

OPEN

Biophysics of high density nanometer regions extracted from super-resolution single particle trajectories: application to voltage-gated calcium channels and phospholipids

P. Parutto¹, J. Heck², M. Heine² & D. Holcman^{1,3*}

The cellular membrane is very heterogenous and enriched with high-density regions forming microdomains, as revealed by single particle tracking experiments. However the organization of these regions remain unexplained. We determine here the biophysical properties of these regions, when described as a basin of attraction. We develop two methods to recover the dynamics and local potential wells (field of force and boundary). The first method is based on the local density of points distribution of trajectories, which differs inside and outside the wells. The second method focuses on recovering the drift field that is convergent inside wells and uses the transient field to determine the boundary. Finally, we apply these two methods to the distribution of trajectories recorded from voltage gated calcium channels and phospholipid anchored GFP in the cell membrane of hippocampal neurons and obtain the size and energy of high-density regions with a nanometer precision.

Single Particle trajectories (SPTs) obtained from super-resolution techniques such as sptPALM 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^{1,2}, interpret and extract relevant biophysical parameters such as flows and arrival time statistics between various subregions³⁻⁸ 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^{9,10}. Similarly, at the pre-synaptic terminal, voltage-gated calcium channels (CaV) can accumulate on membrane subregions, with a size of hundreds of nanometers¹¹. 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^{3,11,12}. However, the physical origin of these wells remains unclear because the length of classical electrostatic interactions is ten time shorter¹³ 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.

¹Group of Data Modeling and Computational Biology, IBENS-PSL, Ecole Normale Supérieure, 75005, Paris, France.

²Research Group Functional Neurobiology at the Institute of Developmental Biology and Neurobiology, Johannes Gutenberg University Mainz, Mainz, Germany. ³DAMPT and Churchill College, University Of Cambridge, Cambridge, CB30DS, United Kingdom. *email: david.holcman@ens.fr

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 (CAV2.2) and lipid anchored GFP (GPI-GFP) trajectories obtained from sptPALM or UPaint experiments.

Methods

Coarse-grained description of stochastic trajectories. In the Smoluchowski's limit of the Langevin equation^{14,15}, the position (t) of a stochastic molecule at time t can be described by

$$\dot{X} = \frac{F(X(t), t)}{\gamma} + \sqrt{2D} \dot{W}, \quad (1)$$

where $F(X, t)$ is a field of force, W is a white noise and γ is the friction coefficient¹⁴ 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^{3,16}

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

where $a(X)$ is the drift field and $B(X)$ the diffusion matrix. The effective diffusion tensor is given by $D(X) = \frac{1}{2}B(X)B^T(X)$ (T denotes the transposition)^{14,17}. The diffusion tensor accounts for impenetrable obstacles of various sizes. Note that the interpretation at the physical level of the stochastic Eq. (2) is from the Ito's sense and not Stratanovich or any other sense, because a physical process has to be non-anticipating¹⁷ (the future cannot interfere with the past).

Potential wells characteristics. The drift field $a(X)$ in Eq. 2 may represent a field force acting on the diffusing particle, that could be due to a potential well¹³. When the diffusion tensor (X) is locally constant and the coarse-grained drift field $a(X)$ is a gradient of a potential

$$a(X) = -\nabla U(X), \quad (3)$$

then the density of particles is given locally by the Boltzmann distribution¹⁸

$$\rho(X) = N_0 e^{-U(X)/D}, \quad (4)$$

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

$$U(x, y) = A \left[\left(\frac{x - \mu_x}{a} \right)^2 + \left(\frac{y - \mu_y}{b} \right)^2 \right], \quad (5)$$

where the center is (μ_x, μ_y) , A is the field amplitude A 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

$$\Gamma_{\mathcal{E}} = \{(x, y) | U(x, y) \leq \mathcal{E}\}. \quad (6)$$

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

$$U(X) = \begin{cases} A \left[\left(\frac{x - \mu_x}{a} \right)^2 + \left(\frac{y - \mu_y}{b} \right)^2 \right], & \text{if } X \in \Gamma_{\mathcal{E}}, \\ \mathcal{E} & \text{otherwise} \end{cases} \quad (7)$$

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

$$\nabla U(X) = -2A \begin{bmatrix} \frac{x - \mu_x}{a^2} \\ \frac{y - \mu_y}{b^2} \end{bmatrix}. \quad (8)$$

The goal of these section is to recover, from empirical single particle trajectories that consists of few successive points acquired with a sampling time Δt , the center (μ_x, μ_y) , the amplitude A and the size of each semi-axis a, b for the boundary \mathcal{E} .

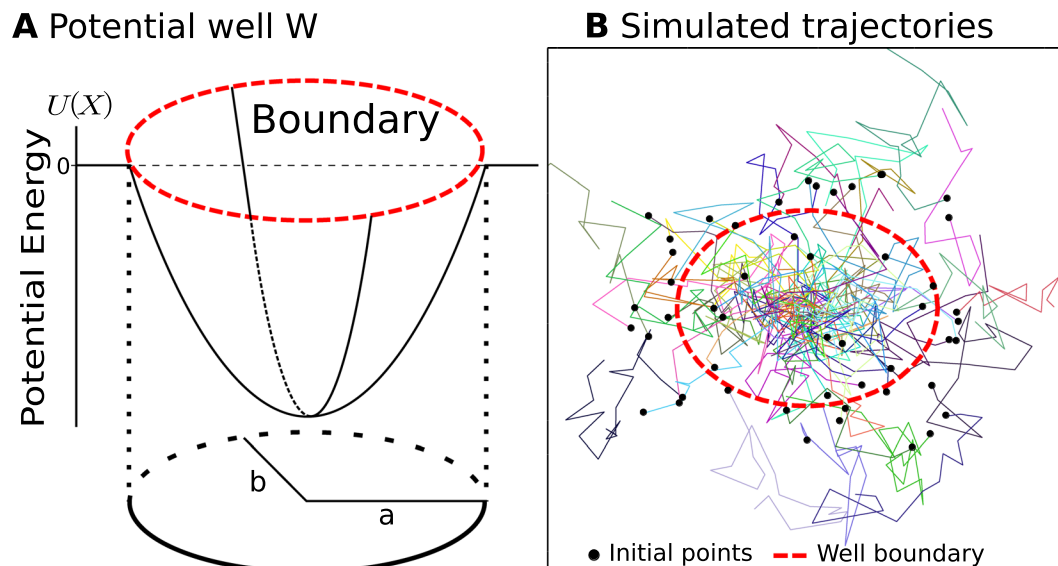


Figure 1. Numerical simulation scheme. **(A)** Model of a truncated potential well with two axes a, b and energy $U(X)$ with a boundary. **(B)** Trajectories generated using Eq. (9) where the initial points (black dots) can either be located inside or outside the boundary of the well (dashed red). Parameters: $D = 0.042 \mu\text{m}^2/\text{s}$, $\lambda_x = 10$, $\lambda_y = 17.78$.

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

$$\dot{X} = -\nabla U(X) + \sqrt{2D} \dot{W}, \quad (9)$$

where the potential U is defined in Eq. (5) (as presented in Fig. 1A), D is the diffusion coefficient and W is a white noise. To reproduce observed trajectories, we keep a fixed lapse time Δt between successive points and generated N trajectories $(X_1(0), \dots, X_N(K\Delta t))$ containing K points ($K = 20$), using the classical Euler's scheme (Fig. 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. 1B). 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.

Empirical estimators. The drift of the stochastic model from Eq. 2 can be recovered from SPTs acquired at any infinitesimal time step Δt by estimating the conditional moments of the trajectory displacements $\Delta X = X(t + \Delta t) - X(t)$ ^{14,16,19–21}

$$a(x) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[\Delta X(t) | X(t) = x]}{\Delta t}, \quad (10)$$

$$D(x) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[\Delta X(t)^T \Delta X(t) | X(t) = x]}{2\Delta t}. \quad (11)$$

The notation $\mathbb{E}[\cdot | X(t) = x]$ represents averaging over all trajectories that are passing at point x at time t . To estimate the local drift $a(X)$ and diffusion coefficients $D(X)$ at each point of the membrane and at a fixed time resolution Δt , we use a similar procedure as the one for the estimation of the density in section 3 based on a square grid. The points of trajectories are first grouped within a lattice of squared bins $S(x_k, \Delta x)$ centered at x_k and of width Δx and the drift and local diffusion coefficient are estimated for each bin.

When there are N trajectories $\{X_i(0), \dots, X_i(K\Delta t)\}$, with $i = 1 \dots N$ acquired with a sampling time Δt , the discretization of Eq. 10 for the drift $a(x_k) = (a_x(x_k), a_y(x_k))$ in a bin centered at position x_k is

$$\begin{aligned} a_x(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_t} \sum_{i=0, x_j(i\Delta t) \in S(x_k, r)}^{N_s-1} \left(\frac{x_j((i+1)\Delta t) - x_j(i\Delta t)}{\Delta t} \right) \\ a_y(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_t} \sum_{i=0, x_j(i\Delta t) \in S(x_k, r)}^{N_s-1} \left(\frac{y_j((i+1)\Delta t) - y_j(i\Delta t)}{\Delta t} \right), \end{aligned} \quad (12)$$

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

$$\begin{aligned} D_{xx}(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_i} \sum_{i=0, x_j(i\Delta t) \in S(x_k, r)}^{N_s-1} \frac{(x_j((i+1)\Delta t) - x_j(i\Delta t))^2}{2\Delta t} \\ D_{yy}(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_i} \sum_{i=0, x_j(i\Delta t) \in S(x_k, r)}^{N_s-1} \frac{(y_j((i+1)\Delta t) - y_j(i\Delta t))^2}{2\Delta t} \\ D_{xy}(x_k) &\approx \frac{1}{N_k} \sum_{j=1}^{N_i} \sum_{i=0, x_j(i\Delta t) \in S(x_k, r)}^{N_s-1} \frac{x_j((i+1)\Delta t) - x_j(i\Delta t)}{2\Delta t} (y_j((i+1)\Delta t) - y_j(i\Delta t)). \end{aligned} \quad (13)$$

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

Estimators for the elliptic boundary geometry. To identify parts of trajectories inside the well, we use the level line ensemble of the density distribution

$$\Gamma_\alpha = \{\mathbf{X}_i \text{ such that } \rho_e(x) > \alpha\}, \quad (14)$$

where ρ_e is the empirical point density, estimated over the bins of the square grid constructed from the ensemble of trajectories (Fig. 2B). The ensemble Γ_α contains all trajectory points falling into a bin, with a density greater than the density threshold α .

To recover the center of the distribution, we consider all points $\mathbf{X}_i = (x_i, y_i)$ located in Γ_α (Fig. 2C) and use the empirical estimators

$$\hat{\mu}_x^{(\alpha)} = \frac{1}{N_p} \sum_{\{k=1, \mathbf{X}_k \in \Gamma_\alpha\}}^{N_p} x_k, \quad \hat{\mu}_y^{(\alpha)} = \frac{1}{N_p} \sum_{\{k=1, \mathbf{X}_k \in \Gamma_\alpha\}}^{N_p} y_k, \quad (15)$$

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

$$U(\mathbf{X}) = (-\mu^{(\alpha)})^T C^{(\alpha)} (-\mu^{(\alpha)}), \quad (16)$$

we use the empirical estimators

$$\hat{C}_{ij}^{(\alpha)} = \frac{1}{N_p - 1} \sum_{\{k=1, \mathbf{X}_k \in \Gamma_\alpha\}}^{N_p} X_{i,k} X_{j,k}, \quad (17)$$

where $X_{i,k}$ is the i^{th} coordinates of \mathbf{X}_k (Fig. 2C).

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 λ and centered at μ (Eq. (9)): we derived in the SI that the drift term at position x and at resolution Δt

$$\tilde{a}_{\Delta t}(x) = -\frac{1 - e^{-\lambda\Delta t}}{\Delta t} (x - \mu). \quad (18)$$

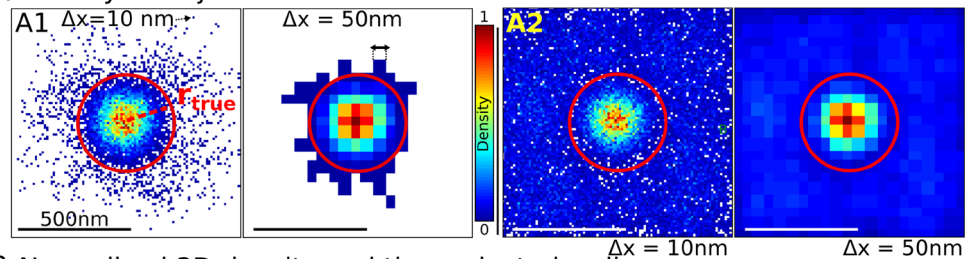
Hence, the first order moment at resolution Δt computed from the displacement $X(t + \Delta t) - X(t)$ from SPTs deviates from the expected drift. When $\lambda\Delta t$ is small, a first order Taylor expansion leads to the approximation

$$\tilde{a}_{\Delta t}(x) = a(x) \left(1 - \frac{1}{2} \lambda \Delta t \right) + o(\lambda^2 \Delta t) \quad (19)$$

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

Processing of CaV2.2 and GPI SPTs. For the experiments related to CaV2.2 data, we refer to²², while the experimental procedure of GPI-GFP data have been described for other molecules in²³. We will first isolate trajectories in non-overlapping time windows of 20s and apply the following procedure to each window. We will construct a square grid with bins size Δx around trajectories and collect the 5% highest density bins. For each of these selected bins, we will detect well as follows: we will first use 90% of the local point density (threshold $\alpha = 0.1$) to detect the center of the well from Eq. (15), then we will apply the procedure described in subsection 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 Eq. (13), estimated for all displacements with an initial points falling inside the well.

A Density of trajectories around a well at two resolutions



B Normalized 3D density and the projected well area

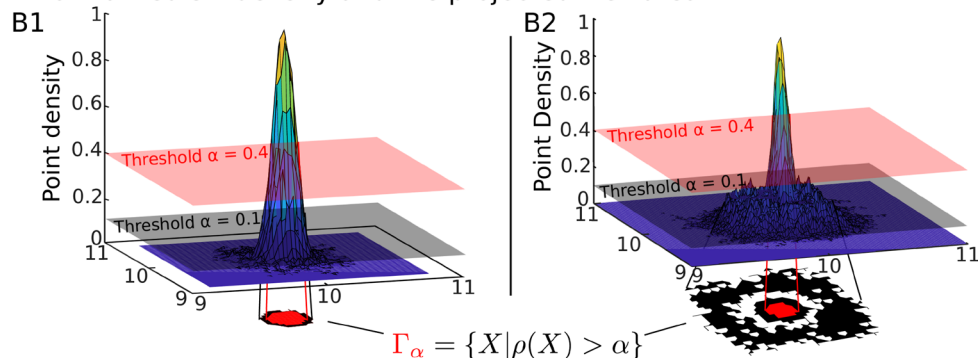
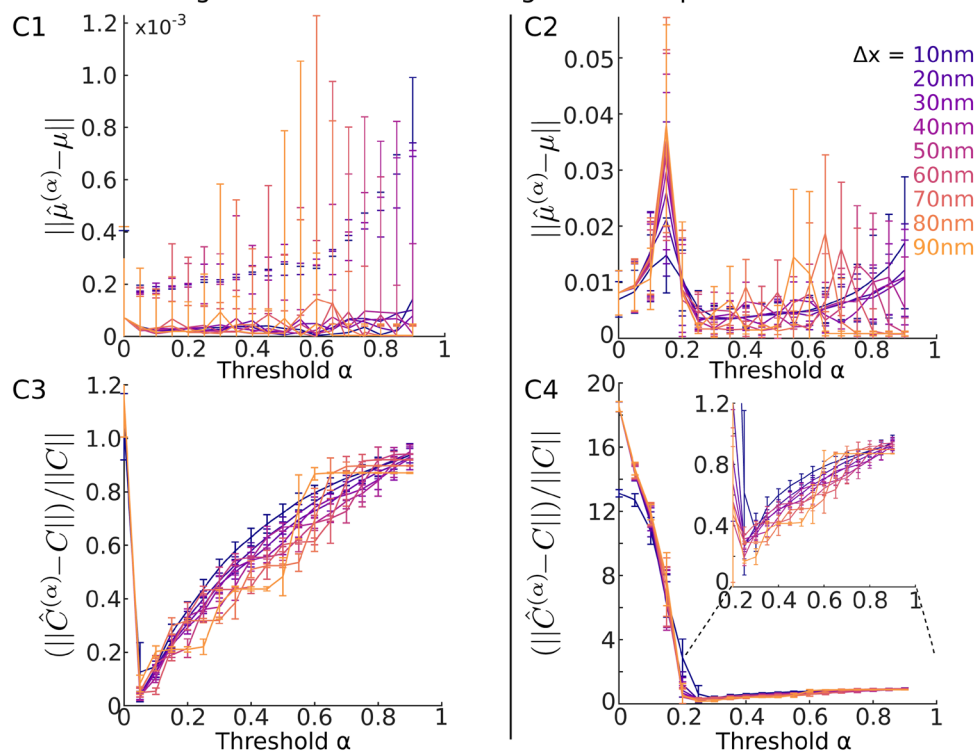
C Influence of grid size Δx on estimating the mean μ and covariance C 

Figure 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 ρ 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 Δx and threshold α on the well characteristics estimations. (C1, C3) (resp. C2, C4) panels are obtained by computing with the initial distribution described in A1 (resp. A2).

Results

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. 2A). We recall that inside a well given by Eq. 5, this is a Boltzmann distribution:

$$\rho(\mathbf{X}) = N_0 \exp \left[- \frac{A \left(\left(\frac{x - \mu_x}{a} \right)^2 + \left(\frac{y - \mu_y}{b} \right)^2 \right)}{D} \right], \quad (20)$$

where N_0 is a normalization coefficient while the other parameters are defined in subsection 2.2. Based on this distribution, we estimated the center using Eq. (15) and the covariance matrix from Eq. (17).

The accuracies of these estimators are analyzed by plotting the errors between the true and the estimated centers $|\hat{\mu}^\alpha - \mu|$ and between the covariance matrices $\|\hat{C}^\alpha - C\|$ (quadratic norm of the matrix) versus the parameter α , which represents the threshold of level line (Eq. 14, Fig. 2B,C) and various grid sizes (from $\Delta x = 10$ to 90 nm). When α 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 α (Fig. 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. 2C2). Below this value, points of the trajectories falling outside the well are also contained in the ensemble Γ_α , thus contaminating the error of the estimation. When the initial points fall inside the well only (Fig. 2C3), the ensemble Γ_0 contains external trajectories that perturb the estimation of the covariance matrix $C^{(\alpha)}$. However, as α increases, these external points disappear from Γ_α and the error becomes minimal at the value $\alpha_{opt} = 0.05$. When α 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. 2C4). As α increases, the estimator converges toward an optimal value $\alpha 0.25$ (75% of the points are used), which minimizes the matrix error. When α continues to increase, the error increases slowly (Fig. 2C4-inset), similar to the case of Fig. 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 (17), the covariance ratio

$$Cv(r) = \sqrt{\frac{C_{1,1}(r)}{C_{2,2}(r)}} \quad (21)$$

estimated over the trajectories located inside the annulus $(r, r + \Delta r)$ (Fig. 3A1). To compute $Cv(r)$ (Fig. 3A1,A2,B1,B2), we recall that the diagonal form of covariance matrix can be found from Eqs. (20) and (16):

$$C = \frac{D}{A} \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}. \quad (22)$$

Thus in that case, we expect that $Cv(r) = \sqrt{\frac{C_{1,1}(r)}{C_{2,2}(r)}} = \frac{a}{b}$, 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. 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 $\hat{\mu}$ (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. 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. 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 $\alpha = 2A/a^2$, $\beta = 2A/b^2$ 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. 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. 3B3), which allows us to recover \hat{r}_0 as the minimum point of the density curve (Fig. 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. 4A1,A2,B1,B2, for a ratio $a/b = 2$). Second, using this ratio, we introduced the elliptic distance $r_e(x, y) = \sqrt{x^2 + C_v(r^*)y^2}$, for a point $P = (x, y)$ from which we generated the point density distribution (Fig. 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. 4A4,B4). The small-semi axis is then given by $\hat{b} = \frac{\hat{a}}{\sqrt{C_v(r^*)}}$.

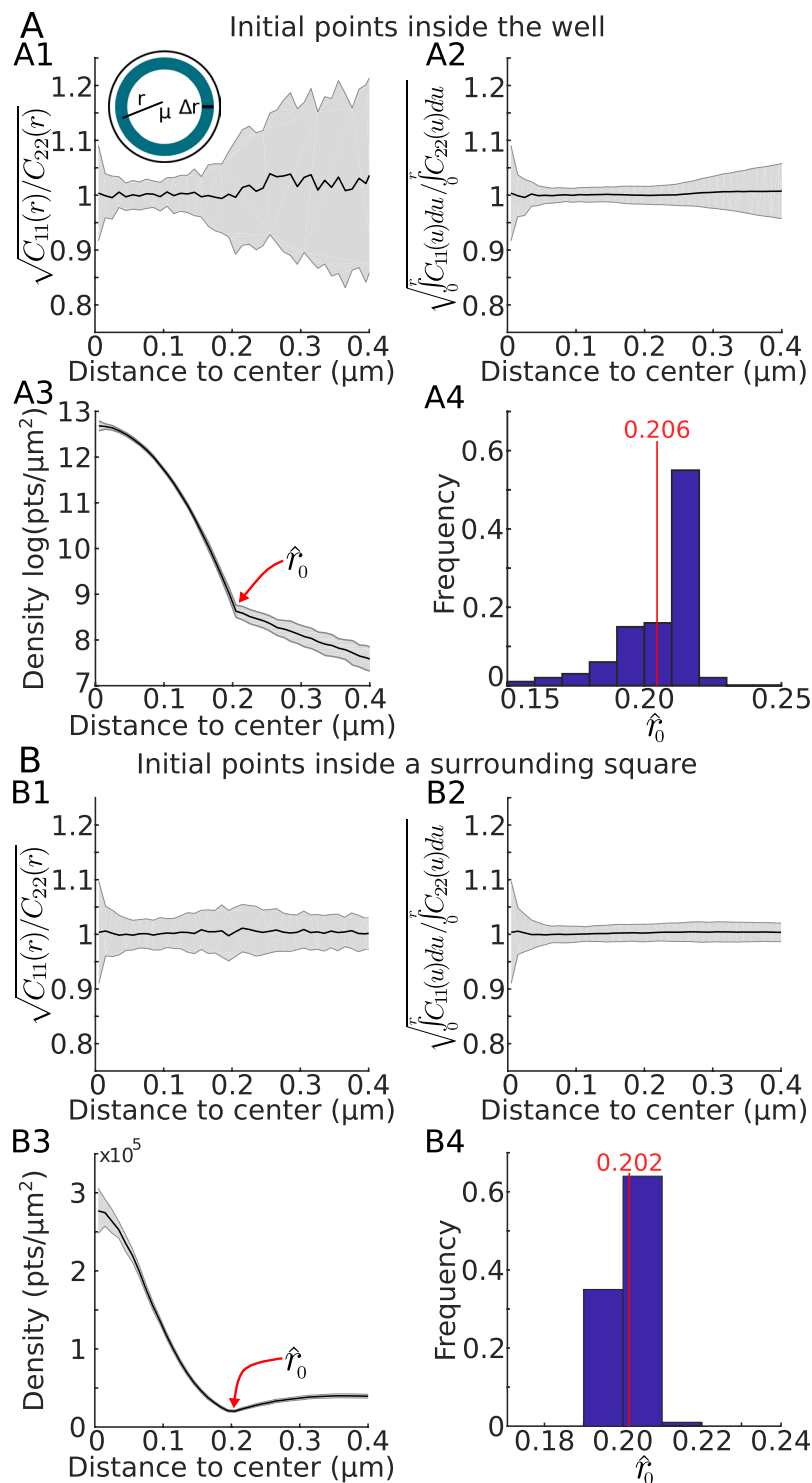


Figure 3. Estimating the potential well boundary. (A) Initial trajectory points are chosen inside the well while in (B) initial points are chosen inside a surrounding square. (A1,B1) Covariance ratio $C_v(r) = \sqrt{\frac{C_{11}(r)}{C_{22}(r)}}$ estimated in the annulus $r, r + \Delta r$. (A2,B2) Cumulative 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 \hat{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).

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. S1 and S2, we compare this density method with the MLE, which is classically used to recover the center and covariance, but not the boundary.

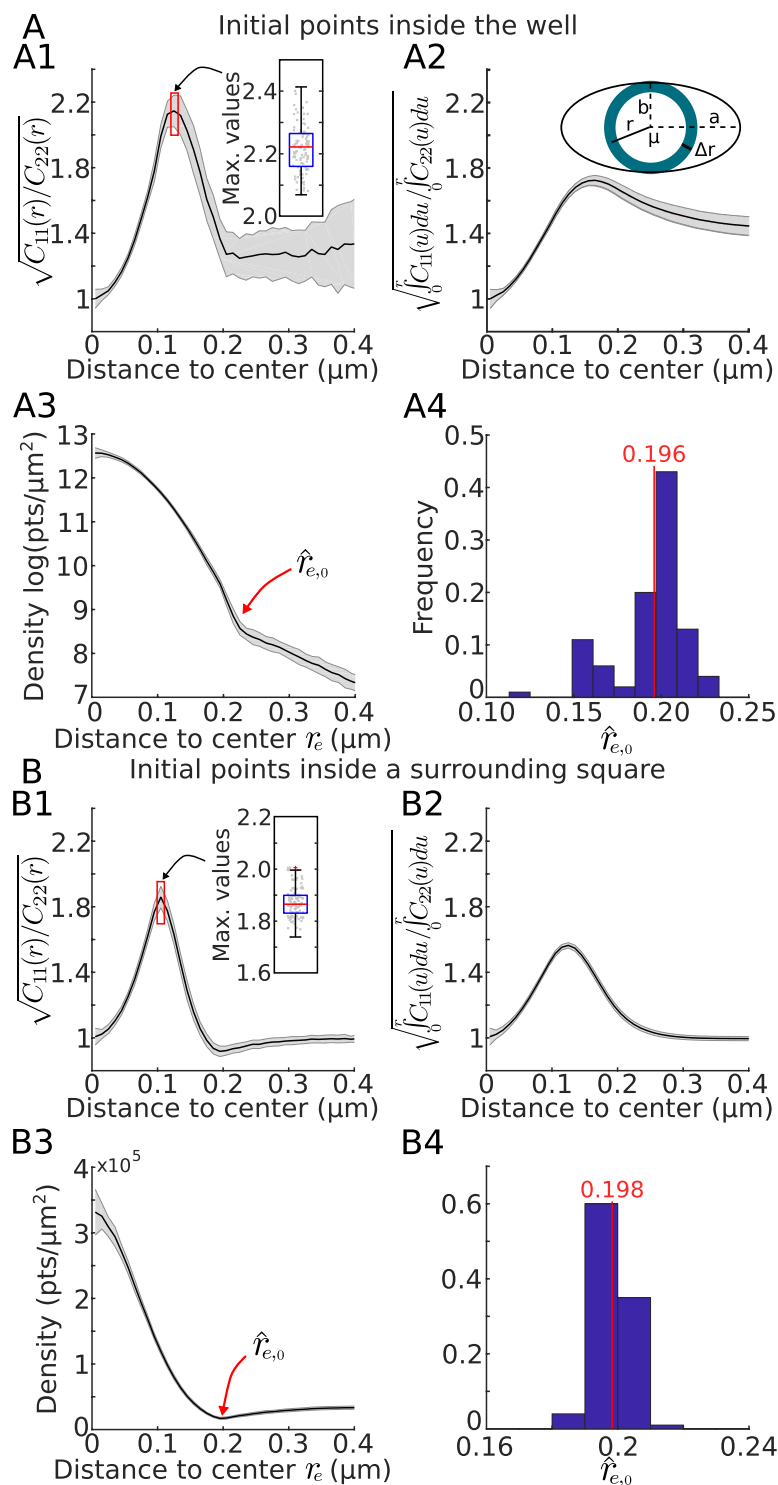
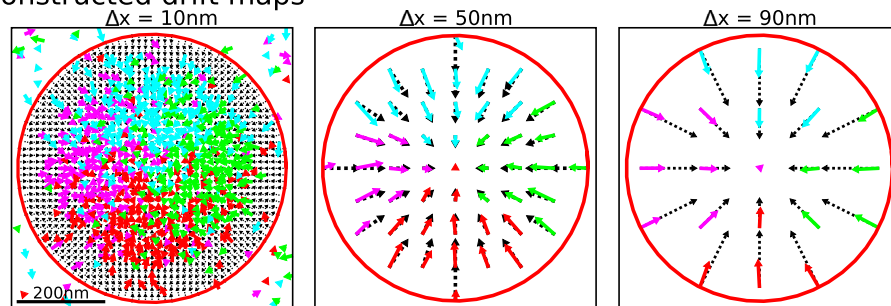


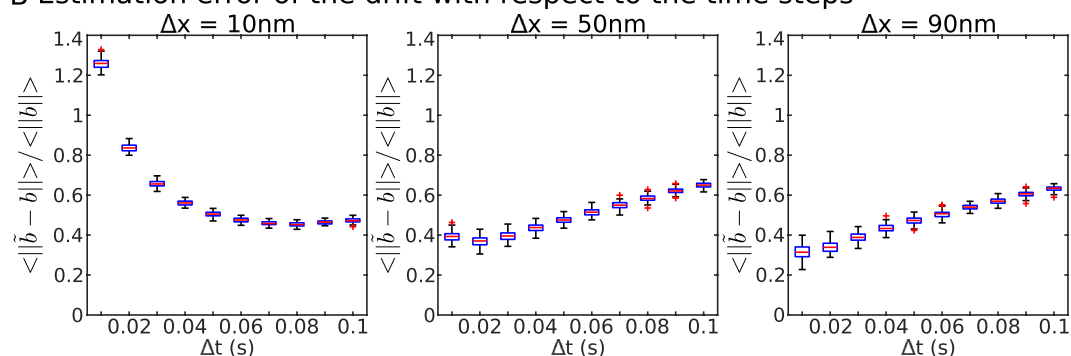
Figure 4. Estimating the potential well boundary for an elliptic well. **(A)** Initial trajectory points are chosen inside the well while in **(B)** Initial points are chosen inside a surrounding square. **(A1,B1)** Covariance ratio $C_v(r) = \sqrt{\frac{C_{11}(r)}{C_{22}(r)}}$ estimated in the annulus $r, r + \Delta r$. **(A2,B2)** Cumulative from A1,B1. **(A3,B3)** Point density (in log for A3) based on the modified distance $r_e(x, y) = \sqrt{x^2 + C_v(r^*)y^2}$ 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).

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 $X(t + \Delta t) - X(t)$. They allow to recover the drift of the vector field and reconstruct the center μ and the two axes a, b of the well boundary. This method is based on the least square quadratic error (LSQE),

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

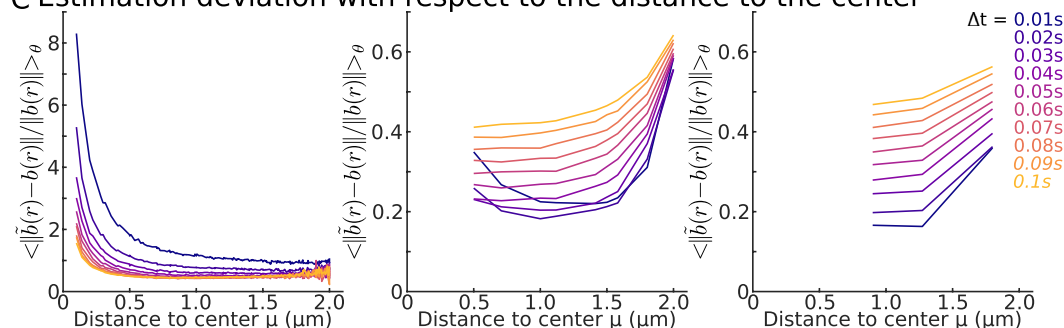


Figure 5. Vector field characteristics. (A) Recovering the local drift field inside a circular well for different grid sizes (10 nm, 50 nm, 90 nm) using numerical simulations with $\Delta t = 20$ ms, with the constraints that at least 10 points falls inside a bin. (B) Error between the true and observed fields averaged over all the square bins inside the well vs the time step Δt . (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 timesteps (see color code).

$$Err_b(\mu_x, \mu_y, \lambda_x, \lambda_y) = \sum_{i=1}^N \|\!-\nabla U(X_i) - (X_i)\|^2 \quad (23)$$

between the empirical drift and the parabolic well U , defined in Eq. 8, with $\lambda_x = -\frac{2A}{a^2}$, $\lambda_y = -\frac{2A}{b^2}$.

Estimating the center and the field coefficients of the potential well. The center μ and the coefficients λ of the potential well can be obtained explicitly from Eqs. 24 and 20 (SI). We compare in Fig. 5A, the reconstructed and the true drift value based on Eq. (12) for various grid sizes. At this stage, we considered the trajectory to be known and estimated the drift only for bins that are falling inside the well. The error of the norm $\frac{\langle \|\tilde{b} - b\| \rangle}{\langle \|b\| \rangle}$ is plotted in Fig. 5B for multiple time steps Δt and for three grid sizes $\Delta x = 10, 50$ and 90 nm. Having both a small grid size and time step Δt produces a large error that quickly decreases with increasing the time step Δt . Interestingly, for a large grid size, we found a slow increase of the error when increasing Δt . To better understand which parts of the field contributed the most to the error, we plotted *Err* versus the distance to the center (Fig. 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. S3 for recovering a drift at a different time resolution Δt and also with some restrictions on the trajectories for which the end point remained inside the well (Fig. S4).

Finally, to estimate the boundary of the well from the drift distribution (Fig. 6A), we plotted the drift amplitude versus the distance to the well center (Fig. 6B, blue crosses representing the drift amplitude in individual bins). From the distribution and the average (Fig. 6B lower panel), we could recover the location of the boundary

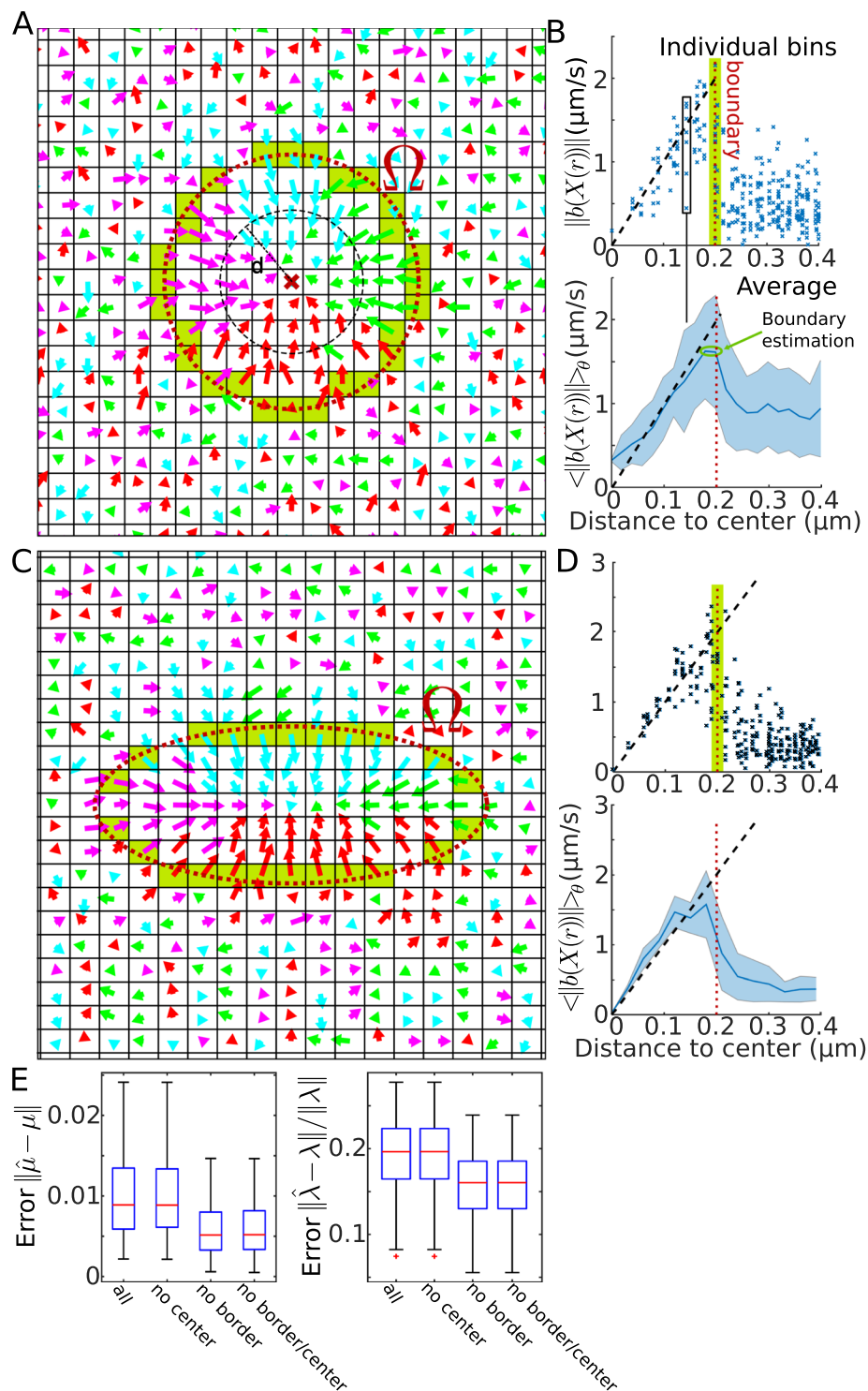


Figure 6. Recovery of the well depth from trajectories. **(A)** Vector field recovered from trajectories including the field generated by the Brownian dynamics outside the well. At the boundary of the well, there are mixed displacements (OU and Brownian), marked by the green band. **(B)** Upper: Drift 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 $\|\hat{\mu} - \mu\| = \sqrt{\mu_x^2 + \mu_y^2}$ and the eigenvalue $\|\hat{\lambda} - \lambda\|/\|\lambda\| = \sqrt{\lambda_x^2 + \lambda_y^2}$ of the reconstructed ellipse in 4 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.

Parameter	GPI dataset	CaV dataset
Δt (exp)	20 ms	33 ms
Δx	40 nm	30 nm
r_{min}	30 nm	20 nm
r_{max}	300 nm	400 nm
Δr	20 (m)	10 nm

Table 1. Parameters used for CaV and GPI analysis.

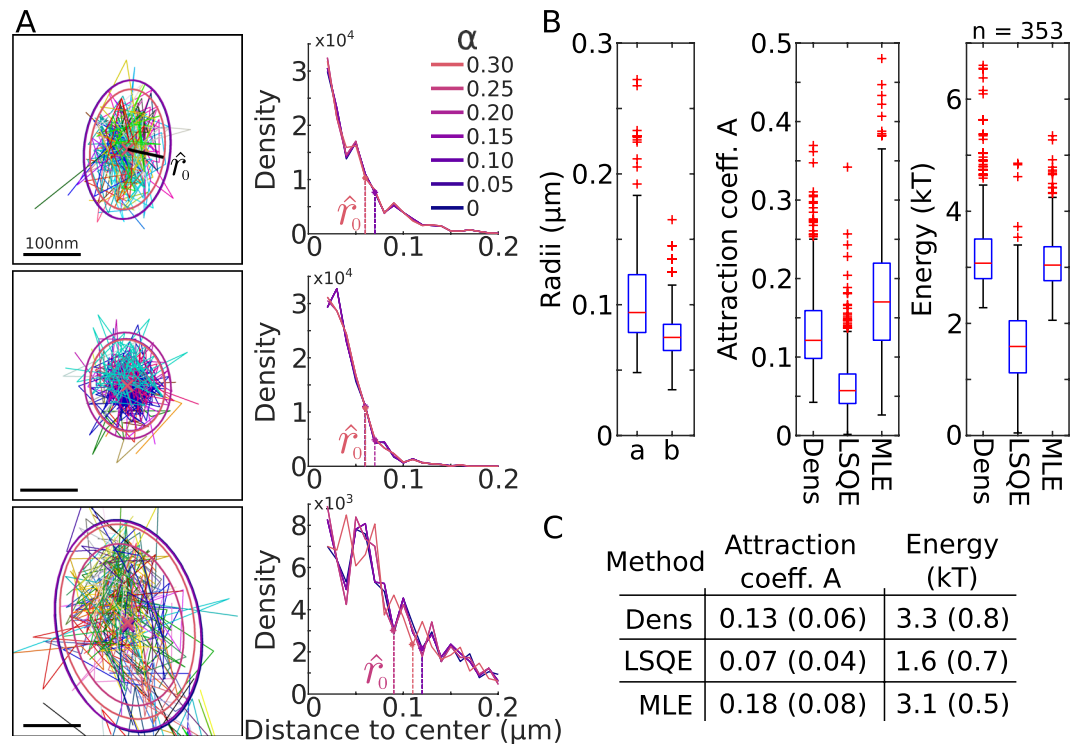


Figure 7. Reconstruction of wells associated to CaV2.2. (A) 3 examples of potential wells (left) obtained from the density analysis on SPTs. The boundary of the well are estimated from various level of density α (right). The estimated radius \hat{r}_0 is obtained using a threshold $T = 4\%$ on the density. (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, LSQ and MLE methods. (C) Summary of mean and SD for the coefficient A and the energy.

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. 6C,D) and recover the boundary after we used the covariance ratio C_V (Eq. 21) 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 μ , and eigenvalues λ_x and λ_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. 6E). We found that the latter case produces the best estimation.

Interpretation high-density regions for CaV 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 CaV2.1 isoform¹¹. We focus here on the isoform CaV2.2 (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.

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 1. We report in Fig. 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

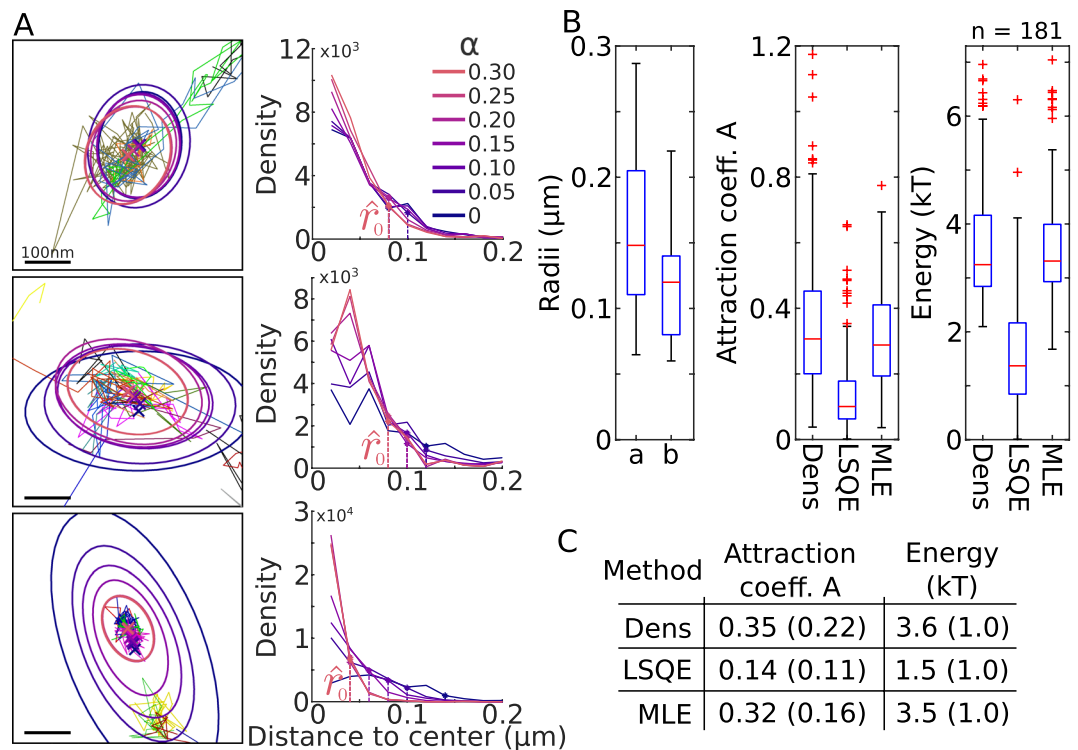


Figure 8. Reconstructed wells associated to GPI-anchored GFP. (A) 3 examples of potential wells (left) obtained from the density analysis on SPTs. The boundary of the well are estimated from various level of density α (right). The estimated radius \hat{r}_0 is computed using a threshold $T=4\%$ on the 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.

3.3 kT estimated for the density method. These results differ from the CaV2.1 isoform¹¹. Note that the distribution of energy varied with the statistical method (Fig. 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.

We also apply our statistical methods to the case of GFP linked to the outer leaflet of the membrane by a GPI-anchor (Fig. 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 ± 1.0 and 3.5 ± 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. rather

Summary and Discussion

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. 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. S1, S2). The second method is based on estimating the vector field distribution at a given bin resolution Δx . 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 trajectories²¹, which provide robustness, reducing the effect of tracking errors or localization noise^{24,25}.

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 Δt and spatial Δx steps influenced the recovery process

as shown in Figs. 3, 4 and 6. The advantage of the first method is that we do not need to introduce an artificial grid of size Δx which is a serious constraint in the second method as the bins size defines the resolution to recover the well and its boundary.

High-density regions contained calcium-voltage channels and GPI SPTs data. We recall that high-density regions revealed by SPTs are not necessarily due physical forces and potential wells³. However, for potential well, the geometry (center, curvature and boundary) can be recovered from our two methods. We applied them to CaV 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 CaV (Fig. 7A) are characterized by two main axes with a length $a = 104 \pm 36$ and $b = 77 \pm 20$ nm (Fig. 7B), with a mean energy of 3.3 ± 0.8 kT (density method, Fig. 7C). We note that hydrogen bonds between calcium channels and phospholipid molecules could participate in the formation of the wells²⁶. Surprisingly, we did not expect to find high-density regions for GPI-GFP, but we found several (Fig. 8A) that were characterized by average semi-axes lengths $a = 158 \pm 57$ and $b = 118 \pm 39$ nm (Fig. 8B), with a mean energy $E = 3.6 \pm 1.0$ kT (density method, Fig. 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 CaV.

Although the interpretation of high-density regions as potential wells for AMPA receptors was first anticipated in²⁷ and discovered in³, the nature of these wells and others, remains unclear¹³. Potential wells were found for membrane proteins such CaV¹¹, GAG^{12,28} and recently for G-protein²⁹. They could be generated by protein clusters, membrane cusps at vesicle fusion points or membrane-membrane contact at location of organelle interactions³⁰. 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^{12,28}: 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 spines¹⁶ or trapping proteins in the endoplasmic reticulum¹¹. 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.

Received: 5 July 2019; Accepted: 7 November 2019;

Published online: 11 December 2019

References

- Caridi, C. P. *et al.* Nuclear f-actin and myosins drive relocalization of heterochromatic breaks. *Nature* **559**(7712), 54 (2018).
- Burov, S. *et al.* Distribution of directional change as a signature of complex dynamics. *Proceedings of the National Academy of Sciences* **110**(49), 19689–19694 (2013).
- Hoze, N. *et al.* Heterogeneity of ampa receptor trafficking and molecular interactions revealed by superresolution analysis of live cell imaging. *Proceedings of the National Academy of Sciences* **109**(42), 17052–17057 (2012).
- Vestergaard, C. L. *et al.* Estimation of motility parameters from trajectory data. *The European Physical Journal Special Topics* **224**(7), 1151–1168 (2015).
- Vestergaard, C. L., Blainey, P. C. & Flyvbjerg, H. Optimal estimation of diffusion coefficients from single-particle trajectories. *Physical Review E* **89**(2), 022726 (2014).
- Calderon, C. P., Weiss, L. E. & Moerner, W. Robust hypothesis tests for detecting statistical evidence of two-dimensional and three-dimensional interactions in single-molecule measurements. *Physical Review E* **89**(5), 052705 (2014).
- Bronshtein, I. *et al.* Loss of lamin a function increases chromatin dynamics in the nuclear interior. *Nature communications* **6**, 8044 (2015).
- Briane, V., Kervrann, C. & Vimond, M. Statistical analysis of particle trajectories in living cells. *Physical Review E* **97**(6), 062121 (2018).
- Huganir, R. & Nicoll, R. Ampars and synaptic plasticity: the last 25 years. *Neuron* **80**(3), 704–717 (2013).
- Nicoll, R. A. A brief history of long-term potentiation. *Neuron* **93**(2), 281–290 (2017).
- Heck, J. *et al.* Transient confinement of cav2. 1 ca2+-channel splice variants shapes synaptic short-term plasticity. *Neuron* (2019).
- Hoze, N. & Holcman, D. 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).
- Holcman, D. Unraveling novel features hidden in superresolution microscopy data. *Communicative & integrative biology* **6**(3), e23893 (2013).
- Schuss, Z. *Theory and Applications of Stochastic Processes: An Analytical Approach* (Springer New York, 2010).
- Holcman, D. & Schuss, Z. 100 years after smoluchowski: stochastic processes in cell biology. *Journal of Physics A: Mathematical and Theoretical* **50**(9), 093002 (2017).
- Hoze, N. & Holcman, D. Residence times of receptors in dendritic spines analyzed by stochastic simulations in empirical domains. *Bio-physical journal* **107**(12), 3008–3017 (2014).
- Schuss, Z. *Diffusion and stochastic processes: an analytical approach* (Springer New York, 2010).
- Holcman, D., Hoze, N. & Schuss, Z. Analysis and interpretation of superresolution single-particle trajectories. *Biophysical journal* **109**(9), 1761–1771 (2015).
- Friedrich, R. & Peinke, J. Description of a turbulent cascade by a fokker-planck equation. *Physical Review Letters* **78**(5), 863 (1997).
- Siegert, S., Friedrich, R. & Peinke, J. Analysis of data sets of stochastic systems. *Physics Letters A* **243**(5–6), 275–280 (1998).
- Hozé, N. & Holcman, D. Statistical methods for large ensembles of super-resolution stochastic single particle trajectories in cell biology. *Annual Review of Statistics and Its Application* **4**, 189–223 (2017).
- Schneider, R. *et al.* Mobility of calcium channels in the presynaptic membrane. *Neuron* **86**(3), 672–679 (2015).
- Giannone, G. *et al.* Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Biophysical journal* **99**(4), 1303–1310 (2010).
- Saxton, M. J. Single-particle tracking: connecting the dots. *Nature methods* **5**(8), 671 (2008).
- Chenouard, N. *et al.* Objective comparison of particle tracking methods. *Nature methods* **11**(3), 281 (2014).
- Zhao, G.-J. & Han, K.-L. Hydrogen bonding in the electronic excited state. *Accounts of chemical research* **45**(3), 404–413 (2011).

27. Holcman, D. & Schuss, Z. Escape through a small opening: receptor trafficking in a synaptic membrane. *Journal of Statistical Physics* **117**(5–6), 975–1014 (2004).
28. Floderer, C. *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* **8**(1), 16283 (2018).
29. Sungkaworn, T. *et al.* Single-molecule imaging reveals receptor-g protein interactions at cell surface hot spots. *Nature* **550**(7677), 543 (2017).
30. Cohen, S., Valm, A. M. & Lippincott-Schwartz, J. Interacting organelles. *Current opinion in cell biology* **53**, 84–91 (2018).

Author contributions

P.P. and D.H. elaborate the research, theory, simulations wrote the manuscript. J.H. and M.H. designed the experiments and collected super-resolution microscopy data.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-55124-8>.

Correspondence and requests for materials should be addressed to D.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019