Robust evaluation of diffusion coefficient against displacement threshold parameter of single particle tracking algorithm

Reiji Motohashi[1], *Itsuo Hanasaki*[1] ✉, *Yuto Ooi*[1], *Yu Matsuda*[2]

[1]*Institute of Engineering, Tokyo University of Agriculture and Technology, Naka-cho 2-24-16, Koganei-shi, Tokyo 184-8588, Japan*
[2]*Institute of Materials and Systems for Sustainability, Nagoya University, Furo-cho, Chikusa, Nagoya, Aichi 464-8603, Japan*
✉ *E-mail: hanasaki@cc.tuat.ac.jp*

Published in Micro & Nano Letters; Received on 21st January 2017; Revised on 14th March 2017; Accepted on 20th March 2017

Single particle/molecule tracking is widely used for the evaluation of diffusion coefficients using diverse tracking algorithms. Many of them require the subtle decision of displacement threshold parameter to enable the appropriate tracking, though it depends on both of the diffusion coefficient itself and the measurement system. A simple and highly transferable technique is proposed to overcome this difficulty based on the evaluation of diffusion coefficient from the peak position of the distribution of squared displacements in the logarithmic scale. In combination with its linear dependence on time, this protocol is remarkably robust against the too large value of the displacement threshold including that covers the whole image. Furthermore, the proposed technique is compatible with many of the existing algorithms by construction.

1. Introduction: Sufficiently small particles in fluid exhibit Brownian motion due to the collisions from the surrounding fluid molecules. The relative importance of Brownian motion compared with the flow field is drastically pronounced in the lab-on-a-chip systems, in general, where the analyte objects are subject to intensive Brownian motion. The Brownian motion leads to the diffusion, which is evaluated by the diffusion coefficient. The importance of Brownian motion is ubiquitous [1–5]. Diffusion coefficients are generally obtained from particle trajectories when they are available. There are several other techniques to evaluate diffusion coefficients of particles in fluids, but the trajectories include more information besides diffusion coefficients. Namely, it is possible to evaluate diffusion anisotropy of prolate particles [6, 7], possible motility and its characteristics [8], and the evaluation of surface effects [9, 10] from the ensemble of trajectories. Therefore, single particle/molecule tracking (SPT/SMT) has been a standard and effective approach to analyse the small particle dynamics suspended in fluids.

Many of the particle tracking algorithms consists of mainly two parts: (i) the detection of particles of interest in each image captured through the camera connected to the microscopy system and (ii) the identification of the same particles between the time-sequential two images [11–13]. The process of (ii) often requires the threshold parameter for possible displacements in different algorithms [14–16]. It precludes the possibility of different particles far away from the particle of interest is examined for the candidate to regard identical. If the frame rate is sufficiently fast and/or the particle motion is sufficiently slow, there is little possibility of confusion of different particles as identical. However, the diffusions of particles are the unknown quantity to evaluate by the SPT/SMT, and the frame rate of the camera is always limited. Even when the sufficiently high frame rate is available, the data storage problem sometimes leads to the compromise of frame rate to cover the long-term dynamics. On the other hand, too small value of this displacement threshold causes the failure of displacement detection, and leads to the evaluation of unreasonably small diffusion coefficient than the actual dynamics. The specificity of the optimal threshold parameter for displacements usually calls for the empirical conclusion *a posteriori* on the value based on the exhaustive parameter study. The ideal solution is the development of a technique to evaluate diffusion coefficient without relying on this threshold parameter. Furthermore, it is even more convenient if one becomes free from the subtlety of this threshold decision even using the existing algorithms.

One of the essential aspects of Brownian motion is the distribution of random displacements. A particle with a specific magnitude of diffusion coefficient can exhibit a large displacement at some time step and then an extremely small displacement. This stochasticity is usually incompatible with a constant threshold value for each of the displacements for the distinction. In other words, the Brownian displacements are the data to be evaluated statistically. We have previously proposed a new technique to distinguish the random displacements without dominant effect of adsorption on the substrate surface and the displacements dominated by adsorptive effect [9]. We have realised it by the change of scale to logarithm. The combination with distinction of ensemble subsets based on the consecutively tracked durations even clarified that the adsorption-dominated motion of short ssDNA on the substrate surface actually exhibited the diffusion [9]. The method is extremely simple yet it has never been reported in spite of the vast amount of SPT/SMT studies in the vicinity of solid/fluid interface often using total internal reflection fluorescence (TIRF). We have found that the principle of logarithmic scale for the evaluation of random displacements can be extended and applied to the distinction of the '*false link*', i.e. the mis-interpretation of different particles in the sequential images as identical (see Fig. 1), from the real displacements. In this Letter, we report our novel technique for the evaluation of diffusion coefficients without being affected by the displacement threshold parameter in the existing particle tracking algorithm.

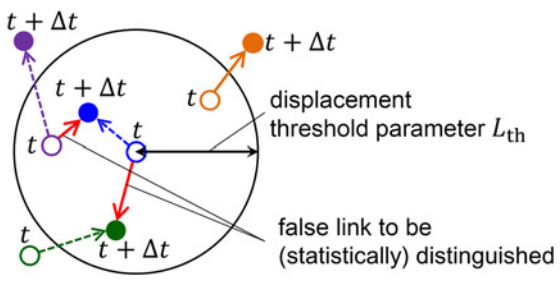

Fig. 1 *Schematic diagram of the false link in the particle tracking*

2. Theoretical basis and our novel technique: The diffusion coefficient D in the diffusion equation can be derived from the particle trajectories as follows:

$$D = \lim_{t\to\infty} \frac{1}{2n_{\mathrm{d}}t} \langle |\boldsymbol{r}(t) - \boldsymbol{r}(0)|^2 \rangle \tag{1}$$

where n_{d} is the dimension of observation, $\boldsymbol{r}(t)$ is the position of particle at time t, and $\langle \dots \rangle$ indicates the ensemble average. While the analytical expression use infinite limit, only finite set of data is available for the experimental (and numerical) trajectories. Then, there can be some options in the evaluation of the diffusion coefficient. One of them is the frame-based diffusion coefficient D_{FB} [9]

$$D_{\mathrm{FB}} = \frac{1}{2n_{\mathrm{d}}\Delta t \sum_{i=1}^{N_{\mathrm{I}}} (N_{\mathrm{F}i} - 1)} \sum_{i=1}^{N_{\mathrm{I}}} \sum_{j=1}^{N_{\mathrm{F}i}-1} \left| \boldsymbol{r}_i(t_{j+1}) - \boldsymbol{r}_i(t_j) \right|^2 \tag{2}$$

where N_{I} is the total number of particles, $\boldsymbol{r}_i(t_j)$ is the position of the ith particle at the jth time step. Δt is the frame interval, $N_{\mathrm{F}i}$ is the number of consecutive frames for the ith particle. D_{FB} treats every displacement with an equal weight in the evaluation of the diffusion coefficient. If $N_{\mathrm{F}i}$ is significantly different among different i and the characteristics depends on the i, the particles with large $N_{\mathrm{F}i}$ has larger weight in the evaluation of D_{FB}. If one intends to avoid it, individual-based diffusion coefficient D_{IB} can be employed [9]

$$D_{\mathrm{IB}} = \frac{1}{N_I} \sum_{i=1}^{N_{\mathrm{I}}} \left[\frac{1}{2n_{\mathrm{d}}\Delta t (N_{\mathrm{F}i} - 1)} \sum_{j=1}^{N_{\mathrm{F}i}-1} \left| \boldsymbol{r}_i(t_{j+1}) - \boldsymbol{r}_i(t_j) \right|^2 \right] \tag{3}$$

D_{IB} is further decomposed into $D_{\mathrm{IS}i}$

$$D_{\mathrm{IS}i} = \frac{1}{2n_{\mathrm{d}}\Delta t (N_{\mathrm{F}i} - 1)} \sum_{j=1}^{N_{\mathrm{F}i}-1} \left| \boldsymbol{r}_i(t_{j+1}) - \boldsymbol{r}_i(t_j) \right|^2 \tag{4}$$

which is the diffusion coefficient for each particle. Then, it is useful to define a quantity $D_{\mathrm{ISF}ij}$ with a dimension of diffusion coefficient for each of the displacements

$$D_{\mathrm{ISF}ij} = \frac{\left| \boldsymbol{r}_i(t_{j+1}) - \boldsymbol{r}_i(t_j) \right|^2}{2n_{\mathrm{d}}\Delta t} \equiv \frac{S_{ij}}{2n_{\mathrm{d}}\Delta t} \tag{5}$$

which we call $D_{\mathrm{ISF}ij}$ the 'diffusion coefficient for individual sample and frame' for descriptive purpose, though this is not a converged statistical quantity, but corresponds to a squared displacement S_{ij} just in a unit of diffusion coefficient. We propose that the diffusion coefficient can be evaluated by the peak position of a histogram for the logarithm of $D_{\mathrm{ISF}ij}$ even when too large threshold value for the possible displacement is assumed in the tracking algorithm.

3. Experimental details: We validate our new technique with the experimental data. We employ acrylic particles with a nominal diameter of 0.9 µm and coefficient of variation (CV) value of ca. 9% [17] (Chemisnow MX80H-3wT, Soken Chemical & Engineering Co., Ltd.) dispersed in water at a concentration of 0.1 wt%. The dispersion is sealed in a cylindrical volume with a diameter of 8 mm and a height of 100 µm made of top and bottom walls of cover glass and surrounded by a silicone sheet. We measure the central region in the horizontal location at a height of ca. 25 µm from the bottom of the cover glass. We capture the motion of particles using the camera (Zyla 5.5, Andor Technology Ltd.) that has a pixel pitch of 6.5 µm from the inverted microscope (IX73, Olympus Co. Ltd.) with objective lens of ×20. The frame rate of the camera was 200 fps and 1000 frames were captured at the image size of 512 × 512 px. While this is sufficient for the main analysis to obtain the diffusion coefficient, we also obtained the movie data at 100 fps for 500 frames, and 10 fps for 500 frames for comparison. We used the SPT algorithm of Sbalzarini and Koumoutsakos [12] with input parameters as shown in Table 1 and varied the threshold parameter for the displacements from 1 to 730 px, where 730 px covers the whole domain of the images. The conditions for the movie data are summarised in Table 2. The diffusion coefficient of a spherical particle with diameter $2a$ in a fluid with a viscosity η at an absolute temperature T under sufficiently dilute condition can be estimated as

$$D_{\mathrm{SE}} = \frac{k_{\mathrm{B}}T}{6\pi\eta a} \tag{6}$$

where k_{B} is the Boltzmann's constant. We validate our technique by comparison with the estimation based on the Stokes–Einstein estimation. The specific value was $D_{\mathrm{SE}} = 5.2 \times 10^{-13}$ m^2/s using the value of water viscosity at 23°C (by linear interpolation from the literature) corresponding to the experimental condition.

Table 1 Parameters for the SPT algorithm of Sbalzarini and Koumoutsakos [12]

Parameter	Value
particle radius, px	3
intensity percentile, %	0.02
cut-off score	0
link range, frames	1

Table 2 Combination of frame rate, number of frames, and the displacement threshold parameter L_{th}

Trajectory label	Frame rate, fps	Number of frames	Displacement threshold, L_{th}, px
A	200	1000	1
B	200	1000	50
C	200	1000	100
D	200	1000	730
E	100	500	730
F	50	250	730
G	10	500	730

4. Results and discussion: In most of the tracking algorithms, the displacement threshold L_{th} needs to be determined in an appropriate range, so that the identification of the same particles in the adjacent frames are correctly determined. On the one hand, too large L_{th} can cause the confusion of different particles as identical. On the other hand, too small value of L_{th} leads to missing the trajectories of the particles captured in the movie data. Depending on the specification of the system, it can lead to the different value of the evaluated diffusion coefficient compared with the case when sufficient fractions of the trajectories are correctly determined. We show that our technique can evaluate the diffusion coefficient from the trajectories based on too large L_{th}. Fig. 2 shows the distribution of $D_{\mathrm{ISF}ij}$ for trajectory D in Table 2, where L_{th} covers the whole range of the image corresponding to the absence of restrictions by L_{th} in the trajectory evaluation. Fig. 2*a* shows the long tail of $D_{\mathrm{ISF}ij}$ and the trajectories include those based on the false link between the different particles of the adjacent frames. It is not clear which displacement is correct or wrong. On the other hand, the logarithmic measure [9] clarifies the difference of $D_{\mathrm{ISF}ij}$ from

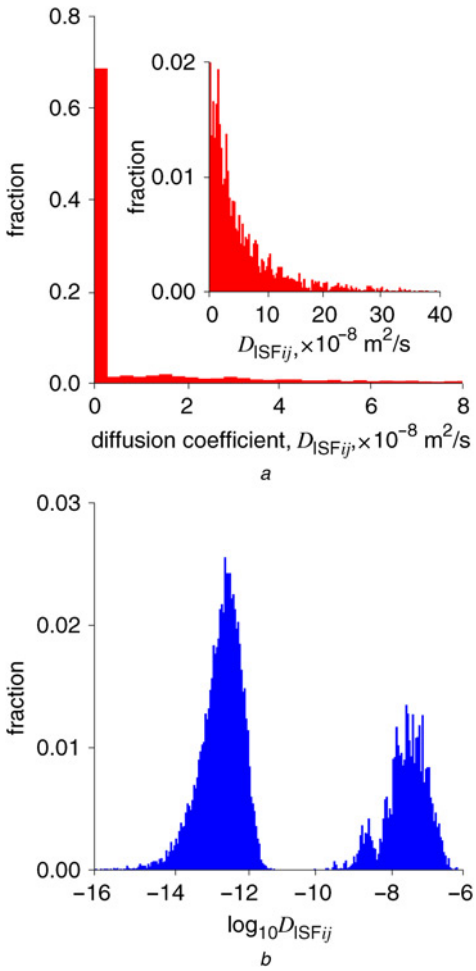

Fig. 2 *Distribution of D_{ISFij} for trajectory D of Table 2 shown by*
a Ordinary histograms
b That for logarithmic scale of the quantity. The number of samples for the histogram itself is 14,559. Note that the vertical axis of the inserted figure in (a) does not cover the whole range of the data

the actual displacements and the confused ones. The first peak reasonably agrees with D_{SE} (6). The actual value of diffusion coefficient can be evaluated by $D_{ISFP} = 10^p$ using the value of the peak position p in the histogram.

Then, the next question is how we can determine the peak that corresponds to the actual dynamics from multiple ones (usually two). It can be clarified by the analysis of time dependence of the ensemble of displacements that the normal diffusion exhibits. Fig. 3 shows the dependence on the peak S_P of distribution for the squared displacements [S_{ij} in (5)]. The first peak corresponding to the actual Brownian motion shows a linear dependence of squared displacements on time. This is the basis of the evaluation of diffusion coefficients by (1). On the other hand, the false-linked trajectories exhibit roughly constant value after some transient time scale. The appearance into and disappearance out of the microscopy observation depth do not have time correlation among the different particles, and hence the link between the frames does not have time correlation either. Thus, there is no ambiguity in the determination of diffusion coefficients. The clear linearity on the extracted actual dynamics also suggests that the false link of trajectory is negligible in the positioning of the peak for actual dynamics. The use of linearity of mean squared displacements (MSDs) against time limits the application of this technique within the normal diffusion. In other words, our technique can be extended to the subdiffusion and superdiffusion if we simply take into account the nonlinearity of MSD in such situations.

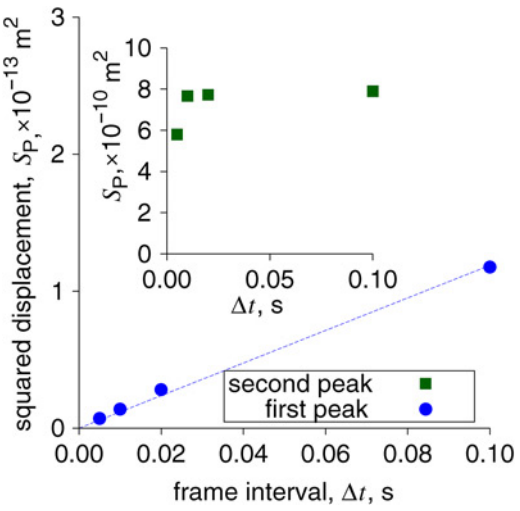

Fig. 3 *Time dependence of the peak S_P of the squared displacements S_{ij} determined by the use of the logarithmic scale. The trajectory data sets are based on the D, E, F, and G in Table 2*

The diffusion coefficient is one of the non-trivial quantities to evaluate precisely. Besides the possible errors in the experimental origin, the value based on the proposed technique fundamentally depends on the bin width of the histogram that is employed in the determination of peak position for the logarithm of the distribution of squared displacements. Therefore, we have examined the dependence of bin width on the evaluated value of diffusion coefficient. Fig. 4 shows that the value of diffusion coefficient converges with the reduction of the bin width (or the increase of the inverse B^{-1} of bin width B). The span of each error bar in this figure corresponds to the range of possible diffusion coefficients based on the two edge values of the bin in the histogram where the count shows the peak.

We have confirmed that our proposed approach works even without the presence of displacement threshold L_{th}, but it is also worth testing the range of L_{th} shorter than the diagonal line of the microscopy image does not deteriorate the performance. Fig. 5 shows the evaluated diffusion coefficients as a function of L_{th} for different definitions of the diffusion coefficients. The smallest test value for L_{th} is 1 px, which is larger than 3σ, where σ denotes

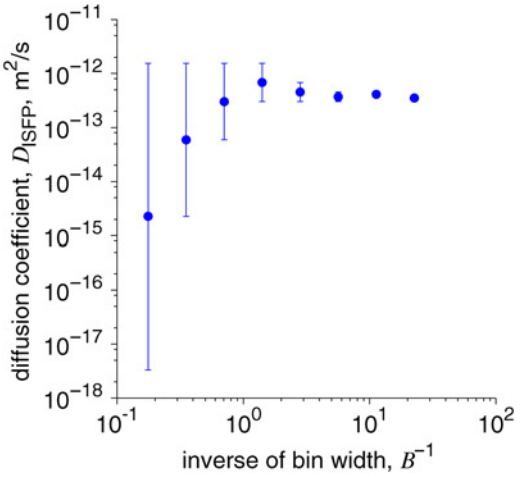

Fig. 4 *Dependence of diffusion coefficients D_{ISFP} on the bin width B of histogram for the evaluation of peak position. The plotted points are based on trajectory D in Table 2. The narrowest bin width in this figure corresponds to the condition of Fig. 2b*

doi: 10.1049/mnl.2017.0044

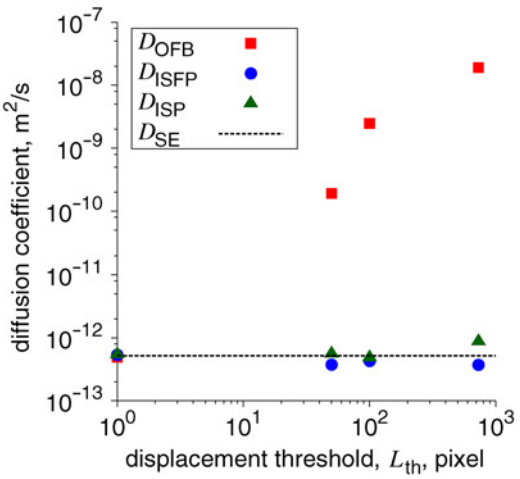

Fig. 5 *Evaluated diffusion coefficients as a function of the displacement threshold L_{th}. The plotted points are based on the trajectories A, B, C, and D in Table 2. Comparison is based on the resolution of 64 bins on the obtained numerical range of data for the histogram of diffusion coefficients*

the standard deviation of the displacements estimated by D_{SE}. Thus, $L_{th} = 1$ px is also an acceptable condition for the orthodox evaluation procedure of the diffusion coefficient. The overall D_{FB} denoted by D_{OFB}, where all the displacements including those based on the false link are used in the evaluation, shows the dependence on L_{th}. On the other hand, the diffusion coefficient D_{ISFP} based on the peak position of the logarithm of squared displacements does not depend on L_{th}. Now, it is fully conclusive that the diffusion coefficient based on our proposed method is not affected by the too large L_{th}. The value of diffusion coefficient itself also shows reasonable agreement with the estimation D_{SE} by the Stokes–Einstein relation.

We also examined the case D_{ISP}, where distributions of diffusion coefficients for individual samples are employed for the histogram construction. The use of (4) instead of (5) causes the difference in the apparent number of samples in the histogram, which affects the possible resolution in the determination of the peak position. Therefore, we have compared the performance of these two types of evaluation based on the same resolution of the histograms. The possible resolution of D_{ISFP} tends to be finer than D_{ISP}, which is

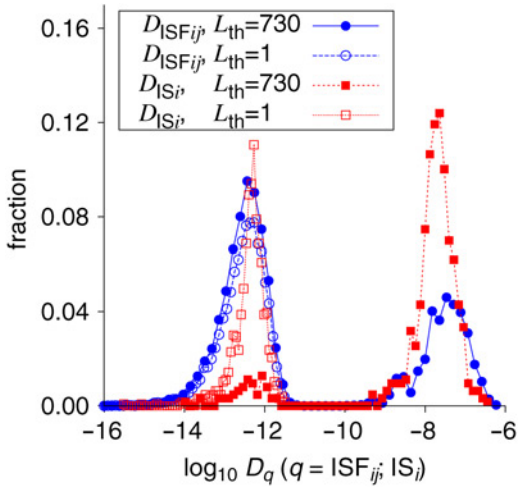

Fig. 6 *Sensitivity of diffusion coefficient histogram on the displacement threshold L_{th}. The D_{ISi} and D_{ISFij} are defined in (4) and (5), respectively. Comparison is based on the resolution of 64 bins on the obtained numerical range of data*

one of the advantages of D_{ISFP} against D_{ISP}. Fig. 5 shows that D_{ISP} tends to be slightly larger than D_{ISFP}. This is attributed to the possibility of including the false link in D_{ISi} for each i. For example, consider the case where particle α gets out of the observation domain in the thickness direction (i.e. perpendicular to the observation plane), and particle $\beta \neq \alpha$ gets into the observation domain in the thickness direction at sufficiently close location in the observation plane. The particles α and β might be regarded as identical while it is not the case in reality. If it happens, the trajectories of α and β are merged into a single one including the unreal displacement based on the false link. On the other hand, D_{ISFP} is based on the single displacements. The unreal trajectories based on the false link are statistically separated in the histogram of the squared displacements. This is effectively confirmed from Fig. 6. The fraction of samples corresponding to the actual dynamics substantially decreases in the case of D_{ISi} when L_{th} is effectively absent, whereas it is not strongly affected in the case of D_{ISFij}.

5. Concluding remarks: There exists diverse types of particle tracking algorithm, and most of them require input parameters including the displacement threshold to enable the analysis. Our proposed approach solves the difficulty of deciding the appropriate value of displacement threshold. This technique is remarkably simple: just plotting the histogram of logarithm for the squared displacements and reading the peak position. Furthermore, this approach is applicable to the wide range of the existing algorithms. Therefore, it is unnecessary to implement or get used to a new tracking algorithm in many cases, but our approach empowers the ability not to be spoiled by the too large displacement threshold parameter in the particle tracking. Once diffusion coefficient is successfully evaluated, it can be used for the further analysis of the particle dynamics.

6. Acknowledgment: This work was partly supported by a Grant-in-Aid for Young Scientists (A), no. 26709008.

7. References

[1] Ma X., Chen W., Wang Z., *ET AL.*: 'Test of the universal scaling law of diffusion in colloidal monolayers', *Phys. Rev. Lett.*, 2013, **110**, pp. 078302-1–078302-5

[2] Peng Y., Chen W., Fischer T.M., *ET AL.*: 'Short-time self-diffusion of nearly hard spheres at an oil–water interface', *J. Fluid Mech.*, 2009, **618**, pp. 243–261

[3] Sickert M., Rondelez F.: 'Shear viscosity of Langmuir monolayers in the low-density limit', *Phys. Rev. Lett.*, 2003, **90**, pp. 126104-1–126104-4

[4] Hanasaki I., Uehara S., Kawano S.: 'Role of time scales for the non-Gaussianity of the Brownian motion combined with intermittent adsorption', *J. Comput. Sci.*, 2015, **10**, pp. 311–316

[5] Marín Á.G., Gelderblom H., Lohse D., *ET AL.*: 'Order-to-disorder transition in ring-shaped colloidal stains', *Phys. Rev. Lett.*, 2011, **107**, p. 085502

[6] Hanasaki I., Isono Y.: 'Detection of diffusion anisotropy due to particle asymmetry from single-particle tracking of Brownian motion by the large-deviation principle', *Phys. Rev. E*, 2012, **85**, pp. 051134-1–051134-9

[7] Matsuda Y., Hanasaki I., Iwao R., *ET AL.*: 'Faster convergence of diffusion anisotropy detection by three-step relation of single-particle trajectory', *Anal. Chem.*, 2016, **88**, pp. 4502–4507

[8] Hanasaki I., Kawano S.: 'Evaluation of bacterial motility from non-Gaussianity of finite-sample trajectories using the large deviation principle', *J. Phys., Condens. Matter*, 2013, **25**, pp. 465103-1–465103-9

[9] Hanasaki I., Uehara S., Arai Y., *ET AL.*: 'Threshold-free evaluation of near-surface diffusion and adsorption-dominated motion from single-molecule tracking data of single-stranded DNA through total internal reflection fluorescence microscopy', *Jpn. J. Appl. Phys.*, 2015, **54**, pp. 125601-1–125601-6

[10] Uehara S., Hanasaki I., Arai Y., *ET AL.*: 'Statistical characterisation of single-stranded DNA motion near glass surface beyond diffusion coefficient', *Micro Nano Lett.*, 2014, **9**, pp. 257–260

[11] Chenouard N., Smal I., Chaumont F., *ET AL.*: 'Objective comparison of particle tracking methods', *Nat. Methods*, 2014, **11**, pp. 281–289
[12] Sbalzarini I.F., Koumoutsakos P.: 'Feature point tracking and trajectory analysis for video imaging in cell biology', *J. Struct. Biol.*, 2005, **151**, pp. 182–195
[13] Meijering E., Dzyubachyk O., Smal I.: 'Methods for cell and particle tracking', *Methods Enzymol.*, 2012, **504**, pp. 183–200
[14] Racine V., Hertzog A., Jouanneau J., *ET AL.*: 'Multiple-target tracking of 3D fluorescent objects based on simulated annealing'. 2006 Third IEEE Int. Symp. on Biomedical Imaging: Nano to Macro, 2006, pp. 1020–1023
[15] Crocker J.C., Grier D.G.: 'Methods of digital video microscopy for colloidal studies', *J. Colloid Interface Sci.*, 1996, **179**, pp. 298–310
[16] Ku T.-C., Huang Y.-N., Huang C.-C., *ET AL.*: 'An automated tracking system to measure the dynamic properties of vesicles in living cells', *Microsc. Res. Tech.*, 2007, **70**, pp. 119–134
[17] http://www.soken-ce.co.jp/en/product/fine_particles/mx/

doi: 10.1049/mnl.2017.0044