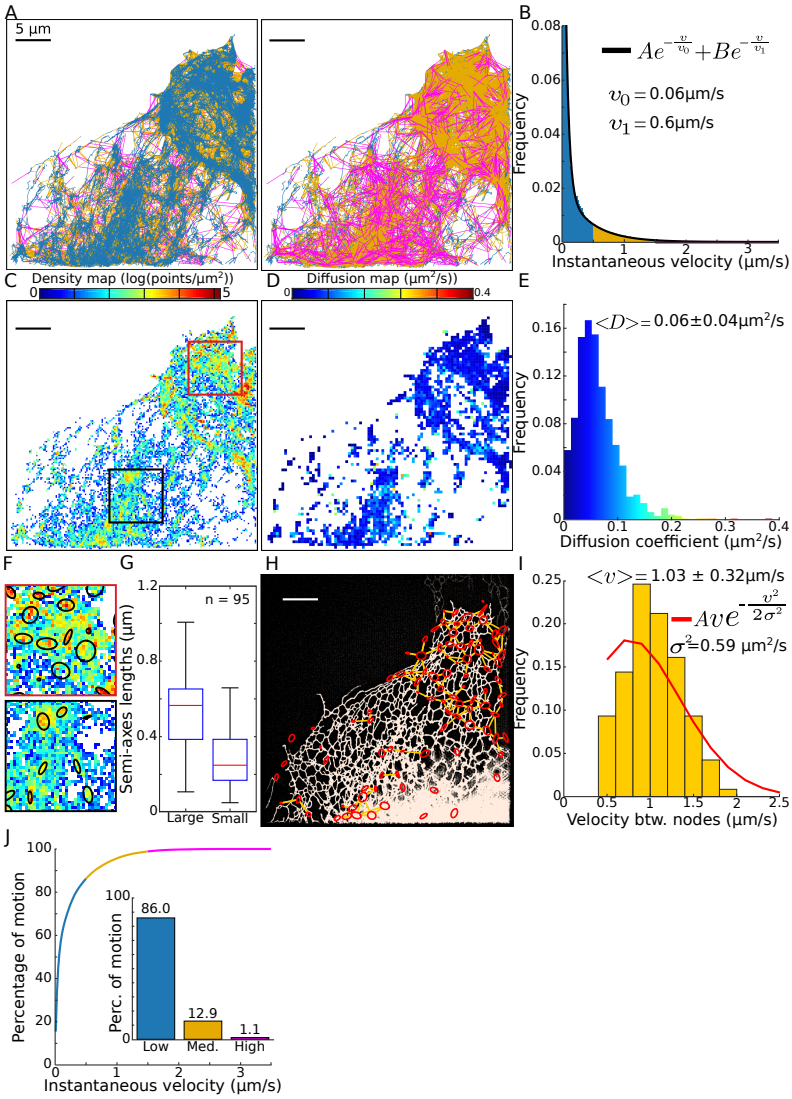
Figure 5:

legend fig. 5

ER network reconstructed in COS-7 dATL cells. **A** Individual trajectories color-coded according to their instantaneous velocity shown in **B**. **C** Density and **D** drift maps. **E** Distribution of diffusion coefficients obtained from the individual bins of the diffusion map presented in D. **F** Reconstructed network, showing the

nodes in red and links in blue overlaid on top of the individual trajectories (black). **G** Distribution of distances (i.e. tubule lengths), between connected nodes. **H** Distribution of the areas covered by nodes.

9.6 Fig.6 Reconstructing ER network from lysosome trajectories

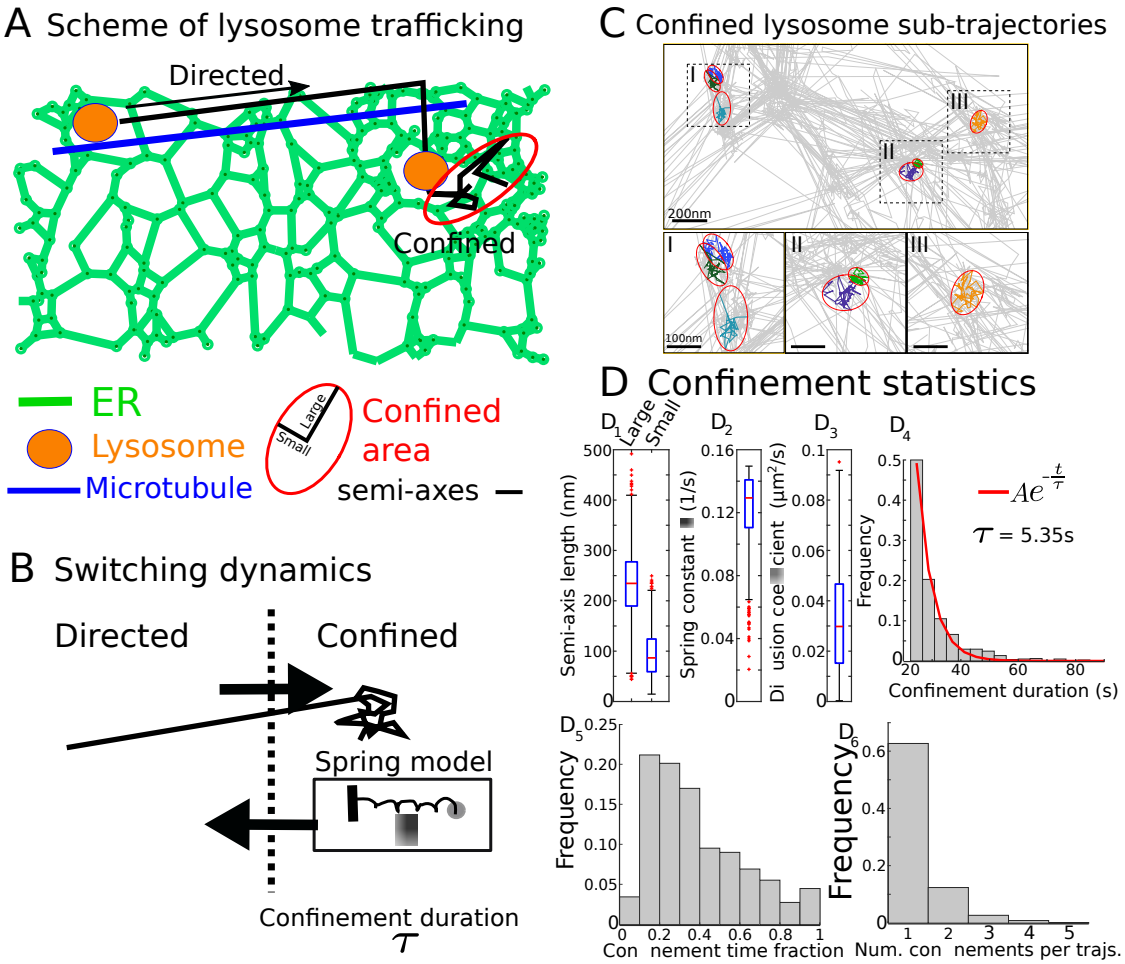


legend fig. 6

Lysosome trajectories analysis. **A**. Lysosome trajectories color-coded according to individual displacement amplitudes (panel B). **B**. Instantaneous velocities, color-coded with respect to the value and fitted by a double exponential distribution. **C**.

Density map of points. **D**. Diffusion map presenting the local diffusion coefficients. **E**. Histogram of the diffusion coefficients obtained in the square bins form panel D. **F**. Magnification of the density map of two regions of interest, showing high-density regions, approximated by ellipses. **G**. Length of semi-axes of high-density regions approximated as ellipses. **H**. Reconstruction of a lysosome graph, where nodes correspond to high-density regions. A link (in yellow) is added when at least one trajectory starts in one node and enter to the other one (in one or two frames).**I**. Average instantaneous velocities between pairs of connected nodes found in panel F. **J**. Percentage of displacements with a specific instantaneous velocity. Inset, percentage of displacements for the velocity regimes defined in panels A-B.

9.7 Fig.7 ER network-Lysosome interaction revealed by confined trajectories

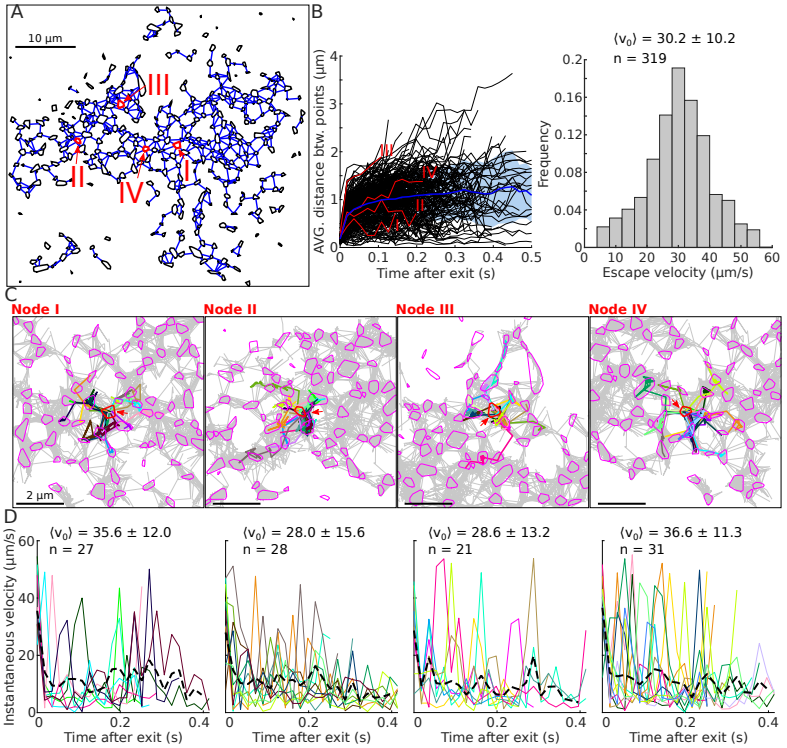


legend fig. 7

Lysosome confinement time **A**. Schematic representation of a lysosome switching between a directed motion along microtubule and confined motion at ER nodes. **B**. Switching dynamics representation: the confined state is characterized by a spring constant λ a diffusion coefficient D and a confined time constant τ. **C**. Regions of confinement for individual lysosome trajectories: three examples (insets). $\mathbf{D}_1$. Semi-axes (small and large) of the ellipse fitted to the confinement regions. $\mathbf{D}_2$.

Spring constants of an Ornstein-Ulenbeck process–confined diffusion–. $\mathbf{D}_3$. Diffusion coefficients estimated inside a confinement region. $\mathbf{D}_4$. Residence times inside a confinement region. $\mathbf{D}_5$. Fraction of time each trajectory spends confined (relative to the trajectory length). $\mathbf{D}_6$. Number of confinement events along individual trajectories.

9.8 Fig.8 Local space exploration after trajectory re-synchronization



legend fig. 8

Local network exploration revealed by spatial re-synchronization of super-resolution SPTs. A. ER reconstructed network connectivity using the GRA with nodes (black polygons) connected by segments (blue). Four nodes (I, II, III, IV) are selected (red arrows) inside a ER segmented network. **B.** Average distance between exit points vs time after escape. We highlight 4 examples (red), average (blue) and individual examples (black), fast ($< 50ms$) and a slow phase. **C.** Local exploration of neighboring nodes from trajectories (various colors) exiting from the chosen node.**D.** Distribution of velocities for trajectories exiting a node for the four examples: the mean velocity decreases from a high amplitude corresponding to the instant of exit.

References

[1] N. B. Cole, C. L. Smith, N. Sciaky, M. Terasaki, M. Edidin, and J. Lippincott-Schwartz, "Diffusional mobility of golgi proteins in membranes of living cells," *Science*, vol. 273, no. 5276, pp. 797–801, 1996.

[2] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[3] S. Cohen, A. M. Valm, and J. Lippincott-Schwartz, "Interacting organelles," *Current opinion in cell biology*, vol. 53, pp. 84–91, 2018.

[4] H. Wu, P. Carvalho, and G. K. Voeltz, "Here, there, and everywhere: The importance of er membrane contact sites," *Science*, vol. 361, no. 6401, 2018.

[5] M. Petkovic, A. Jemaiel, F. Daste, C. G. Specht, I. Izeddin, D. Vorkel, J.-M. Verbavatz, X. Darzacq, A. Triller, K. H. Pfenninger, *et al.*, "The snare sec22b has a non-fusogenic function in plasma membrane expansion," *Nature cell biology*, vol. 16, no. 5, pp. 434–444, 2014.

[6] Y. Wu, C. Whiteus, C. S. Xu, K. J. Hayworth, R. J. Weinberg, H. F. Hess, and P. De Camilli, "Contacts between the endoplasmic reticulum and other membranes in neurons," *Proceedings of the National Academy of Sciences*, vol. 114, no. 24, pp. E4859–E4867, 2017.

[7] M. Lu, F. W. van Tartwijk, J. Q. Lin, W. Nijenhuis, P. Parutto, M. Fantham, C. N. Christensen, E. Avezov, C. E. Holt, A. Tunnacliffe, *et al.*, "The structure and global distribution of the endoplasmic reticulum network is actively regulated by lysosomes," *bioRxiv*, 2020.

[8] D. Axelrod, D. Koppel, J. Schlessinger, E. Elson, and W. W. Webb, "Mobility measurement by analysis of fluorescence photobleaching recovery kinetics," *Biophysical journal*, vol. 16, no. 9, pp. 1055–1069, 1976.

[9] M. J. Saxton and K. Jacobson, "Single-particle tracking: applications to membrane dynamics," *Annual review of biophysics and biomolecular structure*, vol. 26, no. 1, pp. 373–399, 1997.

[10] S. Wang, J. R. Moffitt, G. T. Dempsey, X. S. Xie, and X. Zhuang, "Characterization and development of photoactivatable fluorescent proteins for single-molecule–based superresolution imaging," *Proceedings of the National Academy of Sciences*, vol. 111, no. 23, pp. 8452–8457, 2014.

[11] J. Lippincott-Schwartz, N. Altan-Bonnet, and G. H. Patterson, "Photobleaching and photoactivation: following protein dynamics in living cells.," *Nature cell biology*, pp. S7–14, 2003.

[12] S. Manley, J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz, "High-density mapping of single-molecule trajectories with photoactivated localization microscopy," *Nature methods*, vol. 5, no. 2, p. 155, 2008.

[13] N. Hoze, D. Nair, E. Hosy, C. Sieben, S. Manley, A. Herrmann, J.-B. Sibarita, D. Choquet, and D. Holcman, "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, no. 42, pp. 17052–17057, 2012.

[14] J. Heck, P. Parutto, A. Ciuraszkiewicz, A. Bikbaev, R. Freund, J. Mitlöhner, M. Andres-Alonso, A. Fejtova, D. Holcman, and M. Heine, "Transient confinement of cav2. 1 ca2+-channel splice variants shapes synaptic short-term plasticity," *Neuron*, vol. 103, no. 1, pp. 66–79, 2019.

[15] M. M. Wu, E. D. Covington, and R. S. Lewis, "Single-molecule analysis of diffusion and trapping of stim1 and orai1 at endoplasmic reticulum–plasma membrane junctions," *Molecular biology of the cell*, vol. 25, no. 22, pp. 3672–3685, 2014.

[16] M. Heine and D. Holcman, "Asymmetry between pre-and postsynaptic transient nanodomains shapes neuronal communication," *Trends in Neurosciences*, 2020.

[17] D. Holcman, P. Parutto, J. E. Chambers, M. Fantham, L. J. Young, S. J. Marciniak, C. F. Kaminski, D. Ron, and E. Avezov, "Single particle trajectories reveal active endoplasmic reticulum luminal flow," *Nature cell biology*, vol. 20, no. 10, p. 1118, 2018.

[18] J. Crank, *The mathematics of diffusion*. Oxford university press, 1979.

[19] A. Kusumi, Y. Sako, and M. Yamamoto, "Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). effects of calcium-induced differentiation in cultured epithelial cells," *Biophysical journal*, vol. 65, no. 5, pp. 2021–2040, 1993.

[20] M. J. Saxton, "Single-particle tracking: effects of corrals," *Biophysical Journal*, vol. 69, no. 2, pp. 389–398, 1995.

[21] D. Holcman and Z. Schuss, "Escape through a small opening: receptor trafficking in a synaptic membrane," *Journal of Statistical Physics*, vol. 117, no. 5-6, pp. 975–1014, 2004.

[22] N. Hozé and D. Holcman, "Statistical methods for large ensembles of superresolution stochastic single particle trajectories in cell biology," *Annual Review of Statistics and Its Application*, vol. 4, pp. 189–223, 2017.

[23] S. M. Gasser, "Nuclear architecture: Past and future tense," *Trends in Cell Biology*, 2016.

[24] A. Amitai, A. Seeber, S. M. Gasser, and D. Holcman, "Visualization of chromatin decompaction and break site extrusion as predicted by statistical polymer modeling of single-locus trajectories," *Cell reports*, vol. 18, no. 5, pp. 1200–1214, 2017.

[25] M. H. Hauer, A. Seeber, V. Singh, R. Thierry, R. Sack, A. Amitai, M. Kryzhanovska, J. Eglinger, D. Holcman, T. Owen-Hughes, *et al.*, "Histone degradation in response to dna damage enhances chromatin dynamics and recombination rates," *Nature structural & molecular biology*, vol. 24, no. 2, p. 99, 2017.

[26] J. P. Dupuis and L. Groc, "Surface trafficking of neurotransmitter receptors: From cultured neurons to intact brain preparations," *Neuropharmacology*, vol. 169, p. 107642, 2020.

[27] D. Holcman, N. Hoze, and Z. Schuss, "Analysis and interpretation of super-resolution single-particle trajectories," *Biophysical journal*, vol. 109, no. 9, pp. 1761–1771, 2015.

[28] Z. Schuss, *Theory and applications of stochastic processes: an analytical approach*, vol. 170. Springer Science & Business Media, 2009.

[29] H. A. Kramers, "Brownian motion in a field of force and the diffusion model of chemical reactions," *Physica*, vol. 7, no. 4, pp. 284–304, 1940.

[30] L. Niu, T. Ma, F. Yang, B. Yan, X. Tang, H. Yin, Q. Wu, Y. Huang, Z.-P. Yao, J. Wang, *et al.*, "Atlastin-mediated membrane tethering is critical for cargo mobility and exit from the endoplasmic reticulum," *Proceedings of the National Academy of Sciences*, vol. 116, no. 28, pp. 14029–14038, 2019.

[31] J.-B. Sibarita, "High-density single-particle tracking: quantifying molecule organization and dynamics at the nanoscale," *Histochemistry and cell biology*, vol. 141, no. 6, pp. 587–595, 2014.

[32] F. Levet, E. Hosy, A. Kechkar, C. Butler, A. Beghin, D. Choquet, and J.-B. Sibarita, "Sr-tesseler: a method to segment and quantify localization-based super-resolution microscopy data," *Nature methods*, vol. 12, no. 11, pp. 1065–1071, 2015.

[33] P. Parutto, J. Heck, M. Heine, and D. Holcman, "Biophysics of high density nanometer regions extracted from super-resolution single particle trajectories: application to voltage-gated calcium channels and phospholipids," *Scientific Reports*, vol. 9, no. 1, pp. 1–14, 2019.

[34] V. Briane, M. Vimond, and C. Kervrann, "An overview of diffusion models for intracellular dynamics analysis," *Briefings in bioinformatics*, vol. 21, no. 4, pp. 1136–1150, 2020.

[35] S. Chandrasekhar, "Stochastic problems in physics and astronomy," *Reviews of modern physics*, vol. 15, no. 1, p. 1, 1943.

[36] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise.," in *Kdd*, vol. 96, pp. 226–231, 1996.

[37] N. Hoze and D. Holcman, "Residence times of receptors in dendritic spines analyzed by stochastic simulations in empirical domains," *Biophysical journal*, vol. 107, no. 12, pp. 3008–3017, 2014.

[38] J.-B. Masson, P. Dionne, C. Salvatico, M. Renner, C. Specht, A. Triller, and M. Dahan, "Mapping the energy and diffusion landscapes of membrane proteins at the cell surface using high-density single-molecule imaging and bayesian inference: Application to the multiscale dynamics of glycine receptors in the neuronal membrane," *Biophysical Journal*, vol. 106, no. 1, pp. 74 – 83, 2014.

[39] N. Hoze and D. Holcman, "Coagulation-fragmentation with a finite number of particles: Models, stochastic analysis, and applications to telomere clustering and viral capsid assembly," in *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology*, pp. 205–239, Springer, 2017.

[40] C. Floderer, J.-B. Masson, E. Boilley, S. Georgeault, P. Merida, M. El Beheiry, M. Dahan, P. Roingeard, J.-B. Sibarita, C. Favard, *et al.*, "Single molecule localisation microscopy reveals how hiv-1 gag proteins sense membrane virus assembly sites in living host cd4 t cells," *Scientific reports*, vol. 8, no. 1, pp. 1–15, 2018.

[41] D. Holcman, "Unraveling novel features hidden in superresolution microscopy data," *Communicative & integrative biology*, vol. 6, no. 3, p. e23893, 2013.

[42] Z. Feng, X. Chen, M. Zeng, and M. Zhang, "Phase separation as a mechanism for assembling dynamic postsynaptic density signalling complexes," *Current opinion in neurobiology*, vol. 57, pp. 1–8, 2019.

[43] S. Manley, J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz, "High-density mapping of single-molecule trajectories with photoactivated localization microscopy," *Nature methods*, vol. 5, no. 2, p. 155, 2008.

[44] Z. Schuss, *Theory and Applications of Stochastic Processes: An Analytical Approach.* Springer New York, 2010.

[45] Z. Schuss, *Diffusion and stochastic processes: an analytical approach.* Springer New York, 2010.

[46] R. Friedrich and J. Peinke, "Description of a turbulent cascade by a fokker-planck equation," *Physical Review Letters*, vol. 78, no. 5, p. 863, 1997.

[47] S. Siegert, R. Friedrich, and J. Peinke, "Analysis of data sets of stochastic systems," *Physics Letters A*, vol. 243, no. 5-6, pp. 275–280, 1998.

[48] C. Y. Tang and S. X. Chen, "Parameter estimation and bias correction for diffusion processes," *Journal of Econometrics*, vol. 149, no. 1, pp. 65–81, 2009.