



**HAL**  
open science

# Statistical analysis of single particle trajectories reveals sub-cellular nanodomain organisation and function

Pierre Parutto

► **To cite this version:**

Pierre Parutto. Statistical analysis of single particle trajectories reveals sub-cellular nanodomain organisation and function. Neuroscience. Université Paris sciences et lettres, 2019. English. NNT : 2019PSLEE055 . tel-03040055

**HAL Id: tel-03040055**

**<https://theses.hal.science/tel-03040055>**

Submitted on 4 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**THÈSE DE DOCTORAT**  
**DE L'UNIVERSITÉ PSL**

Préparée à l'École Normale Supérieure

**Statistical analysis of single particle trajectories reveals  
sub-cellular nanodomain organisation and function**

Soutenue par

**Pierre Parutto**

Le 12/11/2019

École doctorale n°158

**ED3C**

Spécialité

**Neurosciences**

Composition du jury :

Denis Grebenkov École polytechnique	<i>Rapporteur</i>
Charles Kervrann INRIA	<i>Rapporteur</i>
Martin Heine Johannes Gutenberg University Mainz	<i>Examineur</i>
Dmitri Rusakov University College London	<i>Président du jury</i>
Pierre Sens Institut Curie	<i>Examineur</i>
Christian Vestergaard Institut Pasteur	<i>Examineur</i>
David Holcman École Normale Supérieure	<i>Directeur de thèse</i>



# Contents

Summary . . . . .	5
Résumé . . . . .	6
Remerciements . . . . .	7
Publications . . . . .	8
<b>1 Introduction</b>	<b>9</b>
1.1 Nanodomains in neurosciences and cell biology . . . . .	11
1.1.1 Receptor dynamics at synapses . . . . .	11
1.1.2 Recovery after photobleaching experiments for characterizing molecular dynamics . . . . .	12
1.1.3 Trajectories of individual post-synaptic receptors . . . . .	14
1.2 Super-resolution microscopy of single particle trajectories . . . . .	14
1.2.1 Abbe’s law for diffraction limit . . . . .	14
1.2.2 Super-resolution microscopy techniques . . . . .	15
1.2.3 Constructing single particle trajectories from localization microscopy experiments . . . . .	17
1.2.4 Biophysical features contained in SPTs . . . . .	19
1.3 Stochastic models of individual particles . . . . .	20
1.3.1 Stochastic theory of diffusion . . . . .	20
1.3.2 Diffusion and active forces . . . . .	20
1.3.3 Nanodomains modeled as parabolic potential wells . . . . .	21
1.3.4 Discrete equation of motion . . . . .	22
1.4 Parameter estimation from SPTs . . . . .	22
1.4.1 Mean squared displacement estimators . . . . .	23
1.4.2 Local estimators constructed from statistical moments . . . . .	24
1.4.3 Maximum likelihood estimators . . . . .	25

2	1.5	Main results of the thesis . . . . .	27
	1.5.1	Statistical approach to recover parabolic potential wells from SPTs . . . . .	27
	1.5.2	Main applications of the well detection method . . . . .	33
	1.5.3	Application of SPTs analysis to reveal the flow in the endoplasmic reticulum lumen . . . . .	37
	1.5.4	Applications of SPTs analysis to lysosome, drosophila neuro-muscular junctions and NuRD complex . . . . .	42
	1.6	General conclusion on the statistical analysis of SPTs . . . . .	44
<b>2</b>		<b>Reconstructing potential wells of high density regions from super-resolution single particle trajectories</b>	<b>45</b>
	2.1	Introduction . . . . .	46
	2.2	Methods . . . . .	47
	2.2.1	Coarse-grained description of stochastic trajectories . . . . .	47
	2.2.2	Potential wells characteristics . . . . .	47
	2.2.3	Simulations of stochastic trajectories . . . . .	48
	2.2.4	Estimators for the elliptic boundary geometry . . . . .	48
	2.2.5	Improved drift estimation . . . . .	49
	2.2.6	Processing of $Ca_v2.2$ and GPI SPTs . . . . .	51
	2.3	Results . . . . .	51
	2.3.1	Recovering a bounded potential well from the point density of trajectories	51
	2.3.2	Estimating the characteristics of the well using the velocity distribution	54
	2.3.3	Estimating the center and the field coefficients of the potential well .	56
	2.3.4	Interpretation high-density regions for $Ca_v2.2$ and GPI-GFP as potential wells . . . . .	56
	2.4	Summary and Discussion . . . . .	59
	2.4.1	Two statistical methods to interpret high-density regions . . . . .	59
	2.4.2	High-density regions contained calcium-voltage channels and GPI SPTs data . . . . .	62
	2.5	Supplementary Information . . . . .	63
	2.5.1	MLE estimator for a potential well . . . . .	63
	2.5.2	Least Square Quadratic Estimator (LSQE) . . . . .	65
	2.5.3	Influence of the time and spatial discretizations on the Least Square Estimation . . . . .	67
	2.5.4	Influence of the time steps in stochastic simulations . . . . .	75
	2.5.5	Conditional drift estimation . . . . .	77

<b>3</b>	<b>Transient Confinement of Ca<sub>v</sub>2.1 Ca<sup>2+</sup>-Channel Splice Variants Shapes Synaptic Short-Term Plasticity</b>	<b>79</b>
3.1	Introduction . . . . .	80
3.2	Results . . . . .	81
3.2.1	Expression and Synaptic Localization of Ca <sub>v</sub> 2.1 C-Terminal Splice Variants . . . . .	81
3.2.2	Functional Differences of Synapses Dominated by Ca <sub>v</sub> 2.1 <sub>+47</sub> or Ca <sub>v</sub> 2.1 <sub>Δ47</sub> Channels . . . . .	83
3.2.3	Dynamics of Ca <sub>v</sub> 2.1 Channels Differ between C-Terminal Splice Variants	84
3.2.4	Light-Triggered Re-organization of Calcium Channels . . . . .	88
3.2.5	C-Terminal Splicing Contributes to Short-Term Plasticity . . . . .	92
3.3	Discussion . . . . .	95
3.4	Supplementary Information . . . . .	98
3.4.1	Single-particle tracking PALM . . . . .	98
3.4.2	Data analysis and modeling . . . . .	99
<b>4</b>	<b>Single particle trajectories reveal active endoplasmic reticulum luminal flow</b>	<b>113</b>
4.1	Introducion . . . . .	114
4.2	Results . . . . .	114
4.3	Discussion . . . . .	122
4.4	Supplementary information . . . . .	125
4.4.1	Mean Squared Displacement and first moment analysis . . . . .	125
4.4.2	Models for the instantaneous velocity distribution . . . . .	128
4.4.3	Reconstruction of the ER Network from SPTs . . . . .	131
4.4.4	Recovery of the local dynamics in the ER lumen from SPTs . . . . .	132
4.4.5	Static ER network analysis . . . . .	133
4.4.6	Transient ER network analysis . . . . .	135
4.4.7	Instantaneous velocities along individual trajectories . . . . .	136
4.4.8	Instantaneous velocity peaks duration and inter-peaks period . . . . .	137
4.4.9	Dynamics of tubular junctions . . . . .	138
4.5	Characterization of ER tubule contractions . . . . .	142
4.5.1	Extraction of tubule contraction statistics . . . . .	142
4.5.2	Elementary model of tubule contraction . . . . .	142
4.5.3	Simultaneous contractions statistics . . . . .	143
4.6	Tables . . . . .	145

<b>5 Discussion: three case studies in cellular trafficking</b>	<b>147</b>
5.0.1 ER-lysosomes interactions participate in maintaining the peripheral ER shape . . . . .	147
5.0.2 Transient confinement detection . . . . .	152
5.0.3 Analysis of $Ca_v2.1$ nanodomains at Drosophila neuro-muscular junctions	152
5.0.4 Local high-density region analysis . . . . .	157
5.0.5 Analysis of the dynamics of the NuRD nuclear remodeler complex . .	159
5.0.6 Model of switching modes of motion . . . . .	159
5.1 Conclusion and perspectives . . . . .	163
<b>Bibliography</b>	<b>165</b>

## Summary

Single-Particle Trajectories (SPTs) recorded through super-resolution microscopy reveal protein motion with a ten of nanometers precision in living cells. In neurons, SPTs have contributed to a better understanding of the link between molecular dynamics and synaptic transmission. However extracting precise quantitative information from SPTs requires a stochastic model, which is often difficult to construct due to the heterogeneity of cellular nanodomains.

In the first part of the thesis, I studied a model of potential well for characterizing certain high-density nanodomains found in SPTs. This model allows computing quantities that are not directly measurable from trajectories, such as the height of the energy barrier or the residence time of a molecule inside the well. The main difficulty in recovering wells from SPTs lies in the interplay between its boundary and its dynamics parameters (attraction and diffusion coefficients). To resolve this, I developed a new hybrid method combining the estimation of the local point density (recovering the boundary) and local drift vector-field (recovering the dynamics) and compared its performance to classical maximum-likelihood using Monte Carlo simulations.

In the second part, I used this new method to characterize potential wells observed in SPTs of two splice variants of the  $\text{Ca}_v2.1$  voltage-gated calcium channels (with/without C-terminal tail), imaged on the membrane of hippocampal neurons. I found that these nanodomains have sizes  $\approx 80\text{nm}$ , are localized at presynaptic terminal's active zones, stabilizing channels there for  $\approx 100\text{ms}$ , much longer than by pure diffusion. In addition, the wells themselves are highly dynamic with a mean lifetime of  $\approx 30\text{s}$ , compatible with the redistribution of  $\text{Ca}_v$  following synaptic vesicle fusion events at active zones. These findings, in conjunction with electrophysiological recordings and calcium/glutamate imaging under various conditions, suggested that  $\text{Ca}_v$  redistribution contributes to shaping the release probability as well as the short-term plasticity dynamics of individual synapses.

In the third part, I analyzed SPTs of calreticulin in the Endoplasmic Reticulum (ER) lumen. There, I found that the high-density nanodomains from SPTs correlate with the tubular junctions of the peripheral ER network while the high-velocity jumps are found in tubules. To interpret the luminal dynamics, I first developed a method to reconstruct the ER network from SPTs and then proposed a stochastic model of motion where particles can switch between confined diffusion at junctions and transient high-velocity jumps in tubules. This analysis revealed that the luminal material can possibly visit the entire network; the network is at equilibrium and that tubules exhibit periods of alternating flow direction. Overall, the possible presence of an active flow in the ER lumen provides an efficient mechanism for luminal content homogenization over long distances.

In the last part, I discussed possible extensions of the presented methods for analyzing 1. lysosome trafficking and interactions with the ER network, 2.  $\text{Ca}_v$  dynamics at neuromuscular junctions of drosophila and 3. the different modes of motion exhibited by the NuRD chromatin remodeling complex in the nucleus.

To conclude, this thesis introduces stochastic models to reconstruct the molecular dynamics deriving from fundamental cellular nanophysiology processes.



## Résumé

Les trajectoires de molécules individuelles obtenues par microscopie super-résolution révèlent le mouvement des protéines avec une précision nanométrique dans des cellules vivantes, améliorant, par exemple, notre compréhension du lien entre dynamiques moléculaires et transmission synaptique entre neurones. Cependant, extraire des informations fines à partir de trajectoires requiert un modèle stochastique qui est souvent difficile à construire à cause de l'hétérogénéité des nanodomains cellulaires.

Dans la première partie de cette thèse, j'ai travaillé sur un modèle de puits de potentiels pour caractériser certains nanodomains de haute densité observés dans des trajectoires. Ce modèle permet de calculer des quantités qui ne sont pas mesurables directement à partir des trajectoires comme la hauteur de la barrière de potentiel ou le temps de résidence des molécules dans un puit. La difficulté principale pour estimer des puits à partir de trajectoires vient des interactions entre leurs bords et leurs paramètres dynamiques (coefficients d'attraction et de diffusion). Pour palier à cela, j'ai proposé une méthode hybride combinant l'estimation de la densité locale de points (pour estimer le bord) et du champ de force (pour la dynamique) dont j'ai comparé les performances vis-à-vis d'approches classiques par maximum de vraisemblance via des simulations de Monté-Carlo.

Dans la seconde partie, j'ai utilisé cette méthode pour caractériser les puits de potentiels dans les nanodomains de haute densité apparaissant dans deux variantes d'épissage des canaux calciques voltage dépendants  $Ca_V2.1$  (avec/sans queue C-terminale) à la membrane de neurones hippocampaux. Ces puits ont des tailles  $\approx 80\text{nm}$ , sont localisés au niveau des zones actives des terminaux présynaptiques où ils y stabilisent les canaux pour  $\approx 100\text{ms}$ , beaucoup plus longtemps que par diffusion. De plus, les puits sont dynamiques et durent  $\approx 30\text{s}$ , compatibles avec la redistribution des  $Ca_V$  après la fusion vésiculaire aux zones actives. Cette analyse associée avec des enregistrements électrophysiologiques et de l'imagerie calcique / glutamique dans diverses conditions suggère que la redistribution des  $Ca_V$  module la probabilité de relâche et la plasticité synaptique à court terme de chaque synapse.

Dans la troisième partie, j'ai analysé des trajectoires de calreticuline dans le lumen du Réticulum Endoplasmique (RE). J'ai trouvé des régions de haute densité corrélées aux jonctions du réseau du RE périphérique et des hautes vitesses corrélées aux tubules. Pour interpréter cette dynamique luminale, j'ai développé une méthode pour re-construire le réseau du RE à partir des trajectoires puis j'ai proposé un modèle stochastique du mouvement où les molécules alternent entre diffusion confinée au niveau des jonctions et sauts à haute vitesse dans les tubules. L'analyse du réseau a révélé que le contenu luminal peut visiter le réseau entier, que le réseau est à l'équilibre et qu'il existe des périodes d'alternance dans la direction du flot dans les tubules. La présence d'un flot actif dans le lumen du RE serait un mécanisme efficace d'homogénéisation de son contenu sur de longues distances.

Dans la dernière partie, je discute de possibles extensions des méthodes présentées pour analyser : 1. le trafic des lysosomes interagissant avec le réseau du RE, 2. la dynamique des  $Ca_V$  au niveau des jonctions neuromusculaires chez la drosophile et 3. les différents modes de mouvement du complexe remodeleur de chromatine NuRD dans le noyau.

En conclusion, cette thèse introduit des modèles stochastiques pour reconstruire la dynamique moléculaire générée par des processus fondamentaux de nanophysologie cellulaire.

## Remerciements

Avant de commencer, je tiens à remercier les personnes qui m'ont accompagné tout au long de cette thèse :

Au niveau professionnel, les collaborateurs avec qui j'ai eu la chance d'interagir : Martin Heine, Jennifer Heck, Edward Avezov, Srinjan Basu et Meng Lu dont les discussions m'ont permis d'apprendre beaucoup sur la biologie, la manière dont ils conduisent leurs expériences ainsi que sur comment produire des analyses qui leur sont pertinentes. Ensuite, les membres du jury qui ont accepté d'évaluer cette thèse : les deux rapporteurs Denis Grebenkov et Charles Kervrann ainsi que les examinateurs : Martin Heine, Dmitri Rusakov, Pierre Sens, Christian Specht et Christian Vestergaard dont la profondeur et la complémentarité des expertises présument d'une soutenance intéressante. Finalement je tiens particulièrement à remercier mon directeur de thèse David Holcman sans lequel je n'aurais pas pu mener à bien ce travail.

Au niveau personnel maintenant, il me semble important de mettre en lumière les personnes sans le soutien desquelles cette thèse n'aurait pas été possible. En commençant par mon père Sylvain qui a toujours été un soutien indéfectible tout au long de ma vie et Céline qui la partage et à qui incombe donc la lourde tâche de me supporter chaque jour. Ensuite mes amis proches (sans ordre particulier (!)) : Gaëtan, Gary, Jérôme, Alexandre, Cédric, Thomas, Cyrille et Claire et tous ceux que j'ai oublié ... Finalement mes collègues du labo présents et passés : Nathanaël, Khanh, Assaf, Jürgen, Nicolas, Ofir, Adrien, Kimsy, Claire et Jérôme.

## Publications & presentations

### Published:

- Jennifer Heck\*, Pierre Parutto\*, Anna Ciuraszkiewicz\*, Arthur Bikbaev, Romy Freund, Jessica Mitlöhner, Maria Alonso, Anna Fejtova, David Holcman and Martin Heine, "*Transient Confinement of  $Ca_v2.1$   $Ca^{2+}$ -Channel Splice Variants Shapes Synaptic Short-Term Plasticity*", Neuron volume 103, pages 66-79 (2019).
- David Holcman\*, Pierre Parutto\*, Joseph E. Chambers, Marcus Fantham, Laurence J. Young, Stefan J. Marciniak, Clemens F. Kaminski, David Ron and Edward Avezov, "*Single particle trajectories reveal active endoplasmic reticulum luminal flow*", Nature Cell Biology volume 20, pages 1118-1125 (2018).
- Pierre Parutto, Jennifer Heck, Martin Heine, and David Holcman, "*Biophysics of high density nanometer regions extracted from super-resolution single particle trajectories: application to voltage-gated calcium channels and phospholipids*", Scientific Reports, 9(1), 1-14 (2019).

### Submitted:

- Meng Lu, Francesca M van Tartwijk\*, Qiaojin Lin\*, Wilco Nijenhuis, Pierre Parutto, Marcus Fantham, Charles N Christensen, Edward Avezov, Christine E Holt, Alan Tunnacliffe, David Holcman, Lukas Kapitein, Gabriele Kaminski Schierle, Clemens F Kaminski, "*The structure and global distribution of the endoplasmic reticulum network is actively regulated by lysosomes*", bioRxiv, 2020 (submitted to Science).

### In preparation:

- Srinjan Basu, Ofir Shukron, Pierre Parruto, Wei Zhang, Wayne Boucher, Tom Drury, Nicola Reynolds, Louisa Sober, Antoine Magre, Alex Carr, David Lando, Jenny Balmer, Alice Aubert, Guillaume Communie, Devina Shah, Keir Murison, Julie Cramard, Robin Floyd, Aleks Ponjavic, Imre Berger, Christiane Berger-Schaffitzel, Steven F. Lee, David Klenerman, David Holcman, Brian Hendrich and Ernest Laue, "*Live-cell 3D single-molecule tracking reveals how NuRD complex modulates enhancer/promoter movement*".
- Pierre Parutto and David Holcman, "*Detection and interpretation of high density nanometer regions embedded in super-resolution single particle trajectories*".

### Oral Presentations:

- Pierre Parutto, David Holcman, "*Characterization of the dynamics of the endoplasmic reticulum luminal protein calreticulin and the  $CaV2.1$  calcium channels at synapses*", Reverse mathematical methods for reconstructing molecular dynamics in single cell workshop, Pisa, October 2018.
- Pierre Parutto, Joseph E. Chambers, Marcus Fantham, Laurence Young, Stefan Marciniak, Clemens F Kaminski, David Ron, David Holcman, Edward Avezov, "*Single Particle Trajectories Reveal Active Endoplasmic Reticulum Luminal Flow*, 63<sup>rd</sup> annual meeting of the biophysical society, Baltimore, March 2019.

(\*: equally contributing)

# Chapter 1

## Introduction

Nowadays more than ever, progresses in technology continuously drive biological discoveries by making once unreachable questions intelligible. During the XXth century, the field of biology has shifted from being qualitative to quantitative, based on precise observations interpreted through models built from physical laws. Indeed, the following quote from R. Feynman, although dating from 1959 is still surprisingly relevant today:

It is very easy to answer many of these<sup>1</sup> fundamental biological questions; you just look at the thing! [...]. Unfortunately, the present microscope sees at a scale which is just a bit too crude. Make the microscope one hundred times more powerful, and many problems of biology would be made very much easier. I exaggerate, of course, but the biologists would surely be very thankful to you—and they would prefer that to the criticism that they should use more mathematics.

(Richard Feynman's December 29th 1959 speech at the Annual meeting of the American Physical society).

"Looking" in the study of life also has a temporal component that becomes more and more a subject of study in modern biology. Spatially speaking, life spans multiple ranges of sizes, from large organisms and populations, down to their building blocks that are DNA, proteins, and cells at the microscopic scale and below. Similarly, on the temporal scale, some biological processes need minutes to months to complete (e.g. human reproduction or bacterial fission) while others such as molecular interactions (e.g. enzymatic reactions) or synaptic transmission take milliseconds or below. Thoroughly understanding a biological process is thus not limited to a static picture, but also requires to capture its dynamic with the correct spatio-temporal scale. For example, Electron Microscopy (EM) techniques allow

---

<sup>1</sup>The previous paragraph of the quote reads: "They are questions like: What is the sequence of bases in the DNA? What happens when you have a mutation? How is the base order in the DNA connected to the order of amino acids in the protein? What is the structure of the RNA; is it single-chain or double-chain, and how is it related in its order of bases to the DNA? What is the organization of the microsomes? How are proteins synthesized? Where does the RNA go? How does it sit? Where do the proteins sit? Where do the amino acids go in? In photosynthesis, where is the chlorophyll; how is it arranged; where are the carotenoids involved in this thing? What is the system of the conversion of light into chemical energy?"

to visualize a sample with a few nanometer resolutions but require fixation of the sample (either chemically or by freezing) thus totally discarding the temporal component. On the other hand, microscopes implementing super-resolution methods are usually able to acquire images every  $\approx 5 - 100$ ms (for example [1]) or even below (for example [2]) but with much lower spatial resolutions of the order of the dozen to hundred nanometers.

Of prime interest in biology is the precise understanding of the location and dynamics of proteins inside cells in normal and perturbed conditions, a subject that has received an increased wave of interest since the 2000s. Supported by multiple (bio)-technological breakthroughs: the characterization of the Green Fluorescent Protein and continuous improvements of subsequent fluorescent labels, the development of photoactivable / photo-bleachable labels and improvements in microscopy technology, in particular, the advent of super-resolution microscopy. Super-resolution microscopy defines an ensemble of techniques allowing to observe individual molecules in live samples using confocal microscopes, circumventing the diffraction limit of light. Among these techniques, the Single Molecule Localization Microscopy (SMLM) family is especially concerned in detecting and following individual molecules of interest through time. The advent of such advanced acquisition methods also requires the development of new data analysis procedures and models in order to quantitatively exploit these data.

SMLM techniques for imaging Photo-Activable Fluorescent Proteins (PAFP) combined with particle detection and tracking algorithms are able to provide large ensembles of Single Particle Trajectories (SPTs) of proteins of interest [3]. They are the continuation of previously existing Quantum Dot (QD) imaging techniques that were able to obtain longer trajectories (thanks to QD's greater photostability, although more prone to blinking) but QDs are larger than FPs and thus more prone to get stuck in sub-structures and are harder to work with, requiring functionalization before usage [4, 5]. Since the inception of PAFP-based SMLM techniques in the late 2000s, technological improvements in microscopes, fluorescent proteins [6], and processing softwares have allowed to target new proteins, reduce motion artifacts caused by the marker, reach better spatio-temporal resolutions and reconstruct more faithful trajectories.

The biophysical study of SPTs relies on specifying a model of motion, the gold standard being the Langevin equation that considers the dynamics as being driven by possible external forces as well as thermal fluctuations (diffusion). Historically, the tool used to estimate the parameters from this equation is the Mean Squared Displacement (MSD) that allows to recover the diffusion coefficient and roughly characterizes the observed motion as transport, purely diffusive or confined. MSD applies to long trajectories that do not change their mode of motion (homogeneous motion), which is problematic as SMLM techniques produce many short trajectories and cells are a thriving, crowded environment full of interactions, making it more likely for a molecule to exhibit a heterogeneous motion. The new generation of tools tailored for the study of large ensembles of short SPTs instead rely on the characterization of the local dynamics.

In this thesis, I developed new computational and statistical procedures, improved existing ones and used Monte-Carlo simulations to compare their performance. Particularly, I was interested in the characterization of local structures called potential wells, on the es-

timation of parameters of motion when molecules are subject to spatiotemporally random forces and on the reconstruction of higher-order structures from SPTs.

The manuscript is structured as follows: In the first chapter, I introduce the field of super-resolution microscopy, its applications in neurosciences and cellular biology as well as the main models of motion used for interpreting SPTs and their parameter estimators. In the second chapter, I present a new method for estimating the boundary of potential wells, discuss different ways of recovering the other well parameters and compare them on both simulated and empirical SPTs from GPI-anchored GFP and  $\text{Ca}_v2.2$  on the membrane of hippocampal neurons. In the third chapter, I present another application where I analyzed SPTs of voltage-gated calcium channels  $\text{Ca}_v2.1$  at hippocampal synapses where I characterized the presence of potential wells at active zones providing a mechanism that stabilizes channels there and has important implications in tuning synaptic release probability and short-term plasticity. In the fourth chapter, I present the analysis of SPTs from multiple Endoplasmic Reticulum (ER) resident proteins where I discovered the presence of a flow dynamics in the peripheral ER lumen contrary to the previous thinking that the motion was diffusive. I then study both static and transient ER network characteristics reconstructed from SPTs. Finally, in the fifth chapter I present as a discussion three other applications of the same type of analyses: the characterization of the interactions between the ER and lysosomes, the study of potential wells at the neuro-muscular junctions of drosophila and their comparison with the hippocampal  $\text{Ca}_v2.1$  wells and finally the characterization of multiple modes of motion for the components of the NurD chromatin remodeling complex.

## 1.1 Nanodomains in neurosciences and cell biology

Except when tethered, molecules in living cells are constantly moving through thermal agitation and are often maintained out of equilibrium by active process preventing some molecules to stay in specific locations (e.g protein export in the endoplasmic reticulum) or enforcing their presence in some place (e.g receptor channels at post-synaptic densities). These local differences in concentrations create nanodomains with specific geometries where molecules are organized to ensure a specific function (e.g. pre-synaptic active zones). The study of molecular dynamics and nanodomains allows a better understanding of the functioning of the biophysical processes present in cells.

### 1.1.1 Receptor dynamics at synapses

Synapses are key structures of brains that connect two neurons through the conversion of an electrical signal arriving at the presynaptic terminal into a chemical signal diffusing through the synaptic cleft and back to an electrical signal in the postsynaptic neuron. On the presynaptic side, the arrival of an electrical current (action potential) triggers the opening of voltage-dependent calcium channels (Fig. 1.1A), allowing entry of calcium ions into the cell leading to the fusion of some synaptic vesicles to the plasma membrane and the release of neurotransmitters into the synaptic cleft (Fig. 1.1B). On the post-synaptic side, ligand-gated channels open upon binding to neurotransmitters allowing ion transfers through the

membrane generating a post-synaptic electrical current (Fig. 1.1C). Synapses thus rely on a large variety of molecular species and their interactions in a very small region to function and is thus really prone to be studied by SPTs.

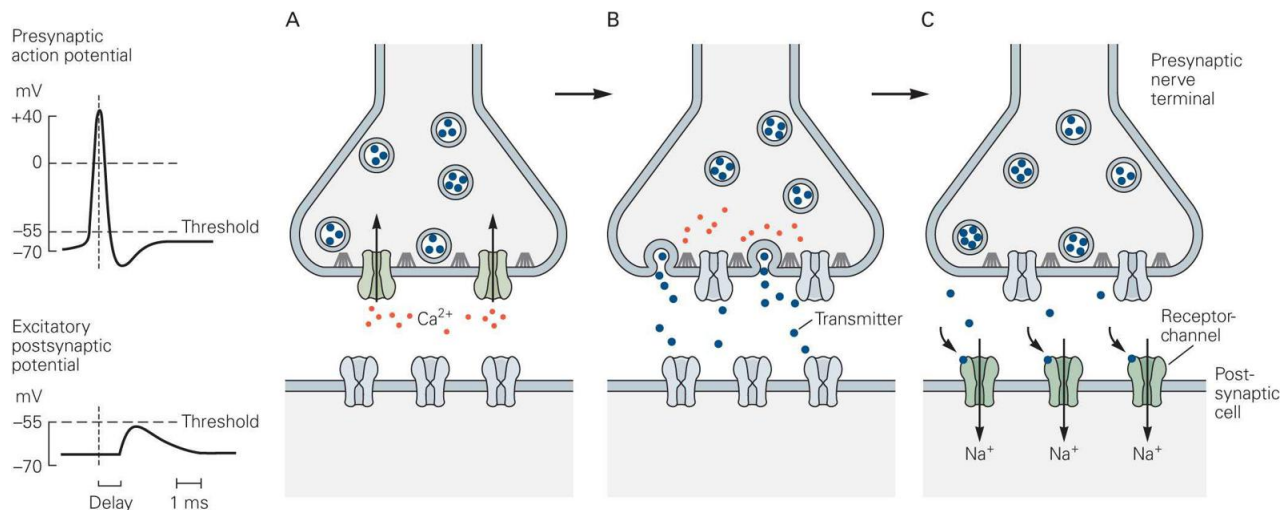


Figure 1.1: **Schematic steps of synaptic transmission of electrical currents (from [7]).** **A.** voltage-gated calcium channels open upon the arrival of an action potential allowing the entry of  $\text{Ca}^{2+}$  ions in the cell. **B.**  $\text{Ca}^{2+}$  ions eventually bind to synaptic vesicles' tethering machinery triggering their fusion with the plasma membrane. **C.** Released neurotransmitter diffuse through the synaptic cleft eventually reaching a postsynaptic receptor.

In the late 1990s technology allowed to study individual synapses constituents (channels, receptors, scaffolds, vesicles, ...) revealing the importance of their dynamics in synaptic plasticity (the modification of a synapse strength depending on its activity). Interestingly, these studies were able to understand the importance of molecular dynamics without directly looking at it but instead using only confocal images, genetic manipulations, application of chemical treatments and electrophysiological recordings. In particular, the following studies investigate the role of trafficking and retention at post-synaptic densities of AMPA receptors and its sub-units in relation to plasticity [8, 9, 10, 11]. In a 2000 review article [12], Scannevin and Hugarir conclude that "In contrast to initial impressions that considered the postsynaptic complex to be a static structure, it is now clear that the [post-synaptic protein] lattice is a very dynamic complex that is critical in the rapid modification of the efficacy of synaptic transmission during synaptic plasticity."

### 1.1.2 Recovery after photobleaching experiments for characterizing molecular dynamics

Before SPTs was a common tools, or in contexts where it cannot be applied, the Fluorescence Recovery After Photobleaching (FRAP) technique gives access to the dynamics from fluorescence labeling, allowing to estimate the diffusion coefficient and immobile molecule

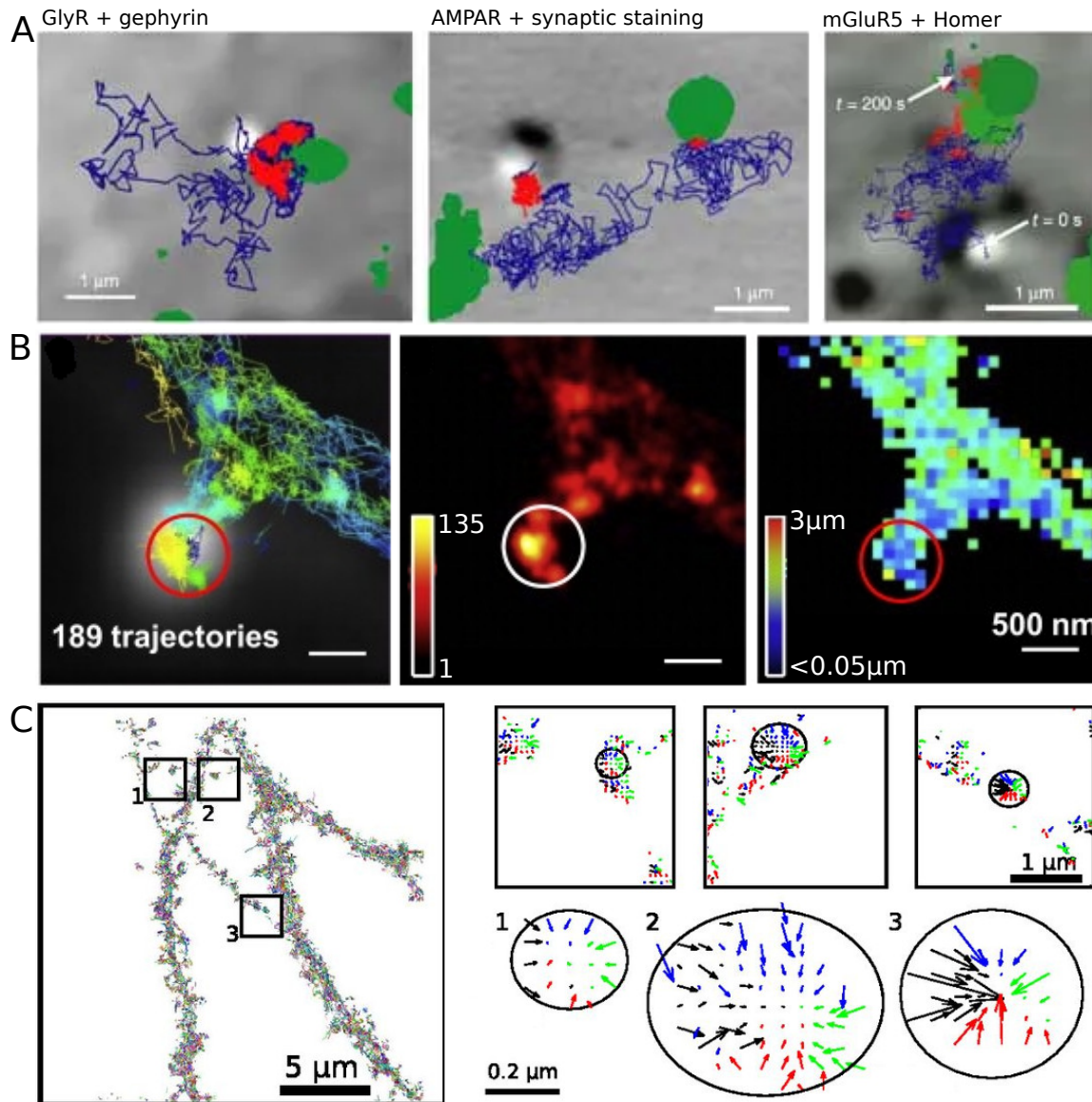


Figure 1.2: **Single particle trajectories of post-synaptic receptors.** **A.** First generation SPTs of post-synaptic receptors using latex-beads (from left [13], middle [14], right [15]). **B.** High-density UPAIN trajectories of AMPA receptors overlaid on a Homer1C::GFP fluorescence signal to detect post-synaptic densities represented here as a circle (left), reconstructed SPT image (middle) and map of local displacement lengths (right), from [16]. **C.** Trajectories of AMPA receptors labeled with mEos2 fluorescent proteins (left) and confinement regions revealed by the analysis of the local drift vector field (arrow color depends on direction) showing areas (circle) with converging drift arrows, from [17].

fraction. This techniques consists in photobleaching a small portion of the experimental plane and measuring the fluorescence recovery corresponding to the rate of replacement of bleached fluorophores by intact ones [18]. Its first uses emerged in the 70s [19] and the theory behind these experiments have been developed in the seminal paper of Axelrod [20]. FRAP



has especially been used to characterize the luminal motion in the Endoplasmic Reticulum (ER) [21]. The ER is an important organelle for neurosciences as it is a calcium store strongly involved in calcium homeostasis, releasing calcium when required and absorbing it when the local concentration is too high. This regulation of calcium by the ER is of particular interest in presynaptic terminals of neurons where calcium plays a crucial role as described in the previous paragraphs.

### 1.1.3 Trajectories of individual post-synaptic receptors

This first generation of SPTs studies, although suffering from multiple technical limitations such as the huge bead size and few numbers of acquired trajectories ( $\approx 40$  reported per condition in these articles), lead to the definition of a model of a dynamic equilibrium at post-synaptic densities with a constant receptor turnover [22]. Later on, another study [23] using quantum dots and fluorescent proteins (much smaller than latex beads) revealed a role for this equilibrium, in compensating for AMPA receptor desensitization by replacing them with functional ones for the fast recovery of synaptic properties at the millisecond timescale.

A breakthrough in the way of conducting SPTs studies came from the development of methods that allow to label a much larger fraction of the population of proteins of interest, and visualizing multiple molecules at the same time [3]. Among these, the UPAINT technique, based on the stochastic binding of extracellular fluorescent ligands to transmembrane molecules, allowed to visualize with high density an endogenous population of AMPA receptors at the surface of hippocampal neurons [16]. This study found reduced diffusion coefficients and increased density of receptors near synapses (Fig. 1.2B). Following articles based on similar high-density methods, used new analysis techniques taking advantage of this high-density, to reveal transient confinement, in the form of potential wells, of postsynaptic receptors both at synapses and on dendrites' necks [17, 24] (Fig. 1.2C).

## 1.2 Super-resolution microscopy of single particle trajectories

I now present the ideas behind the most widely used super-resolution microscopy methods as well as single-particle tracking techniques that allowed the high-density imaging breakthrough presented in the previous paragraph.

### 1.2.1 Abbe's law for diffraction limit

Our capacity to observe small objects is limited by the resolution of the microscopes given by the physical law of diffraction for the minimum resolvable distance  $d$  between two fluorescent molecules on the focal plane [25]:

$$d = \frac{\lambda}{2\text{NA}} \quad (1.1)$$

where  $\lambda$  in *nm* is the wavelength of the emitted beam and NA is the numerical aperture that depends on the immersing medium of the objective and the characteristics of the microscope. This law means that a single fluorescent emitter will generate a scattering pattern (see Fig. 1.3A) with a size proportional to its emission wavelength and the characteristics of the microscope. For optical microscopy (based on photon beams) this pattern size is much larger than the size of the emitter by a factor  $\approx 50$ . For example, a Green Fluorescent Protein [26], is shaped like a  $4.2 \times 2.4$ nm barrel and has an emission peak at 509nm, when imaged on a microscope with a numerical aperture of 1.2 will generate a spot  $\approx 212$ nm radius on the optical plane.

Abbe's law, however, is not about how precisely the position of a molecule can be reconstructed from its observed emission pattern. Indeed, an immobile emitter will generate a certain amount of photon during one acquisition period. These photons will be scattered around the real position of the emitter with a specific distribution called a point spread function (PSF) (see Fig. 1.3B). The reconstruction of the molecule position thus depends on the precision in determining the PSF center based on the photon distribution and will greatly depend on the emitter motion and number of acquired photons, proportional to the acquisition time and the emitter quantum yield (see Fig. 1.3C). On the other hand, Abbe's law says that two molecules closer than the diffraction limit  $d$  will generate overlapping patterns on the optical plane, creating instabilities in their PSF estimation and below a certain distance becoming indistinguishable (see Fig. 1.3D).

## 1.2.2 Super-resolution microscopy techniques

Super-resolution microscopy circumvents the resolution limitation from Abbe's law by making sure that only one molecule is fluorescent at a given time and position. It encompasses three main techniques:

- Single Molecule Localization Microscopy (SMLM) such as Photo-Activated Localization Microscopy (PALM) [28] and Stochastic Optical Reconstructed Microscopy (STORM) [29]. These methods rely on the use of photo-activable fluorophores which activation is tuned such that only a very sparse and spatially spread sub-population is activated at the same time.
- Stimulated Emission Depletion (STED) [30] that uses two laser beams: an excitation laser to activate fluorophores in a circular region and a depletion beam to deplete the fluorescence in a donut shape around the excited region, thus only conserving the signal at the center.
- Structured Illumination Microscopy (SIM) [31] where the idea is to project a certain number of very specific illumination patterns on the sample and reconstruct an image using the interferences between the projected patterns and the recorded fluorescence.

By preventing interferences between multiple molecules patterns these techniques allow to obtain an increase by a factor  $\approx 10$  in spatial resolution, from  $\approx 200$ nm to  $\approx 20$ nm and, being

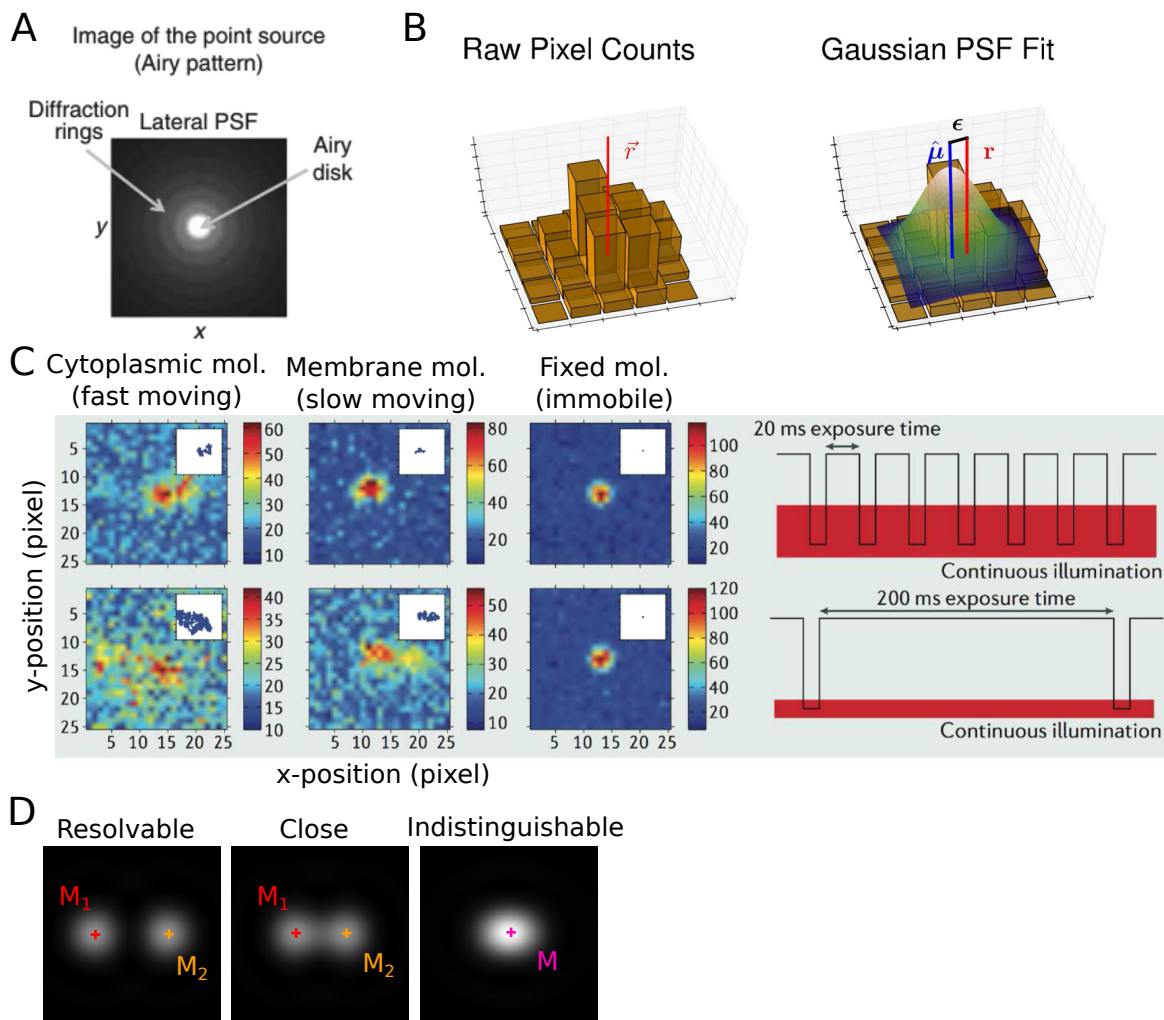


Figure 1.3: **Principles of single particle detection in optical microscopy.** **A.** A single emitter generates a scattering pattern (Airy disk) on the optical plane (left). **B.** Using a digital camera, the optical plane is divided into pixels and the number of photons falling into each pixel during the acquisition period is counted (middle). Finally, a point spread function (here a Gaussian of center  $\mu$  and standard deviation  $\epsilon$ ) is fitted to this discrete photon count distribution to recover the position of the emitter (Right). **C.** Example of recorded photon distributions for three different categories of molecules: a fast-moving cytoplasmic (left column), a slow-moving transmembrane (middle) and a fixed (right) molecule and using either a small (20ms) acquisition time (top) or a larger (200ms) one (bottom). These different scenarios lead to different acquired photon count distributions where faster molecules and/or larger acquisition time lead to less accuracy in the emitter position reconstruction. **D.** Example of two closely positioned emitters that have close but non-interfering (left), interfering (middle) and totally overlapping (right) scattering patterns. Images are adapted from ref. [27].

based on confocal microscopy, they can directly be used on live samples (in culture [32, 33], slice [34] or in vivo [35, 36]). SMLM techniques also allow to visualize the dynamics of individual molecules opening the way for a quantitative analysis of molecular dynamics [3, 17, 24, 16, 37].

It is to note however that most of the articles relying on super-resolution microscopy are based on static image analysis using techniques such as: the characterization of local density distributions [38, 39], the co-localization of multiple fluorescence signals [40, 41, 42], molecule density clustering [43] / density distribution characterization [44, 45], time correlation of molecular detections [46] and others.

### 1.2.3 Constructing single particle trajectories from localization microscopy experiments

The process of recovering an ensemble of Single Particle Trajectories (SPTs) from a stack of single-molecule localization images is called Single Particle Tracking [47] and is composed of two steps:

1. **Fluorescent spots detection** for each frame, groups of pixels with intensities above the background noise, assumed to belong to a single molecule, are extracted. Then the local intensity distribution of each group is fitted to a point spread function (usually a Gaussian) to extract its center. This is made difficult when: fluorescent emitters density becomes high, emitters have a low yield compared to the acquisition time, there is a non-uniform background noise or due to dynamic noise caused by molecule motion during the acquisition [48] (as presented in Fig. 1.3). Examples of methods for detecting fluorescent spots are: intensity thresholding, probabilistic models [49], generalized Laplacian of Gaussians [50], multiscale wavelet decomposition [51, 52] or fast radial symmetry detection [53]. The spots detected this way are associated with a frame, a position (2 or 3D) and possibly other data such as their max./min./AVG. intensity, contrast or signal to noise ratio (Fig. 1.4A)
2. **Detection Linking** linking spots detected from successive frames to form trajectories. The two main quantities used in the decision of linking two spots are their spatial and temporal distances. Examples of tracking algorithms include: nearest neighbor distance [54], multiple hypothesis testing [55] (considering a model of motion), Linear Assignment Problem [56] or simulated annealing heuristic for the multi-frame objects correspondence problem [57]. The resulting ensemble of trajectories can be visualized by neglecting the time component and using one color per trajectory (Fig. 1.4B).

Classical single particle tracking softwares include TrackMate [58], u-track [56] or Icy's tracking plugin [55]. Obtaining good quality trajectories is about making tradeoffs: for the detection, balancing the amount of false positive (misinterpreting a background variation for a fluorescent spot) and of false negatives (not detecting a fluorescent spot). Being too lax will corrupt the results with false background detection while being too conservative prevents

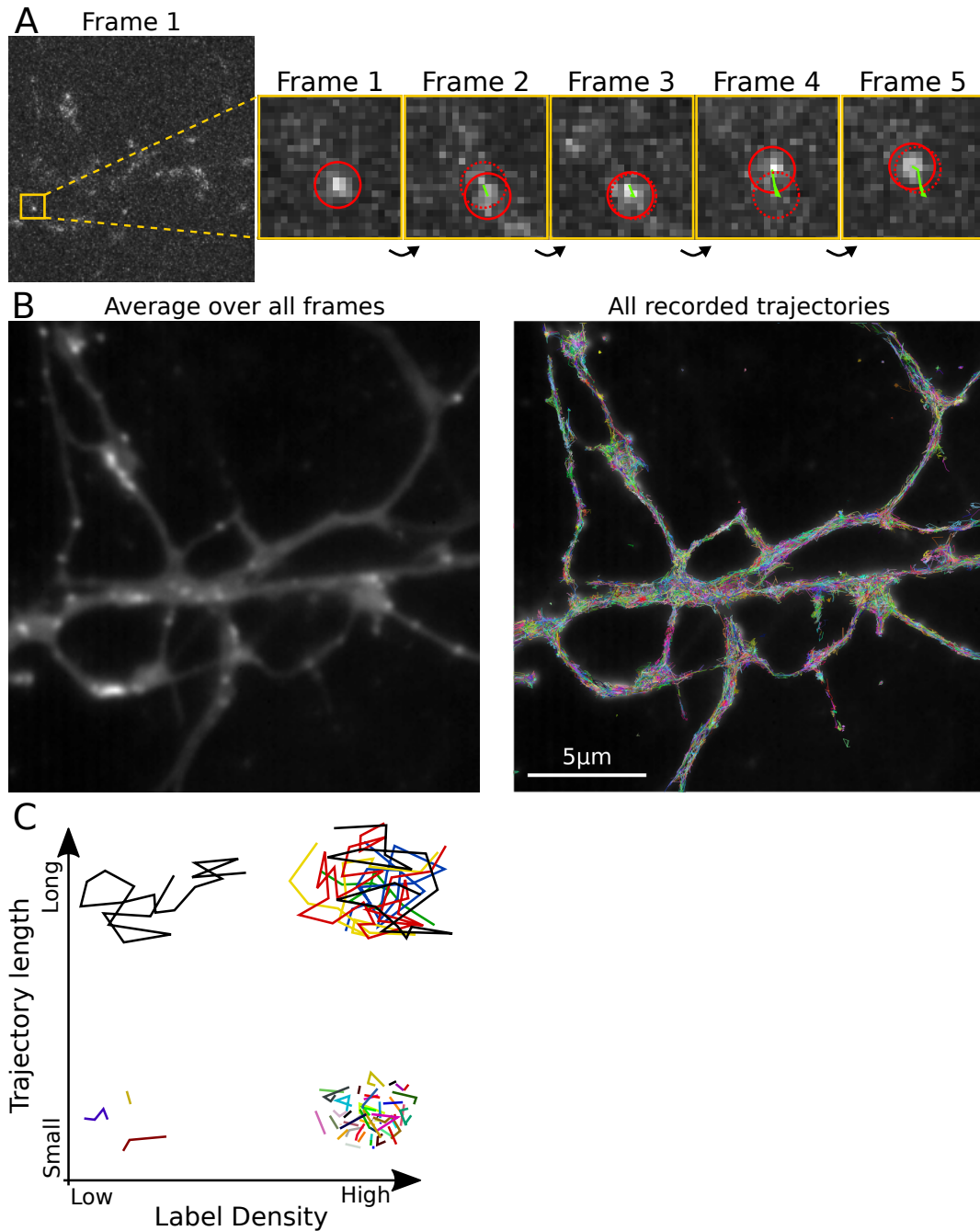


Figure 1.4: **Single particle tracking.** **A.** Example of a single image of meos3.2::Ca<sub>v</sub>2.1 channels on the membrane of hippocampal neurons. Inset: successive frames of the selected region with the detected molecule (filled circle, dashed: detection at the previous frame) and the nascent trajectory (green). **B.** Results from the complete image stack: average image (left) and all recorded trajectories (right). **C.** Classes of tracking problem depending on the expected trajectory length and label density.

forming extended trajectories. For the linking, the balance is now about not linking too many different molecules in the same trajectory but also not to be too restrictive in the linking and miss some parts of the dynamics. Up to my knowledge, it is still not clear how particle tracking can bias the recovered motion but different analysis strategies can be used depending on the density and stability of fluorophores.

Three classes important of SPTs experiments can be distinguished (Fig. 1.4C): low density/long trajectories (as generated by early methods relying on beads) suited for mean squared displacement analysis (see sub-section 1.4.1); Low density/short trajectories (as obtained by sptPALM) suited for local estimators (see sub-section 1.4.2); High density/long trajectories that can be studied both locally and at long range as presented in this thesis (see chapters 4 and 5). Finally, having both a small number of short trajectories is not suitable for analysis.

### 1.2.4 Biophysical features contained in SPTs

SPTs analysis techniques can be decomposed in two categories: the ones based on physical models of motion and those based on computing trajectory features:

**Characteristics based on physical models of motion** The most common piece of information extracted in this category is the diffusion coefficient when approximating molecular motion as diffusive (see section 1.4). To go further, more complicated modes of motion can be considered, such as confined motion for extracting the confinement radius, trapping due to parabolic potential wells for recovering the force acting on the molecule, or switching modes of motion for recovering the different parameters of motion and switching rates (for example alternating between multiple diffusion coefficients).

**Characteristics based on trajectory features** On the other hand, some observations can be obtained without relying on a model of motion, but instead directly computing quantities based on detections and trajectories. These kinds of information include computing the detection points density and measures based on the point spatial distribution, the temporal repetition of a detection or confinement durations based on trajectories classification.

It is worth noting however that scientific reasoning based on SPTs has multiple shortcomings. First, although SPTs have allowed circumventing the diffraction limit of light, reaching single molecule precision is at the cost of losing all contextual information. This first shortcoming can however be alleviated by combining SPTs with confocal images of other landmark proteins and co-localizing both signals. Then, SPTs experiments rely on the stochastic activation of a small subset of a population with unknown spatio-temporal distribution making the number of expected molecules in a given region for a given time impossible to predict, limiting the capacity of SPTs to address any question based on the absence of event. SPTs are thus mostly confined to answering questions that can be formulated by the analysis of observed trajectories. Finally, trajectories lengths, limited by fluorophores characteristics (photo-bleaching) and the increase in the false-linking probability with the trajectory

length, makes it impossible to visualize long events (relative to the acquisition time of the microscope) in individual trajectories.

## 1.3 Stochastic models of individual particles

The field of individual particle motion analysis started with the observation of the disordered movement of individual pollen particles suspended in a liquid by the botanist Robert Brown which coined the term Brownian motion. The theory underlying these observations and its relation with the macroscopic diffusion phenomenon was laid out by Einstein in 1905 [59], Smoluchowski in 1906 [60], Langevin in 1908 [61] and later on adapted for transmembrane proteins by Saffman and Delbrück [62]. All these models are made in the limit of the observation time  $\Delta t$  being much greater than the collision time, such that the resulting motion is generated by enough collisions to be uncorrelated and modeled statistically.

### 1.3.1 Stochastic theory of diffusion

The physical model behind purely diffusive motion is based on the study of molecules in a gas, where the motion is driven only by the stochastic collisions between molecules. At a microscopic level, one collision generates a ballistic motion, while many collisions, due to their random directions, can only be modeled statistically. In this case, the equation of motion for a molecule at position  $\mathbf{X}(t)$  at time  $t$  is

$$\dot{\mathbf{X}} = \sqrt{\frac{2k_B T}{m}} \mathbf{w}, \quad (1.2)$$

where  $T$  is the temperature of the gas,  $k_B$  the Boltzmann constant providing the average kinetic energy as a function of the temperature,  $m$  is the mass of the molecule and  $\mathbf{w}$  is a vector of identically independently distributed Gaussian variables with mean  $E[\mathbf{w}(t)] = 0$  and covariance  $E[\mathbf{w}(t)\mathbf{w}(s)] = \delta(t-s)$  ( $\delta$  is the Dirac function). Now, for molecules immersed in a fluid, the diffusion coefficient can be defined in relation to the physical properties of the molecule and medium through the Stokes-Einstein relation [59]  $D = \frac{k_B T}{\xi}$  with  $\xi = 6\pi r\eta$  the drag coefficient,  $r$  the radius of the (considered spherical) molecule and  $\eta$  the viscosity of the surrounding fluid. This model is correct only for observation times  $\delta t \gg \gamma^{-1}$  (with  $\gamma = \xi/m$  the dynamical friction coefficient) the momentum relaxation time of the molecule [63].

### 1.3.2 Diffusion and active forces

In a more biologically realistic context, molecules in a cell are also subject to local forces that affect their motion. This phenomenon can be modeled from Newton's equations of motion as:

$$m\ddot{\mathbf{X}} = \sum_i \mathbf{F}_i, \quad (1.3)$$

that gives a relation between the acceleration  $\ddot{\mathbf{X}}$  and the sum of the forces  $\mathbf{F}_i$  acting on the molecule. For a diffusive molecule immersed in a viscous laminar flow, and subject to a potential of energy  $U(\mathbf{X})$ , this equation becomes the Langevin equation [61]:

$$m\ddot{\mathbf{X}} + \gamma\dot{\mathbf{X}} + \nabla U(\mathbf{X}) = \sqrt{2D\gamma^2}\dot{\mathbf{w}}, \quad (1.4)$$

where  $\gamma\dot{\mathbf{X}}$  is the frictional force due to the immersing fluid and  $\nabla U(\mathbf{X})$  the force generated by the potential  $U$ . In The limit where the magnitude of the frictional force is larger than the inertia  $\|\gamma\dot{\mathbf{X}}\| \gg \|m\ddot{\mathbf{X}}\|$ , then the acceleration term can be neglected leading to the the overdamped Langevin equation or Smoluchowski limit of the Langevin equation [64]:

$$\dot{\mathbf{X}} = -\frac{\nabla U(\mathbf{X})}{\gamma} + \sqrt{2D}\dot{\mathbf{w}}. \quad (1.5)$$

### 1.3.3 Nanodomains modeled as parabolic potential wells

The potential force  $U(\mathbf{X})$  from the Langevin equation can be approximated, close to the potential center  $\boldsymbol{\mu} = (\mu_x, \mu_y)$ , to the second order as

$$U(\mathbf{X}) = A \left( \frac{(x - \mu_x)^2}{a^2} + \frac{(y - \mu_y)^2}{b^2} \right) + O(x, y)^2, \quad (1.6)$$

where  $A$  is the attraction strength of the well and  $a, b$  define the shape of the well in the  $x$ - and  $y$ -dimension respectively. This equation corresponds to a parabolic potential well that can alternatively be rewritten in matrix form as

$$U(\mathbf{X}) = (\mathbf{X} - \boldsymbol{\mu})\boldsymbol{\Gamma}(\mathbf{X} - \boldsymbol{\mu})^T, \quad (1.7)$$

where  $.^T$  is the matrix transpose,  $\boldsymbol{\mu} = [\mu_x, \mu_y]$  is the center and  $\boldsymbol{\Gamma}$  is a matrix proportional to the characteristics of the well  $A, a, b$ . The force generated by such potential is given by

$$\nabla U(\mathbf{X}) = 2\boldsymbol{\Gamma}(\mathbf{X} - \boldsymbol{\mu}) \quad (1.8)$$

and the equilibrium particles' density for this potential follows the Boltzmann distribution

$$\rho(\mathbf{X}) = C \exp\left(-\frac{U(\mathbf{X})}{k_B T}\right), \quad (1.9)$$

with  $C$  a normalization constant and the  $k_B T$  term corresponds to the thermal energy at the well and can be associated to the diffusion coefficient  $D$ . The motion in such potential well is modeled by the Ornstein-Uhlenbeck stochastic process [65]:

$$\dot{\mathbf{X}} = -\boldsymbol{\Lambda}(\mathbf{X} - \boldsymbol{\mu}) + \sqrt{2D}\dot{\mathbf{w}} \quad (1.10)$$

with  $\boldsymbol{\Lambda} = 2\boldsymbol{\Gamma}$  the coefficient of the attraction strength of the well. Such potential well model allows to compute a quantity that is not obtainable from the direct observation of trajectories: the residence time  $\tau_e$  of a molecule inside the well. Short trajectories will often



not allow to see an entire trapping event (free motion followed by trapping followed by free motion) and thus cannot be used to quantify trapping duration. On the other hand, this value can be approximated for a two-dimensional circular potential well of radius  $r$  as [17]

$$\tau_e = \frac{Dr^2}{4A^2} e^{\frac{A}{D}}. \quad (1.11)$$

This quantity can then be compared to the time  $\tau_B$  needed for a particle to escape the same region only through Brownian motion ( $A = 0$ ), given by [17]

$$\tau_b = \frac{r^2}{4D}. \quad (1.12)$$

The ratio  $\frac{\tau_e}{\tau_B}$  is thus a measure of the strength of the interaction relative to the fluctuation level. This formula cannot be directly extended for elliptic wells however. Instead, for ellipses that are not too much elongated, we can use equation (1.11) with the approximation  $r = \sqrt{ab}$  with  $a, b$  are the two semi-axes lengths.

### 1.3.4 Discrete equation of motion

When working with single particle trajectories, the minimal time between two observations is limited by the image acquisition time  $\Delta t$  of the microscope. Depending on the recording technology,  $\Delta t$  can vary but usually remains of the order of a few to a few hundred milliseconds. This acquisition time being much larger than molecular dynamics, the equation of motion corresponding to the recorded trajectories is better modeled by a discretized version of the Langevin equation

$$\Delta \mathbf{X}(t) = \mathbf{b}(\mathbf{X}(t))\Delta t + \sqrt{2D_e(\mathbf{X}(t))\Delta t}\boldsymbol{\eta}, \quad (1.13)$$

where  $\Delta \mathbf{X}(t) = \mathbf{X}(t + \Delta t) - \mathbf{X}(t)$  is called the displacement,  $\mathbf{b}$  is the local drift,  $D_e$  the effective diffusion coefficient and  $\boldsymbol{\eta} = [\eta_1, \dots, \eta_d]$  is a  $d$ -dimensional vector of iid Gaussian variables of mean zero and variance one. The effective diffusion coefficient  $D_e$  depends on the crowding of the local environment [66], the size and shape of the imaged molecule and the properties of the fluid/membrane it is embedded into [62, 67, 68]. The only fixed quantity being a priori the shape of the molecule, it should be expected for the diffusion coefficient to exhibit spatio-temporal fluctuations.

## 1.4 Parameter estimation from SPTs

Once the equation of motion that followed by a molecule is characterized, the next step is to extract the corresponding parameters. For a purely diffusive motion, the only parameter is the diffusion coefficient  $D$  while in more realistic situations, more parameters are present such that the amplitudes and directions of external forces or membrane geometry.

### 1.4.1 Mean squared displacement estimators

The classical parameter estimation method is based on the computation of the Mean Squared Displacement (MSD) along an ensemble of trajectories and stems from Einstein's 1905 observation that this quantity grows linearly with time [59]. For a trajectory  $\mathbf{X}(t)$  ( $t \geq 0$ ), the squared displacement for a given lag time  $\tau$  is computed as  $\|\Delta_\tau \mathbf{X}(t)\|^2 = \|\mathbf{X}(t+\tau) - \mathbf{X}(t)\|^2$  ( $\|\cdot\|$  is the Euclidean norm) and its average should increase linearly with  $\tau$  as:

$$\langle \|\Delta_\tau \mathbf{X}(t)\|^2 \rangle = 2Dd\tau \quad (1.14)$$

where  $\langle \cdot \rangle$  represents the empirical average over some ensemble and  $d$  is the number of dimensions. Formula (1.14) is called the Mean Squared Displacement (MSD) and possesses multiple variants depending on the ensemble chosen for the average. We talk about trajectories-averaged MSD when the average is taken over multiple trajectories [69]:

$$\langle \|\Delta_\tau \mathbf{X}(t)\|^2 \rangle = \frac{1}{N} \sum_{i=1}^N \|\Delta_\tau \mathbf{X}_i(t)\|^2, \quad (1.15)$$

where  $N$  is the number of available trajectories. On the other hand, we talk about time-averaged MSD (taMSD) when the average is taken over multiple sub-parts of the same trajectory [70, 71, 72]:

$$\langle \|\Delta_\tau \mathbf{X}(t)\|^2 \rangle_T = \frac{1}{K} \sum_{i=1}^{T-\tau} \|\Delta_\tau \mathbf{X}(t_i)\|^2 \quad (1.16)$$

where the trajectory lasts for a time  $T$  and  $K$  is the number of possible sub-parts of size  $\tau$  that can be formed from the trajectory. In both cases, given the number of dimensions of the data, the diffusion coefficient can be recovered from the trajectories by a standard linear fit to the mean squared displacement as a function of the time-lag  $\tau$  curve.

#### MSD and motion type

The MSD curve also contains information about the type of motion. Indeed equation 1.14 assumes that the motion is unconfined and purely diffusive. Any deviation from this motion, affecting the displacement length, such as confinement or transport will thus be reflected in the MSD curve and the MSD can be more generally approximated as

$$\langle \|\Delta_\tau \mathbf{X}(t)\|^2 \rangle = A\tau^\alpha \quad (1.17)$$

where  $\alpha \in [0, 2]$  is called the anomalous exponent and  $A$  is a constant such that  $A = 2dD$  when  $\alpha = 1$ . The value of  $\alpha$  reflects the underlying type of motion:  $\alpha < 1$ , leading to a plateaued curve corresponds to a confined motion while  $\alpha > 1$  leading to a super-linear behavior implies some form of transport motion with ultimately  $\alpha = 2$  corresponds to a purely ballistic motion. An alternative method used especially for estimating the anomalous exponent  $\alpha$  is the Mean Maximal Excursion (MME) method [73] based on the computation of the largest distance from the starting point reached by a particle up to a given time.

## Limitation of MSD

The MSD is a very efficient estimator to the diffusion coefficient of Brownian motion but fails when the motion is mixed or the parameters are not constant [74]. Indeed, the assumption behind the taMSD is that the mode and parameters of motion along a trajectory do not change. For trajectories-averaged MSD, the assumption is similarly that the type of motion and its parameters are the same for different trajectories (separated in space and also possibly in time). These assumptions fail when the parameters of motion are not spatially or temporally constant or any transient event such as trapping or transport occur. Actually, early SPTs articles based on MSDs already used a two lines fit method to account for the short and long-term behavior of the MSD curve [71, 13].

Finally, the MSD estimation relying on the presence of long trajectories, it is not well suited when fluorophores density becomes higher making the particle tracking more difficult and introducing errors in trajectories reconstruction.

### 1.4.2 Local estimators constructed from statistical moments

Another category of estimators called local estimators, allow to recover an estimate of both the diffusion coefficient, and the local drift field, based on the average of the trajectory displacements  $\Delta\mathbf{X}(t) = \mathbf{X}(t + \Delta t) - \mathbf{X}(t)$  falling into a specific region. Indeed the following relations hold for any point  $\mathbf{x}$  of the experimental plane [75, 76, 77, 78]

$$\begin{aligned}\mathbf{b}(\mathbf{x}) &= \lim_{\Delta t \rightarrow 0} \frac{E[\Delta\mathbf{X}(t) | \mathbf{X}(t) = \mathbf{x}]}{\Delta t} \\ \mathbf{D}(\mathbf{x}) &= \lim_{\Delta t \rightarrow 0} \frac{E[\Delta\mathbf{X}(t)^T \Delta\mathbf{X}(t) | \mathbf{X}(t) = \mathbf{x}]}{2\Delta t},\end{aligned}\tag{1.18}$$

where  $\mathbf{X}(t) = [\mathbf{X}^{(1)}(t), \mathbf{X}^{(2)}(t)]$  is a trajectory point,  $\Delta t$  the timestep,  $\mathbf{D}(\mathbf{x})$  the diffusion tensor and  $E[.]$  the expectation over all trajectory displacements passing at point  $\mathbf{x}$ . These relations are based on the Gaussian properties of the noise in the equations (1.5) and (1.13):

$$\begin{aligned}E[\Delta\mathbf{X}(t)] &= E[\mathbf{b}(\mathbf{X}(t))\Delta t + \sqrt{2D_e\Delta t}\boldsymbol{\eta}] = \mathbf{b}(\mathbf{X}(t))\Delta t \\ Cov[\Delta\mathbf{X}(t)] &= Cov[\mathbf{b}(\mathbf{X}(t))\Delta t + \sqrt{2D_e\Delta t}\boldsymbol{\eta}] = 2D_e\Delta t\mathbf{I},\end{aligned}\tag{1.19}$$

where we used the fact that  $E[\eta_k] = 1$  ( $k = 1, 2$ ) and  $Cov[\boldsymbol{\eta}] = \mathbf{I}$  with  $\mathbf{I}$  the identity matrix. One way to use the equations (1.18) for recovering the drift and diffusion coefficient from SPTs consists in computing the first and second order moments of the displacements. Obviously, it is not possible to apply these equations for every possible point of the experimental plane, instead one relies on spatial coarse-graining. The method proposed in [17] consists in decomposing the experimental plane into a regular square grid  $S_{\Delta x}$  with square bins  $S_{\Delta x}(\mathbf{c})$  centered at point  $\mathbf{c}$  and of size  $\Delta x$ . Applying this method, for an ensemble of  $N$  trajectories each of size  $M_i$ :  $\mathbf{X}_i(t_j) = [\mathbf{X}_i^{(1)}(t_j), \mathbf{X}_i^{(2)}(t_j)]$  ( $i = 1 \dots N$ ,  $j = 1 \dots M_i$ ,  $t_{j+1} - t_j = \Delta t$ ) the

equations (1.18) are rewritten as

$$\begin{aligned}\mathbf{b}^{(u)}(\mathbf{x}) &= \frac{1}{N(\mathbf{x})} \sum_{i=1}^N \sum_{(0 \leq j < M_i - 1, \mathbf{X}_i(t_j) \in S_{\Delta x}(\mathbf{x}))} \frac{\Delta \mathbf{X}^{(u)}(t_j)}{\Delta t} \\ \mathbf{D}^{(u,v)}(\mathbf{x}) &= \frac{1}{N(\mathbf{x})} \sum_{i=1}^N \sum_{(0 \leq j < M_i - 1, \mathbf{X}_i(t_j) \in S_{\Delta x}(\mathbf{x}))} \frac{\Delta \mathbf{X}^{(u)}(t_j) \Delta \mathbf{X}_i^{(v)}(t_j)}{2\Delta t},\end{aligned}\quad (1.20)$$

where  $N(\mathbf{x})$  is the number of points falling into the bin  $S_{\Delta x}(\mathbf{x})$ ,  $\Delta \mathbf{X}_i^{(u)}(t_j) = \mathbf{X}_i^{(u)}(t_{j+1}) - \mathbf{X}_i^{(u)}(t_j)$  is the displacement in dimension  $u = 1, 2$  and in an isotropic environment  $D(\mathbf{x}) \approx \frac{1}{2}(\mathbf{D}^{(u,u)}(\mathbf{x}) + \mathbf{D}^{(v,v)}(\mathbf{x}))$ . The drift field obtained in this manner is the result of the correlated motion of different particles and should not be used when only a portion of the molecules falling into the bin are subject to this force.

### Effect of localization noise

Molecules in SPTs experiments are never detected perfectly, but instead with some standard deviation called localization noise  $\sigma$ . The effect of the noise on the overdamped Langevin equation 1.13 can be modeled through the use of an extra process [79]

$$\Delta \mathbf{Y}(t) = \Delta \mathbf{X}(t) + \sigma \boldsymbol{\eta}', \quad (1.21)$$

where  $\boldsymbol{\eta}'$  is a Gaussian vector of center 0 and variance 1 (independent of the vector  $\boldsymbol{\eta}$  used for modeling the Brownian motion) and  $\mathbf{X}$  follows eq. (1.13). Computing the estimators for the local drift field and diffusion coefficients leads to [79]

$$\begin{aligned}\mathbf{b}_Y(\mathbf{x}) &= E \left[ \frac{\Delta \mathbf{Y}(t)}{\Delta t} \middle| \mathbf{Y}(t) = \mathbf{x} \right] = \mathbf{b}_X(\mathbf{x}) + o(\Delta t) + o(\sigma^2) \\ \mathbf{D}_Y(\mathbf{x}) &= E \left[ \frac{\Delta \mathbf{Y}(t)^T \Delta \mathbf{Y}(t)}{2\Delta t} \middle| \mathbf{Y}(t) = \mathbf{x} \right] = \mathbf{D}_X(\mathbf{x}) + \frac{\sigma^2}{\Delta t} + \frac{\sigma^2}{2} \nabla \cdot \mathbf{b}_X(\mathbf{x}) + o(\Delta t),\end{aligned}\quad (1.22)$$

with  $\mathbf{b}_{X/Y}$  and (resp.  $\mathbf{D}_{X/Y}$ ) the drift field (resp. diffusion tensor) computed on process  $X$  or  $Y$  and  $\nabla \cdot \mathbf{b}_X(\mathbf{x})$  the divergence of the local drift field. Thus both the drift vector field and diffusion coefficient are not affected by the localization noise to first order in  $\Delta t$ . Note that the corrections in  $\sigma^2$  in the estimation of  $D$  are hard to compute on real data as they involve the divergence of the drift field.

### 1.4.3 Maximum likelihood estimators

Finally, the last category of routinely used estimation methods are the ones based on Maximum Likelihood (ML) relying on the capacity to formulate the probability transition function  $p(\mathbf{X}(t) | \mathbf{X}'(t'))$  of the process: the probability of the molecule being at location  $\mathbf{X}$  at time

$t$  knowing that it was at location  $\mathbf{X}'$  at time  $t'$  ( $t > t'$ ). For a purely diffusive motion with coefficient  $D$ , the probability is given in two dimensions by

$$p(\mathbf{X}(t)|\mathbf{X}'(t')) = \frac{1}{4\pi D(t-t')} \exp \left[ -\frac{\|\mathbf{X}(t) - \mathbf{X}'(t')\|^2}{4D(t-t')} \right], \quad (1.23)$$

where  $\|\cdot\|$  is the Euclidean norm. This transition probability corresponds to a Gaussian distribution centered at  $\mathbf{X}'$  and of variance  $2D(t-t')$ . On the other hand, when the motion of the molecule follows an Ornstein-Uhlenbeck (model of the motion inside a parabolic potential well) process with attraction rate matrix  $\mathbf{\Lambda}$ , center  $\boldsymbol{\mu} = [\mu_1, \mu_2]$  and diffusion coefficient  $D$ , the transition function is obtained from the Fokker-Planck equation [80] as

$$p(\mathbf{X}(t)|\mathbf{X}'(t')) = \frac{\sqrt{|\mathbf{\Lambda}|}}{2\pi D \sqrt{|(1-e^{-2\mathbf{\Lambda}(t-t')})|}} \exp \left[ -(\mathbf{X}(t) - \boldsymbol{\mu} - \mathbf{r}) \frac{\mathbf{\Lambda}}{2D(1-e^{-2\mathbf{\Lambda}(t-t')})} (\mathbf{X}(t) - \boldsymbol{\mu} - \mathbf{r})^T \right], \quad (1.24)$$

where  $\mathbf{X}(t) = [X_1(t), X_2(t)]$ ,  $\mathbf{X}'(t') = [X'_1(t'), X'_2(t')]$ ,  $|\cdot|$  is the matrix determinant and  $\mathbf{r} = [(X'_1(t') - \mu_1)e^{-\lambda_1(t-t')}, (X'_2(t') - \mu_2)e^{-\lambda_2(t-t')}]$ .

Given the probability transition function  $p$ , the maximum likelihood for the successive displacements over a trajectory of size  $M$  composed of the points  $\mathbf{X}(t_1) \dots \mathbf{X}(t_M)$  following an equation of motion with parameter vector  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K]$  is obtained as

$$L(\boldsymbol{\theta}|\mathbf{X}) = p(\mathbf{X}(t_1)) \prod_{i=1}^{M-1} p(\mathbf{X}(t_{j+1})|\mathbf{X}(t_j)). \quad (1.25)$$

This equation can be extended for an ensemble of  $N$  trajectories  $\mathbf{X}_1 \dots \mathbf{X}_N$  of sizes  $M_i$  ( $i = 1 \dots N$ ), by using the fact that trajectories are independent

$$L(\boldsymbol{\theta}|\{\mathbf{X}_1 \dots \mathbf{X}_N\}) = \prod_{i=1}^N \left[ p(\mathbf{X}_i(t_1)) \prod_{j=1}^{M_i-1} p(\mathbf{X}_i(t_{j+1})|\mathbf{X}_i(t_j)) \right]. \quad (1.26)$$

Then the maximum likelihood estimator  $\hat{\theta}_k$  for the parameter  $\theta_k$  ( $k = 1 \dots K$ ) is the solution of the following equation

$$\frac{\partial}{\partial \theta_k} L(\boldsymbol{\theta}|\{\mathbf{X}_1 \dots \mathbf{X}_N\}) = 0. \quad (1.27)$$

For example, this procedure applied to the transition probability of the Brownian motion from equation (1.23) leads to the following estimator of the diffusion coefficient  $\hat{D}$  [81]

$$\hat{D} = \frac{1}{4C(t-t')} \sum_{i=1}^N \sum_{j=1}^{M_i-1} \|\mathbf{X}_i(t_{j+1}) - \mathbf{X}_i(t_j)\|^2, \quad (1.28)$$

where  $C = \sum_{i=1}^N (M_i - 1)$  is the total number of displacements. A close-form expression also exists for the parameters of the Ornstein-Uhlenbeck process and is given in sub-section 2.5.1.

## 1.5 Main results of the thesis

Early studies using SPTs considered only a coarse and global model of motion (often diffusion) [16] or were applied only when a large number of trajectories were available [17, 24]. New challenges consist in obtaining more precise measurements by detecting, separating and characterizing multiple modes of motion. Possibly in cases where proteins are present in low copy numbers.

### 1.5.1 Statistical approach to recover parabolic potential wells from SPTs

**Main results:** *I developed methods for automatically detecting and estimating parabolic potential wells parameters from trajectories, especially focusing on the detection of the well boundary and on cases where only a few numbers of trajectories are available for which I propose two methods:*

- **The first method** is based on the study of a truncated Ornstein-Uhlenbeck process representing the motion inside the well and consists in fitting the local point density inside the well, using a density filter to remove outliers (see sub-section 2.3.1) and recovering the boundary by looking at the point of discontinuity between the inside and outside density distributions (see paragraph 2.3.1). Here I distinguished between circular and elliptic boundaries, where the later requires to first compute the ratio of semi-axes lengths to define a new distance measure (see Fig. 2.3) for computing the density inside the well.
- **The second method**, is a hybrid iterative procedure (see paragraph 1.5.1) that uses the densities to recover the elliptic boundary and the local drift field for the estimation of the other well parameters (see subsection 1.5.1). The method starts by detecting the center of the well and restricts the analysis to the points in a square region around it, iteratively increasing the size of this square and finally selecting the well with the best score. This approach does not require a priori a size for the wells and improves the stability of the estimation by using two mostly independent sources of information: the density and the local drift.

*Overall, these approaches allow to automatically detect wells in any type of SPTs, can be applied to recover small wells ( $\approx 80\text{nm}$  radius) and fewer trajectories than before as presented in sub-section 2.3.4, chapter 3 and sub-section 5.0.3.*

#### Presentation of the methods

Parabolic potential wells observed in SPTs have the specificity that they occupy a bounded region (considered as elliptic) in space. Characterizing a potential well thus consists in

estimating seven parameters, two "dynamic" ones: the attraction strength of the well  $A$ , the local diffusion coefficient  $D$  and five geometric parameters related to its elliptic boundary  $\varepsilon = (\mu^{(1)}, \mu^{(2)}, a, b, \varphi)$  with  $\boldsymbol{\mu} = [\mu^{(1)}, \mu^{(2)}]$  the well center,  $a \geq b$  the semi-axes lengths and  $\varphi$  its orientation. In addition, these parameters influence each other as determining the boundary of the well determines which parts of trajectories are kept in the estimation of the dynamic parameters  $A$  and  $D$ . The standard method for extracting parabolic wells from SPTs was developed in [17]. It relies on the use of the magnitude of the local drift field to estimate  $A$  but considered only circular wells and did not give a clear method to extract the boundary. Another method was presented in [24] using a Bayesian framework for the estimation of the wells but relies heavily on numerical optimization and does not discuss the determination of the boundary. Lastly, maximum-likelihood estimation methods are also discussed in [79] but never applied to data.

First, I proposed a new method to estimate the parameters of potential wells based only on the local point density, I adapted the least square estimation (LSQE) method from [17] to work for elliptic wells and provided another method, also based on the density, to extract the well boundary. Then, I compared using Monte Carlo simulations, the estimated attraction coefficients  $A$  obtained with these methods with the maximum likelihood estimators [79, 82]. Finally, I proposed an alternative hybrid method based on both the point density and the local drift field.

### Estimation of potential wells from the local point density

On one hand, statistical physics tells us that the observed density of points  $\rho$  for a potential well is given by the Boltzman distribution

$$\rho(\mathbf{X}) = a \exp\left(\frac{-U(\mathbf{X})}{k_B T}\right), \quad (1.29)$$

where  $a$  is a constant depending on the local number of points and  $U$  is the energy function. On the other hand, the motion inside a potential well can be modeled as a two-dimensional truncated Ornstein-Uhlenbeck stochastic process

$$\dot{\mathbf{X}} = \boldsymbol{\Lambda}(\mathbf{X} - \boldsymbol{\mu}) + \sqrt{2D}\dot{\mathbf{w}}, \quad (1.30)$$

with  $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_x & 0 \\ 0 & \lambda_y \end{bmatrix}$  the attraction rate matrix. The steady-state distribution associated to the process (1.30) is given by taking the limit  $(t - t') \rightarrow +\infty$  in equation (1.24) (considering the initial point  $X'(t') = \boldsymbol{\mu}$ )

$$f(\mathbf{X}) = \frac{\sqrt{\lambda_x \lambda_y}}{2\pi D} \exp\left(-\frac{(\mathbf{X} - \boldsymbol{\mu})\boldsymbol{\Lambda}(\mathbf{X} - \boldsymbol{\mu})^T}{2D}\right), \quad (1.31)$$

corresponding to a Gaussian distribution with mean  $\boldsymbol{\mu}$  and covariance  $\frac{D}{\boldsymbol{\Lambda}}$ . Matching equations (1.29) and (1.31) and considering the well as spatially finite of elliptic shape, we obtain that the observed density  $\rho$  must follow a Gaussian distribution of mean  $\boldsymbol{\mu}$  and covariance

$$\mathbf{C} = \frac{D}{2A} \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}, \quad (1.32)$$

where  $A$  is the attraction rate coefficient and  $a, b$  are the coefficients defining the elliptic shape of the well. Thus, given the parameters  $a, b$  and  $D$  are known, then  $A$  can be recovered directly from the covariance matrix of the observed density distribution. The parameters  $a$  and  $b$  can be obtained from the same distribution and are in any case often estimated before  $A$ . The diffusion coefficient  $D$  can be either known *a priori*, estimated through formula (1.20), or through another estimator on another trajectory ensemble. This method has the advantage of not requiring any particle tracking at all to characterize potential wells but is very dependent on the estimation of the size of the boundary region  $a, b$ .

### Refinement to detect the well boundary

In practice, the local point density  $\rho$  can be estimated using a square grid of bins  $S_{\Delta x}(\mathbf{x})$  with bin size  $\Delta x$  and center  $\mathbf{x}$ . The density is then estimated inside each bin for an ensemble of  $N$  trajectories composed of  $M_i$  points each  $\mathbf{X}_i(t_j)$  ( $i = 1 \dots N, j = 1 \dots M_i$ ), we have

$$\rho(\mathbf{x}) = \sum_{i=1}^N \sum_{j=1}^{M_i} \mathbb{1}_{\mathbf{X}_i(t_j) \in S_{\Delta x}(\mathbf{x})}, \quad (1.33)$$

with  $\mathbb{1}_{a \in b} = \begin{cases} 1 & \text{if } a \in b \\ 0 & \text{otherwise} \end{cases}$ . In subsection 1.5.1 we know that if a potential well is present, it should generate a local density maximum and  $\rho$  should locally follow a Gaussian distribution. On the other hand, a Brownian motion, when sufficiently sampled, should generate a uniform density distribution. The problem now lies in determining which bins should be considered inside the well and included in the parameter estimation and which should be left out. As  $\rho$  is Gaussian, by removing the bins with a density below a certain threshold  $\alpha$ , given as a percentage of the local maximum, we can expect to remove parts of trajectories not falling inside the well. We thus define the sub-set of points falling into bins with density greater than  $\alpha$  as

$$\Gamma_\alpha = \{\mathbf{X}_i(t_j) \text{ such that } \mathbf{X}_i(t_j) \in S_{\Delta x}(\mathbf{x}) \text{ and } \rho(\mathbf{x}) > \alpha\}, \quad (1.34)$$

and the estimator for the well center and covariance matrix are computed only for the points in  $\Gamma_\alpha$ .

### Estimation of the well boundary from the local density

A way to estimate the well boundary without enforcing a threshold on the density is to find the point from which  $\rho$  deviates from a Gaussian distribution. Indeed, the well retaining trajectories, and allowing only a few to escape, there should be a difference between the internal and external density distributions. More specifically, I considered two types of surrounding densities based on observations: 1. when the well is surrounded by other particles (undergoing Brownian motion for simplicity) generating a uniform external distribution and 2. when most of the observed trajectories appear inside the well, generating a mostly null external distribution except for a few escaped trajectories.



At this stage of the estimation process, I consider the center to have already been estimated, such that I can rely on a different method than the square grid estimation (equation 1.33) to estimate  $\rho$  in order to prevent any count or geometry artifacts. Instead I rely on counting the number of points falling in concentric annulus of small radius  $r$ , width  $\Delta r$  and centered on  $\boldsymbol{\mu}$

$$\rho(r) = \sum_{i=1}^N \sum_{j=1}^{M_i} \mathbb{1}_{\|\mathbf{X}_i(t_j) - \boldsymbol{\mu}\| \in [r, r + \Delta r]}. \quad (1.35)$$

The first step is then to estimate the ratio of the ellipse semi-axes  $Cv = \frac{a}{b}$  giving an indication on the shape of the elliptic boundary. Using equation (1.32) we find that  $Cv(r)$  is given exactly by the square root of the ratio of the diagonal terms of the estimated covariance matrix  $\tilde{\mathbf{C}}(r)$  in the annulus with small radius  $r$

$$Cv(r) = \sqrt{\frac{\tilde{\mathbf{C}}_{11}(r)}{\tilde{\mathbf{C}}_{22}(r)}}. \quad (1.36)$$

For a circular well the curve  $Cv$  as a function of the distance to the center  $r$  is flat with  $Cv(r) \approx 1$ . For an ellipse however,  $Cv(r)$  should exhibit a maximum at a position  $r^*$  which value  $Cv(r^*)$  is used as the estimate for the semi-axes ratio. Using this ratio, we can define a new distance measure for the points  $\mathbf{X} = [\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]$  inside the well as

$$r' = \sqrt{(\mathbf{X}^{(1)})^2 + Cv(r^*)(\mathbf{X}^{(2)})^2}. \quad (1.37)$$

This new distance measure allows to reshapes the local point density from an ellipse to a circle and has no effect if the base density is already circular as in this case  $Cv(r^*) \approx 1$ . We then recompute the annulus density based on this new distance measure  $\rho(r')$  and use it to extract the well circular boundary.

Finally, the two external density situations defined above lead to different estimation procedures for the boundary: for a uniform external density (case 1. above), we can observe a minimum at position  $r'^*$ , near the boundary of the well, where the two different distributions overlap. For the case where the external density is null (case 2. above), there is no minimum but a small discontinuity in the density that we use as our estimate that can be more easily determined in log scale.

## Comparing various estimators to recover the strength of the potential well

In order to obtain realistic simulations for comparing the different well estimation methods, I use the following procedure. To simulate an individual trajectory, I first choose a boundary ellipse  $\varepsilon$ , attraction coefficients matrix  $\Lambda$  and diffusion coefficient  $D$  and then use the following equation based on the Euler scheme [76]:

$$\mathbf{X}(t + \delta t) = \mathbf{X}(t) - \Lambda(\mathbf{X}(t) - \boldsymbol{\mu})\delta t \mathbb{1}_{\mathbf{X}(t) \in \varepsilon} + \sqrt{2D\delta t}\boldsymbol{\eta} \quad (1.38)$$

where  $\mathbb{1}_{\mathbf{X}(t) \in \varepsilon}$  is 1 when  $\mathbf{X}(t)$  is inside the ellipse and 0 otherwise,  $\delta t$  is the simulation timestep ( $\approx 10^{-5}$ s) and  $\boldsymbol{\eta}$  is a vector of i.i.d Gaussian variables of zero mean and unit

variance. A trajectory is simulated up to a given time  $t_{\max}$  in order to obtain a specific amount of trajectory points (usually 20 points after sub-sampling). To recover an observed trajectory with timestep  $\Delta t = K\delta t$  (usually  $\Delta t \approx 10^{-3}\text{s}$ ), I sub-sampled the trajectory generated by equation (1.38) keeping one point every  $K$ . Finally, the initial point of each trajectory  $\mathbf{X}(0)$  is chosen following two different empirical scenarios (as discussed in the previous paragraph): 1)  $\mathbf{X}(0)$  is chosen uniformly distributed inside the well such that most of the trajectory points are located in the well or 2)  $\mathbf{X}(0)$  is chosen uniformly distributed in a square of width larger than the well in order to simulate a uniform density distribution outside the well. A simulation setup consists in an ensemble of trajectories where new trajectories are generated until reaching a specific amount of displacements inside the well.

Based on these simulations, I compared three different methods for obtaining the attraction coefficient of the well as a function of the density threshold  $\alpha$  (between 0 and 50%). The first method is the density estimation presented in sub-section 1.5.1 and thus only relies on the local point density estimate  $\rho$ . The second method uses a least square estimation (LSQE) of the local drift vector field with the one expected from a well (equation (1.8)), is an extension of the formulas presented by [17] to work with ellipses and relies on the computation of the drift vector field  $\mathbf{b}$ . Finally, the third method uses maximum likelihood estimation and only requires to know the well boundary. None of the methods, however, give an estimation of the well boundary, so in each case, I rely on the procedure given in sub-section 1.5.1. As presented in section 2.3, I found that all methods gave very accurate estimates for any tested value of  $\alpha$  both for circular and elliptic wells. The results for the estimation of the attraction coefficients are presented in sub-section 2.5.2 and show that the density method gave a correct value for a small  $\alpha$  but diverged afterward while the LSQ estimator seemed to always slightly underestimate (this underestimation increasing with increasing attraction strength), finally the MLE estimator was able to recover accurate estimates for all tested  $\alpha$ .

### Hybrid density-drift method for detecting potential wells

Based on the previous considerations about the local point density, I developed an hybrid method that uses the density to detect the boundary and the local drift field for obtaining the  $A$  parameter and scoring the well. The resulting procedure is as follows:

1. Form a square grid  $S_{\Delta x}$  with bins of size  $\Delta x$  over the plane covered by SPTs;
2. Compute the density map  $\rho_{\Delta x}(\mathbf{x})$  where we count the number of data points falling into each bin of from  $S_{\delta x}$ .
3. Keep only the highest  $d\%$  bins from  $S_{\Delta x}$  as possible potential well regions.
4. For each of these select bin centered at  $\mathbf{x}$ , do the following iterative procedure starting with  $N = 1$ :
  - (a) Compute the square grid  $S_{N,\Delta x,\mathbf{c}}$  centered at  $\mathbf{c}$  and composed of  $(2N+1) \times (2N+1)$  bins of width  $\Delta x$ .

- (b) Compute the boundary ellipse  $\varepsilon_N = [\mu_N^{(1)}, \mu_N^{(2)}, a_N, b_N, \varphi_N]$  with center  $\boldsymbol{\mu}_N = [\mu_N^{(1)}, \mu_N^{(2)}]$ , semi-axes lengths  $a_N \geq b_N$  and orientation  $\varphi_N$ , of the well by computing the 95% confidence matrix from the (considered) Gaussian local point density distribution in the square  $S_{N,\Delta x, \mathbf{c}}$ .
- (c) Compute the re-centered grid  $S_{N,\Delta x, \boldsymbol{\mu}_N}$  with center corresponding to the center of the estimated boundary ellipse.
- (d) Compute the local drift map  $\mathbf{b}_{N,\Delta x}(x) = [b_{N,\Delta x}^{(1)}(x), b_{N,\Delta x}^{(2)}(x)]$  for each bin of the grid  $S_{N,\Delta x, \boldsymbol{\mu}}$  using formula (1.20).
- (e) Estimate the attraction coefficient of  $A_N$  the well using the formula

$$A_N^* = \frac{\sum_{i=1}^M \frac{b_{N,\Delta x}^{(1)}(\mathbf{X}_i)x_i^{(1)}}{a^2} + \frac{b_{N,\Delta x}^{(2)}(\mathbf{X}_i)x_i^{(2)}}{b^2}}{2 \sum_{i=1}^M \frac{(x_i^{(1)})^2}{a^4} + \frac{(x_i^{(2)})^2}{b^4}}. \quad (1.39)$$

where  $\mathbf{X}_i = [x_i^{(1)}, x_i^{(2)}]$  ( $i = 1 \dots M$ ) are the centers of the  $M$  bins from  $S_{N,\Delta x, \boldsymbol{\mu}}$  that are contained inside the boundary  $\varepsilon_N$  of the well.

- (f) Estimate the parabolic index of the current well based on the residual least square error:

$$S(\mathbf{b}_{N,\Delta x}, A_N^*) = \frac{Err(\mathbf{b}_{N,\Delta x}, A_N^*)}{\sum_{i=1}^M \|\mathbf{b}_{N,\Delta x}(\mathbf{X}_i)\|^2} = 1 - \frac{\left( \sum_{i=1}^M \frac{b_{N,\Delta x}^{(1)}(\mathbf{X}_i)x_i^{(1)}}{a^2} + \frac{b_{N,\Delta x}^{(2)}(\mathbf{X}_i)x_i^{(2)}}{b^2} \right)^2}{\left( \sum_{i=1}^M \frac{(x_i^{(1)})^2}{a^4} + \frac{(x_i^{(2)})^2}{b^4} \right) \left( \sum_{i=1}^M \|\mathbf{b}_{N,\Delta x}(\mathbf{X}_i)\|^2 \right)} \quad (1.40)$$

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

- (g) If  $N < N_{\max}$ , set the initial center  $\mathbf{c} \leftarrow \boldsymbol{\mu}_N$ ,  $N \leftarrow N + 1$ , and return to step (a).
- (h) The final set of parameters  $(\varepsilon_{N^*}, A_{N^*}^*)$  is given by the iteration  $N^*$  with the smallest parabolic index:  $N^* = \arg \min_{N=1 \dots N_{\max}} S(\mathbf{b}_{N,\Delta x}, A_{N^*}^*)$ .
- (i) The final parameter set  $(\varepsilon_{N^*}, A_{N^*}^*)$  is considered to be a potential well when the index  $S(\mathbf{b}_{N,\Delta x}, A_{N^*}^*) < \frac{1}{2}$ .

The maximal iteration number  $N_{\max}$  can be determined by a locality criterion where we define the maximal allowed region in which to search a potential well  $d_{th}$  and use the relation  $N_{\max} = \arg \max_{N \in \mathbb{N}^+} (2N + 1)\Delta x < d_{th}$ .

The interest of this method is that it relies both on the local density through the estimation of the well boundary  $\varepsilon$  and on the local drift field for the estimation of the attraction coefficient  $A$ .

## 1.5.2 Main applications of the well detection method

**Main results:** *Using the methods described in chapter 2 and paragraph 1.5.1, I was able to recover hundreds of potential wells from different sets of  $Ca_V$  SPTs experiments (see section 2.3.4 and chapter 3) and found that:*

- *In  $Ca_V2.1$  wells had a size ( $\sqrt{ab}$ ,  $a, b$  semi-axes)  $\approx 80\text{nm}$  with potential energy barriers  $\approx 3kT$  that allow to stabilize the channels for  $\approx 100\text{ms}$ . A duration about three times larger than the time needed to escape the same region by pure diffusion (see sub-section 3.2.3).*
- *Colocalization of  $Ca_V2.1$  SPTs with  $\text{syntag1}::\text{CypHer5E}$ , a pre-synaptic marker, revealed that the wells are mostly located at pre-synaptic active zones (see Fig. 3.3).*
- *Modifying channel dynamics by light-induced cross-linking of the channels modified their spatial distribution, increasing channels density (from  $\approx 15$  to  $\approx 20$  wells/ $\mu\text{m}^2$ ) and induced spatial clusters of wells (see sub-section 3.2.4).*
- *By dividing the SPTs into successive time-windows, I was able to recover the dynamics of the wells themselves and found that they had a lifetime  $\approx 30\text{s}$  (see Figs. 3.3H and 3.11E-F), much longer than the residence time of the channels.*
- *The differential study of the dynamics of the long and short  $Ca_V2.1$  splice variants (Figs. 3.3D-H and 3.5) in conjunction with electrophysiological and calcium/glutamate imaging (Figs 3.2, 3.4 and 3.6) suggested that short-tailed  $Ca_V2.1$  channels have a more variable interaction distance with the scaffold proteins of the active zones than the long-tailed variant (see Fig. 3.6E).*

### Presynaptic nanodomains revealed in SPTs of $Ca_V$ VGCCs

The transmission of an electrical current across a chemical synapse from pre to postsynaptic neurons relies on the release of neurotransmitters from the presynaptic terminal. These neurotransmitters are contained inside Synaptic Vesicles (SV) docked at the presynaptic membrane's active zone and their fusion is triggered by the arrival of calcium ions near the membrane-SV junction point. The ions penetrate inside the cell through Voltage-Gated Calcium Channels (VGCC) that open upon sensing an action potential. The coupling between VGCCs and SVs has been found to be of prime importance for the correct functioning of synapses [83] and can also act as a way of implementing plasticity at the pre-synaptic level. In this study, conducted in collaboration with the group of Martin Heine at TU Magdeburg, I analyzed trajectories of two isoforms of the main VGCC at hippocampal synapses:  $Ca_V2.1_{+47}$  possessing a long C-terminal tail and its shorter variant  $Ca_V2.1_{\Delta 47}$ . It is known that  $Ca_V2.1$  channels can interact with scaffold proteins at the active zone and we wanted to characterize these interactions and how they differed between the two isoforms. We also wanted to further understand how the channel dynamics influences the electrical properties of the neurons (by modifying channel dynamics through cross-linking) and in turn how it is modified by the electrical properties of the neurons (*e.g.* by silencing the network activity).

## Detection of potential wells in $\text{Ca}_V2.2$ and GPI-anchored GFP trajectories

First, I compared the results of the three different well estimation methods density, Least Square Estimation (MLE) and Maximum Likelihood Estimation (MLE) presented in chapter 2 (and discussed in the previous sub-section) to two empirical SPTs datasets. The first dataset consisted in SPTs of the  $\text{Ca}_V2.2$  voltage-gated channels at the surface of hippocampal neurons that were acquired with a timestep  $\Delta t = 33\text{ms}$ . There, I was able to detect 353 potential wells from  $\approx 280000$  total trajectories from 13 experiments with average semi-axes lengths of  $\langle a \rangle \approx 90\mu\text{m}$  and  $\langle b \rangle \approx 70\mu\text{m}$  and average attraction coefficients  $\langle A \rangle \approx 0.13$  for the density estimation,  $\approx 0.07$  for LSQE and  $\approx 0.18$  for MLE. The second dataset consisted in SPTs of GPI-anchored GFP also at the surface of hippocampal neurons and were obtained with  $\Delta t = 20\text{ms}$ . There I detected 181 wells from  $\approx 310000$  total trajectories from 10 experiments with average semi-axes lengths of  $\langle a \rangle \approx 0.15\mu\text{m}$  and  $\langle b \rangle \approx 0.12\mu\text{m}$  and attraction coefficients  $\langle A \rangle \approx 0.35$  with the density method,  $\approx 0.14$  with the LSQE and  $\approx 0.32$  with the MLE.

## $\text{Ca}_V2.1$ channels are trapped in potential wells at synapses

Now looking at another dataset composed of SPTs from the two variants:  $\text{Ca}_V2.1_{\Delta 47}$  (short-tailed) and  $\text{Ca}_V2.1_{+47}$  (long-tailed) of the  $\text{Ca}_V2.1$  channels and co-localizing these SPTs with the syntag1::CypHer5E pre-synaptic marker, I could detect potential wells at these regions with an average of 2 wells per synapse (see Fig. 1.5A). Three examples of wells appearing at pre-synaptic terminal are presented in Fig. 1.5B. The size of the wells was  $r \approx 80\text{nm}$  with  $r = \sqrt{ab}$  ( $a, b$  the well semi-axes lengths) with a median of  $r = 76\text{nm}$  and  $r = 81\text{nm}$  for  $\text{Ca}_V2.1_{\Delta 47}$  and  $\text{Ca}_V2.1_{+47}$  respectively which is in agreement with other models of pre-synaptic arrangement such as [84]. Computing the residence time of the channels inside the wells I found average residence times of  $99\text{ms}$  and  $105\text{ms}$  for  $\text{Ca}_V2.1_{\Delta 47}$  and  $\text{Ca}_V2.1_{+47}$  respectively. Although apparently low, these values are more than three times larger than the average time required for a channel to escape a region of  $80\text{nm}$  only by diffusion ( $\approx 30\text{ms}$ ) (see sub-section 1.3.3).

To see how the wells were affected by the network activity, SPTs were acquired after chemically blocking either ionotropic glutamate receptors (using APV and CNQX chemicals) or sodium channels (using TTX). Both treatments globally reduced the diffusion coefficients both at the presynaptic terminals and outside and tended to slightly strengthen confinement by decreasing the well sizes and increasing the residence times, this effect being more pronounced for the  $\text{Ca}_V2.1_{+47}$  isoform.

## Duration of the potential wells

The SPTs data revealed that the wells could move slightly during the experiments. This motion does not correspond to "technical" drift as the displacements between wells appearing at the same time are not correlated. This motion blurs the characteristics of the wells such as the center and local drift field and thus requires the use of a sliding time window over the

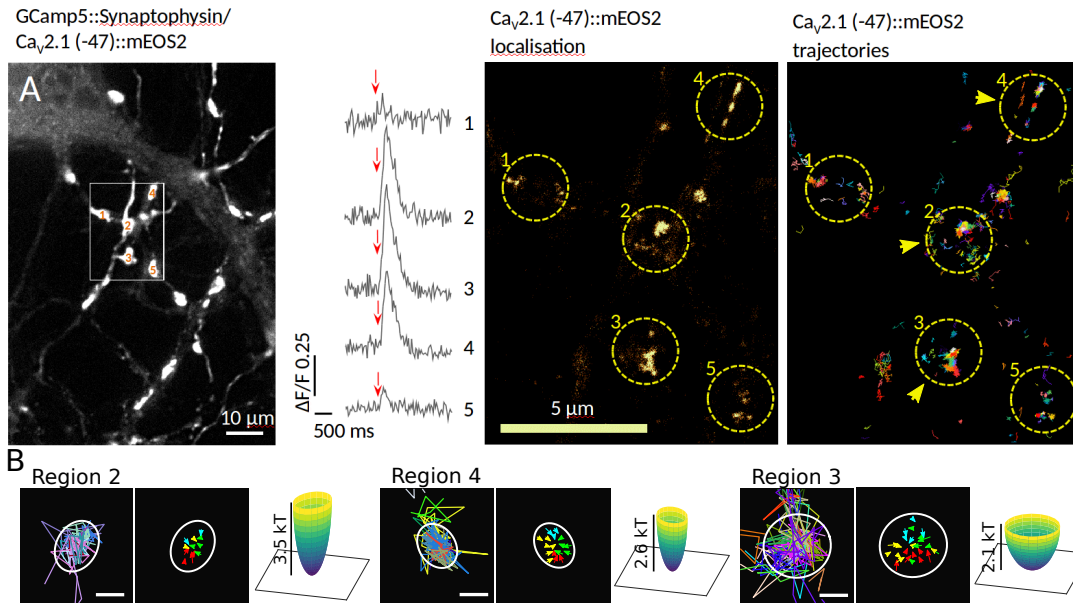


Figure 1.5: **Correlation of potential wells and functional synapses.** **A.** Colocalization between  $\text{Ca}_V2.1$  channels and synaptic vesicles located through synaptophysin and presynaptic calcium transients as reported by changes in GCaMP5G fluorescence after an electrical stimulation (left). Corresponding  $\text{Ca}_V2.1$  fluorescence signal and reconstructed trajectories (right). **B.** Three potential wells located in the functional regions from A.

SPTs. In this new setup, a 20s sliding time-window is passed over the ensemble of SPTs and potential wells are detected independently for each time window. By having a 50% window overlap and by linking together closeby wells in successive windows. In addition I computed that the wells lifetime for both  $\text{Ca}_V2.1$  isoforms was in average  $\approx 30s$  with a few wells lasting  $\approx 60s$ .

### Cross-linking $\text{Ca}_V2.1$ channels leads to clustered wells

In order to understand the effect of the channels mobility on synaptic transmission, I looked at SPTs of both  $\text{Ca}_V2.1$  variants obtained for channels fused at their N-terminal domains with a mutant of the Arabidopsis flavoprotein cryptochrome 2 (CRY2). These CRY2: $\text{Ca}_V2.1$  channels form transient cross-links with other CRY2: $\text{Ca}_V2.1$  channels upon being illuminated with a blue light. The SPTs revealed that the diffusion coefficient outside presynaptic terminals was greatly decreased compared to non-cross linked channels for both isoforms: from  $D = 0.095$  for  $\text{Ca}_V2.1_{\Delta 47}$  to  $0.063\mu\text{m}^2/\text{s}$  for CRY2: $\text{Ca}_V2.1_{\Delta 47}$  and from  $D = 0.091$  for  $\text{Ca}_V2.1_{+47}$  to  $0.070\mu\text{m}^2/\text{s}$  for CRY2: $\text{Ca}_V2.1_{+47}$ . In turn the diffusion coefficients inside the wells, sizes and residence times coefficients were only slightly modified on the different splice variants. Finally, although the biophysical characteristics of the wells were not modified, I observed many more potential wells, the well density increasing from 15 wells/ $\mu\text{m}^2$  in  $\text{Ca}_V2.1$  to 20 wells/ $\mu\text{m}^2$  for CRY2: $\text{Ca}_V2.1$  channels with the wells spatially distributed in clusters supporting the fact that more  $\text{Ca}_V2.1$  channels be integrated in active zones (Fig 1.6A).

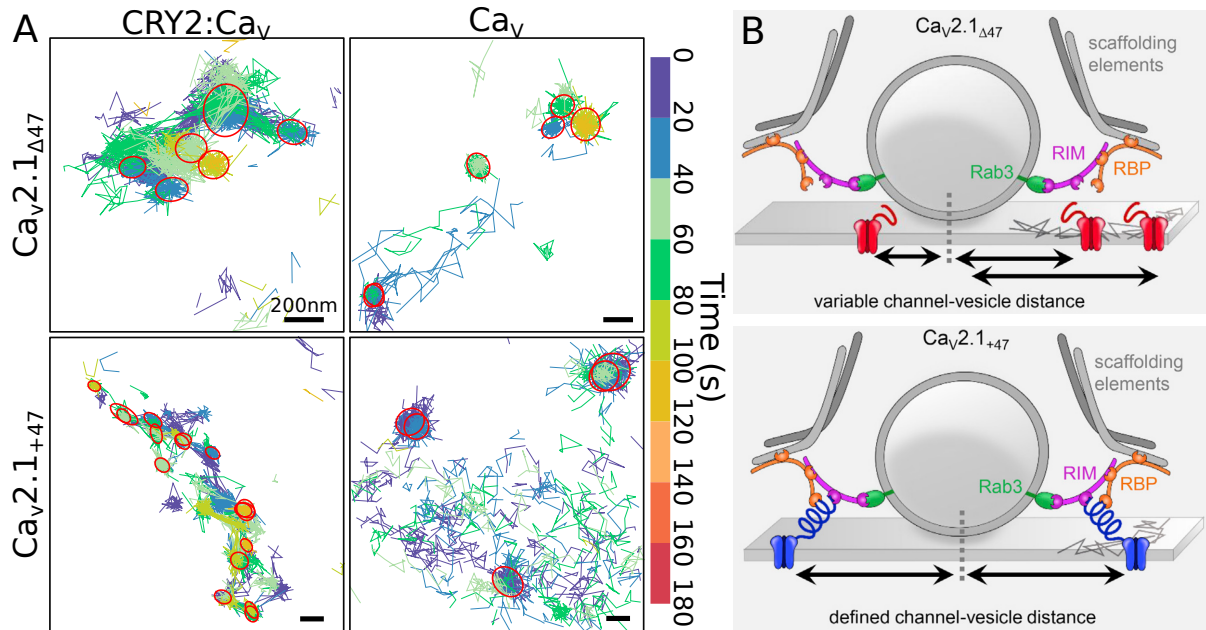


Figure 1.6: **Spatial effect of cross-linking and proposed model of channels organization.** **A.** Sample experimental areas showing the increased well density and spatial clustering of wells after light-induced cross-linking (left column) compared to wild type Ca<sub>v</sub>2.1 channels. **B.** Proposed organization of the two Ca<sub>v</sub>2.1 splice variants at the active zone.

### Main conclusion of the SPT analysis on the function of the active zones

In this study, SPTs are only one of the multiple tools used to better characterize the dynamics of the two Ca<sub>v</sub>2.1 channels variants at presynaptic terminals. The other tools comprise calcium (GCaMP5G::synaptophysin) and glutamate (iGluSnFr) imaging to monitor synapse activity, immunocytochemical staining for co-visualizing multiple protein species, confocal imaging colocalization and electrophysiological recordings. These different results suggest that calcium channels mobility has the ability to tune the synaptic vesicle release probability in response to previous activity. In addition, we proposed that the amount of integration of the different splice variants of Ca<sub>v</sub>2.1 channels contributes to the variability of the release probability and short-term plasticity of individual synapses from different neurons or even along the axon of individual neurons. Indeed, we propose the model that Ca<sub>v</sub>2.1<sub>Δ47</sub> variant lacking C-terminal scaffold interactions, form stochastic coupling with SVs primarily caused by their surface mobility. However, Ca<sub>v</sub>2.1<sub>+47</sub> channels are localized in a transient but defined proximity to SVs.(Fig. 1.6B).

Altogether, these data suggest that alternative splicing can serve as an endogenous mechanism to shape release probability and short-term plasticity of individual synapses, as shown for homeostatic plasticity [85]. We propose that stochastic mobility-driven and use-dependent rearrangement of calcium channels constitute the initial processing of information at the presynaptic site and depends on the flexible association between channels and SVs.

### 1.5.3 Application of SPTs analysis to reveal the flow in the endoplasmic reticulum lumen

**Main results:** *Analyzing for the first time SPTs from Endoplasmic Reticulum (ER) luminal proteins, I found that:*

- *They exhibited a very heterogeneous pattern of motion with high-density regions connected by high-velocity jumps colocalizing respectively with the junctions and tubules of the peripheral ER membrane (see Fig 4.2a,b,d).*
- *The apparent diffusion coefficient (obtained by considering the motion as purely diffusive) was  $\approx 1.1\mu\text{m}^2/\text{s}$  in agreement with FRAP experiments (see paragraph 4.4.4). Using the peculiar pattern formed by SPTs, I developed a procedure to recover the ER network (see sub-section 4.4.3). Then, I decomposed the motion depending on the position in the network into confined diffusion at junctions with  $D \approx 0.2\mu\text{m}^2/\text{s}$  (see paragraph 4.4.4) and high-velocity jumps in tubules with normally distributed amplitudes  $\approx 23\mu\text{m}/\text{s}$  (see paragraph 4.4.4).*
- *This pattern of motion was not found in SPTs of ER membrane proteins (see Fig. 4.3) but was present in the ER lumens from multiple cell types: HEK293, Cos-7, SH-SY5Y (see Table 4.1 and Fig. 4.8) and could not be attributed to the network motion (see sub-section 4.4.9 and Fig. 4.9) suggesting that this flow is a property of the lumen.*
- *After reconstructing the network in the form of a graph, I found that the molecules could visit the entire network (see paragraph 4.4.5). Besides, there was a similar amount of trajectories entering and exiting the junctions suggesting that the network flow is at equilibrium (see sub-section 4.4.5). Finally, I found that trajectories exhibit transient unidirectional periods in the crossing of tubules (see sub-section 4.4.6).*
- *Finally, we proposed a nano-peristalsis model where the flow is locally generated by the transverse contractions of tubules (see sub-section 4.5.2 Fig. 4.5h) providing an efficient mechanism for the homogenization of the ER luminal content across long distances.*

In this study, conducted in collaboration with Edward Avezov and David Ron at the University of Cambridge, we sought to understand the motion of molecules inside the Endoplasmic Reticulum (ER). The ER is an organelle that extends from the nucleus up to the plasma membrane. It is composed of different regions, more or less developed depending on the type of cell, the rough ER close to the nucleus and the smooth peripheral ER near the membrane. The latter has a peculiar shape composed of thin tubes interconnected into three-way junctions or sheets. The main ER functions in cells consist of calcium storage, protein maturation thanks to chaperone proteins and protein addressing. Thus the precise spatio-temporal availability of the ER-resident proteins is crucial for the correct functioning of the cell.

The dynamics of the lumen content have been studied in the 2000s through Fluorescence Recovery After Photobleaching (FRAP) experiments [21] based on population measurements of the recovery of fluorescence after photo-bleaching (see sub-section 1.1.2). In our study, we could address this question for the first time at the single molecule level.



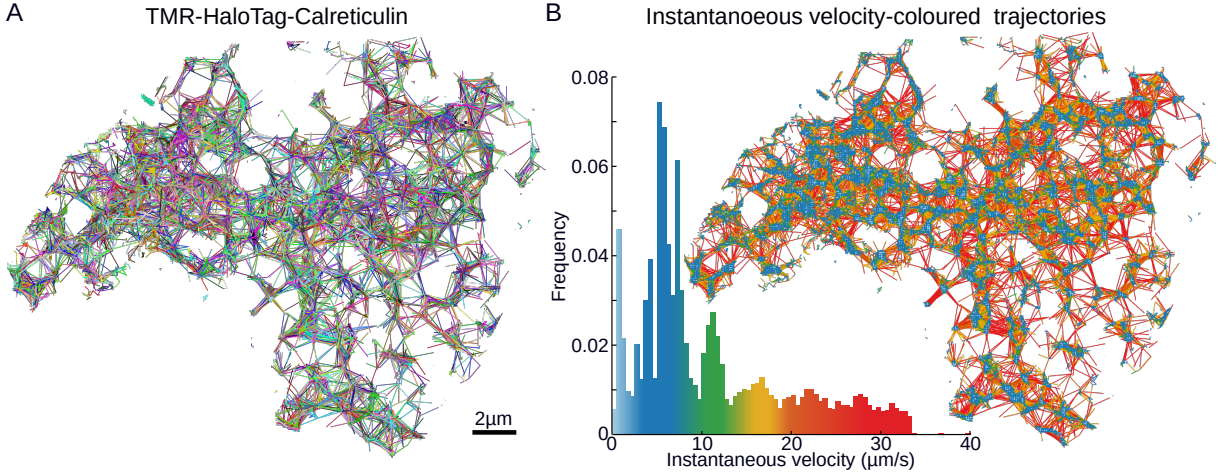


Figure 1.7: **SPTs from ER resident protein calreticulin tagged with TMR dye.** **A.** Individual calreticulin trajectories displaying the characteristic peripheral ER shape. **B.** Instantaneous-velocity color-coded trajectories where each displacement is colored as a function of the corresponding instantaneous velocity as presented in the inset histogram.

### Recovering the ER network from SPTs

SPTs from the ER resident protein calreticulin tagged with HaloTag-TMR depicted the well known peripheral ER shape of pipes and junctions (Fig. 1.7A). In addition, trajectories also exhibited a specific velocity patterns, revealed by the distribution the instantaneous velocities computed for each displacement  $\Delta\mathbf{X}(t_j) = \mathbf{X}(t_{j+1}) - \mathbf{X}(t_j)$  as

$$v(\Delta\mathbf{X}(t_j)) = \frac{\|\Delta\mathbf{X}(t_j)\|}{t_{j+1} - t_j} \quad (1.41)$$

with  $\|\cdot\|$  the Euclidean norm. Spatially mapping where the different velocity regimes appear as in Fig. 1.7B revealed an accumulation of low-velocities at junctions while high-velocities jumps form tubules.

From these observations, I developed a method to construct a graph of the ER network from SPTs, such as the one presented in Fig. 1.8. In this graph, nodes correspond to junctions or sheets and links to the tubules. For an ensemble of  $N$  two-dimensional trajectories of size  $M_i$ ,  $\mathbf{X}_i(t_j)$  ( $i = 1 \dots N$ ,  $j = 1 \dots M_i$ ), the procedure is summarized as follows:

1. Extract the sub-set of  $l$  points participating in low-velocity displacements:

$$l = \{\mathbf{X}_i(t_j) | v(\Delta\mathbf{X}_i(t_j)) < v_{th} \text{ or } j > 1 \text{ and } v(\Delta\mathbf{X}_i(t_{j-1})) < v_{th}\} \quad (1.42)$$

2. Apply the DBSCAN, density-based clustering algorithm (or alternatively another one) to find the clusters present in  $l$ .
3. Approximate the boundary of each cluster by fitting an ellipse around its points and merge any overlapping ellipses. The resulting ensemble of  $K$  ellipses  $E = \{\varepsilon_1 \dots \varepsilon_K\}$  forms the nodes of the graph.

4. Finally, form the connectivity matrix  $\mathbf{C}$  of size  $K \times K$  that counts the number of displacements between each pair of nodes

$$c_{u,v} = \sum_{i=1}^N \sum_{j=1}^{M_i-1} \mathbb{1}_{u,v}(X_i(t_j), X_i(t_{j+1})) + \sum_{i=1}^N \sum_{j=1}^{M_i-2} \mathbb{1}'_{u,v}(X_i(t_j), X_i(t_{j+2})) \quad (1.43)$$

$$\text{with } \mathbb{1}_{u,v}(\mathbf{X}(t_j), \mathbf{X}(t_{j+1})) = \begin{cases} 1 & \text{if } \mathbf{X}(t_j) \in \varepsilon_u \text{ and } \mathbf{X}(t_{j+1}) \in \varepsilon_v \text{ and} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{1}'_{u,v}(\mathbf{X}(t_j), \mathbf{X}(t_{j+1})) = \begin{cases} 1 & \text{if } \mathbf{X}(t_j) \in \varepsilon_u \text{ and } \forall \varepsilon \in E, \mathbf{X}(t_{j+1}) \notin \varepsilon \text{ and } \mathbf{X}(t_{j+2}) \in \varepsilon_v \\ 0 & \text{otherwise} \end{cases}$$

Where we count both direct jumps between the two nodes (left sum) but also allow three-points jumps if the middle point does not belong to any other node (right sum). This latter part was added to be more robust against artifacts and usually contributes only marginally to the matrix.

The final result is the graph defined by the ensemble of (elliptic) nodes  $E$  and the connectivity matrix  $\mathbf{C}$ .

### Using the reconstructed ER network to extract SPTs characteristics

With this particular dynamics, very few points fall in tubules rendering impossible their study with the local estimators (see sub-section 1.4.2). Furthermore, looking at the displacements along individual trajectories revealed that the duration and time interval between high-velocity jumps are both stochastic, following exponential distributions (Fig. 4.4d). Thus the motion is not homogeneous in time nor in space, preventing the use of MSD estimation techniques (see sub-section 1.4.1).

Instead, I developed a method to characterize the dynamics based on the graph constructed in the previous sub-section. First, I characterized the dynamics inside the nodes as a confined diffusion and estimated the diffusion coefficient  $D_{\text{node}}$  using the local estimators from sub-section 1.4.2. Then, I characterized the velocity in tubules by averaging the instantaneous velocities of all the displacements between pairs of connected nodes

$$v_{\text{flow}}(u, v) = E \left[ \frac{\|\mathbf{X}(t_{j+1}) - \mathbf{X}_{t_j}\|}{t_{j+1} - t_j} \mid \begin{array}{l} \mathbf{X}(t_j) \in \varepsilon_u \text{ and } \mathbf{X}(t_{j+1}) \in \varepsilon_v \text{ or} \\ \mathbf{X}(t_j) \in \varepsilon_v \text{ and } \mathbf{X}(t_{j+1}) \in \varepsilon_u \end{array} \right], \quad (1.44)$$

with  $\varepsilon_u, \varepsilon_v$  two nodes,  $\|\cdot\|$  the Euclidean norm and  $E[\cdot]$  the expectation. I found that the distribution of the amplitudes  $v_{\text{flow}}$  for each pair connected nodes is Gaussian. The overall equation of motion is thus given by

$$\dot{\mathbf{X}}(t) = \begin{cases} \mathbf{J} & \text{if the molecule goes through a tubule} \\ \sqrt{2D_{\text{node}}} \dot{\mathbf{w}} & \text{in a node,} \end{cases} \quad (1.45)$$

where  $\mathbf{J}$  is a vector of norm  $\|\mathbf{J}\| \sim \mathbf{N}(\mu_{\text{tubule}}, \sigma_{\text{tubule}})$  and random orientation. For the cell presented in Fig. 1.7, I obtained  $D_{\text{node}} = 0.19 \mu\text{m}^2/\text{s}$ ,  $\mu_{\text{tubule}} = 22.9 \mu\text{m}/\text{s}$  and  $\sigma_{\text{tubule}} = 6.9 \mu\text{m}/\text{s}$  and the parameters for the other cell types are given in Table 4.2.

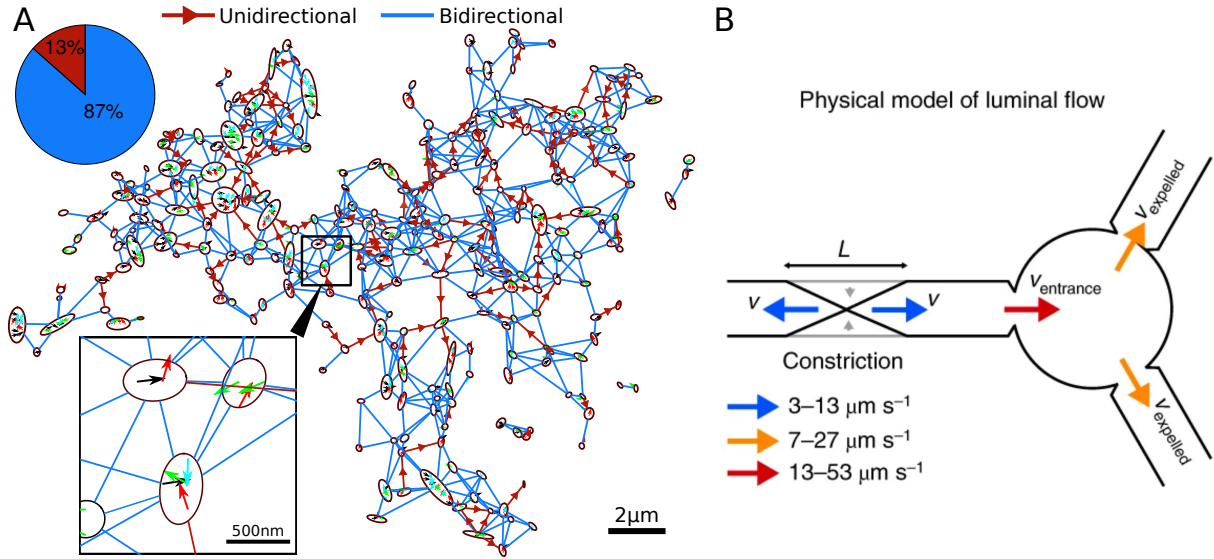


Figure 1.8: **Reconstructed ER graph and nanoperistalsis mechanism.** **A.** ER graph reconstructed from SPTs where the tubule direction was determined with a 75% co-directionality threshold. Inset box presents the local drift dynamics at junctions. **B.** Schematic drawing of the proposed nano-peristalsis mechanism generating the luminal flow.

### Derivation of a jump-diffusion model for ER luminal motion

In order to find whether the proposed jump-diffusion model better explained the observed motion than a purely diffusive model, I looked at which model of motion best approximate the distribution of instantaneous velocities. I recover the the diffusion coefficient associated to an apparent purely diffusive motion by constructing a square grid over the experimental plane, computing the diffusion coefficient in each square (forming a diffusion map, see subsection 1.4.2) and taking the average of the distribution of these diffusion coefficients named  $D_{\text{app}}$ . This coefficient mixes both confined and high-velocity displacements. The expected distribution of instantaneous velocities for a purely diffusive model is obtained from the transition probability of the Brownian motion (equation (1.23)) as

$$Pr \left\{ \frac{\|\Delta \mathbf{X}\|}{\Delta t} = u \right\} = \frac{u}{\sigma^2} \exp \left( -\frac{u^2}{2\sigma^2} \right), \quad (1.46)$$

with  $\sigma^2 = \frac{2D_{\text{app}}}{\Delta t}$ . The distribution of displacement for the "jump-diffusion" process given in equation (1.45) requires the introduction of a new parameter  $\kappa \in [0, 1]$  the proportion of confined to jumping displacements

$$Pr \left\{ \frac{\|\Delta \mathbf{X}(t)\|}{\Delta t} = u \right\} = \kappa Pr \left\{ \frac{\|\Delta \mathbf{X}(t)\|}{\Delta t} = u \mid \text{tubule} \right\} + (1 - \kappa) Pr \left\{ \frac{\|\Delta \mathbf{X}(t)\|}{\Delta t} = u \mid \text{node} \right\}. \quad (1.47)$$

The distribution of jump displacements in tubule was computed in the previous paragraph and approximated as a Gaussian  $\mathcal{N}(\mu_{\text{tubule}}, \sigma_{\text{tubule}})$  while the distribution of displacements

in the nodes is given by equation 1.46 with  $\sigma^2 = \frac{2D_{\text{node}}}{\Delta t}$  and the coefficient  $\kappa$  is a free parameter. Finally, using a two-ways Kolmogorov-Smirnov statistical test, I found that the jump-diffusion model better explains the observed instantaneous velocity distribution than the purely diffusive model (see 4.2c).

To test that the observed "jumps" are a property of the ER lumen, I applied the same procedure from SPTs of calnexin (cnx), an ER transmembrane protein. There, the distribution of instantaneous velocities revealed the absence of jumps, the density map did not contain local high-density regions and the diffusion map was more widespread over the experimental plane, preventing the reconstruction of the network with the method presented above and indicating an absence of confinement. Finally, the apparent diffusion coefficient obtained from the diffusion map is in average  $D_{\text{app, cnx}} = 0.42 \pm 0.09 \mu\text{m}^2/\text{s}$  much lower than the one obtained for the luminal calreticulin  $D_{\text{app, crt}} = 1.13 \pm 0.51 \mu\text{m}^2/\text{s}$  as expected from trans-membrane proteins.

### Network properties of the luminal flow

The reconstructed graph of the ER can also be used to understand the properties of the ER network. The static orientation of a tubule can be obtained by setting a threshold to the ratio of displacements observed in a given direction (see Fig 1.8A). From this oriented graph, I computed the strongly connected components and found only one main component, indicating that the molecules have the potential to visit the entire network, in agreement with previous findings that the ER does not form isolated sub-compartments (by successive photo-bleaching, FLIP) [86]. Finally, by looking at the temporal distribution of the direction of the displacements in each tubule, I also found that individual tubules exhibit transient unidirectionality periods lasting in average  $\approx 3.89\text{s}$ . This number however is surely an overestimate of the true duration due to the scarcity of the observation of displacements in the tubules, a rough correction model estimates these periods to last  $\approx 38\text{ms}$  (see subsection 4.4.7)

### Proposing a nanoperistalsis mechanism for generating the ER luminal flow

Finally, the mechanistic assumption that leads to the realization of this work is that local tubule contraction would generate the observed flow, a sort of nano-peristalsis mechanism (Fig 1.8B). Unfortunately, it is not possible at the moment to track both membrane and luminal proteins at the same time. In addition, observing fast contractions in the nanometer-thin tubules requires a microscope with a high resolution and fast acquisition time. To observe these contractions, we used fast-SIM technology allowing a  $90\text{ms}$  acquisition time and a  $32\text{nm}$  pixel-size (barely enough to visualize them) where multiple contractions on different tubules could be observed and collected. I found that an average contraction duration  $\approx 213 \pm 169\text{ms}$  with a time interval for successive contractions on the same tubule  $\approx 980 \pm 995\text{ms}$  and a contraction length  $\approx 140 \pm 50\text{nm}$ . Using these characteristics together with the morphology of an ER junction, we constructed a small mechanistic model to quantify the flow that could be created by these contractions and found that the order of magnitude of

the flow velocity predicted by the model ( $3-13 \mu\text{m/s}$ ) was coherent with the "jump" velocity recovered from SPTs ( $\approx 23 \mu\text{m/s}$ ).

### 1.5.4 Applications of SPTs analysis to lysosome, drosophila neuromuscular junctions and NuRD complex

**Main results:** *I present here three other applications of the SPT analysis in different cellular contexts:*

1. *Studying SPTs of lysosomes I found that their motion formed a network where nodes correspond to high-density regions and links to jumps between these regions. Interestingly, this network followed the peripheral ER structure suggesting that lysosomes interact with the ER network (see sub-section 5.0.1).*
2. *In the second application, I looked at SPTs of  $\text{Ca}_v2.1$  channels at the neuromuscular junctions (NMJ) of *Drosophila* and found wells with stronger potential energy barriers  $\approx 5kT$  (compared to  $\approx 3kT$  for hippocampal synapses) and longer lifetimes  $\approx 127s$  (compared to  $\approx 30s$ ). All in all these results suggest a more stable organization of active zones at neuromuscular junctions compared to hippocampal synapses (see sub-section 5.0.3).*
3. *The last application is about SPTs of the nuclear proteins composing the NuRD chromatin remodeling complex where I found that the motion of these proteins could be best explained by a model of motion composed of two or three diffusion states. These dynamics states are related to different physical states of the NuRD complex: confined motion when it is fully assembled, slowly moving due to chromatin motion when it is partially formed and moving fast when the molecule is not on the chromatin (see sub-section 5.0.5).*

This chapter shows applications of the present SPTs analysis methods to study trafficking in cell biology. I present extra methods based on the ones developed in previous chapters and apply them in different cellular contexts.

### ER-lysosomes interactions participate in maintaining the peripheral ER shape

The endoplasmic reticulum is a very dynamics organelle that constantly undergoes remodeling of its structure in its peripheral region through the formation and disappearance of tubules and sheets. In particular, the formation of new tubules is driven in part by their attachment to the molecular motors sliding on microtubules. Another organelle that has been shown to interact with both microtubules and the ER membrane are lysosomes that are cytoplasmic vesicles involved in many cellular processes such as sensing and waste disposal. Based on our previous study about ER luminal flow (see chapter 4), we searched in lysosome SPTs for signs of their interactions with the ER.

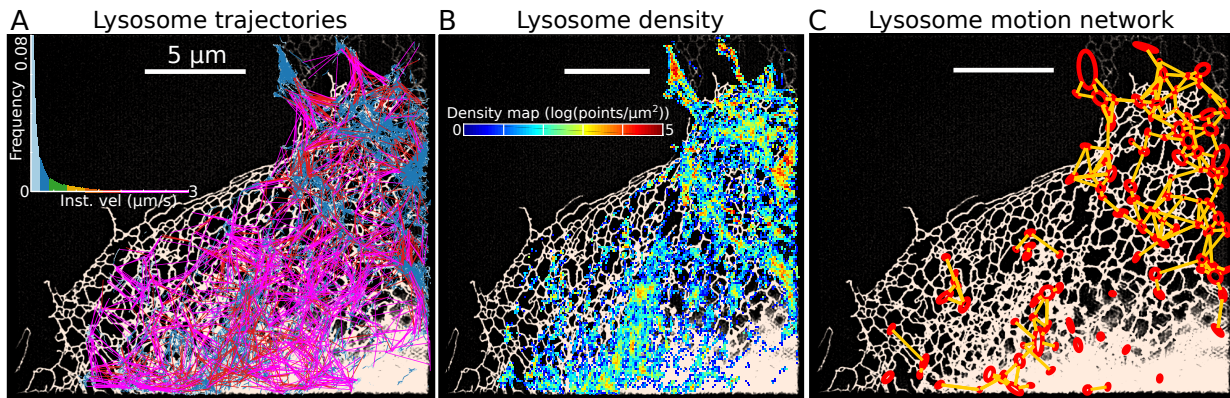


Figure 1.9: **Colocalization of lysosome SPTs with ER network.** **A.** Lysosome SPTs color-coded for each displacement depending on its instantaneous velocity (only displaying slower and faster displacements) as given by the inset distribution. **B.** Density map built from lysosome SPTs overlaid on top of the ER network. **C.** Reconstructed lysosome motion network overlaid on top of the ER network.

We found that lysosome motion followed three observable patterns (Fig. 1.9A): 1. individual trajectories can remain confined for  $\approx 1$ min in small regions of size  $\approx 200$ nm, 2. Ensemble of trajectories aggregate in larger high-density regions of size  $\approx 400$ nm (Fig. 1.9B) and 3. trajectories can exhibit large jumps with an average instantaneous velocity  $\approx 1\mu\text{m/s}$  connecting high-density regions. Furthermore, applying a similar method than for luminal ER proteins, I was able to reconstruct a lysosomal network graph where nodes correspond to high-density regions and links to high-velocity jumps between these regions (Fig. 1.9C). It is worth noting however that the luminal and lysosome modes of motion are quite different as in the first case, the node-tubule topology is associated with a diffusion-drift dynamics, while for lysosomes the motion is linked to transient interactions with the ER and motion along tubules.

Together with other biological experiments, this study suggests that lysosomes, with their sensing capabilities, could serve to reshape the ER helping the cell to adapt to its environment.

### Potential wells found in $\text{Ca}_v2.1$ nanodomains at *Drosophila* neuro-muscular junction

Extending the study of the motion of  $\text{Ca}_v2.1$  channels in hippocampal neurons (see chapter 3) I studied SPTs of these channels at *drosophila*'s Neuro-Muscular Junction (NMJ). Contrary to hippocampal neurons where  $\text{Ca}_v2.1$  trajectories were found uniformly over the axons, most of the trajectories at NMJ were located in high-density regions that follow the organization of the junctions. Some of these regions exhibited patterns of locally converging drift vector field with increasing amplitude with the distance to the center, coherent with the presence of potential wells. These wells had similar sizes ( $\approx 80$ nm) but lower diffusion coefficients inside ( $\approx 0.02$  vs  $\approx 0.04\mu\text{m}^2/\text{s}$ ) and higher energies ( $\approx 5$  vs  $\approx 3kT$ ) and residence times ( $\approx 500$  vs  $\approx 100$ ms) compared to hippocampal wells. Finally, I evaluated the duration of the detected

well structures in time by splitting the SPTs into successive 20s time windows and found that the average duration is  $\approx 120s$  much larger than the  $\approx 30s$  for the hippocampal wells.

These results suggest that potential wells at the NMJ are much more stable than the wells found at hippocampal synapses and are only slightly affected by the amount of extracellular calcium or the addition of philanthotoxin for increasing the number of calcium channels at NMJs.

### **Multiple modes of motions detected in SPTs of proteins from the NuRD nuclear complex**

For this study, I was given multiple SPTs from the molecules composing the Nucleosome Remodeling Deacetylase (NuRD) complex residing in the nucleus and having an ATP-dependent chromatin remodeling activity [87]. This complex is formed by the association of seven proteins: HDAC1, HDAC2, RbAp46, RbAp48, MTA1/2/3, MBD2/3, CHD2/4 [88]. As the NuRD complex is composed of many molecules, we wanted to understand the state of the complex under normal conditions and when some of its constituents are missing. The difficulty lies in the fact that we do not know in the SPTs whether the imaged molecule is alone or in a complex. As the trajectories are sparse, I relied on the analysis of their distribution of displacements and tried to distinguish multiple modes of motions in it. To this end, I used a Markov-Chain model based on either 2 or three states, distinguished by different diffusion coefficients and a null model of a single diffusion. After fitting the different models to the instantaneous velocity distribution, I used the Bayesian Information Criterion (BIC) to find the one that best explain the data (taking into account the number of parameters). I found that a three states diffusion model explained best the motion of CHD4, MBD3, CHD4-MDBD3 (deleted MBD3), MTA2 and MTA2-MBD3 molecules, while a two states model was best to explain the motion of CHD4+ATPase and CHD4-HDAC conditions.

## **1.6 General conclusion on the statistical analysis of SPTs**

The statistical analysis methods for SPTs are very diverse and must be adapted to the dynamics suggested by the dataset at hand. In any case, local analysis methods, as presented in the thesis (see chapter 2), where we start by selecting regions of interests and then isolate and characterize each region, are easier to use than global or static methods. The statistical analysis of empirical SPTs has proven very useful to the fields of cellular biology and neurosciences as exemplified by the two applications of the thesis: characterizing the presence of potential wells at active zones of hippocampal synapses used to retain  $Ca_v2.1$  channels in these regions close to synaptic vesicles (chapter 3) and the discovery of a flow in the peripheric endoplasmic reticulum lumen that has significant implications in calcium homogeneization throughout the ER network (chapter 4). SPT experiments have revealed the great heterogeneity of membrane structure and the way forward is now to understand the functional implications of these heterogeneities.

## Chapter 2

# Reconstructing potential wells of high density regions from super-resolution single particle trajectories

Published as:

Pierre Parutto, Jenifer Heck, Martin Heine, and David Holcman, "*Biophysics of high density nanometer regions extracted from super-resolution single particle trajectories: application to voltage-gated calcium channels and phospholipids*", Scientific Reports, 9(1), 1-14 (2019).

**Abstract** Large amount of super-resolution single particle trajectories has revealed that the cellular environment is enriched in heterogenous regions of high density, which remain unexplained. The biophysical properties of these regions are characterized by a drift and their extension (a basin of attraction) that can be estimated from an ensemble of trajectories. We develop here two statistical methods to recover the dynamics and local potential wells (field of force and boundary) using as a model a truncated Ornstein-Uhlenbeck process. The first method uses the empirical distribution of points, which differs inside and outside the potential well, while the second focuses on recovering the drift field. Finally, we apply these two methods to voltage-gated calcium channels and phospholipids moving on the surface of neuronal cells and recover the energy and size of these high density regions with nanometer precision.



## 2.1 Introduction

Single Particle trajectories (SPTs) obtained from super-resolution techniques such as spt-PALM or UPaint summarize the history of large amount of particles that can be cytoplasmic molecules, membrane receptors or channels in live cells. Over the past decade, statistical methods based on stochastic models have been developed to segment [89, 90], interpret and extract relevant biophysical parameters such as flows and arrival time statistics between various subregions [91, 92, 78, 93, 94, 95] from these large data sets. The most striking and universal characteristic of these trajectories is that they are not homogeneously distributed in cells, but rather are concentrated in sub-regions, a phenomenon that is not fully understood: what are these high-densities regions? What are the underlying physical forces that restrict and confine trajectories? For example, AMPA receptors that traffic on the surface of neuronal cells accumulate specifically at the post-synaptic density (PSD) of synapses, where they are needed for proper synaptic transmission [96, 97]. Similarly, at the pre-synaptic terminal, voltage-gated calcium channels ( $\text{Ca}_v$ ) can accumulate on membrane subregions, with a size of hundreds of nanometers [98]. Retaining these channels guarantee that calcium ions can remain near vesicles to trigger release.

A possible mechanism to retain trajectories is a field of force caused by the presence of an extended potential well. These structures have been detected in a size of hundreds of nanometers [91, 99, 98]. However, the physical origin of these wells remains unclear because the length of classical electrostatic interactions is ten time shorter [100] than the observed wells sizes. These high-density regions are characterized by several features: (1) a converging field of force, whether or not it is the gradient of a potential energy, (2) an energy depth and (3) a boundary. Finding and estimating these geometrical characteristics from trajectories and their statistical distribution remain challenging especially at tens of nanometers below the diffraction limit of light.

Here, we present two methods to detect and reconstruct potential wells from high-density regions contained in SPTs. The first approach is based on estimating the density of points of a truncated Ornstein-Uhlenbeck process (which accounts for a motion driven by a converging force and diffusion). We recover the center of the well, the covariance matrix and the boundary. While the second approach is based on estimating the local drift vector field. We insist that the first approach will clearly reveal the peak of aggregation, while the strength of the second method is its ability to extract a field of force. This field confirms the underlying deterministic structure that maintains the random trajectories together. We will first validate both approaches on stochastic simulations and then apply them to characterize nanodomains appearing in voltage-gated calcium channels ( $\text{Ca}_v2.2$ ) and lipid anchored GFP (GPI-GFP) trajectories obtained from sptPALM or UPaint experiments.

## 2.2 Methods

### 2.2.1 Coarse-grained description of stochastic trajectories

In the Smoluchowski's limit of the Langevin equation [101, 102], the position  $\mathbf{X}(t)$  of a stochastic molecule at time  $t$  can be described by

$$\dot{\mathbf{X}} = \frac{\mathbf{F}(\mathbf{X}(t), t)}{\gamma} + \sqrt{2D}\dot{\mathbf{w}} \quad (2.1)$$

where  $\mathbf{F}$  is a field of force,  $\mathbf{w}$  is a white noise and  $\gamma$  is the friction coefficient [101] and  $D$  is the diffusion coefficient. The source of the noise is the thermal agitation of the ambient lipids and membrane molecules. However, due to the timescale of acquisition of trajectories, which is in general too low to follow the thermal fluctuations, rapid events are not resolved in data, and at this spatiotemporal scale, the motion can be coarse-grained as a stochastic process [91, 103]

$$\dot{\mathbf{X}} = \mathbf{b}(\mathbf{X}) + 2\sqrt{\mathbf{B}(\mathbf{X})}\dot{\mathbf{w}} \quad (2.2)$$

where  $\mathbf{b}$  is the drift field and  $\mathbf{B}(\mathbf{X})$  the position-dependent diffusion matrix. The effective diffusion tensor is given by  $D(\mathbf{X}) = \frac{1}{2}\mathbf{B}(\mathbf{X})\mathbf{B}^T(\mathbf{X})$ , ( $\cdot^T$  denotes the transposition) [104, 101]. The diffusion tensor accounts for impenetrable obstacles of various sizes. Note that the interpretation at the physical level of the stochastic equation (2.42) is from the Ito's sense and not Stratanovich or any other sense, because a physical process has to be non-anticipating [104] (the future cannot interfere with the past).

### 2.2.2 Potential wells characteristics

The drift field  $\mathbf{a}(\mathbf{X})$  in equation (2.42) may represent a field force acting on the diffusing particle, that could be due to a potential well [100]. When the diffusion tensor  $D(\mathbf{X})$  is locally constant and the coarse-grained drift field  $\mathbf{b}(\mathbf{X})$  is a gradient of a potential

$$\mathbf{b}(\mathbf{X}) = -\nabla U(\mathbf{X}), \quad (2.3)$$

then the density of particles is given locally by the Boltzmann distribution [105]

$$\rho(\mathbf{X}) = N_0 e^{-U(\mathbf{X})/D}, \quad (2.4)$$

where  $N_0$  is a normalization constant. An infinite paraboloid potential well with an elliptic base has the analytical representation for  $\mathbf{X} = (x, y)$

$$U(x, y) = A \left[ \left( \frac{x - \mu_x}{a} \right)^2 + \left( \frac{y - \mu_y}{b} \right)^2 \right], \quad (2.5)$$

where the center is  $\boldsymbol{\mu} = (\mu_x, \mu_y)$ ,  $A$  is the field amplitude and  $a, b$  are the lengths of the large and small semi-axes of the ellipse. To account for a finite well, we restricted the influence of the well to the region

$$\Gamma_\epsilon = \{(x, y) | U(x, y) \leq \epsilon\}. \quad (2.6)$$

The truncated energy function  $U$  associated to such parabolic potential well is

$$U(\mathbf{X}) = \begin{cases} A \left[ \left( \frac{x - \mu_x}{a} \right)^2 + \left( \frac{y - \mu_y}{b} \right)^2 \right], & \text{if } \mathbf{X} \in \Gamma_\epsilon \\ 0 & \text{otherwise} \end{cases},$$

from which the drift field is the gradient of the energy, is given by

$$\nabla U(\mathbf{X}) = -2A \begin{bmatrix} \frac{x - \mu_x}{a^2} \\ \frac{y - \mu_y}{b^2} \end{bmatrix}. \quad (2.7)$$

The goal of these section is to recover, from empirical single particle trajectories that consists of few successive points acquired with a sampling time  $\Delta t$ , the center  $\boldsymbol{\mu} = (\mu_x, \mu_y)$ , the amplitude  $A$  and the size of each semi-axis  $a, b$  for the boundary  $\epsilon$ .

### 2.2.3 Simulations of stochastic trajectories

To validate our methods, we first generated synthetic single particle trajectories from the stochastic process

$$\dot{\mathbf{X}} = -\nabla U(\mathbf{X}) + \sqrt{2D}\dot{\mathbf{w}}, \quad (2.8)$$

where the potential  $U$  is defined in equation (2.5) (as presented in Fig. 2.1A),  $D$  is the diffusion coefficient and  $\mathbf{w}$  is a white noise. To reproduce observed trajectories, we keep a fixed lapse time  $\Delta t$  between successive points and generated  $N$  trajectories  $(\mathbf{X}_1(0), \dots, \mathbf{X}_N(K\Delta t))$  containing  $K$  points ( $K = 20$ ), using the classical Euler's scheme (Fig. 2.1B).

We consider two types of numerical simulations depending whether the initial points  $X_i(0)$  are uniformly distributed (1) inside the well or (2) inside a square box surrounding the well. This uniform distribution represents the random activation of fluorophores by a laser (Fig. 2.11B). To guarantee a constant number of points inside the wells across multiple simulations, we did not fix the number  $N$  of trajectories but instead generate new trajectories until a certain quantity of displacements has happened inside the well. This resetting procedure generates a distribution of points which depends on the initial uniform distribution. However, in the limit of large  $N$ , the distribution of points converges toward the steady-state, which is Gaussian inside the well and uniform outside, when trajectories are restricted to a large square domain.

### 2.2.4 Estimators for the elliptic boundary geometry

To identify parts of trajectories inside the well, we use the level line ensemble of the density distribution

$$\Gamma_\alpha = \{\mathbf{X} \text{ such that } \rho_e(\mathbf{X}) > \alpha\} \quad (2.9)$$

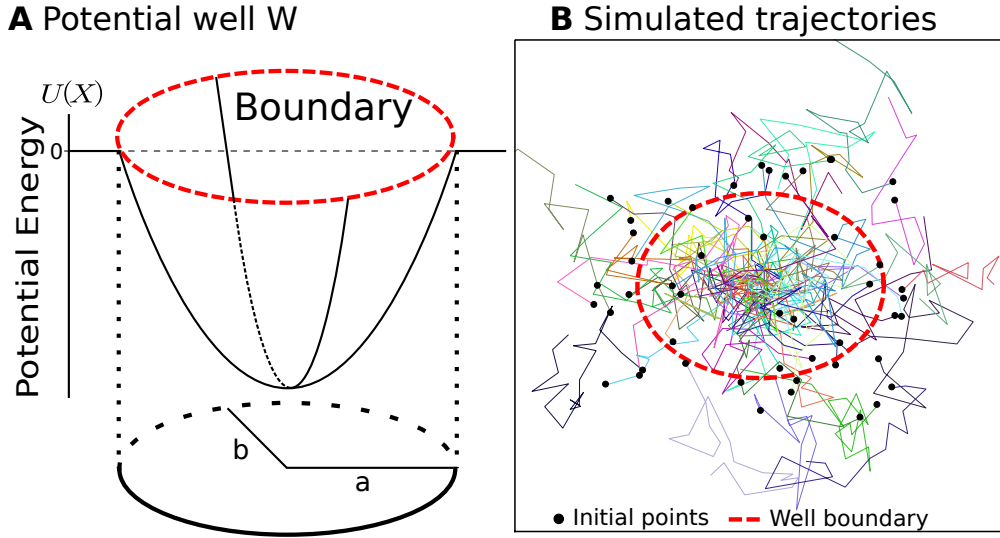


Figure 2.1: **Numerical simulation scheme.** **A.** Model of a truncated potential well with two axes  $a, b$  and energy  $U(\mathbf{X})$  with a boundary. **B.** Trajectories generated using equation (2.67) where the initial points (black dots) can either be located inside or outside the boundary of the well (dashed red). Parameters:  $D = 0.042 \mu\text{m}^2/\text{s}$ ,  $\lambda_x = 10$ ,  $\lambda_y = 17.78$ .

where  $\rho_e$  is the empirical point density, estimated over the bins of the square grid constructed from the ensemble of trajectories (Fig. 2.2B). The ensemble  $\Gamma_\alpha$  contains all trajectory points falling into a bin, with a density greater than the density threshold  $\alpha$ .

To recover the center of the distribution, we consider all points  $\mathbf{X}_i = (x_i, y_i)$  located in  $\Gamma_\alpha$  (Fig. 2.2C) and use the empirical estimators

$$\mu_x^{(\alpha)} = \frac{1}{N_p} \sum_{k=1, \mathbf{X}_k \in \Gamma_\alpha}^{N_p} x_k, \quad \mu_y^{(\alpha)} = \frac{1}{N_p} \sum_{k=1, \mathbf{X}_k \in \Gamma_\alpha}^{N_p} y_k, \quad (2.10)$$

where  $N_p$  is the number of points in the ensemble  $\Gamma_\alpha$ . To estimate the covariance two-by-two matrix  $C^{(\alpha)}$ , defined as

$$U(\mathbf{X}) = (\mathbf{X} - \boldsymbol{\mu}^{(\alpha)})^T \mathbf{C}^{(\alpha)} (\mathbf{X} - \boldsymbol{\mu}^{(\alpha)}), \quad (2.11)$$

we use the empirical estimators

$$C_{ij}^{(\alpha)} = \frac{1}{N_p - 1} \sum_{k=1, \mathbf{X}_k \in \Gamma_\alpha}^{N_p} X_{i,k} X_{j,k}, \quad (2.12)$$

where  $\mathbf{X}_{i,k}$  is the  $i$ th coordinates of  $\mathbf{X}_k$  (Fig. 2.2C).

## 2.2.5 Improved drift estimation

We recall briefly here (see SI) that a correction term has to be added in order to recover an Ornstein-Uhlenbeck process of parameter  $\lambda$  and centered at  $\boldsymbol{\mu}$  (equation (2.67)): we derived

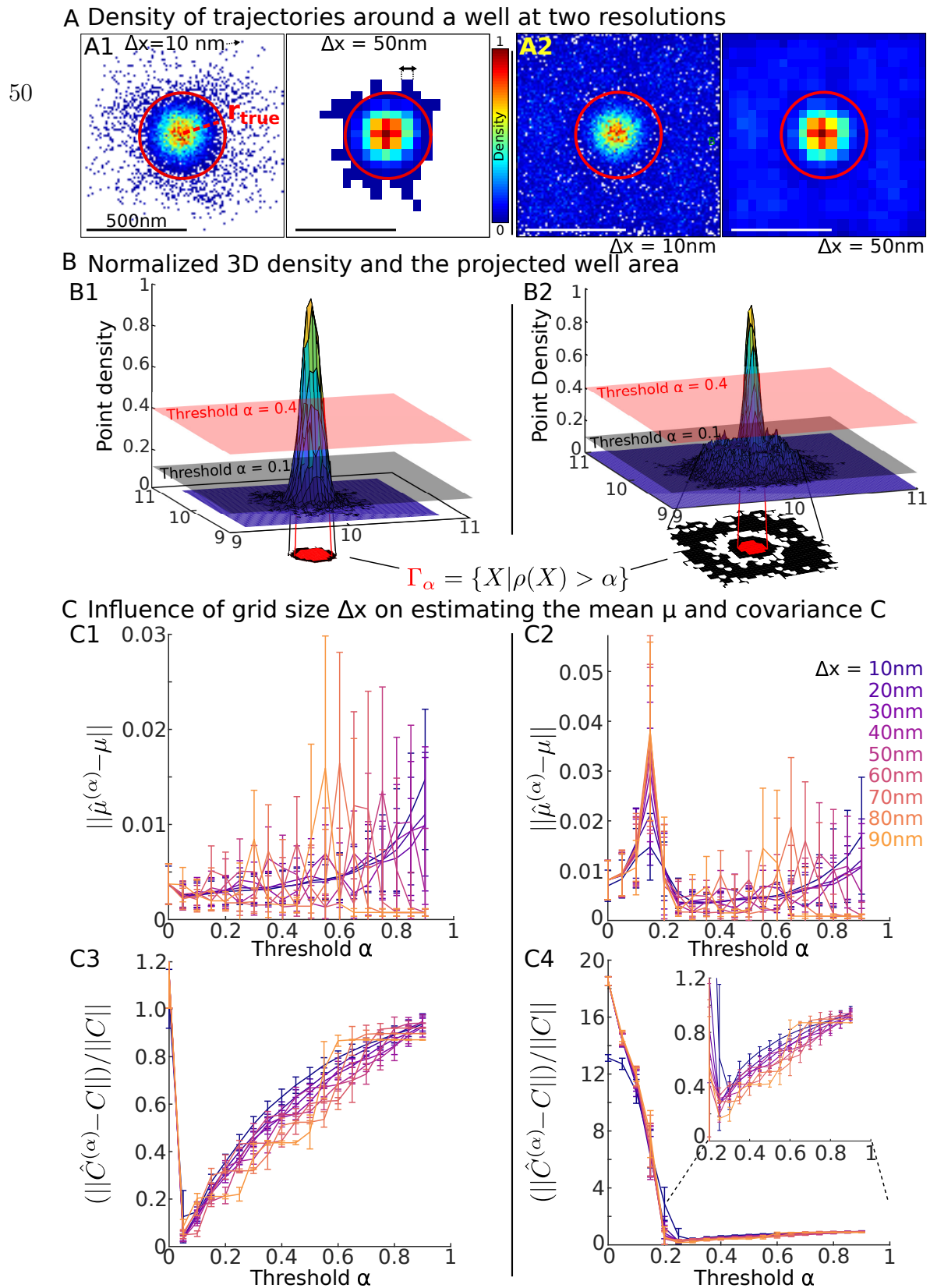


Figure 2.2: **Recovering a truncated potential well from the density of points.** **A.** Density maps (in  $\log(\text{points})/\mu\text{m}^2$ ) for two different grid sizes  $\Delta x = 10$  (left) and  $50$  nm (right) when the initial points are located inside the well A1 or uniformly distributed in a square of size  $1 \mu\text{m}$  A2. **B.** Normalized three-dimensional empirical density function  $\rho$  obtained from A. We plotted the ensemble  $\Gamma_\alpha = \{X | \rho(X) > \alpha\}$  for  $\alpha = 0.1$  (black) and  $\alpha = 0.4$  (pink) and the projected area (red) in the well in the two cases (B1,B2) associated to (A1,A2) respectively. **C.** Influence of the grid size  $\Delta x$  and threshold  $\alpha$  on the well characteristics estimations. (C1,C3) (resp. C2,C4) panels are obtained by computing with the initial distribution described in A1 (resp. A2).

in the SI that the drift term at position  $\mathbf{X}$  and at resolution  $\Delta t$

$$\tilde{\mathbf{b}}_{\Delta t}(\mathbf{X}) = \frac{1 - e^{-\lambda\Delta t}}{\Delta t}(\mathbf{X} - \boldsymbol{\mu}). \quad (2.13)$$

Hence, the first order moment at resolution  $\Delta t$  computed from the displacement  $\mathbf{X}(t + \Delta t) - \mathbf{X}(t)$  from SPTs deviates from the expected drift. When  $\lambda\Delta t$  is small, a first order Taylor expansion leads to the approximation

$$\tilde{\mathbf{b}}_{\Delta t} = \underline{\mathbf{X}}(1 - \frac{1}{2}\lambda\Delta t) + o(\lambda^2\Delta t), \quad (2.14)$$

and hence to recover the drift, we have to use the correction factor  $1 + \frac{1}{2}\lambda\Delta t$  on the estimated drift.

## 2.2.6 Processing of Ca<sub>v</sub>2.2 and GPI SPTs

For the experiments related to Ca<sub>v</sub>2.2 data, we refer to [32], while the experimental procedure of GPI-GFP data have been described for other molecules in [16]. We will first isolate trajectories in non-overlapping time windows of 20 s and apply the following procedure to each window. We will construct a square grid with bins size  $\Delta x$  around trajectories and collect the 5% highest density bins. For each of these selected bins, we will detect wells as follows: we will first use 90% of the local point density (threshold  $\alpha = 0.1$ ) to detect the center of the well from equations (2.10), then we will apply the procedure described in paragraph 2.3.1 (elliptic case), restraining the computation of the semi-axes ratio to a maximum distance from the center  $r_{cov} = 150$  nm and using a threshold  $T_\rho = 35\%$  on the density of points for determining  $\hat{r}_{e,0}$ . Once the center and semi-axes of the well are found, the diffusion coefficient will be determined using equations (??), estimated for all displacements with an initial points falling inside the well.

## 2.3 Results

### 2.3.1 Recovering a bounded potential well from the point density of trajectories

We first reconstruct the characteristics of the potential wells from the distribution of trajectories. This approach ignores the temporal causality between successive points and relies on a truncated paraboloid model. We will first recover the center and covariance matrix of the steady-state density distribution using a square grid (Fig. 2.2A). We recall that inside a well given by equation (2.5), this is a Boltzmann distribution

$$\rho(\mathbf{X}) = N_0 \exp \left\{ -A \frac{\left[ \left( \frac{x - \mu_x}{a} \right)^2 + \left( \frac{y - \mu_y}{b} \right)^2 \right]}{D} \right\} \quad (2.15)$$

where  $N_0$  is a normalization coefficient while the other parameters are defined in subsection 2.2.2. Based on this distribution, we estimated the center using equations (2.10) and the covariance matrix from equation (2.12).

The accuracies of these estimators are analyzed by plotting the errors between the true and the estimated centers  $\|\boldsymbol{\mu}^{(\alpha)} - \boldsymbol{\mu}\|$  and between the covariance matrices  $\|\mathbf{C}^{(\alpha)} - \mathbf{C}\|$  (quadratic norm of the matrix) versus the parameter  $\alpha$ , which represents the threshold of level line (equation (2.17), Fig. 2.2B,C) and various grid sizes (from  $\Delta x = 10$  to 90 nm). When  $\alpha$  decreases from one to zero and the initial points are located inside the well, the iterative sequences of positions of the estimated centers converge to the true value and the fluctuations (SD computed over 100 realizations) decreases with  $\alpha$  (Fig. 2.2C1). However, when the initial points of the simulated trajectories were also chosen outside the well, we found that there was an optimal threshold value  $\alpha \approx 0.3$  for which the error in the estimated and true centers is minimum (Fig. 2.2C2). Below this value, points of the trajectories falling outside the well are also contained in the ensemble  $\Gamma_\alpha$ , thus contaminating the error of the estimation. When the initial points fall inside the well only (Fig. 2.2C3), the ensemble  $\Gamma_0$  contains external trajectories that perturb the estimation of the covariance matrix  $\mathbf{C}(\alpha)$ . However, as  $\alpha$  increases, these external points disappear from  $\Gamma_\alpha$  and the error becomes minimal at the value  $\alpha_{\text{opt}} = 0.05$ . When  $\alpha$  continues to increase, the estimators become less accurate. However, when the initial points are chosen also outside the well, the error starts by decreasing because trajectories that are not inside the well affects the estimation (Fig. 2.2C4). As  $\alpha$  increases, the estimator converges toward an optimal value  $\alpha_{\text{opt}} = 0.25$  (75% of the points are used), which minimizes the matrix error. When  $\alpha$  continues to increase, the error increases slowly (Fig. 2.2C4-inset), similar to the case of Fig. 2.2C3.

To conclude, depending whether or not trajectories are falling inside the well or could also escape the high-density regions, the statistical estimators give different results: using as many points as possible increases the estimate of the center, but not necessarily of the covariance matrix.

## Estimating the boundary of the well

None of the estimators described above can be used to reconstruct the location of the well boundary. We now present a method to recover first a circular and then an elliptic boundary in two cases: when the initial points falls only inside the well and when they can also fall outside. The first step consists in discriminating between a circular and an elliptical boundary. To do so, we computed from the matrix (2.12), the covariance ratio

$$Cv(r) = \sqrt{\frac{\mathbf{C}_{1,1}(r)}{\mathbf{C}_{2,2}(r)}} \quad (2.16)$$

estimated over the trajectories located inside the annulus  $(r, r+\Delta r)$  (Fig. 2.3A1). To compute  $Cv(r)$  (Fig. 2.3A1,A2,B1,B2), we recall that the diagonal form of covariance matrix can be found from equations (2.15) and (2.11)

$$\mathbf{C} = \frac{D}{A} \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}.$$

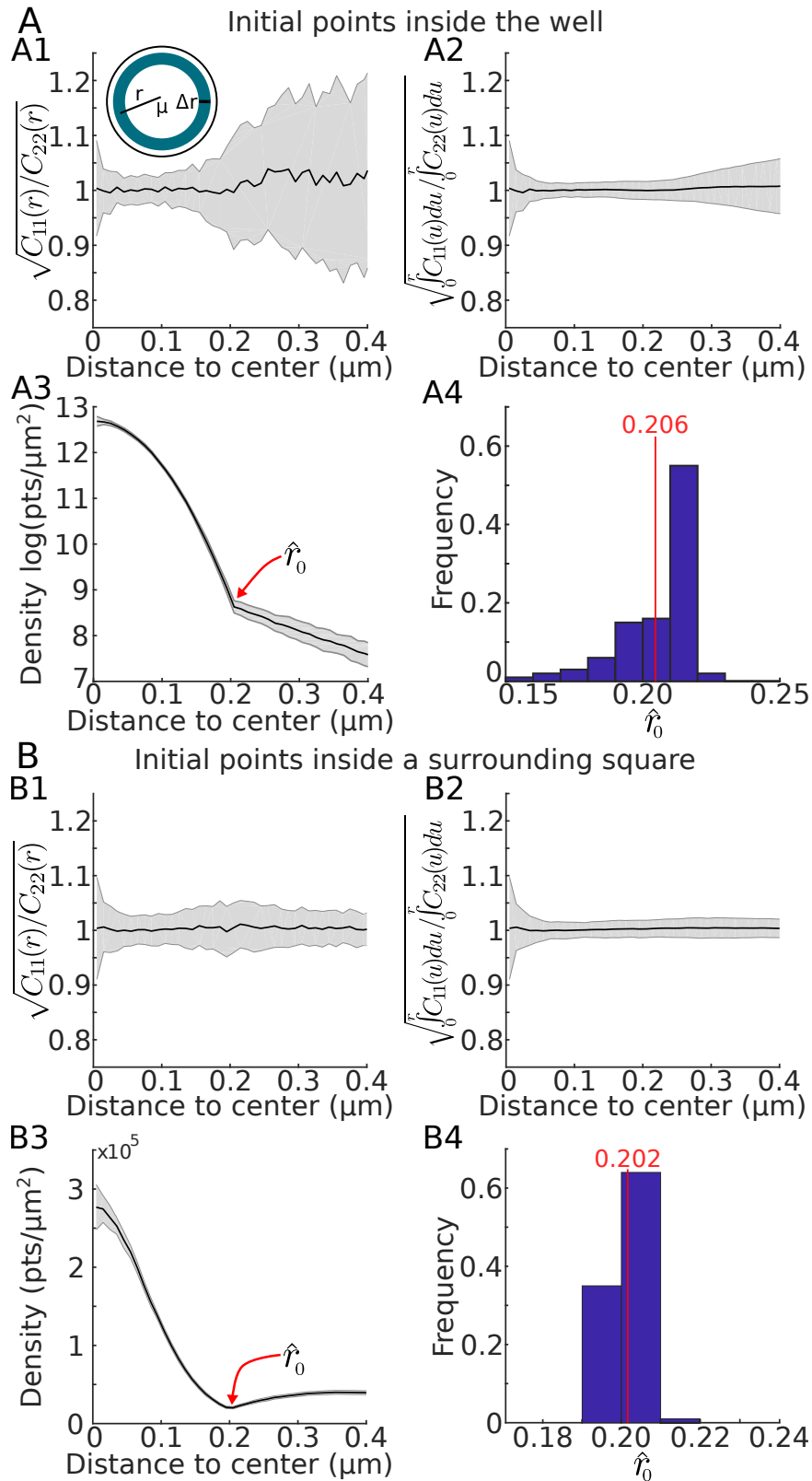


Figure 2.3: **Estimating the potential well boundary.** **A.** Initial trajectory points are chosen inside the well while in **B.** initial points are chosen inside a surrounding square. (A1,B1) Covariance ratio  $Cv$  estimated in the annulus  $[r, r + \Delta r]$ . (A2,B2) Cumulative from A1,B1. (A3,B3) Point density (in log for A3) estimated from the distance  $r$  to center showing a clear inflection point at the boundary of the well (criteria of selection). (A4,B4) Estimation of the radius  $r^0$  using the inflection point for A4 (as presented in A3) or the minimum value of the density for B4 (as presented in B3).



Thus in that case, we expect that  $Cv(r) \approx \sqrt{\frac{C_{1,1}(r)}{C_{2,2}(r)}} = \frac{a}{b}$ , the ratio of the large to the small elliptic semi-axes lengths does not depend on any other parameters. In the case of a disk,  $Cv(r) = 1$  as shown in the simulation cases (Fig. 2.3A1,A2,B1,B2).

Once the well boundary has been identified as circular, to estimate its radius  $r_0$ , we plotted the density of points  $\rho(r)$  versus  $r$ , the radial distance with respect to the center  $\hat{\boldsymbol{\mu}}$  (see Method). Interestingly, this procedure reveals the location of the boundary between the Boltzmann (inside the well) and the uniform (surrounding it) density distributions of the trajectories (Fig. 2.3A3,A4). When the initial points falls inside the well, the density of points decays with the radius  $r$  and the boundary can be identified by plotting  $-\log \rho(r)$  (Fig. 2.3A3). Indeed, for points inside the well, we have  $\log \rho(r) \sim C_0 - (\alpha x^2 + \beta y^2)$ , where  $r^2 = x^2 + y^2$ , with  $\alpha = \frac{2A}{a^2}$ ,  $\beta = \frac{2A}{b^2}$  and  $\exp(C_0)$  is the maximum value of the distribution. In practice, we find  $r_0$  as the first value for which the error  $\int_0^r (C_0 - C_1 s^2 + \log \rho(s))^2 ds$  starts to increase. The distribution of  $\hat{r}_0$  for 100 simulations is shown Fig. 2.3A4. When the initial points are now also chosen outside the well, the trajectories are either attracted inside the well or leave, thus the distribution of points is minimal at the boundary (Fig. 2.3B3), which allows us to recover  $r^0$  as the minimum point of the density curve (Fig. 2.3B4).

In the case of an elliptic well, we modified the previous method as follows: first, the ratio of the semi-axes lengths  $a/b$  is recovered as the maximal value of  $Cv(r)$  (Fig. 2.4A1,A2,B1,B2, for a ratio  $a/b = 2$ ). Second, using this ratio, we introduced the elliptic distance  $r_e(x, y) = \sqrt{x^2 + Cv(r^*)y^2}$ , for a point at position  $(x, y)$  from which we generated the point density distribution (Fig. 2.4A3,B3) and used on this curve the procedure described for the disk case to recover the large semi-axis  $\hat{a} = \hat{r}_{e,0}$  (Fig. 2.4A4,B4). The small-semi axis is then given by  $\hat{b} = \frac{\hat{a}}{\sqrt{Cv(r^*)}}$ .

To conclude, the present method based on the density of points allows to reconstruct the geometrical parameters of a bounded parabolic potential well: center, boundary, small and large semi-axes. In SI Figs. 2.9 and 2.10, we compare this density method with the MLE, which is classically used to recover the center and covariance, but not the boundary.

### 2.3.2 Estimating the characteristics of the well using the velocity distribution

In this section, we describe a second approach to reconstruct the potential well associated to a nanodomain, using the statistics of displacements  $\mathbf{X}(t + \Delta t) - \mathbf{X}(t)$ . They allow to recover the drift of the vector field and reconstruct the center  $\boldsymbol{\mu}$  and the two axes  $a, b$  of the well boundary. This method is based on the least square quadratic error (LSQE),

$$Err_b(\mu_x, \mu_y, \lambda_x, \lambda_y) = \sum_{i=1}^N \left\| -\nabla U(\mathbf{X}_i) - \begin{bmatrix} \lambda_x x_i \\ \lambda_y y_i \end{bmatrix} \right\|^2,$$

between the empirical drift and the parabolic well  $U$ , defined in equation (2.7), with  $\lambda_x = -\frac{2A}{a^2}$ ,  $\lambda_y = -\frac{2A}{b^2}$ .

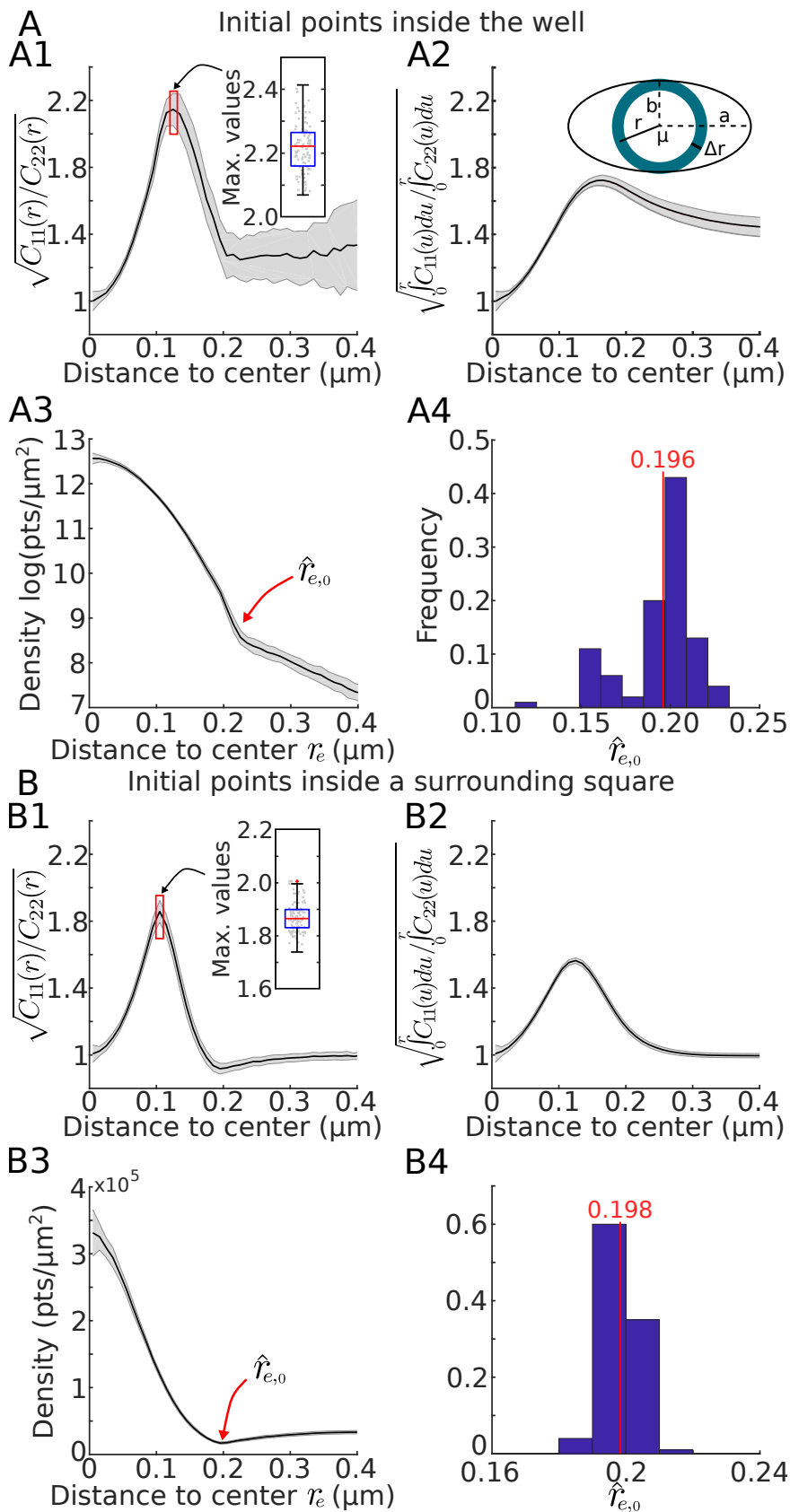


Figure 2.4: **Estimating the potential well boundary for an elliptic well.** **A.** Initial trajectory points are chosen inside the well while in. **B.** Initial points are chosen inside a surrounding square. (A1,B1) Covariance ratio  $Cv$  estimated in the annulus  $[r, r + \Delta r]$ . (A2,B2) Cumulative from A1,B1. (A3,B3) Point density (in log for A3) based on the modified distance  $r_e$  to the center showing a clear inflection point at the boundary of the well. (A4,B4) Estimation of the radius  $\hat{r}_{e,0}$  using the inflection point for A4 (as presented in A3) or the minimum value of the density for B4 (as presented in B3).

### 2.3.3 Estimating the center and the field coefficients of the potential well

The coefficient  $\lambda$  and the center  $\mu$  of the potential well can be obtained explicitly from equations (2.35) and (2.40), (2.41) respectively. We compare in Fig. 2.5A, the reconstructed and the true drift value based on equation (1.20) for various grid sizes. At this stage, we considered the boundary to be known and estimated the drift only for bins that are falling inside the well. The error of the norm  $\frac{\langle ||\bar{b}-b|| \rangle}{\langle ||\bar{b}|| \rangle}$  is plotted in Fig. 2.5B for multiple time steps  $\Delta t$  and for three grid sizes  $\Delta x = 10, 50$  and  $90$  nm. Having both a small grid size and time step  $\Delta t$  produces a large error that quickly decreases with increasing the time step  $\Delta t$ . Interestingly, for a large grid size, we found a slow increase of the error when increasing  $\Delta t$ . To better understand which parts of the field contributed the most to the error, we plotted *Err* versus the distance to the center (Fig. 2.5C). This result shows that for small size  $\Delta x = 10$  nm, a major contribution came from the center, while for large step  $\Delta x = 50, 90$  nm, an error came also from the boundary. We refer to Fig. 2.11 for recovering a drift at a different time resolution  $\Delta t$  and also with some restrictions on the trajectories for which the end point remained inside the well (Fig. 2.12).

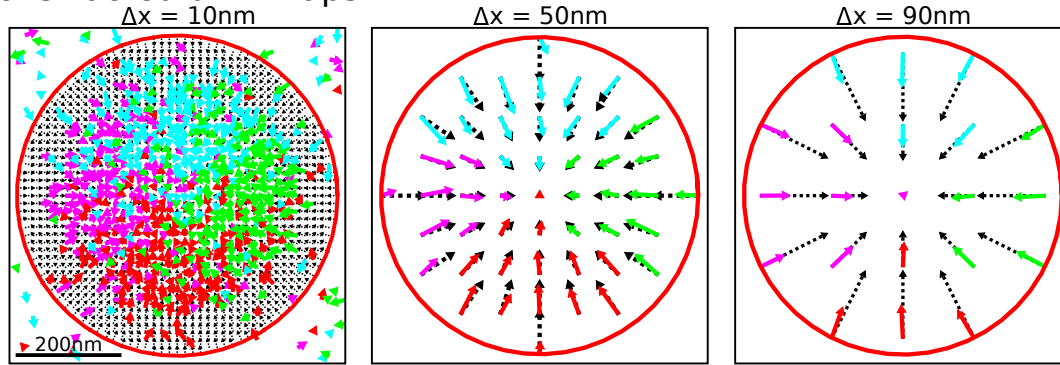
Finally, to estimate the boundary of the well from the drift distribution (Fig. 2.6A), we plotted the drift amplitude versus the distance to the well center (Fig. 2.6B, blue crosses representing the drift amplitude in individual bins). From the distribution and the average (Fig. 2.6B lower panel), we could recover the location of the boundary at the local maximum. Indeed, after the boundary is passed, the contribution of the deterministic field disappears and only fluctuations due to the Brownian motion remains in the statistics. We apply the same procedure for the case of an ellipse (Fig. 2.6C,D) and recover the boundary after we used the covariance ratio  $Cv$  (equation (2.16)) to plot the drift amplitude versus the elliptic distance to the boundary.

To evaluate the influence of the bins located at the center or the ones near the boundary, we estimated the center  $\mu$ , and eigenvalues  $\lambda_x$  and  $\lambda_y$  in four cases: for all bins falling inside the well, all bins except the ones at the center, all bins except the ones intersecting the boundary and finally removing the center and the boundary bins (Fig. 2.6E). We found that the latter case produces the best estimation.

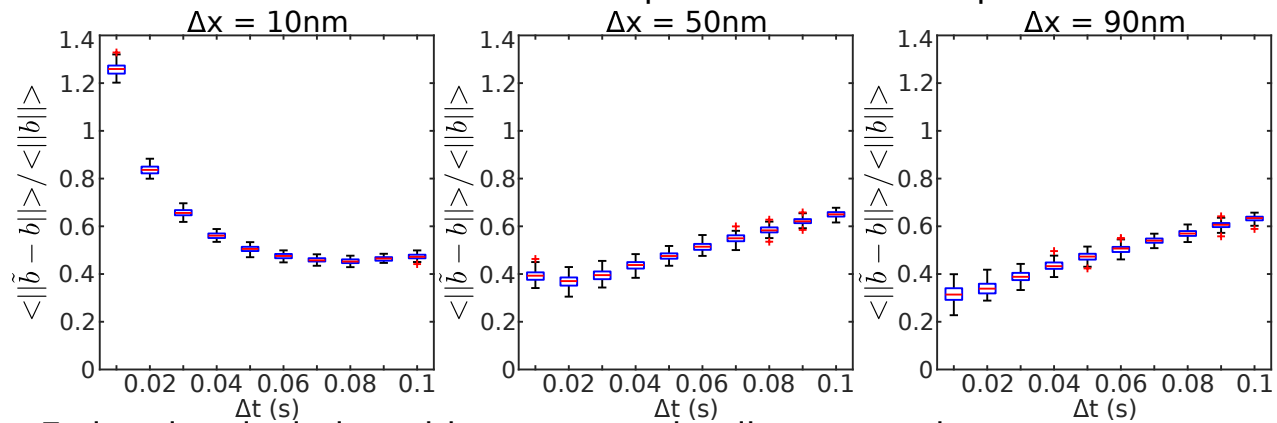
### 2.3.4 Interpretation high-density regions for Ca<sub>v</sub>2.2 and GPI-GFP as potential wells

In this section, we will apply the methodology developed in the present article to characterize high-density regions found in SPTs of voltage-gated calcium channels and phospholipids. We recently reported that these regions could be associated with potential wells, as revealed from the voltage-gated calcium channels Ca<sub>v</sub>2.1 isoform [98]. We focus here on the isoform Ca<sub>v</sub>2.2 (N-type channel) by using the density of points, the least-square estimation (SI Section 2) and the maximum-likelihood method (SI section 1). For the analysis, we use only wells that contain at least 50 points with a minimum of 5 different trajectories.

### A Reconstructed drift maps



### B Estimation error of the drift with respect to the time steps



### C Estimation deviation with respect to the distance to the center

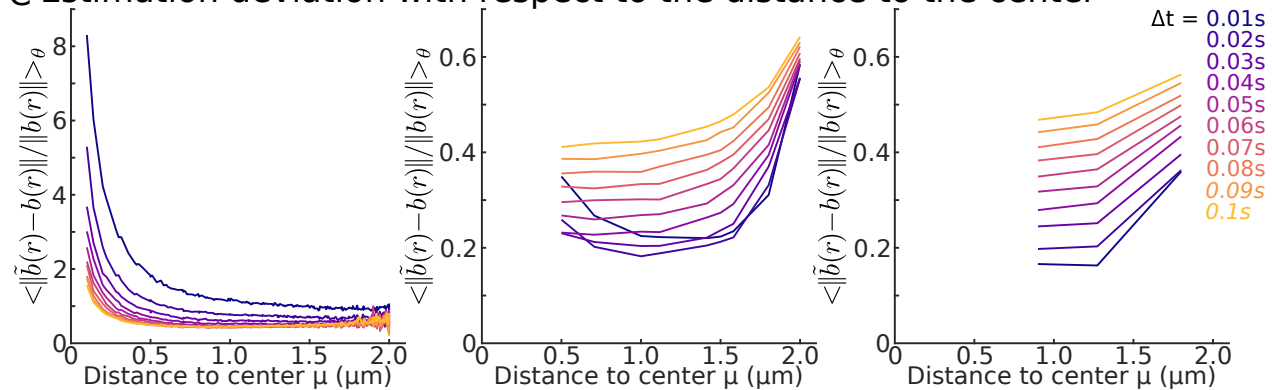


Figure 2.5: **Vector field characteristics.** **A.** Recovering the local drift field inside a circular well for different grid sizes (10 nm, 50 nm, 90 nm) using numerical simulations with  $\Delta t = 20$  ms, with the constraints that at least 10 points falls inside a bin. **B.** Error between the true and observed fields averaged over all the square bins inside the well vs the time step  $\Delta t$ . **C.** Error between the true and observed fields averaged over the radial angle vs the distance  $r$  to the center of the well center for various timestep (see color code).

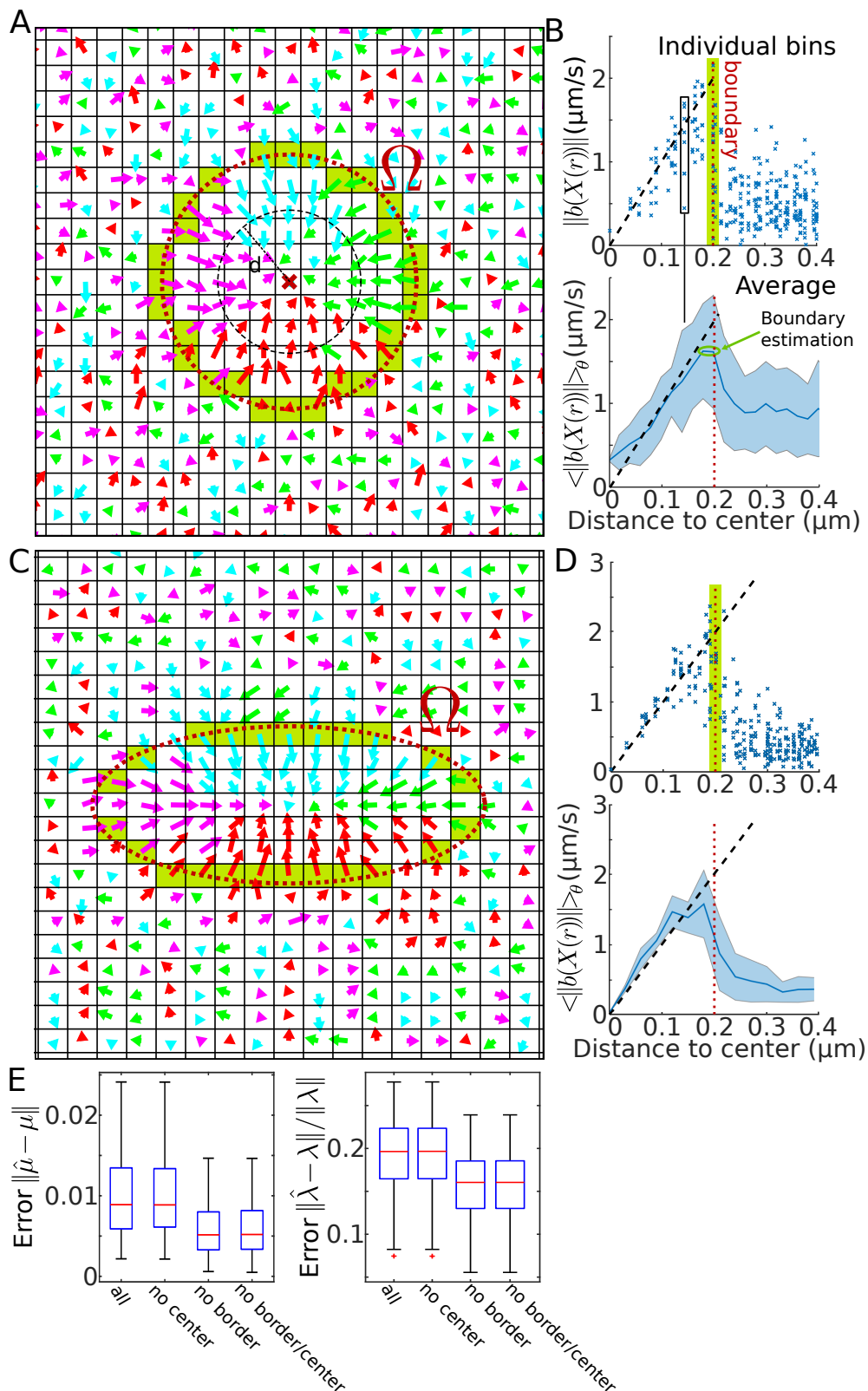


Figure 2.6: **Recovery of the well depth from trajectories.** **A.** Vector field recovered from trajectories including the field generated by the Brownian dynamics outside the well. At the boundary of the well, there are mixed displacements (OU and Brownian), marked by the green band. **B.** Upper: Drift amplitudes in each grid square versus the distance  $r$  to the center of the well. The expected amplitude is marked by a dashed line and the boundary with a narrow green band. Lower: Average and SD of the upper panel. **C.,D.** Same as in A-B for the case of an ellipse where  $a/b = 2$ . **(E)** Error of the center  $\|\mu\| = \sqrt{\mu_x^2 + \mu_y^2}$  and the eigenvalue  $\|\lambda\| = \sqrt{\lambda_x^2 + \lambda_y^2}$  of the reconstructed ellipse in four cases: (1) bins falling inside the boundary are considered, (2) the center bin has been omitted (3) boundary bins are omitted and (4) when the center bin plus the ones intersecting the boundary (green bin in C) are not considered.

We find that all three approaches produce reasonable values of the coefficient  $A$  and the energy (we restricted to wells with energies  $< 7$  kT). The values of the parameters are summarized in Table. 2.1. We report in Fig. 2.7A–C that the high-density regions can be characterized as potential wells with the following characteristics: the two main axes have average lengths ( $\pm$ SD) of  $a = 104 \pm 36$  nm,  $b = 77 \pm 20$  nm associated with a mean energy of 3.3 kT estimated for the density method. These results differ from the  $\text{Ca}_v2.1$  isoform [98]. Note that the distribution of energy varied with the statistical method (Fig. 2.7C), as we reported  $E = 3.1 \pm 0.5$  kT for the MLE and  $E = 1.6 \pm 0.7$  kT for the LSQ. To conclude, this statistical analysis suggests that to trap calcium channels, specific long-range molecular mechanisms should be present in the active zone of the pre-synaptic terminal, probably associated to vesicular release molecules such as synaptotagmin. These sites retain channels for a long time, enough to trigger vesicular release.

Parameter	GPI dataset	$\text{Ca}_v2.2$ dataset
$\Delta t$ (exp)	20 ms	33 ms
$\Delta x$	40 nm	30 nm
$r_{min}$	30 nm	20 nm
$r_{max}$	300 nm	400 nm
$\Delta r$	20 nm	10 nm

Table 2.1: **Parameters used for  $\text{Ca}_v2.2$  and GPI analysis.**

We also apply our statistical methods to the case of GFP linked to the outer leaflet of the membrane by a GPI-anchor (Fig. 2.8A–C), which are considered to be non-interacting molecules. However, we found many high-density regions ( $N = 181$ ), which are characterized as potential wells. The elliptic axes are  $a = 158 \pm 57$  nm and  $b = 118 \pm 39$  nm, associated with an energy of  $E = 3.6 \pm 1.0$ ,  $1.5 \pm 1.0$  and  $3.5 \pm 1.0$  kT for the density, LSQ and MLE methods respectively. To conclude, although it is surprising to detect high-density regions in GPI-GFP SPTs, we found here that they can be characterized as potential wells. Possibly they correspond to places where local signaling complexes or other transmembrane proteins are present. The exact nature of these regions remain unclear and should be further investigated.

## 2.4 Summary and Discussion

### 2.4.1 Two statistical methods to interpret high-density regions

We presented here two methods to extract the biophysical characteristics of high-density regions explored by SPTs. Interestingly, these regions are associated with bounded potential wells. The first method exploits the density of points of the trajectories, ignoring the causality between the successive points. It assumes that the nanodomain is a parabolic potential well with an elliptic base and a constant diffusion coefficient. In that case, the distribution of points inside the well is given by a Boltzmann distribution and should be uniform outside.

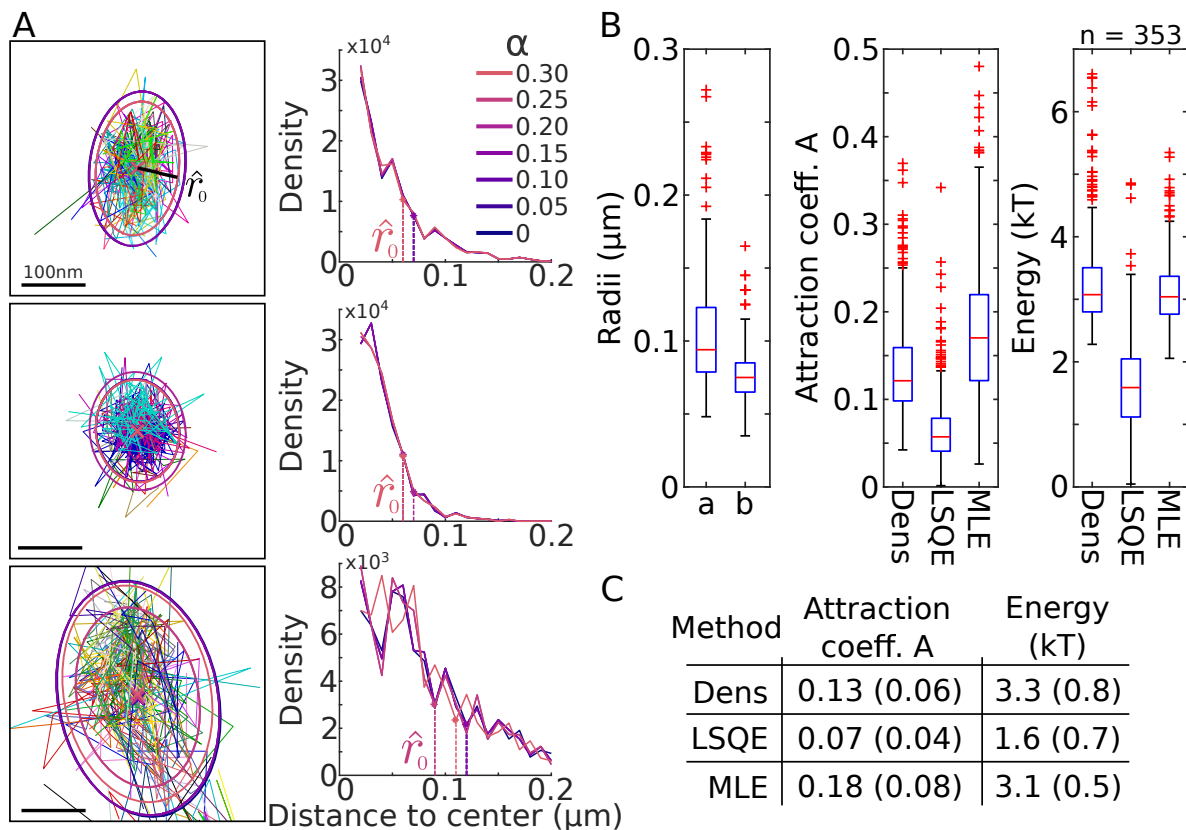


Figure 2.7: **Reconstruction of wells associated to  $\text{Ca}_v2.2$ .** **A.** Three examples of potential wells (left) obtained from the density analysis on SPTs. The boundary of the well are estimated from various level of density  $\alpha$  (right). The estimated radius  $\hat{r}_0$  is obtained using a threshold  $T = 4\%$  on the density. **B.** Box plots for the statistics computed over 353 detected wells for the two semi-axes a and b of the ellipse, the coefficient A and the energy (in kT). Results are obtained for the Density, LSQE and MLE methods. **C.** Summary of mean and SD for the coefficient A and the energy.

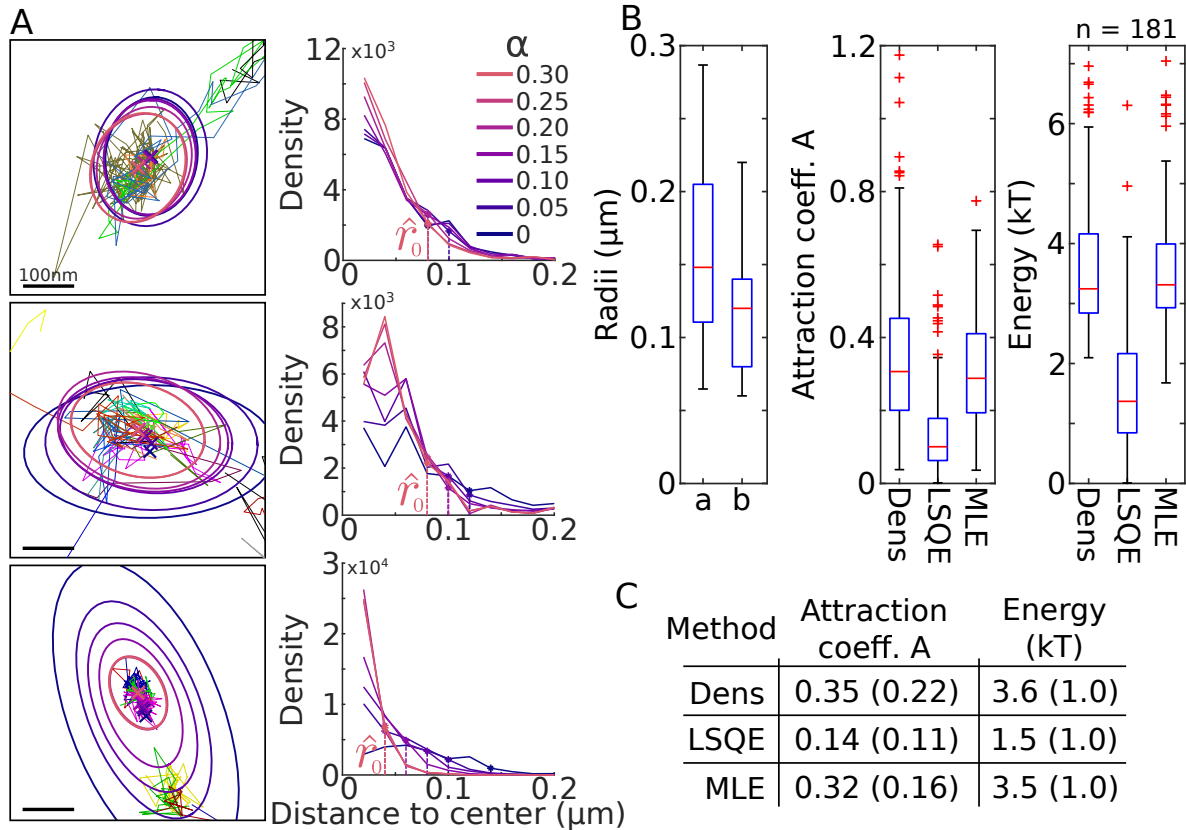


Figure 2.8: **Reconstructed wells associated to GPI-anchored GFP.** **A.** Three examples of potential wells (left) obtained from the density analysis on SPTs. The boundary of the well are estimated from various level of density  $\alpha$  (right). The estimated radius  $\hat{r}_0$  is computed using a threshold  $T = 4\%$  on the density. **B.** Box plots of the two semi-axes  $a$  and  $b$  (of the ellipse), estimated over 181 detected wells, the coefficient  $A$  and the energy (in kT), compared for the Density, LSQ and MLE methods. **C.** Table of mean and SD for the coefficient  $A$  and the energy.



We use this key observation to recover the main physical parameters and the location of the boundary. We compared also our result to the classical MLE (see Figs. 2.9, 2.10). The second method is based on estimating the vector field distribution at a given bin resolution  $\Delta x$ . We used an optimal estimator to recover the characteristic of the field and we found that the boundary is located at the discontinuity between the converging field of the well (Ornstein-Uhlenbeck) and the random field generated the surrounding Brownian motion. Finally, the present methods are based on multiple averaging over many trajectories [64], which provide robustness, reducing the effect of tracking errors or localization noise [106, 107].

The two methods are complementary and provide certain advantages compared to the MLE and PCA. In all cases, the center of the well could be retrieved. The quality of the estimators of the covariance parameters, however, were dependent on the method: changing the time  $\Delta t$  and spatial  $\Delta x$  steps influenced the recovery process as shown in Figs. 2.3, 2.4 and 2.6. The advantage of the first method is that we do not need to introduce an artificial grid of size  $\Delta x$  which is a serious constraint in the second method as the bins size defines the resolution to recover the well and its boundary.

## 2.4.2 High-density regions contained calcium-voltage channels and GPI SPTs data

We recall that high-density regions revealed by SPTs are not necessarily to due physical forces and potential wells [91]. However, for potential well, the geometry (center, curvature and boundary) can be recovered from our two methods. We applied them to  $\text{Ca}_V$  channels that mediate vesicular release at neuronal synapses and to phospholipid anchored GFP (GPI-GFP) moving on the cellular membrane. We found that the high-density regions for  $\text{Ca}_V2.2$  (Fig. 2.7A) are characterized by two main axes with a length  $a = 104 \pm 36$  and  $b = 77 \pm 20$  nm (Fig. 2.7B), with a mean energy of  $3.3 \pm 0.8$  kT (density method, Fig. 2.7C). We note that hydrogen bonds between calcium channels and phospholipid molecules could participate in the formation of the wells [108]. Surprisingly, we did not expect to find high-density regions for GPI-GFP, but we found several (Fig. 2.8A) that were characterized by average semi-axes lengths  $a = 158 \pm 57$  and  $b = 118 \pm 39$  nm (Fig. 2.8B), with a mean energy  $E = 3.6 \pm 1.0$  kT (density method, Fig. 2.8C). Possibly the higher energies of GPI-GFP wells can be due to the large variance caused by the lower number of trajectories restricted inside the wells as compared to  $\text{Ca}_V$ .

Although the interpretation of high-density regions as potential wells for AMPA receptors was first anticipated in [109] and discovered in [91], the nature of these wells and others, remains unclear [100]. Potential wells were found for membrane proteins such  $\text{Ca}_V$  [98], GAG [99, 110] and recently for G-protein [111]. They could be generated by protein clusters, membrane cusps at vesicle fusion points or membrane-membrane contact at location of organelle interactions [112]. In general, potential wells are characterized by long-range forces of the order of hundreds of nanometers.

The wells could have multiple roles: they could retain receptors for hundreds of milliseconds to seconds at specific locations in order to increase the probability of a robust signal transduction, such as during synaptic transmission. Transient wells allow to trap

proteins to create aggregates as proposed for capsid assembly [99, 110]: once the energy of the well decreases, molecules are not interacting with the well anymore. Other possible roles for wells could be regulating the flow of receptors in micro-compartments such as dendritic spines [103] or trapping proteins in the endoplasmic reticulum [113]. Finally, correlating undefined membrane geometry with an energy landscape remains difficult, because a physical model is needed to interpret them. Thus, the dynamics of receptors outside potential wells that deviates from trapped Brownian motion is still challenging to comprehend.

## 2.5 Supplementary Information

In the SI, we first present the MLE estimator that we adapted to study potential wells. In the second part, we summarize some properties of the drift estimator that depends on the time step  $\Delta t$  and the bin size  $\Delta x$ .

### 2.5.1 MLE estimator for a potential well

We modified the Maximum likelihood Estimator (MLE) procedure to reconstruct from SPTs, the geometrical parameters (center and covariance matrix) of a well. Using a Ornstein-Uhlenbeck process, we apply the MLE procedure to the points of the trajectories falling inside the ensemble

$$\Gamma_\alpha = \{\mathbf{X}_i \text{ such that } \rho(\mathbf{x}) > \alpha\}, \quad (2.17)$$

where  $\rho$  is the steady-state probability density function of the OU process. The advantage of the Maximum-likelihood approach is that no spatial discretization is needed. We recall that the transition probability density of an OU process centered at  $\mu$ , with diffusion coefficient  $D$  and spring constant  $\lambda$

$$\dot{x} = -\lambda(x - \mu) + \sqrt{2D}\dot{w} \quad (2.18)$$

is

$$p(x, t|y, s) = \sqrt{\frac{\lambda}{2\pi D(1 - e^{-2\lambda(t-s)})}} e^{-\frac{\lambda}{2D} \frac{(x - \mu - (y - \mu)e^{-\lambda(t-s)})^2}{(1 - e^{-2\lambda(t-s)})}}. \quad (2.19)$$

The transition probability is

$$p(y_{i+1}|y_i) = \sqrt{\frac{\lambda}{2\pi D(1 - e^{-2\lambda\Delta t})}} e^{-\frac{\lambda}{2D} \frac{(y_{i+1} - y_i e^{-\lambda\Delta t})^2}{1 - e^{-2\lambda\Delta t}}}, \quad (2.20)$$

where a trajectory is discretized in  $y_1, \dots, y_M$ , with a fixed time step  $\Delta t$ . The log likelihood is

$$l(y_1, \dots, y_M | \lambda, \mu, D) = \sum_{i=0}^{M-1} \log p(y_{i+1} | y_i) \quad (2.21)$$

$$= \frac{M}{2} \log \frac{\lambda}{2\pi D(1 - e^{-2\lambda\Delta t})} - \frac{\lambda}{2D(1 - e^{-2\lambda\Delta t})} \sum_{i=0}^{M-1} (y_{i+1} - y_i e^{-\lambda\Delta t})^2. \quad (2.22)$$

The maximum-likelihood approach consists in estimating  $\lambda$  that maximizes the log-likelihood  $l(y_1, \dots, y_M)$ . We change variables  $x = e^{-\lambda\Delta t}$  and  $v = \frac{\lambda}{2\pi D(1 - e^{-2\lambda\Delta t})}$  so that

$$\tilde{l}(y_1, \dots, y_M | x, \mu, v) = \frac{M}{2} \log \frac{v}{\pi} - v \sum_{i=1}^{M-1} (y_{i+1} - \mu - (y_i - \mu)x)^2. \quad (2.23)$$

At the maximum,

$$\frac{\partial \tilde{l}}{\partial x} = \frac{\partial \tilde{l}}{\partial v} = \frac{\partial \tilde{l}}{\partial \mu} = 0, \quad (2.24)$$

leads to the coupled equations

$$\begin{aligned} \hat{\mu} &= \frac{\sum_0^{M-1} y_{i+1} - \hat{x} y_i}{M(1 - \hat{x})} \\ \hat{x} &= \frac{\sum_0^{M-1} (y_{i+1} - \hat{\mu})(y_i - \hat{\mu})}{\sum_0^{M-1} (y_i - \hat{\mu})^2} \\ \hat{v} &= \frac{M}{2 \sum_0^{M-1} ((y_{i+1} - \hat{\mu}) - (y_i - \hat{\mu})\hat{x})^2}. \end{aligned} \quad (2.25)$$

This system of equation can be solved leading to the following estimators [82]

$$\begin{aligned} \hat{x} &= \frac{\frac{1}{M} \sum_{i=0}^{M-1} y_{i+1} y_i - \frac{1}{M^2} \sum_{i=0}^{M-1} y_{i+1} \sum_{i=0}^{M-1} y_i}{\frac{1}{M} \sum_{i=1}^{M-1} y_i^2 - \frac{1}{M^2} \left( \sum_{i=1}^{M-1} y_i \right)^2} + \frac{4}{M} \\ \hat{\mu} &= \frac{\frac{1}{M} \sum_{i=1}^{M-1} (y_{i+1} - \hat{x} y_i)}{1 - \hat{x}}. \end{aligned} \quad (2.26)$$

The diffusion coefficient can be found from the third equation of eq. (2.25):

$$\hat{v} = \frac{1}{M} \sum_0^{M-1} (y_{i+1} - \hat{x} y_i - \hat{\mu}(1 - \hat{x}))^2. \quad (2.27)$$

This procedure can be generalized in two dimensions and in addition, we apply the estimator of eq. (2.25) to the ensemble  $\Gamma_\alpha$ , defined by (2.17). We thus obtain the following estimators

$$\begin{aligned}\hat{x}^\alpha &= \frac{\frac{1}{M_\alpha} \sum_{y_i \in \Gamma_\alpha} y_{i+1} y_i - \frac{1}{M_\alpha^2} \sum_{y_i \in \Gamma_\alpha} y_{i+1} \sum_{y_i \in \Gamma_\alpha} y_i}{\frac{1}{M_\alpha} \sum_{y_i \in \Gamma_\alpha} y_i^2 - \frac{1}{M_\alpha^2} \left( \sum_{y_i \in \Gamma_\alpha} y_i \right)^2} + \frac{4}{M_\alpha} \\ \hat{\mu}^\alpha &= \frac{\frac{1}{M_\alpha} \sum_{y_i \in \Gamma_\alpha} (y_{i+1} - \hat{x} y_i)}{1 - \hat{x}},\end{aligned}\tag{2.28}$$

where  $M_\alpha$  is the number of points  $y_i \in \Gamma_\alpha$ . We apply this estimator to numerical simulations in Figs. 2.9 and 2.10 and to Ca<sub>v</sub>2.2 and GPI-GFP SPTs in Figs. 2.7 and 2.8 respectively.

## 2.5.2 Least Square Quadratic Estimator (LSQE)

### Estimating the center and the curvature of the well

To recover the potential well from the drift distribution, we use a least square estimator

$$\begin{aligned}Err_{\mathbf{b}}(\mu_x, \mu_y, \lambda_x, \lambda_y) &= \sum_{i=1}^N \left\| -\nabla U(\mathbf{X}_i) - \mathbf{b}(\mathbf{X}_i) \right\|^2 \\ &= \sum_{i=1}^N \left( \mathbf{b}_x^i + \lambda_x(x_i - \mu_x) \right)^2 + \left( \mathbf{b}_y^i + \lambda_y(y_i - \mu_y) \right)^2,\end{aligned}\tag{2.29}$$

where  $N$  is the number of points  $\mathbf{X}_i = (x_i, y_i)$  and the potential well is

$$U(\mathbf{X}) = \lambda_x(x - \mu_x)^2 + \lambda_y(y - \mu_y)^2,\tag{2.30}$$

so that

$$\mathbf{b}(\mathbf{X}) = -\nabla U(\mathbf{X}) = 2 \begin{bmatrix} \lambda_x(x - \mu_x) \\ \lambda_y(y - \mu_y) \end{bmatrix}.\tag{2.31}$$

The minimizers are given by

$$\frac{\partial}{\partial \lambda_x} Err_{\mathbf{b}} = \frac{\partial}{\partial \lambda_y} Err_{\mathbf{b}} = 0\tag{2.32}$$

and

$$\frac{\partial}{\partial \mu_x} Err_{\mathbf{b}} = \frac{\partial}{\partial \mu_y} Err_{\mathbf{b}} = 0\tag{2.33}$$

from which, we obtain the center

$$\tilde{\mu}_x = \frac{\sum_{i=1}^N \mathbf{b}_x^i + \lambda_x x_i}{N \lambda_x}, \quad \tilde{\mu}_y = \frac{\sum_{i=1}^N \mathbf{b}_y^i + \lambda_y y_i}{N \lambda_y}.\tag{2.34}$$

and the eigenvalues of the covariance matrix:

$$\tilde{\lambda}_x = \frac{\sum_{i=1}^N \mathbf{b}_x^i (x_i - \mu_x)}{\sum_{i=1}^N (x_i - \mu_x)^2}, \quad \tilde{\lambda}_y = \frac{\sum_{i=1}^N \mathbf{b}_y^i (y_i - \mu_y)}{\sum_{i=1}^N (y_i - \mu_y)^2}. \quad (2.35)$$

In practice, we computed the center  $\mu_x, \mu_y$  and the eigenvalues  $\lambda_x, \lambda_y$  over the points  $\mathbf{X}_i$  falling inside the well.

### Center location and semi-axes from optimal fit

We derive here a close formula for the eigenvalues and the center associated to the optimal estimators of the drift. For an OU process, we recall that the eigenvalues are given by

$$\begin{aligned} \lambda_y &= \frac{\sum_{i=1}^N \mathbf{b}_y^i (y_i - \mu_y)}{\sum_{i=1}^N (y_i - \mu_y)^2} = \frac{\sum_{i=1}^N \mathbf{b}_y^i y_i - \mu_y \sum_{i=1}^N \mathbf{b}_y^i}{\sum_{i=1}^N y_i^2 - 2\mu_y \sum_{i=1}^N y_i + N\mu_y^2} \\ &= \frac{A_y - B_y \mu_y}{C_y - 2\mu_y D_y + N\mu_y^2}, \end{aligned} \quad (2.36)$$

where  $A_y = \sum_{i=1}^N \mathbf{b}_y^i y_i$ ,  $B_y = \sum_{i=1}^N \mathbf{b}_y^i$ ,  $C_y = \sum_{i=1}^N y_i^2$  and  $D_y = \sum_{i=1}^N y_i$ . Using

$$\mu_y = \frac{\sum_{i=1}^N \mathbf{b}_y^i + \lambda_y \sum_{i=1}^N y_i}{N\lambda_y} = \frac{B_y + \lambda_y D_y}{N\lambda_y}, \quad (2.37)$$

and eq. (2.36) in eq. (2.37), we obtained

$$\mu_y = \frac{B_y + \frac{A_y - B_y \mu_y}{C_y - 2\mu_y D_y + N\mu_y^2} D_y}{N \frac{A_y - B_y \mu_y}{C_y - 2\mu_y D_y + N\mu_y^2}} \quad (2.38)$$

where  $\mu_y$  is solution of the quadratic equation

$$2NB_y\mu_y^2 - (NA_y + 3B_y D_y)\mu_y + B_y C_y + A_y D_y = 0. \quad (2.39)$$

With  $\Delta_y = N^2 A_y^2 - 2NA_y B_y D_y + 9B_y^2 D_y^2 - 8NB_y^2 C_y$ , we retain the positive solution

$$\mu_y = \frac{(NA_y + 3B_y D_y) + \sqrt{\Delta_y}}{4NB_y}. \quad (2.40)$$

Similarly, we get

$$\mu_x = \frac{(NA_x + 3B_x D_x) + \sqrt{\Delta_x}}{4NB_x} \quad (2.41)$$

Relations (2.40) and (2.41) lead to a close expression of the eigenvalues (eq. (2.35)).

## Comparing ML, LSQ and density estimators

To recover the center and eigenvalues of an Ornstein-Uhlenbeck process with the potential well given in eq. (2.30), we apply the MLE procedure (subsection 2.5.1) for various values of the parameter  $\alpha$ . We compare the results with the density (section 3 main text) and the LSQ (section 2.5.2) methods. We find that the MLE and density approaches are quite robust and give similar results, as shown in Fig. 2.9: Interestingly, all three estimators allow to recover the center  $(\mu_x, \mu_y)$  with high accuracy for the disk and the ellipse, when 100% to 50% of the points are taken into account  $\alpha \in [0; 0.5]$  (Fig. 2.9). However, the estimation of the eigenvalues is acceptable for the MLE only, in the range  $\alpha \in [0; 0.5]$ , because it diverges in the two other cases, except when  $\alpha = 0.1$  (90% of the distribution is used).

When the time step  $\delta t$  of the numerical simulations of eq. (2.18) and the sampling time  $\Delta t$  are equal, the three methods lead to a good recovery of the center  $(\hat{\mu}_x, \hat{\mu}_y)$  (Fig. 2.10), but differ for recovering the eigenvalues  $(\hat{\lambda}_x, \hat{\lambda}_y)$ . The least square approach is less dependent on the parameter  $\alpha$  than the two others. Probably because the distribution was generated with a large time step so that the statistics are calculated on trajectories far from the equilibrium. This result shows that the least square approach does not require to sample over a steady state distribution and thus recovering the parameters from the drift is possible for a large range of the parameter  $\alpha$ . In section 2.5.4, we will estimate the effect of changing the time steps.

### 2.5.3 Influence of the time and spatial discretizations on the Least Square Estimation

In this section, we shall estimate the impact of the initial points distribution of the trajectories on the estimation of the drift. For the stochastic equation [103]

$$\dot{\mathbf{X}} = \mathbf{b}(\mathbf{X}) + \sqrt{2}B(\mathbf{X})\dot{\mathbf{w}}, \quad (2.42)$$

the optimal estimator for the drift  $\mathbf{b}$  at a time resolution  $\Delta t$  is obtained by the formula [114, 64],

$$\begin{aligned} \mathbf{b}_{\Delta t}(\mathbf{x}) &= \mathbb{E} \left[ \frac{\mathbf{X}_{n+1} - \mathbf{X}_n}{\Delta t} \middle| \mathbf{X}_n = \mathbf{x} \right] = \frac{1}{\Delta t} \int_{\mathbb{R}} (\mathbf{y} - \mathbf{x}) p(\mathbf{X}_{n+1} = \mathbf{y} | \mathbf{X}_n = \mathbf{x}) d\mathbf{y} \\ &= a(\mathbf{x}) + o(1), \end{aligned} \quad (2.43)$$

where  $\mathbf{X}_n = \mathbf{X}(n\Delta t)$  and  $\mathbb{E}[\cdot]$  in eq. (2.43) is the expectation. When a grid of size  $\Delta x$  is used to estimate the drift map, all points in bin  $k$  leads to the same drift  $\mathbf{b}_{\Delta t}(\mathbf{x}_k)$ , where  $\mathbf{x}_k$  is the center of the bin. For long trajectories, the stochastic process samples the steady-state distribution  $p(x)$ . Such distribution might influence the computation of the drift inside bin  $k$ . To estimate this contribution, we normalize the steady-state distribution  $q(x)$  of the stationary process by

$$\begin{aligned} q_{\Delta x}(x) &= Pr\{\mathbf{X} \in \Delta(x) | \mathbf{X} \text{ has a steady state } p(x)\} \\ &= \frac{p(x)}{\int_{\Delta(x)} p(x) dx}, \end{aligned} \quad (2.44)$$

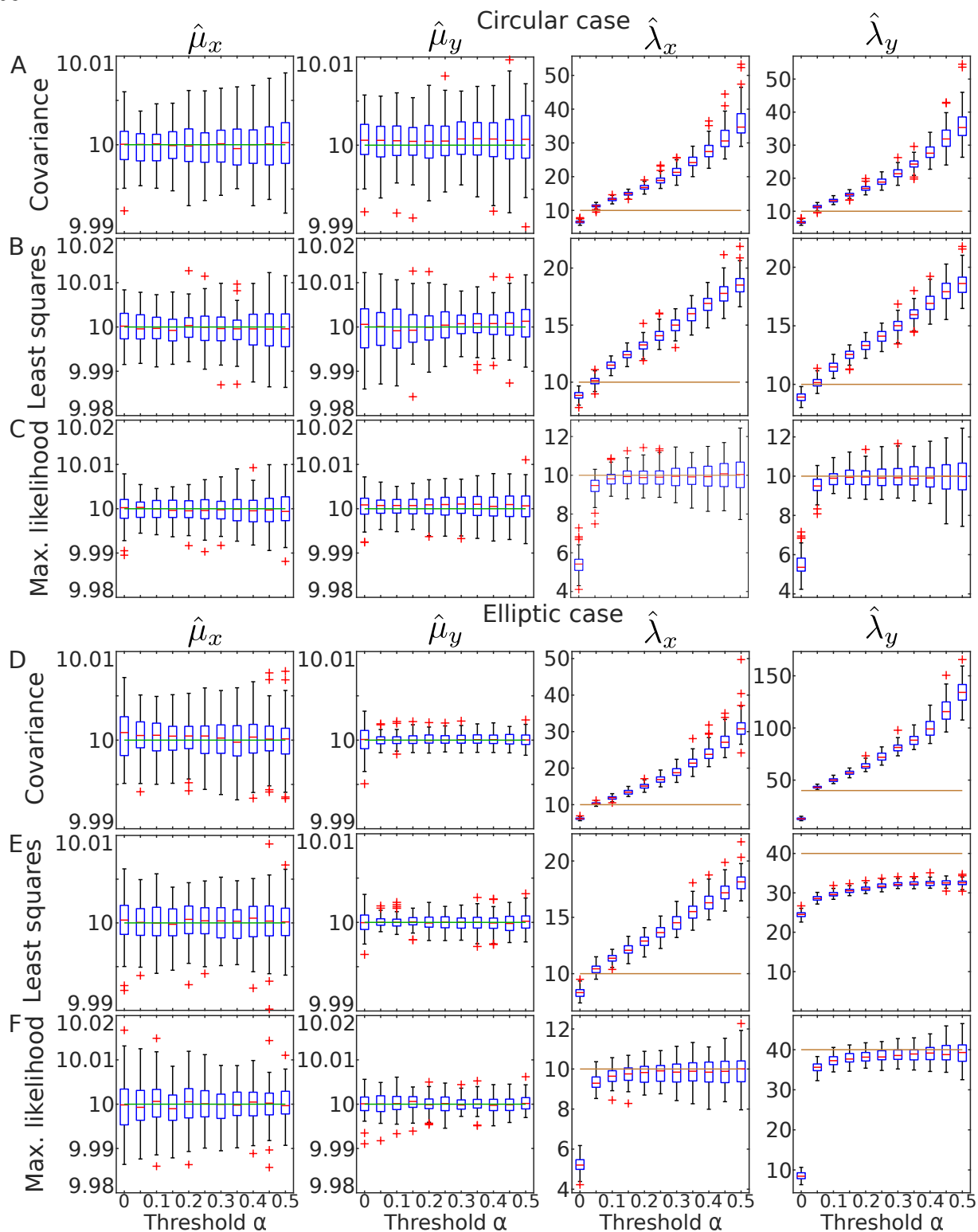
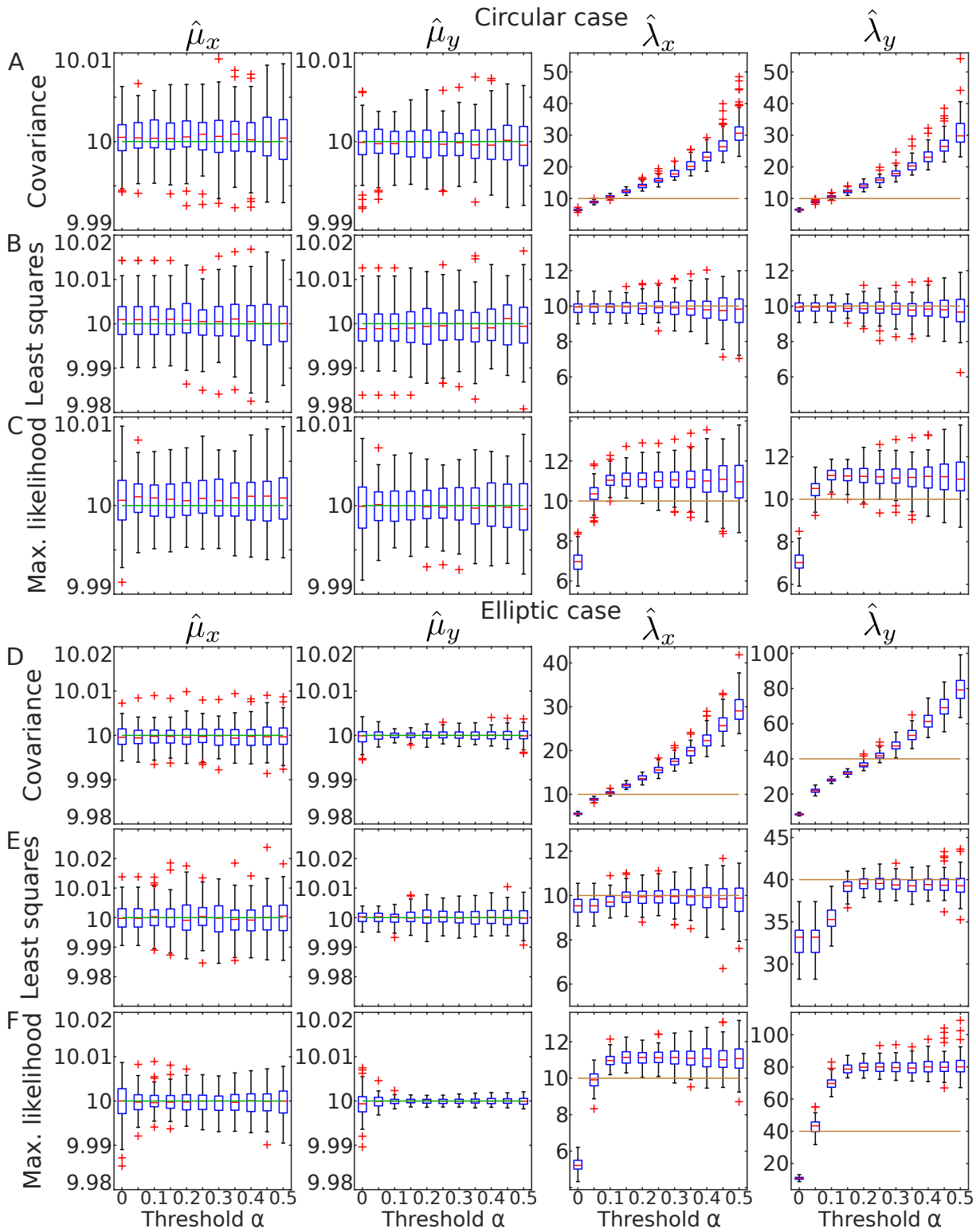


Figure 2.9: **Comparing MLE and covariance estimators.** **A.-B.** estimation of  $\lambda_x, \lambda_y$  for the case of a disk for the MLE and covariance estimator respectively. The true value of  $\lambda_x = \lambda_y = 10$  corresponds to the purple line. **C.-D.** same as in A-B for the case of an ellipse with  $\lambda_x = 10$  and  $\lambda_y = 40$ . Data were obtained from stochastic simulations with a time step  $\delta t = 10^{-4}s$  whereas the sampling time was  $\Delta t = 0.02s$ .





where  $\Delta(x) = [x - \Delta x/2, x + \Delta x/2]$ . Note that  $q_{\Delta x}(x) \rightarrow \delta(x)$ , when  $\Delta x \rightarrow 0$ . The estimated drift  $a_{\Delta t}(x)$  depends on the distribution of points falling into the bin  $\Delta(x)$  as follows

$$b_{\Delta t}(x)\Delta t + o(\Delta t) = \lim_{N \rightarrow \infty} \sum_{k=1}^N \mathbb{E} [Y_{n+1}^k - Y_n^k | Y_n^k \in \Delta(x)] \quad (2.45)$$

$$= \int_{\Delta_k} \int_{\mathbb{R}} (y - x) p_{\Delta t}(y|x) q_{\Delta x}(x) dy dx, \quad (2.46)$$

where  $p_{\Delta t}(y|x)$  is the pdf to find  $\mathbf{X}(t)$  at the point  $y$  at time  $t + \Delta t$  when it started at point  $x$  at time  $t$ . In the small time limit,

$$\begin{aligned} b_{\Delta x}(x) &= \lim_{\Delta t \rightarrow 0} \mathbb{E} \left[ \frac{Y_{n+1} - Y_n}{\Delta t} | Y_n \in \Delta_k \right] = \int_{\Delta_k} b(x) q(x) dx = \int_{\Delta_k} b(x) \frac{p(x)}{\int_{x-\Delta x/2}^{x+\Delta x/2} p(x) dx} \\ &= \frac{\int_{\Delta(x)} p(y) b(y) dy}{\int_{\Delta(x)} p(y) dy}. \end{aligned}$$

We shall now obtain a further approximation by using a Taylor's expansion of the function  $F(x) = \int_0^x p(s) a(s) ds$ . We obtain that

$$\int_{\Delta_k} p(x) b(x) dx = F(x - \Delta x/2) - F(x + \Delta x/2) = F'(x)(\Delta x) + \frac{2}{6} F^{(3)}(x) (\Delta x/2)^3 + o((\Delta x)),$$

which leads to the approximation:

$$b_{\Delta x}(x) = \frac{p(x)b(x)\Delta x + \frac{1}{3}(p(x)b(x))''(x)(\Delta x/2)^3}{p(x)\Delta x + \frac{1}{3}p''(x)(\Delta x/2)^3} \quad (2.47)$$

$$= b(x) \left( 1 + \left( \frac{b''(x)}{b(x)} + 2 \frac{b'(x)b'(x)}{b(x)p(x)} \right) \frac{(\Delta x)^2}{24} + o((\Delta x)^2) \right). \quad (2.48)$$

The formula in higher dimension is given by

$$\mathbf{b}_{\Delta x}(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \left( 1 + \left( \frac{\Delta \mathbf{b}(\mathbf{x})}{\mathbf{b}(\mathbf{x})} + 2 \frac{\nabla \partial(\mathbf{x}) \cdot \nabla \mathbf{b}(\mathbf{x})}{\mathbf{b}(\mathbf{x}) \partial(\mathbf{x})} \right) \frac{(\Delta x)^2}{24} + o((\Delta x)^2) \right). \quad (2.49)$$

We conclude that a discretization in bins of size  $\Delta x$  perturbs the drift recovery by a term  $(\Delta x)^2$ .

### Influence of the time discretization $\Delta t$ on the drift estimation

To study the consequences of a discrete sampling time on the reconstruction of the drift from SPTs, we focus on the one-dimensional OU-process

$$dx = -\lambda(x - \mu)dt + \sqrt{2D}dw. \quad (2.50)$$

where  $\lambda, \mu$  are fixed. A direct integration of equation (2.50) for  $s \leq t$  leads to

$$x(t) = x(s)e^{-\lambda(t-s)} + \mu(1 - e^{-\lambda(t-s)}) + e^{-\lambda(t-s)} \int_s^t \sqrt{2D}e^{\lambda u} dW_u. \quad (2.51)$$

and for two consecutive points  $x(t)$  and  $x(t + \Delta t)$ , we have

$$x(t + \Delta t) - x(t) = -(x(t) - \mu)(1 - e^{-\lambda\Delta t}) + e^{-\lambda\Delta t} \int_t^{t+\Delta t} \sqrt{2D} e^{\lambda u} dW_u. \quad (2.52)$$

Thus, when  $\Delta t$  is small, the drift at position  $x$  and resolution  $\Delta t$  is

$$\begin{aligned} b_{\Delta t}(x)\Delta t + o(\Delta t) &= E\left(x(t + \Delta t) - x(t) | x(t) = x\right) \\ &= \int_{\mathbb{R}} (y - x) p(x(t + \Delta t) = y | x(t) = x) dy \\ &= -(1 - e^{-\lambda\Delta t})(x - \mu). \end{aligned} \quad (2.53)$$

We conclude that at resolution  $\Delta t$ , the approximation error is

$$F(t) = \frac{1 - e^{-\lambda\Delta t}}{\lambda\Delta t} = 1 - \frac{\lambda}{2}\Delta t + o(\Delta t) \quad (2.54)$$

suggesting that the drift of an OU process is always under-estimated using the displacement estimator.

### Time and space discretization for Ornstein-Uhlenbeck process

We shall now evaluate the cumulative effect of a temporal  $\Delta t$  and spatial  $\Delta x$  discretization on the recovery of an OU-process. The spatial grid of size  $\Delta x$  and the drift at position  $x$  are estimated empirically using the points falling in the bin  $\Delta(x) = [x - \Delta x/2, x + \Delta x/2]$ . We start with the conditional steady-state distribution  $q_{\Delta}(x)$  of points falling in  $\Delta(x)$ , which is linked to the pdf  $p(x)$  of the OU-stationary process by

$$q_{\Delta}(x) = \frac{p(x)}{\int_{\Delta(x)} p(y) dy}. \quad (2.55)$$

The drift term from eq. (2.53) can be approximated as

$$\begin{aligned} b_{\Delta t, \Delta x}(x) &= \mathbb{E}\left(\frac{x(t + \Delta t) - x(t)}{\Delta t} | x(t) \in \Delta(x)\right) \\ &= \int_{\Delta} \int_{\mathbb{R}} \frac{y - x}{\Delta t} p(x(t + \Delta t) = y | x(t) = x) dy dx \\ &= -\frac{1 - e^{-\lambda\Delta t}}{\Delta t} \int_{\Delta(x)} (x - \mu) q_{\Delta}(x) dx, \end{aligned} \quad (2.56)$$

where  $\mu$  is the center of the OU-process and the stationary pdf is given by

$$p(x) = \sqrt{\frac{\lambda}{2\pi D}} e^{-\frac{\lambda}{2D}(x - \mu)^2}. \quad (2.57)$$

To estimate eq. (2.56), we use eq. (2.47). For this computation, we set  $\mu = 0$ . In that case, we have

$$\int_{\Delta(x)} y q_{\Delta(x)}(y) dy = \frac{\int_{\Delta(x)} y e^{-\frac{\lambda}{2D} y^2} dx}{\int_{\Delta(x)} e^{-\frac{\lambda}{2D} y^2} dy}. \quad (2.58)$$

In the small  $\Delta x$  approximation, we have

$$\begin{aligned} \int_{\Delta} y q_{\Delta}(y) dy &= x + \frac{\int_{-\Delta x/2}^{\Delta x/2} h e^{-\frac{\lambda}{2D}(x+h)^2} dh}{\int_{-\Delta x/2}^{\Delta x/2} e^{-\frac{\lambda}{2D}(x+h)^2} dh} \\ &= x + \frac{\int_{-\Delta x/2}^{\Delta x/2} h e^{-\frac{\lambda}{2D} h^2} e^{-\frac{\lambda}{D} x h} dh}{\int_{-\Delta x/2}^{\Delta x/2} e^{-\frac{\lambda}{2D} h^2} e^{-\frac{\lambda}{D} x h} dh} \\ &= x + \frac{-\frac{\lambda}{4D} x (\Delta x)^3 + o((\Delta x)^3)}{\Delta x + (\frac{\lambda^2 x^2}{D^2} - \frac{\lambda}{D})(\Delta x)^3 + o(\Delta x^3)} \\ &= x + \frac{-\frac{\lambda}{6D} x \Delta x^2 + o((\Delta x)^2)}{1 + (\frac{\lambda^2 x^2}{D^2} - \frac{\lambda}{D})(\Delta x)^2 + o((\Delta x)^2)} \\ &= x(1 - \frac{\lambda}{12D} (\Delta x)^2) + o((\Delta x)^2). \end{aligned} \quad (2.59)$$

Using eqs. (2.56) and (2.59), we obtain an approximation for the drift at finite time step  $\Delta t$  and grid size  $\Delta x$

$$b_{\Delta t, \Delta x}(x) = -\frac{1 - e^{-\lambda \Delta t}}{\Delta t} (x - \mu - \frac{\lambda}{12D} (\Delta x)^2 + o((\Delta x)^2)). \quad (2.60)$$

To conclude, relation (2.60) reveals that the empirical displacements  $x(t + \Delta t) - x(t)$  collected over trajectories for an OU, can be used to recover the drift, with an additional exponential order correction in  $\Delta t$  and a second order in  $\Delta x$ .

### Empirical estimations of the drift

The empirical estimator  $\tilde{b}$  of the drift at position  $x$  for finite time  $\Delta t$  and spatial steps  $\Delta x$ , is defined

$$\tilde{b}_{\Delta t, \Delta x}(x) = \frac{1}{N} \sum_{i=1}^N \sum_{x^i(t_j) \in \Delta(x)} \frac{x^i(t_{j+1}) - x^i(t_j)}{\Delta t}, \quad (2.61)$$

where  $N$  is the number of points  $x^i(t_j)$  located in the segment  $[x - \Delta x/2, x + \Delta x/2]$ . Using eq. (2.60), the approximation at second order in  $\Delta x$  gives that

$$\tilde{b}(x) = -\frac{1 - e^{-\lambda \Delta t}}{\Delta t} (x_k - \mu - \frac{\lambda}{12D} \Delta x^2) + o(\Delta x^2). \quad (2.62)$$

To recover the parameter  $\lambda$  at various order of  $\Delta x$ , we can express  $\tilde{\lambda}$  using a regular series expansion

$$\tilde{\lambda}(\Delta t) = \lambda_0 + \lambda_2 (\Delta x)^2 + \lambda_4 (\Delta x)^4 + \dots \quad (2.63)$$

The first term is obtained by setting  $\Delta x = 0$  in eq. (2.62), leading for any bin centered around  $x_k$  for  $k = ..N_b$  to

$$\tilde{b}(x_k) = -\frac{1 - e^{-\lambda_0 \Delta t}}{\Delta t}(x_k - \mu). \quad (2.64)$$

In that case, a linear regression method can be used using the two coordinates  $x_k - \mu$  and  $\tilde{b}$  and to fit the distribution with a line and invert eq. (2.64). If necessary, the next in the expansion can be found. Note that a formal inversion of eq. (2.64) shows that for each k, we can obtain an estimation for  $\lambda_0$ :

$$\lambda_0^k(\Delta t) = -\frac{1}{\Delta t} \log\left(1 - \frac{\tilde{b}(x_k)\Delta t}{x_k - \mu}\right). \quad (2.65)$$

so that

$$\hat{\lambda}_0(\Delta t) = -\sum_{k=1}^{N_b} \frac{1}{\Delta t} \log\left(1 - \frac{\tilde{b}(x_k)\Delta t}{x_k - \mu}\right), \quad (2.66)$$

showing that numerical fluctuations in  $\tilde{b}(x_k)$  for  $|x_k - \mu|$  small can drastically affect the estimation. We use this result to study the recovery of the parameters in Fig. 2.5C and Fig. 2.11C, where we indeed observe larger errors near the center of the well than inside. We refer to Fig. 2.6 for the estimation of the eigenvalue with and without the center bin.

### Effect of the grids intersecting the boundary in the estimation of the drift

The recovery of a truncated OU involves estimating several parameters that depend on the accurate detection of the boundary. We focus here on the drift estimation for the part of the square grid that intersects the boundary (green bins in Fig. 2.6). In that case, for the interior part that intersects the elliptic domain, the empirical estimation recovers the local vector, while outside, it fluctuates around zero, due to the nature of the Brownian motion (no drift). Thus the error of the drift estimation at the boundary increases with the area fraction of the bin falling outside the domain. To estimate this error, we recall that the truncated OU-process is defined by

$$\dot{X} = -\nabla U(\mathbf{X})dt + \sqrt{2}B(\mathbf{X})\dot{w}, \quad (2.67)$$

where

$$U(\mathbf{X}) = \begin{cases} A \left[ \left(\frac{x-\mu_x}{a}\right)^2 + \left(\frac{y-\mu_y}{b}\right)^2 \right], & \text{if } \mathbf{X} \in \Gamma_{\mathcal{E}} \\ \mathcal{E} & \text{otherwise} \end{cases}. \quad (2.68)$$

Since the drift is zero for a diffusion process, located outside  $\Gamma_{\mathcal{E}}$ , we have

$$\begin{aligned} \mathbf{b}_{\Delta x}(\mathbf{x}) &= \lim_{\Delta t \rightarrow 0} \mathbb{E} \left[ \frac{Y_{n+1} - Y_n}{\Delta t} \mid Y_n \in \Delta(\mathbf{x}) \cap \Gamma_{\mathcal{E}} \right] \\ &= \int_{\Delta_{\frac{1}{k}}(\mathbf{x})} \mathbf{b}(\mathbf{y})q(\mathbf{y})d\mathbf{y}, \end{aligned} \quad (2.69)$$

where  $\Delta^1(\mathbf{x}) = \Delta(\mathbf{x}) \cap \Gamma_\varepsilon$  is the part of the grid interior to the ellipse. Indeed, the drift a diffusion process is zero. In addition, we suppose that the sample is made according to a normalized distribution

$$q_{\Delta x}(\mathbf{x}) = \frac{p(\mathbf{x})}{\int_{\Delta(x)} p(\mathbf{y}) d\mathbf{y}}, \quad (2.70)$$

where  $p(\mathbf{x})$  is any distribution that could be the steady-state distribution of a truncated OU inside the well and is uniform outside. From eq. (2.69), we get

$$\mathbf{b}_{\Delta x}(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \frac{\int_{\Delta^1(x)} p(\mathbf{x}) d\mathbf{x}}{\int_{\Delta(x)} p(\mathbf{x}) d\mathbf{x}}. \quad (2.71)$$

In first approximation,

$$\mathbf{b}_{\Delta x}(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \frac{\Delta^1(x)}{\Delta^1(x) + \Delta^2(x)}, \quad (2.72)$$

where  $\Delta^2(x) = \Delta(\mathbf{x}) \cap \Gamma_\varepsilon \Delta(\mathbf{x})$ . We shall now estimate  $\Delta^1(x)$ . We first note that the conservation of surfaces:  $\Delta^1(x) + \Delta^2(x) = (\Delta x)^2$ . For a square centered at a boundary of the ellipse  $(x, y)$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (2.73)$$

we consider the square grid with integer coordinates  $k = \lfloor \frac{x}{\Delta x} \rfloor$  and  $q = \lfloor \frac{y}{\Delta x} \rfloor$ . To compute  $\Delta^2(x)$ , we subtract the total area of the rectangle

$$S^{(2)} = [\Delta x][(q+1)\Delta x] \quad (2.74)$$

to the surface underneath the ellipse between the point  $k\Delta x$  and  $\Delta x + k\Delta x$ :

$$S^{(1)} = \int_{k\Delta x}^{(k+1)\Delta x} b \sqrt{1 - \left(\frac{u}{a}\right)^2} du \quad (2.75)$$

$$\approx b \sqrt{1 - \frac{k^2(\Delta x)^2}{a^2}} \left(1 - \frac{1}{2} \frac{k\Delta x^2}{a^2}\right). \quad (2.76)$$

Thus the computations lead to

$$S^{(2)} - S^{(1)} = \frac{1}{2} \frac{k\Delta x^2}{a^2} + o(\Delta x^2) \quad (2.77)$$

and for  $x > \Delta x$ , we get

$$\Delta^2(x) = S^2 - S^1 = \frac{xy}{2a^2} (\Delta x)^2 + o(\Delta x^2). \quad (2.78)$$

A similar computation leads for  $0 < x \leq \Delta x$  to

$$\Delta^2(x) = S^1 - S^2 = (\Delta x - x)y + o(\Delta x^2). \quad (2.79)$$

To conclude, except for the bin at the four extreme positions of the ellipse, the error is of order  $O((\Delta x)^2)$  for each grid bin, leading to a cumulative error along the total length of  $O((\Delta x))$ . The error contribution is shown in Fig. 2.6E.

### 2.5.4 Influence of the time steps in stochastic simulations

In the Smoluchowski's limit of the Langevin equation, the first order stochastic equation from which trajectories are generated is obtained by choosing a time step  $\delta t$ . This time step should not be smaller than the reciprocal of the friction coefficient  $\gamma$  so that the successive points  $x(\delta t), x(2\delta t), \dots, x(n\delta t), \dots$  should approximate the physical trajectory. When the sampling rate is such that  $\Delta t \gg \delta t$ , we can compare the drift estimation in that case and also study the extreme case when the sampling and simulation time steps are identical, leading to a jump process. When  $\Delta t \gg \delta t$ , the drift is computed after  $n$  steps. Using the empirical estimator, in the limit of a large number of trajectories  $N$ , we get

$$\begin{aligned} b_{\Delta t}^{est}(x)\Delta t + o(\Delta t) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=1}^N (x^m(t + p\delta t) - x^m(t + (p-1)\delta t)) + \dots (x^m(t + \delta t) - x^m(x)) \\ &\approx \int_0^{n\delta t} \mathbb{E}[b(x(s)) | x(0) = x] ds \\ &= \int_0^{\Delta t} \int_y b(y) p_s(y|x) ds dy, \end{aligned} \quad (2.80)$$

where  $p_s(y|x)$  is the pdf of the process  $x(t)$  starting at  $x$  at time  $t = 0$  and ending at  $y$  at time  $s$ . This result is quite different from the classical estimator of equation (2.45) and is very different from the estimation from a single observation time step  $\Delta t$ . Using the pdf an Ornstein-Uhlenbeck process

$$p_s(y|x) = \sqrt{\frac{\lambda}{2\pi D(1 - e^{-2\lambda s})}} \exp\left\{-\frac{\lambda}{2D} \frac{(y - xe^{-\lambda s})^2}{1 - e^{-2\lambda s}}\right\}, \quad (2.81)$$

and the change of variable  $u = \frac{(y - xe^{-\lambda s})^2}{\sqrt{\frac{2D}{\lambda}(1 - e^{-2\lambda s})}}$  we get

$$\int_0^{\Delta t} \int_y b(y) p_s(y|x) ds dy = \int_0^{\Delta t} \int_u b(xe^{-\lambda s} + \sqrt{\frac{2D}{\lambda}(1 - e^{-2\lambda s})}u) \frac{1}{\sqrt{\pi}} \exp\{-u^2\} du ds. \quad (2.82)$$

Using a Taylor's expansion in the drift term:

$$\begin{aligned} b(xe^{-\lambda s} + \sqrt{\frac{2D}{\lambda}(1 - e^{-2\lambda s})}u) &= \\ b(xe^{-\lambda s}) + \sqrt{\frac{2D}{\lambda}(1 - e^{-2\lambda s})}ub'(xe^{-\lambda s}) &+ \frac{1}{2} \frac{2D}{\lambda} (1 - e^{-2\lambda s})u^2 b''(xe^{-\lambda s}) + o((1 - e^{-2\lambda s}))^2. \end{aligned}$$

Thus the estimator of the drift with many time steps is at first order in  $\Delta t$  given by

$$\begin{aligned} \int_0^{\Delta t} \int_y b(y) p_s(y|x) ds dy &= \int_0^{\Delta t} \left( b(xe^{-\lambda s}) + \frac{1}{2} \frac{D}{\lambda} (1 - e^{-2\lambda s}) b''(xe^{-\lambda s}) \right) ds \\ &\approx b(x)\Delta t + (-\lambda x b'(x) + D b''(x)) \frac{(\Delta t)^2}{2} + o((\Delta t)^2) \\ &\approx -\lambda x \Delta t + \lambda^2 x \frac{(\Delta t)^2}{2} + o((\Delta t)^2). \end{aligned}$$

We now evaluate the consequences of simulating a process with many time steps  $\delta t = 10^{-4}s$  whereas the sampling time was  $\Delta t = 0.02s$  in Fig. 2.5. We show in Fig. 2.11 how the drift field can be recovered. This situation corresponds to large jumps of the underlying physical process. To conclude, we find that the center and peripheral grid bins are generating most of the error, especially for large grid sizes ( $\Delta x = 50$  and  $90$  nm).

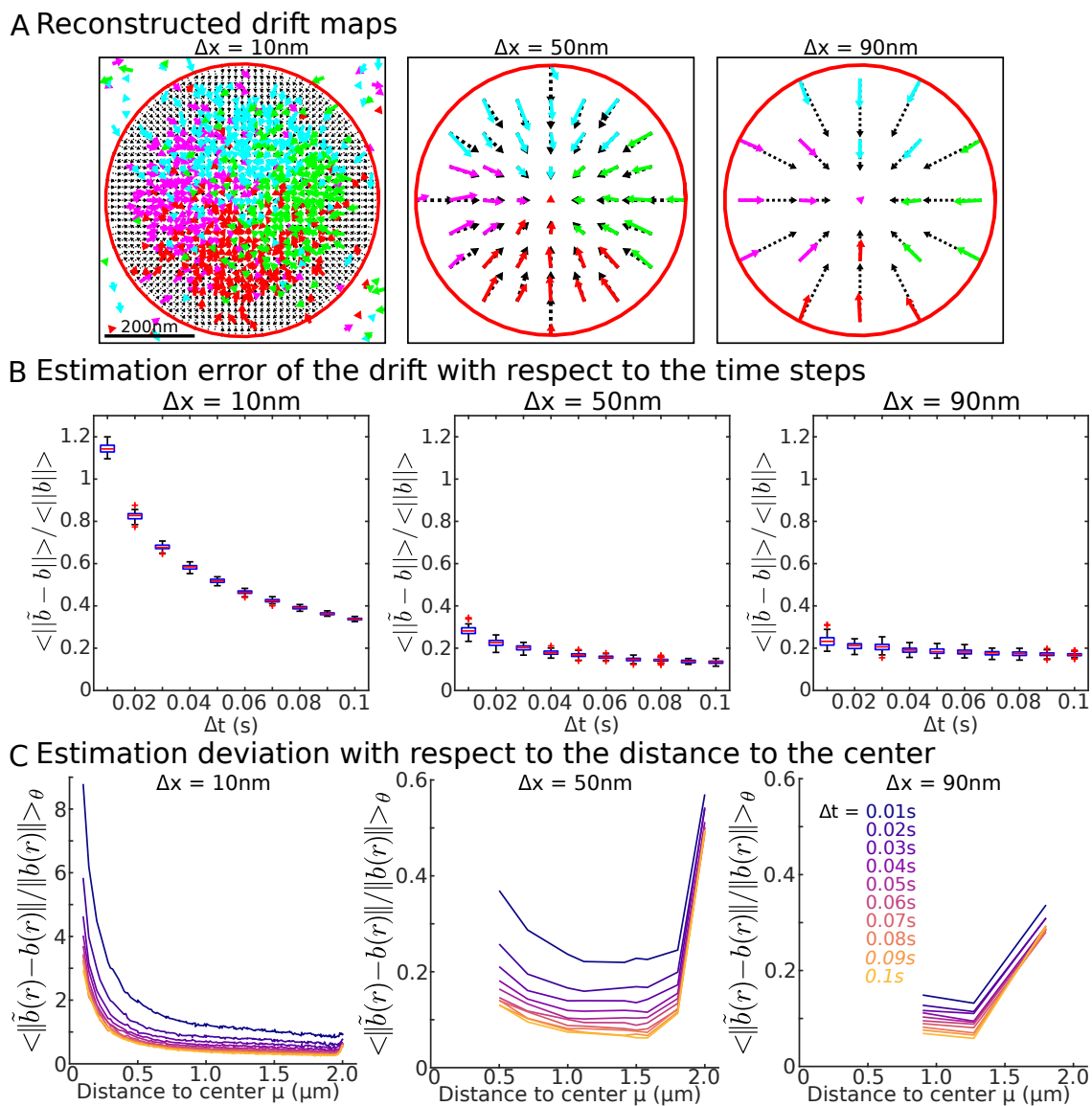


Figure 2.11: **Recovering the vector field with an equal simulated and sampled time step  $\delta t = \Delta t$ .** **A.** Recovering the local drift field inside a circular well for different grid sizes (10 nm, 50 nm, 90 nm) using numerical simulations with a sampling  $\Delta t = 20$  ms, with the constraints that at least 10 points falls inside a bin. **B.** Error between the true and observed fields averaged over all the square bins inside the well versus the time step  $\Delta t$ . **C.** Error between the true and observed fields averaged over the radial angle versus the distance  $r$  to the center for various timestep (see color code).

### 2.5.5 Conditional drift estimation

In this last section, we discuss the effect of estimating the drift by conditioning the end points of a displacement to stay inside the potential well. Computing the displacement  $\Delta X$  by selecting only trajectories that stay inside the well gives a bias estimator of the drift. This situation appears when the trajectories never reach the boundary. The drift estimator is computed from the conditional pdf  $p^*$  of the process that stays inside the potential well. To find such a drift, we introduce the probability that a stochastic particle hits a ball of radius  $\epsilon$  centered on the well before escaping from the well [115], then

$$p^*(x, y, t) = p(x, y, t) \frac{q(x)}{q(y)}, \quad (2.83)$$

where  $p$  is the pdf in the entire space.  $q$  is solution of

$$L^*(q) = 0 \quad (2.84)$$

$$q = 0 \text{ on } \partial W \quad (2.85)$$

$$q = 1 \text{ on } \partial B_\epsilon, \quad (2.86)$$

$L^*$  is the backward Fokker-Planck equation associated the process  $X(t)$  (p.77 [115]), defined by

$$dx = b(x)dt + \sqrt{2D}dw. \quad (2.87)$$

In that case,

$$\lim_{\Delta t \rightarrow 0} E \left[ \frac{x(t + \Delta t) - x(t)}{\Delta t} \middle| x(t) = x \right] = b(x) + \sqrt{2D} \frac{\nabla q(x)}{q(x)}. \quad (2.88)$$

To conclude, by restricting the computation of the displacements to empirical trajectories that only remain in the well, an additional term has to be accounted for, which diverges as the distance from the point  $x$  to the boundary tends to zero (Fig. 2.12).



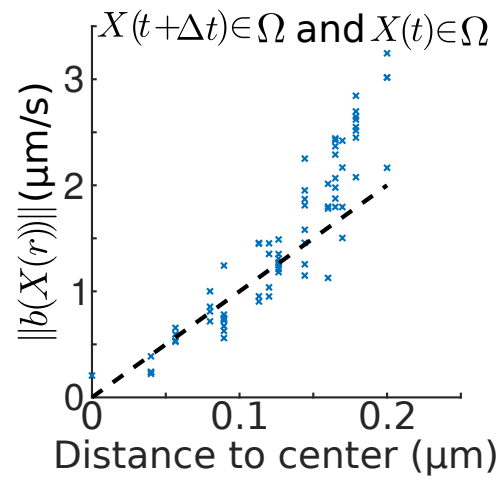


Figure 2.12: **Conditional reconstructed drift versus distance to the center.** We estimated the drift for displacements that do not exit the well (assuming the boundary is known).

## Chapter 3

# Transient Confinement of $\text{Ca}_V2.1$ $\text{Ca}^{2+}$ -Channel Splice Variants Shapes Synaptic Short-Term Plasticity

Published as:

Jennifer Heck\*, Pierre Parutto\*, Anna Ciuraszkiewicz\*, Arthur Bikbaev, Romy Freund, Jessica Mitlöhner, Maria Alonso, Anna Fejtova, David Holcman and Martin Heine, "*Transient Confinement of  $\text{Ca}_V2.1$   $\text{Ca}^{2+}$ -Channel Splice Variants Shapes Synaptic Short-Term Plasticity*", *Neuron* volume 103, pages 66-79 (2019).

(\*: equally)

**Abstract** The precision and reliability of synaptic information transfer depend on the molecular organization of voltage-gated calcium channels (VGCCs) within the presynaptic membrane. Alternative splicing of exon 47 affects the C-terminal structure of VGCCs and their affinity to intracellular partners and synaptic vesicles (SVs). We show that hippocampal synapses expressing VGCCs either with exon 47 ( $\text{Ca}_V2.1_{+47}$ ) or without ( $\text{Ca}_V2.1_{\Delta 47}$ ) differ in release probability and short term plasticity. Tracking single channels revealed transient visits ( $\approx 100$  ms) of presynaptic VGCCs in nanodomains ( $\approx 80$  nm) that were controlled by neuronal network activity. Surprisingly, despite harboring prominent binding sites to scaffold proteins,  $\text{Ca}_V2.1_{+47}$  persistently displayed higher mobility within nanodomains. Synaptic accumulation of  $\text{Ca}_V2.1$  was accomplished by optogenetic clustering, but only  $\text{Ca}_V2.1_{+47}$  increased transmitter release and enhanced synaptic short-term depression. We propose that exon 47-related alternative splicing of  $\text{Ca}_V2.1$  channels controls synapse-specific release properties at the level of channel mobility-dependent coupling between VGCCs and SVs.

### 3.1 Introduction

Neurons communicate via chemical synapses that can change their transmission properties on a short timescale, thereby directly affecting the computational performance of neuronal networks [116, 117]. Following the release of synaptic vesicles (SVs), the number of activated postsynaptic receptors is the effective readout. Thus, the weight of a given synapse within a neuronal network strongly depends on two parameters: the number of release sites for SVs within the active zone and the release probability of individual SVs. The number of release sites harboring readily releasable SVs within the active zone mainly depends on the molecular composition and interactions within the presynaptic bouton [118, 119, 120]. The release probability in response to an action potential (AP) is determined by the coupling distance between voltage-gated calcium channels (VGCCs) and vesicular calcium sensors [121, 122]. Upon AP-mediated transient opening of VGCCs, the local influx of calcium ions triggers the fusion of SVs and neurotransmitter release. With extremely tight coupling, the action of a single VGCC may be sufficient to initiate this event [123, 124, 125]. Several parameters such as conductance of VGCCs, average width of the AP, calcium affinity of vesicular calcium sensors, abundance, affinity of endogenous buffering proteins, and calcium extrusion mechanisms constrain the spatial signaling capacity of VGCCs [122]. Hence, the local density of VGCCs and their distance to releasable SVs define the probability of transmitter release. Synapses characterized by loose coupling (20-100 nm) between VGCCs and SVs have been described to exhibit a low release probability [126, 121, 127, 128, 129], whereas synapses with tight coupling ( $< 20$  nm) were found to have high release probabilities [123, 83, 130]. Direct measurements and modeling of calcium channel densities within the presynaptic membrane, in combination with functional analysis, have identified a heterogeneous distribution of VGCCs, suggesting the existence of local channel clusters assigned to specific release sites [131, 132, 133, 134, 135, 136, 124]. However, despite the need for nanometer-precise coupling between VGCCs and SVs, the stability of molecular complexes changes over time, considering ongoing fusion and retrieval of SVs. This permanent reconstruction of the presynaptic plasma membrane suggests a flexible arrangement of SVs and associated proteins embedded in a more stable core scaffold [137]. Direct indicators for presynaptic glutamate release revealed that several release sites exist within the active zone of small cortical and hippocampal synapses, providing the physical basis for multiple vesicular release [138, 119]. Altogether, the small number of synaptic VGCCs, their kinetic properties, and the fluidity of the membrane argue against a stable nanometer-tight connection between VGCCs and SVs over time. In previous experiments, we have shown that  $Ca_v2.1$  and  $Ca_v2.2$ , the predominant presynaptic VGCCs in the mammalian brain [139], are dynamically organized within the synapse, suggesting a flexible contribution of individual VGCCs to SV release and short-term synaptic plasticity [32]. It has been proposed that the attachment of VGCCs to individual release sites is, besides interactions of the synprint region [140, 141], mainly achieved via C-terminal binding of VGCCs to the Rab3-interacting molecule (RIM) and the RIM binding protein (RBP) [142, 143, 144]. However, alternative splicing of VGCCs can significantly change C-terminal structure and thereby alter their connectivity to scaffold proteins [145, 146]. Calcium channels can either express (+47) or lack ( $\Delta 47$ ) exon 47, which encodes about 150-250 amino acid residues, depending on the species and channel isoform.

Here, we found that these differences between the C termini of  $\text{Ca}_V2.1$  splice variants are not critical for the incorporation of channel complexes into the presynaptic membrane as proposed for  $\text{Ca}_V2.2$  channels [147] but rather influence their dynamic organization within the presynaptic membrane and the release probability of SVs. Acute changes in the dynamics of individual  $\text{Ca}_V2.1$  channels affect release probability and short-term plasticity exclusively for the long C-terminal variant, suggesting a transient character of splice variant-specific coupling of  $\text{Ca}_V2.1$  channels to SVs.

## 3.2 Results

### 3.2.1 Expression and Synaptic Localization of $\text{Ca}_V2.1$ C-Terminal Splice Variants

Previous investigations [148, 149] have indicated that both, the short and the long C-terminal splice variant of  $\text{Ca}_V2.1$  ( $\text{Ca}_V2.1_{\Delta 47}$  and  $\text{Ca}_V2.1_{+47}$ , respectively), are expressed in the mammalian brain. Western blot analysis performed on the plasma membrane fraction of rat hippocampus and cortex confirmed these results (Figure 3.1A-C). The molecular weight of heterologously expressed  $\text{Ca}_V2.1_{+47}$  and  $\text{Ca}_V2.1_{\Delta 47}$  splice variants corresponds to the predicted molecular weight of  $\approx 243$  kDa for rat  $\text{Ca}_V2.1_{\Delta 47}$  and  $\approx 260$  kDa for rat  $\text{Ca}_V2.1_{+47}$  (according to the amino acid sequences NCBI: NP\_037050.2). Comparing the ratios between  $\text{Ca}_V2.1_{\Delta 47}$  and  $\text{Ca}_V2.1_{+47}$  during development revealed that the long C-terminal  $\text{Ca}_V2.1$  splice variant is dominant in older animals.

To explore the synaptic distribution of the long and short  $\text{Ca}_V2.1$  splice variant, we created N-terminally GFP-tagged constructs to observe the localization of calcium channels within the synapse. The species specificity of  $\text{Ca}_V2.1$  antibodies allowed us to distinguish between the tagged channel population (derived from human  $\text{Ca}_V2.1$ ) and the endogenous  $\text{Ca}_V2.1$  channels in rat hippocampal neurons (Figures 3.7A-F). Testing N-terminally tagged  $\text{Ca}_V2.1$  channels in HEK293T cells confirmed that the  $\text{Ca}_V2.1$  C-terminal splice variants did not differ in their voltage-dependent kinetic properties; in addition, channel surface expression was unaffected by N-terminal tagging (Figure S2) [32]. In general, the expression of  $\text{Ca}_V2.1_{+47}$  or  $\text{Ca}_V2.1_{\Delta 47}$  in rat hippocampal neurons increased the total population of  $\text{Ca}_V2.1$  channels within the synapse (Figure 3.1F). Using a specific antibody detecting the endogenous rat  $\text{Ca}_V2.1$  channel revealed that 50% of transfected synapses were populated exclusively by the overexpressed  $\text{Ca}_V2.1$  channel. The other 50% of synapses expressed a mixed population of endogenous and tagged  $\text{Ca}_V2.1$  channels (Figures 3.7F-H). Presynaptic compartments were identified by co-staining of Bassoon or RIM. Both  $\text{Ca}_V2.1$  C-terminal splice variants were found to co-localize with the presynaptic marker Bassoon, similar to non-transfected synapses, whereby the  $\text{Ca}_V2.1_{+47}$  splice variant showed slightly better synaptic localization (Figure 3.1D-F; Figure 3.7G), confirming previous studies [150, 149, 151, 32, 85]. Normalization of the fluorescence intensity of pre- and postsynaptic marker proteins to non-transfected synapses revealed a general accumulation of these proteins when expressing the  $\text{Ca}_V2.1_{\Delta 47}$  or  $\text{Ca}_V2.1_{+47}$  splice variant. This enlargement of presynaptic boutons supports previous observations showing that the calcium channel number in the presynapse scales with scaffold

proteins, and most likely vice versa [132, 32]. However, the expression of the short  $Ca_v2.1_{\Delta 47}$  splice variant resulted in stronger up-scaling of synapses compared to  $Ca_v2.1_{+47}$ -expressing synapses (Figure 3.1F). This stronger accumulation of presynaptic proteins and even the postsynaptic scaffold PSD95 in  $Ca_v2.1_{\Delta 47}$  dominated synapses was surprising, because this splice variant does not express canonical C-terminal scaffold interaction sites (Figure 3.1A).

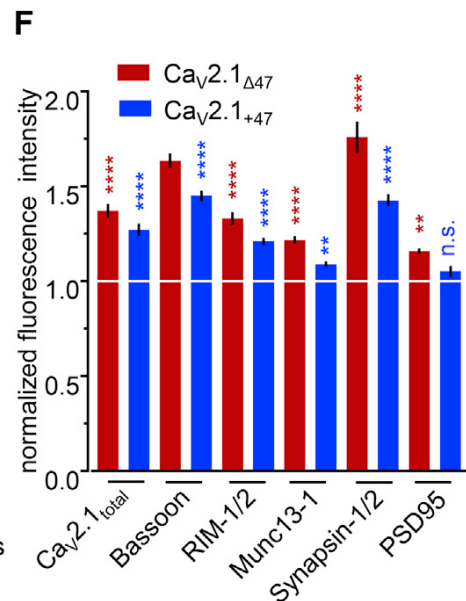
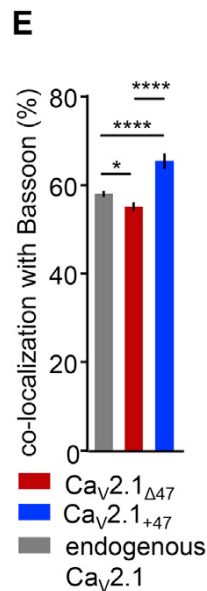
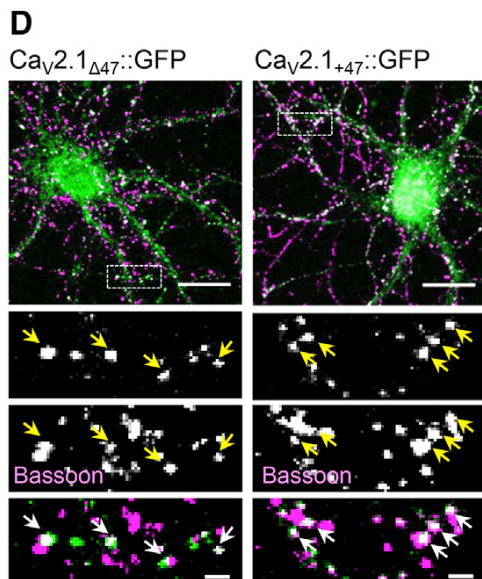
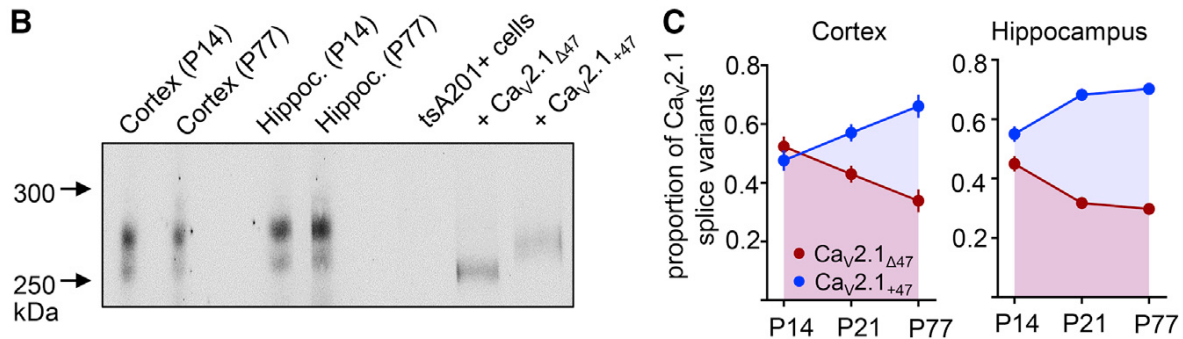
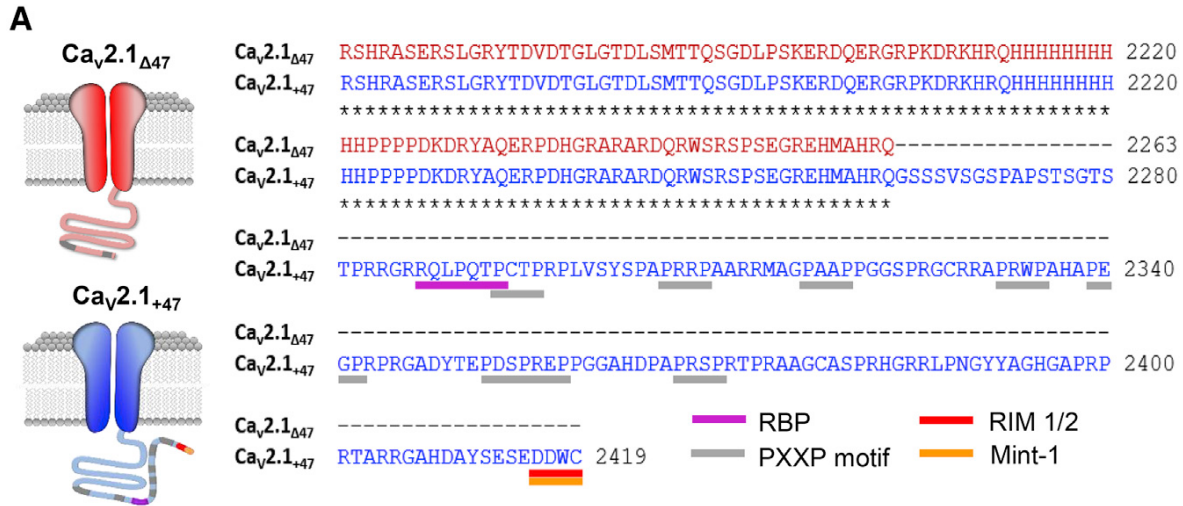


Figure 3.1: (*previous page*) **Expression of C-Terminal Ca<sub>v</sub>2.1 Splice Variants and Synaptic Targeting of Overexpressed Ca<sub>v</sub>2.1<sub>Δ47</sub> and Ca<sub>v</sub>2.1<sub>+47</sub>.** **A.** Depiction of C-terminal Ca<sub>v</sub>2.1 splice variants Δ47 (red) and +47 (blue), illustrating the loss of distal C-terminal interaction sites for the Δ47 splice variant. **B.** Western blot showing the endogenous expression of Ca<sub>v</sub>2.1<sub>Δ47</sub> and Ca<sub>v</sub>2.1<sub>+47</sub> in the plasma membrane fraction of juvenile and adult rat cortex (postnatal day [P] 14 and P77, respectively) and hippocampus (P14 and P77, respectively). Antibody specificity was tested on lysates from a tsa201 cell line, expressing Ca<sub>v</sub>2.2/β3/α2δ1 as negative control, which was transfected with Ca<sub>v</sub>2.1<sub>Δ47</sub> or Ca<sub>v</sub>2.1<sub>+47</sub> to probe the differences in the molecular weight. **C.** Quantification of the Ca<sub>v</sub>2.1<sub>Δ47</sub> to Ca<sub>v</sub>2.1<sub>+47</sub> ratios in male P14, P21, and P77 rat cortex and hippocampus. For data and statistics, see Table S1 in original article. **D.** Images of hippocampal neurons 16 days in vitro (DIV) transfected with either Ca<sub>v</sub>2.1<sub>Δ47</sub>::GFP (left) or Ca<sub>v</sub>2.1<sub>+47</sub>::GFP (right). Presynaptic boutons are identified by the intensity of anti-Bassoon immunoreactivity (magenta). Arrows indicate co-localization of calcium channels (green) and presynaptic structures. Scale bars correspond to 10 mm (overview) or 2 mm (selection). **E.** Quantification of the co-localization of Ca<sub>v</sub>2.1 channels (Ca<sub>v</sub>2.1<sub>Δ47</sub>, red; Ca<sub>v</sub>2.1<sub>+47</sub>, blue; Ca<sub>v</sub>2.1, gray) with Bassoon in hippocampal neurons 16 DIV. For data and statistics, see Table S1 in original article. **F.** Effects of the expression of Ca<sub>v</sub>2.1<sub>Δ47</sub> (red) or Ca<sub>v</sub>2.1<sub>+47</sub> (blue) on the total expression of Ca<sub>v</sub>2.1, Bassoon, RIM, Munc13, Synapsin, and PSD95. The fluorescence intensities are normalized to non-transfected synapses within the same field of view. For data and statistics, see Table S1 in original article.

### 3.2.2 Functional Differences of Synapses Dominated by Ca<sub>v</sub>2.1<sub>+47</sub> or Ca<sub>v</sub>2.1<sub>Δ47</sub> Channels

We measured presynaptic calcium responses, glutamate release, and transmission properties of monosynaptically coupled neurons to probe the effect of C-terminal Ca<sub>v</sub>2.1 splice variants (Figure 3.2). Co-expression of RFP-tagged Ca<sub>v</sub>2.1 channels with the presynaptic-targeted calcium sensor GCaMP5G::synaptophysin (Figures 3.9A-D) enhanced the evoked presynaptic calcium responses as expected from the increase of immunofluorescence labeling of synaptically localized Ca<sub>v</sub>2.1 channels (Figures 3.1F and 3.2A-C). Here, synapses expressing Ca<sub>v</sub>2.1<sub>+47</sub> channels showed the strongest increase of calcium in response to a pattern of AP-like electrical field stimuli (Figure 3.2C). Next, we co-expressed iGluSnFR [152], a fluorescent glutamate sensor, with RFP-tagged Ca<sub>v</sub>2.1 splice variants to directly probe presynaptic glutamate release. Despite recoding the strongest calcium signal in Ca<sub>v</sub>2.1<sub>+47</sub>-dominated synapses, the fluorescence response of the glutamate sensor to a single stimulus was similar to that of non-transfected synapses (Figures 3.2D-F). More surprisingly, the amplitude of the iGluSnFR fluorescence signal of synapses expressing Ca<sub>v</sub>2.1<sub>Δ47</sub> channels was significantly enhanced, indicating increased glutamate release in comparison to non-transfected or Ca<sub>v</sub>2.1<sub>+47</sub>-dominated synapses (Figure 3.2F). These results were confirmed by measuring excitatory postsynaptic currents (EPSCs) between monosynaptically coupled transfected and non-transfected neurons in paired whole-cell patch-clamp recordings (Figures 3.2H,I). We isolated glutamatergic responses by bath application of 5 mM bicuculline to examine whether the overexpression of either Ca<sub>v</sub>2.1<sub>Δ47</sub> or Ca<sub>v</sub>2.1<sub>+47</sub> alters the ampli-

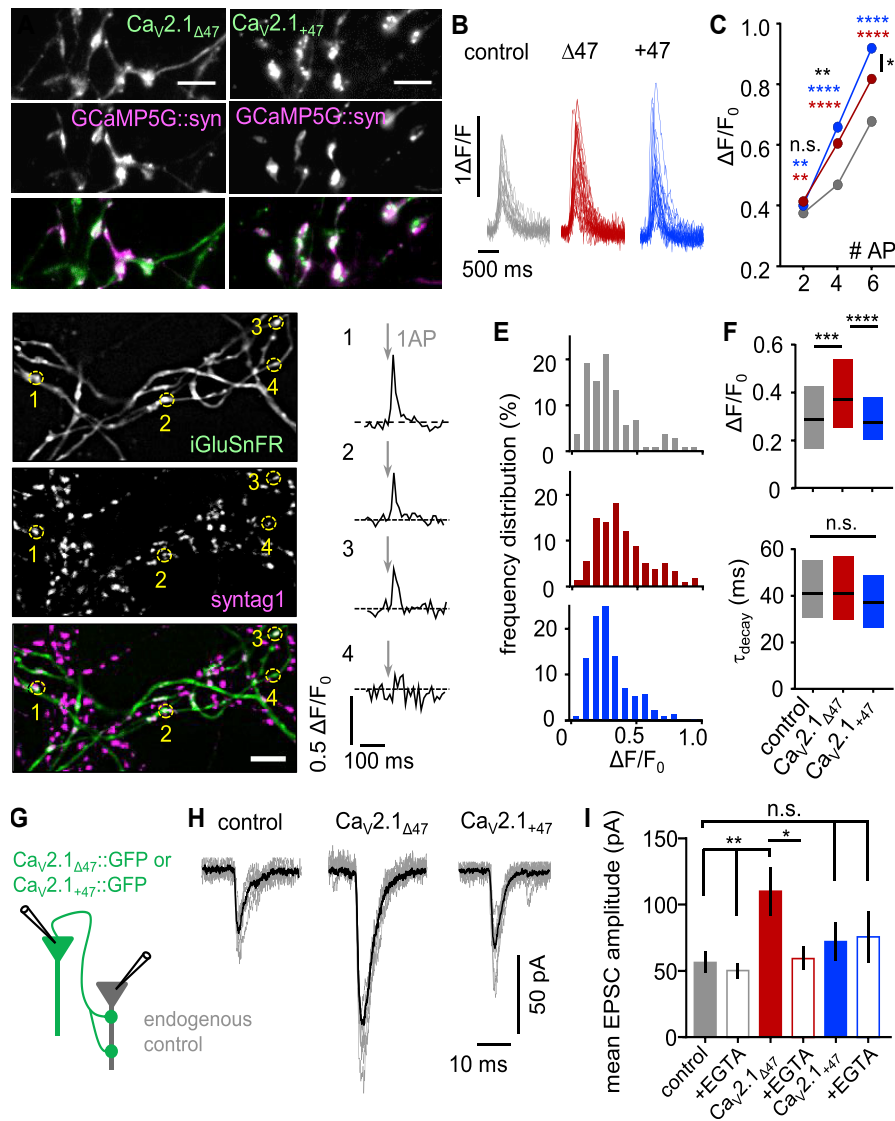
tude or variability of EPSCs. The EPSC amplitude was larger for  $\text{Ca}_v2.1_{\Delta 47}$ -expressing synapses, supporting an enhancement of neurotransmitter release as demonstrated by glutamate imaging. EPSC amplitudes of synapses dominated by  $\text{Ca}_v2.1_{+47}$  were comparable with non-transfected neurons. Application of the  $\text{Ca}^{2+}$  chelator EGTA (10 mM in the pipette solution) resulted in a reduction of the  $\text{Ca}_v2.1_{\Delta 47}$ -mediated rise of the EPSC amplitude to control level, suggesting more distant localization of  $\text{Ca}_v2.1_{\Delta 47}$  channels to SVs. In contrast, no buffering effect was observed for  $\text{Ca}_v2.1_{+47}$ -dominated synapses or the control condition (Figures 3.2G-I). The kinetic properties of EPSCs and their variability within trains of stimuli were not affected (Figures 3.9E-F).

Furthermore, we tested whether channel overexpression affects the size of the available vesicular pool and vesicle recycling. We used the activity-driven uptake of CypHer5E-labeled antibodies against the luminal domain of synaptotagmin1 (syntag1-CypHer5E) [153] and monitored the fluorescence intensity, the quenching of the fluorescence during evoked SV release, and the fluorescence recovery as a readout for SV endocytosis [153]. The SV release was induced by 200 AP-like stimuli (pulse duration of 1 ms at 20 Hz; 2 mM  $\text{Ca}^{2+}$ ) (Figure 3.10) in the presence of 1  $\mu\text{M}$   $\omega$ -conotoxin-GVIA to isolate  $\text{Ca}_v2.1$ -mediated SV release. Analysis of the amplitude and kinetics of the fluorescence signal revealed that time to peak, as well as maximal quenching of the fluorescence, differed between transfected and control synapses but was similar between the two C-terminal splice variants. The fluorescence recovery was not altered (Figure 3.10G), indicating no change in the endocytosis of SVs.

Altogether, the overexpression of  $\text{Ca}_v2.1$  C-terminal splice variants generally increased the presynaptic calcium influx. However, only  $\text{Ca}_v2.1_{\Delta 47}$  was potent to induce larger glutamate release and EPSCs. These data suggest a different initial release probability for synapses depending on the expression of  $\text{Ca}_v2.1$  splice variants. We hypothesize that the difference in the C-terminal structure of  $\text{Ca}_v2.1$  influences their coupling to scaffold proteins and thus to SVs and is eventually accompanied by changes in their local surface mobility.

### 3.2.3 Dynamics of $\text{Ca}_v2.1$ Channels Differ between C-Terminal Splice Variants

Using single-particle tracking photo-activation localization microscopy (sptPALM), we analyzed the dynamics of individual  $\text{Ca}_v2.1$  variants in living neurons to test the idea that their different affinities to scaffold proteins affect their local dynamics. Expression of mEOS-tagged  $\text{Ca}_v2.1$  variants, together with  $\text{GCaMP5G}::\text{synaptophysin}$  or activity-driven uptake of anti-syntag1 antibodies (Figures 3.3A-C) allowed us to identify presynaptic structures, localize tagged channels, and evaluate their function within the presynaptic membrane. Although, the N-terminal tagging of  $\text{Ca}_v2.1$  channels does not allow to distinguish between intracellular and surface-expressed channels, several aspects indicate that most tagged channels are located in the presynaptic membrane. As reported earlier, in 50% of synapses expressing tagged channels, we found full replacement of the endogenous channel population (Figure 3.7H). Second, an increase in the number of synaptic  $\text{Ca}_v2.1$  channels induced up-scaling of presynaptic scaffold proteins (Figure 3.1F), which is in line with previous data showing a positive correlation between VGCC number and scaffold proteins [132]. Third, elevated



**Figure 3.2: Effects of Alternative Splicing of the  $Ca_v2.1$  C Terminus on Synaptic Transmission.** **A.** Images of axonal expression of  $Ca_v2.1_{\Delta 47}::RFP$  (left) or  $Ca_v2.1_{+47}::RFP$  (right), together with  $GCaMP5G::synaptophysin$  in 15 DIV hippocampal neurons.  $GCaMP5G::synaptophysin$  (magenta) indicates presynaptic compartments. Scale bar corresponds to  $5\mu m$ . **B.** Fluorescence changes of synapses expressing  $GCaMP5G::synaptophysin$  in response to 4 AP-like field stimuli, transfected with  $GCaMP5G::synaptophysin$  only (gray) or with  $Ca_v2.1_{\Delta 47}$  (red) or  $Ca_v2.1_{+47}$  (blue). **C.** Medians of presynaptic calcium signals ( $\Delta F/F_0$ ) from synapses expressing  $GCaMP5G::synaptophysin$  only or with  $Ca_v2.1$  splice variants (color code as indicated in B). Synaptic calcium influx was evoked with 2, 4, and 6 AP-like stimuli, which are in the linear range of the used calcium sensor  $GCaMP5G$  (Figures 3.9A-D). For values and statistics, see Table S2 in original article. **D.** Axonal expression of the glutamate sensor  $iGluSnFR$  and presynaptic labeling using  $syntag1::CypHer5E$  antibody (AB) uptake in hippocampal cultures 15 DIV. Labeled regions show  $iGluSnFR$ -expressing presynaptic compartments with corresponding changes in fluorescence intensity in response to 1 AP-like field stimulation. Scale bar corresponds to  $10\mu m$ . **E.** Frequency distribution of evoked (1 AP) glutamate responses from synapses expressing  $iGluSnFR$  only (gray) or with  $Ca_v2.1_{\Delta 47}::RFP$  (red) or  $Ca_v2.1_{+47}::RFP$  (blue). **F.** Median and IQR for amplitude ( $\Delta F/F_0$ ) and decay time constant ( $\tau_{decay}$ ). For values and statistics, see Table S2 in original article. **G.** Illustration of EPSCs of synaptic coupled neurons. **H.** Example traces of EPSCs from neurons (15-16 DIV) in control conditions or presynaptic expression of  $Ca_v2.1_{\Delta 47}::GFP$  or  $Ca_v2.1_{+47}::GFP$ . Average responses of repetitive stimuli (0.1 Hz) are given in black; individual responses in gray. **I.** Mean first EPSC amplitude and kinetic parameters of the EPSC (Figures 3.9E-F). For values and statistics, see Table S2 in original article.



presynaptic calcium transients (Figure 3.2C) likely reflect a larger population of functional channels in the presynaptic membrane. In addition, we examined whether translocation of channels in transport vesicles might bias our mobility analysis. Based on individual mean square displacement (MSD) curves, we identified a minor fraction of axonal  $\text{Ca}_V2.1$  channels (14%- 17%) showing directed movement, which could indicate an association of VGCCs with transport vesicles (Figures 3.10D- 3.11A). Within synapses, the fraction was reduced to only 3%-8%. Although the contribution of a confined intracellular fraction cannot be ruled out, the functional and structural data demonstrate that most tagged  $\text{Ca}_V2.1$  channels are expressed within the presynaptic membrane.

The diffusion coefficient differed between channels in the extrasynaptic and synaptic membrane but was comparable for both splice variants (Figure 3.3D). To explore  $\text{Ca}_V2.1$  dynamics within synapses in more detail, we applied a step analysis algorithm [17] and quantified the local confinement and dwell time of VGCCs. Within presynaptic membranes, we identified regions that fulfill the requirements of parabolic potential wells: regions that exhibit a drift vector fields converging toward a center point [154] given by equation (1.8) (Figures 3.3B,C). Potential wells, which we call nanodomains, are characterized as elliptic regions, defined by the following parameters: size, diffusion coefficient inside the well, confinement energy, and residence time of the channels within this region (estimation methods describes in paragraph 3.4.2). Surprisingly, the diffusion coefficient inside nanodomains was significantly larger for  $\text{Ca}_V2.1_{+47}$  channels compared to  $\text{Ca}_V2.1_{\Delta 47}$  channels, while the energies within the well were similar (Figure 3.3G) (Figures 3.11I-J). In general, the formation of nanodomains was observed with a higher incidence inside synaptic compartments but also occurred in the extrasynaptic membrane. The density of nanodomains along the axon was independent of the expressed splice variant (Figure 3.11J). On average, 2 nanodomains were detected within synapses ( $\text{Ca}_V2.1_{\Delta 47} = 2.1 \pm 1.8$  SD;  $\text{Ca}_V2.1_{+47} = 2.0 \pm 1.7$  SD). The size of confinement (characterized by the harmonic mean  $R$  of the two elliptic axes  $a > b$ :  $R = \sqrt{ab}$ ) was around  $R \approx 80$  nm and slightly varied between  $\text{Ca}_V2.1_{+47}$  and  $\text{Ca}_V2.1_{\Delta 47}$  ( $\text{Ca}_V2.1_{\Delta 47}$  median = 76, interquartile range [IQR] = 67/89 nm;  $\text{Ca}_V2.1_{+47}$  median = 81, IQR = 70/97 nm). Notably, the characteristic length scale of the detected synaptic wells agrees with the proposed trans-synaptic arrangement of pre- and postsynaptic scaffold proteins [84]. From the energy and the size of a nanodomain, the residence time of a calcium channel inside the well was calculated from equation (1.11). Individual  $\text{Ca}_V2.1$  channels transiently dwell for around 100 ms within these nanodomains, with no difference between the splice variants (Figure 3.3G). We compared these residence times of calcium channels inside nanodomains with the theoretical mean time  $\tau_B$  of a freely diffusing (non-interacting) channel to escape a region that was 80 nm in diameter. Given the MSD relation  $\langle ||X(t)||^2 \rangle = 4Dt$ , it resulted in a shorter mean diffusion residence time of  $\tau_B \approx 30ms$ . This result demonstrates that nanodomains stabilize individual channels in small regions about three times longer than pure diffusion. Furthermore, we applied the step analysis algorithm in a temporal sliding window (with a width of 20 s and an overlap of 10 s) to access the stability of nanodomains as calcium channel-confining structures. We found that nanodomains were transiently formed with a mean lifetime of about 30 s (Figure 3.3H; Figures 3.11E-F). In addition, the nanodomains could deform and slightly move within the presynaptic membrane (Figure 3.3H). One could speculate that these temporary nanodomains correspond to

local presynaptic calcium channel densities, as seen in freeze fracture electron microscopy pictures of the presynaptic membrane [132, 133]. Finally, we attempted to estimate the number of channels inside a potential well. For this purpose, the single-particle tracking (SPT) trajectories cannot be used directly. We estimated that the density of  $Ca_v2.1$  channels at active zones was  $\approx 400$  channels/ $\mu m^2$  [132]. Using the estimation of the potential well size (Figure 3.3G), we obtained  $\approx 9$  calcium channels inside one nanodomain (paragraph 3.4.2). If there are, on average,  $\approx 3$ -5 docking sites per active zone of hippocampal synapses [132], we conclude that on average, 2-5 channels will be associated with surrounding vesicles.

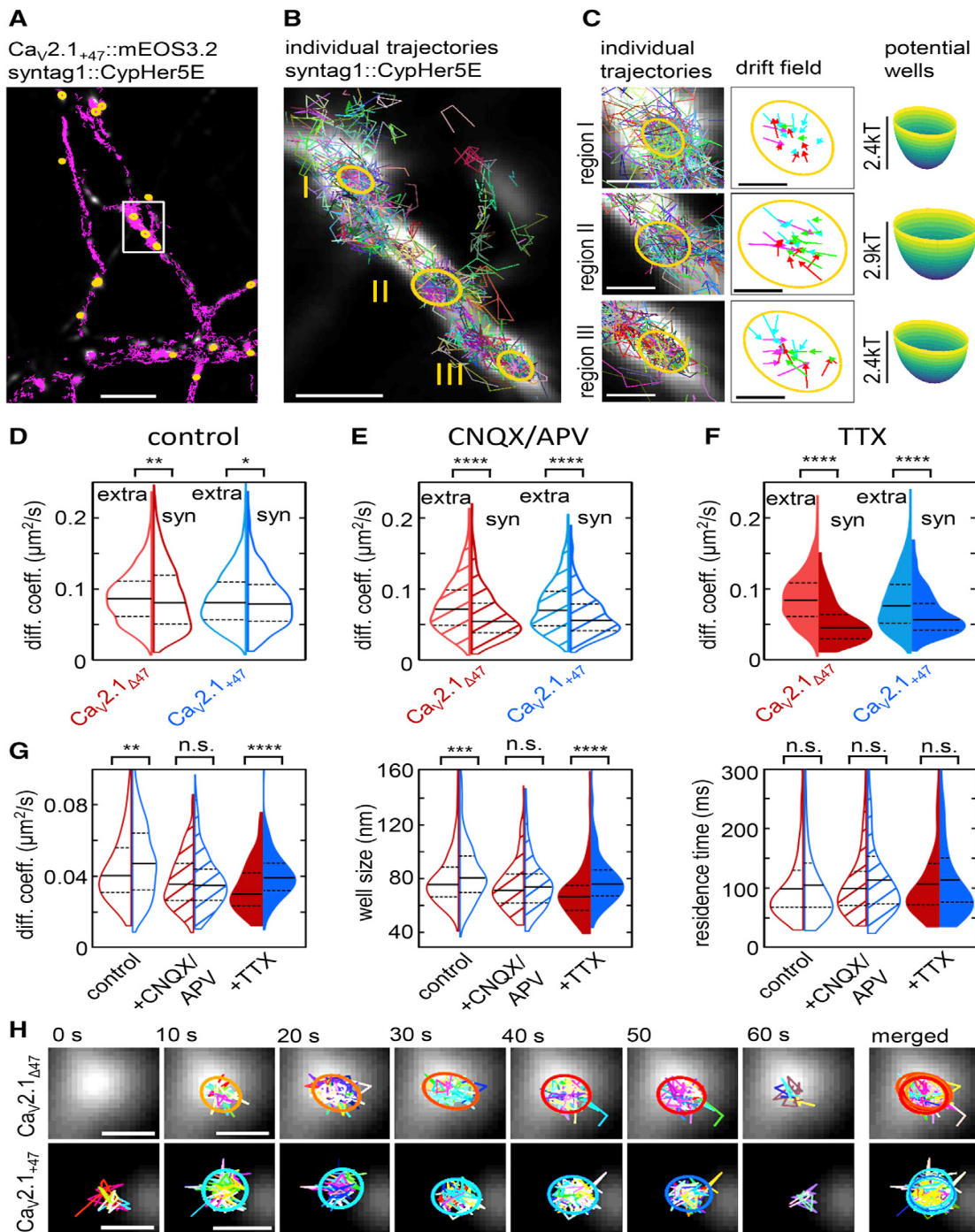


Figure 3.3: (*previous page*) **Dynamic Organization of  $\text{Ca}_v2.1_{\Delta 47}$  and  $\text{Ca}_v2.1_{+47}$  in Nanodomains and Effects of Network Silencing.** **A.** Image of axonal branches transfected with  $\text{Ca}_v2.1_{+47}::\text{mEOS3.2}$  at 15 DIV. Presynapses are labeled via  $\text{syntag1}::\text{CypHer5E}$  antibody (AB) uptake. Scale bar corresponds to  $2\mu\text{m}$ . **B.** Zoomed image shows recorded trajectories and identified nanodomains (yellow ellipses). Scale bar corresponds to  $0.5\mu\text{m}$ . **C.** Zoomed images show detected nanodomains presenting the individual trajectories (left panel), reconstructed drift vector field (middle panel), and a 3D representation of the potential well (right panel). Scale bar corresponds to  $0.1\mu\text{m}$ . **D.** Distribution and median of the diffusion coefficient outside (extra) and inside (syn) synapses. **E.** Distribution and median of the diffusion coefficient outside (extra) and inside (syn) synapses after adding CNQX/APV ( $10\mu\text{M}$ ) to the extracellular solution. **F.** Similar to (E) after adding TTX ( $1\mu\text{M}$ ) to the extracellular solution. **G.** Changes of nanodomain characteristics (diffusion coefficient, well size, and residence time) comparing the two  $\text{Ca}_v2.1$  splice variants and activity conditions as indicated in the figure. For values and statistics, see Table S3 in original article. **H.** Example of two potential wells over time. Scale bar corresponds to  $0.2\mu\text{m}$ .

To test whether the surprisingly low dwell time of individual calcium channels within the nanodomains and their diffusion properties outside nanodomains are influenced by ongoing activity, we altered the neuronal network activity. We repeated sptPALM experiments and acutely reduced synaptic activity by blocking either ionotropic glutamate receptors ( $10\text{ mM APV}/10\text{ mM CNQX}$ ) or sodium channels ( $1\text{ mM TTX}$ ) for 10-15 min. These treatments significantly reduced the dynamics of both  $\text{Ca}_v2.1$  splice variants, particularly inside synapses (Figures 3.3E-G). Similarly, the TTX treatment showed the most prominent decrease of  $\text{Ca}_v2.1$  dynamics within nanodomains (Figure 3.3G). These results demonstrate that the mobility of  $\text{Ca}_v2.1$  channels, regardless of their C-terminal structure, is driven by AP-evoked synaptic activity. Stronger confinement of  $\text{Ca}_v2.1$  in silenced synapses, particularly in the presence of TTX, suggests that additional scaffold interaction sites, which are present in the  $\text{Ca}_v2.1_{+47}$  splice variant, do not directly support tighter coupling of channels and SVs. To test the impact of local diffusion on synaptic transmitter release and the association between channels and SVs, we next used optogenetic oligomerization of the two  $\text{Ca}_v2.1$  splice variants to directly interfere with the calcium channel nanoscale organization.

### 3.2.4 Light-Triggered Re-organization of Calcium Channels

We adopted the optogenetic approach reported earlier to effectively cluster transmembrane proteins in a fast and reversible manner [155]. Fusing a mutant of the Arabidopsis flavoprotein cryptochrome 2 (CRY2<sub>olig</sub>, hereafter referred to as CRY2) [155] to the N-terminal domain of the  $\text{Ca}_v2.1$  splice variants allowed transient clustering of  $\text{Ca}_v2.1$  channels in response to brief illumination with blue light (Figures 3.4A-D). The light-induced cross-linking of  $\text{Ca}_v2.1$  channels was evident within the first minutes after light stimulus and decayed within the following 30 min (Figures 3.4C,D). CRY2- mCherry-tagged  $\text{Ca}_v2.1$  splice variants were co-expressed with either iGluSnFR or GCaMP5G::synaptophysin to monitor synaptic glutamate release or calcium influx, respectively, while triggering channel clustering. The first image sequence in combination with an electrical stimulus was considered as control for fol-

lowing stimuli. One minute after triggering the channel cross-link (x-link),  $\text{Ca}_V2.1_{\Delta 47}::\text{CRY2-mCherry}$ -expressing synapses showed no change in the glutamate response compared to the initial control condition, even though the presynaptic calcium influx was elevated 1 min after x-link induction (Figures 3.4E,F). In contrast, in  $\text{Ca}_V2.1_{+47}::\text{CRY2-mCherry}$ -expressing synapses, the iGluSnFR fluorescence response was significantly enhanced within the first minute and reversibly decayed within longer time intervals in the dark. The synaptic calcium response was not enhanced within the first minute after cross-linking (x-linking) of  $\text{Ca}_V2.1_{+47}::\text{CRY2-mCherry}$ . Instead, it decreased in comparison with the initial control condition (Figure 3.4G). Performing immunocytochemical staining, we found that the CRY2-mediated cross-linking of both  $\text{Ca}_V2.1$  splice variants increased the accumulation of  $\text{Ca}_V2.1$  channels in the synapse when compared to synapses expressing GFP-tagged channels (Figures 3.6A-C). The fluorescence intensities of presynaptic markers, such as Bassoon, did not scale up. This supports our interpretation that  $\text{Ca}_V2.1$  channels are mobile within the axonal membrane and are only transiently coupled to presynaptic interaction partners. Given the increase in density of nanodomains for both cross-linked (x-linked)  $\text{Ca}_V2.1$  splice variants (Figure 3.11J), we assumed transient recruitment of axonal  $\text{Ca}_V2.1$  channels into presynaptic boutons upon cross-linking. However, the additional recruitment of  $\text{Ca}_V2.1$  channels via cross-linking is not the only driving factor to enhance synaptic release. The C-terminal binding affinities to scaffold proteins in  $\text{Ca}_V2.1_{+47}$  channels are likely needed to optimize the positioning of  $\text{Ca}_V2.1$  with respect to SVs and thus effectively enhance SV release.

To examine whether massive recruitment of calcium channels could outcompete the C-terminal structural differences between  $\text{Ca}_V2.1$  splice variants, we evaluated both parameters, presynaptic calcium response and glutamate release, after prolonged illumination of transfected synapses (>500 ms and 488 nm) and a longer incubation time (2 min) (Figure 3.4H). Repetitive stimulation of hippocampal cultures with AP-like stimuli in the presence of 10  $\mu\text{M}$ MCNQX and 10  $\mu\text{M}$ MAPV induced a strong rise in the presynaptic calcium response for synapses expressing x-linked  $\text{Ca}_V2.1::\text{CRY2}$  splice variants (Figure 3.4K), confirming a significant increase of the functional, presynaptic  $\text{Ca}_V2.1$  channel population on the membrane surface. Recordings of the glutamate response after prolonged induction of cross-linking showed a similar increase in the amplitude of the glutamate-induced fluorescence signal in synapses expressing the CRY2-tagged  $\text{Ca}_V2.1_{+47}$  splice variant, as seen after fast CRY2-mediated cross-linking. In the case of the CRY2-tagged  $\text{Ca}_V2.1_{\Delta 47}$  splice variant, the amplitude was not affected but the decay of the fluorescence signal was significantly prolonged (Figures 3.4H-J), which could reflect asynchronous release.

Paired recordings from synaptically coupled neurons were used to verify the functional imaging data and further explore the impact of cross-linking on synaptic transmission. CRY2-induced clustering of  $\text{Ca}_V2.1$  splice variants increased the EPSC amplitude in synapses dominated by  $\text{Ca}_V2.1_{+47}::\text{CRY2}$  but showed no effect on the EPSC amplitude in synapses expressing CRY2-tagged  $\text{Ca}_V2.1_{\Delta 47}$  channels (Figures 3.3I and 3.4L,M). The kinetic properties of EPSCs were not altered (Figures 3.3E,F). The change in EPSC amplitude for synapses expressing  $\text{Ca}_V2.1_{+47}::\text{CRY2}$  was suppressed to control level in the presence of the calcium chelator EGTA (10 mM). In contrast, the EPSC amplitude recorded in  $\text{Ca}_V2.1_{\Delta 47}::\text{CRY2}$ -expressing neurons was EGTA insensitive (Figure 3.4M). The variability in the postsynaptic

response was not affected by cross-linking (Figure 3.4N).

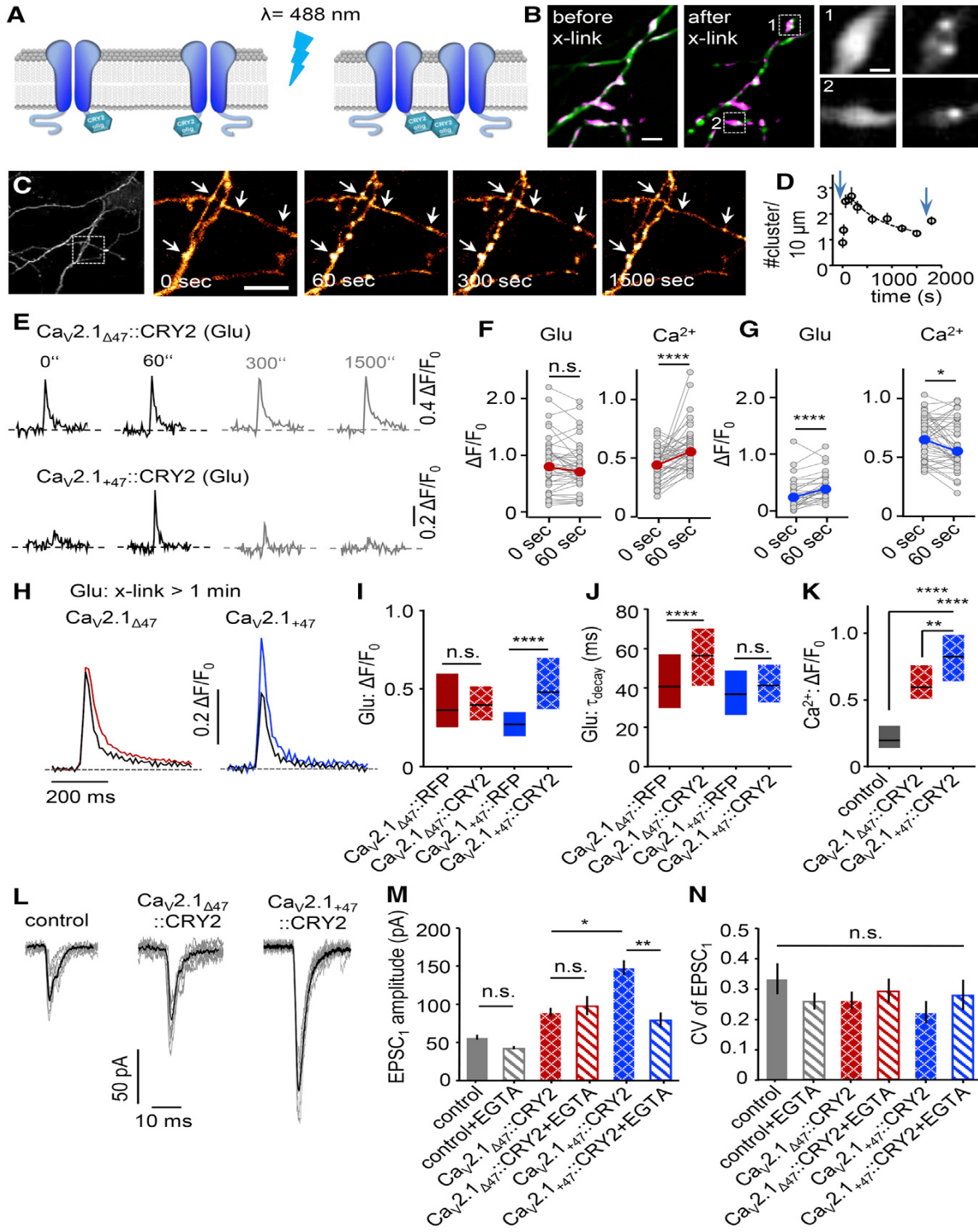


Figure 3.4: (*previous page*) **Light-Induced Cross-Linking of Calcium Channels within Axons and Consequences for Synaptic Transmission Properties.** **A.** Illustration of  $\text{Ca}_v2.1$  clustering using light-induced ( $\lambda = 488 \text{ nm}$ ) oligomerization via a N-terminally tagged cryptochrome 2 mutant (CRY2<sub>olig</sub>, hereafter referred to as CRY2). **B.** Example images show axonal  $\text{Ca}_v2.1_{\Delta 47}::\text{CRY2}$  expression (green) before and after (5 min) initiation of a cross-link. Synaptic compartments are identified by GCaMP5G::synaptophysin coexpression (magenta). Scale bar corresponds to  $2 \mu\text{m}$ . Sections show  $\text{Ca}_v2.1_{\Delta 47}::\text{CRY2}$  distribution in synapses before and after light-induced crosslinking. Scale bar corresponds to  $500 \text{ nm}$ . **C.** Time course of  $\text{Ca}_v2.1_{+47}::\text{CRY2}$  clustering along axons of hippocampal neurons (15 DIV) after 100 ms illumination with blue light ( $\lambda = 488 \text{ nm}$ ). Arrows indicate positions of transient calcium channel aggregates. Scale bar corresponds to  $5 \mu\text{m}$ . **D.** Channel cluster density along axons over time after oligomerization (see Table S4 in original article). **E.** Exemplary traces for the fluorescence signal of iGluSnFR in response to 1 AP-like field stimulation of single synapses co-transfected with  $\text{Ca}_v2.1_{\Delta 47}::\text{CRY2}$ -mCherry (top) or  $\text{Ca}_v2.1_{+47}::\text{CRY2}$ -mCherry (bottom). The fluorescence signal changes over time for  $\text{Ca}_v2.1_{+47}::\text{CRY2}$ -mCherry-expressing synapses but is not altered for  $\text{Ca}_v2.1_{\Delta 47}::\text{CRY2}$ -mCherry-expressing synapses. **F., G.** Quantification of fluorescence signals ( $\Delta F/F_0$ ) for the first and second light pulse (interval 1 min) of individual synapses co-expressing GCaMP5G or iGluSnFR and splice variant  $\text{Ca}_v2.1_{\Delta 47}::\text{CRY2}$ -mCherry (F, red) or  $\text{Ca}_v2.1_{+47}::\text{CRY2}$ -mCherry (G, blue). For, data, see Table S4 in original article. (**H-K**) Synaptic calcium and glutamate signals in response to 1 AP-like electrical stimulus after prolonged cross-linking of  $\text{Ca}_v2.1$  splice variants compared to RFP-tagged channels. **H.** Average traces of the glutamate response to 1 AP-like stimulation for synapses expressing  $\text{Ca}_v2.1_{\Delta 47}::\text{CRY2}$ -mCherry (red) or  $\text{Ca}_v2.1_{+47}::\text{CRY2}$ -mCherry (blue). **I., J.** Median and IQR for maximal amplitude ( $\Delta F/F_0$ ; I) and time constant of the glutamate signal decay ( $\tau_{\text{decay}}$ ; J) as indicated. **K.** Medians and IQR for synaptic calcium signals ( $\Delta F/F_0$ ) of synapses transfected with GCaMP5G::synaptophysin alone and together with clustered  $\text{Ca}_v2.1::\text{CRY2}$  splice variants as indicated. For values and statistics, see Table S4 in original article. **L.** Example traces of EPSCs from neurons expressing cross-linked  $\text{Ca}_v2.1_{\Delta 47}::\text{CRY2}$  or  $\text{Ca}_v2.1_{+47}::\text{CRY2}$  in comparison to control EPSCs. **M.** Average EPSC amplitude from paired recordings that are presynaptic transfected with the CRY2-tagged channel variants as indicated with 1 or 10mM EGTA in the pipette solution. **N.** Coefficient of variation determined for EPSC amplitudes (10 pulses at 0.1 Hz). For values and statistics, see Table S4 in original article.

Whether the functional consequences of cross-linking are reflected in the local organization of  $\text{Ca}_v2.1$  channels was tested by sptPALM experiments (Figure 3.5A). First, we measured the diffusion coefficient of  $\text{Ca}_v2.1::\text{CRY2}$  splice variants outside nanodomains and found a significant decrease of channel mobility after light-triggered cross-linking (Figure 3.5B). The most striking difference between mEOStagged and CRY2-mEOS-tagged channels was the increased density of detected nanodomains along the axonal membrane and their packed distribution within synapses (Figure 3.5C; Figure 3.11J). Channel dynamics and dwell time within nanodomains, as well as nanodomain size, were not changed after cross-linking (Figure 3.5D). This indicates maximal packing of channels within nanodomains and suggests that drift of individual channels in and out of nanodomains is an important parameter affecting SV release probability. To explore the consequences of nanodomain aggregation,

we developed a stochastic model (paragraph 3.4.2 and Figure 3.12D) in which channels can enter and subsequently escape a nanodomain. In the case of two neighboring nanodomains,  $W_1$  and  $W_2$ , we can define the effective well as a collection of the two wells:  $W_{12} = W_1 \cup W_2$  (Figures 3.12D and 3.12E). Channels that recurrently fall into neighboring nanodomains can significantly increase the time they are trapped. The total residence time of a channel depends on the escape probability  $p_e$ . When two wells are located closely, mobile channels will fall immediately into the neighboring nanodomain after escaping the previous one. To give an example, if the residence time in each well is  $t_1 = t_2 = 100$  ms and the escape probability is  $p_e = 0.5$ , we find a total trapping time of  $\bar{\tau}_{12} = 50$  ms. With proximity of the nanodomains so that the escape probability reduces to  $p_e = 0.1$ , the time of confinement inside the two nanodomains increases to  $\bar{t}_{12} = 1.5$  s. These estimations show that the residence time strives toward a very high value in the case of very low escape probability  $p_e$  (Figure 3.12E).

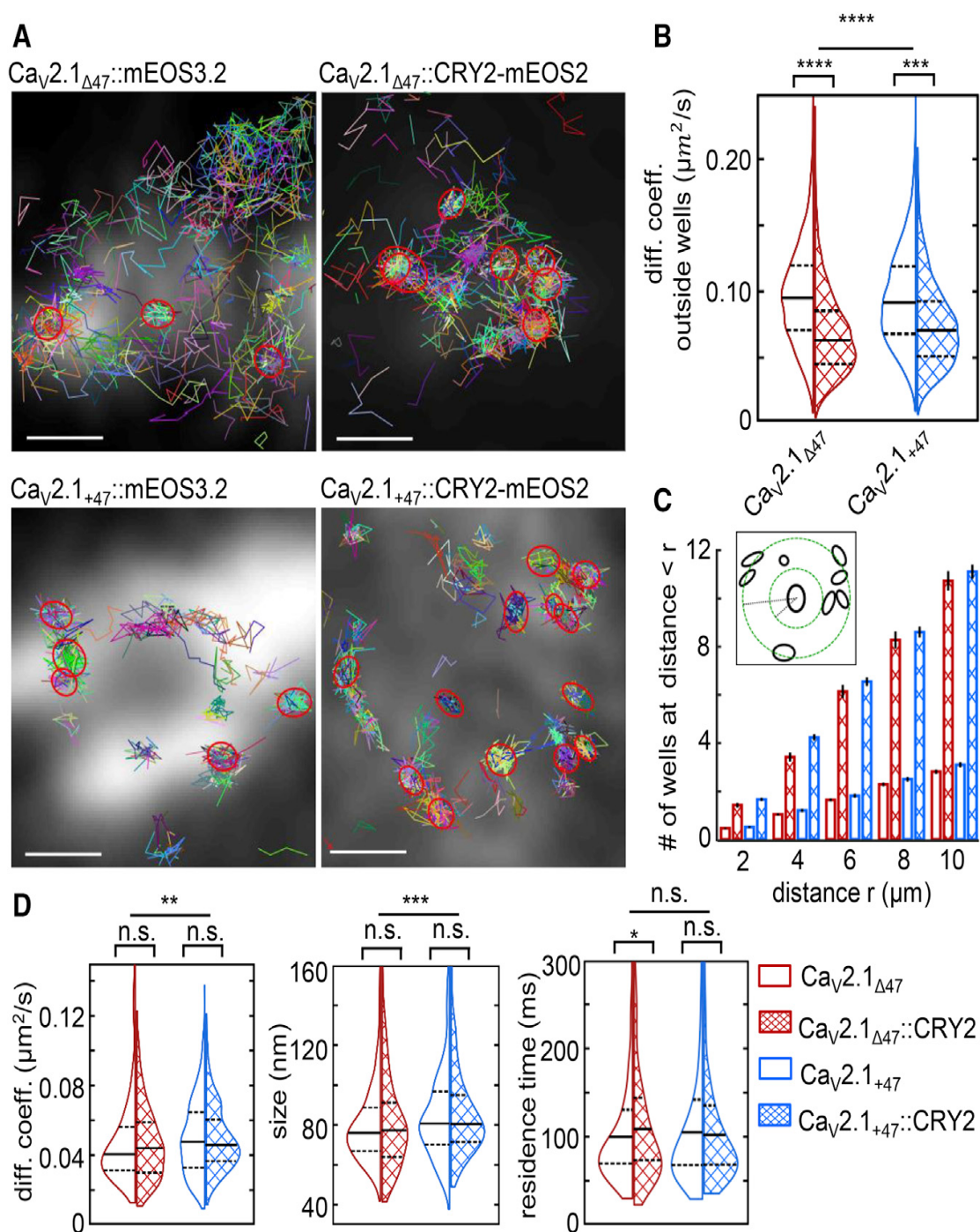
Altogether, the overall increase in nanodomains along the axonal membrane confirms our assumption that  $\text{Ca}_V2.1$  channels are only transiently stabilized inside the synapse and probably can fluctuate between neighboring synapses along the axonal membrane. This could potentially serve as a mechanism to tune synaptic calcium channel density or the ratio of short and long C-terminal splice variants in an activity-dependent manner.

### 3.2.5 C-Terminal Splicing Contributes to Short-Term Plasticity

The described differences in glutamate release and EPSC amplitudes between the two C-terminal  $\text{Ca}_V2.1$  splice variants could indicate variations in the initial vesicular release probability between synapses populated by  $\text{Ca}_V2.1_{+47}$  or  $\text{Ca}_V2.1_{\Delta 47}$  splice variants. In turn, this can directly affect the reliability and temporal coding in the synapse [156].

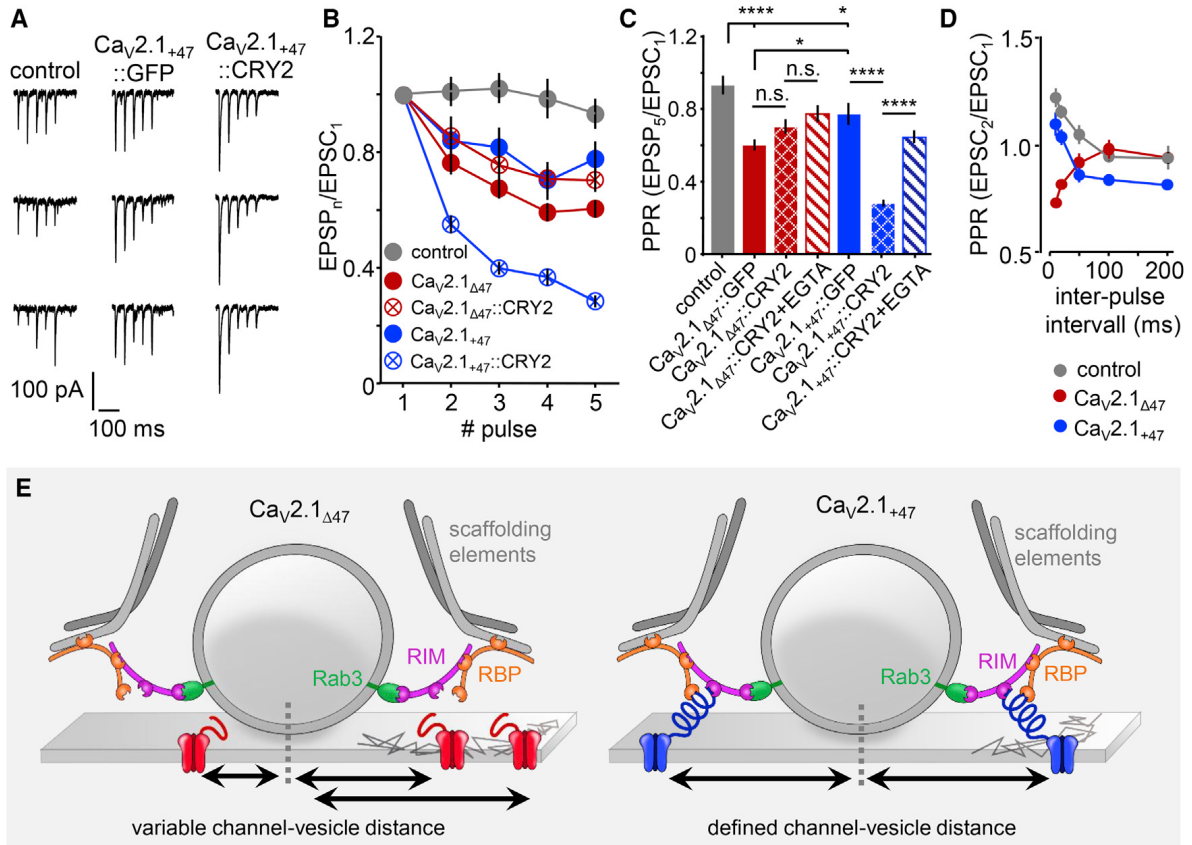
Because channels are mobile and trapped in nanodomains only for short intervals ( $\approx 80$  ms) (Figure 3.3G), we used a repetitive stimulation protocol (5 pulses at 20 Hz) (Figure 3.6A) in which the exocytosis of vesicles outcompetes compensation by SV endocytosis [157]. Assuming that  $\text{Ca}_V2.1$  splice variants are differently coupled to SVs, we expected to find varying paired-pulse ratios (PPRs) between the first and consecutive EPSCs. Repetitive stimulation led to less pronounced depression for GFP-tagged  $\text{Ca}_V2.1_{+47}$  than measured for GFP-tagged  $\text{Ca}_V2.1_{\Delta 47}$  (Figures 3.6B,C). Using cross-linking, we directly tested whether channel dynamics are sufficient to alter PPR. For CRY2-tagged  $\text{Ca}_V2.1_{+47}$  channels, we observed strong paired-pulse depression (PPD) of consecutive responses, whereas synapses expressing the CRY2-tagged  $\text{Ca}_V2.1_{\Delta 47}$  splice variant showed no difference with respect to the GFP-tagged  $\text{Ca}_V2.1_{\Delta 47}$  channel (Figures 3.6A-C). The pronounced PPD of EPSCs in synapses populated by CRY2-tagged  $\text{Ca}_V2.1_{+47}$  channels was diminished by adding EGTA (10 mM) to the intracellular solution. In contrast, no effect of EGTA was evident in synapses expressing CRY2-tagged  $\text{Ca}_V2.1_{\Delta 47}$  upon cross-linking (Figure 3.6C). These data suggest that the intrinsic dynamics of  $\text{Ca}_V2.1_{+47}$  channels control SV release probability. Immobilization of  $\text{Ca}_V2.1_{+47}$  channels by CRY2-induced cross-linking probably results in a more stable interaction between  $\text{Ca}_V2.1$  channels and SVs. In contrast, cross-linking of  $\text{Ca}_V2.1_{\Delta 47}$  channels is not effective to change the coupling between channels and SVs.

The initial release probability of a given synapse has direct consequences for tempo-



**Figure 3.5: Nanoscale Organization of Light-Induced Cross-Linking of  $\text{Ca}_v2.1$  Channels.** **A.** Individual trajectories of  $\text{Ca}_v2.1_{\Delta 47}$  channels (upper row) and  $\text{Ca}_v2.1_{+47}$  channels (lower row) within synapses before and after cross-linking. **B.** Distribution of the diffusion coefficients of both splice variants along the axon, excluding nanodomains, of mEOS2-tagged and mEOS2::CRY2-tagged channels after light-induced cross-linking. For values and statistics, see Table S5 in original article. **C.** Distribution of calcium channel nanodomains without and with cross-linking along the axon. Insert illustrates the method of quantification. For data, see Table S5 in original article. **D.** Nanodomain characteristics of diffusion coefficient, well size, and residence time without and with CRY2-mediated cross-linking. Color-coding and labeling are indicated. For values and statistics, see Table S5 in original article.





**Figure 3.6: Impact of Dynamic Calcium Channel Organization on Synaptic Short-Term Plasticity.** **A.** Example traces of EPSCs (five stimuli at 20 Hz) recorded from monosynaptically coupled neurons (15-16 DIV) under control conditions or presynaptic expression of  $Ca_V2.1_{+47}::GFP$  or  $Ca_V2.1_{+47}::CRY2$  after blue light stimulation. **B.** Ratio of the  $n$ th EPSC amplitude to the 1st EPSC amplitude as indicated in the plot. **C.** Mean ratios between the 5th and the 1st EPSC are plotted. For values and statistics, see Table S6 in original article. **D.** Interevent interval-dependent modulation of the paired-pulse ratio for paired recordings from neurons expressing presynaptic  $Ca_V2.1_{\Delta 47}::GFP$  (red) or  $Ca_V2.1_{+47}::GFP$  (blue) in comparison to control neurons (gray). For values and statistics, see Table S6 in original article. **E.** Schematic illustration for the organization of calcium channels in the presynaptic membrane.  $Ca_V2.1_{\Delta 47}$  channels can closely approach SV (left sketch,  $Ca_V2.1_{\Delta 47}$  in red) but can also be located away from SV.  $Ca_V2.1_{+47}$  are either tethered to scaffold proteins or free diffusive. The specific but transient interactions keep  $Ca_V2.1_{+47}$  channels (blue) in a defined distance to the SV and hence contribute to the control of the release probability, in addition to their kinetic properties and other molecular interactions. The transient nature of channel-SV interactions allows the steady exchange of channels around the SV.

ral information transfer. To test whether the time resolution of synaptic activation affects synapses dominated by the one or the other C-terminal  $\text{Ca}_v2.1$  splice variant differently, we analyzed PPRs recorded from for time intervals between 10 and 200 ms. Control synapses showed paired-pulse facilitation (PPF) at pulse intervals  $<50$  ms and equal amplitudes for intervals  $>50$  ms. Synapses expressing  $\text{Ca}_v2.1_{+47}$  channels followed the trend of control synapses, being facilitated at high frequency and running into slight depression at longer interpulse intervals (Figure 3.6D). In contrast, synapses expressing  $\text{Ca}_v2.1_{\Delta 47}$  channels showed the opposite short-term behavior. Here, fast stimulation with short time intervals resulted in PPD, whereas longer activation intervals ( $>100$  ms) did not affect the second EPSC amplitude (Figure 3.6D).

These functional results indicate that structural differences induced by C-terminal splicing of the  $\text{Ca}_v2.1$  channel are relevant for temporal encoding of information within neuronal networks dominated by the local nanoscale dynamics of the channels. We propose that the difference in the structure of the calcium channel's C terminus sets the minimal distance between channel and SV. This distance might be variable for  $\text{Ca}_v2.1_{\Delta 47}$  channels, whereas  $\text{Ca}_v2.1_{+47}$  channels associated with scaffold proteins have a defined distance to the SV to ensure fast and precise single-vesicular release (Figure 3.6E).

### 3.3 Discussion

Here, we tested the hypothesis that differences in the structure and affinity of the  $\text{Ca}_v2.1$  C terminus to scaffold proteins affect local synaptic channel organization and mobility, as well as synaptic transmission. Our outcomes of sptPALM in living neurons show that most calcium channels at the presynapse are mobile and undergo transient confinement within nanodomains (Figure 3.3). The statistical power of thousands of trajectories generated from sptPALM experiments allowed us to uncover that individual  $\text{Ca}_v2.1$  channels are transiently confined in 1-2 nanodomains per synapse. Within these nanodomains, channels remain for  $\approx 80$  ms (Figure 3.3G), long enough to sustain AP-evoked transmitter release but still flexible enough to influence the release properties of the synapse during short-term plasticity. Studies using freeze fracture replica-labeling electron microscopy have demonstrated clustered organization of VGCCs within the active zone [132, 133, 135], leading to the prediction of a perimeter release model [135]. In theory, the nanodomains described here match the requirements of a central cluster of VGCCs. However, electron microscopy performed on fixed samples provides only a snapshot of the dynamic presynaptic organization. Our results confirm that calcium channels are confined in nano-sized zones. However, this stabilization is transient considering the diffusive character of a single channel, which allows channels to enter and escape nanodomains and most likely enables an exchange between synapses. The fluctuation and small numbers of presynaptic  $\text{Ca}_v2.1$  channels suggest that mobility of VGCCs contributes to the effective coupling distance between VGCC and SV.

Modeling and super-resolution microscopy data from pre- and postsynaptic scaffold proteins propose the existence of trans-synaptic nanocolumns [158, 84] that ensure the alignment of pre- and postsynaptic membranes [159]. Our data suggest that these structures influence

the position of calcium channels within nanodomains and support current models of channel arrangement in synapses [127, 135]. The dynamics of postsynaptic  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (AMPA receptors), analyzed with similar analytical tools, revealed that their nanodomain organization is more stable compared with VGCCs in the presynaptic membrane [17, 160, 161, 32]. The difference in nanodomain stability of presynaptic VGCCs and postsynaptic AMPARs probably reflects their molecular environment. Ongoing turnover of vesicles at the presynaptic membrane may destabilize VGCCs, whereas postsynaptic receptor populations are less disturbed by exo- and endocytotic events that mainly occur outside the postsynaptic density [162, 163]. The question of how presynaptic calcium channel nanodomains are defined at a molecular level remains open. One possibility is the transient association of VGCCs to individual release sites within the active zone [119]. Observed stabilization of  $\text{Ca}_v2.1$  channels after acute silencing of network activity (Figure 3.3) suggests that ongoing transmitter release prevents stable connections between individual calcium channels and release sites and hence promotes fast turnover of channels inside the nanodomain. Specific transsynaptic adhesion complexes that are known to influence presynaptic release properties and to promote the accumulation of channels in the presynaptic membrane could contribute to the formation of nanodomains as confining elements [164, 165]. Channels outside nanodomains may be kept inside the presynaptic compartments by repulsive interactions with the endocytotic machinery on the edge of the active zone [166]. Individual release sites occur at distinct places within the active zone [138] and are initialized by Munc13 [119], which suggests that these molecular complexes may determine the fluctuations of individual nanodomains of VGCCs inside synapses. In addition, molecular crowding of the presynaptic membrane will contribute to keeping a sufficient population of channels within the presynaptic membrane.

The distal C-terminal binding sites in relation to scaffold proteins do not influence the general synaptic targeting of  $\text{Ca}_v2.1$  channels [150, 149, 167] (Figures 3.1D,E). However, depending on the splicing of the C terminus,  $\text{Ca}_v2.1$  channels influence synaptic transmission properties and short-term plasticity differently (Figures 3.2 and 3.6D). Using functional imaging and electrophysiology, we show that surprisingly, the loss of the distal C terminus ( $\text{Ca}_v2.1_{\Delta 47}$ ) increases the release probability of SV, although these channels seem to be loosely coupled, as indicated by the sensitivity of the EPSC amplitude to EGTA (Figures 3.2D-I). In contrast, synapses dominated by  $\text{Ca}_v2.1_{+47}$  channels were proposed to be tightly associated to SVs via the C-terminal scaffold interaction [143] and were insensitive to EGTA. Opposite to our initial expectations,  $\text{Ca}_v2.1$  channels expressing the full C terminus ( $\text{Ca}_v2.1_{+47}$ ) are significantly more dynamic within presynaptic nanodomains and less potent to scale the presynaptic molecular composition compared with  $\text{Ca}_v2.1_{\Delta 47}$  channels (Figures 3.1F and 3.3G). Stronger up-scaling of presynaptic proteins and the total presynaptic  $\text{Ca}_v2.1$  population upon expression of  $\text{Ca}_v2.1_{\Delta 47}$  channels can involve several mechanisms. First, the loss of the distal C terminus could allow the  $\text{Ca}_v2.1_{\Delta 47}$  variant to get closer to the SV due to reduced scaffold interaction. Second, the absence of the distal part of the C terminus could uncover additional binding sites, which are otherwise sterically blocked or simply impeded by other interaction partners of the distal C terminus. Such a mechanism has been described for a  $\text{Ca}_v2.1$  exon 47 knockout mouse model expressing exclusively the short C-terminal splice variant [148] found impaired binding of  $\text{Ca}_v2.1_{\Delta 47}$  channels to RIM

and the  $\text{Ca}_V\beta 4$  subunit, but more interestingly, they uncovered an enhanced  $\text{GABA}_{\text{B}2}$  interaction. Considering the observed deficits in motor coordination and the occurrence of absence seizures in this mouse model, these findings support the importance of the distal C-terminal domain for multiple protein interactions not only by providing binding sites but also by masking other ones. Other studies focusing on the C terminus of  $\text{Ca}_V1.3$  channels and its interaction to calmodulin also showed the expression of the  $\text{Ca}_V1.3$  exon 47 in connection with impeded interaction [168, 169]. Here, only the  $\Delta 47$  variant has been reported to bind calmodulin via a C-terminal pre-IQ domain [168, 169]. In the context of calmodulin interaction, the authors proposed a regulatory mechanism that allows activity-dependent cooperative gating of calcium channels by the formation of channel clusters using calmodulin as a linker. Our functional and electrophysiological data, demonstrating that the short  $\text{Ca}_V2.1$  splice variant is more potent in synaptic release and at the same time only loosely coupled (Figures 3.2E,F,H,I), could be interpreted as cooperative gating of clustered channels, leading to add-on of local  $\text{Ca}^{2+}$  transients. Consequently, we hypothesize that the distal part of  $\text{Ca}_V2.1$  not only sets a specific VGCC-to-SV distance but also prevents binding to other potential interaction partners and probably to calcium channels so that the expression of exon 47 helps to maintain single-channel properties.

Using light-induced cross-linking, we directly explored the impact of calcium channel surface dynamics on synaptic transmission properties. Here, we observed effective recruitment of  $\text{Ca}_V2.1$  channels to the synaptic compartment and reduction of their overall mobility (Figures 3.5A-C; Figure 3.11J). Keeping these effects in mind, the dramatic effect of exclusively  $\text{Ca}_V2.1_{+47}$  cross-linking on the PPR and its sensitivity to EGTA supports our idea of a defined connection between SVs and  $\text{Ca}_V2.1_{+47}$  channels. We propose that C-terminal scaffold interactions define a certain distance between  $\text{Ca}_V2.1_{+47}$  channels and SVs. Thus, cross-linking of  $\text{Ca}_V2.1_{+47}$  might recruit channels close to readily releasable SVs reflected in an enhanced EPSC amplitude. The cross-linking of  $\text{Ca}_V2.1_{+47}$  channels reduced their dynamics and could increase the possibility of  $\text{Ca}_V2.1_{+47}$  channels being tethered to SVs. Due to the N-terminal x-link and the likely limited number of interaction sites to SVs, additional x-linked  $\text{Ca}_V2.1_{+47}$  channels might be not perfectly localized but rather interlinked to channels already associated with the SV.

We conclude that the emergent effect of mobile channels within the synapse is the ability to tune SV release in response to previous activity, which will influence the association between channels and SVs depending on their C-terminal configuration and the effective affinities to interaction partners. Hence, quick rearrangements within the active zone can affect the reliability of synaptic transmission, as seen in synapses overpopulated with one of the C-terminal splice variants or after cross-linking (Figures 3.6D,F,G).

We propose that alternative splicing of exon 47 of  $\text{Ca}_V2.1$  channels contributes to release probability and short-term plasticity of individual synapses from different neurons or even along the axon of individual neurons [85, 170]. When the  $\text{Ca}_V2.1_{\Delta 47}$  variant lacks C-terminal scaffold interactions, we assume stochastic coupling of calcium channels to SVs primarily caused by their surface mobility. However,  $\text{Ca}_V2.1_{+47}$  channels are likely localized in defined proximity to SVs, which is transient in nature. The impact of alternative splicing on the developmental profile of short-term plasticity remains to be explored. The clinical manifestation of SCA6, a developmental late-onset spinocerebellar ataxia, was reported to

be caused by abnormal expansion of a CAG nucleotide triplet repeat in the CACNA1A gene expressed only in the +47 splice variant [171].

Altogether, our data suggest that alternative splicing can serve as an endogenous mechanism to shape release probability and short-term plasticity of individual synapses, as shown for homeostatic plasticity [85]. We propose that stochastic mobility-driven and use-dependent rearrangement of calcium channels constitutes initial processing of information at the presynaptic site and depends on flexible association between channels and SVs.

## 3.4 Supplementary Information

*[Irrelevant parts of the SI with respect to the thesis were left out. Please see the original article for the complete SI].*

### 3.4.1 Single-particle tracking PALM

#### Experimental setup

Hippocampal neurons were transfected at DIV4 either with the synaptic calcium sensor synaptophysin::GCaMP5G and the N-terminally tagged Ca<sub>v</sub>2.1 splice variants (mEOS2 or mEOS3.2) in a ratio of 1:1, or Ca<sub>v</sub>2.1 splice variants alone. When Ca<sub>v</sub>2.1 splice variants were transfected alone, cells were incubated in culture medium with 2 mg/ml synaptotagmin1 antibody labeled with CypHer5E (Synaptic Systems; Cat #: 105311CpH) for 1 h prior to the experiment to indicate synaptic compartments. All sptPALM experiments were performed at DIV15-19 within a whole microscope incubator at 37°C using an inverted TIRF setup. The eclipse Ti microscope (Nikon GmbH) was equipped with a 100 x Apo TIRF oil objective (1.49 NA; Nikon). Fluorescence was excited by oblique illumination of the probe with a combined laser system (Coherent; MPB communications Inc.) and image sequences (up to 9999 frames) were captured by an EMCCD camera (iXon+ 897, Andor Technology) controlled by NIS-Elements Advanced Research acquisition software (Nikon) at a frame rate of 50 Hz. If not mentioned otherwise, cells were perfused with extracellular solution containing (in mM): 145 NaCl, 2.5 KCl, 10 HEPES, and 10 D-glucose 2 MgCl<sub>2</sub>, 2 CaCl<sub>2</sub> (pH 7.4). Prior to acquisition, the red channel was bleached to reduce background and unspecific signals. The mEOS was photo-activated using a continuous illumination of the sample with a 405 nm laser set to 4%-5% of maximal laser power (100 mW). Simultaneously, the photo-converted mEOS single-molecule fluorescence signal was excited with a 561 nm laser using 20% of maximal laser power (300 mW). To further improve the separation of the mEOS signal from auto fluorescence and background signals, an emission bandpass filter (ET620/60 nm; AHF analysentechnik) was used. In oligomerization experiments, CRY2olig-mCherry-positive cells were exposed to a 488 nm light pulse (> 500 ms) and recorded 2 min later. To dissect the impact of neuronal activity on Ca<sub>v</sub> 2.1 surface mobility, the network activity was acutely suppressed by 1) blockade of postsynaptic glutamate receptors (through application of 10 mM APV and 10 mM CNQX to the extracellular solution) and 2) blockade of sodium channels (through application 1 mM tetrodotoxin to the extracellular solution).

Localization and trajectory reconnection of mEOS signals was carried out using wavelet-based algorithm [52]. Trajectories of mEOS-tagged molecules were reconstructed by a simulated annealing algorithm [57], which takes into account molecule localization and total intensity. It has been reported that mEOS2 molecules can show blinking-like behavior [172, 32]. To avoid false reconnections between trajectories, all sub-trajectories were analyzed as individual trajectories. The experimental limit to resolve the dynamic behavior of single molecules was computed from the mean square displacement (MSD) of immobile mEOS3.2 molecules in fixed samples. Diffusion coefficients ( $D$ ) were calculated by linear fitting of the first four points of the MSD plots. MSD plots of immobilized molecules (on fixed samples) revealed that, under our imaging conditions, molecules with  $D \geq 0.002\mu\text{m}^2/\text{s}$  can be considered as mobile.

### 3.4.2 Data analysis and modeling

#### Estimation of potential wells characteristics

The characteristics of the parabolic well are estimated from SPTs using the drift map computed from equation (1.20) with a square grid of resolution  $r = 30$  nm considering only bins containing at least 3 points. The center of the wells corresponds to local maxima of the density map. The boundary of each well corresponds to the error ellipse of the covariance matrix of the points located around the center of the well and obtained by principal component analysis. We then estimate the depth of the each well (equation (1.8)) by computing the optimal gradient field corresponding to the observed drift field inside the ellipse. We use a least-squares approach for optimizing the following error function

$$Err_N(\mathbf{b}, A) = \sum_{i=1}^N \|\nabla U(\mathbf{X}_i) - \mathbf{b}(\mathbf{X}_i)\|^2 = \sum_{i=1}^N \left( b_x(\mathbf{X}_i) + 2A \frac{x_i}{a^2} \right)^2 + \left( b_y(\mathbf{X}_i) + 2A \frac{y_i}{b^2} \right)^2 \quad (3.1)$$

where  $N$  is the number of bins contained in the ellipse,  $\mathbf{X}_i = (x_i, y_i)$  is the center of each bin (considering the center of the ellipse as the origin and rotating the coordinate system such that the large semi-axis aligns with the x axis),  $\mathbf{b}(\mathbf{X}_i) = (b_x(\mathbf{X}_i), b_y(\mathbf{X}_i))$  drift field estimated for bin  $\mathbf{X}_i$ ,  $a > b$  are the semi-axes of the ellipse and  $A$  is the depth of the well. The only free parameter of equation (1.8) is the depth  $A$  that we obtain as:

$$A_N^* = \underset{A \in \mathbb{R}}{\operatorname{argmin}} Err_N(\mathbf{b}, A) \quad (3.2)$$

Leading to the unbiased estimator

$$A_N^* = \frac{\sum_{i=1}^N \frac{b_x(\mathbf{X}_i)x_i}{a^2} + \frac{b_y(\mathbf{X}_i)y_i}{b^2}}{2 \sum_{i=1}^N \frac{x_i^2}{a^4} + \frac{y_i^2}{b^4}}. \quad (3.3)$$

Finally, we defined a similarity index between the observed drift field and the corresponding optimal parabolic field, based on the residual error as

$$S_N(\mathbf{b}, A_N^*) = \frac{Errr_N(\mathbf{b}, A_N^*)}{\sum_{i=1}^N \|\mathbf{b}(\mathbf{X}_i)\|^2} = 1 - \frac{\left( \sum_{i=1}^N \frac{b_x(\mathbf{X}_i)x_i}{a^2} + \frac{b_y(\mathbf{X}_i)y_i}{b^2} \right)^2}{\left( \sum_{i=1}^N \frac{x_i^2}{a^4} + \frac{y_i^2}{b^4} \right) \left( \sum_{i=1}^N \|\mathbf{b}(\mathbf{X}_i)\|^2 \right)} \quad (3.4)$$

The index  $S$  has the following characteristics:  $S \in [0, 1]$  such that  $S \rightarrow 0$  for an observed drift field corresponding to equation (1.8) and  $S \rightarrow 1$  for a random vector field as observed for diffusive motion. We consider as parabolic potential wells the regions for which  $S < \frac{1}{2}$ .

### Temporal segmentation of trajectories

In order to capture the temporal dynamics of the wells and to prevent possible motion of well centers to contaminate the estimations, we performed a temporal segmentation of the trajectories. For each dataset, we segmented the raw trajectories into successive time windows of 20 s duration with either 0% or 50% overlap between successive windows. Trajectories appearing across two time windows were cut into two separate trajectories. Then, for each time window, we run the potential well detection algorithm and collected the detected wells. Finally, we regroup into families the wells from successive time windows which centers are less than 100 nm apart. Families allow to track the temporal evolution of the same nanodomain through multiple time-windows. In Figures 3.3C,E, and 3.5B-D, the reported potential wells were detected using 0% overlap between time windows. In Figures 3.3A,B, and 3.5A, we used a 50% overlap between time windows and the reported ellipses are averages for the different detected families.

### Colocalization of the potential wells with presynaptic markers

We use the following criteria to determine whether or not a potential well colocalize with either synaptotagmin1::CypHer5E or GCaMP5G::synaptophysin when the threshold of the pixel intensity was above 50 (in absolute value in the scale from 0 to 255) for at least 30% of the surface area of the well. Except specified otherwise, all the potential wells reported colocalize with either synaptic marker.

### Estimation of the residence time for two close-by potential wells

Cross-linking channels result in aggregation of potential wells. To assess how this clustering affects the residence time of a channel in a cluster, we consider the case of two close-by wells  $W_1$  and  $W_2$  with residence times  $\tau_1$  and  $\tau_2$  respectively. We consider that when channels escape from  $W_1$  they can either escape from the cluster with probability  $p_e$  or enter the second well with probability  $1 - p_e$ . The probability  $p_e$  depends on the distance between the

two wells. To compute the total residence time  $\tau_T$  of a channel inside the cluster  $W_1 \cup W_2$ , we use Bayes law by conditioning on the number of jumps  $k$  between the wells

$$\bar{\tau}_T = \sum_{k=0}^{+\infty} E\{\tau_T|k\}Pr\{k\}, \quad (3.5)$$

where  $Pr\{k\}$  is the probability of  $k$  transitions between wells before exit and  $E\{\tau_T|k\}$  is the conditional exit time for  $k$  transition. By definition,

$$Pr\{k\} = p_e(1 - p_e)^k, \quad (3.6)$$

and

$$E\{\tau_T|2p\} = (p + 1)\tau_1 + p\tau_2, \quad E\{\tau_T|2p - 1\} = p(\tau_1 + \tau_2). \quad (3.7)$$

The sum is thus

$$\bar{\tau}_T = \frac{1}{2(2 - p_e)^2} \left( (1 - p_e)(\tau_1 + \tau_2) + \tau_2(1 - p_e)^2 + \tau_1 \right). \quad (3.8)$$

The value of  $\bar{\tau}_T$  as a function of  $p_e$  is presented in Figure 3.12E for different values of  $\tau_1$  and  $\tau_2$ . In Figures 3.12F-G, we present a numerical scheme to estimate the probability  $p_e$  using Brownian simulations. In this scheme, we consider two wells in the active zone separated by a distance  $d$ , that we vary. To estimate the escape probability  $p_e$ , we ran simulations and computed the first time that a Brownian trajectory starting at a point uniformly distributed outside the well (modeled as disk of radius  $r_{in}$ ) hits the active zone boundary, which is defined as outer circle of radius  $r_{out}$  (Figure 3.12F, red circle), before any other one of the well. The result is presented in Figure 3.12G.

To conclude, this stochastic simulation shows how the splitting probability can be computed from a model of potential wells located inside a round Active Zone. For Active Zones with different geometrical features (elliptic shape), with potential wells located at a shorter distance from the boundary could largely modulate the present result.



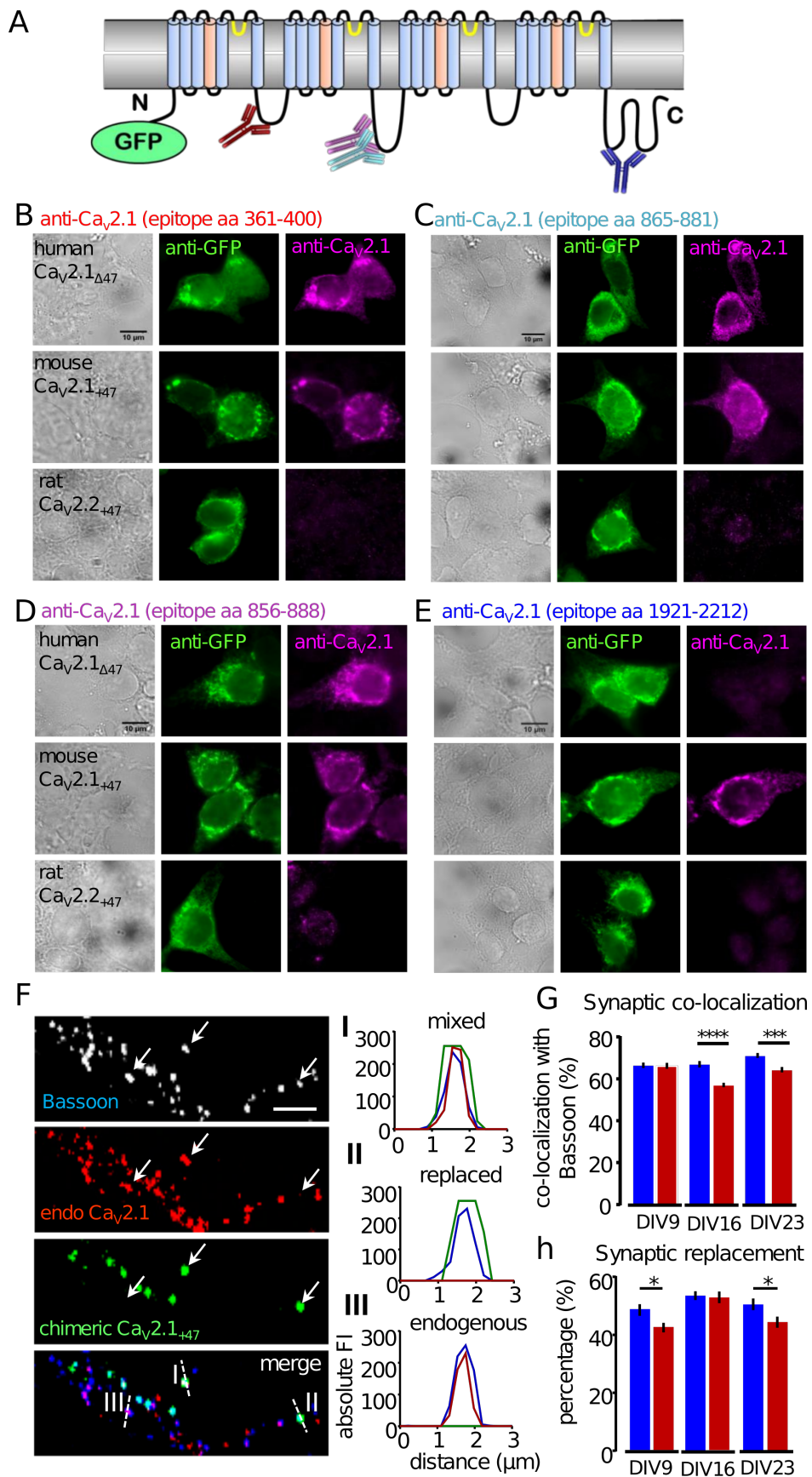
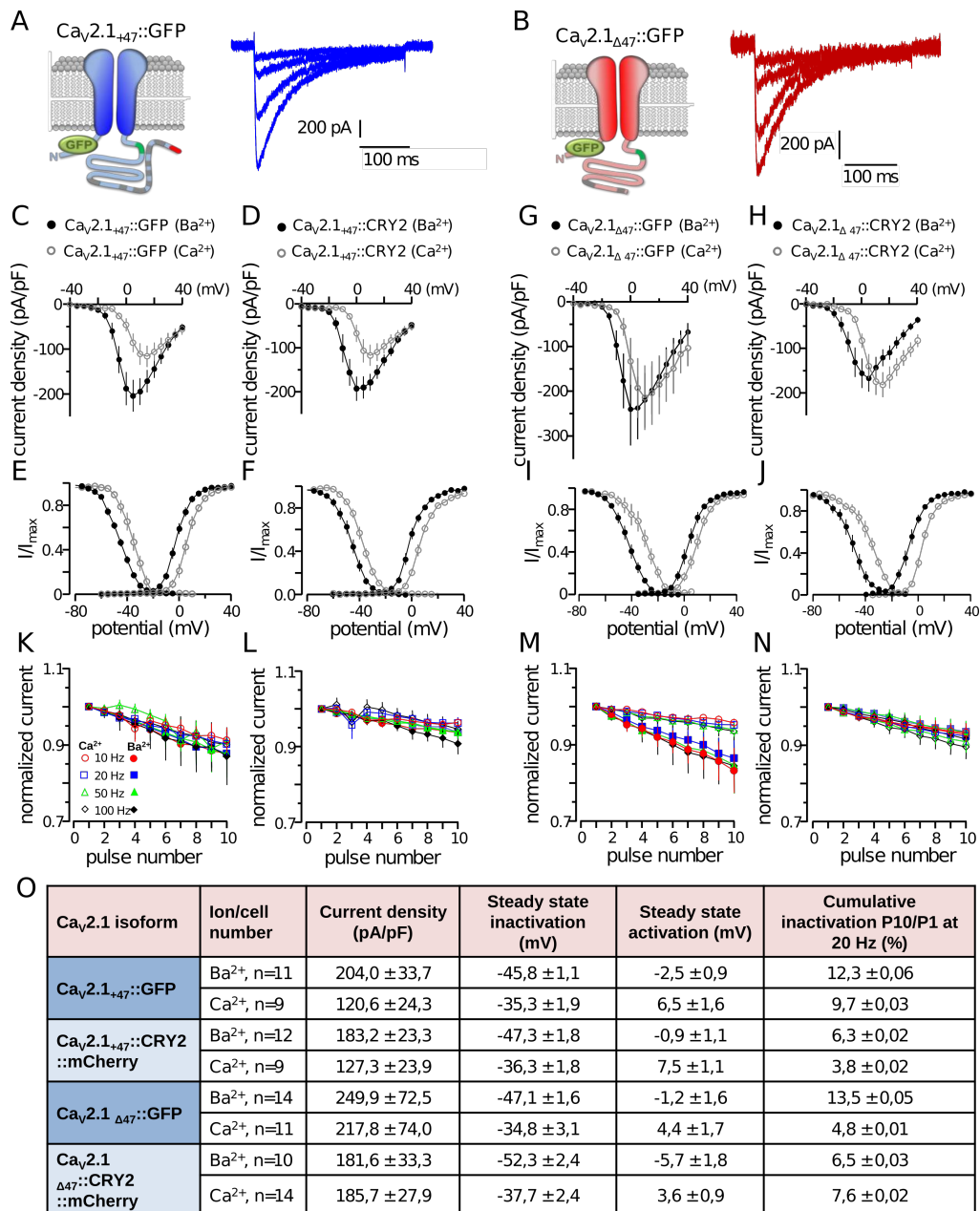


Figure 3.7: (*previous page*) Species specific immune reactivity of commercially anti-Ca<sub>V</sub>2.1 antibodies and replacement of endogenous Ca<sub>V</sub>2.1 channel population by overexpressed Ca<sub>V</sub>2.1 Antibodies against Ca<sub>V</sub>2.1 channel epitopes were tested in transfected HEK293T cells to determine their specificity for Ca<sub>V</sub>2.1 and species specific differences in the C-terminus. Different subunits and C-terminal splice variants from rat and human were used to determine the specificity of used antibodies. **A.** Schematic representation of Ca<sub>V</sub>2.1  $\alpha$ 1 subunit indicating the binding sites of the four tested Ca<sub>V</sub>2.1 antibodies against intracellular epitopes. **B.** Guinea pig anti-VDCC  $\alpha$ 1A (Frontier Institute, VDCC $\alpha$ 1A-GP-Af810). **C.** Rabbit anti-Ca<sub>V</sub>2.1 (Alomone Labs, Cat #: ACC-001). **D.** Rabbit anti-voltage-gated Ca<sup>2+</sup> channel antibody 1 (Synaptic Systems, Cat #: 152103). **E.** rabbit anti-voltage-gated Ca<sup>2+</sup> channel antibody 2 (Synaptic Systems, Cat #: 152203). The experiments were repeated for 3 independent transfections. Scale bar 10  $\mu$ m. **F.** Representative fluorescence images of over-expressed chimeric Ca<sub>V</sub>2.1<sub>+47</sub>::GFP clusters (green) in cultured hippocampal neurons (DIV16). Endogenous Ca<sub>V</sub>2.1 channels (red) were labeled with the anti-Ca<sub>V</sub>2.1 antibody (Synaptic Systems, Cat #152203). Anti-Bassoon labeling (blue) was used as presynaptic marker. Scale bar 5  $\mu$ m. I) Line scan of a mixed synapse (characterized by positive staining for Bassoon, the chimeric Ca<sub>V</sub>2.1<sub>+47</sub> and endogenous Ca<sub>V</sub>2.1 channels. II) Line scan of a replaced synapse lacking the endogenous Ca<sub>V</sub>2.1 population. III) Line scan of an endogenous synapse showing staining for Bassoon and endogenous Ca<sub>V</sub>2.1 only. **G.** Synaptic co-localization analyzed for overexpressed Ca<sub>V</sub>2.1 splice variants (Ca<sub>V</sub>2.1 $\Delta$ <sub>47</sub> (red) and Ca<sub>V</sub>2.1<sub>+47</sub> (blue)) transfected at DIV 4 and stained at DIV9, 16 and 23. Data represent 3 hippocampal cultures per condition. **H.** Power of Ca<sub>V</sub>2.1 splice variants for synaptic replacement of endogenous Ca<sub>V</sub>2.1 with overexpressed Ca<sub>V</sub>2.1::GFP, as indicated in G). Significance for synaptic co-localization was determined for the splice variants: DIV9, p-value=0.887778 (not significant); DIV16, p-value=0.000001 (\*\*\*\*), DIV23, p-value=0.000207 (\*\*\*). H) Synaptic replacement has been analyzed for overexpressed Ca<sub>V</sub>2.1 splice variants transfected at DIV 4 at stained at DIV9, 16 and 23. Significance for strength of synaptic replacement was determined for the splice variants: DIV9, p-value=0.023746 (\*); DIV16, p-value=0.738954 (ns), DIV23, p-value=0.015258 (\*). All data shown in g and h are means calculated from 35 (n=35) 512  $\times$  512 px images. The mean number of analyzed GFP-positive synapses per image was 80 for DIV9; 150 for DIV16 and 180 for DIV23. The experiment performed on three individual hippocampal cultures in case of the DIV9 and DIV16 condition and on two hippocampal cultures in case of DIV23 condition. Error bars indicate SEM.

### Estimation of the number of channels in a well

SPTs are obtained by random activation and it is not possible to obtain this data to estimate the number of channels located in a single potential well at the active zone. Indeed, the method is based on a sparse excitation of the fluorophores at a random time. To estimate the number of receptors inside potential wells, it is necessary to use the density of Ca<sub>V</sub>2.1 channels at the active zone  $\approx 403$  channels/ $\mu$ m<sup>2</sup> [132, 133]. Using the estimation of potential well size,  $S = \pi ab = 0.022\mu$ m<sup>2</sup> ( $a, b$  are the main axes of the ellipses), we obtain around 9 calcium channels in an average well. If there are  $\approx 3$  docking sites (in average) per active zone [133], we conclude that there are  $\approx 3$  channels per vesicle, although the key parameter is



**Figure 3.8: Electrophysiological characterization of tagged Ca<sub>v</sub>2.1 splice variants in HEK293T cells.** **A.**, **B.** Schematic representation of Ca<sub>v</sub>2.1 C-terminal splice variants: Ca<sub>v</sub>2.1<sub>+47</sub>::GFP (blue) and Ca<sub>v</sub>2.1<sub>Δ47</sub>::GFP (red) with corresponding whole-cell currents using barium as charge carrier and induced by a step protocol from holding potential at -90 mV to test potentials between -10 - 5 mV. **C.-J.** Current/voltage relationship, activation and steady inactivation properties of tagged Ca<sub>v</sub>2.1 splice variants (as indicated) using barium (black) or calcium (gray) as charge carrier. **K.-N.** Fast AP-like voltage ramps from -90 mV to 0 mV and back to -90 mV were used to probe the cumulative inactivation of tagged channel subunits. Trains of 10-100 Hz for 10 repetitive stimuli were used. Currents were recorded either with barium (filled symbols) or calcium (bordered symbols) as charge carrier. **O.** Averaged values and cell number for whole-cell patch clamp analysis of tagged Ca<sub>v</sub>2.1 splice variants.

the distance of the channel to the center of the vesicle or the place where synaptotagmin are located. We note that the present estimations are very different from the ones obtained from a packing density calculation where we find that there are 47 round channels (in projection) of radius  $r = 10$  nm that fill an ellipse (region of a potential well) of size  $a = 88$  nm and  $b = 70$  nm.

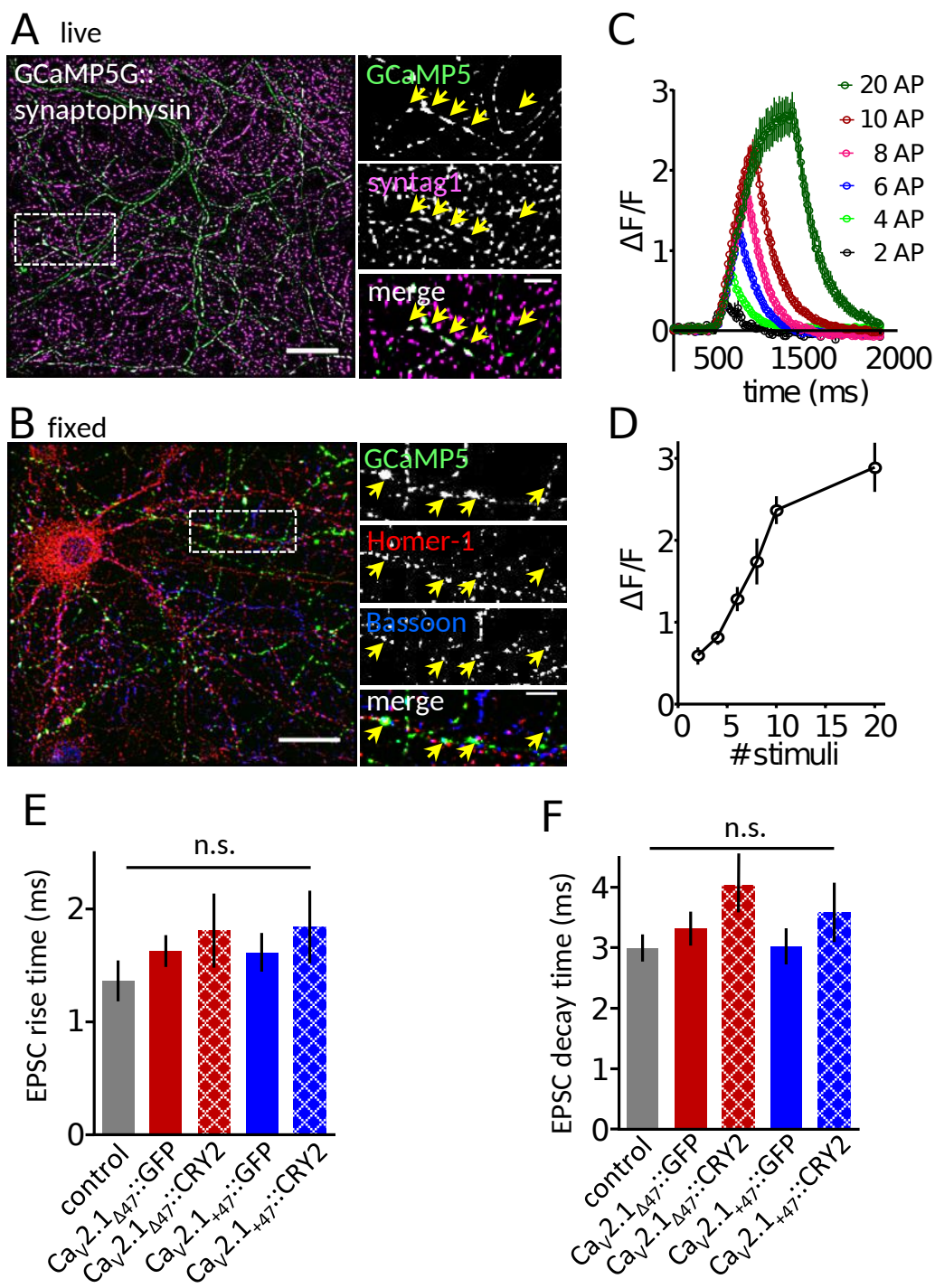
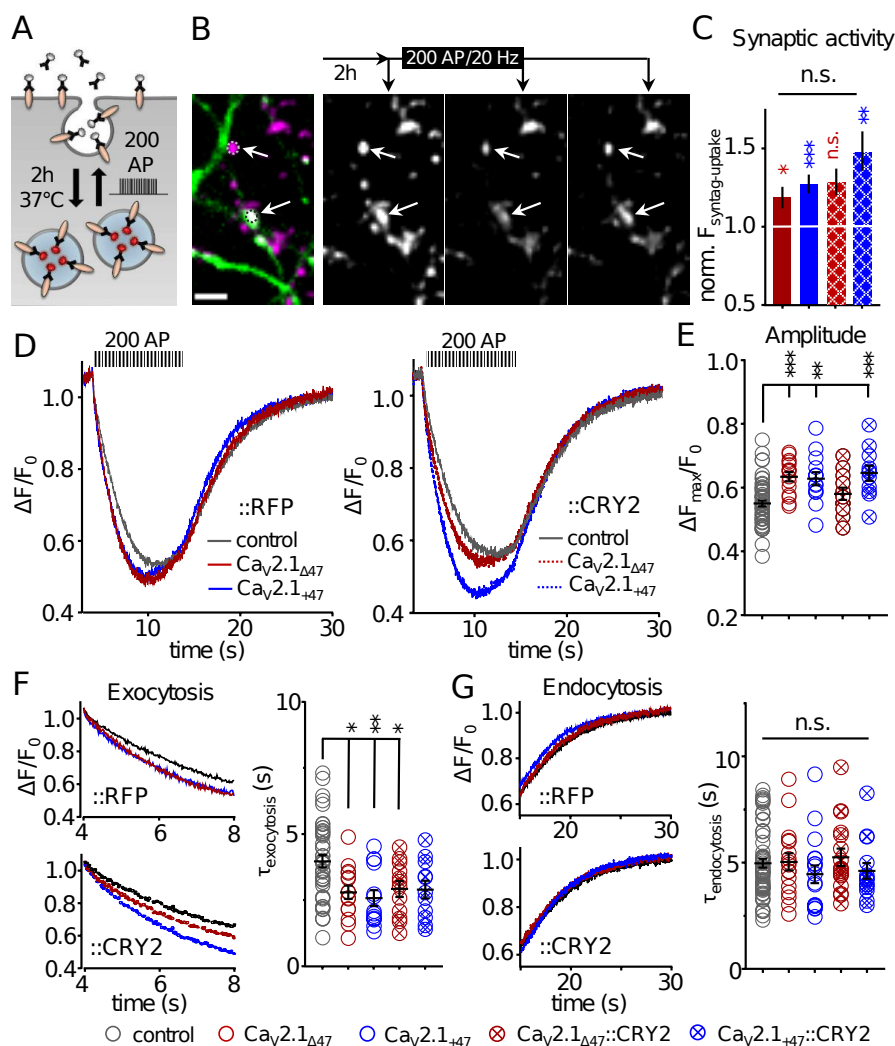


Figure 3.9: (*previous page*) **Characterization of the synaptic localization and function of the genetically encoded calcium sensor GCaMP5G::synaptophysin; Kinetic of EPSCs from paired recordings of monosynaptic coupled neurons.** **A.** Representative image of live rat hippocampal neurons (DIV 19) expressing GCaMP5G::synaptophysin (green). Active synapses were labeled by uptake of anti-synaptotagmin1::CypHer5E (magenta). Expanded regions show co-localization of both presynaptic markers as indicated (arrows). Scale bars correspond to 20  $\mu\text{m}$  (overview) and 5  $\mu\text{m}$  (expanded view). **B.** Representative image of fixed rat hippocampal neurons (DIV17) expressing GCaMP5G::synaptophysin (green). Presynaptic compartments are labelled by anti-Bassoon immunoreactivity (blue) postsynaptic compartments are labelled by anti-Homer-1 immunoreactivity (red). Expanded regions show co-localization of both labels and enrichment of expressed GCaMP5G::synaptophysin. Scale bars correspond to 20  $\mu\text{m}$  (overview) and 5  $\mu\text{m}$  (expanded view). **C.** Fluorescence changes over time for repetitive electrical field stimuli (at 20 Hz) as indicated in the color code. Data are from 3 independent cultures transfected at 3-4 DIV and recorded 15-17 DIV at 37 °C in extracellular solution containing 2 mM MgCl<sub>2</sub> and 2 mM CaCl<sub>2</sub>. **D.** Averaged peak amplitude  $\pm$  SEM of the GCaMP5G::synaptophysin as a function of the number of stimuli. **E.** Mean rise time of EPSCs recorded from postsynaptic neurons that were connected to presynaptic neurons expressing tagged Ca<sub>v</sub>2.1 channels as indicated. **F.** Mean decay time-constant for EPSCs for conditions as indicated. Data correspond to Fig. 3.2I, 3.4M and 3.4N.



**Figure 3.10: Synaptic activity and vesicle exo- and endocytosis in synapses expressing different tagged  $Ca_V2.1$  splice variants.** **A.** Schematic representation of anti-Synaptotagmin1 antibody uptake assay as well as experimental conditions. **B.** Overlay of synaptotagmin1-antibody::CypHer5E labeled synapses (magenta) and  $Ca_V2.1_{+47}::GFP$  expressing axonal structures (green). Next pictures show anti-synaptotagmin1 labelling, before, during and after stimulation with 200 action potential (AP)-like electrical field stimulation at 20 Hz and 37 °C. The scale bar corresponds to 5  $\mu m$ . **C.** Bleach corrected mean quenching curves of the CypHer5E signal in response to 200 AP-like stimuli. **D.** Normalized fluorescent intensity of synaptotagmin1::CypHer5E labelled synapses expressing either  $Ca_V2.1_{\Delta 47}::GFP$  (red),  $Ca_V2.1_{+47}::GFP$  (blue),  $Ca_V2.1_{\Delta 47}::CRY2::GFP$  (red-crossed) and  $Ca_V2.1_{+47}::CRY2::GFP$  (blue-crossed) to not transfected neighbor synapses. **E.** The quenching of the CypHer5E fluorescence correlates with the neutralization of the fluorophore environment through ongoing fusion of SV. The re-uptake is seen as a steady increase in the fluorescence after the indicated stimulus train. Bleach corrected quenching of synaptotagmin1::CypHer5E labelled synapses is shown for GFP and CRY2 tagged  $Ca_V2.1$  channels and non-transfected synapses as indicated in the graphs. **F.** Maximal CypHer5E quenching response of transfected and non-transfected synapses as indicated. **G.** Time constant  $\tau$  of synaptic vesicle exocytosis are calculated by one-phase exponential fitting of the fluorescent quenching during ongoing stimulation. Synapses transfected either with GFP or CRY2 tagged  $Ca_V2.1$  splice variants did show faster exocytosis time constant in comparison to control synapses. **H.** Time constant  $\tau$  of synaptic vesicle endocytosis as recovery of fluorescent signal in quenched synapses. The kinetic of fluorescent recovery was not different between transfected and non-transfected synapses.

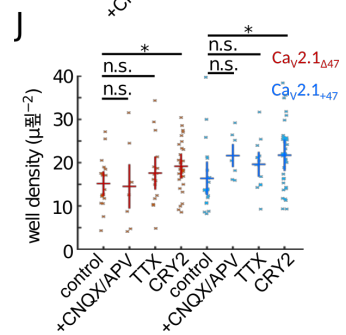
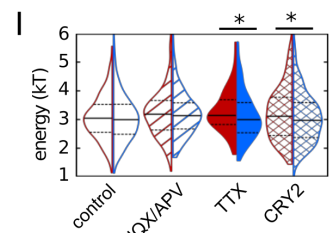
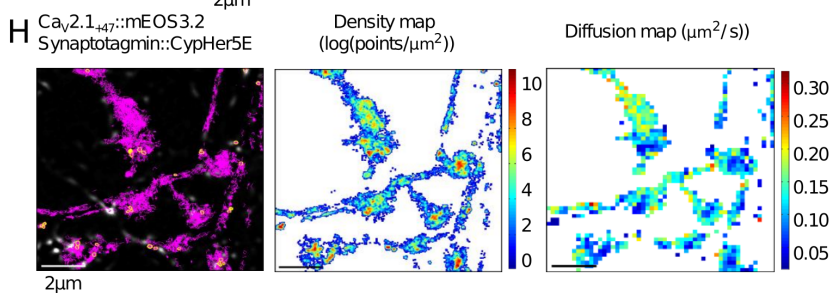
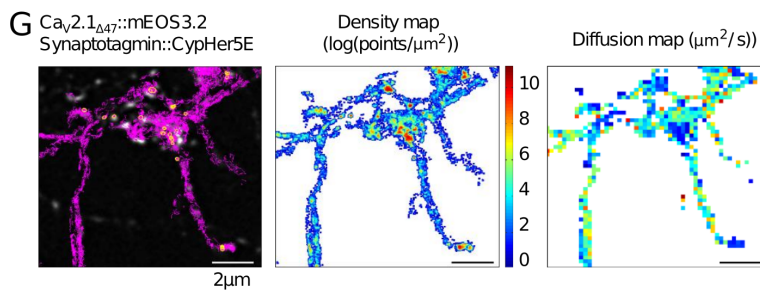
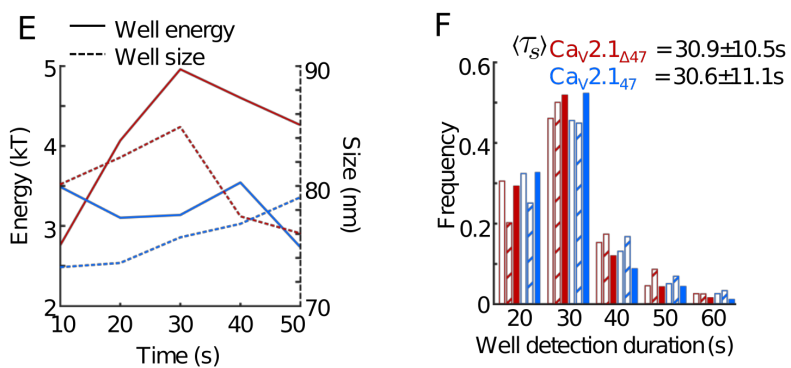
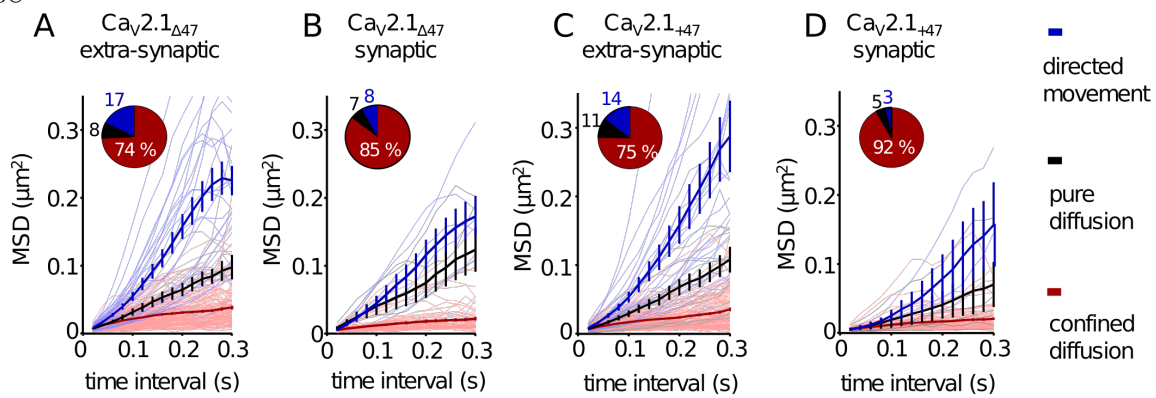
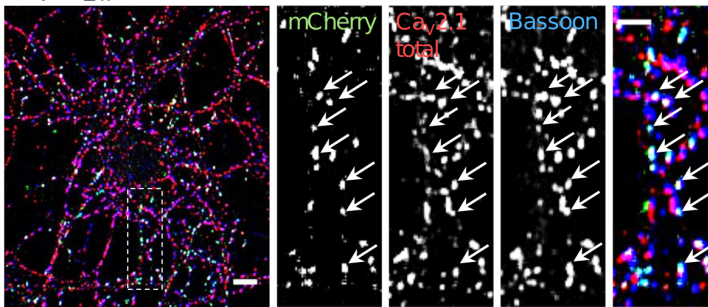


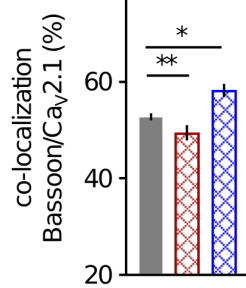
Figure 3.11: (*previous page*) Proportions of the different mobility modes for Ca<sub>v</sub>2.1 splice variants; Potential well stability and number per synapse; Heterogeneous distribution of SPTs and high-density regions of Calcium channels MSD curves of individual Ca<sub>v</sub>2.1 channels were classified according to their modes of motion using a log-log representation of the proportionality between MSD and  $t^\alpha$ , where the  $\alpha$  parameter was then extracted as the curve's slope. The value of  $\alpha$  corresponds to direct motion  $\alpha > 1$  (blue), pure diffusion  $\alpha = 1$  (black) or confined diffusion  $\alpha < 1$  (red) [173]. Shown are the mean MSD curves determined for the modes of motion and their calculated proportion (considering all MSD curves determined as mobile ( $D > 0.002\mu\text{m}^2/\text{s}$ )). **A.** Extra synaptic Ca<sub>v</sub>2.1 $\Delta$ <sub>47</sub>. **B.** Synaptic Ca<sub>v</sub>2.1 $\Delta$ <sub>47</sub>. **C.** Extra synaptic Ca<sub>v</sub>2.1<sub>+47</sub>. **D.** Synaptic Ca<sub>v</sub>2.1<sub>+47</sub>. **E.** Energy (solid) and size (dashed) of the two potential well presented in Fig. 3.3H (same color-code) as a function of the time. **F.** Well longevity vs time for Ca<sub>v</sub>2.1 $\Delta$ <sub>47</sub> red and Ca<sub>v</sub>2.1<sub>+47</sub> (blue) splice variants and different treatments: control (void), CNQX+APV (dashed), TTX (plain). **G., H.** Selected regions presenting trajectory maps of Ca<sub>v</sub>2.1 $\Delta$ <sub>47</sub> (G) and Ca<sub>v</sub>2.1<sub>+47</sub> (H) channels overlaid on top of a synaptotagmin fluorescence images (left panels); the corresponding density map (middle panels) presenting the local point density (in log units); and the corresponding diffusion coefficients map presenting the local diffusion coefficients (right panels). **I.** Energy distribution in potential wells corresponding to the expressed Ca V 2.1 splice variant (red = Ca<sub>v</sub>2.1 $\Delta$ <sub>47</sub>, blue = Ca<sub>v</sub>2.1<sub>+47</sub>) network activity (CNQX/APV, TTX) and cross-link (CRY2). The differences between treatments were tested by a Wilcoxon rank test, data are from 2-7 independent cultures) Density of potential wells along the axonal membranes. Differences between treatments were tested by a unpaired t-test, data are from 2-9 independent cultures.



A

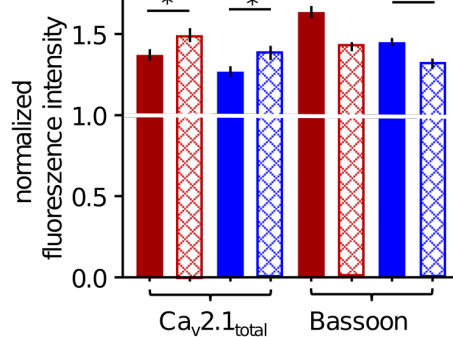
Ca<sub>v</sub>2.1<sub>Δ47</sub>::CRY2::mCherry

B



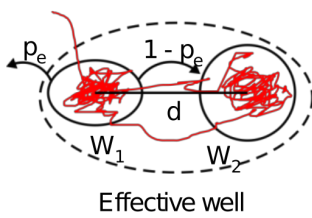
■ endo Ca<sub>v</sub>2.1  
 ▨ Ca<sub>v</sub>2.1<sub>Δ47</sub>::CRY2  
 ▨ Ca<sub>v</sub>2.1<sub>+47</sub>::CRY2

C

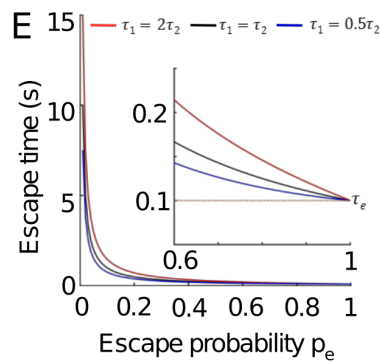


■ Ca<sub>v</sub>2.1<sub>Δ47</sub>::GFP  
 ■ Ca<sub>v</sub>2.1<sub>+47</sub>::GFP

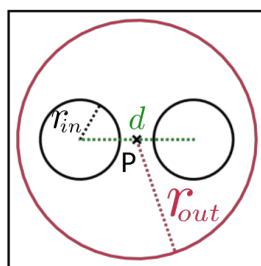
D



E



F



G

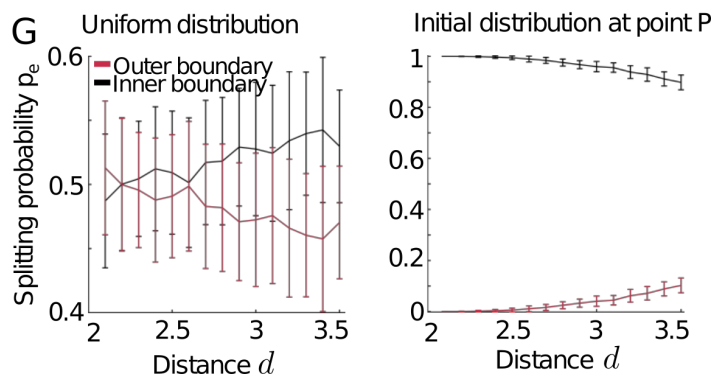


Figure 3.12: (*previous page*) **Effect of CRY2olig-mediated Ca<sub>v</sub>2.1 channels cross-link on synaptic expression of channels and bassoon; Effective potential wells formed by multiple close by potential wells.** **A.** Example image of hippocampal cultured neurons 16 DIV transfected with Ca<sub>v</sub>2.1<sub>Δ47</sub>::CRY2::mCherry (green). Cells were exposed to blue light for 1 min prior fixation to induce Ca<sub>v</sub>2.1 cross-linking. Total Ca<sub>v</sub>2.1 population was labeled via anti-Ca<sub>v</sub>2.1 antibody (red). Presynaptic terminals were labelled with anti-Bassoon (blue). Co-localization of calcium channel clusters and presynaptic structures are indicated by arrows. Scale bars correspond to 10μm (overview) or 5μm (picture section). **B.** Quantification of the co-localization of cross-linked Ca<sub>v</sub>2.2::CRY2 (red, crossed) or Ca<sub>v</sub>2.1<sub>+47</sub>::CRY2 (blue, crossed) with Bassoon in hippocampal neurons 16 DIV in comparisons with the synaptic co-localization of endogenously expressed Ca<sub>v</sub>2.1 (gray). **C.** Effects of the cross-linking of Ca<sub>v</sub>2.1 splice variants (as indicated in B) on the total amount of presynaptic Ca<sub>v</sub>2.1 and expression of Bassoon. The individual fluorescence intensities are normalized to non-transfected synapses from the same field of view and compared to GFP-tagged Ca<sub>v</sub>2.1 splice variants used as non-clustered control. Data are from 3 independent cultures, significance were tested by ANOVA followed by a Holm-Sidak test. **D.** Illustration of the effective wells that can be theoretically formed when cross-linked calcium channels repetitively fall into close-by potential wells, effectively increasing its residence time in the region. **E.** Simulation of the residence time as a function of the escape probability of the effective well presented in D) for different values of the residence times in each well. **F.** Simulation scheme for evaluating the splitting probability  $p_e$  between the boundary of the Active Zone, modeled as a circle of radius  $r_{out}$  and two potential wells of radius  $r_{in}$  and separated by a distance  $d$ . Brownian trajectories are generated outside the wells but inside the Active Zone (AZ). **G.** Estimation of the Escape (splitting) probability  $p_e$  (red curve) and  $1-p_e$  (black curve) for a Brownian channel to hit the boundary of the AZ (red) before one of the wells vs the distance  $d$ . Two cases are presented: when the initial distribution is uniformly distributed (Left) and at the center P(Right) of the AZ. The Results are given as mean and standard deviation over 100 realizations with 100 trajectories ( $\Delta t = 10^{-4}$ s, number of trajectories  $n_{traj} = 100$ , inner radius  $r_{in} = 1\mu m$ , outer radius  $r_{out} = 3\mu m$  and diffusion coefficient  $D = 0.041\mu m^2/s$ ).



## Chapter 4

# Single particle trajectories reveal active endoplasmic reticulum luminal flow

Published as:

David Holcman\*, Pierre Parutto\*, Joseph E. Chambers, Marcus Fantham, Laurence J. Young, Stefan J. Marciniak, Clemens F. Kaminski, David Ron and Edward Avezov, "*Single particle trajectories reveal active endoplasmic reticulum luminal flow*", Nature Cell Biology volume 20, pages 1118-1125 (2018).

(\*: equally)

**Abstract** The endoplasmic reticulum (ER), a network of membranous sheets and pipes, supports functions encompassing biogenesis of secretory proteins and delivery of functional solutes throughout the cell [174, 175]. Molecular mobility through the ER network enables these functionalities, but diffusion alone is not sufficient to explain luminal transport across supramicrometre distances. Understanding the ER structure-function relationship is critical in light of mutations in ER morphology-regulating proteins that give rise to neurodegenerative disorders [176, 177]. Here, super-resolution microscopy and analysis of single particle trajectories of ER luminal proteins revealed that the topological organization of the ER correlates with distinct trafficking modes of its luminal content: with a dominant diffusive component in tubular junctions and a fast flow component in tubules. Particle trajectory orientations resolved over time revealed an alternating current of the ER contents, while fast ER super-resolution identified energy-dependent tubule contraction events at specific points as a plausible mechanism for generating active ER luminal flow. The discovery of active flow in the ER has implications for timely ER content distribution throughout the cell, particularly important for cells with extensive ER-containing projections such as neurons.

## 4.1 Introduction

The endoplasmic reticulum (ER) is a contiguous network of membranous sheet-like reservoirs and tubes extending throughout the cell. Maintained by membrane-shaping proteins [174, 175], this morphology supports the distribution of ER luminal content to distant sites. The ER-content distribution rate affects the efficiency of ER-mediated intracellular connectivity. Perturbation of this fundamental process may contribute to diseases caused by mutations in ER-shaping proteins [176, 177]. Measurements of ER-luminal protein mobility using fluorescence recovery after photobleaching (FRAP) have previously uncovered an energy dependence that is difficult to reconcile with passive diffusion [178, 21]. An increase in luminal crowding due to the incapacitation of ATP-dependent ER chaperones has been suggested as a plausible explanation for this effect (direct crowding measurements here do not detect such an effect). Addressing this paradox remained challenging since FRAP measurements report on bulk mobility, and do not inform as to the nature of forces driving mobility at a molecular level. A passive diffusion model for luminal transport is also challenged by the notion that traversal time of random walking molecules increases sharply with distance. This poses kinetic limits for material exchange in an expanded ER network. Recent advances in super-resolution microscopy afford a basis for development of a single particle tracking (SPT) approach to provide a detailed description of molecular motion in the ER lumen, with potential to generate a wealth of information regarding directionality and velocity from a large number of simultaneous single-molecule displacement events. Using live-cell super-resolution microscopy, we visualized and analyzed single molecule trajectories traversing tubular ER, and the organelle's real-time morphological dynamics.

## 4.2 Results

First we examined, in our experimental system, the energy dependence of luminal protein mobility, previously demonstrated for green fluorescent protein (GFP) using FRAP [21]. The escape rate of photoconvertible fluorescent protein (Dendra2-ER) from a region of activation was attenuated by ATP depletion (Fig. 4.1a). This may reflect either an increase in resistance to motion or a decrease in active transport of proteins on energy starvation. The former is contradicted by measurements of ER crowdedness, using a sensitive fluorescence resonance energy transfer (FRET)-based probe [179], that, over a broad range of expression levels, remained unaltered by ATP depletion (Fig. 4.1b-d and Fig. 4.6a). Furthermore, displacement of Dendra2 proteins (initially localized as a 'packet' in a small volume of the tubular ER that had been subjected to a colour-photoswitching laser pulse) occurred with variable speed and had a conspicuous unidirectional component (Fig. 4.1e). These features are inconsistent with Brownian motion and suggest, instead, active transport.

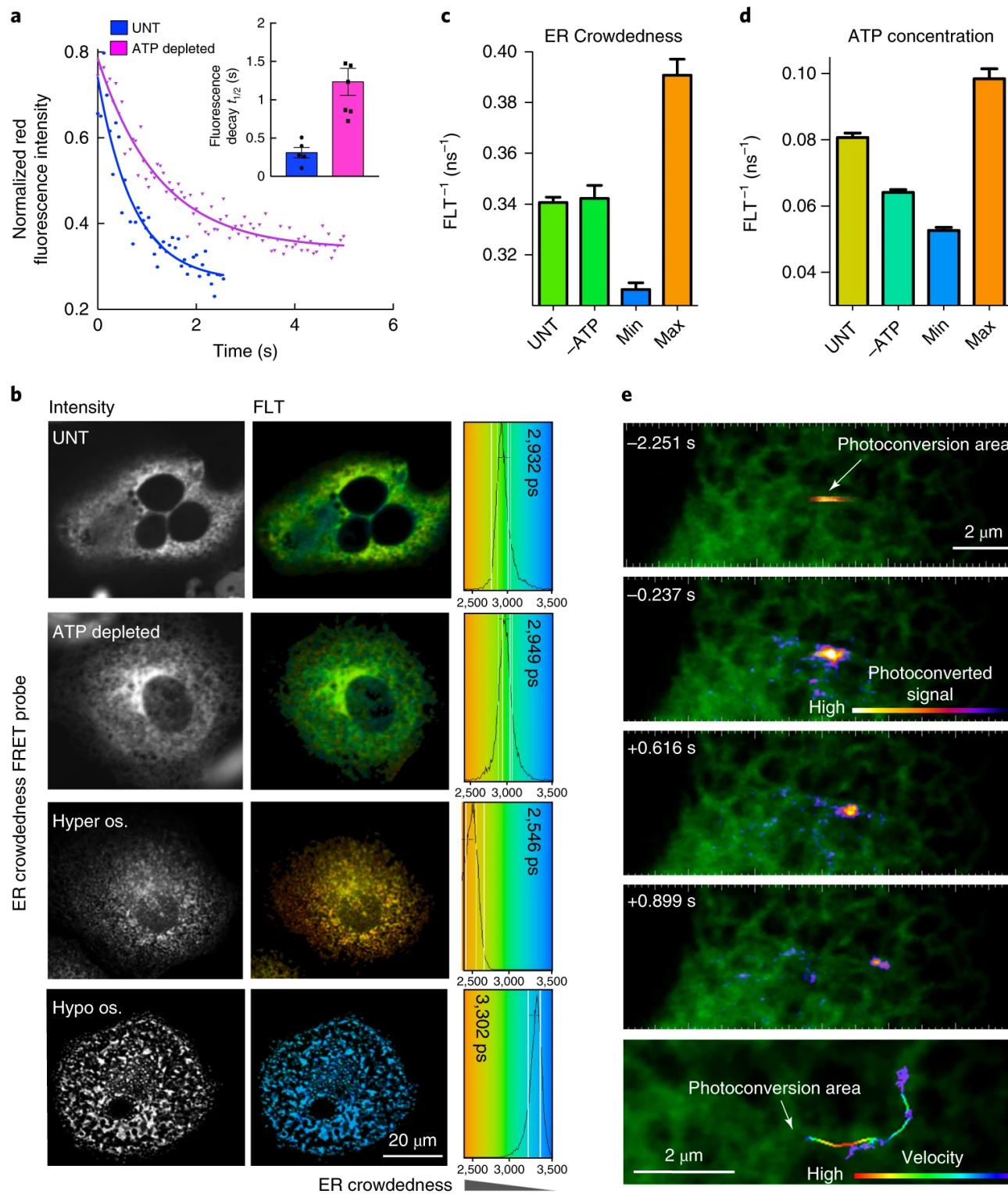


Figure 4.1: (*previous page*) **A.** A trace of time-dependent decay in the intensity of the fluorescence signal from an ER-localized photoconvertible protein, Dendra2-ER, after a pulse of photoconverting illumination delivered to a small patch of untreated or ATP-depleted COS7 cells. The inset denotes mean  $\pm$  s.e.m. ( $n = 5$  traces per condition) of fluorescence decay half-time, reflecting the probe's escape from the photoconversion area. **B.** Fluorescence intensity (left) and colour-coded fluorescence lifetime (FLT) images (right) of COS7 cells transfected with an ER-localized molecular crowding probe<sup>7</sup>. The FLT distribution within imaged cells is displayed in colour-coded histograms with mean FLT noted (in picoseconds, right). Cells were left untreated (UNT), ATP depleted or treated with hyper-osmotic (Hyper os.) or hypo-osmotic buffers that induce cell shrinking or swelling to obtain maximal and minimal crowding values, respectively. Shown are characteristic images observed in three independent repeats. **C.** A bar diagram of FRET-donor FLT values (in nano seconds, ns) measured as in b (mean values  $\pm$  s.d.,  $n = 22$  independently sampled cells). **D.** A bar diagram of relative intracellular ATP concentration measured with a FRET-based ATP probe (ATeam) [180] in cells untreated or ATP-depleted as in **a** and **B**. The minimum and maximum values represent the probe readings in ATP-depleted or saturating conditions, respectively, imposed in semi-permeabilized cells. Shown are mean values  $\pm$  s.d.,  $n = 10$  independently sampled cells. **E.** Images of a COS7 cell expressing Dendra2-ER. A brief pulse of illumination-photoconverted Dendra2 from green to red in a restricted region of the ER. The progression of the photoconverted packet of proteins is revealed by the time series and summated in the magnified image in the bottom panel with its velocity colour-coded.

To characterize motion in the ER lumen, we performed real-time SPT in live cells, acquiring trajectories at optical super-resolution. Imaging at 56 Hz, we recorded single molecule fluorescent signals from functional (Halo-tagged ER luminal chaperone, calreticulin, Crt) and inert (ER-targeted HaloTag) proteins, both sparsely labelled with a fluorescent ligand (chloroalkane-tetramethylrhodamine, TMR). The brightness and photostability of the TMR ligand enabled the tracking of single particles over longer trajectories than genetically encoded fluorophores (Fig. 4.6b), thus establishing spatio-temporal correlations in motion patterns over extended periods of time. This approach offered information on directionality and instantaneous velocity at a single molecule level, not available in bulk methods such as FRAP. By implementing a single molecule localization algorithm, images reconstructed from the SPT series revealed a pattern typical of the ER network, confirming ER-localization of the HaloTag-Crt and washout of the unbound dye (Fig. 4.2a,b).

Trajectories were generated from single-molecule time series by sequentially implementing spot detection and tracking algorithms [181, 107]. ER geometry constrains particle movement to a narrow tubular network, which limits overlapping trajectories, contributing to the tracking algorithm's ability to faithfully trace many molecules simultaneously. Resulting trajectories reconstruct a map recognizable as a pattern of ER tubes and their connecting reservoirs (Fig. 4.2c), reflecting tracking fidelity. Notably, spatio-temporal particle distributions were non-uniform, with higher time-integrated abundance in the tube-connecting reservoirs (Fig. 4.2d). This heterogeneity correlated with spatial distribution of instantaneous velocities, revealing distinct subgroups: relatively slow-moving particles predominantly detectable in segments of trajectories mapped to the nodes; and particles with relatively high and variable velocities mapped to node-connecting tubes (Fig. 4.2c). These

characteristics could be observed for two different markers (Crt and HaloTagER) and three different cell types (HEK-293T, COS-7 and SH-SY5Y, Table 4.1). The long, single-tailed velocity distribution observed (histogram, Fig. 4.2c) is incompatible with Brownian motion (Fig. 4.2c, modelled by a solid line). Furthermore, ATP depletion led to selective loss of the fast-moving population (Fig. 4.2e). These observations suggest that diffusive motion is manifested by the slow-moving particles in the tube-connecting reservoirs (nodes), while the rapidly moving particles in the tubes are subjected to an ATP-dependent propulsive force, resulting in an ER luminal flow. The displacement profile observed fits significantly better (Fig. 4.2c inset) a bi-modal diffusion and flow model (Fig. 4.2c, dashed line) than a purely diffusive one (Fig. 4.2c, solid line).

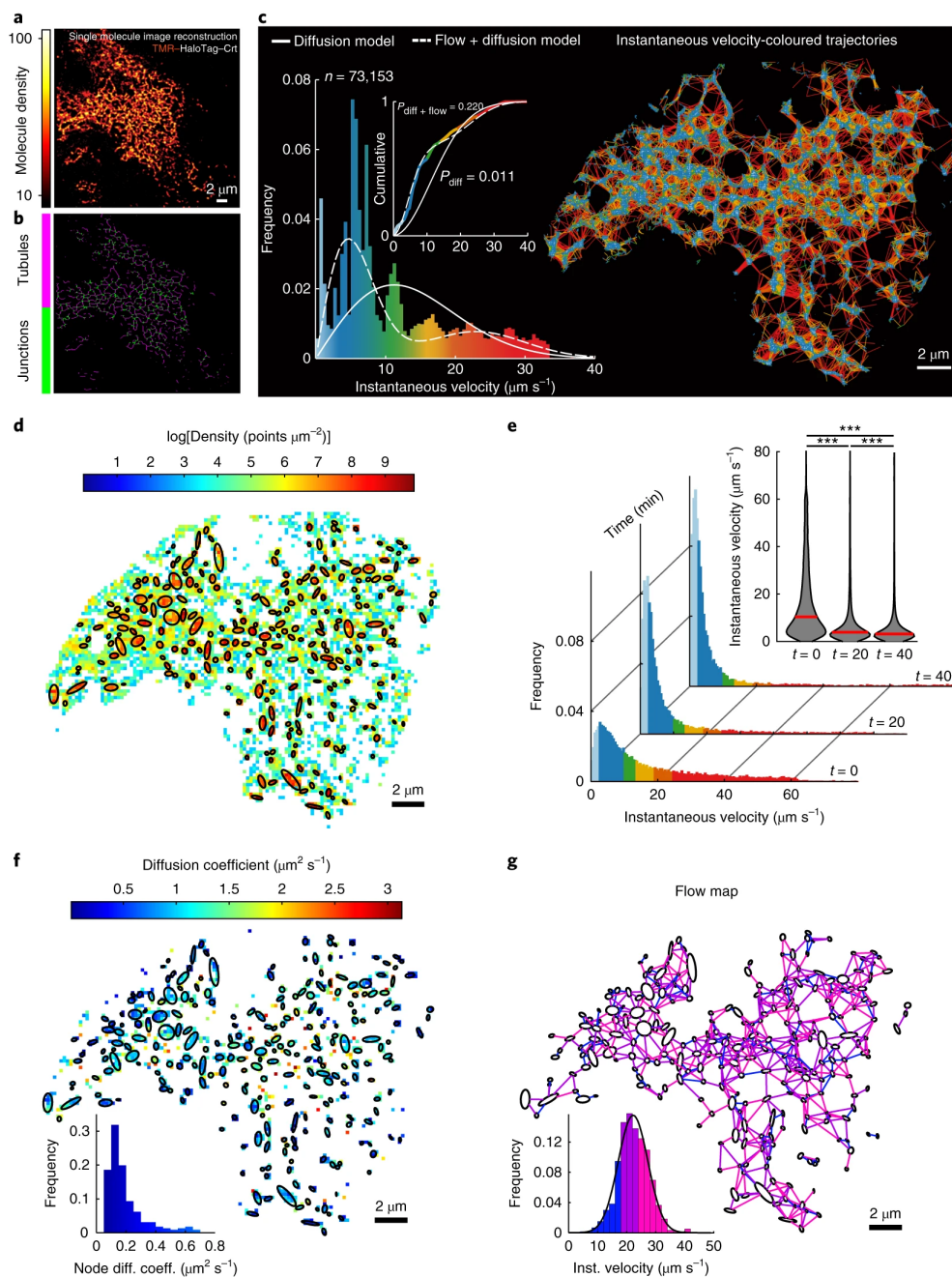




Figure 4.2: (*previous page*) **A.** An image reconstructed from single molecule localizations of TMR-labelled Halo-tagged Crt, in HEK-293 cells, rendered with a molecular density colour code. **B.** A skeletonization view of the image in **A.** Shown are representatives of  $n = 3$  independent experiments. **C.** Single molecule trajectories generated using a particle-tracking algorithm from the time series of image **A.**, colour-coded according to the instantaneous velocity distribution shown in the histogram. Overlaid traces: the velocity distribution simulated assuming exclusively diffusion-driven motion (solid line, using apparent  $D$  from **F**), or a combination of diffusion and flow (using  $D$  and flow rate from **F** and **G**). Inset: cumulative distribution, Kolmogorov-Smirnov test of observed versus expected distributions. **D.** A density map computed for a grid of square bins (sides of  $0.2\mu\text{m}$ ) imposed on the particle displacement map. The ellipses mark boundaries of higher-density regions (correspond to tube-connecting reservoirs/junctions). **E.**, Histograms of instantaneous velocity frequency distributions of SPT from a cell before/after ATP depletion (as in 4.1a-c). Inset: a violin plot presenting the medians (red bars) and density (grey) of the distributions. A two-sided Mann-Whitney U-test was used to compare the median of each pair of distributions ( $***P < 1 \times 10^{-3}$ ),  $P(0 - 20\text{min}) = 1 \times 10^{-80}$ ,  $P(20 - 40\text{min}) = 9.889 \times 10^{-64}$ ,  $P(0 - 40\text{min}) = 1 \times 10^{-80}$ ;  $n = 20, 526$ ,  $n = 14, 591$  and  $n = 10, 108$  trajectory displacements respectively. **F.**, A diffusion map extracted from the empirical estimator of the Langevin equation (equations (1.13) and (1.20)) and computed from a square grid as in **d**. Inset: distributions of the diffusion coefficients inside nodes (avg  $\pm$  s.d. =  $0.19 \pm 0.13$ ,  $n = 226$  nodes). **G.**, Flow map computed by averaging non-Brownian velocity jumps of particles moving between pairs of neighbouring nodes identified in **D** and colour-coded according to the inset histogram. Inset: the distribution of the average instantaneous velocity between pairs of neighbouring nodes ( $n = 705$  node-pairs; avg  $\pm$  s.d. =  $22.90 \pm 6.92$ ).

Next we quantified temporal coordinate changes of HaloTag-Crt by analysing SPT data using the overdamped limit of the Langevin model (where velocity is described as the sum of diffusional and drift forces, and motion parameters are estimated from local statistics of the displacement, see Methods) [76, 79, 17, 61]. Motion is described by the stochastic model (a sum of directed and diffusional motion terms):

$$\dot{\mathbf{X}} = \mathbf{b}(\mathbf{X}) + \sqrt{2D(\mathbf{X})}\dot{\mathbf{w}} \quad (4.1)$$

where  $\mathbf{w}$  is a white noise which source is the ambient thermal agitation, while the term  $\mathbf{b}$  is the local drift vector field and  $D$  is the effective local diffusion coefficient. The local estimators (sub-section 1.4.2) allow to recover  $D$  and  $\mathbf{b}$  from large numbers ( $\sim 10^4$  per cell acquisition) of single particle trajectories.

The global nature of this statistical approach considers not only particle displacement speed but also direction patterns, extracted from a large number of trajectories repeatedly traversing the same regions, thereby unmixing the contribution of Brownian and deterministic forces. This computation identifies a slow diffusional ( $D = 0.19 \pm 0.13\mu\text{m}^2/\text{s}$ ) component that maps predominantly onto the nodes (Fig. 4.2f). The relatively fast movement of internode particles required an additional component to account for their directionality and persistence, consistent with a propulsive force with normal distribution of velocity amplitudes ( $22.9 \pm 6.92\mu\text{m}/\text{s}$ , Fig. 4.2g). The super-diffusive nature of particle motion in ER tubules was further confirmed by analysis of time-averaged mean squared displacement (MSD) of

SPT. The time-averaged MSD can be described by  $\sim t^\alpha$ , where the anomalous exponent  $\alpha$  defines motion as sub-diffusive if  $\alpha < 1$ , Brownian/diffusive if  $\alpha = 1$  and super-diffusive if  $\alpha > 1$ . Conducted on the entire ensemble of trajectories, the MSD analysis revealed a broad range of particle behaviours ( $0 < \alpha < 1.5$ , Fig. 4.7b), whereas the same analysis restricted to trajectory fragments located in nodes revealed clearly confined diffusion dynamics ( $\alpha < 0.8$ , Fig. 4.7c). Trajectories of particles moving outside the nodes exhibited super-diffusive dynamics ( $\alpha > 1$ , Fig. 4.7d). These results are consistent with the active motion mode identified in the analysis of Fig. 4.2.

Similar observations were made in green monkey kidney (COS-7) and human neuroblastoma (SH-SY5Y) cell lines, attesting to their broad validity in describing ER flow dynamics and its spatial organization (Fig. 4.8, Table 4.1). Measurements of the motion parameters of a lower-mass ER-localized protein, HaloTag-KDEL, showed similar values to those observed with the tagged Crt (Fig. 4.8 and Table 4.1). Flow velocity was slightly higher in COS cells than in HEK-293T and SH-SY5Y cells. The behaviour of the luminal ER markers, HaloTag-KDEL and Crt, contrasted with that of a membrane-associated analogue of the latter, mEOS2-calnexin: its velocities were distributed relatively homogeneously through the ER network (Fig. 4.3a,b), lacked the thick tail of high values in distribution of instantaneous velocities and fitted well to a purely diffusional model (Fig. 4.3a-d).

To establish whether recently reported ER macrostructure motion characteristics [1] are reflected in the SPT analysis, we focused on their numerical parameters. Motion of ER tubules characterized by their relatively slow transverse oscillation (4 Hz, with an amplitude  $< 50$  nm, which translates to velocity  $< 0.2\mu\text{m/s}$  [1]), does not significantly contribute to the relatively fast velocities of flow-assisted marker particles moving along the tubules ( $27 - 42\mu\text{m/s}$ , Table 4.1). Junction fluctuations contribute a similarly insignificant component to the diffusional motion inside the junction since the diffusion coefficient calculated from tracking of whole junctions was 69 times slower than the mean diffusion coefficient computed for single molecules [1] (Fig. 4.9). Furthermore, contribution of tubule growth to single particle trajectories was found to be negligible, with a mean percentage of tubules growing at any given time of  $0.14 \pm 0.04\%$ .

Consistently, considering trajectory motion as a purely diffusive process yielded an apparent diffusion coefficient of  $1.13\mu\text{m}^2/\text{s}$  (Fig. 4.7a) similar to that previously estimated by FRAP [178, 182, 183].

Although most individual trajectories visited only a limited number of nodes (Fig. 4.4a), an oriented network graph analysis, which identifies directly or proxy interconnected junctions through trajectory directions [184], revealed that, regardless of their starting point, particles have the potential to visit almost the entire ER network (Fig. 4.4b; the disconnected periphery is probably contributed by signals from neighbouring cell(s)). This analysis is consistent with the notion that the ER network maintains a luminal continuum. The ER appeared to be in a state of equilibrium, with nodes, on average, connected by an equal number of inward and outward trajectories (Fig. 4.4c,d). These findings are consistent with an interconnected system of flows that preserves the content across the ER.

Closer scrutiny of the directionality of individual tube-traversing particles suggested a pattern whereby the direction of visible flow alternates with variable frequency (switching on

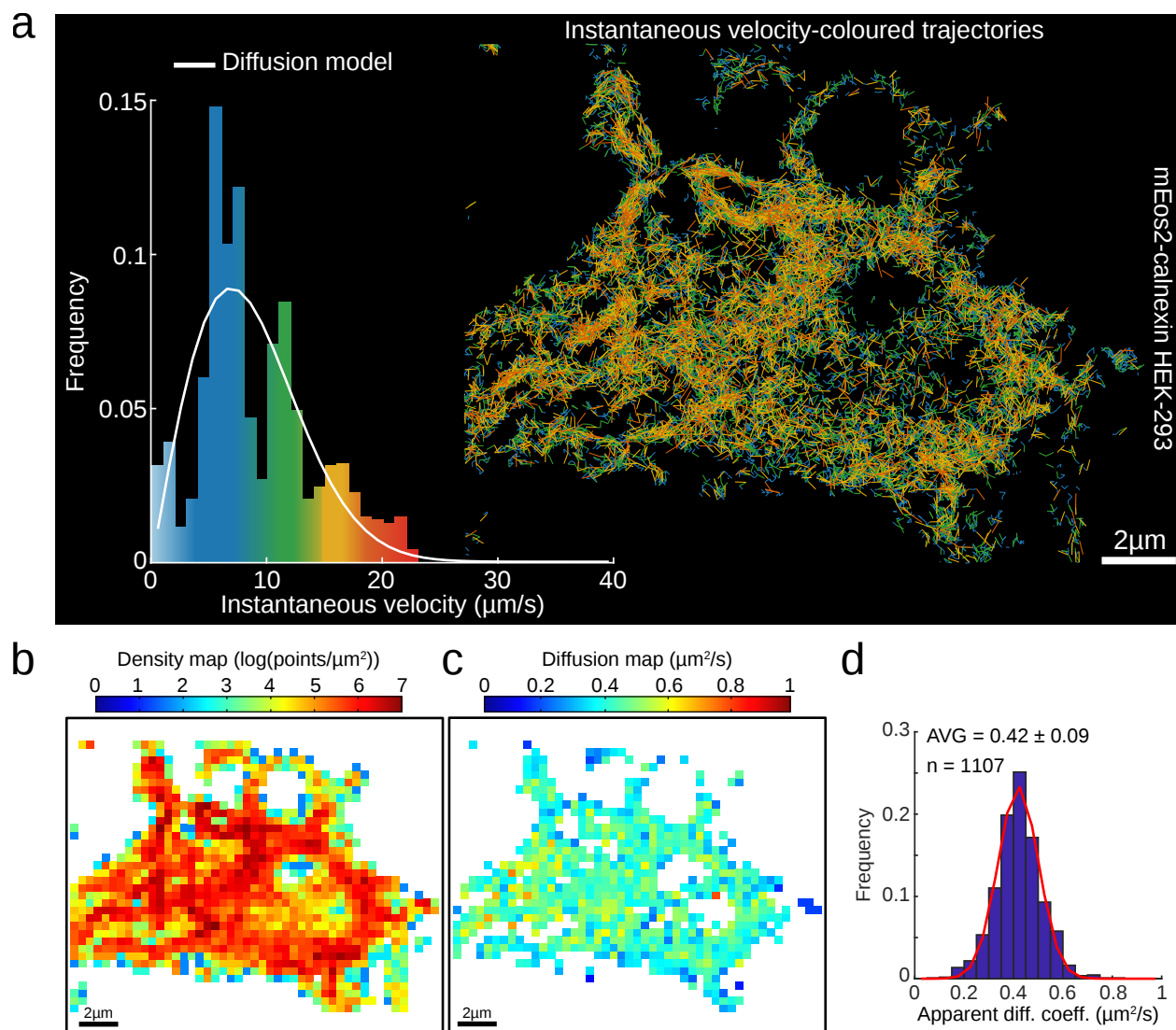


Figure 4.3: **A.** Single molecule trajectories of mEOS2-calnexin expressed in HEK-293 cells generated as in Fig. 4.2, colour-coded according to instantaneous velocity distribution. Inset: instantaneous velocity distribution histograms computed from the displacements extracted from the trajectories and overlaid by the expected distribution for a purely diffusive motion with the diffusion coefficient extracted from **D.** **B.**, Density map computed for a grid of square bins (sides of  $0.4\mu\text{m}$ ) imposed on the particle displacement map. **C.**, A diffusion map extracted from the empirical estimator of the Langevin equation (equation (1)) and computed from the same square grid as in **B.** **D.**, Histograms of diffusion coefficients computed from individual square bins, pooled from two cells, for the entire domain as presented in **C.** The red curve on top of the diffusion histogram corresponds to a fit (trust region reflective algorithm) to a Gaussian distribution with  $\mu_D = 0.42\mu\text{m}^2/\text{s}$ ,  $\sigma_D = 0.12\mu\text{m}^2/\text{s}$  and a determination coefficient  $R^2 = 0.986$ . Descriptive statistics are given as avg  $\pm$  s.d.

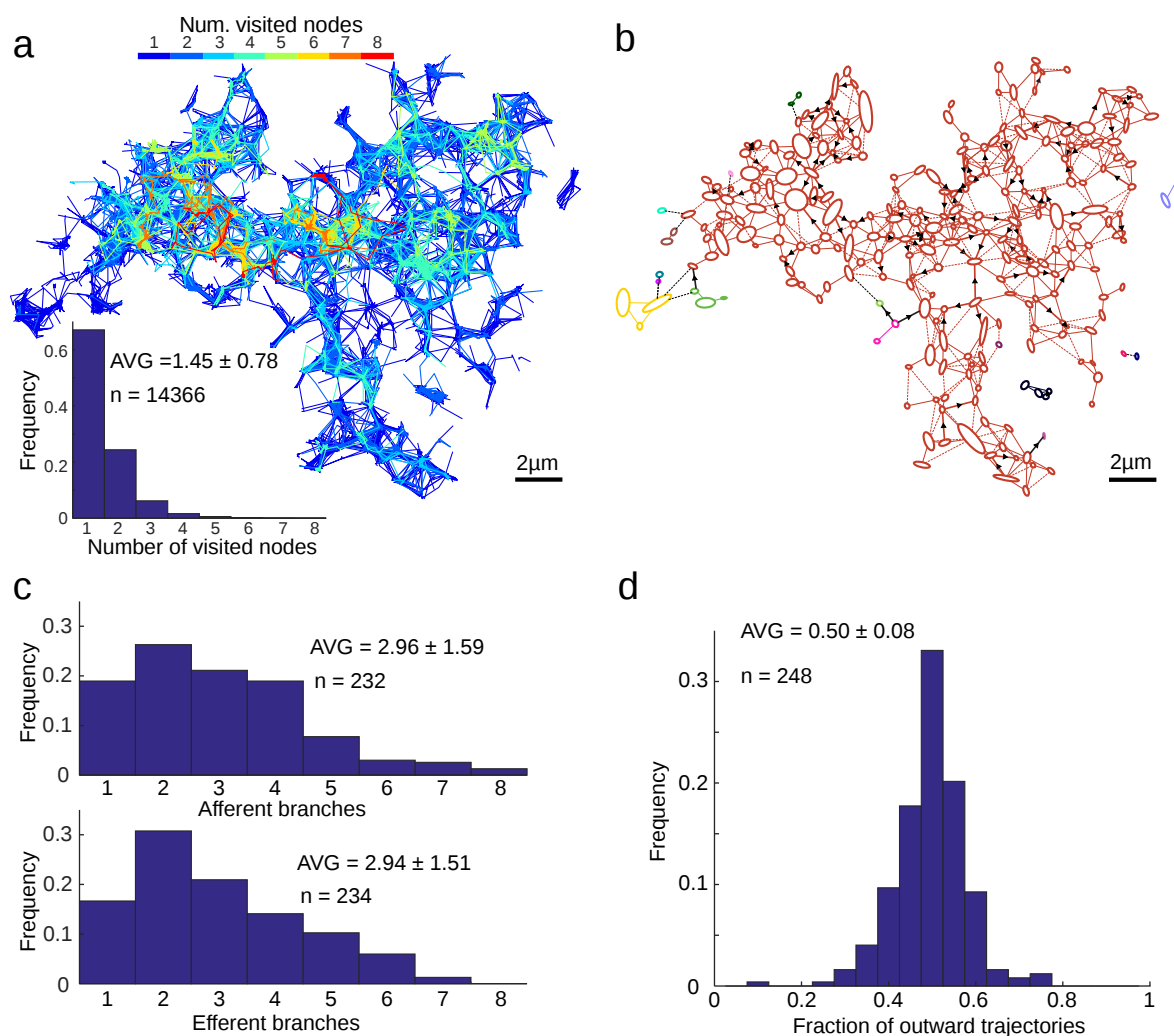


Figure 4.4: **A.** Number of nodes visited by individual particles. Trajectories map, as in Fig. 4.2, colour-coded according to the number of nodes visited by a particle; and the distribution of the number of nodes visited by each individual trajectory (excluding trajectories visiting zero nodes). **B.** Vectorial representation of the ER network from Fig. 4.2 analysed using oriented network graph analysis, to assess the direct or proxy, uni/bi-lateral trajectory-connectivity of the nodes, assigning a single colour for each interconnected area. Note a strong connected component resulting in a monochromatic appearance of almost the entire network. The arrows denote prevalent displacement directionality (detected in 18% of tubes), defined as such if the steady-state ratio of flow in one direction versus the total flow exceeds 0.75. The dashed links represent flows whose directionality could not be determined due to an insufficient number of displacement events. **C.** Distribution histogram of the number of outward-directed (efferent) and inward-directed (afferent) branching for individual nodes. Efferent branches were defined as the number of nodes, reached by the outward trajectories originating in the examined node, in the time-integrated map; accordingly, afferent branches reflect the number of nodes-of-origin for the trajectories arriving at the examined node. **D.** Distribution of the fraction of exiting trajectories for each node. All values are given as  $\text{avg} \pm \text{s.d.}$

average every 4s, maintained for up to 14s, Fig. 4.5a,b); and particles accelerated periodically following their exit from a node, reaching brief velocity peaks that lasted up to 120 ms (Fig. 4.5c,d). Intervals between velocity peaks and flow-directionality alternations were distributed stochastically (Fig. 4.5b,d), suggesting that flow-inducing events (for example, transient tubule contractions, discussed below) are not produced by synchronized oscillators and are therefore not centrally coordinated. However, we cannot exclude the contribution of synchronization processes whose phase is lost, as trajectories are recorded asynchronously. Note that temporal profiles of directionality are not available using low spatio-temporal resolution approaches (for example, FRAP or photoconversion pulse-chase, Fig. 4.1e).

### 4.3 Discussion

The oscillatory luminal motion suggested the possibility of nanoperistalsis-like [185] propulsion, attainable by tubule contractions. To test this, we obtained high-resolution time-series images of the ER tube structure of live cells by fast structured illumination microscopy (SIM) [1, 186, 187]. These revealed transient, asynchronous constriction of the tubes at specific locations (Fig. 4.5e, Fig. 4.10a,b), consistent with a role for tube constriction in generating flow. Constriction-driven propulsion is also consistent with the observed velocity values and variation of packets of photoconverted Dendra2 during their deterministic traversal of the tubular ER ( $V_{\max} = 19.9\mu\text{m/s}$ , Fig. 4.1e). Furthermore, the frequency of contraction events decreased fourfold on ATP depletion (Fig. 4.5f). It is expected that following an individual contraction event (with a frequency of hundreds of milliseconds), both deterministic and acceleratory displacement of multiple particles would be detected, as SPT acquisition operates at approximately ten times the contraction frequency. Assuming uncoordinated contractions throughout the tubular network, consecutive contractions have the same probability to preserve or invert the direction of the next set of detectable SPT events (consistent with observed distribution of directionality preservation time, Fig. 4.5b).

The notion that tubule contractions generate high-velocity peaks in luminal particles is supported by the fact that their temporal distributions are both Poissonian (Fig. 4.5d,g), indicating compatible physical processes. The larger time constant of contractions ( $\sim 900$  ms) compared to that of high-velocity peaks ( $\sim 80$  ms) is expected since several contraction points may contribute to the particles' acceleration incidence.

Furthermore, a physical model simulating forces resulting from tubule contraction, and based on their empirical characteristics (Fig. 4.5h and sub-section 4.5.2), predicts flow velocities of  $10 - 40\mu\text{m/s}$ , in agreement with the high velocities observed in SPT (Fig. 4.5c). Notably, the contraction frequency is low enough to avoid coinciding proximal contractions that may cancel the local flows (probability of simultaneous contractions of two points = 0.022, sub-section 4.5.3). While the existence of a mechanism for spatio-temporal coordination of the contraction events cannot be ruled out, our findings indicate that an uncoordinated system, inducing fast local currents with alternating directionality inside the tubular network, is sufficient to ensure a rapid luminal content homogenization/distribution and thereby overcome a critical kinetic limitation of passive diffusion as a mechanism for ER content mixing in large cells.

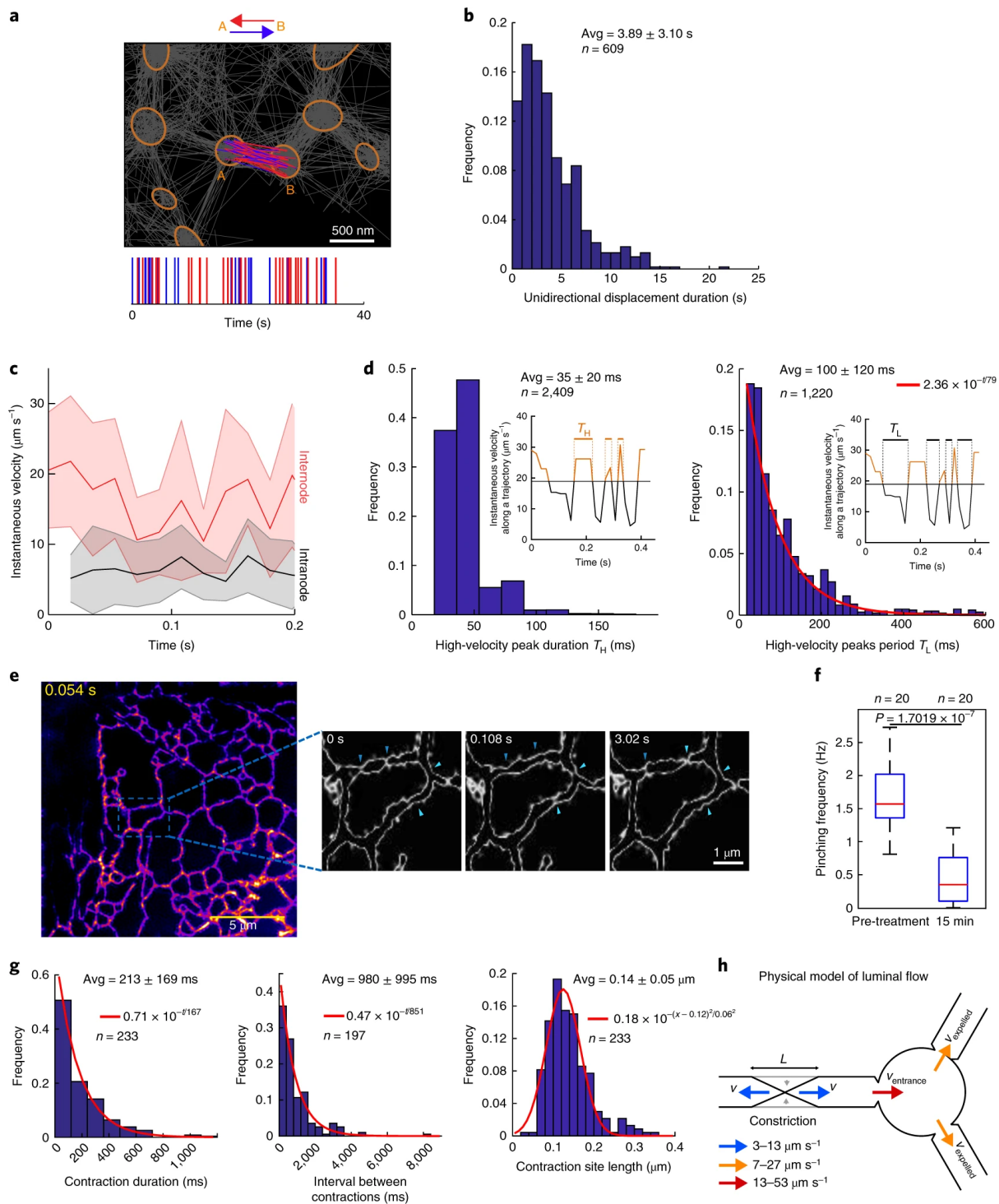


Figure 4.5: (*previous page*) **A.** Analysis of particle trajectories’ directionality. Tubular junctions/nodes are denoted by orange ellipses; grey lines denote all particle trajectories. Trajectories connecting two nodes indicated as A and B are colour-coded according to their direction either in red (B to A), or blue (A to B). The lower graph represents the temporal pattern of traversal-directionality. The results shown are representative of  $n = 108$  node-pairs. **B.** Distribution of time periods of unidirectional inter-node displacement. **C.** Plots of instantaneous particle velocities fluctuations. Velocities of particles following departure from nodes and travelling along tubules (between nodes, red), and those of particles residing within nodes (black). The solid lines represent mean values for all trajectories, and shaded regions represent s.d. of total sample size:  $n = 111$  internode and  $n = 140$  intranode trajectory displacements. **D.** Analysis of time duration  $T_H$  of high-velocity ( $v > 20\mu\text{m/s}$ ) peaks (left) and time interval  $T_L$  (right) between high-velocity peaks. Insets denote  $T_H$  and  $T_L$  of a trajectory. The red line represents an exponential fit ( $R^2 = 0.998$ ). **E.** High-speed SIM super-resolved images of the tubular ER in live COS7 cells stained with an ER membrane dye (ER tracker green). Images were acquired in 54 ms intervals and processed as described in Methods. The resulting SIM reconstructions were colour-coded according to intensity. The magnified area shows the contours of ER tubules at higher magnification. The arrows denote positions where transient contraction events occur repeatedly. Tubule contractions are better visualized in COS7 cells, but detectable in HEK-293 cells too (Fig. 4.10). The results shown are representative of  $n = 5$  independent experiments. **F.** Box plot of tubule contraction frequencies extracted from high-speed SIM time series as shown in **E** before and after ATP depletion. Red line, median; boxes’ bottom/top edges, the 25th and 75th percentiles, respectively; whiskers, extreme data points. Two-sided Mann-Whitney U-test  $P = 1.7019 \times 10^{-7}$ ,  $n = 20$  ER tubules. **G.** Distributions of contraction duration, intervals and lengths from SIM videos as in **e** and Fig. 4.10. Red curves: exponential (left and middle) and Gaussian (right) fits ( $R^2 = 0.988$ ,  $R^2 = 0.969$ ,  $R^2 = 0.937$ , respectively). **H.**, Schematic representation of the model for estimating tubule contraction-induced particle velocity. All values are given as  $\text{avg} \pm \text{s.d.}$  for the noted  $n$ .

Localized contraction of ER tubules, leading to ER deformation, was observed during calcium manipulation [188] (Fig. 4.10c), or phototoxicity (Fig. 4.10e), both reversible processes affecting ER morphology (Fig. 4.10e). These super-resolved images of the ER’s structural dynamics under severe experimental perturbation highlight a potential for ER tubes to contract, revealing that fragments of the perturbed ER that had lost their tubular structure displayed characteristically slow-velocity diffusional motion (Fig. 4.10d). Other physiological membrane dynamic processes involving molecular motors, vesicular fusion and budding, network oscillation [1] and even tube elongation/withdrawal may also contribute to flow and warrant further investigation.

Regardless of their origin(s), the alternating luminal currents described here are well suited to serve as a mixing device, enhancing distribution of ER content throughout the cell. Given that diffusion-driven connectivity (matter exchange rate) decreases sharply with distance, it is expected that the active process described here would be especially important in cells with extensive ER projections, such as motor neurons. It is therefore tempting to speculate that perturbed luminal flow might contribute to diseases such as hereditary spastic paraplegia, associated with defective ER membrane-shaping proteins [176, 177, 189].

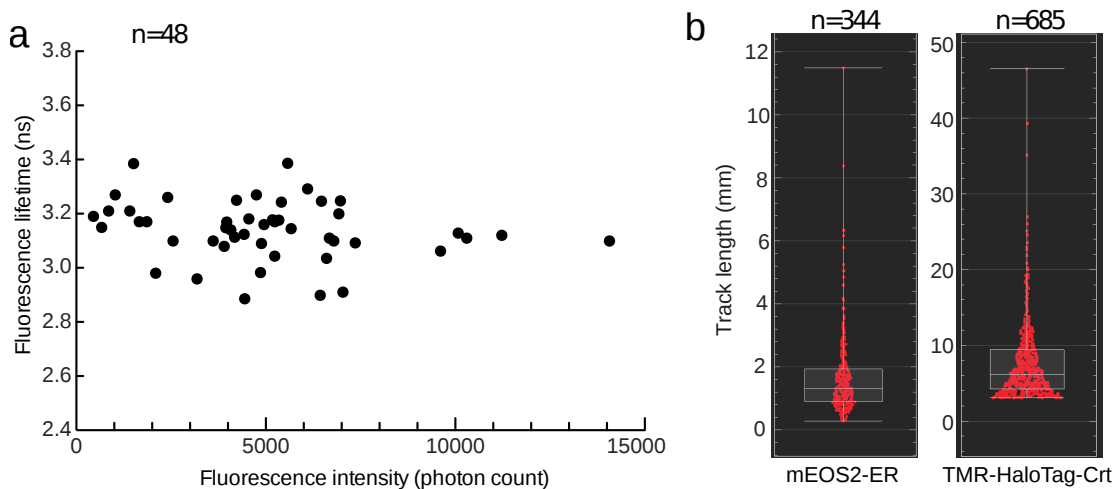


Figure 4.6: Exogenous expression levels of a fluorescent protein do not affect macromolecular crowding in the ER lumen; TMR photostability affords longer SPT trajectories. **A.** Plot of fluorescence intensity versus fluorescence lifetime of individual cells expressing an ER-localised FRET-based macromolecular crowding probe as in Fig. 4.1b. Fluorescence intensity serves as a measure of exogenous protein expression levels. Fluorescence lifetime correlates inversely with macromolecular crowding as described in Fig. 4.1b. Note, no correlation (Pearson correlation coefficient  $\rho = -0.2109$ ) was observed between the broad range of exogenous protein expression levels and macromolecular crowding,  $n=48$  independently sampled cells. **B.** Distribution of trajectory lengths (central bar indicate the median, bottom and top edges of the box indicate the 25th and 75th percentiles respectively and whiskers extend to the most extreme data points), generated as in Figure 4.2, the correlation between fluorophore photostability and trajectory length attest to the fidelity of the tracking procedure.

## 4.4 Supplementary information

When not stated otherwise, the following analysis was performed using MATLAB version 9.0 (MathWorks).

### 4.4.1 Mean Squared Displacement and first moment analysis

The Mean Squared Displacement (MSD) approach has been used to detect deviations from classical free diffusive motion of SPTs. Indeed, the MSD at time lag  $\tau$  is defined by

$$MSD(\tau) = \langle |\mathbf{X}(t + \tau) - \mathbf{X}(t)|^2 \rangle, \quad (4.2)$$

where the average is computed either over realizations or time. This quantity is expected to grow like  $MSD(\tau) \approx A\tau^\alpha$  for  $\tau$  small where  $A > 0$  and  $\alpha$  is called the anomalous exponent. An exponent  $\alpha > 1$  is a signature of a super-diffusive regime while  $\alpha < 1$  indicates a sub-diffusive regime [190]. This analysis however does not provide any explanation for the mechanisms causing a specific regime because it is usually computed by averaging over time



(time-averaged MSD) or space (trajectory-averaged MSD) thus disregarding the possible heterogeneous geometrical organization of STPs.

The motion exhibited by the recorded trajectories analyzed in Fig. 4.2 is neither time nor space homogeneous. Indeed, the motions in nodes and tubules are different and trajectories exhibit high-velocity peaks of random duration (Fig. 4.5d left), separated by random time intervals (Fig. 4.5d right), making the MSD analysis hard to perform on these data. For that reason, we have chosen to conduct here an analysis based on the Langevin equation of motion (equation (1.13)) through which local parameters of motion  $\mathbf{b}$  and  $D$  are extracted from trajectories.

It is also possible to draw the similar conclusions about the observed dynamics by conducting an MSD analysis on trajectories in the nodes and tubules. For this analysis we relied on the time averaged MSD (taMSD) at time lag  $\tau = 1 \dots (M_i - 1)$  defined, for a two-dimensional trajectory  $\mathbf{X}_i$  composed of  $M_i$  points

$$taMSD_i(\tau) = \frac{1}{M_i - \tau} \sum_{k=1}^{M_i - \tau} \|\mathbf{X}_i(t_{k+\tau}) - \mathbf{X}_i(t_k)\|^2 \quad (4.3)$$

To recover the anomalous exponent from a taMSD curve, we fitted to the first 8 points of this curve in the log-log space the function:

$$\log(A) + \alpha \log(\tau) \quad (4.4)$$

using the `fit` function from MATLAB version 9.0 (MathWorks). Fig. 4.7b presents the taMSD curves in log-log space (left) and the corresponding anomalous exponents  $\alpha$  (right) obtained by computing the taMSD on each trajectory, longer than 20 points (360ms) and visiting at least two reconstructed network nodes, from the experiment presented in Fig. 4.2 and found a broad distribution of anomalous exponents ( $0 \leq \alpha \leq 1.5$ ) in agreement the proposed two-states dynamics. To investigate the type of dynamics inside the nodes, we determined for each node the ensemble of  $N$  sub-trajectories  $\mathbf{X}_1 \dots \mathbf{X}_N$  located inside the node and obtained the node MSD by averaging:

$$taMSD(\tau) = \frac{1}{N} \sum_{i=1}^N taMSD_i(\tau) \quad (4.5)$$

In Fig. 4.7c we report for each node of the experiment presented in Fig. 4.2 its taMSD curve in log-log space (left) and the corresponding distribution of anomalous exponents (right) observing exponents  $\alpha < 0.8$ , indicating a sub-diffusive behavior. Finally, to investigate the type of dynamics outside nodes, we synchronized for each nodes, the trajectories on their exit from the node (as presented in sub-section 4.4.7 and Fig. 4.7e) and computed an average MSD using equation (4.5). In Fig. 4.7d we present three examples of nodes for which this synchronization result in an averaged taMSD characterized by an anomalous exponent  $\alpha > 1$  indicating in these cases a super-diffusive behavior.

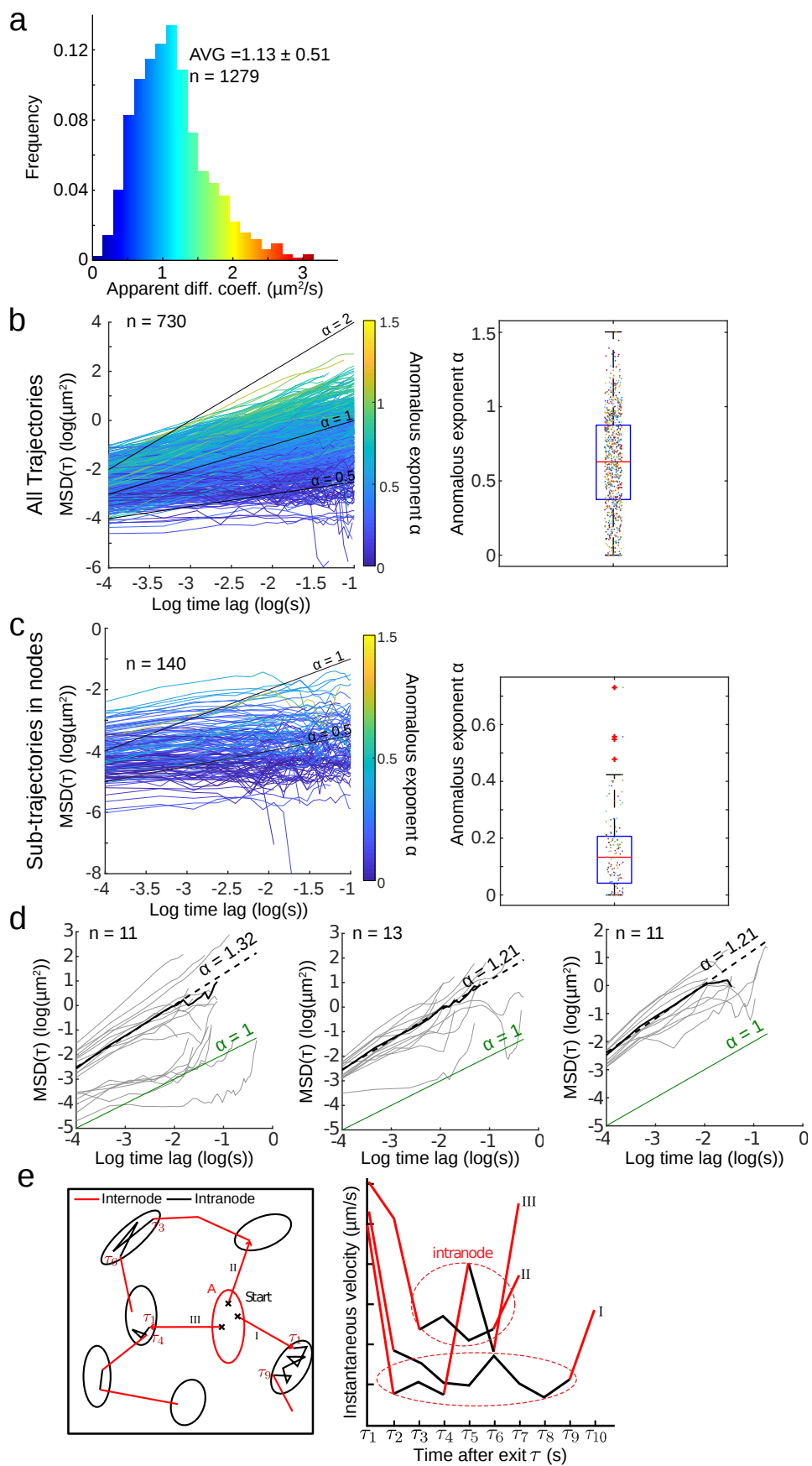


Figure 4.7: (*previous page*) Analysis of SPTs Diffusive dynamics and schematic illustration of internode and intranode motion. **A.** Distributions of diffusion coefficients computed on the individual square bins presented in Fig. 4.2f for the entire domain, including internode high-velocity displacements and intranode displacements. Numbers correspond to  $\text{AVG} \pm \text{SD}$ . **B.** Mean Squared Displacement analysis applied on the SPTs from the experiment presented in Fig. 4.2. Left panel: individual MSD curves color-coded by their anomalous exponent  $\alpha$  and right panel: distribution of the corresponding anomalous exponents (central bar indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles respectively and whiskers extend to the most extreme data points). **C.** Average MSD computed on trajectory fragments contained in nodes. Left panel: average MSD curve for each node, color-coded according to its anomalous exponent and right panel: distribution of the corresponding anomalous exponents (central bar indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively and whiskers extend to the most extreme data points). **D.** Averaged MSD curves computed from trajectories exiting three different nodes with anomalous exponents  $\alpha > 1$  indicative of a super-diffusive regime (sub-section 4.4.1). **E.** Synchronization procedure: Left: Trajectories are synchronized on their last point spent in the indicated node (red arrow). Trajectories can then either jump from node to node or move inside a node. Right: For each time  $\tau$  after escaping the node, internode (red segments) and intranode (black segments) displacements (as identified from left panel) are grouped into separate ensembles and an average instantaneous velocity computed from equations (4.31) and (4.32); analysis performed on  $n$  trajectories as denoted in individual panels.

## 4.4.2 Models for the instantaneous velocity distribution

### Pure Diffusion model

To analyze the velocity histograms presented in Fig. 4.2c, 4.3a and 4.8, we use a two-dimensional random walk model:

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \sqrt{2D\Delta t}\boldsymbol{\eta}, \quad (4.6)$$

where  $\boldsymbol{\eta} = [\eta_1, \eta_2]$  with  $\eta_1, \eta_2 \sim \mathcal{N}(0, 1)$  is a white noise. The distribution of the displacement lengths is given for  $l > 0$  by [69]:

$$Pr_{\text{diff}} \left\{ \frac{\|\Delta X(t)\|}{\Delta t} = l \right\} = \frac{l}{\sigma_d^2} \exp\left(\frac{-l^2}{2\sigma_d^2}\right), \quad (4.7)$$

where  $\|\cdot\|$  is the Euclidean norm and

$$\sigma_d = \sqrt{\frac{2D}{\Delta t}}. \quad (4.8)$$

We use equation (4.6) with the parameters presented in Table 4.2, using the apparent diffusion coefficient  $D = D_{app}$  defined in paragraph 4.4.4 to generate the distributions shown in Fig. 4.2c, Fig. 4.3a and 4.8 (solid lines).

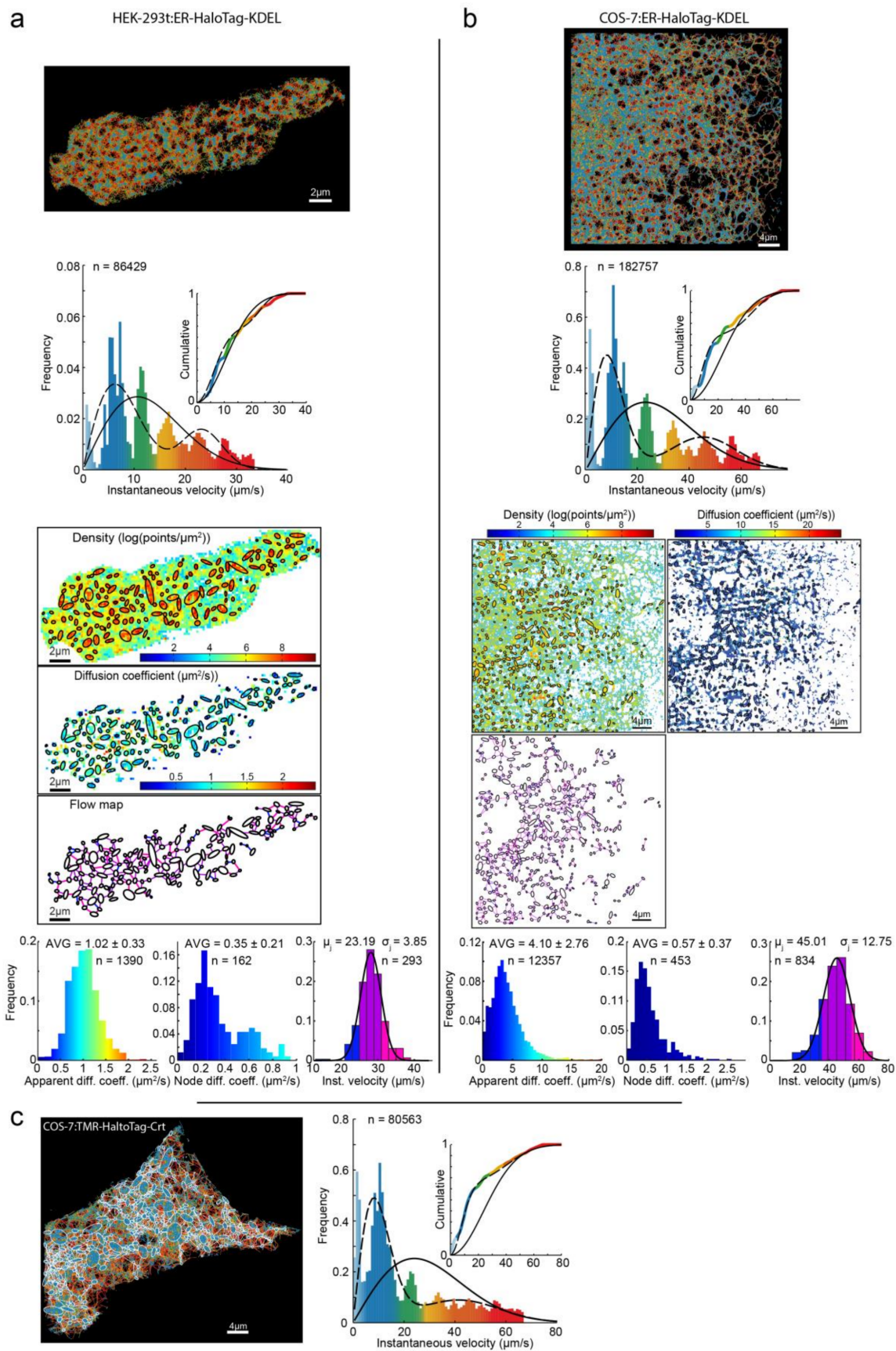


Figure 4.8: (*previous page*) Characteristics of ER luminal SPT displacements for several cell types and markers. Statistics of SPTs motion obtained as in Fig. 4.2, but for a HEK-293t cell **A.** and a COS-7 cell **B.** expressing an inert tracer-ER-targeted HaloTag with a C-terminal ER-retention signal, KDEL. Each column is organized similarly to Fig. 4.2. Top: SPTs are presented color-coded according to the instantaneous velocity distributions presented below. The overlaid curves on the instantaneous velocity distributions correspond to the distributions expected for an exclusively diffusion-driven motion (solid lines, using the apparent  $D$ ), or a combination of diffusion and flow (dashed lines, using  $D$  in nodes and the average instantaneous velocity between node). Inset are the corresponding cumulative distributions (color-coded as in the histogram). Middle: density and diffusion maps computed for a grid of square bins (sides of  $0.2 \mu\text{m}$ ) imposed on the particle displacements and flow map computed by averaging the non-Brownian velocity jumps of particles moving between pairs of neighbour nodes and color-coded according to their velocity. Bottom: distributions of apparent diffusion coefficients (left), node diffusion coefficients (middle) and average instantaneous velocity between neighbouring nodes (right). **C.** SPTs recorded for a COS-7 cell expressing an ER-localised TMR labeled Halo-tagged Calreticulin (left) color-coded according to the instantaneous velocity distribution (right). Overlaid curves on the histogram are as for (a-b). Shown are characteristic images observed in three independent repeats. All values are given as  $\text{AVG} \pm \text{SD}$ , for  $n$  trajectories as denoted in individual panels.

### Flow-diffusion switching model

To account for the fast (faster than the acquisition time  $\Delta t = 18\text{ms}$ ) internode dynamics observed in Fig. 4.2c, we use a jump-diffusion model [76] defined by the following rule

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \begin{cases} \mathbf{J}\Delta t & \text{w. p. } \lambda\Delta t \\ \sqrt{2D\Delta t}\boldsymbol{\eta} & \text{w. p. } 1 - \lambda\Delta t \end{cases}, \quad (4.9)$$

where the statistics of the jumps  $\mathbf{J}$  is approximated from the observed flow velocity distribution, presented in the inset histogram of Fig. 4.2g and in Fig. 4.8 as  $\|\mathbf{J}\| \sim \mathcal{N}(\mu_{\text{jump}}, \sigma_{\text{jump}})$  (these distributions are always positive in practice). Although the jump angle should follow the ER architecture, for the present model, we draw the angles from a uniform distribution in  $[0, 2\pi]$ . This simplification holds true as long as we are considering only the norm of the jumps. To estimate the distribution of displacement lengths of process (4.9), we use Bayes' law and condition the displacement on each state of the process:

$$\begin{aligned} Pr_{\text{switch}} \left\{ \frac{\|\Delta\mathbf{X}(t)\|}{\Delta t} = l \right\} &= Pr \left\{ \frac{\|\Delta\mathbf{X}(t)\|}{\Delta t} = l | \text{Jump} \right\} Pr\{\text{Jump}, t\} \\ &+ Pr \left\{ \frac{\|\Delta\mathbf{X}(t)\|}{\Delta t} = l | \text{Diff} \right\} Pr\{\text{Diff}, t\}. \end{aligned} \quad (4.10)$$

By definition the steady-state probability of being in a jump or diffusion state are:

$$\begin{aligned} Pr\{\text{Jump}\} &= \lim_{t \rightarrow \infty} Pr\{\text{Jump}, t\} = \frac{\lambda\Delta t}{1 - \lambda\Delta t + \lambda\Delta t} = \kappa. \\ Pr\{\text{Diff}\} &= 1 - Pr\{\text{Jump}\} = 1 - \kappa, \end{aligned} \quad (4.11)$$

where  $\kappa = \lambda\Delta t \in [0, 1]$ . Approximating equation (4.10) by (4.11) and replacing the displacement length distributions by a normal distribution for the jumps and a Rayleigh distribution (equation (4.7)) for diffusive displacements, we obtain

$$\begin{aligned} Pr_{\text{switch}} \left\{ \frac{\|\Delta \mathbf{X}(t)\|}{\Delta t} = l \right\} &= \kappa Pr \left\{ \frac{\|\Delta \mathbf{X}(t)\|}{\Delta t} = l | \text{Jump} \right\} + (1 - \kappa) Pr \left\{ \frac{\|\Delta \mathbf{X}(t)\|}{\Delta t} = l | \text{Diff} \right\} \\ &= \kappa \frac{1}{\sigma_j \sqrt{2\pi}} \exp \left( -\frac{(l - \mu_{\text{jump}})^2}{2\sigma_j^2} \right) + (1 - \kappa) \frac{l}{\sigma_d^2} \exp \left( -\frac{l^2}{2\sigma_d^2} \right), \end{aligned} \quad (4.12)$$

where  $\sigma_d$  is given by equation (4.8). To estimate the switching probability  $\kappa$  in equation (4.12), we use a Maximum-Likelihood Estimation (MLE) approach based on the probability

$$p(l|\kappa) = Pr_{\text{switch}} \left\{ \frac{\|\Delta X(t)\|}{\Delta t} = l \right\}, \quad (4.13)$$

of observing  $l$  given  $\kappa$ . The MLE  $\hat{\kappa}$  for  $N$  observed displacements  $l_1, \dots, l_N$  is [191]

$$\hat{\kappa} = \arg \max_{\kappa \in [0,1]} \sum_{i=1}^N \ln(p(l_i|\kappa)). \quad (4.14)$$

We compute  $\hat{\kappa}$  using the `mle` function of MATLAB version 9.0 (MathWorks) applied to the trajectory displacements extracted from the SPTs described in methods.

To generate the distributions shown in Fig. 4.2c and 4.8 (dashed lines), we use equation (4.12), with the parameters presented in Table 4.2, using the diffusion coefficient in the nodes  $D = D_{\text{node}}$  defined in section 4.4.4.

### 4.4.3 Reconstruction of the ER Network from SPTs

Detecting nodes (tubule junctions) boundary and inter-junction stretches (tubules) is based on the heterogeneity of the time-integrated particle spatial density presented in Fig. 4.2d and 4.8. The algorithmic procedure uses the large amount of recorded SPTs described in methods and proceeds as follows. We first construct clusters of points (nodes) defined as regions of aggregation of short displacements (aggregation regions appear to co-localize with short displacements as presented in Fig. 4.2c and 4.8:

1. Define a threshold  $V_L$  (in  $\mu\text{m}/\text{s}$ ), and discard from the analysis any point  $X_i(t)$  such that  $\frac{\|\Delta X_i(t)\|}{\Delta t} \geq v_L$  ( $\Delta X_i(t)$  is the displacement as defined in methods).
2. Apply the `dbscan` [192] clustering algorithm to cluster the remaining points (implementation from `scikit-learn` [193] through Python3 provided by the Anaconda Distribution version 4.3.8 (Anaconda Inc.)).
3. Approximate the boundary of each cluster as an ellipse with semi-axes  $a > b$  using a principal component analysis. Remove ellipses with an area  $\pi ab > 4\mu\text{m}^2$  or an eccentricity  $\frac{a}{b} > 4$ . Merge overlapping ellipses by fitting a new ellipse to the union of their points.

4. Assign each points discarded in step 1 to the cluster corresponding to the ellipse in which they fall, if any.

The `dbscan` algorithm used in step 2 allows to generate clusters based on the local point density and requires two parameters:

1. The maximum distance  $R$  (in  $\mu\text{m}$ ) below which two points are considered to be neighbors.
2. The minimum number of points  $N$  at a distance  $\leq R$  of a point to start a cluster.

These two parameters define a minimal density  $\frac{N}{R}$  of points/ $\mu\text{m}^2$  inside each cluster. The values of  $R$  and  $N$  depend on the morphology of the imaged ER and the local number of recorded trajectories. For each dataset, these values were determined empirically such that the computed clusters overlap with the ER structure formed by the trajectories.

Once nodes are found, we defined tubules by constructing a connectivity matrix  $C$  of size  $K \times K$  ( $K$  number of detected nodes) where  $c_{i,j}$  ( $1 \leq i, j \leq K$ ) contains the number of trajectory displacements starting in node  $i$  and arriving in node  $j$ . Specifically, we increment the coefficient  $c_{i,j}$  by one for each data point  $X_k(t_l)$  ( $1 \leq k \leq N_t$ ,  $0 \leq l < M_k - 1$ ) in either of 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  $0 \leq l < M_k - 2$ ).

Finally, we removed from the graph any disconnected node. The different parameters used for reconstructing the graphs presented in the main text and supplementary data are given in Table 4.3. Graphs from Table 4.1 (Main Text) were constructed using the parameter ranges:  $V_L = [9, 19]\mu\text{m/s}$  (following the histogram of instantaneous velocities),  $R = [0.08, 0.26]\mu\text{m}$ ,  $N = [10, 45]$  points and  $\pi ab \leq [4, 8]\mu\text{m}^2$ . Ellipses representing the nodes of the graphs are shown in Fig. 4.2dfg and 4.8.

#### 4.4.4 Recovery of the local dynamics in the ER lumen from SPTs

##### Estimation of the apparent diffusion coefficient

We define the apparent diffusion coefficient  $D_{app}$  as the diffusion coefficient estimated using the entire distribution of displacements. Note that this distribution also contains large values, that cannot be attributed to diffusion. We estimated  $D_{app}$  by averaging the values obtained from equation (1.20) on each bin over the entire map (only for bins containing at least 20 points). The distribution of diffusion coefficients from individual bins of the diffusion map from Fig. 4.2f is presented in Fig. 4.7a, in Fig. 4.3d for the map of Fig. 4.3c and in Fig. 4.8a,b for the other datasets.

## Estimation of the diffusion coefficients in the nodes

The diffusion coefficient  $D_{node}$  inside each node is computed from equation (1.20), but in addition we constrained both ends of the displacement be located inside the node. For a node delimited by an ellipse  $E$ , we get:

$$D_{node}(\mathbf{X}(t)) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[\Delta X(t)^T \Delta X(t) | X(t) = \mathbf{X} \in E \text{ and } X(t + \Delta t) \in E]}{2\Delta t}. \quad (4.15)$$

We use equation (4.15) to estimate the diffusion coefficient  $D_{node}$  in each node and report their distributions in Fig. 4.2f (inset) for the main text dataset and Fig. 4.8a,b.

## Estimation of the internodes displacements

We define the instantaneous velocity (in  $\mu m/s$ ) between two successive points of the same trajectory  $\mathbf{X}(t_1)$  and  $\mathbf{X}(t_2)$  ( $t_2 > t_1$ ) as the ratio of the distance to elapsed time

$$v(\mathbf{X}(t_1), \mathbf{X}(t_2)) = \frac{\|\mathbf{X}(t_2) - \mathbf{X}(t_1)\|}{t_2 - t_1}, \quad (4.16)$$

where  $\|\cdot\|$  is the Euclidean distance. We define the flow velocity  $v_{flow}$  between two nodes  $A$  and  $B$ , as the average of the instantaneous velocities for the displacements connecting the two nodes:

$$v_{flow}(A, B) = \mathbb{E} \left[ v(\mathbf{X}(t_1), \mathbf{X}(t_2)) \left| \begin{array}{l} (\mathbf{X}(t_1) \in A, \mathbf{X}(t_2) \in B) \text{ or} \\ (\mathbf{X}(t_1) \in B, \mathbf{X}(t_2) \in A) \end{array} \right. \right], \quad (4.17)$$

where by construction  $t_2 - t_1 \in \{\Delta t, 2\Delta t\}$  (see tubule reconstruction from sub-section 4.4.3). We discretize equation (4.17) and obtain the estimator:

$$v_{flow}(A, B) \approx \frac{1}{N} \sum_{i=1}^N v(\mathbf{X}_i(t_1), \mathbf{X}_i(t_2)), \quad (4.18)$$

where  $N$  is the number of trajectory displacements connecting the nodes  $A, B$  and  $\mathbf{X}_i$  is the  $i^{th}$  such displacements. The velocity  $v_{flow}$  characterizes the jump flow between nodes. As shown in Fig. 4.2c, this internode flow is associated with the thick tail of the velocity distribution. The distributions of jump velocities between each pair of neighbor nodes for the different datasets are presented in Fig. 4.2g and Fig. 4.8a,b.

### 4.4.5 Static ER network analysis

#### Flow directionality and strongly connected components

To determine whether the ensemble of observed displacements between two neighbor nodes  $i$  and  $j$  form a uni- or bi-directional flow, we define a uni-directionality score  $r$  as the ratio of



the number of observed displacements between the two nodes with direction  $i \rightarrow j$  divided by the total number of displacements in both directions:

$$r_{i,j} = \frac{c_{i,j}}{c_{i,j} + c_{j,i}}, \quad (4.19)$$

where  $C$  is the connectivity matrix of the graph as defined in sub-section 4.4.3. Using  $r_{i,j}$ , we build a binary version  $C'$  of  $C$  defined as

$$c'_{i,j} = \begin{cases} 1 & \text{when } r_{i,j} \geq 0.25 \\ 0 & \text{otherwise} \end{cases}. \quad (4.20)$$

$C'$  has the property that  $c'_{i,j} = c'_{j,i} = 1$  when  $0.25 \leq r_{i,j} \leq 0.75$  (bi-directional flow) and  $c'_{i,j} = 1, c'_{j,i} = 0$  when  $r_{i,j} > 0.75$  (uni-directional flow). The ratio  $r_{i,j}$  (and thus  $c'_{i,j}$ ) is computed only for pairs of nodes connected by at least three displacements ( $c_{i,j} + c_{j,i} \geq 3$ ). To find the connected components of the reconstructed network, we use the classical Tarjan Strongly Connected Components (SCCs) detection algorithm [184] on the binary connectivity matrix  $C'$  (`graphconncomp` function from MATLAB version 9.0 (MathWorks)). SCCs are a partition of the ensemble of nodes such that there exists a path (taking into account the directionality) from each node of a subset to any other node of the same subset. The presence in Fig. 4.4b of a SCC encompassing almost the entire graph shows that the observed flows have the potential to move particles through almost the entire network. In addition Fig. 4.4b also displays the flows directionality on the links as: arrows for uni-directional, solid (no arrow) for bi-directional and dashed for undecided.

### Count of Afferent and Efferent branches

We further characterize the structure of the reconstructed ER graph by computing for each node  $k$  the number of afferent and efferent branches connected to it. An afferent (resp. efferent) branch is a link  $l \rightarrow k$  (i.e.  $c_{l,k} > 0$ ) (resp.  $k \rightarrow l, c_{k,l} > 0$ ) where  $l$  is any other node of the graph. Based on the count of afferent and efferent nodes, we define the out and in-degree of node  $k$  using the connectivity matrix  $C$  as:

$$\text{outdeg}(k) = \sum_{i=1}^N \mathbb{1}_{c_{k,i} > 0} \quad \text{and} \quad \text{indeg}(k) = \sum_{i=1}^N \mathbb{1}_{c_{i,k} > 0}, \quad (4.21)$$

where  $\mathbb{1}_{a>b} = \begin{cases} 1 & \text{if } a > b \\ 0 & \text{otherwise} \end{cases}$ , and  $N$  is the number of nodes in the graph. In- and out-

degrees are computed only for nodes  $k$  such that  $\sum_{i=1}^N c_{i,k} > 2$  and  $\sum_{i=1}^N c_{k,i} > 2$  respectively. The distribution of in- and out-degrees are presented in Fig. 4.4c.

### Fraction of entering and exiting displacements in nodes

To study the passing dynamics of trajectories through nodes we define a retention score for the nodes defined as the ratio of the number of exiting displacements to the total number of

entering and exiting displacements for a node  $k$ :

$$\phi_k = \frac{\sum_{i=1}^N c_{k,i}}{\sum_{i=1}^N c_{i,k} + \sum_{i=1}^N c_{k,i}}, \quad (4.22)$$

where  $C$  is the connectivity matrix of the graph and  $N$  the number of nodes. We have  $\varphi \in [0, 1]$  such that  $\varphi_k \rightarrow 0$  indicates that the node retains trajectories,  $\varphi_k \rightarrow 1$  indicates that trajectories originate from this node and  $\varphi_k \approx 0.5$  indicates that trajectories pass through the node.  $\varphi_k$  was computed only for nodes  $k$  such that  $\sum_{i=1}^N c_{k,i} + c_{i,k} > 2$ . The distribution of  $\phi$  for the reconstructed ER network is shown in Fig. 4.4d.

#### 4.4.6 Transient ER network analysis

At steady-state, we identified the uni- and bi-directional flows inside tubules (Fig. 4.4b), we now investigate how the direction of these flows evolve with time. To this end, we determine the distribution of durations of uni-directionality periods of the flow between two nodes as presented in Fig. 4.5a. The analysis is performed on each pair or neighbouring (directly linked by trajectories) nodes, linked by at least 20 displacements registering one of the two possible directionalities (either node  $A \rightarrow B$  or  $B \rightarrow A$ ). The following algorithm groups the successive displacement events as a function of their directionality and determine the duration of these groups:

1. Collect the next displacement event observed at time  $t_{\text{first}}$ , identify its direction say  $A \rightarrow B$  and form a new group containing this event.
2. Accumulate in the group created in step.1 the following jump events with direction  $A \rightarrow B$  and stop when there are no more event or after encountering two events with direction  $B \rightarrow A$  (in this case the event or the two successive  $B \rightarrow A$  events when it occurs, are not collected). The last event considered in the group occurs at time  $t_{\text{end}}$  and always has the same direction as the group (see Fig. 4.5a).
3. Compute the duration of the group as  $\tau = t_{\text{end}} - t_{\text{first}}$ .
4. Return to step 1.

We discarded groups formed by less than 3 displacements with the same direction. The distribution of the uni-directional flow durations  $\tau$  is presented in Fig. 4.5b.

To interpret the mean of this distribution, we recall that this statistics is partly contaminated by the fraction of activated particles located in nodes and the track lengths. Indeed, only photo-activated particles appear in the statistics. To recover the uni-directional duration of the fluxes from the empirical distribution, we use a model taking into account these two characteristics. The probability that the flux between nodes  $A$  and  $B$  lasts a duration  $\tau_{Uni} = t$  is computed by conditioning on having  $k$  events ( $k$  particles flowing in the same direction) during that time

$$Pr\{\tau_{Uni} = t\} = \sum_{k=1}^{\infty} Pr\{\tau_1 + \dots + \tau_k = t|k\} Pr_{Uni}\{k\}, \quad (4.23)$$

where the probability that  $k$  event occurs in the same direction is by symmetry,

$$Pr_{U_{ni}}\{k\} = \frac{1}{2^k} \quad (4.24)$$

and  $\tau_k$  is the arrival time of the  $k^{th}$  event after the first one. We consider that the distributions of arrival times are Poissonian with same rate  $\lambda$ , independent of the initial node  $A$  or  $B$ . We conclude that

$$Pr\{\tau_1 + \dots + \tau_k = t|k\} = \lambda \frac{(\lambda t)^{k-1}}{(k-1)!} \exp(-\lambda t). \quad (4.25)$$

Computing the sum from equation (4.23), we obtain for at least 2 events

$$Pr\{\tau_{U_{ni}} = t\} = \frac{\lambda}{2} \left( \exp\left(-\frac{\lambda t}{2}\right) - \exp(-\lambda t) \right). \quad (4.26)$$

Equation (4.26) is the probability density function when all particles are labeled and its average is  $\frac{3}{2\lambda}$ . When a particle is activated with probability  $p$ , which represents the steady-state fraction of labeled particles, the statistics of unidirectional flow is still given by equation (4.26) but with a rate  $\tilde{\lambda}$ . Using Bayes' law, the probability of observing a displacement event is

$$\begin{aligned} Pr\{\tau_{U_{ni}} = t\} &= Pr\{\tau_{U_{ni}} = t, \text{activated}\} Pr\{\text{activated}\} \\ &+ Pr\{\tau_{U_{ni}} = t, \text{notactivated}\} Pr\{\text{notactivated}\}, \end{aligned} \quad (4.27)$$

where  $Pr\{\text{activated}\} = p$  is the fraction of activated molecules. The second term is zero because we cannot see displacement events from non-activated molecules. Thus the mean unidirectional flow duration becomes

$$\langle \tau \rangle = p\bar{\tau}, \quad (4.28)$$

where the rates are related by  $\lambda = p\tilde{\lambda}$ . Considering a fraction  $p = 1\%$  of activated molecules, we predict that the mean unidirectional time should be around  $\langle \tau \rangle = 38ms$  instead of the observed  $\bar{\tau} = 3.89s$ .

#### 4.4.7 Instantaneous velocities along individual trajectories

To obtain statistics of the velocity fluctuations of trajectories (recorded at different times over a period of seconds) and possibly located either in tubules or nodes, we introduced a 'synchronization' procedure: the velocity fluctuations of individual trajectories were monitored starting from the last time point where the particle was in a given node prior to its exit. The velocity fluctuations following the node exit was plotted as a function of time elapsed since the synchronization event (a universal time scale for all trajectories regardless when they were detected). In details, the synchronisation event is defined for a trajectory  $X_i$ , its last recorded point inside some node  $A$ ,

$$t_i^* = \max_{0 \leq j < M_i - 1} \{ \mathbf{X}_i(t_j) \in A \text{ and } \mathbf{X}_i(t_{j+1}) \notin A \}. \quad (4.29)$$

Considering the  $n_A$  trajectories going through node  $A$ , the ensemble of last points inside  $A$  is

$$\mathbf{X}_A = \{\mathbf{X}_i(t_i^*) | i = 1..n_A\}. \quad (4.30)$$

We now divide displacements along individual trajectories into two subsets based on the starting node  $A$  and appearing at a time  $\tau > 0$  after exit:  $\mathbf{X}_{A,node}(\tau)$  containing displacements inside nodes and  $\mathbf{X}_{A,tubule}(\tau)$  containing displacements connecting two nodes:

$$\begin{aligned} \mathbf{X}_{A,node}(\tau) &= \{\mathbf{X}_i(t_i^* + \tau) | \mathbf{X}_i(t_i^*) \in \mathbf{X}_A \text{ and } \mathbf{X}_i(t_i^* + \tau - \Delta t) \in N \text{ and } \mathbf{X}_i(t_i^* + \tau) \in N\} \\ \mathbf{X}_{A,tubule}(\tau) &= \{\mathbf{X}_i(t_i^* + \tau) | \mathbf{X}_i(t_i^*) \in \mathbf{X}_A \text{ and } \mathbf{X}_i(t_i^* + \tau - \Delta t) \in N_1 \text{ and } \mathbf{X}_i(t_i^* + \tau) \in N_2\}, \end{aligned}$$

where  $N, N_1, N_2 \in \mathcal{N}$  the ensemble of nodes and  $N_1 \neq N_2$  ( $N_1$  or  $N_2$  can be not a node). We computed the average velocities  $v_{Inter}(\tau)$  (resp.  $v_{Intra}(\tau)$ ) at time  $\tau \geq \Delta t$  for each subset:

$$v_{Intra}(\tau) = \frac{1}{n_{A,intra}(\tau)} \sum_{\mathbf{X}_i(t_i^* + \tau) \in \mathbf{X}_{A,node}(\tau)} v(\mathbf{X}_i(t_i^* + \tau - \Delta t), \mathbf{X}_i(t_i^* + \tau)) \quad (4.31)$$

and

$$v_{Inter}(\tau) = \frac{1}{n_{A,inter}(\tau)} \sum_{\mathbf{X}_i(t_i^* + \tau) \in \mathbf{X}_{A,tubule}(\tau)} v(\mathbf{X}_i(t_i^* + \tau - \Delta t), \mathbf{X}_i(t_i^* + \tau)), \quad (4.32)$$

where  $n_{A,intra}(\tau)$  (resp.  $n_{A,inter}(\tau)$ ) is the number of displacements from synchronized trajectories that fall inside (resp. between two nodes) at time  $\tau$  and  $v(.,.)$  is the instantaneous velocity (equation (4.16)). The synchronization process and the intra and inter node velocities along synchronized trajectories are presented schematically in Fig. 4.7e. In this figure, schematic trajectories synchronized on their exit of a node  $A$  are presented on the left panel and their associated instantaneous velocity as a function of the time since exit  $\tau$  from  $A$  are presented on the right panel. Fig. 4.5c presents the average (line) and standard deviation (shade) of the instantaneous velocity for  $v_{intra}$  (black) and  $v_{inter}$  (red) for a selected node  $A$  of the main-text network for  $\tau \leq 0.2s$ .

#### 4.4.8 Instantaneous velocity peaks duration and inter-peaks period

To further analyze the fluctuations of the instantaneous velocities along individual trajectories we proceed as in sub-section 4.4.7 and synchronize the trajectories leaving a given node  $A$ . We then estimated the time spent by individual trajectories in a high velocity regime (above a threshold  $v_{\text{high}} = 19\mu m/s$ ) as the number of successive time steps  $\tau = j\Delta t$  ( $j = 0 \dots$ ) for which

$$v(\mathbf{X}_i(t_i^* + \tau + j\Delta t), \mathbf{X}_i(t_i^* + \tau + (j+1)\Delta t)) > v_{\text{high}}, \quad (4.33)$$

where  $\mathbf{X}_i$  is a synchronized trajectory and  $v(.,.)$  is the instantaneous velocity (equation (4.16)). The distribution of the high velocity regime durations is shown in Fig. 4.5d and the distribution of periods between two high velocity regimes in Fig. 4.5e, computed for trajectories synchronized for each node of the ER-network. Note that we considered only trajectories that visited at least three different network nodes.

#### 4.4.9 Dynamics of tubular junctions

##### Extraction of junctions from SIM images

We extracted the features of tubular junctions from a SIM images stack (50 images) through a procedure similar to [1]: first, the contrast of the entire image stack was manually modified to make the ER network more apparent. Then each image was skeletonized, junctions were extracted from the skeleton [194] (AnalyzeSkeleton plugin, ImageJ) and only junctions covering more than 3 pixels were kept. For each image of the stack, a new grayscale image was generated where only the pixels belonging to selected junctions appear white and afterward a Gaussian blur with  $\sigma = 1\text{px}$  was applied to produce a single-particle like image. On this stack of grayscale images, we applied a particle detection and tracking algorithm [55] to follow junctions through successive images (Spot Detection and Tracking plugins, Icy). This procedure produces an ensemble of  $N$  trajectories  $\mathbf{X}_1 \dots \mathbf{X}_N$  such that  $\mathbf{X}_i = \mathbf{X}_i(t_0) \dots \mathbf{X}_i(t_{M_i})$  and  $\Delta t = t_i - t_{i-1} = 90\text{ms}$ . From this ensemble, we keep only trajectories possessing at least 40 points. These trajectories are displayed in Fig. 4.9ab overlaid on top of the average stack image.

##### Mean Squared Displacement (MSD) analysis

We first characterize the type of diffusive motion exhibited by trajectories using a Mean Squared Displacement (MSD) analysis. For an ensemble of  $T$  trajectories the MSD at time  $t_k$  is defined as:

$$MSD(t_k) = \frac{1}{T} \sum_{i=1}^T \|\mathbf{X}_i(t_k) - \mathbf{X}_i(t_0)\|^2 \approx A(t_k - t_0)^\alpha \quad (4.34)$$

where  $\|\cdot\|$  is the Euclidean distance,  $A$  is a coefficient and  $\alpha > 0$  is the anomalous exponent characterizing the type of diffusive motion. To estimate  $\alpha$ , we fit the first 20 points of the MSD curve to the line  $y(t_k) = \log(A) + \alpha \log(t_k - t_0)$  in the log-log space using a MATLAB version 9.0 (MathWorks) script. Fig. 4.9c shows the distribution of  $\alpha$  obtained applying this procedure to 100 independent samples of  $T = 20$  randomly selected trajectories from the ensemble of trajectories possessing at least 40 points and keeping only the fits for which the coefficient of determination  $R^2 > 0.75$ . We found that  $\alpha = 0.60 \pm 0.24$  suggesting a sub-diffusive behavior or diffusion in the presence of confinement forces [195].

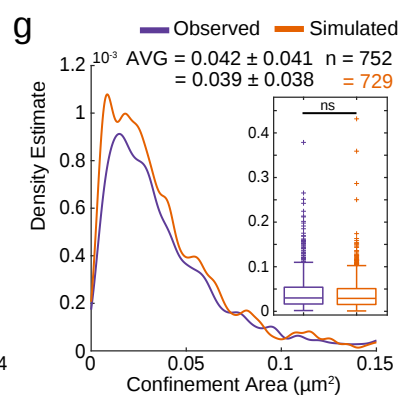
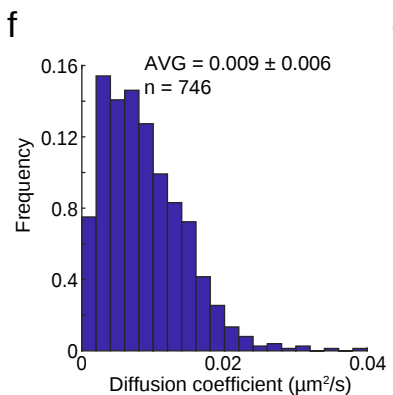
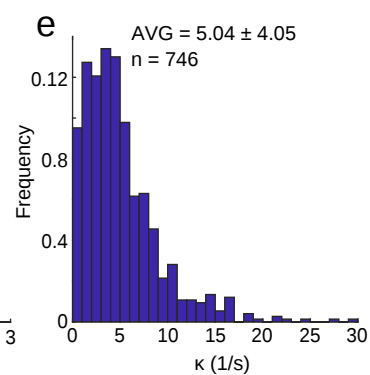
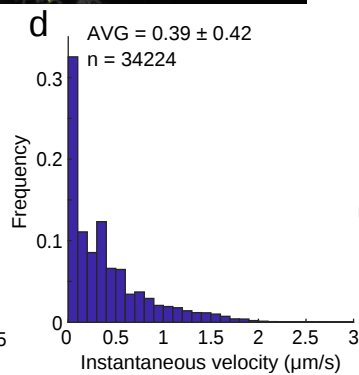
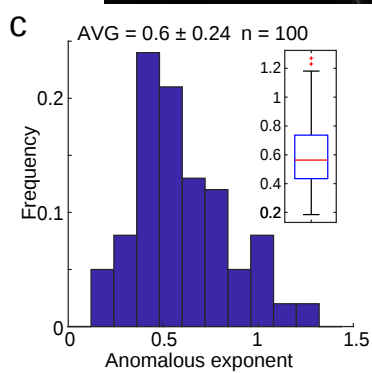
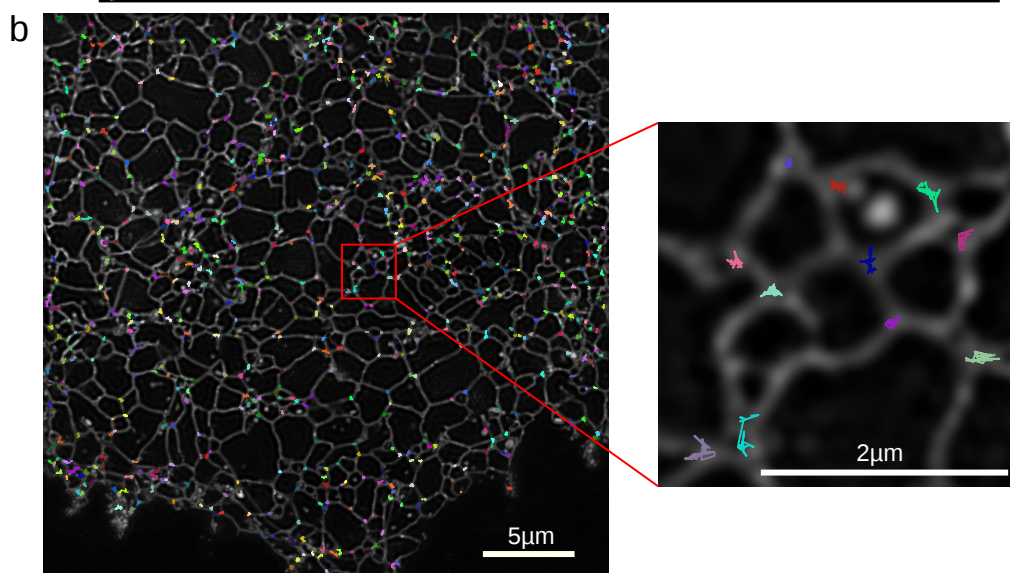
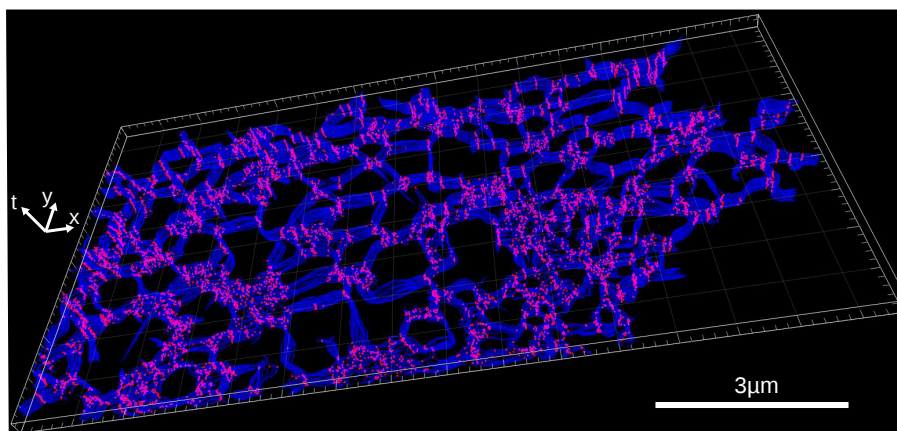


Figure 4.9: (*previous page*) Characteristics of ER structural components mobility. **A.** 3D kymograph presenting the motion of tubules and junctions across time (z-axis). **B.** SPTs extracted from a stack of 50 images. Inset: magnification of an ER region showing junction trajectories. **C.** Distribution of anomalous exponents obtained from SPTs (central bar indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively and whiskers extend to the most extreme data). **D.** Distribution of instantaneous velocities of the SPTs. (e-f) Estimation of the characteristics of the stochastic dynamics associated to the trajectories: distribution of spring constants **E.** and diffusion coefficients **F.** **G.** Observed and simulated junctions confinement areas. Inset: corresponding box plots (central bar indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively and whiskers extend to the most extreme data points). A Man-Whitney U-test was used to compare the two distributions, returning a non-significant p-value ( $p=0.1927$ ), see also video S5. (a,b) images are representative of three independent repeats. Analyses performed on  $n$  tubular junctions as denoted in individual panels.

### Confined motion of tubular junctions

The distribution of instantaneous velocities (Fig. 4.9d) suggests to model junction's dynamics as a diffusion process confined by active forces generated by a parabolic potential well. The corresponding equation of motion is expressed as an Ornstein-Uhlenbeck stochastic process

$$\dot{\mathbf{X}} = -\kappa(\mathbf{X} - \boldsymbol{\mu}) + \sqrt{2D}\dot{\mathbf{w}}, \quad (4.35)$$

where  $\boldsymbol{\mu}$  is the center of the potential well,  $\kappa$  (1/s) the spring coefficient,  $D$  ( $\mu\text{m}^2/\text{s}$ ) the diffusion coefficient and  $\mathbf{w}$  is a white noise. We estimate for each recorded trajectory  $\mathbf{X}_i$  the three parameters:  $\hat{\boldsymbol{\mu}}_i$ ,  $\hat{\kappa}_i$  and  $\hat{D}_i$ . The center of the well  $\boldsymbol{\mu}_i$  is approximated by the center of mass of the trajectory:

$$\hat{\boldsymbol{\mu}}_i = \frac{1}{M_i} \sum_{j=0}^{M_i-1} \mathbf{X}_i(t_j). \quad (4.36)$$

The parameters  $\hat{\kappa}_i$  and  $\hat{D}_i$  are estimated using the maximum-likelihood estimators [82]. For a trajectory  $\mathbf{X}_i(t_0) \dots \mathbf{X}_i(t_n)$  we compute for each dimension  $\mathbf{X}_i(t) = (x_i^1(t), x_i^2(t))$ ,  $d = 1, 2$

$$\hat{\beta}_1^d = \frac{n^{-1} \sum_{k=1}^n x_i^d(t_k) x_i^d(t_{k-1}) - n^{-2} \sum_{k=1}^n x_i^d(t_k) \sum_{k=1}^n x_i^d(t_{k-1})}{n^{-1} \sum_{i=1}^n x_i^d(t_{k-1})^2 - n^{-2} (\sum_{i=1}^n x_i^d(t_{k-1}))^2} + \frac{4}{n}, \quad (4.37)$$

$$\hat{\beta}_2^d = \frac{n^{-1} \sum_{k=1}^n (x_i^d(t_k) - \hat{\beta}_1^d x_i^d(t_{k-1}))}{1 - \hat{\beta}_1^d}, \quad (4.38)$$

$$\hat{\beta}_3^d = n^{-1} \sum_{k=1}^n (x_i^d(t_k) - \hat{\beta}_1^d x_i^d(t_{k-1}) - \hat{\beta}_2^d (1 - \hat{\beta}_1^d))^2, \quad (4.39)$$

from which we obtain the estimators:

$$\hat{\kappa}_i^d = \frac{\log(\hat{\beta}_1^d)}{\Delta t} \quad \text{and} \quad \hat{D}_i^d = \frac{\hat{\kappa}_i \hat{\beta}_3^d}{1 - (\hat{\beta}_1^d)^2} \quad (4.40)$$

where  $\Delta t = t_k - t_{k-1}$  is the time-step and the term  $\frac{4}{n}$  in  $\hat{\beta}_1^d$  is a correction for the low number of points. In practice, we obtain a symmetric tensor and force field computed as the averages:

$$\hat{\kappa}_i = \frac{\hat{\kappa}_i^1 + \hat{\kappa}_i^2}{2} \quad \text{and} \quad \hat{D}_i = \frac{\hat{D}_i^1 + \hat{D}_i^2}{2}. \quad (4.41)$$

We apply this estimation only on trajectories possessing at least 40 points. Fig. 4.9e shows the distribution of the estimated spring constants  $\hat{\kappa}$  and Fig. 4.9f the distribution of the estimated diffusion coefficients  $\hat{D}$ .

### Estimating the area of confinement of tubular junctions

In this section, we define and compute the area of confinement of each junction based on the statistics of the trajectories. For each trajectory  $\mathbf{X}_i$  we compute the 95% confidence ellipse  $e_i = (c_i, a_i, b_i, \varphi_i)$  of center  $c_i$ , largest (resp. smallest) semi-axis  $a_i$  (resp.  $b_i$ ) and angle (with x-axis)  $\varphi_i$ , of the spatial spreading of its points, considering this distribution as normal. We obtained the ellipse as follows [196]: first, we collected all points  $\mathbf{X}_i(t)$  of the trajectory into a  $2 \times n$  matrix  $O^i$ , then applied a Singular Value Decomposition algorithm to the covariance matrix of  $O^i$ :  $U^i \Sigma^i (V^i)^* = \text{cov}(O^i)$  and finally recovered the ellipse as:

$$c_i = \frac{1}{M_i} \sum_{j=0}^{M_i-1} \mathbf{X}_i(t_j), \quad a_i = \sqrt{5.991 \sigma_{1,1}^i}, \quad b_i = \sqrt{5.991 \sigma_{2,2}^i}, \quad \varphi_i = \arctan \left( \frac{u_{2,1}^i}{u_{1,1}^i} \right), \quad (4.42)$$

where  $\sigma_{1,1}^i$  and  $\sigma_{2,2}^i$  are the two eigenvalue of the matrix  $\text{cov}(O^i)$ . We define the confinement area  $A_i$  for a trajectory  $\mathbf{X}_i$  as the area of the estimated ellipse:  $A_i = \pi a_i b_i$ . The distribution of observed confinement areas is presented in Fig. 4.9g.

We now compare this distribution, with the expected distribution for a particle moving in a potential well. To this end for each trajectory  $\mathbf{X}_i$ , we simulated equation (4.35) with the estimated parameters  $\hat{\boldsymbol{\mu}}_i, \hat{\kappa}_i, \hat{D}_i$  using Euler's scheme to obtain a trajectory  $Y_i$ :

$$\mathbf{Y}_i(t_k) = \mathbf{Y}_i(t_{k-1}) + \hat{\kappa}_i (\mathbf{Y}_i(t_{k-1}) - \boldsymbol{\mu}_i) \delta t + \sqrt{2 \hat{D}_i} \delta t \boldsymbol{\eta}, \quad (4.43)$$

where  $\delta t = 0.0001$ s is the simulation time-step,  $\boldsymbol{\eta} = [\eta_1, \eta_2]$  with  $\eta_1, \eta_2 \sim \mathcal{N}(0, 1)$  is a white noise and  $\mathbf{Y}_i(t_0) = \boldsymbol{\mu}_i$ . To prevent the choice of the first point to influence the statistics, we run the simulation for 45000 time steps before recording the trajectory. We then sub-sampled  $Y_i$  by keeping one every 900 points to match the experimental acquisition time  $\Delta t = 0.09$ s and computed the confinement area of  $Y_i$  using the same procedure as for  $\mathbf{X}_i$ . This procedure was only applied to junctions for which  $\hat{\kappa}_i > 0$ . The distribution of confinement areas obtained from simulated trajectories is shown in Fig. 4.9g.



## 4.5 Characterization of ER tubule contractions

### 4.5.1 Extraction of tubule contraction statistics

Fast SIM images of ER in live cells, acquired and reconstructed as described in methods, were rendered using the Edges look-up-table of the Fiji software, with contrast settings to visualise one-pixel wide boundaries of the tubules. Contraction sites were identified as such if the tubule edges bended to merge more than once at the same position. In Fig. 4.5g we report the distributions of three observable characteristics of contraction events extracted from SIM images. The duration of a contraction event (Fig. 4.5g left) is computed as the difference between the last and first frames for which the event is detected; The time interval between successive contractions (Fig. 4.5g middle) is computed for each individual tubule and pair of successive contraction events as the difference between the first frame of the second contraction and the last frame of the first contraction; Finally the length of a contraction (Fig. 4.5g right) is computed by counting the number of pixels along the tubule axis involved in the contraction.

### 4.5.2 Elementary model of tubule contraction

To increase the evidences of a correlation between ER constrictions inside the tubules and the observed flow, we propose an elementary computation to link the contractions and flow. Considering an incompressible ER luminal fluid, the conservation of the mass is

$$\frac{\partial \rho}{\partial t} = \text{div}(\mathbf{v}\rho), \quad (4.44)$$

where  $\rho$  is the fluid density and  $\mathbf{v}$  the velocity at position  $\mathbf{x}$ . When a constriction occurs, we suppose that it leads to a decreased volume  $V_c$ , that generates a local flow. This flow can be obtained by integrating equation (4.44) inside the tubule and we get:

$$\frac{dV}{dt} = S_2 v_2 - S_1 v_1, \quad (4.45)$$

where  $S_2, S_1$  are the cross-sections at the right and left of the constriction associated with a constant velocity  $v_1$  (resp.  $v_2$ ) on the surface (note that we assume that there is no flow through the lateral surface of the tubule). When  $S_2 = S_1$ , we obtain since  $v_2 = -v_1$  (by symmetry), so that the initial flow is given during the constriction phase by

$$v = \frac{1}{2S_1} \frac{dV}{dt}. \quad (4.46)$$

Thus a constriction occurring in a cylinder of constant section  $\pi r^2$  along a segment of size  $L = 300 \text{ nm}$  during  $15 \text{ ms}$ , leads to a velocity  $V = 0.3/0.03 = 10 \mu\text{m}/\text{s}$ . To recover the velocity at the junction, we need now to model how this change in the tubule shape contributes to the velocity of ejection  $v_{\text{expelled}}$ .

A possible model is that the flow enters the node through a smaller section than the radius

of the tubule. Suppose that the size is  $r = 0.5r_{tub}$ , then using the mass conservation with this ratio of surfaces, we obtain a velocity at the entrance of a junction of

$$v_{\text{entrance}} = \frac{S_1}{S_{\text{entrance}}}v, \quad (4.47)$$

leading to a factor 4. Thus we obtain a velocity that could reach  $v_{\text{entrance}} = 40\mu\text{m}/\text{s}$ , compatible with the maximum velocity we find for the ejection of trajectories (Fig. 4.2).

If the cross-section in the node at the two opposite tubules from the one where the flow is generated are identical to the one receiving the flow, we finally get the relation

$$v_{\text{expelled}} = \frac{v_{\text{entrance}}}{2} \approx 20\mu\text{m}/\text{s}. \quad (4.48)$$

Fig. 4.5h illustrates the consequence of a constriction: a local constriction generates a flow  $v$  in both directions. The flow leads to an acceleration at the entrance of a tubule, if the entrance has a small surface. Due to the flow conservation, the velocity of the expelled trajectory is of the order  $20\mu\text{m}/\text{s}$ . Considering a pinch length of  $L = 100\text{nm}$  we obtain  $v \approx 3\mu\text{m}/\text{s}$ ,  $v_{\text{entrance}} = 13\mu\text{m}/\text{s}$ ,  $v_{\text{expelled}} = 7\mu\text{m}/\text{s}$  while for  $L = 400\text{nm}$  we obtain  $v \approx 13\mu\text{m}/\text{s}$ ,  $v_{\text{entrance}} \approx 53\mu\text{m}/\text{s}$  and  $v_{\text{expelled}} = 27\mu\text{m}/\text{s}$ .

### 4.5.3 Simultaneous contractions statistics

Under the assumption that contractions are spatio-temporally independent events, the probability of a contraction follows a Poissonian distribution of rate  $\lambda$  such that

$$P\{\text{one contraction in } [t, t + \Delta t]\} = \lambda\Delta t \quad (4.49)$$

Thus the probability of  $n$  contractions during that time interval is a rare event which probability is given by

$$\begin{aligned} P\{n \text{ contractions in } [t, t + \Delta t]\} &= (P\{\text{one contraction in } [t, t + \Delta t]\})^n \\ &= (\lambda\Delta t)^n. \end{aligned}$$

Considering  $\lambda = 1/f_{\text{contraction}}$  where  $f_{\text{contraction}} = 1.5\text{Hz}$  as given in Fig. 4.5f during a time step of  $\Delta t = 100\text{ms}$  (the acquisition time of SIM images), then  $P\{\text{one contraction in } [t, t + \Delta t]\} = 0.15$  and  $P\{\text{two contractions in } [t, t + \Delta t]\} = 0.0225$  and thus can be neglected compared to one contraction.

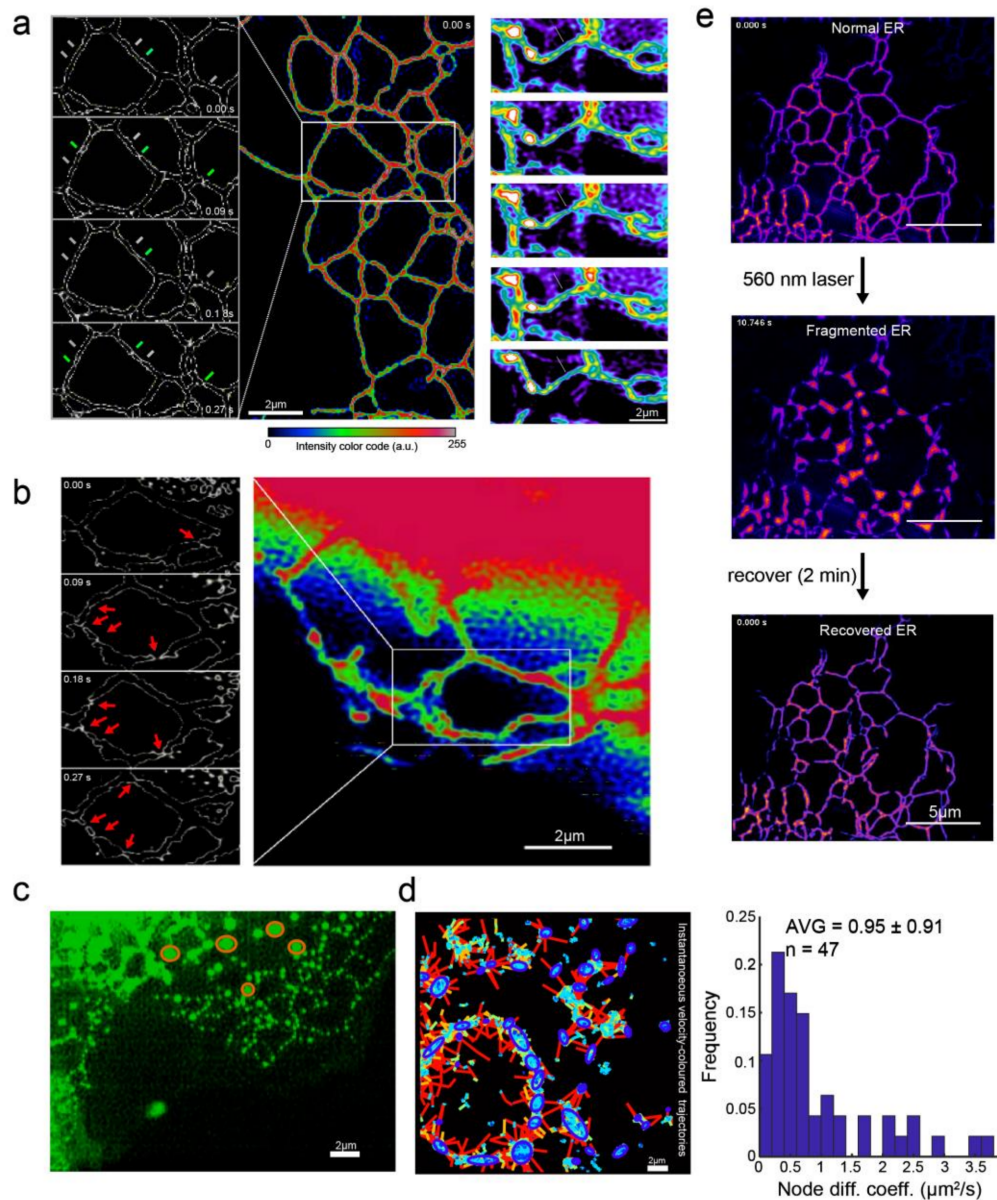


Figure 4.10: Contraction points in ER tubules and ER fragmentation. High-speed Structured Illumination Microscopy (SIM) super-resolved images of the tubular ER for a COS7 cell **A.** and a HEK293 cell **B.**, intensity color-coded. The observed contraction is unlikely to reflect Z-dimension tubule bending at a stiff angle, as sharp bending events have not been seen along the tubules in X-Y, and given the tubule dimensions the observed phenomenon is inconsistent with Z-bending at an angle sufficient to take a small tubule-fragment out of the focal plane. Tubule contractions are visible also in 3D-SIM (Video S9). **C.** Confocal image of the ER, following 20 minutes exposure to 5  $\mu$ M Thapsigargin (an ER calcium uptake inhibitor), Note ER fragmentation accompanied by apparent tube-contraction, through elevation in cytoplasmic calcium [188]. **D.** velocity analysis, performed as in Fig. 4.2, of cell treated to induce ER fragmentation as in c, values given as AVG  $\pm$  SD, n=47 nodes. Ovals exemplify the nearly spherical structures of fragmented ER. **E.** Fragmentation, with notable contracted tubule intermediates, was induced by intense 561 nm laser illumination of cells expressing ER-targeted HaloTag and loaded with excess TRM. (a) characteristic images observed in five independent repeats, (b-e) characteristic images observed in three independent repeats.

## 4.6 Tables

Cell type	Marker	$D_{\text{app}}$ ( $\mu\text{m}^2/\text{s}$ )	$D_{\text{node}}$ ( $\mu\text{m}^2/\text{s}$ )	$v_{\text{flow}}$ ( $\mu\text{m}/\text{s}$ )	$n_{\text{trajs}}$	$n_{\text{exps}}$
		AVG $\pm$ SD	AVG $\pm$ SD	AVG $\pm$ SD	AVG	n
HEK 293T	Calreticulin	$1.64 \pm 0.66$	$0.32 \pm 0.14$	$27.10 \pm 5.69$	14177	7
	Halotag ER	$1.70 \pm 0.94$	$0.39 \pm 0.08$	$26.97 \pm 6.36$	11094	5
COS-7	Calreticulin	$1.96 \pm 1.09$	$0.38 \pm 0.20$	$30.93 \pm 7.44$	13485	6
	Halotag ER	$3.94 \pm 1.03$	$0.57 \pm 0.29$	$42.72 \pm 2.32$	13068	6
SH-SY5Y	Calreticulin	$2.01 \pm 0.43$	$0.40 \pm 0.10$	$30.01 \pm 3.00$	18789	4
	Halotag ER	$2.59 \pm 0.65$	$0.42 \pm 0.19$	$31.70 \pm 3.42$	12925	5

Table 4.1: Single particle motion parameters across cell types, ER tracers.  $D_{\text{node}}$  - diffusion coefficient in tubule junctions;  $v_{\text{flow}}$  - flow velocity in tubules,  $D_{\text{app}}$  - apparent diffusion coefficient computed from without diffusion/flow-deconvolution of single particle displacement profiles.

	Pure Diffusion Model	Flow-Diffusion Model
Dataset 1 (Main text)	$D_{\text{app}} = 1.13 \mu\text{m}^2/\text{s}$	$D_{\text{node}} = 0.19 \mu\text{m}^2/\text{s}$ $\mu_j = 22.9 \mu\text{m}/\text{s}$ $\sigma_j = 6.92 \mu\text{m}/\text{s}$ $\kappa = 0.33$
Dataset 2 (Fig. 4.8a)	$D_{\text{app}} = 1.02 \mu\text{m}^2/\text{s}$	$D_{\text{node}} = 0.35 \mu\text{m}^2/\text{s}$ $\mu_j = 23.19 \mu\text{m}/\text{s}$ $\sigma_j = 3.85 \mu\text{m}/\text{s}$ $\kappa = 0.31$
Dataset 3 (Fig. 4.8b)	$D_{\text{app}} = 4.10 \mu\text{m}^2/\text{s}$	$D_{\text{node}} = 0.57 \mu\text{m}^2/\text{s}$ $\mu_j = 45.01 \mu\text{m}/\text{s}$ $\sigma_j = 12.75 \mu\text{m}/\text{s}$ $\kappa = 0.41$

Table 4.2: Estimated motion parameters of the pure diffusion and flow-diffusion models for the datasets presented in the main text and supplementary figures.

Symbol	Description	Datasets		
		Main text	Fig. 4.8a	Fig. 4.8b
$V_L$	Max. inst. vel. ( $\mu\text{m}/\text{s}$ )	9.5	10	19
$R$	Max. neighbor distance ( $\mu\text{m}$ )	0.1	0.12	0.18
$N$	Min. num. points in cluster	25	35	30

Table 4.3: Parameters used for the network reconstructions.



## Chapter 5

# Discussion: three case studies in cellular trafficking

This chapter shows applications of the SPT analysis methods to study trafficking in cell biology. I also present some extra methods not discussed in the previous chapters and apply them in different cellular contexts: 1. to investigate ER-lysosome interactions; 2. to characterize the motion of  $\text{Ca}_v2.1$  channels at the neuromuscular junction of *Drosophila* and 3. to determine the different modes of motion exhibited by the nuclear proteins composing the NuRD complex.

### 5.0.1 ER-lysosomes interactions participate in maintaining the peripheral ER shape

In this study, in collaboration with Meng Lu and Clemens Kaminski at the University of Cambridge, we looked at the relations between lysosomes and the Endoplasmic Reticulum (ER) network organelles. Lysosomes have many different functions in the cell, the principal one being waste disposal, but they also participate in sensing intracellular elements [197], are involved in cell metabolism [198], protein homeostasis [199] and orchestrate multiple cellular pathways [200]. To ensure these functions, lysosomes are among the most dynamic organelles in the cell [201] and make contact with the ER network [202] but can also be carried by molecular motors along microtubules [203, 204]. On the other hand, the ER itself is a very dynamic organelle with constant tubule and junctions reorganization in its peripheral part [205]. These reorganizations are known to occur in part through microtubules either by sliding of the network along them or pulling out of new microtubules through the attachment of the network to molecular motors [206]. It is also known that the latter mechanism can be mediated by attachment of the ER network first to a lysosome that is itself attached to a molecular motor [207]. In this context, I studied SPTs of lysosomes to understand how their motion is correlated with the ER network.

## Quantifying lysosomes motion

To analyze how lysosomes could interact with the ER, I overlaid images of the ER with the lysosome trajectories (Fig. 1.9) observing a high degree of colocalization between the two signals, suggesting that lysosome motion follows paths along the ER network. Then, I looked at the instantaneous velocity distribution of lysosomes trajectories that is characterized by a spectrum of velocities in the range  $[0 - 3.5]\mu\text{m}/\text{s}$ . To investigate the possible relation between the velocity and the topology of the ER, as is the case in luminal ER motion (see chapter 4), I overlaid on the ER network, the fast (red and purple) and slow (blue) velocity displacements (Fig. 5.1A-B) and found in some cases that the fast velocities occur between low-velocity regions. Interestingly, the distribution  $f(v)$  of velocities (Fig. 5.1B,J), can be fitted by a sum of two exponentials

$$f(v) = A \exp\left(-\frac{v}{v_0}\right) + B \exp\left(-\frac{v}{v_1}\right), \quad (5.1)$$

where a fit gives velocities of  $v_0 = 0.06\mu\text{m}/\text{s}$  (95% confidence interval  $[0.057, 0.072]$ ) and  $v_1 = 0.6\mu\text{m}/\text{s}$  (95% confidence interval  $[0.383, 1.322]$ ), with the constants  $A = 0.20\text{s}/\mu\text{m}$ ,  $B = 0.013\text{s}/\mu\text{m}$  and a coefficient of determination  $R^2 = 0.95$ . This fit shows that lysosomes velocity distribution is mostly driven by the low velocities with rare high-velocity components, suggesting a possible switch between slow and fast motions.

To further study how lysosomes explore the cytoplasm in correlation with the ER, I computed the diffusion map (Fig. 5.1D) and obtained an average apparent diffusion coefficient  $D_{\text{app}} = 0.062 \pm 0.040\mu\text{m}^2/\text{s}$  (Fig. 5.1E). Interestingly regions of low diffusion coefficient colocalize with the region of high density in the density map  $[103, 105]$  (Fig. 5.1C). I then isolated the regions of high density (see method), revealing an ensemble of  $n = 95$  sub-domains, approximated by ellipses (magenta in Fig. 5.1F) of semi-axes sizes  $a = 516 \pm 196$  nm (large semi-axis) and  $b = 278 \pm 143$  nm (Fig. 5.1G). Then, by considering the displacements connecting two high-density regions, I reconstructed (see method 5.0.4) a network explored by the lysosomes (Fig 5.1H). Interestingly, the histogram of average velocities between these regions (Fig. 5.1I) has a mean velocity  $v = 1.03 \pm 0.32\mu\text{m}/\text{s}$ , which is not symmetric and deviated from the Rayleigh distribution (distribution of displacements in dimension two for a random walk). This deviation suggests that the velocities between these regions could be driven by an active motion.

To conclude, lysosomes travel along a network that strongly colocalizes with the ER network with high velocity often occurring in similar regions, suggesting highways when multiple displacement overlap.

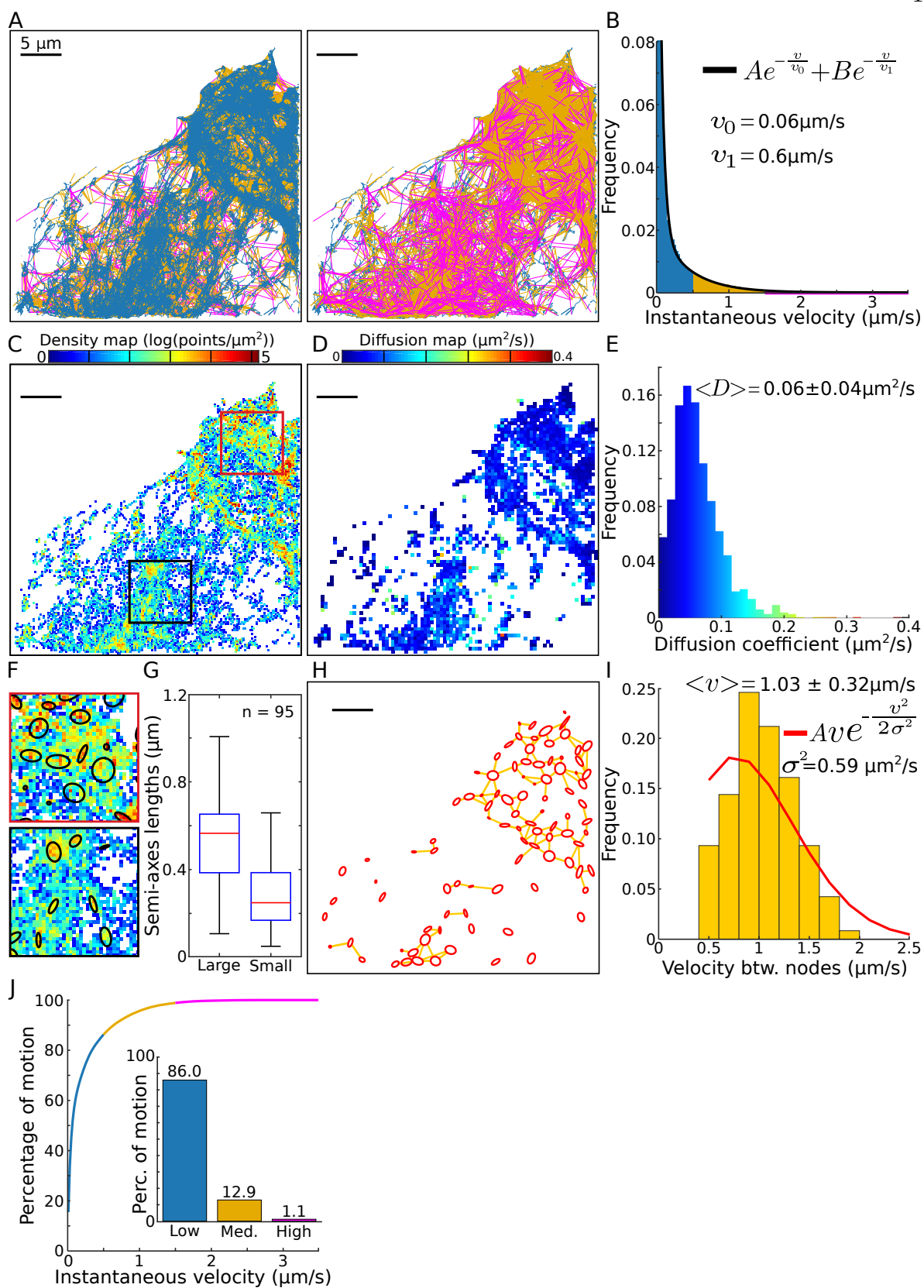




Figure 5.1: (*previous page*) **Lysosome trajectories analysis.** **A.** Lysosome trajectories color-coded according to individual displacement amplitudes (see panel B). **B.** Instantaneous velocities, color-coded with respect to the value and fitted by a double exponential distribution. **C.** Density map of points. **D.** Diffusion map presenting the local diffusion coefficients. **E.** Histogram of the diffusion coefficients obtained from the individual bins of the diffusion map from 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 and enter to the other (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.

## ER-lysosomes interaction

To study the possible interactions between lysosomes and the ER network, I focused on confinement regions found along individual trajectories (see method 5.0.2, and Fig. 5.2A). To recover the size of these confinements, I fitted ellipses over these regions and obtained average semi-axes lengths (Fig. 5.2B)  $a = 232 \pm 77$  nm (large semi-axis) and  $b = 94 \pm 47$  nm (small semi-axis). Furthermore, this approach allows estimating the confinement strength  $\lambda$  by considering that the confined motion could be generated by spring force, as classically described by an Ornstein-Uhlenbeck process [101]. I found an attraction constant of  $\lambda = 0.123 \pm 0.025 s^{-1}$  (Fig. 5.2C), associated with an average local diffusion coefficient of  $D = 0.032 \pm 0.002 \mu m^2/s$  (Fig. 5.2D) 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. 5.2E). The average residence time of lysosomes in these regions is  $\bar{\tau} = 30 \pm 12 s$  that could correspond to a time where lysosomes interact with the ER.

## Discussion

The present analysis of lysosome trajectories reveals that their motion is organized as a network that follows the ER network. This analysis further reveals clear segregation of this lysosome network into regions of low and high velocities, separated by paths that are used multiple times (parallel pink lines in Fig. 5.1), that could be seen as a highway along the ER.

I found that this network is characterized by various types of motion, although a double exponential was enough to account for the entire instantaneous velocity distribution (Fig. 5.1B). This fit revealed two main velocities at 0.6 and 0.06  $\mu m/s$ . However, a more detailed analysis revealed that these velocities can be further subdivided into 1) confined motion (Fig. 5.2) characterized by residence time  $\approx 1$ min. and 2) ballistic motion between regions of high density (Fig. 5.1 F-G), characterized by a distribution of velocity with an average of 1.03  $\mu m/s$ .

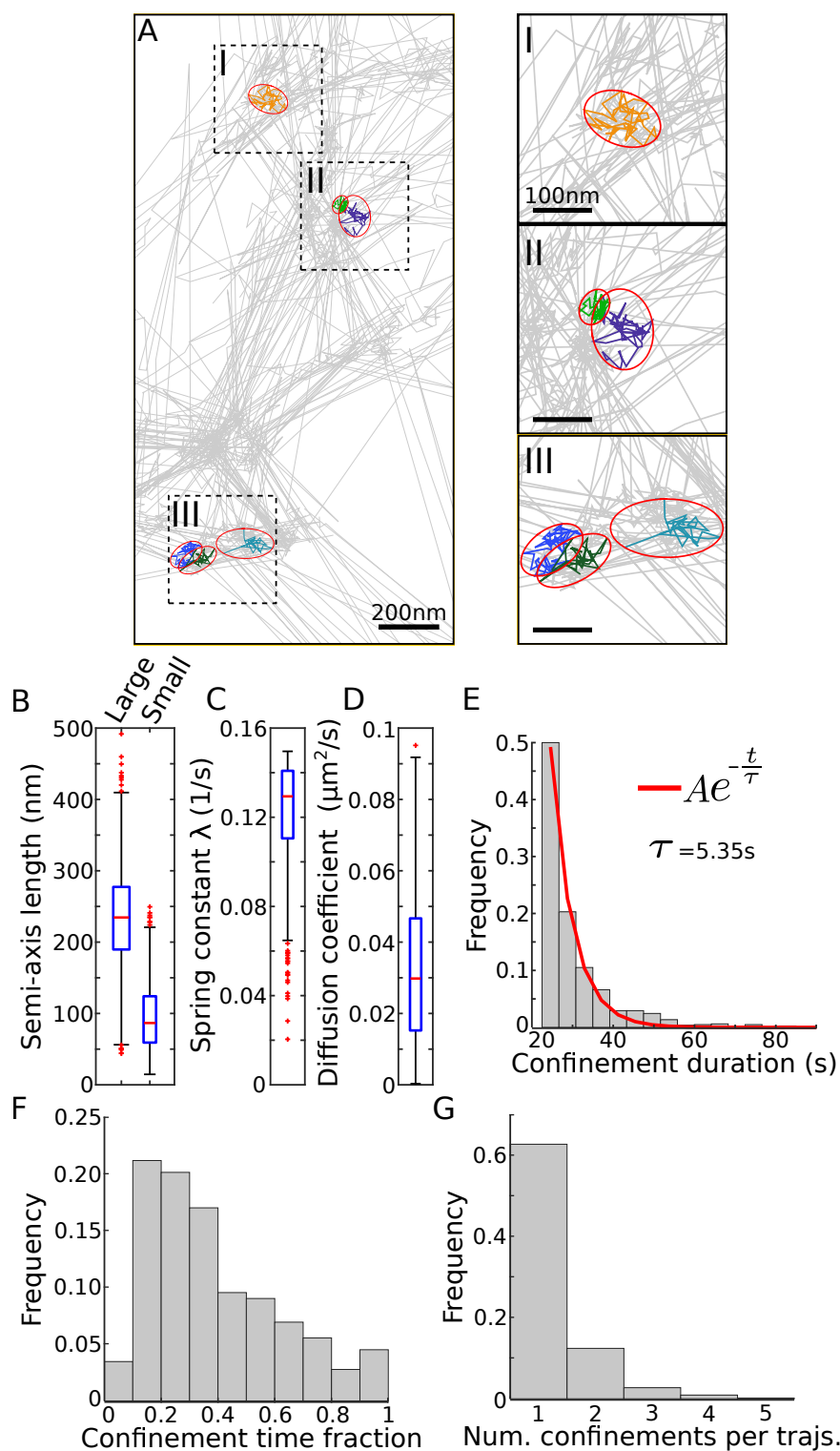


Figure 5.2: **Lysosome confinement time** **A**. Regions of confinement for individual lysosome trajectories: three examples (insets). **B**. Semi-axes (small and large) of the ellipse fitted to the confinement regions. **C**. Spring constants of an Ornstein-Uhlenbeck to confine diffusion. **D**. Diffusion coefficients estimated inside a confinement region. **E**. Residence times inside a confinement region. **F**. Fraction of time each trajectory spend confined (relative to the trajectory length). **G**. Number of confinement events along individual trajectories.

Interestingly 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. I found, however, some regions characterized by a high density of trajectories and low instantaneous velocity, suggesting that there are possible trapping mechanisms to retain lysosome in specific subregions of the ER. This mode of motion is quite different from the 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 is mostly diffusing [113].

To conclude, one striking point of the present analysis is the observed confinement time around 1min of lysosomes, suggesting that lysosomes could be trapped to interact with the ER. Future works should reveal more specifically the interaction time between lysosomes and the ER.

## 5.0.2 Transient confinement detection

Now, when the trajectories are sufficiently long and exhibit extended low instantaneous velocity periods, a confinement detection method based on the distance between successive points of the same trajectories can be used. For each point  $\mathbf{X}_i(t_j)$  ( $j < N_{nh}$ ) of a trajectory  $X_i$ , I compute the center of mass  $\mu_{t_i, N_{nh}}$  based on the ensemble  $e_{\mathbf{X}_i(t_j), N_{nh}} = \{\mathbf{X}_i(t_j), \dots, \mathbf{X}_i(t_{j+N_{nh}})\}$  where  $N_{nh}$  is the number of points kept in the ensemble. I then check the condition  $C_{X_i(t_j), N_{nh}}: \forall X \in e_{t_i, N_{nh}}, \|X - \mu_{\mathbf{X}_i(t_j), N_{nh}}\| < R_{nh}$ , where  $\|\cdot\|$  is the Euclidean norm and  $R_{nh}$  is a distance threshold. I iterate this procedure, increasing each time the size of the ensemble  $N_{nh} \leftarrow N_{nh} + 1$  until either reaching the end of the trajectory or the condition  $C_{\mathbf{X}_i(t_j), N_{nh}}$  is violated. The confinement duration is computed by taking the time difference between the two endpoints the final ensembles  $t_{i+N_{nh}} - t_i$ . The spring constant  $\lambda$  and diffusion coefficient  $D$  of the confinement are obtained by MLE estimation as presented in chapter 4.

## 5.0.3 Analysis of Ca<sub>v</sub>2.1 nanodomains at Drosophila neuro-muscular junctions

In this study, in collaboration with the groups of Martin Heine at U. Mainz, Stephan Sigrist at FU Berlin and Ulrich Thomas at TU Magebeburg, I analyzed SPTs of Ca<sub>v</sub>2.1 channels at the neuromuscular junction of Drosophila. The drosophila has been for decades a model organism in neurosciences due to its easy handling and the large toolbox of genetic manipulations available [208]. Neuro-muscular junctions are synapses that develop between motor neurons and muscle fibers that are essential to control muscle contractions [209]. This synapse is particularly well studied, especially for the relation between presynaptic calcium channels and synaptic vesicles locations in the active zone [210]. I report here several structures that could be extracted in the SPTs of Ca<sub>v</sub>2.1 at the drosophila Neuro-Muscular Junction (NMJ).

## Structures revealed by $\text{Ca}_V$ SPTs

Most of the trajectories are located in regions of high density that are forming structures that recall the EM organization of the NMJ (Fig. 5.3A-B). I extracted the diffusion and density maps (Fig. 5.3A-B (insets)) as well as the drift map at the high-density regions, showing converging arrows, confirming that these regions can be characterized as potential wells. Interestingly, I found that 60% of trajectories are located in the 20% denser regions (Fig. 5.3C), suggesting that  $\text{Ca}_V$  channels are quite confined at NMJ.

## Statistical properties of $\text{Ca}_V$ HDRs

I found that potential wells are formed in regions of high density under different experimental conditions (Fig. 5.4A): control, using different extracellular calcium concentrations (0.2, 1.2 and 1.9mM) or under blocked neuronal activity (addition of PhTX). The characteristics of the wells are presented in Table 5.1. Overall, wells have elliptic boundaries with average semi-axes lengths  $a \approx 80\text{nm}$  (large semi-axis) and  $b \approx 75\text{nm}$  (small semi-axis) (Fig. 5.4B) while the average diffusion coefficient in the wells (see Table 5.1) varies between 0.05 to  $0.18\mu\text{m}^2/\text{s}$  depending on the condition (Fig. 5.4C). Interestingly their average potential energy barrier is  $\approx 5kT$  and mostly independent of the experimental conditions (Fig. 5.4D), associated to a residence time of the channels inside the wells between 600 and 700ms (Fig. 5.4E).

Condition	large semi-axis (nm)	small semi-axis (nm)	Diff. coeff. in well ( $\mu\text{m}^2/\text{s}$ )	Energy (kT)	Residence time (ms)	Num. wells
0.2mM Ca	$88.9 \pm 22.3$	$67.3 \pm 12.2$	$0.016 \pm 0.006$	$5.1 \pm 1.0$	$629 \pm 377$	57
1.2mM Ca	$75.7 \pm 20.0$	$59.0 \pm 14.4$	$0.013 \pm 0.007$	$5.1 \pm 1.0$	$674 \pm 426$	119
1.9mM Ca	$78.5 \pm 26.5$	$58.2 \pm 16.7$	$0.013 \pm 0.007$	$4.7 \pm 1.0$	$578 \pm 338$	52
1.2mM Ca +PhTX	$70.7 \pm 20.4$	$47.3 \pm 8.7$	$0.009 \pm 0.005$	$4.8 \pm 1.1$	$635 \pm 403$	283
1.9mM Ca +PhTX	$76.3 \pm 23.5$	$51.4 \pm 11.7$	$0.011 \pm 0.005$	$4.8 \pm 1.2$	$540 \pm 403$	84

Table 5.1: **Potential well characteristics under different experimental conditions.**

## HDRs are organized in structures

A closer investigation revealed that the HDRs at the NMJ could appear in clusters where trajectories are exchanged (Fig. 5.5A). In that case, I found that the energies of the wells in these structures are  $\approx 3kT$  (Fig. 5.5B), lower than the average energy ( $\approx 5kT$ ) found in other regions. These lower energy could explain that trajectories can escape these potential to travel to the neighboring ones. Another observed feature is that the size of the axes is also reduced compared to isolated wells.

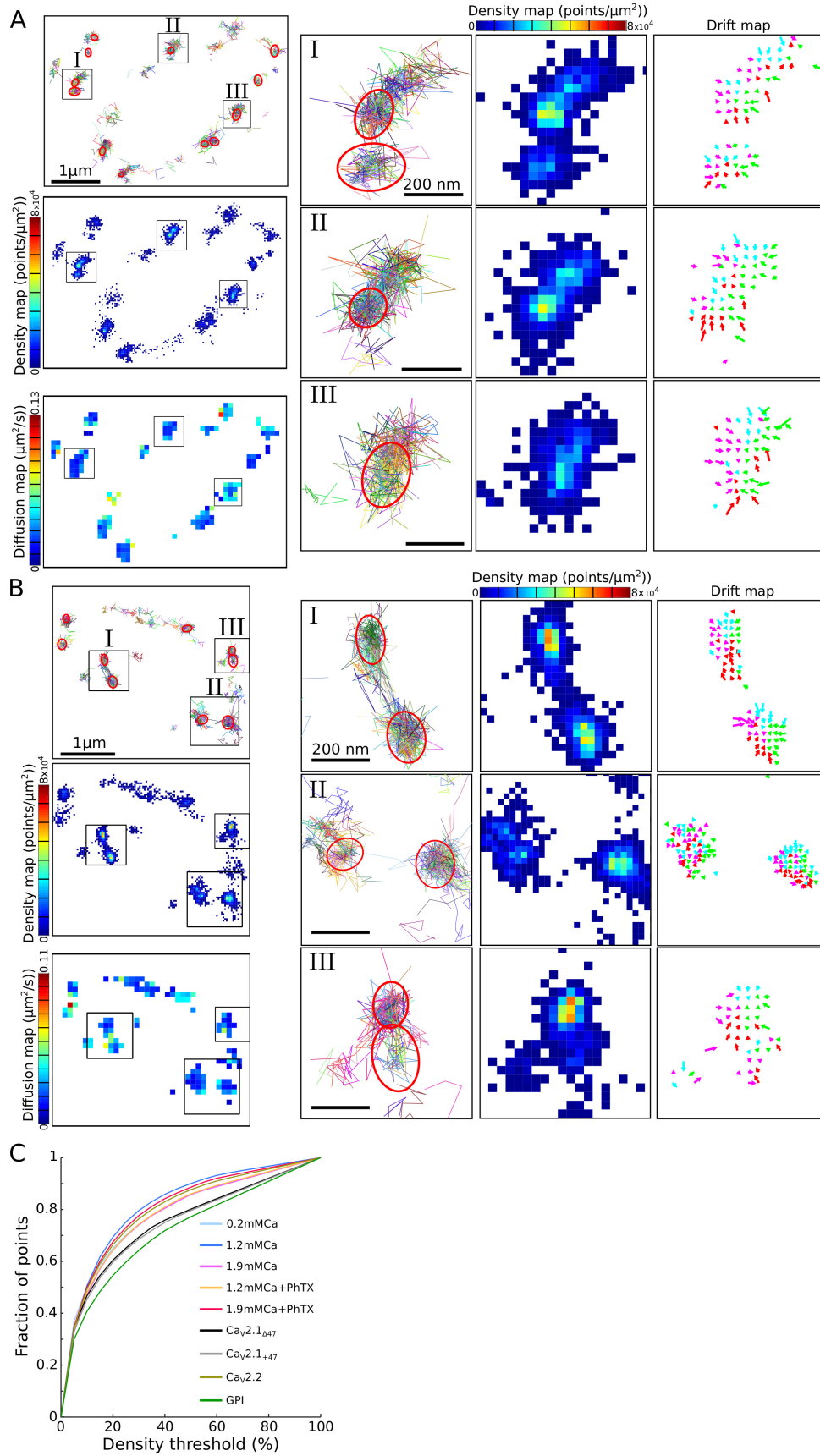


Figure 5.3: (previous page) **Organization of HDRs at the drosophila NMJ.A.-B.** Trajectories, density and diffusion maps of two sample regions where HDRs are organized along the NMJ and for each, a zoom on three regions of interest I, II and III where we show the local density and drift maps. **C.** Fraction of trajectory points falling inside high-density bins as a function of the density threshold, colored by experiments. Bin size  $\Delta x = 30\text{nm}$ .

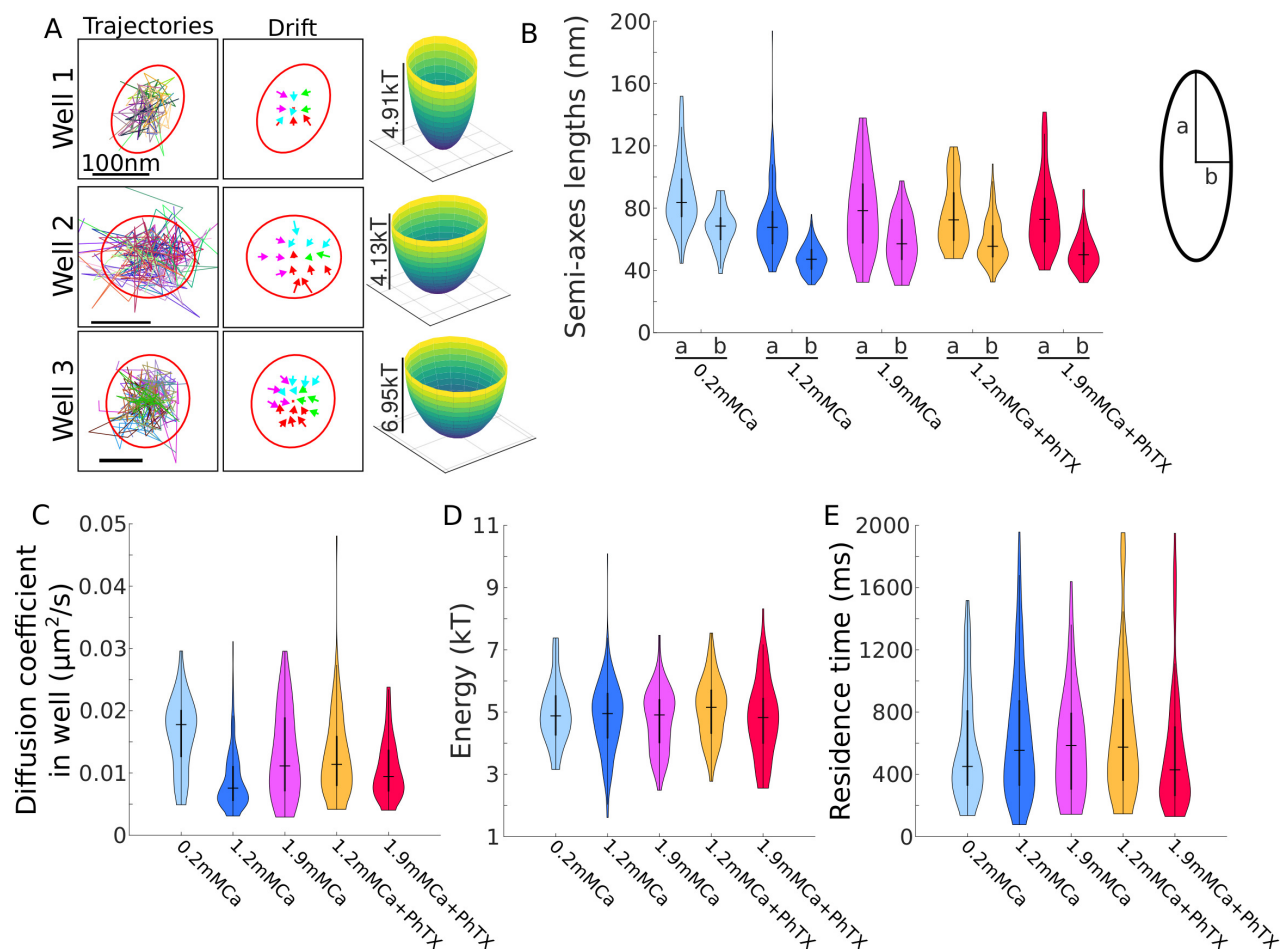


Figure 5.4: **Statistics of potential wells.** **A.** Three examples of HDRs showing trajectories, converging drift and a reconstructed potential well. **B.-E.** Distributions of the semi-axes ( $b$  small and  $a$  large) lengths of the elliptic well boundaries (B), diffusion coefficients the wells (C), Energies (in kT) of the wells (D) and residence times of the channels in the wells (E) for the five experimental conditions.

### Stability of potential wells over time

To evaluate the stability of potential wells over time, I performed a time-lapse analysis with a sliding time window of 20s duration (no overlap between successive windows). For example, the trajectories obtained during a 250s experiment are split into 13, 20s windows (0 – 20s, 20 – 40s, . . . , 240 – 260s) depending on the time of appearance of the detections. For each time windows, I searched for the presence of potential wells (examples are presented in Fig 5.6A)

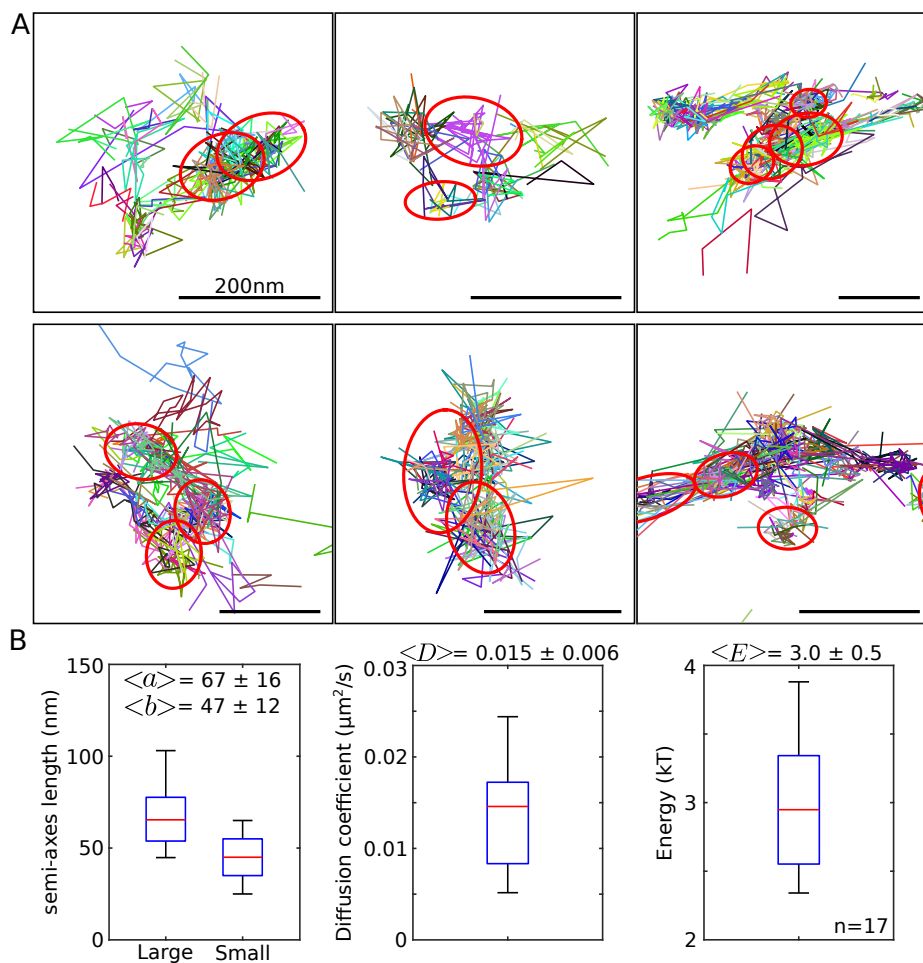


Figure 5.5: **Organized high density regions.** **A.** Six examples of HDR, characterized by potential wells (red ellipses). **B.** Characteristics of the wells detected in Panel A: distribution of large and small semi-axis lengths of the fitted ellipses, the diffusion coefficient and the energy using the algorithm developed in chapter 3.

and then connected through time windows the wells at distances below 250nm of each other, forming chains (allowing multiple frames gap between successive wells). I found that 55% of the wells were present for more than 20s (one time window) and that their average duration is  $127 \pm 84$ s (Fig. 5.6B). This time is much longer than the one of  $\approx 30$ s found for  $\text{Ca}_V2.1$  channels in hippocampal neurons [211].

## Discussion

I found for  $\text{Ca}_V2.1$  SPTs at the NMJ junction of *Drosophila* that most of the trajectories are located in HDRs, this is in contrast with other neuronal cells, such as hippocampal neurons, where  $\text{Ca}_V$  trajectories are found uniformly over the axons (Fig 5.3C). Besides, these HDRs are following the general organization of the NMJ (Fig. 5.3A-B) and are associated to high energy ( $\approx 5$ kT) potential wells (Fig. 5.4), showing that the residence time of  $\text{Ca}_V$  in these regions is  $\approx 600$ ms, which is much longer than  $\text{Ca}_V$  inside the active zones, where the equivalent residence time was about 100ms (see chapter 3). Altogether, these results suggest that HDRs forming potential wells at the NMJ are much more stable than the similar ones in other neurons (Fig. 5.6). Interestingly, the geometry of these regions and energy are mostly independent of the experimental conditions, confirming again their stability (Fig. 5.4).

### 5.0.4 Local high-density region analysis

Here is an alternative method to the one presented in chapter 2 to extract the extent of high-density regions, latter applied to lysosome SPTs. I started by constructing the density map  $\rho(x)$  based on a square grid  $S_{\Delta x}$  with bin size  $\Delta x = 480$  nm covering the whole experimental plane. From this map, I extract the seed grid cells, corresponding to the 5% highest density bins. When multiple seeds appeared within a distance of two bins of each other I kept only the bin with the highest density. For each seed with density  $\rho^*$ , I computed a new density map  $\rho'(x)$  based on a square grid  $S'_{\Delta x'}$  of size  $5 \times 5$ , centered on the center of the seed bin and with bin size  $\Delta x' = 200$  nm. From this local map, I collected the ensemble of points falling into bins with at least a value at 80% of the center bin:  $X' = \{X_i(t_j) \in S'_{\Delta x'}(x) \text{ and } \rho'(x) > 0.8\rho^*\}$ . I then constructed the elliptic boundary of  $X'$  by computing the corresponding 95% confidence ellipse based on the covariance matrix  $Cov(X')$ . Finally, I applied an iterative procedure that merges two overlapping ellipses by fusing the two ensembles of points and computing the elliptic boundary of this new ensemble, until no ellipse overlap anymore.



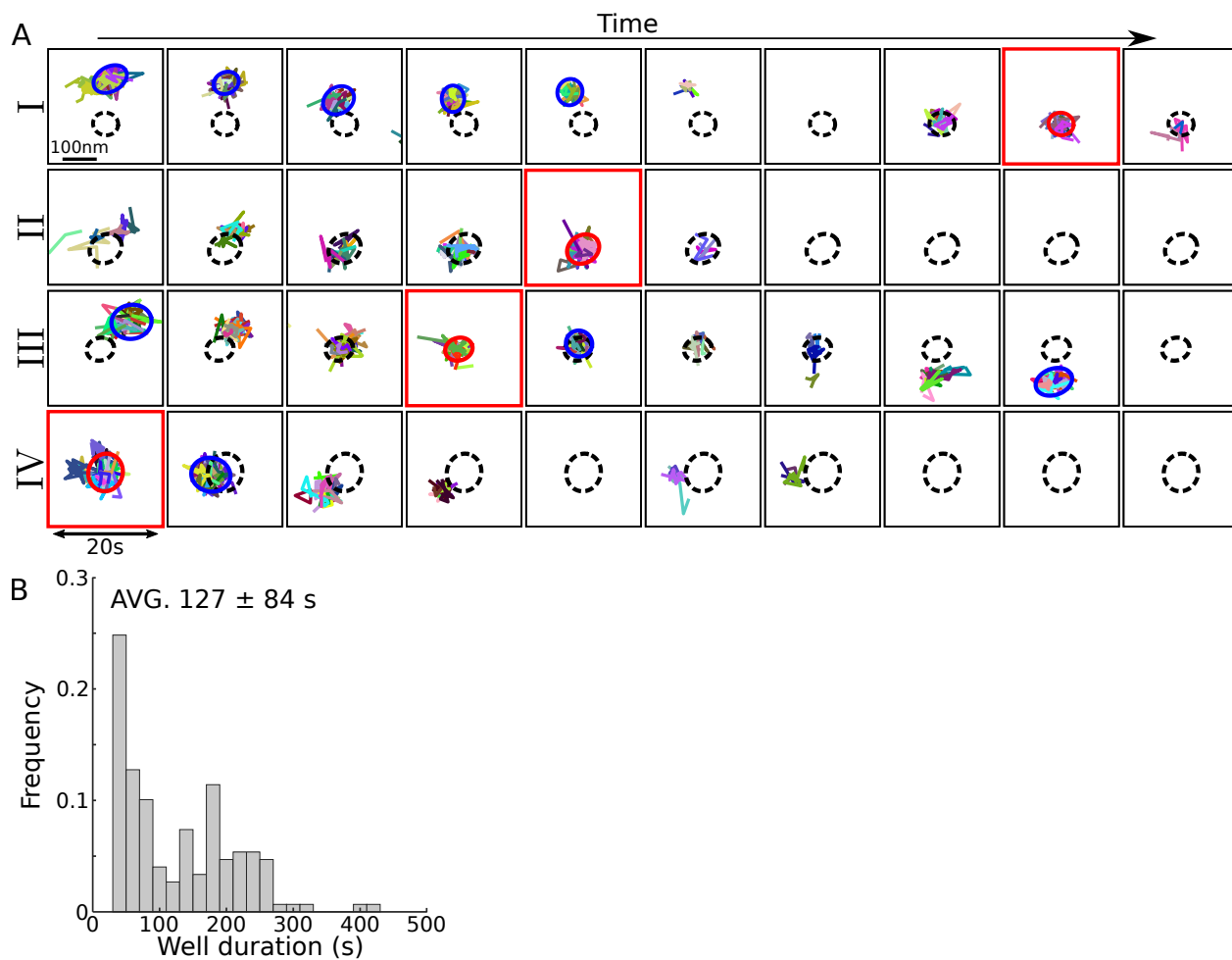


Figure 5.6: **Timelapse analysis.** **A.** Reconstruction of wells from trajectory segmented in non-overlapping time windows of 20s and chaining. The starting point of the chain is the red well in the red square, and it is extended forward and backward in time (left and right). The black dashed well represents the starting well (red) in other windows and the blue wells correspond to other potential wells. **B.** Distribution of well chain duration with average and standard deviation.

### 5.0.5 Analysis of the dynamics of the NuRD nuclear remodeler complex

In this study, in collaboration with the groups of S. Basu and E. Laue at the University of Cambridge, I studied the multiple types of dynamics exhibited by the nuclear proteins composing the NuRD nuclear chromatin remodeling complex. The NuRD complex is formed by the association of seven proteins: HDAC1, HDAC2, RbAp46, RbAp48, MTA1/2/3, MBD2/3, CHD2/4 and I investigated the motion of some of its constituents either in normal conditions or in cells lacking NuRD component molecules. To this end, I used the multiple states model presented in equation (5.9) (method sub-section 5.0.6) where each state is characterized by a diffusion coefficient. Then, one, two and three states models were fitted to each displacement histogram, and the model that best explained the data (using the Bayesian information criterion (see sub-section 5.0.6)) was kept (bold numbers in Table 5.2).

I applied this procedure to the following SPTs datasets: CHD4, CHD4+ATPase (with added ATPase), CHD4-MBD3 (deleted MBD3), CHD4-HDAC (deleted HDAC), MBD3, MTA2, MTA2-MBD3 (deleted MBD3) and MBD3coverslip (MBD3 imaged on a coverslip as a control). The histogram of displacements for these datasets are presented in Figure 5.7A. I found that the number of states in the model of motion that best explain the observed distribution of displacements varies between two states for CHD4+ATPase, CHD4-HDAC, MBD3coverslip and three states for CHD4, MBD3, CHD4-MBD3 (deleted MBD3), MTA2 and MTA2-MBD3 molecules. The parameters of motion are summarized in Table 5.2 and presented in Fig. 5.7B.

Experiments	Single Diff.		Double Diff.				Triple Diff.					$N_d$	$N_e$	
	D	BIC	$\kappa_c$	$D_1$	$D_2$	BIC	$\kappa_c$	$\kappa_{u1}$	$D_1$	$D_2$	$D_3$			BIC
CHD4	0.005	-1868	0.55	0.003	0.010	-3134	0.18	0.65	0.002	0.005	0.017	<b>-3244</b>	118169	9
CHD4+ATPase	0.004	-1674	0.63	0.002	0.013	<b>-2778</b>	0.20	0.55	0.001	0.004	0.017	-2640	3227	3
CHD4-MBD3	0.004	-1852	0.52	0.002	0.008	-3141	0.21	0.63	0.002	0.004	0.015	<b>-3219</b>	63164	4
CHD4-hdac	0.004	-1877	0.55	0.003	0.008	<b>-3018</b>	0.20	0.67	0.002	0.004	0.017	-2952	17377	3
MBD3	0.003	-1679	0.48	0.001	0.006	-2933	0.19	0.58	0.001	0.003	0.011	<b>-3111</b>	105984	8
MTA2	0.004	-1792	0.49	0.002	0.007	-3006	0.14	0.66	0.001	0.004	0.013	<b>-3192</b>	27551	3
MTA2-MBD3	0.004	-1757	0.47	0.002	0.008	-2918	0.14	0.65	0.001	0.004	0.015	<b>-2962</b>	24790	3
MBD3_cslp	0.002	-1784	0.51	0.001	0.003	<b>-2900</b>	0.01	0.55	0.000	0.001	0.003	-2790	5786	1

Table 5.2: **Parameters of the different models of motion.** Diffusion coefficients ( $D, D_1, D_2, D_3$ ) are given in  $\mu\text{m}^2/\text{s}$ ,  $N_d$  number of displacements in the histogram,  $N_e$  number of experiments. In bold are the smallest BIC for each dataset.

### 5.0.6 Model of switching modes of motion

I consider here the case where a trajectory can undergo a Brownian motion with multiple states characterized by different diffusion coefficients. To estimate these coefficients and the switching ratios between the states, I use the distribution of displacements  $\Delta\mathbf{X}_i(t_j) = \mathbf{X}_i(t_{j+1}) - \mathbf{X}_i(t_j)$ . For a Brownian motion with a single diffusion coefficient  $D$ , the proba-

bility of observing a displacement of length  $u$  is given by

$$Pr \{ \|\Delta \mathbf{X}\| = u \} = \frac{u}{2D\Delta t} \exp\left(-\frac{u^2}{4D\Delta t}\right), \quad (5.2)$$

where  $\|\cdot\|$  is the Euclidean norm [212].

A switching model consists of a Markov chain with several states  $k$ , characterized by transition rates between pairs of states  $u$  and  $v$ . For a single molecule and two possible states given at time  $t$  by  $S(t) \in \{1, 2\}$  and characterized by the diffusion coefficients  $D_1$  and  $D_2$ , the associated discrete diffusion process is

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \begin{cases} \sqrt{2D_1\Delta t}\boldsymbol{\eta} & \text{if } S(t) = 1 \\ \sqrt{2D_2\Delta t}\boldsymbol{\eta} & \text{if } S(t) = 2, \end{cases} \quad (5.3)$$

where the Markov chain is given by

$$S(\cdot) = 1 \xrightarrow{\lambda} S(\cdot) = 2. \quad (5.4)$$

For this model, the pdf  $p_d$  for the displacement lengths depends on the state of the process and can be computed by conditioning on the state using Bayes' law:

$$\begin{aligned} p_d(u) = Pr \{ \|\Delta \mathbf{X}\| = u \} &= Pr \{ \|\Delta \mathbf{X}\| = u | S(t) = 1 \} Pr \{ S(t) = 1 \} \\ &+ Pr \{ \|\Delta \mathbf{X}\| = u | S(t) = 2 \} Pr \{ S(t) = 2 \}, \end{aligned} \quad (5.5)$$

where the probabilities  $p_i(t) = Pr \{ S(t) = i \}$  ( $i = 1, 2$ ) are the solutions of the following master equation:

$$\begin{aligned} \frac{d}{dt} p_1 &= -\lambda p_1 + \mu p_2 \\ \frac{d}{dt} p_2 &= \lambda p_1 - \mu p_2, \end{aligned} \quad (5.6)$$

which are given by

$$\begin{aligned} p_1(t) &= \kappa + C e^{-(\lambda+\mu)t} \\ p_2(t) &= 1 - \kappa - C e^{-(\lambda+\mu)t}, \end{aligned} \quad (5.7)$$

with  $\kappa = \frac{\mu}{\lambda+\mu}$ . Thus in two-dimension, for a long time  $t$  and short-time  $\Delta t$ , using equations (5.7) and (5.2) the pdf (5.5) can be written as

$$p_d(u) = \kappa \frac{u}{\sigma_1^2} \exp\left(-\frac{u^2}{2\sigma_1^2}\right) + (1 - \kappa) \frac{u}{\sigma_2^2} \exp\left(-\frac{u^2}{2\sigma_2^2}\right), \quad (5.8)$$

where  $\sigma_k^2 = 2D_k\Delta t$  with  $k = 1, 2$ . More generally, for  $K$  states, the pdf for the displacement is given by

$$Pr_{\text{switch}} \{ \|\Delta \mathbf{X}\| = u \} = \sum_{k=1}^K \left[ \kappa_k \frac{u}{\sigma_k^2} \exp\left(-\frac{u^2}{2\sigma_k^2}\right) \right], \quad (5.9)$$

with  $0 \leq \kappa_k \leq 1$  and  $\sum_{k=1}^K \kappa_k = 1$ .

**Estimation of the model parameters** In order to obtain the ensemble of parameters for a  $K$  state model:  $\{D_1, \dots, D_K, \kappa_1, \dots, \kappa_{K-1}\}$ , I fit equation (5.9) to the histogram of displacements computed from the trajectories. To account for the variability in the optimization procedures, I repeated the fits 50 times with random starting parameter values and kept the parameters set with the highest adjusted coefficient of determination  $R^2$ .

**Model selection** To determine the number of states  $K$  from eq. 5.9 that best explains the data taking into account the quality of the fit and the number of model parameters, I use of the Bayesian Information Criterion (BIC) defined as [213]:

$$BIC = (p + 1)n(\ln n) \left[ \ln \left( \frac{2\pi RSS}{n} \right) + 1 \right], \quad (5.10)$$

with  $p$  the number of parameters of the model,  $n$  the number of data points, and  $RSS$  the residual sum of squares between after fitting the model to the data. For each dataset, the model with the smallest BIC value is chosen as the best fit.

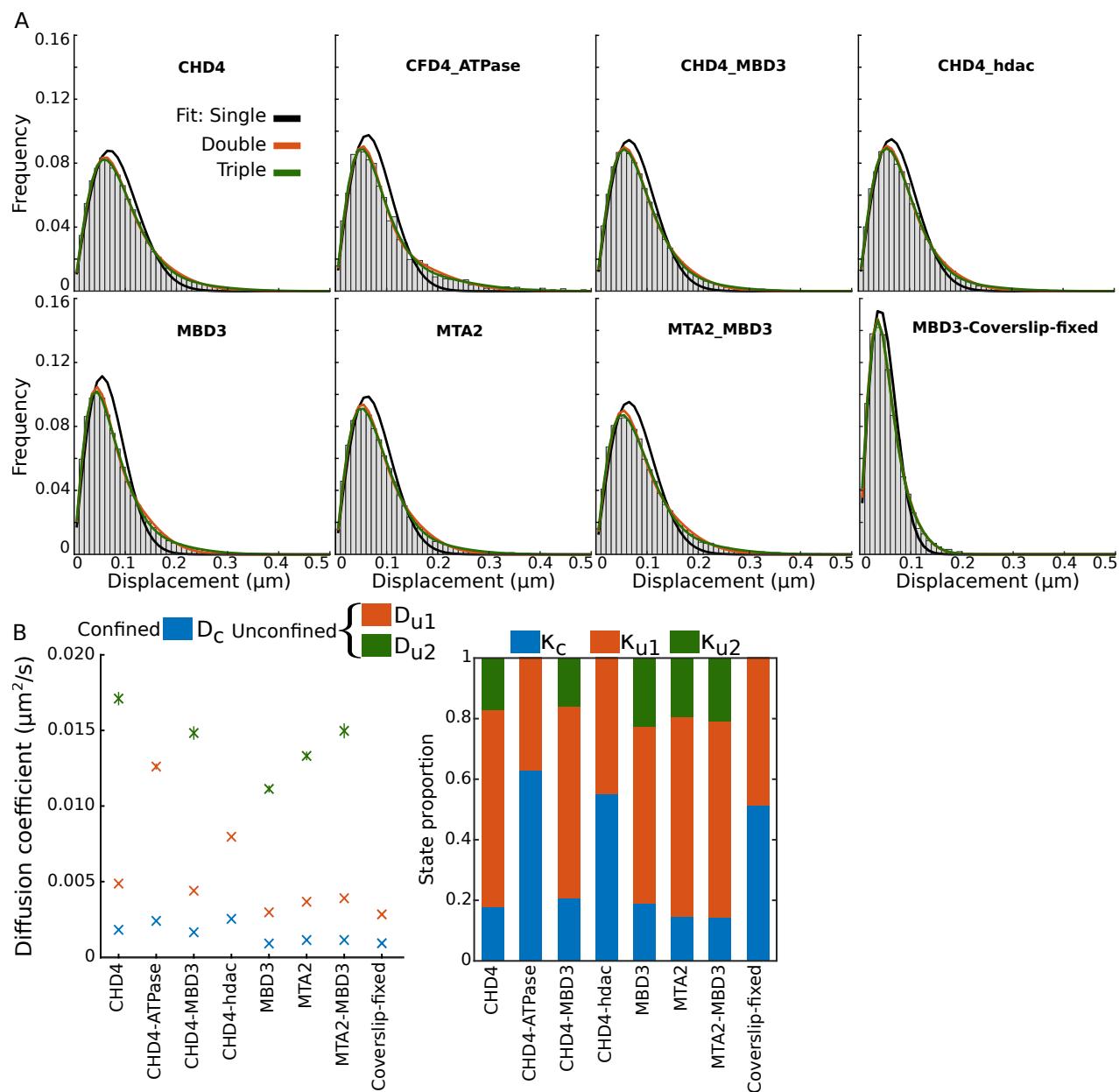


Figure 5.7: **Motion analysis for the proteins composing the NuRD complex.** **A.** Displacement histograms for the considered molecules and conditions together with the corresponding 1,2,3 diffusion states models optimal fits. **B.** Parameters of the selected motion model for each displacement histogram from A, left diffusion coefficients and right proportion of states. Only the parameters for the model with the smallest BIC value are reported (see Table 5.2).

## 5.1 Conclusion and perspectives

Understanding the type of motions of individual molecules and their underlying physical mechanisms is critical in biology. Indeed many fundamental biological processes (eg. signaling, cell cycle) in cells are mediated by the binding and release of molecules with precise spatio-temporal regulation. But molecules, even macromolecules such as proteins, are so small that observing them, especially in live cells, has remained a challenge for years, before the advent of super-resolution microscopy in the late 2000s.

Nowadays, Single Particle Trajectories (SPTs) obtained from Single Molecule Localization Microscopy (a kind of super-resolution technique) methods allow the direct observation of molecular trajectories. But the analysis of SPTs is challenging as trajectories are usually short ( $\approx 5 - 20$  points), strongly affected by thermal fluctuation (modeled as diffusion) and often exhibit multiple modes of motion. The characterization of the latter is of prime interest as they contain information about the underlying physical mechanisms acting on the molecule, such as transient interactions with other proteins, changes in the local environment (crowding or viscosity) or the overlying structure in which the molecules are embedded (e.g. cytoskeleton, organelle, plasma membrane, etc).

In this thesis, I studied two of these modes of motion: potential wells that trap molecules at specific positions at neuronal synapses and active flows that pushes molecules along the tubules of the peripheral endoplasmic reticulum network. Both types of motion are easily observed by eyes from trajectories but the extraction of quantitative measures from them is challenging. The work pursued here followed two axes: 1. the development of new analysis methods for SPTs and 2. applying these method to make contribution in different areas of cellular biology. When dealing with multiple type of motion, classical estimation procedures (either local or MSD) cannot be applied directly anymore and an extra classification step is required to separate trajectory fragment undergoing different types of motion. As I have found, this step is crucial in the correct estimation of the parameters of motion and can be done in different manner: determined from trajectory density (as for the potential wells), determined from the dynamics (as for the junction in the ER network reconstruction) or through other methods such as machine-learning when no clear classification criterion can be extracted directly. This kind of refined models of motion are, in my opinion, what separate the work presented here from the first wave of papers analyzing SPTs and I can foresee that these models will become more and more refined with time. The limiting factor for the diffusion of such analysis method lies, in my opinion, in their complexity and the lack of tools available to easily perform them. A problem that I will address in the near future for the methods presented in this thesis through the development of an ImageJ plugin.

More generally, the future of the single-particle microscopy field is bright with constant improvements in the chemistry of fluorophores and linkers allowing to track new proteins with better signal to noise ratio and over longer periods of time. Associated to this, new microscopes will make 3D recordings more widespread (with reduced difference between lateral and axial resolutions) as well as "multicolor" imaging for concurrently tracking multiple protein species. These new developments will have to be accompanied by new analysis techniques tailored for these type of data.



# Bibliography

- [1] J. Nixon-Abell, C. J. Obara, A. V. Weigel, D. Li, W. R. Legant, C. S. Xu, H. A. Pasolli, K. Harvey, H. F. Hess, E. Betzig, *et al.*, “Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral er,” *Science*, vol. 354, no. 6311, p. aaf3928, 2016.
- [2] H. Mikami, J. Harmon, H. Kobayashi, S. Hamad, Y. Wang, O. Iwata, K. Suzuki, T. Ito, Y. Aisaka, N. Kutsuna, *et al.*, “Ultrafast confocal fluorescence microscopy beyond the fluorescence lifetime limit,” *Optica*, vol. 5, no. 2, pp. 117–126, 2018.
- [3] 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.
- [4] L. Groc, M. Lafourcade, M. Heine, M. Renner, V. Racine, J.-B. Sibarita, B. Lounis, D. Choquet, and L. Cognet, “Surface trafficking of neurotransmitter receptor: comparison between single-molecule/quantum dot strategies,” *Journal of Neuroscience*, vol. 27, no. 46, pp. 12433–12437, 2007.
- [5] L. Cognet, C. Leduc, and B. Lounis, “Advances in live-cell single-particle tracking and dynamic super-resolution imaging,” *Current opinion in chemical biology*, vol. 20, pp. 78–85, 2014.
- [6] 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.
- [7] E. R. Kandel, J. H. Schwartz, T. M. Jessell, D. of Biochemistry, M. B. T. Jessell, S. Siegelbaum, and A. Hudspeth, *Principles of neural science*, vol. 4. McGraw-hill New York, 2000.
- [8] D. V. Lissin, S. N. Gomperts, R. C. Carroll, C. W. Christine, D. Kalman, M. Kitamura, S. Hardy, R. A. Nicoll, R. C. Malenka, and M. Von Zastrow, “Activity differentially regulates the surface expression of synaptic ampa and nmda glutamate receptors,” *Proceedings of the National Academy of Sciences*, vol. 95, no. 12, pp. 7097–7102, 1998.



- [9] R. C. Carroll, D. V. Lissin, M. von Zastrow, R. A. Nicoll, and R. C. Malenka, “Rapid redistribution of glutamate receptors contributes to long-term depression in hippocampal cultures,” *Nature neuroscience*, vol. 2, no. 5, p. 454, 1999.
- [10] L. Chen, D. M. Chetkovich, R. S. Petralia, N. T. Sweeney, Y. Kawasaki, R. J. Wenthold, D. S. Brecht, and R. A. Nicoll, “Stargazin regulates synaptic targeting of ampa receptors by two distinct mechanisms,” *Nature*, vol. 408, no. 6815, p. 936, 2000.
- [11] P. Osten, L. Khatri, J. L. Perez, G. Köhr, G. Giese, C. Daly, T. W. Schulz, A. Wensky, L. M. Lee, and E. B. Ziff, “Mutagenesis reveals a role for abp/grip binding to glur2 in synaptic surface accumulation of the ampa receptor,” *Neuron*, vol. 27, no. 2, pp. 313–325, 2000.
- [12] R. H. Scannevin and R. L. Huganir, “Postsynaptic organisation and regulation of excitatory synapses,” *Nature Reviews Neuroscience*, vol. 1, no. 2, p. 133, 2000.
- [13] J. Meier, C. Vannier, A. Serge, A. Triller, and D. Choquet, “Fast and reversible trapping of surface glycine receptors by gephyrin,” *Nature neuroscience*, vol. 4, no. 3, p. 253, 2001.
- [14] A. J. Borgdorff and D. Choquet, “Regulation of ampa receptor lateral movements,” *Nature*, vol. 417, no. 6889, p. 649, 2002.
- [15] A. Sergé, L. Fourgeaud, A. Hémar, and D. Choquet, “Receptor activation and homer differentially control the lateral mobility of metabotropic glutamate receptor 5 in the neuronal membrane,” *Journal of Neuroscience*, vol. 22, no. 10, pp. 3910–3920, 2002.
- [16] G. Giannone, E. Hosy, F. Levet, A. Constals, K. Schulze, A. I. Sobolevsky, M. P. Rosconi, E. Gouaux, R. Tampé, D. Choquet, *et al.*, “Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density,” *Biophysical journal*, vol. 99, no. 4, pp. 1303–1310, 2010.
- [17] 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.
- [18] 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.
- [19] M.-M. Poo and R. A. Cone, “Lateral diffusion of rhodopsin in the photoreceptor membrane,” *Nature*, vol. 247, no. 5441, p. 438, 1974.
- [20] 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.

- [21] S. Nehls, E. L. Snapp, N. B. Cole, K. J. Zaal, A. K. Kenworthy, T. H. Roberts, J. Ellenberg, J. F. Presley, E. Siggia, and J. Lippincott-Schwartz, “Dynamics and retention of misfolded proteins in native er membranes,” *Nature cell biology*, vol. 2, no. 5, p. 288, 2000.
- [22] D. Choquet and A. Triller, “The role of receptor diffusion in the organization of the postsynaptic membrane,” *Nature Reviews Neuroscience*, vol. 4, no. 4, p. 251, 2003.
- [23] M. Heine, L. Groc, R. Frischknecht, J.-C. Béïque, B. Lounis, G. Rumbaugh, R. L. Huganir, L. Cognet, and D. Choquet, “Surface mobility of postsynaptic ampars tunes synaptic transmission,” *Science*, vol. 320, no. 5873, pp. 201–205, 2008.
- [24] 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.
- [25] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.
- [26] R. Y. Tsien, “The green fluorescent protein,” *Annual Review of Biochemistry*, vol. 67, no. 1, pp. 509–544, 1998. PMID: 9759496.
- [27] A. Lee, K. Tsekouras, C. Calderon, C. Bustamante, and S. Pressé, “Unraveling the thousand word picture: an introduction to super-resolution data analysis,” *Chemical reviews*, vol. 117, no. 11, pp. 7276–7330, 2017.
- [28] 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.
- [29] M. J. Rust, M. Bates, and X. Zhuang, “Stochastic optical reconstruction microscopy (storm) provides sub-diffraction-limit image resolution,” *Nature methods*, vol. 3, no. 10, p. 793, 2006.
- [30] S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy,” *Optics letters*, vol. 19, no. 11, pp. 780–782, 1994.
- [31] M. G. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *Journal of microscopy*, vol. 198, no. 2, pp. 82–87, 2000.
- [32] R. Schneider, E. Hosy, J. Kohl, J. Klueva, D. Choquet, U. Thomas, A. Voigt, and M. Heine, “Mobility of calcium channels in the presynaptic membrane,” *Neuron*, vol. 86, no. 3, pp. 672–679, 2015.

- [33] X. Yang, Z. Lyu, A. Miguel, R. McQuillen, K. C. Huang, and J. Xiao, “Gtpase activity–coupled treadmilling of the bacterial tubulin ftsz organizes septal cell wall synthesis,” *Science*, vol. 355, no. 6326, pp. 744–747, 2017.
- [34] B. Biermann, S. Sokoll, J. Klueva, M. Missler, J. Wiegert, J.-B. Sibarita, and M. Heine, “Imaging of molecular surface dynamics in brain slices using single-particle tracking,” *Nature communications*, vol. 5, p. 3024, 2014.
- [35] S. Berning, K. I. Willig, H. Steffens, P. Dibaj, and S. W. Hell, “Nanoscopy in a living mouse brain,” *Science*, vol. 335, no. 6068, pp. 551–551, 2012.
- [36] H. Steffens, W. Wegner, and K. I. Willig”, “”in vivo sted microscopy: A roadmap to nanoscale imaging in the living mouse”,” *Methods*”, ”2019”.
- [37] I. I. Cisse, I. Izeddin, S. Z. Causse, L. Boudarene, A. Senecal, L. Muresan, C. Dugast-Darzacq, B. Hajj, M. Dahan, and X. Darzacq, “Real-time dynamics of rna polymerase ii clustering in live human cells,” *Science*, vol. 341, no. 6146, pp. 664–667, 2013.
- [38] F. Levet, E. Hosity, 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, p. 1065, 2015.
- [39] J. Neef, N. T. Urban, T.-L. Ohn, T. Frank, P. Jean, S. W. Hell, K. I. Willig, and T. Moser, “Quantitative optical nanophysiology of ca 2+ signaling at inner hair cell active zones,” *Nature communications*, vol. 9, no. 1, p. 290, 2018.
- [40] T. Lagache, N. Sauvonnnet, L. Danglot, and J.-C. Olivo-Marin, “Statistical analysis of molecule colocalization in bioimaging,” *Cytometry Part A*, vol. 87, no. 6, pp. 568–579, 2015.
- [41] P. R. Nicovich, D. M. Owen, and K. Gaus, “Turning single-molecule localization microscopy into a quantitative bioanalytical tool,” *Nature protocols*, vol. 12, no. 3, p. 453, 2017.
- [42] S. J. Sahl, S. W. Hell, and S. Jakobs, “Fluorescence nanoscopy in cell biology,” *Nature reviews Molecular cell biology*, vol. 18, no. 11, p. 685, 2017.
- [43] A. Mazouchi and J. Milstein, “Fast optimized cluster algorithm for localizations (focal): a spatial cluster analysis for super-resolved microscopy,” *Bioinformatics*, vol. 32, no. 5, pp. 747–754, 2015.
- [44] M. A. Kiskowski, J. F. Hancock, and A. K. Kenworthy, “On the use of riple’s k-function and its derivatives to analyze domain size,” *Biophysical journal*, vol. 97, no. 4, pp. 1095–1103, 2009.
- [45] T. Lagache, G. Lang, N. Sauvonnnet, and J.-C. Olivo-Marin, “Analysis of the spatial organization of molecules with robust statistics,” *PLoS One*, vol. 8, no. 12, p. e80914, 2013.

- [46] W.-K. Cho, N. Jayanth, B. P. English, T. Inoue, J. O. Andrews, W. Conway, J. B. Grimm, J.-H. Spille, L. D. Lavis, T. Lionnet, *et al.*, “Rna polymerase ii cluster dynamics predict mrna output in living cells,” *Elife*, vol. 5, p. e13617, 2016.
- [47] Z. Liu, L. D. Lavis, and E. Betzig, “Imaging live-cell dynamics and structure at the single-molecule level,” *Molecular cell*, vol. 58, no. 4, pp. 644–659, 2015.
- [48] T. Savin and P. S. Doyle, “Static and dynamic errors in particle tracking microrheology,” *Biophysical journal*, vol. 88, no. 1, pp. 623–638, 2005.
- [49] C. S. Smith, S. Stallinga, K. A. Lidke, B. Rieger, and D. Grunwald, “Probability-based particle detection that enables threshold-free and robust in vivo single-molecule tracking,” *Molecular biology of the cell*, vol. 26, no. 22, pp. 4057–4062, 2015.
- [50] H. Kong, H. C. Akakin, and S. E. Sarma, “A generalized laplacian of gaussian filter for blob detection and its applications,” *IEEE transactions on cybernetics*, vol. 43, no. 6, pp. 1719–1733, 2013.
- [51] J.-C. Olivo-Marin, “Extraction of spots in biological images using multiscale products,” *Pattern recognition*, vol. 35, no. 9, pp. 1989–1996, 2002.
- [52] I. Izeddin, J. Boulanger, V. Racine, C. Specht, A. Kechkar, D. Nair, A. Triller, D. Choquet, M. Dahan, and J. Sibarita, “Wavelet analysis for single molecule localization microscopy,” *Optics express*, vol. 20, no. 3, pp. 2081–2095, 2012.
- [53] G. Loy and A. Zelinsky, “Fast radial symmetry for detecting points of interest,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 8, pp. 959–973, 2003.
- [54] T.-C. Ku, Y.-N. Huang, C.-C. Huang, D.-M. Yang, L.-S. Kao, T.-Y. Chiu, C.-F. Hsieh, P.-Y. Wu, Y.-S. Tsai, and C.-C. Lin, “An automated tracking system to measure the dynamic properties of vesicles in living cells,” *Microscopy research and technique*, vol. 70, no. 2, pp. 119–134, 2007.
- [55] N. Chenouard, I. Bloch, and J.-C. Olivo-Marin, “Multiple hypothesis tracking for cluttered biological image sequences,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 11, pp. 2736–3750, 2013.
- [56] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid, and G. Danuser, “Robust single-particle tracking in live-cell time-lapse sequences,” *Nature methods*, vol. 5, no. 8, p. 695, 2008.
- [57] V. Racine, A. Hertzog, J. Jouanneau, J. Salamero, C. Kervrann, and J.-B. Sibarita, “Multiple-target tracking of 3d fluorescent objects based on simulated annealing,” in *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006.*, pp. 1020–1023, IEEE, 2006.
- [58] J.-Y. Tinevez, N. Perry, J. Schindelin, G. M. Hoopes, G. D. Reynolds, E. Laplantine, S. Y. Bednarek, S. L. Shorte, and K. W. Eliceiri, “Trackmate: An open and extensible platform for single-particle tracking,” *Methods*, vol. 115, pp. 80–90, 2017.

- [59] A. Einstein, “Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen,” *Annalen der physik*, vol. 322, no. 8, pp. 549–560, 1905.
- [60] M. Von Smoluchowski, “Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen,” *Annalen der physik*, vol. 326, no. 14, pp. 756–780, 1906.
- [61] P. Langevin, “Sur la théorie du mouvement brownien,” *Compt. Rendus*, vol. 146, pp. 530–533, 1908.
- [62] P. G. Saffman and M. Delbrück, “Brownian motion in biological membranes,” *Proceedings of the National Academy of Sciences*, vol. 72, no. 8, pp. 3111–3113, 1975.
- [63] C. W. Gardiner *et al.*, *Handbook of stochastic methods*, vol. 3. springer Berlin, 1985.
- [64] N. Hozé and D. Holcman, “Statistical methods for large ensembles of super-resolution stochastic single particle trajectories in cell biology,” *Annual Review of Statistics and Its Application*, vol. 4, pp. 189–223, 2017.
- [65] G. E. Uhlenbeck and L. S. Ornstein, “On the theory of the brownian motion,” *Physical review*, vol. 36, no. 5, p. 823, 1930.
- [66] D. Holcman, N. Hoze, and Z. Schuss, “Narrow escape through a funnel and effective diffusion on a crowded membrane,” *Physical Review E*, vol. 84, no. 2, p. 021906, 2011.
- [67] Y. A. Domanov, S. Aimon, G. E. Toombes, M. Renner, F. Quemeneur, A. Triller, M. S. Turner, and P. Bassereau, “Mobility in geometrically confined membranes,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 31, pp. 12605–12610, 2011.
- [68] F. Quemeneur, J. K. Sigurdsson, M. Renner, P. J. Atzberger, P. Bassereau, and D. Lacoste, “Shape matters in protein mobility within membranes,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 14, pp. 5083–5087, 2014.
- [69] X. Michalet, “Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium,” *Physical Review E*, vol. 82, no. 4, p. 041914, 2010.
- [70] H. Qian, M. P. Sheetz, and E. L. Elson, “Single particle tracking. analysis of diffusion and flow in two-dimensional systems,” *Biophysical journal*, vol. 60, no. 4, pp. 910–921, 1991.
- [71] M. J. Saxton, “Single-particle tracking: the distribution of diffusion coefficients,” *Biophysical journal*, vol. 72, no. 4, pp. 1744–1753, 1997.
- [72] A. Andreadov and D. S. Grebenkov, “Time-averaged msd of brownian motion,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2012, no. 07, p. P07001, 2012.

- [73] V. Tejedor, O. Bénichou, R. Voituriez, R. Jungmann, F. Simmel, C. Selhuber-Unkel, L. B. Oddershede, and R. Metzler, “Quantitative analysis of single particle trajectories: mean maximal excursion method,” *Biophysical journal*, vol. 98, no. 7, pp. 1364–1372, 2010.
- [74] M. Renner, L. Wang, S. Levi, L. Hennekinne, and A. Triller, “A simple and powerful analysis of lateral subdiffusion using single particle tracking,” *Biophysical journal*, vol. 113, no. 11, pp. 2452–2463, 2017.
- [75] 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.
- [76] Z. Schuss, *Theory and applications of stochastic processes: an analytical approach*, vol. 170. Springer Science & Business Media, 2009.
- [77] A. J. Berglund, “Statistics of camera-based single-particle tracking,” *Physical Review E*, vol. 82, no. 1, p. 011917, 2010.
- [78] C. L. Vestergaard, P. C. Blainey, and H. Flyvbjerg, “Optimal estimation of diffusion coefficients from single-particle trajectories,” *Physical Review E*, vol. 89, no. 2, p. 022726, 2014.
- [79] 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.
- [80] H. Risken, *Fokker-planck equation*. Springer, 1996.
- [81] R. Das, C. W. Cairo, and D. Coombs, “A hidden markov model for single particle tracks quantifies dynamic interactions between lfa-1 and the actin cytoskeleton,” *PLoS computational biology*, vol. 5, no. 11, p. e1000556, 2009.
- [82] 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.
- [83] C. Guerrier and D. Holcman, “The first 100 nm inside the pre-synaptic terminal where calcium diffusion triggers vesicular release,” *Frontiers in synaptic neuroscience*, vol. 10, 2018.
- [84] A.-H. Tang, H. Chen, T. P. Li, S. R. Metzbower, H. D. MacGillavry, and T. A. Blanpied, “A trans-synaptic nanocolumn aligns neurotransmitter release to receptors,” *Nature*, vol. 536, no. 7615, p. 210, 2016.
- [85] A. Thalhammer, A. Contestabile, Y. S. Ermolyuk, T. Ng, K. E. Volynski, T. W. Soong, Y. Goda, and L. A. Cingolani, “Alternative splicing of p/q-type ca<sup>2+</sup> channels shapes presynaptic plasticity,” *Cell reports*, vol. 20, no. 2, pp. 333–343, 2017.
- [86] 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.

- [87] A. Y. Lai and P. A. Wade, “Nurd: A multi-faceted chromatin remodeling complex in regulating cancer biology,” *Nature reviews. Cancer*, vol. 11, no. 8, p. 588, 2011.
- [88] W. Zhang, A. Aubert, J. G. de Segura, M. Karuppasamy, S. Basu, A. Murthy, A. Diamante, T. Drury, J. Balmer, J. Cramard, *et al.*, “The nucleosome remodeling and deacetylase complex nurd is built from preformed catalytically active sub-modules,” *Journal of molecular biology*, vol. 428, no. 14, pp. 2931–2942, 2016.
- [89] C. P. Caridi, C. D’Agostino, T. Ryu, G. Zapotoczny, L. Delabaere, X. Li, V. Y. Khodaverdian, N. Amaral, E. Lin, A. R. Rau, *et al.*, “Nuclear f-actin and myosins drive relocalization of heterochromatic breaks,” *Nature*, vol. 559, no. 7712, p. 54, 2018.
- [90] S. Burov, S. A. Tabei, T. Huynh, M. P. Murrell, L. H. Philipson, S. A. Rice, M. L. Gardel, N. F. Scherer, and A. R. Dinner, “Distribution of directional change as a signature of complex dynamics,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 49, pp. 19689–19694, 2013.
- [91] 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.
- [92] C. L. Vestergaard, J. N. Pedersen, K. I. Mortensen, and H. Flyvbjerg, “Estimation of motility parameters from trajectory data,” *The European Physical Journal Special Topics*, vol. 224, no. 7, pp. 1151–1168, 2015.
- [93] C. P. Calderon, L. E. Weiss, and W. Moerner, “Robust hypothesis tests for detecting statistical evidence of two-dimensional and three-dimensional interactions in single-molecule measurements,” *Physical Review E*, vol. 89, no. 5, p. 052705, 2014.
- [94] I. Bronshtein, E. Kepten, I. Kanter, S. Berezin, M. Lindner, A. B. Redwood, S. Mai, S. Gonzalo, R. Foisner, Y. Shav-Tal, *et al.*, “Loss of lamin a function increases chromatin dynamics in the nuclear interior,” *Nature communications*, vol. 6, p. 8044, 2015.
- [95] V. Briane, C. Kervrann, and M. Vimond, “Statistical analysis of particle trajectories in living cells,” *Physical Review E*, vol. 97, no. 6, p. 062121, 2018.
- [96] R. Huganir and R. Nicoll, “Ampars and synaptic plasticity: the last 25 years,” *Neuron*, vol. 80, no. 3, pp. 704–717, 2013.
- [97] R. A. Nicoll, “A brief history of long-term potentiation,” *Neuron*, vol. 93, no. 2, pp. 281–290, 2017.
- [98] H. Jenifer, P. Pierre, C. Anna, B. Arthur, F. Romy, M. Jessica, A. Maria, F. Anna, H. David, and H. Martin, “Transient confinement of cav2.1 ca2+-channel splice variants shapes synaptic short-term plasticity,” *Neuron*, 2019.

- [99] 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.
- [100] D. Holcman, “Unraveling novel features hidden in superresolution microscopy data,” *Communicative & integrative biology*, vol. 6, no. 3, p. e23893, 2013.
- [101] Z. Schuss, *Theory and Applications of Stochastic Processes: An Analytical Approach*. Springer New York, 2010.
- [102] D. Holcman and Z. Schuss, “100 years after smoluchowski: stochastic processes in cell biology,” *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 9, p. 093002, 2017.
- [103] 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.
- [104] Z. Schuss, *Diffusion and stochastic processes: an analytical approach*. Springer New York, 2010.
- [105] D. Holcman, N. Hoze, and Z. Schuss, “Analysis and interpretation of superresolution single-particle trajectories,” *Biophysical journal*, vol. 109, no. 9, pp. 1761–1771, 2015.
- [106] M. J. Saxton, “Single-particle tracking: connecting the dots,” *Nature methods*, vol. 5, no. 8, p. 671, 2008.
- [107] N. Chenouard, I. Smal, F. De Chaumont, M. Maška, I. F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, *et al.*, “Objective comparison of particle tracking methods,” *Nature methods*, vol. 11, no. 3, p. 281, 2014.
- [108] G.-J. Zhao and K.-L. Han, “Hydrogen bonding in the electronic excited state,” *Accounts of chemical research*, vol. 45, no. 3, pp. 404–413, 2011.
- [109] 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.
- [110] 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, p. 16283, 2018.
- [111] T. Sungkaworn, M.-L. Jobin, K. Burnecki, A. Weron, M. J. Lohse, and D. Calebiro, “Single-molecule imaging reveals receptor–g protein interactions at cell surface hot spots,” *Nature*, vol. 550, no. 7677, p. 543, 2017.



- [112] S. Cohen, A. M. Valm, and J. Lippincott-Schwartz, “Interacting organelles,” *Current opinion in cell biology*, vol. 53, pp. 84–91, 2018.
- [113] 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.
- [114] N. Hoze and D. Holcman, “Recovering a stochastic process from super-resolution noisy ensembles of single-particle trajectories,” *Physical Review E*, vol. 92, no. 5, p. 052109, 2015.
- [115] S. Karlin and H. E. Taylor, *A second course in stochastic processes*. Elsevier, 1981.
- [116] L. Abbott and W. G. Regehr, “Synaptic computation,” *Nature*, vol. 431, no. 7010, p. 796, 2004.
- [117] R. S. Zucker, “Short-term synaptic plasticity,” *Annual review of neuroscience*, vol. 12, no. 1, pp. 13–31, 1989.
- [118] M. A. Böhme, A. T. Grasskamp, and A. M. Walter, “Regulation of synaptic release-site  $ca_{2+}$  channel coupling as a mechanism to control release probability and short-term plasticity,” *FEBS letters*, vol. 592, no. 21, pp. 3516–3531, 2018.
- [119] H. Sakamoto, T. Ariyoshi, N. Kimpara, K. Sugao, I. Taiko, K. Takikawa, D. Asanuma, S. Namiki, and K. Hirose, “Synaptic weight set by munc13-1 supramolecular assemblies,” *Nature neuroscience*, vol. 21, no. 1, p. 41, 2018.
- [120] R. Schneggenburger and C. Rosenmund, “Molecular mechanisms governing  $ca_{2+}$  regulation of evoked and spontaneous release,” *Nature Neuroscience*, vol. 18, no. 7, p. 935, 2015.
- [121] E. Eggermann, I. Bucurenciu, S. P. Goswami, and P. Jonas, “Nanodomain coupling between  $ca_{2+}$  channels and sensors of exocytosis at fast mammalian synapses,” *Nature Reviews Neuroscience*, vol. 13, no. 1, p. 7, 2012.
- [122] R. S. Zucker and W. G. Regehr, “Short-term synaptic plasticity,” *Annual review of physiology*, vol. 64, no. 1, pp. 355–405, 2002.
- [123] I. Bucurenciu, A. Kulik, B. Schwaller, M. Frotscher, and P. Jonas, “Nanodomain coupling between  $ca_{2+}$  channels and  $ca_{2+}$  sensors promotes fast and efficient transmitter release at a cortical gabaergic synapse,” *Neuron*, vol. 57, no. 4, pp. 536–545, 2008.
- [124] A. Scimemi and J. S. Diamond, “The number and organization of  $ca_{2+}$  channels in the active zone shapes neurotransmitter release from schaffer collateral synapses,” *Journal of Neuroscience*, vol. 32, no. 50, pp. 18157–18176, 2012.
- [125] E. F. Stanley, “Single calcium channels and acetylcholine release at a presynaptic nerve terminal,” *Neuron*, vol. 11, no. 6, pp. 1007–1011, 1993.

- [126] D. Baur, G. Bornschein, D. Althof, M. Watanabe, A. Kulik, J. Eilers, and H. Schmidt, “Developmental tightening of cerebellar cortical synaptic influx-release coupling,” *Journal of Neuroscience*, vol. 35, no. 5, pp. 1858–1871, 2015.
- [127] D. Keller, N. Babai, O. Kochubey, Y. Han, H. Markram, F. Schürmann, and R. Schneggenburger, “An exclusion zone for  $ca_2+$  channels around docked vesicles explains release control by multiple channels at a cns synapse,” *PLoS computational biology*, vol. 11, no. 5, p. e1004253, 2015.
- [128] E. F. Stanley, “The nanophysiology of fast transmitter release,” *Trends in Neurosciences*, vol. 39, no. 3, pp. 183–197, 2016.
- [129] N. P. Vyleta and P. Jonas, “Loose coupling between  $ca_2+$  channels and release sensors at a plastic hippocampal synapse,” *Science*, vol. 343, no. 6171, pp. 665–670, 2014.
- [130] H. Schmidt, S. Brachtendorf, O. Arendt, S. Hallermann, S. Ishiyama, G. Bornschein, D. Gall, S. N. Schiffmann, M. Heckmann, and J. Eilers, “Nanodomain coupling at an excitatory cortical synapse,” *Current biology*, vol. 23, no. 3, pp. 244–249, 2013.
- [131] C. Guerrier and D. Holcman, “Hybrid markov-mass action law model for cell activation by rare binding events: application to calcium induced vesicular release at neuronal synapses,” *Scientific reports*, vol. 6, p. 35506, 2016.
- [132] N. Holderith, A. Lorincz, G. Katona, B. Rózsa, A. Kulik, M. Watanabe, and Z. Nusser, “Release probability of hippocampal glutamatergic terminals scales with the size of the active zone,” *Nature neuroscience*, vol. 15, no. 7, p. 988, 2012.
- [133] T. Miki, W. A. Kaufmann, G. Malagon, L. Gomez, K. Tabuchi, M. Watanabe, R. Shigemoto, and A. Marty, “Numbers of presynaptic  $ca_2+$  channel clusters match those of functionally defined vesicular docking sites in single central synapses,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 26, pp. E5246–E5255, 2017.
- [134] S. Nadkarni, T. M. Bartol, C. F. Stevens, T. J. Sejnowski, and H. Levine, “Short-term plasticity constrains spatial organization of a hippocampal presynaptic terminal,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 36, pp. 14657–14662, 2012.
- [135] Y. Nakamura, H. Harada, N. Kamasawa, K. Matsui, J. S. Rothman, R. Shigemoto, R. A. Silver, D. A. DiGregorio, and T. Takahashi, “Nanoscale distribution of presynaptic  $ca_2+$  channels and its impact on vesicular release during development,” *Neuron*, vol. 85, no. 1, pp. 145–158, 2015.
- [136] C. A. Reid, J. D. Clements, and J. M. Bekkers, “Nonuniform distribution of  $ca_2+$  channel subtypes on presynaptic terminals of excitatory synapses in hippocampal cultures,” *Journal of Neuroscience*, vol. 17, no. 8, pp. 2738–2745, 1997.
- [137] S. Tsuruel, A. Fisher, N. Wittenmayer, T. Dresbach, C. C. Garner, and N. E. Ziv, “Exchange and redistribution dynamics of the cytoskeleton of the active zone molecule bassoon,” *Journal of Neuroscience*, vol. 29, no. 2, pp. 351–358, 2009.

- [138] D. Maschi and V. A. Klyachko, "Spatiotemporal regulation of synaptic vesicle fusion sites in central synapses," *Neuron*, vol. 94, no. 1, pp. 65–73, 2017.
- [139] C. A. Reid, J. M. Bekkers, and J. D. Clements, "Presynaptic  $ca^{2+}$  channels: a functional patchwork," *Trends in neurosciences*, vol. 26, no. 12, pp. 683–687, 2003.
- [140] J. Rettig, C. Heinemann, U. Ashery, Z.-H. Sheng, C. T. Yokoyama, W. A. Catterall, and E. Neher, "Alteration of  $ca^{2+}$  dependence of neurotransmitter release by disruption of  $ca^{2+}$  channel/syntaxin interaction," *Journal of Neuroscience*, vol. 17, no. 17, pp. 6647–6656, 1997.
- [141] Z. Szabo, G. J. Obermair, C. B. Cooper, G. W. Zamponi, and B. E. Flucher, "Role of the synprint site in presynaptic targeting of the calcium channel  $cav2.2$  in hippocampal neurons," *European Journal of Neuroscience*, vol. 24, no. 3, pp. 709–718, 2006.
- [142] Y. Han, P. S. Kaeser, T. C. Südhof, and R. Schneggenburger, "Rim determines  $ca^{2+}$  channel density and vesicle docking at the presynaptic active zone," *Neuron*, vol. 69, no. 2, pp. 304–316, 2011.
- [143] P. S. Kaeser, L. Deng, Y. Wang, I. Dulubova, X. Liu, J. Rizo, and T. C. Südhof, "Rim proteins tether  $ca^{2+}$  channels to presynaptic active zones via a direct pdz-domain interaction," *Cell*, vol. 144, no. 2, pp. 282–295, 2011.
- [144] K. S. Liu, M. Siebert, S. Mertel, E. Knoche, S. Wegener, C. Wichmann, T. Matkovic, K. Muhammad, H. Depner, C. Mettke, *et al.*, "Rim-binding protein, a central part of the active zone, is essential for neurotransmitter release," *Science*, vol. 334, no. 6062, pp. 1565–1569, 2011.
- [145] T. W. Soong, C. D. DeMaria, R. S. Alvania, L. S. Zweifel, M. C. Liang, S. Mittman, W. S. Agnew, and D. T. Yue, "Systematic identification of splice variants in human p/q-type channel  $\alpha12.1$  subunits: implications for current density and  $ca^{2+}$ -dependent inactivation," *Journal of Neuroscience*, vol. 22, no. 23, pp. 10142–10152, 2002.
- [146] M. Hirano, Y. Takada, C. F. Wong, K. Yamaguchi, H. Kotani, T. Kurokawa, M. X. Mori, T. P. Snutch, M. Ronjat, M. De Waard, *et al.*, "C-terminal splice variants of p/q-type  $ca^{2+}$  channel  $cav2.1 \alpha1$  subunits are differentially regulated by rab3-interacting molecule proteins," *Journal of Biological Chemistry*, vol. 292, no. 22, pp. 9365–9381, 2017.
- [147] A. Maximov and I. Bezprozvanny, "Synaptic targeting of n-type calcium channels in hippocampal neurons," *Journal of Neuroscience*, vol. 22, no. 16, pp. 6939–6952, 2002.
- [148] T. Aikawa, T. Watanabe, T. Miyazaki, T. Mikuni, M. Wakamori, M. Sakurai, H. Aizawa, N. Ishizu, M. Watanabe, M. Kano, *et al.*, "Alternative splicing in the c-terminal tail of  $cav2.1$  is essential for preventing a neurological disease in mice," *Human molecular genetics*, vol. 26, no. 16, pp. 3094–3104, 2017.

- [149] Q. Hu, H. Saegusa, Y. Hayashi, and T. Tanabe, “The carboxy-terminal tail region of human cav2. 1 (p/q-type) channel is not an essential determinant for its subcellular localization in cultured neurones,” *Genes to Cells*, vol. 10, no. 2, pp. 87–96, 2005.
- [150] Y.-Q. Cao and R. W. Tsien, “Different relationship of n-and p/q-type ca<sup>2+</sup> channels to channel-interacting slots in controlling neurotransmission at cultured hippocampal synapses,” *Journal of Neuroscience*, vol. 30, no. 13, pp. 4536–4546, 2010.
- [151] L. Matthias, G. R. Oliver, R. Satterfield, P. Travis, M. van den Maagdenberg Arn, K. Naomi, and S. M. Young Jr, “A novel region in the ca v 2.1  $\alpha$  1 subunit c-terminus regulates fast synaptic vesicle fusion and vesicle docking at the mammalian presynaptic active zone,” *eLife*, vol. 6, 2017.
- [152] J. S. Marvin, B. G. Borghuis, L. Tian, J. Cichon, M. T. Harnett, J. Akerboom, A. Gordus, S. L. Renninger, T.-W. Chen, C. I. Bargmann, *et al.*, “An optimized fluorescent probe for visualizing glutamate neurotransmission,” *Nature methods*, vol. 10, no. 2, p. 162, 2013.
- [153] Y. Hua, R. Sinha, C. S. Thiel, R. Schmidt, J. Hüve, H. Martens, S. W. Hell, A. Egner, and J. Klingauf, “A readily retrievable pool of synaptic vesicles,” *Nature neuroscience*, vol. 14, no. 7, p. 833, 2011.
- [154] N. Hozé and D. Holcman, “Statistical methods for large ensembles of super-resolution stochastic single particle trajectories in cell biology,” *Annual Review of Statistics and Its Application*, vol. 4, pp. 189–223, 2017.
- [155] A. Taslimi, J. D. Vrana, D. Chen, S. Borinskaya, B. J. Mayer, M. J. Kennedy, and C. L. Tucker, “An optimized optogenetic clustering tool for probing protein interaction and function,” *Nature communications*, vol. 5, p. 4925, 2014.
- [156] S. Rudolph, M.-C. Tsai, H. Von Gersdorff, and J. I. Wadiche, “The ubiquitous nature of multivesicular release,” *Trends in neurosciences*, vol. 38, no. 7, pp. 428–438, 2015.
- [157] T. Fernández-Alfonso and T. A. Ryan, “The kinetics of synaptic vesicle pool depletion at cns synaptic terminals,” *Neuron*, vol. 41, no. 6, pp. 943–953, 2004.
- [158] D. Freche, U. Pannasch, N. Rouach, and D. Holcman, “Synapse geometry and receptor dynamics modulate synaptic strength,” *PloS one*, vol. 6, no. 10, p. e25122, 2011.
- [159] T. Biederer, P. S. Kaeser, and T. A. Blanpied, “Transcellular nanoalignment of synaptic function,” *Neuron*, vol. 96, no. 3, pp. 680–696, 2017.
- [160] H. D. MacGillavry, Y. Song, S. Raghavachari, and T. A. Blanpied, “Nanoscale scaffolding domains within the postsynaptic density concentrate synaptic ampa receptors,” *Neuron*, vol. 78, no. 4, pp. 615–622, 2013.
- [161] D. Nair, E. Hosy, J. D. Petersen, A. Constals, G. Giannone, D. Choquet, and J.-B. Sibarita, “Super-resolution imaging reveals that ampa receptors inside synapses are dynamically organized in nanodomains regulated by psd95,” *Journal of Neuroscience*, vol. 33, no. 32, pp. 13204–13224, 2013.

- [162] E. M. Petrini, J. Lu, L. Cognet, B. Lounis, M. D. Ehlers, and D. Choquet, “Endocytic trafficking and recycling maintain a pool of mobile surface ampa receptors required for synaptic potentiation,” *Neuron*, vol. 63, no. 1, pp. 92–105, 2009.
- [163] B. Rácz, T. A. Blanpied, M. D. Ehlers, and R. J. Weinberg, “Lateral organization of endocytic machinery in dendritic spines,” *Nature neuroscience*, vol. 7, no. 9, p. 917, 2004.
- [164] T. Éltes, T. Kirizs, Z. Nusser, and N. Holderith, “Target cell type-dependent differences in ca<sup>2+</sup> channel function underlie distinct release probabilities at hippocampal glutamatergic terminals,” *Journal of Neuroscience*, vol. 37, no. 7, pp. 1910–1924, 2017.
- [165] E. L. Sylwestrak and A. Ghosh, “Elfn1 regulates target-specific release probability at ca1-interneuron synapses,” *Science*, vol. 338, no. 6106, pp. 536–540, 2012.
- [166] N. Gimber, G. Tadeus, T. Maritzen, J. Schmoranzner, and V. Haucke, “Diffusional spread and confinement of newly exocytosed synaptic vesicle proteins,” *Nature communications*, vol. 6, p. 8392, 2015.
- [167] M. Lübbert, R. O. Goral, R. Satterfield, T. Putzke, A. M. van den Maagdenberg, N. Kamasawa, and S. M. Young Jr, “A novel region in the cav2. 1  $\alpha$ 1 subunit c-terminus regulates fast synaptic vesicle fusion and vesicle docking at the mammalian presynaptic active zone,” *Elife*, vol. 6, p. e28412, 2017.
- [168] R. E. Dixon, C. Yuan, E. P. Cheng, M. F. Navedo, and L. F. Santana, “Ca<sup>2+</sup> signaling amplification by oligomerization of l-type cav1. 2 channels,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 5, pp. 1749–1754, 2012.
- [169] C. M. Moreno, R. E. Dixon, S. Tajada, C. Yuan, X. Opitz-Araya, M. D. Binder, and L. F. Santana, “Ca<sup>2+</sup> entry into neurons is facilitated by cooperative gating of clustered cav1. 3 channels,” *Elife*, vol. 5, p. e15744, 2016.
- [170] P. F. Vincent, Y. Bouleau, G. Charpentier, A. Emptoz, S. Safieddine, C. Petit, and D. Dulon, “Different cav1. 3 channel isoforms control distinct components of the synaptic vesicle cycle in auditory inner hair cells,” *Journal of Neuroscience*, vol. 37, no. 11, pp. 2960–2975, 2017.
- [171] O. Zhuchenko, J. Bailey, P. Bonnen, T. Ashizawa, D. W. Stockton, C. Amos, W. B. Dobyns, S. Subramony, H. Y. Zoghbi, and C. C. Lee, “Autosomal dominant cerebellar ataxia (sca6) associated with small polyglutamine expansions in the  $\alpha$  1a-voltage-dependent calcium channel,” *Nature genetics*, vol. 15, no. 1, p. 62, 1997.
- [172] P. Annibale, S. Vanni, M. Scarselli, U. Rothlisberger, and A. Radenovic, “Identification of clustering artifacts in photoactivated localization microscopy,” *Nature methods*, vol. 8, no. 7, p. 527, 2011.
- [173] T. Li, S. Kheifets, D. Medellin, and M. G. Raizen, “Measurement of the instantaneous velocity of a brownian particle,” *Science*, vol. 328, no. 5986, pp. 1673–1675, 2010.

- [174] R. E. Powers, S. Wang, T. Y. Liu, and T. A. Rapoport, “Reconstitution of the tubular endoplasmic reticulum network with purified components,” *Nature*, vol. 543, no. 7644, p. 257, 2017.
- [175] G. K. Voeltz, W. A. Prinz, Y. Shibata, J. M. Rist, and T. A. Rapoport, “A class of membrane proteins shaping the tubular endoplasmic reticulum,” *Cell*, vol. 124, no. 3, pp. 573–586, 2006.
- [176] C. A. Hübner and I. Kurth, “Membrane-shaping disorders: a common pathway in axon degeneration,” *Brain*, vol. 137, no. 12, pp. 3109–3121, 2014.
- [177] C. Blackstone, C. J. O’kane, and E. Reid, “Hereditary spastic paraplegias: membrane traffic and the motor pathway,” *Nature Reviews Neuroscience*, vol. 12, no. 1, p. 31, 2011.
- [178] M. J. Dayel, E. F. Hom, and A. Verkman, “Diffusion of green fluorescent protein in the aqueous-phase lumen of endoplasmic reticulum,” *Biophysical journal*, vol. 76, no. 5, pp. 2843–2851, 1999.
- [179] A. J. Boersma, I. S. Zuhorn, and B. Poolman, “A sensor for quantification of macromolecular crowding in living cells,” *Nature methods*, vol. 12, no. 3, p. 227, 2015.
- [180] H. Imamura, K. P. H. Nhat, H. Togawa, K. Saito, R. Iino, Y. Kato-Yamada, T. Nagai, and H. Noji, “Visualization of atp levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 37, pp. 15651–15656, 2009.
- [181] M. K. Cheezum, W. F. Walker, and W. H. Guilford, “Quantitative comparison of algorithms for tracking single fluorescent particles,” *Biophysical journal*, vol. 81, no. 4, pp. 2378–2388, 2001.
- [182] C. W. Lai, D. E. Aronson, and E. L. Snapp, “Bip availability distinguishes states of homeostasis and stress in the endoplasmic reticulum of living cells,” *Molecular biology of the cell*, vol. 21, no. 12, pp. 1909–1921, 2010.
- [183] E. L. Snapp, A. Sharma, J. Lippincott-Schwartz, and R. S. Hegde, “Monitoring chaperone engagement of substrates in the endoplasmic reticulum of live cells,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 17, pp. 6536–6541, 2006.
- [184] R. Tarjan, “Depth-first search and linear graph algorithms,” *SIAM journal on computing*, vol. 1, no. 2, pp. 146–160, 1972.
- [185] S. Nadeem and E. Maraj, “The mathematical analysis for peristaltic flow of nano fluid in a curved channel with compliant walls,” *Applied Nanoscience*, vol. 4, no. 1, pp. 85–92, 2014.
- [186] L. J. Young, F. Ströhl, and C. F. Kaminski, “A guide to structured illumination tiff microscopy at high speed with multiple colors,” *Journal of visualized experiments: JoVE*, no. 111, 2016.

- [187] P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, and M. G. Gustafsson, “Super-resolution video microscopy of live cells by structured illumination,” *Nature methods*, vol. 6, no. 5, p. 339, 2009.
- [188] K. Subramanian and T. Meyer, “Calcium-induced restructuring of nuclear envelope and endoplasmic reticulum calcium stores,” *Cell*, vol. 89, no. 6, pp. 963–971, 1997.
- [189] T. L. Giudice, F. Lombardi, F. M. Santorelli, T. Kawarai, and A. Orlandi, “Hereditary spastic paraplegia: clinical-genetic characteristics and evolving molecular mechanisms,” *Experimental neurology*, vol. 261, pp. 518–539, 2014.
- [190] R. Metzler and J. Klafter, “The random walk’s guide to anomalous diffusion: a fractional dynamics approach,” *Physics reports*, vol. 339, no. 1, pp. 1–77, 2000.
- [191] M. Kendall and A. Stuart, “The advanced theory of statistics: Inference and relationship, vol. 2,” *Griffin, London*, 1961.
- [192] 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.
- [193] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, “Scikit-learn: Machine learning in python,” *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.
- [194] I. Arganda-Carreras, R. Fernández-González, A. Muñoz-Barrutia, and C. Ortiz-De-Solorzano, “3d reconstruction of histological sections: Application to mammary gland tissue,” *Microscopy research and technique*, vol. 73, no. 11, pp. 1019–1029, 2010.
- [195] 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.
- [196] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.
- [197] C. Settembre, A. Fraldi, D. L. Medina, and A. Ballabio, “Signals from the lysosome: a control centre for cellular clearance and energy metabolism,” *Nature reviews Molecular cell biology*, vol. 14, no. 5, p. 283, 2013.
- [198] Y. Rabanal-Ruiz and V. Korolchuk, “mTORC1 and nutrient homeostasis: The central role of the lysosome,” *International journal of molecular sciences*, vol. 19, no. 3, p. 818, 2018.
- [199] E.-L. Eskelinen and P. Saftig, “Autophagy: a lysosomal degradation pathway with a central role in health and disease,” *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, vol. 1793, no. 4, pp. 664–673, 2009.
- [200] R. E. Lawrence and R. Zoncu, “The lysosome as a cellular centre for signalling, metabolism and quality control,” *Nat. Cell Biol*, 2019.

- [201] Q. Ba, G. Raghavan, K. Kiselyov, and G. Yang, “Whole-cell scale dynamic organization of lysosomes revealed by spatial statistical analysis,” *Cell reports*, vol. 23, no. 12, pp. 3591–3606, 2018.
- [202] A. M. Valm, S. Cohen, W. R. Legant, J. Melunis, U. Hershberg, E. Wait, A. R. Cohen, M. W. Davidson, E. Betzig, and J. Lippincott-Schwartz, “Applying systems-level spectral imaging and analysis to reveal the organelle interactome,” *Nature*, vol. 546, no. 7656, p. 162, 2017.
- [203] R. Matteoni and T. E. Kreis, “Translocation and clustering of endosomes and lysosomes depends on microtubules.,” *The Journal of cell biology*, vol. 105, no. 3, pp. 1253–1265, 1987.
- [204] J. Pu, C. M. Guardia, T. Keren-Kaplan, and J. S. Bonifacino, “Mechanisms and functions of lysosome positioning,” *J Cell Sci*, vol. 129, no. 23, pp. 4329–4339, 2016.
- [205] D. S. Schwarz and M. D. Blower, “The endoplasmic reticulum: structure, function and response to cellular signaling,” *Cellular and Molecular Life Sciences*, vol. 73, no. 1, pp. 79–94, 2016.
- [206] C. M. Waterman-Storer and E. Salmon, “Endoplasmic reticulum membrane tubules are distributed by microtubules in living cells using three distinct mechanisms,” *Current Biology*, vol. 8, no. 14, pp. 798–807, 1998.
- [207] B. Cabukusta and J. Neefjes, “Mechanisms of lysosomal positioning and movement,” *Traffic*, vol. 19, no. 10, pp. 761–769, 2018.
- [208] H. J. Bellen, C. Tong, and H. Tsuda, “100 years of drosophila research and its impact on vertebrate neuroscience: a history lesson for the future,” *Nature Reviews Neuroscience*, vol. 11, no. 7, p. 514, 2010.
- [209] L. Li, W.-C. Xiong, and L. Mei, “Neuromuscular junction formation, aging, and disorders,” *Annual review of physiology*, vol. 80, pp. 159–188, 2018.
- [210] M. A. Böhme, C. Beis, S. Reddy-Alla, E. Reynolds, M. M. Mampell, A. T. Grasskamp, J. Lützkendorf, D. D. Bergeron, J. H. Driller, H. Babikir, *et al.*, “Active zone scaffolds differentially accumulate unc13 isoforms to tune ca<sup>2+</sup> channel-vesicle coupling,” *Nature neuroscience*, vol. 19, no. 10, p. 1311, 2016.
- [211] J. Heck, P. Parutto, A. Ciuraszkiewicz, A. Bikbaev, R. Freund, J. Mitlöhner, M. Alonso, A. Fejtova, D. Holcman, and M. Heine, “Transient confinement of cav2.1 ca<sup>2+</sup>-channel splice variants shapes synaptic short-term plasticity,” *Neuron*, 2019.
- [212] S. Karlin and H. E. Taylor, *A second course in stochastic processes*. Elsevier, 1981.
- [213] G. Schwarz *et al.*, “Estimating the dimension of a model,” *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.







## RÉSUMÉ

---

Les trajectoires de molécules individuelles obtenues par microscopie super-résolution permettent de suivre des protéines avec une précision nanométrique dans des cellules vivantes. Dans cette thèse, j'ai étudié les régions de hautes densités présentes dans ces trajectoires, dont un modèle possible est celui des puits de potentiel. Pour les caractériser à partir de trajectoires, j'ai développé une nouvelle méthode hybride basée sur la densité de points et le champ de force local puis je l'ai comparé aux méthodes d'état de l'art. Ensuite, j'ai utilisé celle-ci pour caractériser les puits maintenant les canaux calciques Cav au niveau des zones actives des terminaux présynaptiques ce qui a permis de mieux comprendre le rôle des variantes d'épissage de ces canaux dans la transmission synaptique. Dans une autre étude, j'ai analysé des trajectoires de protéines résidant dans le lumen du Réticulum Endoplasmique (RE). J'ai créé une méthode pour reconstruire le réseau du RE à partir des trajectoires que j'ai utilisé pour caractériser le mouvement de ces molécules par un modèle de saut-diffusion qui a pour conséquence une meilleure redistribution du contenu luminal par rapport à un mouvement diffusif. Enfin, je discute d'autres analyses de trajectoires pour les interactions lysosome-ER, les canaux Cav à la jonction neuro-musculaire de la drosophile et les protéines composant le complexe NuRD.

## MOTS CLÉS

---

neurosciences, analyse de données, modélisation, trajectoires de molécules uniques, microscopie super-résolution.

## ABSTRACT

---

Single-Particle Trajectories (SPTs) obtained from super-resolution microscopy allow to track proteins with nanometer precision in living cells and are used in neuroscience and cellular biology. In this thesis, I was interested in the high-density nanodomains found in these trajectories that can be modeled as potential wells. To characterize them, I developed a new hybrid method based on the point density and local drift field and compared it to the other state-of-the-art methods. Then, I used it to identify transient potential wells in SPTs of voltage-gated calcium channels ( $Ca_V$ ) contributing to a better understanding of the role of the different  $Ca_V$  splice variants in synaptic transmission. In another study, I looked at SPTs from Endoplasmic Reticulum (ER) luminal resident proteins where I developed a method to reconstruct the network from trajectories and used it to characterize the luminal motion as a jump-diffusion process, which allows for a better redistribution of the luminal content than the previously assumed diffusive model. Finally, I discuss other analyses of motions for lysosome-ER interactions,  $Ca_V2.1$  channels at drosophila's neuromuscular junctions and the description of the motion of the constituent proteins of the NuRD chromatin remodeling complex.

## KEYWORDS

---

neurosciences, data analysis, modelling, single particle trajectories, super-resolution microscopy.