Main Manuscript for

Enhancing the localization precision of single molecule localization microscopy by analyzing Brownian motion of individual fluorescent probe

Jielei Ni [a, 1], Bo Cao [a, 1], Gang Niu [b], Chen Xu [a], Yanxiang Ni [a, 2],

[a] Nanophotonics Research Center, Shenzhen Key Laboratory of Micro-Scale Optical Information Technology & Institute of Microscale Optoelectronics, College of Physics and Optoelectronic Engineering, College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China

[b] PhilRivers Technology, Beijing, China

[1] J. N. and B. C. contributed equally to this work

[2] To whom correspondence should be addressed. **Email:** niyanxiang0000@gmail.com

Abstract

The localization precision is a crucial parameter for single-molecule localization microscopy which directly influences the achievable spatial resolution. Although it can be optimized with photon number, system noise and system drift, we demonstrate that, Brownian motion of the fluorescent probe can be another factor that needs to be taken into consideration. By calculating the z-score of the displacement between adjacent frames, fluorescent probes that are undergoing large amplitude of Brownian motion can be identified. By filtering out these molecules, distinct improvement of localization precision can be achieved.

Introduction

Single-molecule localization microscopy (SMLM) such as stochastic optical reconstruction microscopy (STORM) (1), photo-activated localization microscopy (PALM) (2), DNA-PAINT (3) or MINFLUX (4-6) has circumvented the diffraction barrier, offering 2-30 nm lateral resolution and the possibility of 3D imaging (1-8). SMLM has allowed for unprecedented high-resolution visualization of various biological structures such as microtubules, actin, clathrin-coated pits, mitochondria, chromatin complexes, neurons, ER, and focal adhesion complexes, which are previously considered impossible due to the diffraction limit (9).

In SMLM, fluorophores can be temporally separated, so that a sparse set of emitters can be captured in one camera frame, avoiding spatial overlap of their point spread functions (PSF). The temporal separation of the fluorophores is achieved via photoactivation (2), stochastic switching of fluorophores between a fluorescent and nonfluorescent state (1), or the reversible binding of fluorophore to the target site (3). The sparse set of activated fluorophores in each frame is then localized by fitting Gaussians to individual intensity peaks. These detected peaks can then be grouped over space and time to precisely estimate the location of the individual emitters, forming a high-resolution image. Experimental factors, such as the number of photons collected at each point, pixilation of the detector and noise sources in the detection process including shot noise of the photons, background noise created by out-of-focus fluorescence, readout noise and dark current of detector, will influence the localization precision. Theoretical calculation of the best localization precision is given by the square root of the Cramér-Rao lower bound (CRLB) based on a statistical theory concerning the Fisher information matrix. Analytical expressions are also given to show how the localization error can be reduced by experimental conditions [10-11].

Although localization error can be minimized using the above-mentioned theoretical model, it is hard to reach its theoretical limit. There are many factors that could cause the discrepancy. Sample or instrument drift during acquisition is critical one. Drift caused by temperature changes or mechanical relaxation effects can occur in the range of several hundred nanometers over the course of a few minutes. Approaches for drift compensation have been proposed including reference markers (1, 12-13) and cross-correlation correction (14-15) in the past years. However, inadequate drift correction can still exist.

Brownian motion of the fluorescent probe is another key factor that is unfortunately impossible to be corrected due to its unpredictable trajectory. The influence of Brownian motion to localization precision has been reported in the measurement of the diffusion coefficient by single particle tracking (SPT). It is shown that Brownian motion during imaging process can lead to distortion of the shape of the observed intensity distribution, resulting in a deterioration of the localization precision compared to the stationary one (16-17). However, in SMLM experiments, Brownian motion of the fluorescent probe has been largely neglected. Its influence on the localization precision has not been reported.

Here, we present a statistical method for the analysis of the molecule motion of fluorescent probes. Displacement between two localizations from adjacent frames and that from random frames are calculated, a statistically significant difference was found between the two groups of displacement. Further analysis of the displacement at different time lag shows the existence of Brownian motion and inadequate drift correction. Our method provides ways to identify fluorescent probes that are

undergoing large amplitude of Brownian motion. By simply eliminating these molecules from the data set, the localization precision can be further improved. Different from current drift compensation approaches that correct the systematic drift of molecules, our results reveal the possibility of improving localization precision by recognition of the motion of single molecule.

Results and discussion

Fluorescent probe with different DNA tether length

In our experiment, Alexa Fluor 647 dye was labeled to the 3' end of a short DNA oligonucleotides (18 nt), which has been immobilized to a glass surface. Since the fixed point of the DNA is non-selected, the effective length R between the dye and the fixed point is a random value smaller than 6.12 nm, as illustrated in Fig. 1a. Due to the Brownian motion of the DNA, the localization of the dye fluctuates with time. The amplitude of motion increases with the effective length R. Three types of molecules will be distinguished here.

In the first type of molecules, the oligonucleotides are fixed close to the 5' end so that the effective length R is large, as depicted by molecule 1 in Fig. 1a. Fig. 1b illustrates the molecule undergoing Brownian motion. The green circle denotes the bead for drift correction. The red circle depicts the Alexa Fluor 647 dye attached to the free end of a DNA molecule undergoing Brownian motion in a domain confined by its effective length. The positions at different time (T0, T1, T2, …, Tn) are illustrated by the red dotted circle. In Fig. 1c, the bead and the sample undergo drift motion during the time interval between T0 and T1' as indicated by the black solid arrow. Drift compensation corrects the position of bead back to the T0 position as shown by the black dotted arrow. However, for the dye molecule, due to the existence of Brownian motion, the position cannot be compensated back to the position at T0 but another position. Analyzing the trajectories after drift compensation would give access to the dynamics of Brownian motion.

Taking the localization uncertainty into consideration, if the localization uncertainty is too large compared to the amplitude of the Brownian motion, the information of the Brownian motion would be undetectable. These molecules are fixed to the glass slide with a fixed point close to the dye, leading to a small effective length, as illustrated by the molecule 2 in Fig. 1a.

Apart from the two types of molecules mentioned above, there are a few molecules with its fixed point near the dye. Since the effective length is close to zero, there would be no Brownian motion.

Using z-score to study Brownian motion

The statistical method to explore the Brownian motion is explained as follows.

The raw data of STORM are obtained by identifying and fitting single fluorescent spots, which generates a "localization list" that contains spatial coordinates, frame number. We define each localization of the same molecule as $Loc_i(x_i, y_i, f_i), i = 1, 2, 3, \dots, N$, where N is the total number of localizations. The spread of the localizations follows a gaussian distribution.

We define the displacement between two localizations in adjacent frames as d_a, which can be calculated as

$$d_a = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \text{ , if } f_{i+1} - f_i = 1. \quad (1)$$

Similarly, the displacement between two localizations in random frames is defined as,

$$d_r = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, i, j \in \{Loc\} \tag{2}$$

To establish statistical significance, we sampled randomly from the data set $\{Loc\}$ with the same number of d_a to get a series of d_r. The mean of both are calculated by

$$\overline{d_a} = mean(d_{ai}) = \frac{1}{L}\sum_{i=1}^{L} d_{ai} \tag{3}$$

$$\overline{d_r} = mean(d_{ri}) = \frac{1}{L}\sum_{i=1}^{L} d_{ri} \tag{4}$$

The sampling is repeated 5000 times to get a series of $\overline{d_r}$, the histogram of which should follows Gaussian distribution according to Central-Limit Theorem. To calculate the statistical difference between d_a and d_r, we define the z-score of d_a as

$$z = \frac{\overline{d_a} - mean(\overline{d_{rn}})}{std(\overline{d_{rn}})} \tag{5}$$

The z-score defines how many standard deviations between the adjacent displacement and the random displacement. A negative z-score indicates the $\overline{d_a}$ is below the mean of $\overline{d_r}$.

Note that the position measurement in SMLM can be treated as a random process, as long as the molecule is static or drift is completely corrected, the displacement between two localizations in adjacent frames and that between two random frames should have no statistically significant difference. However, for Brownian motion or drift motion, since the mean squared displacement should increase with its time interval, which gives a negative z-score that will distinctly deviate from zero.

In our experiment, the z-score before drift correction and that after drift correction are both calculated for 2108 molecules. It is found that the z-scores of most molecules situated near the value of -2.5 before correction, while the value shifted to -0.5 after correction, as presented in Figs. 2a-2b. This shows that drift correction is effective for most molecules.

The uncorrected z-score is plotted against the corrected z-score in Fig. 3a. it can be clear seen that the distribution of corrected z-score is not symmetric. The molecules that are well-corrected should have a distribution with mean of zero and a certain standard deviation. For a quick examination, we plotted the center of zero as the black solid line and a deviation of plus or minus two as the black dash line in Fig. 2(a). It can be clearly seen that there are more counts on the negative side and few counts on the positive side. This means that while the majority of molecules can be well-corrected, there are a group of molecules that cannot be corrected by the fluorescent bead.

To provide more information of the molecules, Figs. 2b and 2c give the z-scores after correction at different time lag. In this case, Eq. (1) and Eq. (5) are revised as,

$$d_{am} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \text{ , if } f_k - f_i = m. \quad (6)$$

$$z_m = \frac{\overline{d_{am}} - mean(\overline{d_{rn}})}{std(\overline{d_{rn}})} \quad (7)$$

Fig. 3b shows the z-score z_m of a molecule situated at the far negative regime in Fig. 3a. As can be seen in Fig. 3b, the z-score z_m first increases with time lag m, starting with small value z_1 =-*3.3687*, indicating the fluorescent probe is undergoing diffusion motion. And then it reaches a saturating value after a certain time lag.

Fig. 3c displays the z-score z_m of a molecule situated near the zero line (black solid line) in Fig. 3a. The z-score z_m keep fluctuating around zero with time lag m, with a start point $z_1 = 0.06$. This suggests the displacements that are irrelevant of time lag.

To further explore the statistical feature in Fig. 3a, we study two group of the molecules, one with z-score z_1 ranging from -2 to 2, the other with z-score z_1 smaller than -2. Fig. 4a displays the z-score z_m of all the 182 molecules with z_1 smaller than -2, at different time lag m. The overall behavior can be clearly seen that the z-score firstly increase with time lag and then saturate near zero after a certain time lag m_0. This should correspond to the time required for the dye to explore the domain within which it is confined by the DNA tether. The data can be fitted with a saturation function $y(x) = y_0 - p_1 exp(-p_2 x)$ with $p_1 = 2.827$, $p_2 = 0.3444$ and $y_0 = -0.3791$ (red line). This group of molecules have typical characteristic of Brownian motion, corresponding to those molecules that have large effective length in Fig. 1a. Fig. 4b shows the z-score z_m of all the 1913 molecules with z_1 ranging from -2 to 2, at different time lag m. We can see that the z-score z_m this group of molecules fluctuate around zero. Most of these molecules can be attributed to the molecule 2 or molecule 3 in Fig. 1a. They either have small effective length so that the amplitude of Brownian motion is too small to be detected due to localization uncertainty, or have zero effective length so that no Brownian motion exists.

Simulation of Brownian motion

To further confirm our speculations, a 2D Brownian motion confined to a circular disc is performed. We used a constant time interval Δt between frames. Disregarding the localization uncertainty of each frame, if positions could be recorded instantaneously, the position recorded at each frame will be,

$$x_n = \sum_{i=1}^{n} \sqrt{2D\Delta t} \cdot randn(1, N) \quad (8)$$

$$y_n = \sum_{i=1}^{n} \sqrt{2D\Delta t} \cdot randn(1, N) \quad (9)$$

Here, D is the diffusion constant and N is the number of frames that collected. The $randn(1, N)$ generates a 1-by-N array of random numbers from the standard normal distribution.

Since the finite length of the DNA strand, we set an upper limit to the scalar distance $R_n = \sqrt{x_n^2 + y_n^2} < R_0$.

In our simulation, we set $\sqrt{2D\Delta t} = 10.3537\, nm$, $R_0 = 10\, nm$, $N = 8000$. Fig. 5a shows the simulated z-score changing with different time lag. The data is fitted by a saturation function $y(x) = y_0 - p_1 exp(-p_2 x)$ with $p_1 = 5.701$, $p_2 = 0.7497$ and $y_0 = 0.1046$ (black solid

line). A distinct deviation from the experimental data in Fig. 4d is that, the maximum z-score in the simulation is close to zero in simulation, while it is around -0.4 in experiment. Due to the restriction of DNA tether, the displacements at large time lag should dominantly be R_0, resulting to $\overline{d_{am}} \approx mean(\overline{d_{rn}})$. In such case, the z-score at large time lag should approximate zero. The deviation from zero in the experiment might suggest drift correction residual, which adds a linear term to the curve.

To confirm this, we added a linear term to the position,

$$x'_n = x_n + kn \tag{10}$$

$$y'_n = x_n + kn \tag{11}$$

In our simulation, we set $k = 0.02$. The simulated z-score with different time lag is plotted in Fig. 5b and fitted by a saturation function $y(x) = y_0 - p_1 exp(-p_2 x)$ with $p_1 = 3.019$, $p_2 = 0.292$ and $y_0 = -0.52$ (black solid line). We can see that this simulation result is in good agreement with our experimental result.

Calculation of localization precision

Brownian motion of the labeled dye in SMLM experiment strongly degrades localization precision. We calculated the experimental localization precision from the positions of multiple molecules overlaid in their position center. 722 molecules with $-0.5 < z_1 < 0.5$ are also calculated, yielding a standard deviation of 7.9582 nm in the x-direction and 7.4179 nm in the y-direction. In comparison, 182 molecules that have $z_1 < -2$ are calculated, yielding a deteriorated standard deviation of 12.9665 nm in the x-direction and 11.1103 nm in the y-direction.

Noticing that the molecules with $2 < z_1 < 2$ take up a dominant percentage (~90%) in this experiment, filtering out molecules beyond this range resulting to a standard deviation improved from 8.9239 nm to 8.1940 nm in the x-direction and 8.1726 nm to 7.6695 nm in the y-direction. Potential ways to suppress Brownian motion including water cooling or using semiconductor cooler. But vibration might be introduced during the cooling process.

Conclusion

The localization precision is a crucial and important parameter for single-molecule localization microscopy (SMLM) and directly influences the achievable spatial resolution. We show that the localization precision depends not only on the photon number, shot noise and background noise, but also on the amplitude of Brownian motion of the fluorescent probe. We developed a statistical method to extract Brownian motion from the trajectory of the localizations of single molecule.

Materials and Methods

System:

A STORM system based on an inverted optical microscope (IX-71, Olympus) with a 100 oil immersion objective lens (Olympus) was used for the nano-imaging as previously described (Ni et al., 2017). Astigmatism imaging method was adopted for 3D-STORM, in which a weak cylindrical lens (1 m focal length) was introduced into the imaging path. A 641 nm laser (CUBE 640–100C;Coherent) was used to excite fluorescence and switch Alexa-647 to the dark state. The illumination used the highly inclined and laminated optical sheet (HILO) configuration (Tokunaga et al., 2008). The laser power densities were approximately 1.45 kW/cm2 for the 641 nm laser unless otherwise indicated. A dichroic mirror (ZT647rdc, Chroma) was used to separate the fluorescence from the laser and a band-pass filter (FF01-676/37, Semrock) on the imaging path was used to filter the fluorescence. Raw images of the fluorescent signals in each nuclear field were acquired with an EMCCD (DU-897U-CV0, Andor) at 33 Hz for 8000 frames. To avoid focal drift, an anti-drift system was used to sustain the focal position within 10 nm during image processing (Huo et al., 2015).

Buffer:

A standard STORM imaging buffer containing 50Mm Tris (pH 8.0), 10mM NaCl,1% b-mercaptoethanol (v/v), 10% glucose (w/v), 0.5 mg/mL glucose oxidase (G2133, Sigma), and 40mg/mL catalase (C30, Sigma) (Dempsey et al., 2011; Olivier et al., 2013) was used.

Label Protocal:

Each coverslip was incubated at RT with 25nM Alexa-647-conjugated 10T (a 10T oligonucleotide labeled with one Alexa-647 at its 5’ end) for 8 min and then incubated with 0.2 μm red fluorescent microspheres (F8810, Thermo Fisher) as fiducial markers to correct for sample drift in the x-y plane during sample imaging.

Method:

For imaging data analyses, a freely available plugin for Image J named Thunderstorm was applied to analyze raw images. The precise localization data of single fluorescent molecules obtained from the fits.

Acknowledgments

This work was supported by Shenzhen Science and Technology Planning Project (JCYJ20170817095211560).

References

1. Rust M J, Bates M, Zhuang X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)[J]. Nature methods, 2006, 3(10): 793-796.
2. Betzig E, Patterson G H, Sougrat R, et al. Imaging intracellular fluorescent proteins at nanometer resolution[J]. Science, 2006, 313(5793): 1642-1645.
3. Jungmann R, Avendaño M S, Woehrstein J B, et al. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT[J]. Nature methods, 2014, 11(3): 313-318.
4. Balzarotti F, Eilers Y, Gwosch K C, et al. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes[J]. Science, 2017, 355(6325): 606-612.
5. Gwosch K C, Pape J K, Balzarotti F, et al. MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells[J]. Nature methods, 2020, 17(2): 217-224.
6. Schmidt R, Weihs T, Wurm C A, et al. MINFLUX nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope[J]. Nature Communications, 2021, 12(1): 1-12.
7. Van de Linde S, Löschberger A, Klein T, et al. Direct stochastic optical reconstruction microscopy with standard fluorescent probes[J]. Nature protocols, 2011, 6(7): 991.
8. Huang B, Wang W, Bates M, et al. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy[J]. Science, 2008, 319(5864): 810-813.
9. Huang B, Babcock H, Zhuang X. Breaking the diffraction barrier: super-resolution imaging of cells[J]. Cell, 2010, 143(7): 1047-1058.
10. Thompson R E, Larson D R, Webb W W. Precise nanometer localization analysis for individual fluorescent probes[J]. Biophysical journal, 2002, 82(5): 2775-2783.
11. Deschout H, Zanacchi F C, Mlodzianoski M, et al. Precisely and accurately localizing single emitters in fluorescence microscopy[J]. Nature methods, 2014, 11(3): 253.
12. Bon P, Bourg N, Lécart S, et al. Three-dimensional nanometre localization of nanoparticles to enhance super-resolution microscopy[J]. Nature communications, 2015, 6(1): 1-8.
13. Lee S H, Baday M, Tjioe M, et al. Using fixed fiduciary markers for stage drift correction[J]. Optics express, 2012, 20(11): 12177-12183.
14. Wang Y, Schnitzbauer J, Hu Z, et al. Localization events-based sample drift correction for localization microscopy with redundant cross-correlation algorithm[J]. Optics express, 2014, 22(13): 15982-15991.
15. Mlodzianoski M J, Schreiner J M, Callahan S P, et al. Sample drift correction in 3D fluorescence photoactivation localization microscopy[J]. Optics express, 2011, 19(16): 15009-15019.
16. Deschout H, Neyts K, Braeckmans K. The influence of movement on the localization precision of sub‐resolution particles in fluorescence microscopy[J]. Journal of biophotonics, 2012, 5(1): 97-109.
17. Michalet X. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium[J]. Physical Review E, 2010, 82(4): 041914.

Figures and Tables

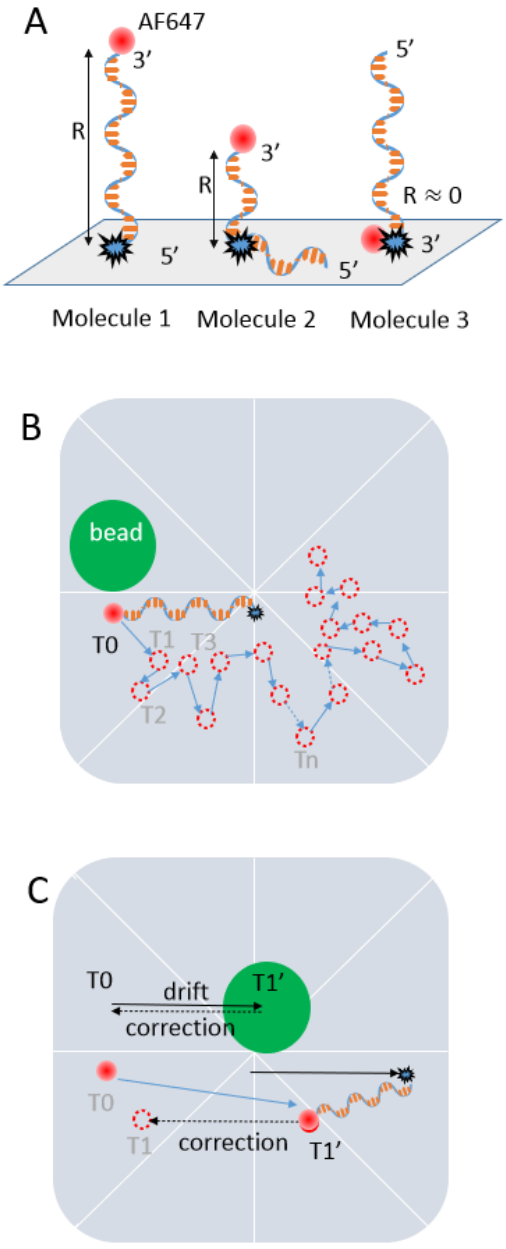

Figure 1. (A) Three types of molecules with different length of DNA tether. (B) Molecules that have long DNA tether exhibit Brownian motion. (C) Brownian motion cannot be corrected by drift correction.

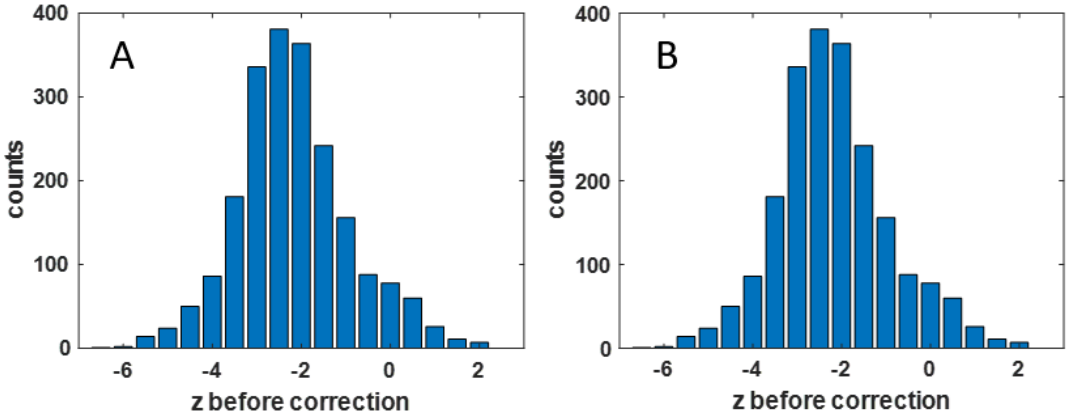

Figure 2. The histogram of z-score before correction (A) and after correction (B).

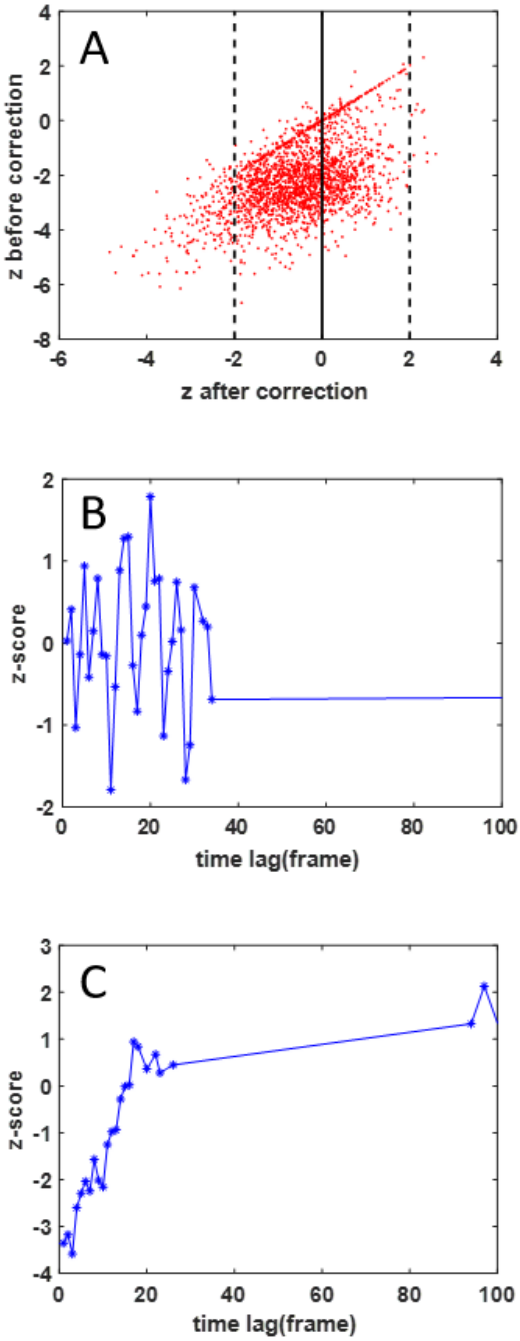

Figure 3. (A) A plot of the z-score before and after correction, which shows that some molecules can be corrected to a z-score around zero, but some molecules still have a large negative z-score after correction. The black solid line indicates zero z-score after correction. The black dashed line represents a z-score of -2 and 2 after correction. (B) the z-score after correction of the molecules near the black solid line in Fig. A with different time lag. (C) the z-score after correction of the molecules at the far negative regime of Fig. A with different time lag.

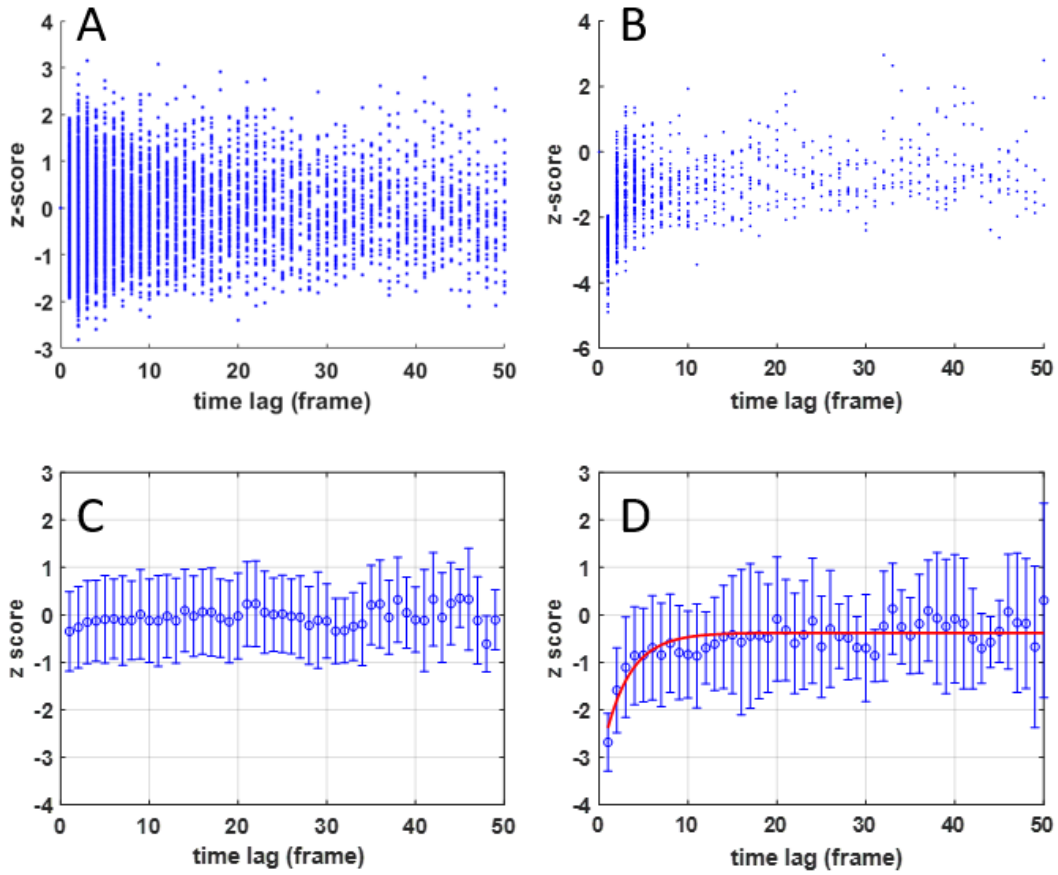

Figure 4. (A) The time-dependence of z-score of 1913 molecules with z_1 ranging from -2 to 2. (B) The time-dependence of z-score of 182 molecules with z_1 smaller than -2. (C) the mean z-score averaged by the 1913 molecules in (A). (D) the mean z-score averaged by the 182 molecules in (B).

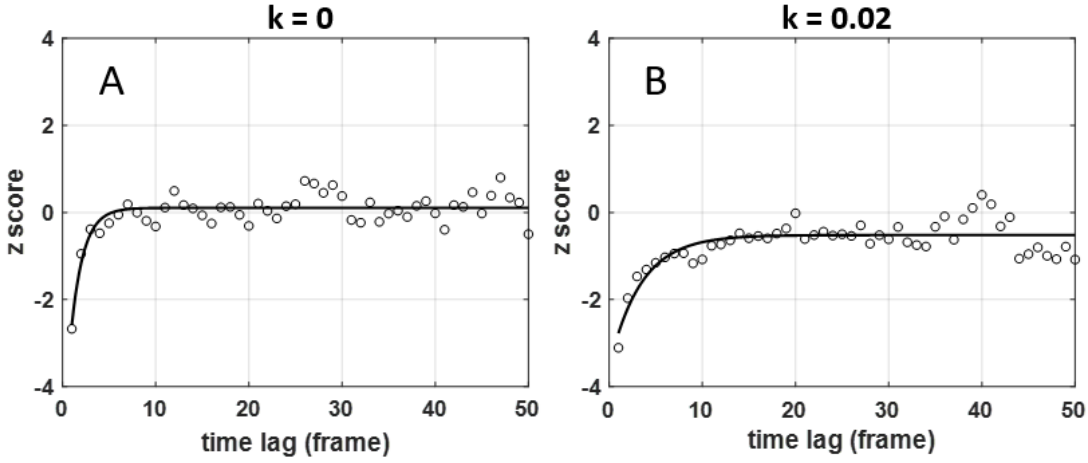

Figure 5. Simulated z-score (black circle) with $z_1 < -2$ at different time lag and the fitted saturation function (black line) without (A) and with (B) drift correction residual.