Editors' Note: Following the in-principle acceptance of this Stage 1 Registered Report, the timelines have been updated as follows: the Competition will run from December 1, 2023, to July 15, 2024; the "Development Phase" will run from December 1, 2023, to April 30, 2024; and the "Challenge Phase" will run from July 1, 2024, to July 15, 2024.

Quantitative evaluation of methods to analyze motion changes in single-particle experiments

Gorka Muñoz-Gil,[1] Harshith Bachimanchi,[2] Jesús Pineda,[2] Benjamin Midtvedt,[2] Maciej Lewenstein,[3,4] Ralf Metzler,[5,6] Diego Krapf,[7] Giovanni Volpe,[2] and Carlo Manzo[8,9,*]

[1] *Institute for Theoretical Physics, University of Innsbruck, Technikerstr. 21a, A-6020 Innsbruck, Austria*
[2] *Department of Physics, University of Gothenburg, Origovägen 6B, SE-41296 Gothenburg, Sweden*
[3] *ICFO – Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*
[4] *ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*
[5] *Institute for Physics & Astronomy, University of Potsdam, Karl-Liebknecht-Str 24/25, D-14476 Potsdam-Golm, Germany*
[6] *Asia Pacific Centre for Theoretical Physics, Pohang 37673, Republic of Korea*
[7] *Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, Colorado 80523, USA*
[8] *Facultat de Ciències, Tecnologia i Enginyeries, Universitat de Vic – Universitat Central de Catalunya (UVic-UCC), C. de la Laura,13, 08500 Vic, Barcelona, Spain*
[9] *Institut de Recerca i Innovació en Ciències de la Vida i de la Salut a la Catalunya Central (IRIS-CC), 08500 Vic, Barcelona, Spain*

(Dated: October 3, 2023)

The analysis of live-cell single-molecule imaging experiments can reveal valuable information about the heterogeneity of transport processes and interactions between cell components. These characteristics are seen as motion changes in the particle trajectories. Despite the existence of multiple approaches to carry out this type of analysis, no objective assessment of these methods has been performed so far. Here, we have designed a competition to characterize and rank the performance of these methods when analyzing the dynamic behavior of single molecules. To run this competition, we have implemented a software library to simulate realistic data corresponding to widespread diffusion and interaction models, both in the form of trajectories and videos obtained in typical experimental conditions. The competition will constitute the first assessment of these methods, provide insights into the current limits of the field, foster the development of new approaches, and guide researchers to identify optimal tools for analyzing their experiments.

INTRODUCTION

Physiological processes occurring in living cells rely on encounters and interactions between molecules. Archetypal examples include gene regulation, transduction of biological signals, and protein delivery to specific locations. All these processes involve the active or passive transport of biomolecules in highly complex, time-varying, and far-from-equilibrium environments, such as the cell membrane (Fig. 1(a)). One of the most powerful tools to study these transport phenomena is the combination of live-cell single-molecule imaging with single-particle tracking [1, 2] because it can provide the time and location when and where single events take place (Fig. 1(b-c)). Alternative ensemble methods (e.g., fluorescence correlation spectroscopy or fluorescence recovery after photobleaching [3]) usually provide limited information because they lose track of crucial information when averaging out spatial and temporal fluctuations.

Methods for single-molecule imaging and single-particle tracking have seen tremendous progress in the last decade, in terms of both experimental acquisition and data analysis [1, 2, 4, 5]. The abundance of experimental single-particle trajectories, encompassing molecules, protein complexes, vesicles, and organelles, has led to the development of numerous methods dedicated to the reliable detection of changes in their motion patterns (as summarized in Table I). These changes serve as valuable indicators for the occurrence of interactions within the system. For instance, diffusing particles may exhibit variations in diffusion coefficients (due to processes like dimerization, ligand binding, or conformational changes) or shifts in their mode of motion (attributed to transient immobilization or confinement at specific scaffolding sites) (Fig. 1(a)) [6]. These interactions can also result in deviations from standard Brownian motion, as characterized by Einstein's free diffusion model, which includes a linear mean-squared displacement (MSD) and a Gaussian distribution of displacements [7]. This is the case, e.g., of spatiotemporal heterogeneities producing transient subdiffusion at specific timescales [8–19]. Other mechanisms can instead produce asymptotically-anomalous diffusion [2, 20–22]. Anomalous diffusion compatible with models such as fractional Brownian motion [23–28], continuous-time random walk [29, 30], scaled Brownian motion [31], and Lévy walk [32] has been observed for telomers, macromolecular complexes, proteins, and organelles in living cells. Several approaches have been recently proposed to

* carlo.manzo@uvic.cat

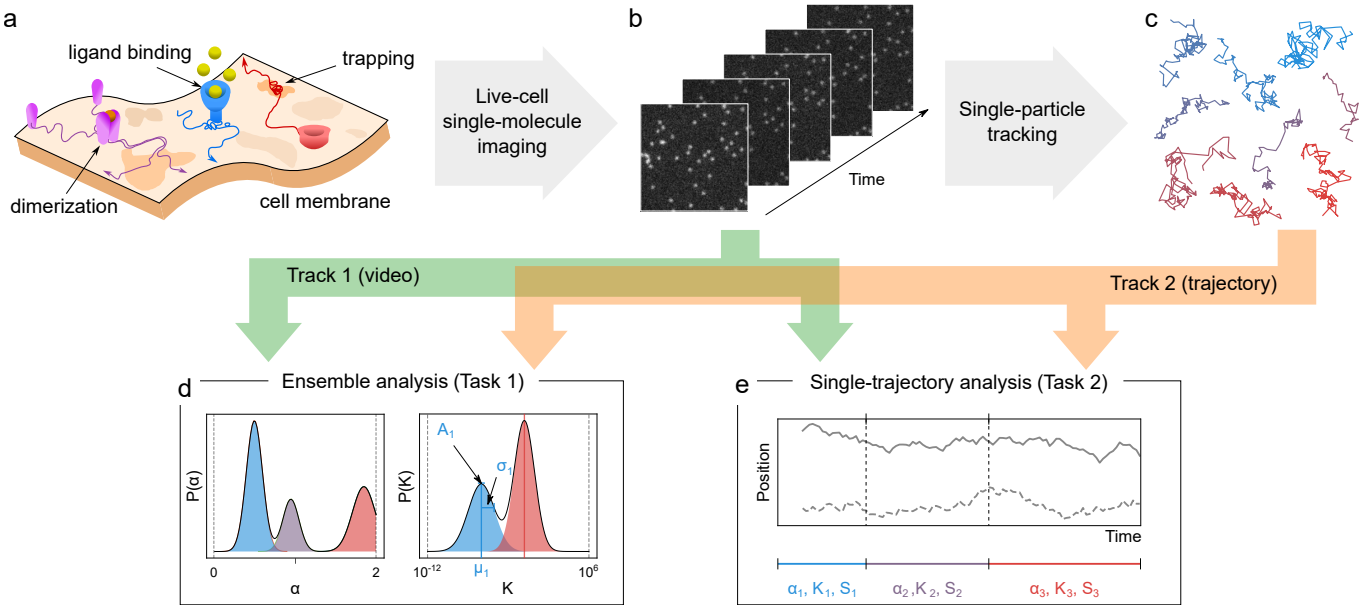

FIG. 1. **Rationale for the challenge organization. a**, The interactions of biomolecules in complex environments, such as the cell membrane, regulate physiological processes in living systems. These interactions produce changes in molecular motion that can be used as a proxy to measure interaction parameters. **b-c**, Time-lapse single-molecule imaging allows us to visualize these processes with high spatiotemporal resolution (**b**) and, in combination with single-particle tracking methods, provide trajectories of individual molecules (**c**). **d-e**, Analytical methods can be applied to imaging data, either raw (**b**) or processed in the form of trajectories (**c**), to infer interaction kinetics and quantify their dynamic properties at the ensemble (e.g., probability distributions, **d**) or single-trajectory level (e.g., changepoints, **e**).

detect and quantify these behaviors [33, 34], also involving machine-learning techniques [35–41].

To gain insights into the performance of methods to detect anomalous diffusion from individual trajectories, in 2021 we successfully ran the 1st AnDi Challenge [42]. The discussion that developed between members of diverse research communities working on biology, microscopy, single-particle tracking, and anomalous diffusion (including experimentalists, theoreticians, data analysts, and computer scientists) emphasized the necessity for deeper insights into biologically relevant phenomena. First, there is a need to evaluate methods to determine the switch between different diffusive behaviors, as often observed in experiments. Second, it is necessary to assess the methods' crosstalk in detecting inherent anomalous diffusion from nonlinearity in the MSD due to motion constraints or heterogeneity. Third, there is a need to determine whether the bottleneck of the analysis process is at the level of the analysis of the single trajectories or associated with their extraction from the experimental videos. These needs shaped the design of the 2nd AnDi Challenge, defining its scope with a focus on characterizing and ranking the performance of methods that analyze changes of dynamic behavior. While we have retained the name of the 1st AnDi Challenge to build upon its already-established community, we would like to emphasize that the main focus of this 2nd AnDi Challenge is on revealing heterogeneity rather than anomalous diffusion. In the simulated datasets, anomalous diffusion will either emerge from heterogeneity itself or be intentionally introduced for evaluation purposes.

There is a multitude of methods that have been designed to identify and characterize heterogeneous diffusion (Table I). They can be classified based on the heterogeneity they aim to identify or on the kind of analysis they perform. We consider three heterogeneity classes that these methods aim to identify: (i) changes in the value of the diffusion coefficient D; (ii) changes in the anomalous diffusion exponent α (often classified as subdiffusion, diffusion, or superdiffusion); and (iii) changes in the phenomenological behavior associated with interactions with the environment (often classified as immobilization, confinement, (free) diffusion, and directed motion). While changes in the diffusion coefficient and in the phenomenological behavior have been widely reported, the exploration of changes in the anomalous diffusion exponent is a more recent development [43–46], which is attracting increasing interest also from the theoretical point of view [47–50]. The introduction of new methods for data analysis, as promoted by the Challenge, could push the performance for detecting subtle changes in these diffusion properties in systems where they have so far been overlooked. Along this line, it must be pointed out that the traditional analysis based on the calculation of the scaling exponent of the mean-squared displacement (MSD) can create some ambiguity between the last two classes. Just to provide an example, a particle performing Brownian diffusion in a confined region has an

exponent $\alpha = 1$ in terms of the generating motion, but its MSD features a horizontal asymptote at long times, corresponding to $\alpha = 0$. In the following, we will refer to the exponent α as the characteristic feature of the generating motion.

From the analysis point of view, we can identify two classes of methods: (i) ensemble methods, meant to determine characteristic features out of an ensemble of trajectories (Fig. 1(d)) and (ii) single-trajectory methods, meant to identify changepoint (CP) locations through trajectory segmentation (Fig. 1(e)). While most available methods rely on the analysis of trajectories obtained from video processing [51], recent advances in computer vision have led to methods capable of directly extracting information from raw movies without requiring the explicit extraction of trajectories [52, 53]. Each method has its own set of advantages and disadvantages, and its performance may depend on the specific problem under consideration. However, there is currently no universally accepted gold standard for determining which method to use to address each specific problem.

To cater to these more advanced needs, we decided to run an open competition as the 2$^{\text{nd}}$ Anomalous Diffusion (AnDi) Challenge. The rationale described above shaped the scope of the challenge, defining the choice of the datasets and the design of the tasks. To rely on objective ground truth, we will assess the methods' performance on simulated datasets inspired by models of diffusion and interactions documented in biological systems. Datasets will describe particles undergoing fractional Brownian motion (FBM, [54]) with piecewise-constant parameters. FBM-type motion has been widely observed in biological systems by means of microrheology, a technique that uses large tracer particles as probes to study the properties of the environment [55]. Anomalous diffusion compatible with FBM has also been reported for telomers and macromolecular complexes in living cells [20, 23–28, 56]. Beyond this evidence, in the context of the Challenge, FBM serves as a tool to enable the tuning of diffusion parameters. The combination of parameter values and interaction models might produce situations that do not correspond to previously documented biological scenarios but will be valuable to test the methods' performance in a wide range of conditions. In biological experiments, other kinds of motion and even non-Gaussian behavior have been reported [21]. However, the choice of FBM does not limit the generality of the Challenge since other models of diffusion and non-Gaussian behavior can be obtained by properly tuning the parameters of the simulations. Datasets provided for the last phase of the competition will also include actual experiments for their comparative analysis with the Challenge methods (these data will not be used for the ranking).

The standard and straightforward approach in live-cell single-molecule imaging primarily captures information related to lateral motion. In cases involving flat membranes or isotropic systems, employing 2D imaging and tracking techniques suffices for obtaining accurate motion-related parameters. However, when dealing with motion on non-flat surfaces or within anisotropic 3D environments, relying solely on 2D projections can result in critical information being overlooked, potentially leading to the misinterpretation of diffusion coefficients or the appearance of apparent anomalous diffusion effects [57, 58]. Consequently, drawing definitive conclusions under such circumstances should be avoided or approached with caution. To study motion occurring in 3D space, it is advisable to employ 3D tracking methods, such as off-focus imaging (i.e., the analysis of ring patterns in the defocused point spread function) [59], interference/holographic approaches [60], multifocus imaging [61], or point spread function engineering [62]. Although more challenging, these methods can measure also the motion along the axial dimension, facilitating a more accurate characterization. For the purposes of the Challenge, we have chosen to concentrate on studying changes in diffusion behavior occurring within a 2D context, driven by particle interactions of various types.

While this challenge focuses on data from biological systems, the use of regime-switching detection and trajectory segmentation extends well beyond the domain of living cells. Particularly interesting applications also include, e.g., the analysis of biomedical signals [63], speech [64], traffic flows [65], seismic signals [66], econometrics [67, 68], ecology [69], and river flows [70].

METHODS

Datasets and ground truth

In order to benchmark the different methods on data with known ground truth, we rely on numerical simulations. We developed the `andi-datasets` Python package [71] to generate the required datasets to train and evaluate the various methods. Details about available functions can be found in the hosting repository [71].

Particle motion is simulated according to fractional Brownian motion (FBM, [54]), a model that reproduces Brownian and anomalous diffusion processes by tuning the correlation of the increments through the Hurst exponent H. FBM is a Gaussian process with a covariance function

$$\mathrm{E}[B_H(t)B_H(s)] = K\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \qquad (1)$$

where E[·] denotes the expected value and K is a constant with units length$^2 \cdot$ time$^{-2\mathrm{H}}$. In order to generalize FBM in two dimensions (2D), a trajectory $\mathbf{R}(t)$ is represented as $\mathbf{R}(t) = \{X(t), Y(t)\}$, where $X(t)$ and $Y(t)$ are independent FBM along the x and y axes, respectively [33]. The anomalous diffusion exponent is related to the Hurst exponent as $\alpha = 2H$ [54], and the MSD for an unconstrained FBM in 2D scales with time t as

$$\mathrm{MSD}(t) = 4Kt^{\alpha}. \qquad (2)$$

When $\alpha = 1$, FBM reverts to Brownian motion and K corresponds to the diffusion coefficient D. FBM can thus produce subdiffusion for $0 < H < 1/2$ ($0 < \alpha < 1$), Brownian diffusion for $H = 1/2$ ($\alpha = 1$), and superdiffusion for $1/2 < H < 1$ ($1 < \alpha < 2$).

We will consider the following physical models of motion and interactions(Fig. 2(a)):

- Single-state model (SSM) — Particles diffusing according to a single diffusion state, as observed for some lipids in the plasma membrane [14, 15, 72]. This model also serves as a negative control to assess the false positive rate of detecting diffusion changes.

- Multi-state model (MSM) — Particles diffusing according to a time-dependent multi-state (2 or more) model of diffusion undergoing transient changes of K and/or α. Examples of changes of K have been observed in proteins as induced by, e.g., allosteric changes or ligand binding [73–76].

- Dimerization model (DIM) — Particles diffusing according to a 2-state model of diffusion, with transient changes of K and/or α induced by encounters with other diffusing particles. Examples of changes of K have been observed in protein dimerization and protein-protein interactions [77–81].

- Transient-confinement model (TCM) — Particles diffusing according to a space-dependent 2-state model of diffusion, observed for example in proteins being transiently confined in regions where diffusion properties might change, e.g., the confinement induced by clathrin-coated pits on the cell membrane [82]. In the limit of a high density of trapping regions, this model reproduces the picket-and-fence model used to describe the effect of the actin cytoskeleton on transmembrane proteins [9, 83].

- Quenched-trap model (QTM) — Particles diffusing according to a space-dependent 2-state model of diffusion, representing proteins being transiently immobilized at specific locations as induced by binding to immobile structures, such as cytoskeleton-induced molecular pinning [17, 84].

While the interaction mechanisms producing the heterogeneous diffusion are inspired by biological scenarios, some of the combinations of diffusion parameters and models lead to situations that may not correspond to previously documented biological contexts. Nevertheless, this approach holds substantial value as it enables the comprehensive assessment of method performance across a broad spectrum of conditions.

In the simulations, each dynamic state is characterized by a distribution of values for the parameters K and α. For each trajectory, the values of K and α for each state are randomly drawn from Gaussian distributions with bounds $\alpha \in (0, 2)$ and $K \in [10^{-12}, 10^{6}]$ pixel^2/frame$^\alpha$. The interaction distance and the radius of confinement or trapping have constant values across each experiment. Simulations are provided in generalized units (i.e., pixels and frames) that can be rescaled to meaningful temporal and spatial scales.

A detailed description of the simulation procedure is presented in Appendix A. Table II summarizes all the parameter values used to generate the Pilot Dataset that can be downloaded at https://drive.google.com/drive/folders/1R41I5Y8ynXdXbUFLWOtQXx1nkUh3bWw3?usp=share_link. The Pilot Dataset provides typical numerical experiments that will be made available to the participants in the different phases of the challenge.

Datasets provided for the last phase of the competition will also include results from actual experiments that have reported the occurrence of heterogeneous diffusion but for which the ground truth is not established beyond any doubt. Therefore, these data will not contribute to the challenge scoring. Nevertheless, the predictions provided by different methods will be comparatively analyzed and discussed also with respect to the conclusions reached in the original publications. Together with the quantitative results obtained from simulations, these analyses will assess the applicability of the methods to real-world experimental data.

Competition design

To enable the assessment of the performance of previously established methods while fostering the development of new approaches and the participation from diverse disciplines, the challenge is organized along two tracks:

- Track 1 — based on the analysis of raw videos.
- Track 2 — based on the analysis of trajectories.

For each track, datasets are provided according to a hierarchical structure (Fig. 2(b-c)) that includes:

- Experiment — A given biological scenario defined by a model of interactions and a set of parameters describing the dynamic interplay of the particles and the environment.
- Field of view (FOV) — A region of the sample where the recording takes place. Particles within the same FOV can undergo interactions among themselves and/or with the environment.
- Video (Track 1 only) — Videos corresponding to each FOV.
- Trajectory (Track 2 only) — Trajectory corresponding to the motion of an individual particle.

For both tracks, all the particles used in the simulations are provided/visualized (i.e., full labeling conditions). The effect of blinking or photobleaching is not taken into

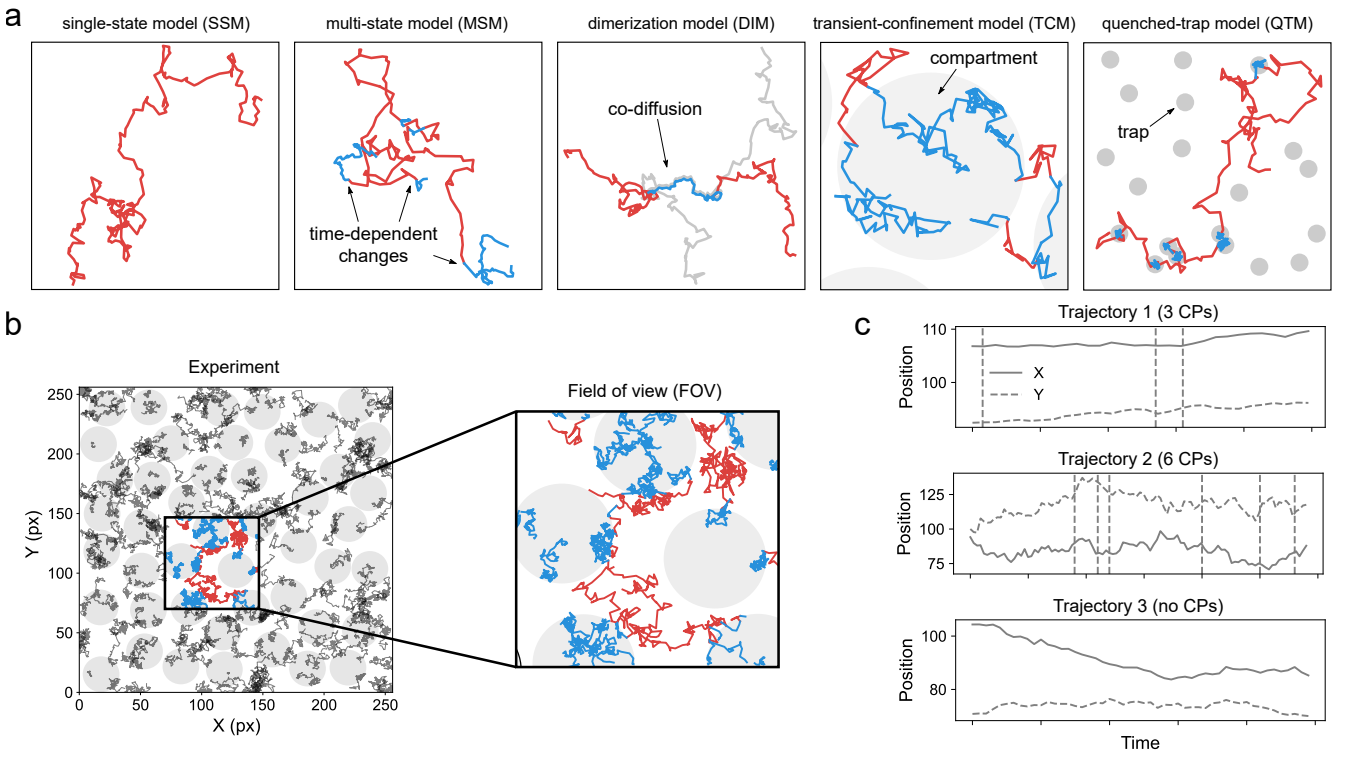

FIG. 2. **Physical models of interaction and structure of the simulated datasets.** **a**, Examples of 2-dimensional trajectories undergoing interactions inducing changes in their motion. From left to right: *single-state model* (SSM) without changes of diffusion; *multi-state model* (MSM) with time-dependent changes between different diffusive states (red and blue); *dimerization model* (DIM) where a particle (red) selectively interacts with another particle (gray) and the two transiently co-diffuse with a different motion (blue trajectory); *transient-confinement model* (TCM) where a particle diffuses inside (blue) and outside (red) compartments with osmotic boundaries (gray area); *quenched-trap model* (QTM) where a particle is transiently immobilized (blue) at specific loci through interactions with static features of the environment (gray areas). **b** An experiment (left panel) consists of simulations performed according to one of the models of interactions described in **a** (here shown a TCM experiment), with a set of parameters describing the dynamic interplay of the particles and/or the environment. From the same experiment, several fields of view (FOVs) are selected. Particles within the same FOV (right panel) diffuse and undergo interactions among themselves and/or with the environment (gray areas) that affect their trajectories. **c**, Time traces of the coordinates of exemplary trajectories from the experiment depicted in **b** displaying changes of diffusion properties at specific times (dashed vertical lines). For the challenge, the motion analysis can be either performed directly from the video recording of the FOV (Track 1), or from detected trajectories linking the coordinates of individual particles at different times (Track 2).

account. In each track, participants can compete in two different tasks, as typically done in the analysis of experimental data:

- Ensemble Task — Ensemble-level predictions providing for each experimental condition the model used to simulate the experiment, the number of states, and the fraction of time spent in each state. For each identified state, participants should determine the mean and standard deviation of the distribution of the generalized diffusion coefficients K, and the mean and standard deviation of the distribution of the anomalous diffusion exponent α corresponding to the underlying motion.

- Single-trajectory Task — Trajectory-level predictions providing for each trajectory a list of M inner CPs delimiting $M+1$ segments with different dynamic behavior. For each segment, participants should identify the generalized diffusion coefficient K, the anomalous diffusion exponent α corresponding to the underlying motion, and an identifier of the kind of constraint imposed by the environment (0 = immobile, 1 = confined, 2 = free (unconstrained), 3 = directed ($\alpha \leq 1.9$)). For Track 1, predictions must be provided for a subset of particles (in the following, we will refer to them as VIP, very important particles) identified through a label map of the first frame of the movie. For Track 2, predictions must be provided for all trajectories in the FOV.

For each task, several metrics will be evaluated (see section Scoring and evaluations). Participants can provide partial submissions, e.g., including predictions for a limited subset of experiments or for specific parameters. For ranking purposes of the challenge, missing predictions

will be scored with the worst possible value of the corresponding metric.

Competition timeline

We plan to run the 2nd AnDi Challenge as a time-limited competition from October 1, 2023, to June 15, 2024 (the competition dates may be shifted depending on the duration of the Stage 1 review). The competition will be hosted on the Codalab platform and will be divided into three phases, namely *Development*, *Validation*, and *Challenge*.

The Development Phase (December 1, 2023, to March 31, 2024) is intended for the participants to set up their methods, test them, and familiarize themselves with the datasets and the scoring platform. An unlabeled dataset will be available and the public leaderboard will show scores obtained on this dataset. The code to generate further data will also be available.

The Validation Phase (May 1, 2024, to June 30, 2024) will be a test of the actual final challenge. A new dataset will be provided and the public leaderboard will show the scores for each task.

The Challenge Phase (July 1, 2024, to June 15, 2024) will be the final stage of the competition. Only teams in the top 10 ranking of each task will be allowed to participate in this phase. A new dataset will be provided and the number of submissions will be limited to 1 per day. Results obtained by other participants will not be publicly disclosed, as the leaderboard will be made publicly available only after the deadline.

The analysis of the final results will be performed from July to October 2024. If Stage 1 review is successful, the manuscript will be submitted by the end of October 2024. Members of the teams in the top-5 ranking of each task will be invited to contribute to the article as authors.

The results of the challenge will be discussed with the participants and other experts from the field during the 2nd Anomalous Diffusion Workshop that will be held in December 2024.

Dataset structure

The datasets used in the challenge (Fig. 3) include different experiments, each contained in a folder labeled with a sequential number (`EXP_[exp number]`) and corresponding to a specific model and a fixed set of parameters. The information about the model and the parameters is unknown to challenge participants. Each experiment folder contains a list of files labeled with a sequential number (`FOV_[fov number]`) associated with 30 FOVs. Each FOV reports data from a variable number of particles diffusing on a 128×128 pixel^2 area.

For Track 1 (Fig. 3(a)), the coordinates of particles in the same FOV are used to generate 200-frame videos as a series of 8-bit images in the multi-tiff format using Deeptrack 2.1 [5]. Noise is added to the synthetic images to account for background fluorescence and shot noise. A map corresponding to the segmentation of VIP particles at the first frame for which CPs and diffusion parameters must be detected is also provided as a tiff file. Connected components of the map are labeled with unique integer values that correspond to the particle index.

For Track 2 (Fig. 3(a)), we provide a csv file for each FOV containing a table whose columns contain trajectory index, time step, x-coordinate, and y-coordinate. Coordinates of simulated trajectories are corrupted with Gaussian noise corresponding to finite (subpixel) localization precision. The trajectories have a maximum length of 200 frames.

Besides localization precision, motion blur can introduce a significant contribution to noise, in particular if the camera frame rate is slow compared to particle motion [85]. However, this aspect will not be included in the Challenge datasets since it would introduce complexities in the definition of the ground truth that could detract from the focus of the work. Nevertheless, the simulation software incorporates the capability to introduce the effect of motion blur both in videos and trajectories.

Exemplary data for all the models are shown in Fig. 4. Files in different tracks labeled with the same experiment and the FOV index (e.g., `Track_1/EXP_4/FOV_3.tiff` and `Track_2/EXP_4/FOV_3.csv`) include simulations obtained with the same set of dynamics parameters but do not correspond to the motion of the same set of particles.

A. Scoring and evaluation

The performance of the methods will be evaluated using specific metrics for each Task. For ranking purposes in the challenge, composite metrics will be used, as described below.

1. *Ensemble Tasks*

Participation in an ensemble task requires predictions of the type of model used for simulating each experiment, the number of states S of the model, and the parameters of each state. The type of model will be simply evaluated as correct or wrong. The prediction of the number of states will be assessed by measuring the difference with the ground truth. For both the generalized diffusion coefficient and the anomalous diffusion exponent, predictions must include the mean, the standard deviation, and the relative weight of each state. From these values, we will compute the associated multi-modal distributions P_α and P_D. The similarity of these distributions to the ground-truth distributions Q_α and Q_D will be assessed by means of the first Wasserstein distance (W_1),

$$W_1(P,Q) = \int_{\mathrm{supp}(Q)} |\mathrm{CDF}_P(x) - \mathrm{CDF}_Q(x)| dx \quad (3)$$

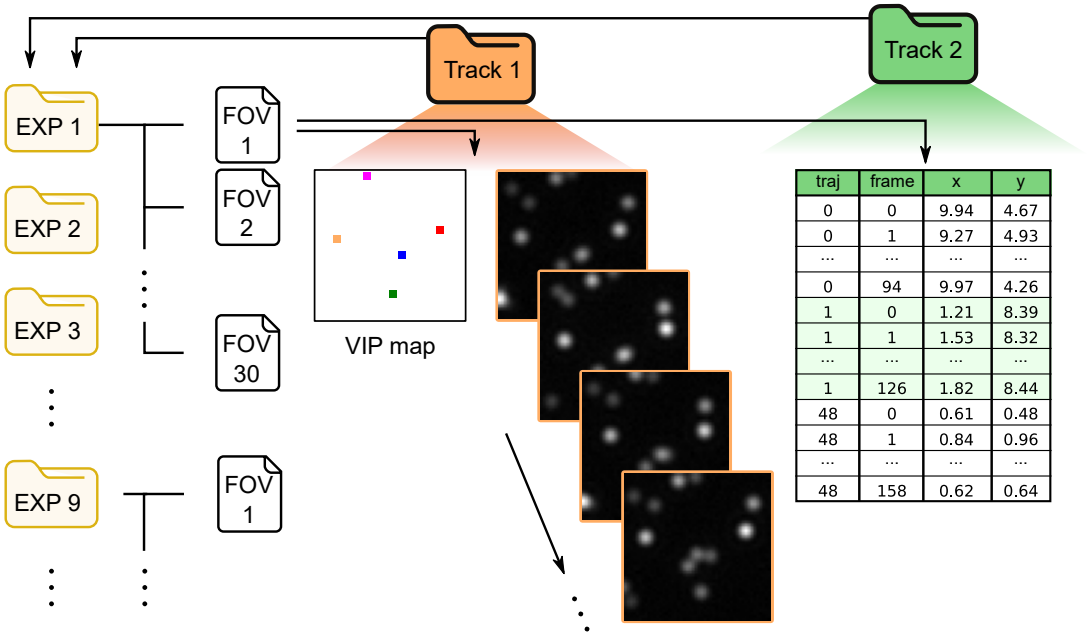

FIG. 3. **Structure of the dataset.** Datasets for each track are contained in the respective directories, namely, `Track 1` and `Track 2`. These include a folder for each experiment labeled with a sequential number (`EXP_[exp number]`). Each experiment folder contains a list of 30 files labeled with a sequential number (`FOV_[fov number]`) associated with different non-overlapping FOVs. For Track 1, the files consist of 200-frame videos and an additional map corresponding to the segmentation of VIP particles at the first frame. For Track 2, the files contain a table whose columns correspond to the trajectory index, time step, x-coordinate, and y-coordinate.

where CDF_Q is the cumulative distribution function of the distribution Q and $\mathrm{supp}(Q)$ is the support ($\alpha \in (0,2)$ and $K \in [10^{-12}, 10^{6}]$ pixel$^2 \cdot$ frame$^{-\alpha}$).

2. *Single-trajectory Tasks*

Participation in a single-trajectory task requires predictions of the M CPs and the dynamic properties, i.e., the generalized diffusion coefficient K, the anomalous exponent α, and diffusive-type identifiers of the resulting $M+1$ segments. Different metrics will be used to evaluate the methods' performance.

a. CP detection metrics Following Ref. [51], given a ground-truth CP at locations $t_{(\mathrm{GT}),i}$ and a predicted CP at locations $t_{(\mathrm{P}),j}$, we define the gated absolute distance:

$$d_{i,j} = \min(|t_{(\mathrm{GT}),i} - t_{(\mathrm{P}),j}|, \varepsilon_{\mathrm{CP}}), \tag{4}$$

where $\varepsilon_{\mathrm{CP}}$ is used as a fixed maximum penalty for CPs located more than $\varepsilon_{\mathrm{CP}}$ apart. For a set of M_{GT} ground-truth CPs and M_{P} predicted CPs, we can solve a rectangular assignment problem using the Hungarian algorithm [86] by minimizing the sum of distances between paired CPs:

$$d_{\mathrm{CP}} = \min_{\text{paired CP}} \Big(\sum d_{i,j} \Big). \tag{5}$$

The distance d_{CP} allows to define a pairing metrics:

$$\alpha_{\mathrm{CP}} = 1 - \frac{d_{\mathrm{CP}}}{d_{\mathrm{CP}}^{\max}}, \tag{6}$$

where $d_{\mathrm{CP}}^{\max} = M_{\mathrm{GT}}\ \varepsilon_{\mathrm{CP}}$ is the distance associated with having all predicted CPs unpaired or at a distance larger than $\varepsilon_{\mathrm{CP}}$ from all ground-truth CP. The metric α_{CP} is bound in $[0,1]$, taking a value of 1 if all ground-truth and predicted CPs are matching exactly. Similarly, we define a CP localization metric:

$$\beta_{\mathrm{CP}} = \frac{d_{\mathrm{CP}}^{\max} - d_{\mathrm{CP}}}{d_{\mathrm{CP}}^{\max} + \overline{d_{\mathrm{CP}}}}, \tag{7}$$

where $\overline{d_{\mathrm{CP}}}$ is the distance associated with having all unassigned predicted CPs at a distance larger than $\varepsilon_{\mathrm{CP}}$ from all ground-truth CPs. This metric measures the presence of spurious CPs and is bound in $[0, \alpha_{\mathrm{CP}}]$, taking value α_{CP} if no spurious CPs are present. We also calculate the number of true positive (TP), i.e., the paired true and predicted CPs with a distance smaller than $\varepsilon_{\mathrm{CP}}$. Spurious predictions, i.e., not associated with any ground truth or having a distance larger than $\varepsilon_{\mathrm{CP}}$ are counted as false positive (FP). Ground truth CPs not having an associated prediction at a distance shorter than $\varepsilon_{\mathrm{CP}}$ are considered false negative (FN). Given an experiment containing N trajectories, we compute the overall number of TP, FP, and FN. We then use these values to calculate the Jaccard similarity coefficient (JSC) over the

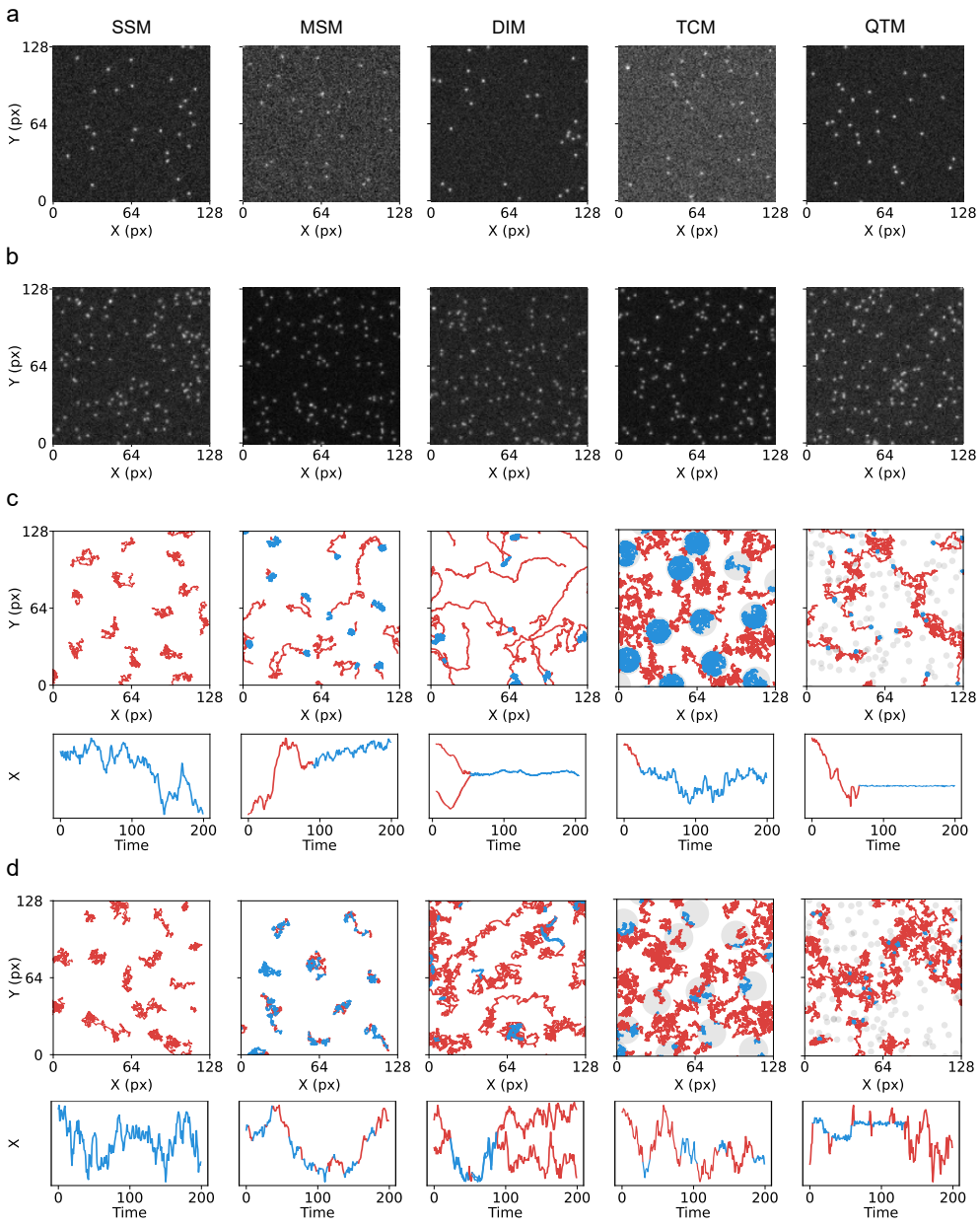

FIG. 4. **Examples of challenge datasets.** **a, b**, Frames of exemplary videos (FOVs) provided for Track 1 corresponding to the different diffusion models and reproducing low (**a**) and high number density (**b**). **c, d**, The upper rows show plots of the trajectories for exemplary FOVs provided for Track 2 corresponding to the different diffusion models and different sets of parameters exhibiting evident (**c**) or subtle changes (**d**) of diffusion, i.e., data in **c** reproduce conditions were diffusion changes are easier to identify as compared to **d**. For example, for the SSM, all trajectories in **c** have the same generalized diffusion coefficient, whereas in **d** they have a broad distribution of K. For the MSM, trajectories in **c** undergo large changes of diffusivity and show long dwell times in each state, as compared to **d**. For the DIM, trajectories in **c** undergo large changes of anomalous diffusion exponent and show long dwell times in each state, as compared to **d**. For the TCM, trajectories cannot leave the confinement zone once they get inside **c**, whereas in **d** the compartments have a finite transmittance. For the QTM, the dwell times in the traps are much longer in **c** than in **d**. The lower rows show time traces of the x-coordinate of representative trajectories. The different colors indicate different diffusive states.

whole experiment as:

$$\mathrm{JSC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP}}. \tag{8}$$

For the predicted CPs classified as TP, we also compute the root mean square error (RMSE), defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{\substack{\text{paired CP} \\ d_{i,j} < \varepsilon_{\mathrm{CP}}}} \left(t_{(\mathrm{GT}),i} - t_{(\mathrm{P}),j} \right)^2}. \tag{9}$$

b. Metrics for the estimation of dynamic properties For the evaluation of methods performance for the estimation of the dynamic properties, we first follow a procedure similar to the one described above for the pairing of the CPs. Predicted CPs are used to define the predicted trajectory segments. We define a distance between predicted and ground-truth segments based on the JSC calculated with respect to their temporal support, where time points at which predicted and ground-truth segments overlap are considered as TP, predicted time points not corresponding to the ground truth as FP, and ground-truth time points not predicted as FN. The Hungarian algorithm is used to pair segments by maximizing the sum of the JSC. Only paired segments are used to calculate metrics assessing methods performance for the estimation of dynamics properties. For the generalized diffusion coefficient K, we use the mean squared logarithmic error (MSLE) defined as:

$$\mathrm{MSLE} = \frac{1}{N} \sum_{\substack{\text{paired} \\ \text{segments}}} \Big(\log(K_{(\mathrm{GT}),i} + 1) - \log(K_{(\mathrm{P}),j} + 1) \Big)^2. \tag{10}$$

For the anomalous diffusion exponents α, we use the mean absolute error (MAE):

$$\mathrm{MAE}_\alpha = \frac{1}{N} \sum_{\substack{\text{paired} \\ \text{segments}}} |\alpha_{(\mathrm{GT}),i} - \alpha_{(\mathrm{P}),j}|, \tag{11}$$

where N is the total number of paired segments in the experiment, $\alpha_{(\mathrm{GT}),i}$ and $\alpha_{(\mathrm{P}),j}$ represent the ground-truth and predicted values of the anomalous exponent of paired segments, respectively. For the classification of the type of diffusion, we use the F_1-score:

$$\mathrm{F}_1 = \frac{2\mathrm{TP_c}}{2\mathrm{TP_c} + \mathrm{FP_c} + \mathrm{FN_c}}, \tag{12}$$

where $\mathrm{TP_c}$, $\mathrm{FP_c}$, and $\mathrm{FN_c}$ represent true positives, false positives, and false negatives with respect to segment classification. The metric is calculated as a micro-average, which aggregates the contributions of all classes to compute the average metric and is generally preferable when class imbalance is present.

B. Metrics for challenge ranking

For ranking purposes, we will use the mean reciprocal rank (MRR) as a summary statistic for the overall evaluation of software performance [42]:

$$\mathrm{MRR} = \frac{1}{N} \cdot \sum_{i=1}^{N} \frac{1}{\mathrm{rank}_{\mathrm{M_i}}}, \tag{13}$$

where $\mathrm{rank}_{\mathrm{M_i}}$ corresponds to the position in an ordered list based on the value of the corresponding metrics $\mathrm{M_i}$.

For Task 1, the metrics involved in the calculation will be the F_1-score of the model, the MAE of the distributions of K and α. For Task 2, the JSC and the RMSE of CPs, the MSLE of K, and the MAE of α.

DATA AVAILABILITY

A Pilot Dataset is available for download at https://drive.google.com/drive/folders/1R41I5Y8ynXdXbUFLWOtQXx1nkUh3bWw3?usp=share_link

CODE AVAILABILITY

All software used for the Challenge is available at https://github.com/AnDiChallenge/andi_datasets.

[1] H. Shen, L. J. Tauzin, R. Baiyasi, W. Wang, N. Moringo, B. Shuang, and C. F. Landes, Single particle tracking: from theory to biophysical applications, Chemical Reviews **117**, 7331 (2017).

[2] C. Manzo and M. F. Garcia-Parajo, A review of progress in single particle tracking: from methods to biophysical insights, Reports on Progress in Physics **78**, 124601 (2015).

[3] Y. Chen, B. C. Lagerholm, B. Yang, and K. Jacobson, Methods to measure the lateral diffusion of membrane lipids and proteins, Methods **39**, 147 (2006).

[4] K. Norregaard, R. Metzler, C. M. Ritter, K. Berg-Sørensen, and L. B. Oddershede, Manipulation and motion of organelles and single molecules in living cells, Chemical Reviews **117**, 4342 (2017).

[5] B. Midtvedt, S. Helgadottir, A. Argun, J. Pineda, D. Midtvedt, and G. Volpe, Quantitative digital microscopy with deep learning, Applied Physics Reviews **8**, 011310 (2021).

[6] J. A. Torreno-Pina, C. Manzo, and M. F. Garcia-Parajo, Uncovering homo-and hetero-interactions on the cell membrane using single particle tracking approaches, Journal of Physics D: Applied Physics **49**, 104002 (2016).

[7] A. Einstein, Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, Annalen der Physik **322**, 549 (1905).

[8] M. Weiss, H. Hashimoto, and T. Nilsson, Anomalous protein diffusion in living cells as seen by fluorescence correlation spectroscopy, Biophysical journal **84**, 4043

(2003).
[9] K. Ritchie, R. Iino, T. Fujiwara, K. Murase, and A. Kusumi, The fence and picket structure of the plasma membrane of live cells as revealed by single molecule techniques, Molecular Membrane Biology **20**, 13 (2003).
[10] F. Daumas, N. Destainville, C. Millot, A. Lopez, D. Dean, and L. Salomé, Confined diffusion without fences of a G-protein-coupled receptor as revealed by single particle tracking, Biophysical Journal **84**, 356 (2003).
[11] K. Ritchie, X.-Y. Shan, J. Kondo, K. Iwasawa, T. Fujiwara, and A. Kusumi, Detection of non-brownian diffusion in the cell membrane in single molecule tracking, Biophysical journal **88**, 2266 (2005).
[12] D. S. Banks and C. Fradin, Anomalous diffusion of proteins due to molecular crowding, Biophysical journal **89**, 2960 (2005).
[13] M. J. Saxton, A biological interpretation of transient anomalous subdiffusion. i. qualitative model, Biophysical journal **92**, 1178 (2007).
[14] C. Eggeling, C. Ringemann, R. Medda, G. Schwarzmann, K. Sandhoff, S. Polyakova, V. N. Belov, B. Hein, C. Von Middendorff, A. Schönle, *et al.*, Direct observation of the nanoscale dynamics of membrane lipids in a living cell, Nature **457**, 1159 (2009).
[15] C. Manzo, T. S. van Zanten, and M. F. Garcia-Parajo, Nanoscale fluorescence correlation spectroscopy on intact living cell membranes with nsom probes, Biophysical journal **100**, L8 (2011).
[16] H. Soula, B. Caré, G. Beslon, and H. Berry, Anomalous versus slowed-down Brownian diffusion in the ligand-binding equilibrium, Biophysical Journal **105**, 2064 (2013).
[17] K. M. Spillane, J. Ortega-Arroyo, G. De Wit, C. Eggeling, H. Ewers, M. I. Wallace, and P. Kukura, High-speed single-particle tracking of GM1 in model membranes reveals anomalous diffusion due to interleaflet coupling and molecular pinning, Nano Letters **14**, 5390 (2014).
[18] A. Mosqueira, P. A. Camino, and F. J. Barrantes, Antibody-induced crosslinking and cholesterol-sensitive, anomalous diffusion of nicotinic acetylcholine receptors, Journal of Neurochemistry **152**, 663 (2020).
[19] Y.-J. Chai, C.-Y. Cheng, Y.-H. Liao, C.-H. Lin, and C.-L. Hsieh, Heterogeneous nanoscopic lipid diffusion in the live cell membrane and its dependency on cholesterol, Biophysical Journal **121**, 3146 (2022).
[20] F. Höfling and T. Franosch, Anomalous transport in the crowded world of biological cells, Reports on Progress in Physics **76**, 046602 (2013).
[21] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking, Physical Chemistry Chemical Physics **16**, 24128 (2014).
[22] D. Krapf, Mechanisms underlying anomalous diffusion in the plasma membrane, Current Topics in Membranes **75**, 167 (2015).
[23] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter, Fractional brownian motion versus the continuous-time random walk: A simple test for subdiffusive dynamics, Physical review letters **103**, 180602 (2009).
[24] E. Kepten, I. Bronshtein, and Y. Garini, Ergodicity convergence test suggests telomere motion obeys fractional dynamics, Physical Review E **83**, 041919 (2011).
[25] 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 **6**, 8044 (2015).
[26] S. C. Weber, A. J. Spakowitz, and J. A. Theriot, Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm, Physical review letters **104**, 238102 (2010).
[27] S. A. Tabei, S. Burov, H. Y. Kim, A. Kuznetsov, T. Huynh, J. Jureller, L. H. Philipson, A. R. Dinner, and N. F. Scherer, Intracellular transport of insulin granules is a subordinated random walk, Proceedings of the National Academy of Sciences **110**, 4911 (2013).
[28] T. J. Lampo, S. Stylianidou, M. P. Backlund, P. A. Wiggins, and A. J. Spakowitz, Cytoplasmic rna-protein particles exhibit non-gaussian subdiffusive behavior, Biophysical journal **112**, 532 (2017).
[29] A. V. Weigel, B. Simon, M. M. Tamkun, and D. Krapf, Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking, Proceedings of the National Academy of Sciences of the United States of America **108**, 6438 (2011).
[30] C. Manzo, J. A. Torreno-Pina, P. Massignan, G. J. Lapeyre Jr, M. Lewenstein, and M. F. G. Parajo, Weak ergodicity breaking of receptor motion in living cells stemming from random diffusivity, Physical Review X **5**, 011021 (2015).
[31] P. R. Smith, I. E. Morrison, K. M. Wilson, N. Fernandez, and R. J. Cherry, Anomalous diffusion of major histocompatibility complex class i molecules on hela cells determined by single particle tracking, Biophysical Journal **76**, 3331 (1999).
[32] M. S. Song, H. C. Moon, J.-H. Jeon, and H. Y. Park, Neuronal messenger ribonucleoprotein transport follows an aging lévy walk, Nature communications **9**, 1 (2018).
[33] D. Krapf, N. Lukat, E. Marinari, R. Metzler, G. Oshanin, C. Selhuber-Unkel, A. Squarcini, L. Stadler, M. Weiss, and X. Xu, Spectral content of a single non-Brownian trajectory, Physical Review X **9**, 011019 (2019).
[34] V. Sposini, D. Krapf, E. Marinari, R. Sunyer, F. Ritort, F. Taheri, C. Selhuber-Unkel, R. Benelli, M. Weiss, R. Metzler, *et al.*, Towards a robust criterion of anomalous diffusion, Communications Physics **5**, 305 (2022).
[35] N. Granik, L. E. Weiss, E. Nehme, M. Levin, M. Chein, E. Perlson, Y. Roichman, and Y. Shechtman, Single-particle diffusion characterization by deep learning, Biophysical Journal **117**, 185 (2019).
[36] P. Kowalek, H. Loch-Olszewska, and J. Szwabiński, Classification of diffusion modes in single-particle tracking data: Feature-based versus deep-learning approach, Physical Review E **100**, 032410 (2019).
[37] S. Bo, F. Schmidt, R. Eichhorn, and G. Volpe, Measurement of anomalous diffusion using recurrent neural networks, Physical Review E **100**, 010102 (2019).
[38] G. Muñoz-Gil, M. A. Garcia-March, C. Manzo, J. D. Martín-Guerrero, and M. Lewenstein, Single trajectory characterization via machine learning, New Journal of Physics **22**, 013010 (2020).
[39] H. Seckler and R. Metzler, Bayesian deep learning for error estimation in the analysis of anomalous diffusion, Nature Communications **13**, 6717 (2022).

[40] J. Pineda, B. Midtvedt, H. Bachimanchi, S. Noé, D. Midtvedt, G. Volpe, and C. Manzo, Geometric deep learning reveals the spatiotemporal features of microscopic motion, Nature Machine Intelligence **5**, 71 (2023).

[41] H. Seckler, J. Szwabiński, and R. Metzler, Machine-learning solutions for the analysis of single-particle diffusion trajectories, The Journal of Physical Chemistry Letters **14**, 7910 (2023).

[42] G. Muñoz-Gil, G. Volpe, M. A. Garcia-March, E. Aghion, A. Argun, C. B. Hong, T. Bland, S. Bo, J. A. Conejero, N. Firbas, *et al.*, Objective comparison of methods to decode anomalous diffusion, Nature Communications **12**, 6253 (2021).

[43] J. Janczura, M. Balcerek, K. Burnecki, A. Sabri, M. Weiss, and D. Krapf, Identifying heterogeneous diffusion states in the cytoplasm by a hidden markov model, New Journal of Physics **23**, 053018 (2021).

[44] G. Lee, G. Leech, M. J. Rust, M. Das, R. J. McGorty, J. L. Ross, and R. M. Robertson-Anderson, Myosin-driven actin-microtubule networks exhibit self-organized contractile dynamics, Science Advances **7**, eabe4334 (2021).

[45] B. Requena, S. Masó, J. Bertran, M. Lewenstein, C. Manzo, and G. Muñoz-Gil, Inferring pointwise diffusion properties of single trajectories with deep learning, arXiv preprint arXiv:2302.00410 (2023).

[46] D. Han, N. Korabel, R. Chen, M. Johnston, A. Gavrilova, V. J. Allan, S. Fedotov, and T. A. Waigh, Deciphering anomalous heterogeneous intracellular transport with neural networks, Elife **9**, e52224 (2020).

[47] W. Wang, F. Seno, I. M. Sokolov, A. V. Chechkin, and R. Metzler, Unexpected crossovers in correlated random-diffusivity processes, New Journal of Physics **22**, 083041 (2020).

[48] M. Balcerek, K. Burnecki, S. Thapa, A. Wyłomańska, and A. Chechkin, Fractional brownian motion with random hurst exponent: Accelerating diffusion and persistence transitions, Chaos: An Interdisciplinary Journal of Nonlinear Science **32** (2022).

[49] W. Wang, M. Balcerek, K. Burnecki, A. V. Chechkin, S. Janušonis, J. Ślęzak, T. Vojta, A. Wyłomańska, and R. Metzler, Memory-multi-fractional brownian motion with continuous correlations, Physical Review Research **5**, L032025 (2023).

[50] J. Ślęzak and R. Metzler, Minimal model of diffusion with time changing hurst exponent, Journal of Physics A: Mathematical and Theoretical **56**, 35LT01 (2023).

[51] 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 **11**, 281 (2014).

[52] J. P. Bohnslav, N. K. Wimalasena, K. J. Clausing, Y. Y. Dai, D. A. Yarmolinsky, T. Cruz, A. D. Kashlan, M. E. Chiappe, L. L. Orefice, C. J. Woolf, *et al.*, Deepethogram, a machine learning pipeline for supervised behavior classification from raw pixels, Elife **10**, e63377 (2021).

[53] B. Midtvedt, E. Olsén, F. Eklund, F. Höök, C. B. Adiels, G. Volpe, and D. Midtvedt, Fast and accurate nanoparticle characterization using deep-learning-enhanced off-axis holography, ACS Nano **15**, 2240 (2021).

[54] B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Review **10**, 422 (1968).

[55] M. Weiss, Single-particle tracking data reveal anticorrelated fractional Brownian motion in crowded fluids, Physical Review E **88**, 010101 (2013).

[56] I. Golding and E. C. Cox, Physical nature of bacterial cytoplasm, Physical review letters **96**, 098102 (2006).

[57] A. Dupont, M. Gorelashvili, V. Schüller, F. Wehnekamp, D. Arcizet, Y. Katayama, D. Lamb, and D. Heinrich, Three-dimensional single-particle tracking in live cells: news from the third dimension, New journal of physics **15**, 075008 (2013).

[58] V. M. Sukhorukov and J. Bereiter-Hahn, Anomalous diffusion induced by cristae geometry in the inner mitochondrial membrane, PLoS One **4**, e4604 (2009).

[59] M. Speidel, A. Jonáš, and E.-L. Florin, Three-dimensional tracking of fluorescent nanoparticles with subnanometer precision by use of off-focus imaging, Optics letters **28**, 69 (2003).

[60] L. Maurice and A. Bilenca, Three-dimensional single particle tracking using 4π self-interference of temporally phase-shifted fluorescence, Light: Science & Applications **12**, 58 (2023).

[61] E. Toprak, H. Balci, B. H. Blehm, and P. R. Selvin, Three-dimensional particle tracking via bifocal imaging, Nano letters **7**, 2043 (2007).

[62] L. Holtzer, T. Meckel, and T. Schmidt, Nanometric three-dimensional tracking of individual quantum dots in cells, Applied Physics Letters **90** (2007).

[63] R. V. Andreao, B. Dorizzi, and J. Boudy, ECG signal analysis through hidden Markov models, IEEE Transactions on Biomedical Engineering **53**, 1541 (2006).

[64] V. Khanagha, K. Daoudi, O. Pont, and H. Yahia, Phonetic segmentation of speech signal using local singularity analysis, Digital Signal Processing **35**, 86 (2014).

[65] M. Cetin and G. Comert, Short-term traffic flow prediction with regime switching models, Transportation Research Record **1965**, 23 (2006).

[66] J. Bulla and A. Berzel, Computational issues in parameter estimation for stationary hidden Markov models, Computational Statistics **23**, 1 (2008).

[67] J. Janczura and R. Weron, Goodness-of-fit testing for the marginal distribution of regime-switching models with an application to electricity spot prices, AStA Advances in Statistical Analysis **97**, 239 (2013).

[68] T. Lux and L. Morales-Arias, Forecasting volatility under fractality, regime-switching, long memory and student-T innovations, Computational Statistics & Data Analysis **54**, 2676 (2010).

[69] H. Edelhoff, J. Signer, and N. Balkenhol, Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns, Movement ecology **4**, 1 (2016).

[70] K. Vasas, P. Elek, and L. Márkus, A two-state regime switching autoregressive model with an application to river flow analysis, Journal of Statistical Planning and Inference **137**, 3113 (2007).

[71] G. Muñoz-Gil, B. Requena, G. Volpe, M. A. Garcia-March, and C. Manzo, AnDiChallenge/ANDI_datasets: Challenge 2020 release (2021).

[72] A. Honigmann, V. Mueller, H. Ta, A. Schoenle, E. Sezgin, S. W. Hell, and C. Eggeling, Scanning STED-FCS reveals spatiotemporal heterogeneity of lipid interaction

in the plasma membrane of living cells, Nature Communications **5**, 1 (2014).
[73] D. Mainali and E. A. Smith, The effect of ligand affinity on integrins' lateral diffusion in cultured cells, European Biophysics Journal **42**, 281 (2013).
[74] M. Yanagawa, M. Hiroshima, Y. Togashi, M. Abe, T. Yamashita, Y. Shichida, M. Murata, M. Ueda, and Y. Sako, Single-molecule diffusion-based estimation of ligand effects on g protein–coupled receptors, Science Signaling **11**, eaao1917 (2018).
[75] B. da Rocha-Azevedo, S. Lee, A. Dasgupta, A. R. Vega, L. R. de Oliveira, T. Kim, M. Kittisopikul, Z. A. Malik, and K. Jaqaman, Heterogeneity in vegf receptor-2 mobility and organization on the endothelial cell surface leads to diverse models of activation by vegf, Cell reports **32** (2020).
[76] A. M. Achimovich, T. Yan, and A. Gahlmann, Dimerization of ilid optogenetic proteins observed using 3d single-molecule tracking in live e. coli, Biophysical Journal **122**, 3254 (2023).
[77] S. T. Low-Nam, K. A. Lidke, P. J. Cutler, R. C. Roovers, P. M. van Bergen en Henegouwen, B. S. Wilson, and D. S. Lidke, Erbb1 dimerization is promoted by domain co-confinement and stabilized by ligand binding, Nature structural & molecular biology **18**, 1244 (2011).
[78] C. C. Valley, D. J. Arndt-Jovin, N. Karedla, M. P. Steinkamp, A. I. Chizhik, W. S. Hlavacek, B. S. Wilson, K. A. Lidke, and D. S. Lidke, Enhanced dimerization drives ligand-independent activity of mutant epidermal growth factor receptor in lung cancer, Molecular Biology of the Cell **26**, 4087 (2015).
[79] A. Tabor, S. Weisenburger, A. Banerjee, N. Purkayastha, J. M. Kaindl, H. Hübner, L. Wei, T. W. Grömer, J. Kornhuber, N. Tschammer, *et al.*, Visualization and ligand-induced modulation of dopamine receptor dimerization at the single molecule level, Scientific Reports **6**, 1 (2016).
[80] 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 **550**, 543 (2017).
[81] J. Grimes, Z. Koszegi, Y. Lanoiselee, T. Miljus, S. L. O'Brien, T. M. Stepniewski, B. Medel-Lacruz, M. Baidya, M. Makarova, D. M. Owen, *et al.*, Single-molecule analysis of receptor-beta-arrestin interactions in living cells, bioRxiv , 2022 (2022).
[82] A. V. Weigel, M. M. Tamkun, and D. Krapf, Quantifying the dynamic interactions between a clathrin-coated pit and cargo molecules, Proceedings of the National Academy of Sciences **110**, E4591 (2013).
[83] S. Sadegh, J. L. Higgins, P. C. Mannion, M. M. Tamkun, and D. Krapf, Plasma membrane is compartmentalized by a self-similar cortical actin meshwork, Physical Review X **7**, 011031 (2017).
[84] O. Rossier, V. Octeau, J.-B. Sibarita, C. Leduc, B. Tessier, D. Nair, V. Gatterdam, O. Destaing, C. Albiges-Rizo, R. Tampé, *et al.*, Integrins $\beta 1$ and $\beta 3$ exhibit distinct dynamic nanoscale organizations inside focal adhesions, Nature cell biology **14**, 1057 (2012).
[85] A. J. Berglund, Statistics of camera-based single-particle tracking, Physical Review E **82**, 011917 (2010).
[86] D. F. Crouse, On implementing 2D rectangular assignment algorithms, IEEE Transactions on Aerospace and Electronic Systems **52**, 1679 (2016).
[87] P. J. Slator, C. W. Cairo, and N. J. Burroughs, Detection of diffusion heterogeneity in single particle tracking trajectories using a hidden Markov model with measurement noise propagation, PLOS One **10**, e0140759 (2015).
[88] P. J. Slator and N. J. Burroughs, A hidden Markov model for detecting confinement in single-particle tracking trajectories, Biophysical Journal **115**, 1741 (2018).
[89] J. A. Helmuth, C. J. Burckhardt, P. Koumoutsakos, U. F. Greber, and I. F. Sbalzarini, A novel supervised trajectory segmentation algorithm identifies distinct types of human adenovirus motion in host cells, Journal of Structural Biology **159**, 347 (2007).
[90] M. Arts, I. Smal, M. W. Paul, C. Wyman, and E. Meijering, Particle mobility analysis using deep learning and the moment scaling spectrum, Scientific Reports **9**, 1 (2019).
[91] A. Sosa-Costa, I. K. Piechocka, L. Gardini, F. S. Pavone, M. Capitanio, M. F. Garcia-Parajo, and C. Manzo, PLANT: a method for detecting changes of slope in noisy trajectories, Biophysical Journal **114**, 2044 (2018).
[92] D. Montiel, H. Cang, and H. Yang, Quantitative characterization of changes in dynamical behavior for single-particle tracking studies, The Journal of Physical Chemistry B **110**, 19763 (2006).
[93] S. Yin, N. Song, and H. Yang, Detection of velocity and diffusion coefficient change points in single-particle trajectories, Biophysical Journal **115**, 217 (2018).
[94] P. Dosset, P. Rassam, L. Fernandez, C. Espenel, E. Rubinstein, E. Margeat, and P.-E. Milhiet, Automatic detection of diffusion modes within biological membranes using back-propagation neural network, BMC Bioinformatics **17**, 1 (2016).
[95] A. R. Vega, S. A. Freeman, S. Grinstein, and K. Jaqaman, Multistep track segmentation and motion classification for transient mobility analysis, Biophysical Journal **114**, 1018 (2018).
[96] Y. Lanoiselée and D. S. Grebenkov, Unraveling intermittent features in single-particle trajectories by a local convex hull method, Physical Review E **96**, 022144 (2017).
[97] G. Sikora, A. Wyłomańska, J. Gajda, L. Solé, E. J. Akin, M. M. Tamkun, and D. Krapf, Elucidating distinct ion channel populations on the surface of hippocampal neurons via single-particle tracking recurrence analysis, Physical Review E **96**, 062404 (2017).
[98] Y. Matsuda, I. Hanasaki, R. Iwao, H. Yamaguchi, and T. Niimi, Estimation of diffusive states from single-particle trajectory in heterogeneous medium using machine-learning methods, Physical Chemistry Chemical Physics **20**, 24099 (2018).
[99] C. Metzner, C. Mark, J. Steinwachs, L. Lautscham, F. Stadler, and B. Fabry, Superstatistical analysis and modelling of heterogeneous random walks, Nature Communications **6**, 1 (2015).
[100] G. J. Schütz, H. Schindler, and T. Schmidt, Single-molecule microscopy on model membranes reveals anomalous diffusion, Biophysical Journal **73**, 1073 (1997).
[101] R. Simson, E. D. Sheets, and K. Jacobson, Detection of temporary lateral confinement of membrane proteins using single-particle tracking analysis, Biophysical Journal **69**, 989 (1995).

[102] N. Meilhac, L. Le Guyader, L. Salome, and N. Destainville, Detection of confinement and jumps in single-molecule membrane trajectories, Physical Review E **73**, 011915 (2006).
[103] 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 **113**, 2452 (2017).
[104] Z. Chen, L. Geffroy, and J. S. Biteen, NOBIAS: Analyzing anomalous diffusion in single-molecule tracks with nonparametric Bayesian inference, Frontiers in Bioinformatics , 40 (2021).
[105] S. Matsuoka, T. Shibata, and M. Ueda, Statistical analysis of lateral diffusion and multistate kinetics in single-molecule imaging, Biophysical Journal **97**, 1115 (2009).
[106] C. P. Calderon and K. Bloom, Inferring latent states and refining force estimates via hierarchical Dirichlet process modeling in single particle tracking experiments, PLOS One **10**, e0137633 (2015).
[107] A. Weron, K. Burnecki, E. J. Akin, L. Solé, M. Balcerek, M. M. Tamkun, and D. Krapf, Ergodicity breaking on the neuronal surface emerges from random switching between diffusive states, Scientific Reports **7**, 5404 (2017).
[108] T. Wagner, A. Kroll, C. R. Haramagatti, H.-G. Lipinski, and M. Wiemann, Classification and segmentation of nanoparticle diffusion trajectories in cellular micro environments, PLOS One **12**, e0170165 (2017).
[109] P. J. Bosch, J. S. Kanger, and V. Subramaniam, Classification of dynamical diffusion states in single molecule tracking microscopy, Biophysical Journal **107**, 588 (2014).
[110] F. Persson, M. Lindén, C. Unoson, and J. Elf, Extracting intracellular diffusive states and transition rates from single-molecule tracking data, Nature Methods **10**, 265 (2013).
[111] N. Monnier, Z. Barry, H. Y. Park, K.-C. Su, Z. Katz, B. P. English, A. Dey, K. Pan, I. M. Cheeseman, R. H. Singer, *et al.*, Inferring transient particle transport dynamics in live cells, Nature Methods **12**, 838 (2015).
[112] H. Kabbech and I. Smal, Identification of diffusive states in tracking applications using unsupervised deep learning methods, in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)* (IEEE, 2022) pp. 1–4.
[113] X. Qu, Y. Hu, W. Cai, Y. Xu, H. Ke, G. Zhu, and Z. Huang, Semantic segmentation of anomalous diffusion using deep convolutional networks, arXiv preprint arXiv:2305.05618 (2023).
[114] Stochastic Python package, https://stochastic.readthedocs.io/en/stable/.

ACKNOWLEDGMENTS

G.M-G. acknowledges support from the European Union (ERC, QuantAI, Project No. 10105529) and the Austrian Science Fund (FWF) through the SFB BeyondC F7102. M.L. acknowledges support from: ERC AdG NOQIA; Ministerio de Ciencia y Innovación Agencia Estatal de Investigaciones (PGC2018-097027-B-I00/10.13039/501100011033, CEX2019-000910-S/10.13039/501100011033, Plan Nacional FIDEUA PID2019-106901GB-I00, FPI, QUANTERA MAQS PCI2019-111828-2, QUANTERA DYNAMITE PCI2022-132919, Proyectos de I+D+I "Retos Colaboración" QUSPIN RTC2019-007196-7); MICIIN with funding from European Union NextGenerationEU(PRTR-C17.I1) and by Generalitat de Catalunya; Fundació Cellex; Fundació Mir-Puig; Generalitat de Catalunya (European Social Fund FEDER and CERCA program, AGAUR Grant No. 2021 SGR 01452, QuantumCAT \ U16-011424, co-funded by ERDF Operational Program of Catalonia 2014-2020); Barcelona Supercomputing Center MareNostrum (FI-2022-1-0042); EU Horizon 2020 FET-OPEN OPTOlogic (Grant No 899794); EU Horizon Europe Program (Grant Agreement 101080086 — NeQST), National Science Centre, Poland (Symfonia Grant No. 2016/20/W/ST4/00314); ICFO Internal "QuantumGaudi" project; European Union's Horizon 2020 research and innovation program under the Marie-Skłodowska-Curie grant agreement No 101029393 (STREDCH) and No 847648 ("La Caixa" Junior Leaders fellowships ID100010434: LCF/BQ/PI19/11690013, LCF/BQ/PI20/11760031, LCF/BQ/PR20/11770012, LCF/BQ/PR21/11840013). Views and opinions expressed in this work are, however, those of the author(s) only and do not necessarily reflect those of the European Union, European Climate, Infrastructure and Environment Executive Agency (CINEA), nor any other granting authority. Neither the European Union nor any granting authority can be held responsible for them. R.M. acknowledges DFG grant ME 1535/12-1. D.K. acknowledges funding from the National Science Foundation grant 2102832. G.V., H.B., J.P., and B.M. acknowledge funding from ERC StG ComplexSwimmers. C.M. acknowledges support through grant RYC-2015-17896 funded by MCIN/AEI/10.13039/501100011033 and "ESF Investing in your future", grants BFU2017-85693-R and PID2021-125386NB-I00 funded by MCIN/AEI/10.13039/501100011033/ and FEDER "ERDF A way of making Europe".

COMPETING INTEREST

The authors declare no competing interests.

Method	Output level	Type of Approach	Diffusion parameters	Diffusion classes
Hidden Markov method with measurement noise propagation [87, 88]	Single trajectory	Stats	D	confined/Brownian
Hidden Markov method and simulations [43]	Single trajectory	Stats	D, α	
Supervised trajectory segmentation algorithm with supported vector machine [89]	Single trajectory	ML		confined/drift/directed
Deep learning and moment scaling spectrum [90]	Single trajectory	ML + Stats	D	immobile/Brownian (fast/slow)
Piecewise linear approximation [91]	Single trajectory	Stats	velocity	
Detection of velocity and diffusion coefficient[92, 93]	Single trajectory	Stats	D, velocity	Brownian/directed
Back-propagation neural network [94]	Single trajectory	ML		confined/directed
Divide-and-conquer moment scaling spectrum [95]	Single trajectory	Stats	D, α, confinement radius	immobile/confined/Brownian/directed
Local convex hull [96]	Single trajectory	Stats	D,α	
Recurrence analysis [97]	Single trajectory	Stats	D,α	
Gamma mixture and a hidden Markov method [98]	Single trajectory	Stats + ML	D	
Superstatistical framework [99]	Single trajectory	Stats	activity, persistence	
Probability distribution of square displacements [100]	Ensemble	Stats	D, α	immobile/confined/Brownian/directed
Probability of transient confinement zones [101]	Single trajectory	Stats		confined/Brownian
Probability of jumps and transient confinement zones [102]	Single trajectory	Stats		confined/Brownian + jumps
Packing coefficient [103]	Single trajectory	Stats		immobile/confined/Brownian
Nonparametric Bayesian inference [104]	Ensemble	Stats + ML	D	
Displacement distribution and autocorrelation of squared displacements [105]	Ensemble	Stats	D	
Hierarchical Dirichlet process modeling [106]	Single trajectory	Stats	force	
Sliding window of time-averaged MSD [107]	Single trajectory	Stats	D,α	
Random forest [108]	Single trajectory	Stats		confined/Brownian/directed/anomalous
Gyration quantification and a Bayesian statistics [109]	Single trajectory	Stats	D	fast/slow
Variational Bayesian treatment of hidden Markov method [110]	Ensemble	Stats	D	
Bayesian model selection of hidden Markov method [111]	Single trajectory	Stats	D	Brownian/directed
Measurement of anomalous diffusion using recurrent neural networks [37]	Single trajectory	ML	α	
Pointwise diffusion properties with transformers [45]	Single trajectory	ML	D and α	
Graph-neural network with attention [40]	Single trajectory & Ensemble	ML	D, α	
Recurrent neural network-based autoencoders [112]	Single trajectory	ML	D, α	
Semantic segmentation with convolutional networks [113]	Single trajectory	ML	α	

TABLE I. **Methods for the analysis of heterogeneous diffusion.** Methods for the analysis of heterogeneous diffusion are classified based on the level of output provided (single-trajectory methods vs. ensemble methods) and on the type of approach (machine learning (ML), classical statistics (Stats), or a combination of the two). We also indicate the diffusion parameters that each method estimates and the diffusion classes that each method uses to classify the trajectories. Only methods accounting for switches of diffusive behavior within the same trajectory are included.

Experiment	Model	Numb. of states	State	K (pixel^2/frame$^\alpha$)		α distrib.		Diffusion class	Model-specific parameters
				μ	σ	μ	σ		
0	SSM	1	1	1	0.01	0.5	0.01	2	-
1	SSM	1	1	0.1	0.01	1.9	0.01	3	-
2	MSM	2	1	1	0.01	1.5	0.01	2	$M = \begin{pmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{pmatrix}$
			2	0.05	0.01	0.5	0.01	2	
3	MSM	3	1	1	0.01	1.5	0.01	2	$M = \begin{pmatrix} 0.98 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 \\ 0.01 & 0.01 & 0.98 \end{pmatrix}$
			2	0.5	0.01	0.5	0.01	2	
			3	0.01	0.01	0.75	0.01	2	
4	QTM	2	1	1	0.01	0.8	0.01	2	$r_t = 0.6,\ N_t = 300,\ P_b = 1,\ P_u = 0.01$
			2	0	0	0	0	0	
5	QTM	2	1	1	0.01	1.5	0.01	2	$r_t = 1.0,\ N_t = 150,\ P_b = 1,\ P_u = 0.05$
			2	0	0	0	0	0	
6	DIM	2	1	1	0.01	1.2	0.01	2	$r = 0.6,\ N = 100,\ P_b = 1,\ P_u = 0.01$
			2	1	0.01	0.8	0.01	2	
7	DIM	2	1	1	0.01	1.2	0.01	2	$r = 1.0,\ N = 80,\ P_b = 1,\ P_u = 0.01$
			2	3	0.01	0.5	0.01	2	
8	TCM	2	1	1	0.01	0.8	0.01	2	$r_c = 5,\ N_c = 30,\ T = 0.1$
			2	1	0.01	0.4	0.01	1	
9	TCM	2	1	1	0.01	1	0.01	2	$r_c = 10,\ N_c = 30,\ T = 0.0$
			2	0.1	0.01	1	0.01	1	

TABLE II. **Parameters of the pilot dataset.** Example of a dataset composed of 10 experiments reproducing the models of diffusion employed in the 2$^{\text{nd}}$ Anomalous Diffusion Challenge: *single-state model* (SSM); *multi-state model* (MSM); *dimerization model* (DIM); *transient-confinement model* (TCM); *quenched-trap model* (QTM). The diffusion class correspond to 0 = immobile, 1 = confined, 2 = free (unconstrained), 3 = directed. The parameters specific to each model are: for MSM, the transition matrix M; for QTM, the trap radius r_{t}, the number of traps N_{t}, the probability of trapping P_{b} and untrapping P_{u}; for DIM, the interactions radius r, the number of particles N, the probability of binding P_{b} and unbinding P_{u}; for TCM, the compartment radius r_{c}, the number of compartments N_{c}, and the boundary transmittance T. Simulations are provided in generalized units (i.e., pixels and frames) that can be rescaled to meaningful temporal and spatial scales. See Appendix A for further details. For all experiments, we simulated $N = 100$ particles (Experiment 7 has $N = 80$) in a box of size $L = 230$ pixel with a FOV size $L_{FOV} = 128$ pixel, and a maximum trajectory length of 200 frames. For Track 1, movies were rendered using a $\text{FWHM}_{\text{PSF}} = 2.1$ pixel and a SNR = 7.1. For Track 2, trajectories were corrupted with Gaussian localization noise with $\sigma_N = 0.12$ pixel.

Appendix A: Simulations of diffusion and interaction models

Trajectories are simulated according to a 2-dimension fractional Brownian motion (FBM) [54]. FBM is a continuous-time Gaussian process $B_H(t)$ with stationary increments and a covariance function $E[B_H(t)B_H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$, where H represents the Hurst exponent and is related to the anomalous diffusion exponent α as $H = \alpha/2$ [54]. The FBM features three regimes: one where the increments are positively correlated ($1/2 < H < 1$, i.e., $1 < \alpha < 2$, superdiffusive); one where the increments are negatively correlated ($0 < H < 1/2$, i.e., $0 < \alpha < 1$, subdiffusive); and one where the increments are uncorrelated ($H = 1/2$, i.e., $\alpha = 1$, diffusive Brownian motion).

The models included in the challenge describe trajectories where diffusion properties are piecewise constant along segments of varying duration T_s and undergo sudden changes. To obtain a trajectory segment of length T_s with given anomalous diffusion exponent α and generalized diffusion coefficient K, a set of $T_s - 1$ displacements for each dimension are sampled from a fractional Gaussian noise generator [114]. The displacements are then standardized to have variance $\sigma^2 = 2K\Delta t$, where Δt is the sampling time.

Simulations are performed considering particles diffusing in a square box of size L with reflecting boundary conditions. However, to avoid boundary effects, the fields of view used for the challenge datasets correspond to a square region of size $L_{\mathrm{FOV}} \ll L$ within the central part of the original box (Fig. 2(b)).

For Track 1, trajectory coordinates are used as sub-pixel localizations of individual particles to simulate movie frames as in single-molecule fluorescence experiments [5]. Each particle has a random intensity I_i that corresponds to the total number of photons collected by the detector. I_i is drawn from a uniform distribution in the interval $[I_{\min}, I_{\max}]$ and fluctuates over time according to a normal distribution with mean I_i and standard deviation σ_I. Each particle is rendered as a diffraction-limited spot using an Airy disk as a point-spread function (PSF) with full width at half maximum $\mathrm{FWHM}_{\mathrm{PSF}} = 2.1$ px. A constant background of $I_{\mathrm{bg}} = 100$ counts is added to each frame. Images are corrupted with Poisson noise.

For Track 2, trajectory coordinates are corrupted with noise from a Gaussian distribution with zero mean and standard deviation σ_N to take into account the finite localization precision obtained in tracking experiments.

All the models share a set of parameters required for the simulations that are described here. Model-specific parameters are defined when describing the details of the models in the following sections.

- $[K_1, K_2, \ldots, K_n]$: average values of the (Gaussian) distribution of the generalized diffusion coefficient for each of the n diffusive states considered in a given experiment, with support $[10^{-12}, 10^6]$ pixel$^2 \cdot$ frame$^{-\alpha}$.
- $[\sigma_{K_1}, \sigma_{K_2}, \ldots, \sigma_{K_n}]$: standard deviations of the (Gaussian) distribution of the generalized diffusion coefficient for each of the n diffusive states considered in a given experiment. If not provided, it is considered to be equal to 0 (i.e., the distribution is $\delta(K - K_i)$).
- $[\alpha_1, \alpha_2, \ldots, \alpha_n]$: average values of the (Gaussian) distribution of the anomalous diffusion exponent for each of the n diffusive states considered in a given experiment, with support $(0, 2)$.
- $[\sigma_{\alpha_1}, \sigma_{\alpha_2}, \ldots, \sigma_{\alpha_n}]$: standard deviations of the (Gaussian) distribution of the anomalous diffusion exponent for each the n diffusive states considered in a given experiment. If not provided, it is considered to be equal to 0 (i.e., the distribution is $\delta(\alpha - \alpha_i)$).
- L: size of the box where trajectories are simulated with reflecting boundary conditions.
- L_{FOV}: size of the box defining the FOV used for the challenge datasets. The same particles can enter and exit the FOV over time but, for evaluation purposes, they will be considered as generating different trajectories.
- Δt: sampling time at which the original motion of the particle is tracked. For the challenge datasets, we consider $\Delta t = 1$.
- T: duration of the recording over each FOV, given as the number of time steps Δt. It also corresponds to the maximum trajectory duration. For the challenge, we set $T = 500$;
- $T_{\min}$: minimum duration of a trajectory to be included in the dataset. For the challenge, we use $T = 20$;
- I_{bg} (Track 1): background level of noise (counts) used in the simulation of videos.
- $\mathrm{FWHM}_{\mathrm{PSF}}$ (Track 1): full width at half maximum in pixels of the point-spread function used to render fluorescent particles.
- I_{tot} (Track 1): mean value in counts of the total fluorescence collected for the detected particles.

- σ_{tot} (Track 1): standard deviation in counts of the distribution of total fluorescence collected for the detected particles.
- I_{peak} (Track 1): mean value in counts of the peak fluorescence collected for the detected particles. Can be calculated as $I_{\text{peak}} = I_{\text{tot}} \frac{4 \ln 2}{\pi \text{FWHM}_{\text{PSF}}{}^2}$
- SNR (Track 1): typical signal-to-noise ratio of the movies, calculated as the average peak intensity over the standard deviation of the noise [51] and thus equal to

$$SNR = \frac{I_{\text{peak}}}{\sqrt{I_{\text{peak}} + I_{\text{bg}}}}. \tag{A1}$$

- σ_N (Track 2): standard deviation of the Gaussian localization noise used to corrupt trajectory coordinates.
- t_{min}: minimum distance between changepoints, corresponding to the minimum amount of time that a particle spends in a state. Shorter segments are eliminated by smoothing the time trace of the state label using a majority filter with a window of 5 steps. For the challenge, we set $t_{\text{min}} = 3$.;

A schematic representation of each of the models presented below is shown in Fig. 2(a).

1. Model 1 - Single-state model (SSM)

This model simply corresponds to particles diffusing according to FBM with constant generalized diffusion coefficient K and anomalous diffusion exponent α. For each trajectory, a value of K and a value of α are sampled from the corresponding distribution. Data corresponding to these models are necessary to establish the false positive rate of the methods toward the detection of changes of diffusion properties.

2. Model 2 - Multi-state model (MSM)

The multi-state model is a Markov model describing particles undergoing FBM whose diffusion properties can change at random times. The number of states S is fixed for a given experiment as are the parameters defining the distributions of K and α for each state. For each trajectory, S values of α and S values of K are sampled from the distribution of the corresponding states, i.e., one per state. At every time step, a diffusing particle has a given probability to undergo a change in one of its diffusive parameters (either α or K). The probability of switching is given by a transition matrix M. Namely, $M_{i,j}$ is the probability of switching from state i to state j at each time step. In the same sense, $M_{i,i}$ is the probability of remaining in state i. The residence time in a given state i can be directly calculated from the previous probability as

$$\tau_i = \frac{1}{\sum_{j \neq i} M_{ij}} = \frac{1}{1 - M_{ii}}. \tag{A2}$$

Model 2 (MSM) parameters

- M: transition matrix between diffusive states.

3. Model 3 - Dimerization (DIM)

This model considers the case in which dimerization, i.e., the transient binding of two particles, may occur and produce changes in the diffusion properties of both particles. In particular, we consider the case of N circular particles of radius r. For each trajectory, a value of α and a value of K are sampled from the corresponding distributions associated with the monomeric state. If two particles are at a distance $d < 2r$, then they have a probability P_{b} of binding. The two particles forming a dimer move with equal displacements, according to a generalized diffusion coefficient K and an anomalous diffusion exponent α drawn from the distributions associated with the dimeric state. At each time step, the dimer has a probability P_{b} of breaking its bond, freeing the two particles to go back to their original motion parameters. The particles cannot form any new dimer until taking a new step. Only dimers are allowed and subsequent hits with other particles will not affect either the particles or the dimers.

Model 3 (DIM) parameters

- N: number of diffusing particles in the box of size L.
- r: interaction radius, corresponding to the radius of the diffusing particles.
- P_b: probability that two particles bind to form a dimer in each time step. For this to happen, the particles must be at a distance $d < 2r$.
- P_u: probability that a dimer breaks up at each time step so that the two particles go back to diffusing independently.

4. Model 4 - Transient-confinement model (TCM)

This model considers an environment with N_c circular compartments of radius r_c. The compartments are distributed randomly throughout the environment such that they do not overlap. We consider that the compartments are *osmotic*, i.e., a particle reaching their boundary from the exterior has a probability 1 of entering them, but a particle reaching the boundary from the interior of a compartment has a probability T of exiting it (and $1 - T$ of being reflected back to the interior of the compartment). The diffusion inside and outside the compartment is different, hence defining two diffusive states. For each trajectory, two values of α and two values of K are sampled from the corresponding distributions, corresponding to the motion outside and inside the compartments.

Model 4 (TCM) parameters

- N_c: number of compartments in the box of size L.
- r_c: radius of the compartments.
- T: transmittance of the boundary. Probability that a particle reaching the boundary from inside the compartment exits the compartment.

5. Model 5 - Quenched-trap model (QTM)

This model considers the diffusion of particles in an environment with N_t immobile traps of radius r_t. The values of α and K are sampled for each trajectory from the corresponding distributions and define its unrestrained motion. A particle that enters the domain defined by a trap has a probability P_b of binding to the trap and, hence, getting temporarily immobilized ($K = 0$, $\alpha = 0$). At each time step, a trapped particle has a probability P_u of unbinding and being released from the trap, going back to its unrestrained motion. A particle cannot be trapped again until taking a new step.

Model 5 (QTM) parameters

- N_t: number of traps in the box of size L.
- r_t: radius of the traps.
- P_b: probability that a particle binds to a trap and gets immobilized. For that to happen, a particle must be at a distance $d < r_t$ from the trap.
- P_u: probability that a trapped particle unbinds from a trap and starts diffusing independently at each time δt.

LIST OF ACRONYMS

RMSE: root mean square error

MAE: mean absolute error

MSLE: mean squared logarithmic error

JSC: Jaccard similarity coefficient

FP: false positive

TP: true positive

FN: false negative

CP: changepoint

FOV: Field of view

SSM: Single-state model

MSM: Multi-state model

DIM: Dimerization model

TCM: Transient-confinement model

QTM: Quenched-trap model