Drift estimation for single marker switching based imaging schemes

Claudia Geisler, Thomas Hotz, Andreas Schönle, Stefan W. Hell, Axel Munk, and Alexander Egner

1Department of Optical Nanoscopy, Laser-Laboratory-Göttingen, Hans-Adolf-Krebs-Weg 1, D-37077 Göttingen, Germany
2Institute for Mathematical Stochastics, University of Göttingen, Goldschmidtstrasse 7, D-37077 Göttingen, Germany
3Department of NanoBiophotonics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
4Statistical Inverse Problems Group, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
5These authors contributed equally to this work

Abstract: In recent years, the diffraction barrier in fluorescence imaging has been broken and optical nanoscopes now routinely image with resolutions of down to 20 nm, an improvement of more than 10 fold. Because this allows imaging much smaller features and because all super-resolution approaches trade off speed for spatial resolution, mechanical instabilities of the microscopes become a limiting factor. Here, we propose a fully data-driven statistical registration method for drift detection and drift correction for single marker switching (SMS) imaging schemes, including a guideline for parameter choice and quality checks of the drift analysis. The necessary assumptions about the drift are minimal, allowing a model-free approach, but more specific models can easily be integrated. We determine the resulting performance on standard SMS measurements and show that the drift determination can be routinely brought to the range of precision achievable by fiducial marker-tracking methods.

©2012 Optical Society of America

OCIS codes: (350.5730) Resolution; (180.2520) Fluorescence microscopy; (100.6640) Superresolution; (100.3010) Image reconstruction techniques.

References and links


1. Introduction

Far-field optical microscopy plays a key role in a number of areas, particularly in the life sciences. The reason is a series of major advantages of imaging with propagating fields of light, such as the non-invasive access to the interior of non-living and living cells and the sensitive detection of macromolecules by fluorescence tagging. Nevertheless, due to the laws of diffraction, conventional far-field microscopy cannot discern details that are substantially smaller than half the wavelength of light [1]. Optical nanoscopy uses the photo-physical
properties of the label itself switching it typically between a signal-giving (on) and a no-signal giving (off) state to break this barrier and to address the information inside a diffraction limited spot with a much higher spatial resolution [2, 3]. The key is to ensure that markers closer than the optical resolution limit never contribute to the signal synchronously thus making their signals separable in time. Due to this time sequential readout of markers inside a diffraction limited volume an increase in resolution almost always comes at the cost of a longer acquisition time. This and the higher resolution and thus finer structures under scrutiny make nanoscopes more sensitive to thermodynamic and mechanical instabilities of the imaging setup, i.e. drift, vibration and deformation of the sample within the field of view.

All nanoscopy methods either switch ensembles of molecules as e.g. in STED microscopy [2] or they are based on (stochastic) switching of single markers [4–7]. In the latter approach the positions of single markers are time sequentially recorded and registered in a position histogram which constitutes the final image. To this end typically a single fluorescent marker out of many markers lying within a diffraction-sized region (which are initially all in an off-state) is transferred to the on-state in which it is repeatedly excited and emits a bunch of photons. A fraction of those is detected on an array detector e.g. on a CCD camera forming a diffraction blob of a single marker from which its position is determined. After 1-100 ms, the fluorophores are either switched off or bleached, and adjacent molecules can be read out in the same way. The final image is then assembled by registering all recorded locations in a position histogram with an average precision which is typically in the range from 20 to 80 nm. Undetected movement of the sample during the acquisition of a single frame will obviously lead to blurring of the image and/or imaging artifacts.

Although image acquisition times as short as 1 s have been reported [8], these cannot be considered to be typical. Such short acquisition times are only achievable in exceedingly sparse samples and consist of the molecular positions from 100 to 1000 individual frames, implying that image quality is limited by the density of localized fluorophores rather than by the localization precision [9]. Exposure times of 1 min-1 h [4, 7, 10, 11] are much more typical for single marker switching (SMS) images which usually consist of positions determined from 10,000 to 100,000 individual camera frames with exposure times of 1-100 ms per frame.

This makes SMS microscopy a long-term measurement [12] and similarly sensitive to instrumental drift as e.g. time-lapse studies in conventional microscopy [13]. Lateral drifts from several tens of nm up to 1 µm, on time scales of seconds to several minutes, have been reported for self-constructed as well as for commercial microscopes [5, 7, 14, 15]. To some extent hardware adaptions can compensate for focus drift e.g. by capacitance based or interferometric distance measurements. However, most of these systems merely control axial drift and all of them are restricted to measuring the drift at a fixed, pre-defined position which can be quite different from the sample movement at the region currently measured. Therefore, fiducial markers are usually incorporated into the sample, since they can be used either for actively compensating the drift during the measurement [16] or for post-processing correction [4, 5, 17–20]. However, even setting aside the increased experimental effort, this strategy has several drawbacks: there has to be at least one fiducial marker inside the field of view of the detector at all times and the reliability of the drift deduction can be severely hampered if the marker moves relative to the sample or if it degrades or bleaches during the measurement. In its simplest implementation, i.e. when utilizing particles strongly emitting in the same spectral range as the fluorophores, additional drawbacks apply: if the particle gets too close to the structure of interest, the single molecule signals will be outshined and will therefore not be discernable. Furthermore, the particle’s signal must not lie outside the dynamic range of the detector.

A reliable method for drift correcting SMS-based imaging data that does not rely on additionally incorporated markers but uses the information given by the sample itself would solve these problems. It has been reported in the literature that the spatial correlation of temporal subsets of position histograms with a reference histogram can be used for this purpose [21–23] and that a drift compensation down to a sub-5 nm level can be achieved for
2. Statistical model, estimation procedures and simulation

2.1 Problem description

The “true” image to be reconstructed is given by the distribution of markers over the image domain. From this distribution, a random marker from a marker distribution is switched on at a random time point, it emits a photon burst of $N$ photons which is recorded by a camera with a single frame exposure time $t_{exp}$. The marker’s position $(x, y)$ is estimated from the approximately Gaussian photon distribution with a precision proportional to $1/\sqrt{N}$ (when background noise and camera noise are neglected) [25] and the corresponding time point $t$ is identified with a precision $t_{exp}$ from the respective camera frame number. Repeating the cycle of switching on, reading out and switching off yields a list of $K$ events $(x_k, y_k, t_k)$, $k = 1, ..., K$. These are then binned in space and represented in a 2D position histogram for algorithmic tractability. If there was a movement over time of the marker distribution in space, the position histogram does not represent the marker distribution, but a motion-blurred version of it. An additional temporal binning yields a series of typically less motion-blurred, but sparser, representations of the marker distribution. For this, not all events are used for the “image”, but only events which were detected in a certain time period which will be described by, e.g., its midpoint. These “images” can then be mathematically described as follows.

We observe $n$ images at time points $t_i$, $i = 1, ..., n$, each comprising $m$ pixels located at positions $(u_j, v_j)$, $j = 1, ..., m$ where $Z_{ij}$ events have been counted, i.e.

$$Z_{ij} = \#\{k : t_k = t_i \text{ and } (x_k, y_k) \text{ in pixel } j\},$$

where $\#$ denotes the number of events. We assume that in good approximation each fluorophore has a small and constant probability to be detected and registered within any given frame. The number of events $K_i = \sum Z_{ij}$ counted at times $t_i$ are then independent Poisson variables and the observations $Z_{ij}$ (the image) result form $K_i$ independent draws from the marker distribution $f$ (more precisely, the convolution of the marker distribution with some kernel of width proportional to $1/\sqrt{N}$); hence, $\sum K_i = K$. That means, according to the thinning of this Poisson process, that the $Z_{ij}$ are also independent Poisson random variables and their mean is given by

$$\mu_{ij} = \langle Z_{ij} \rangle = \langle K_i \rangle f(u_j - \delta_{x,i}, v_j - \delta_{y,i}),$$

where we already incorporated a possible drift described by the shift of $(\delta_{x,i}, \delta_{y,i})$ of the marker distribution $f$ at time $t_i$. We ignore binning effects which could easily be made explicit by integrating over each bin in space and time but can be neglected if the bin and pixel sizes are small enough. Note that the factor $\langle K_i \rangle$, i.e. the expected total number of events per frame, is constant in space, hence the images are independent while their expected counts per pixel are proportional to each other (up to the drift); if the assumption of constant intensity over time is violated, the estimates below should be adapted with the appropriate weights. It is now our aim to reconstruct the drift $\delta$ given the observations $Z_{ij}$, $i = 1, ..., n$, $j = 1, ..., m$. 
2.2 Estimation procedures

There are many alignment techniques to determine the drift between two images, see e.g. the reviews [26–28], or also [29]. The particular method we chose in this manuscript is based on a cross-correlation analysis of the two images, followed by a Gaussian mask-fitting algorithm to determine the center position of the cross-correlation’s maximum with sub-pixel precision [25]. To remedy the heteroscedasticity of the $Z_{ij}$, we slightly modified this well-established and efficient technique by using an Anscombe-like variance stabilizing transform, working with $\sqrt{(Z_{ij} + 1/4)}$ instead, cf [30]. Note that this transform is applied because we assumed the $Z_{ij}$ to be Poisson random variables; would they be Gaussian with constant variance, this would be unnecessary.

It is important to state that our reconstruction approach can be based on any method which takes the observations $Z_{ij}$ and $Z_{lj}$ of images $i$ and $l$, and returns an estimate $\Delta_{il}$ of the drift $\delta_{i} - \delta_{l}$ from time $t_{i}$ to $t_{l}$. In fact, we only need to assume that they stem from independent images $i$ and $l$, with the variance of the $\Delta_{ij}$ being the same for all pairs of images; the assumptions on the $Z_{ij}$ made above, combined with our cross-correlation approach, clearly ensure this. It is natural to ask for this method to be symmetric, i.e. $\Delta_{il} = -\Delta_{li}$, and hence $\Delta_{il} = 0$, for $1 \leq i, l \leq n$, as is the case for the cross-correlation we employed. From now on, we will thus directly use the pair-wise drift estimates $\Delta_{il}$ and the important question that remains is how to obtain a reliable estimate $\delta^*$ of the drift $\delta$ from these.

2.2.1. Compare each image with a single image

Two approaches appear natural: e. g., one might correlate each image with the first one (or the second one or any other arbitrary frame of the series), i.e. estimating $\delta_{l}$ by $\delta^*_{il} = \Delta_{il}$ for $i = 2, ..., n$, $\delta^*_{1l} = \Delta_{1l} = 0$. Alternatively, one might correlate each image with its predecessor, i.e. recursively setting $\delta^*_{i,j} = 0$, and then $\delta^*_{i} = \delta^*_{i,j} + \Delta_{i,j}$, for $i = 2, ..., n$. In the following, we will abbreviate these methods by “first” (“second”, “third”, …) and ’pre’ respectively. Note that the model in Eq. (2) does not allow to identify an absolute drift, since adding a constant to the drift may be incorporated into $f$. Hence, only a relative drift can be reconstructed, allowing us to freely choose the origin, e.g. by letting $\delta^*_{1} = 0$ as above.

2.2.2. Compare each image with all other images

Comparing each image with just a single image has the obvious shortcoming of not using information from all $n(n-1)/2$ pair-wise drift estimates $\Delta_{il}$, $1 \leq i < l \leq n$ available. In order to use all available information we view the $\Delta_{il}$ as noisy drift estimates,

$$\Delta_{il} = D_{il} + \varepsilon_{il}$$

with zero-mean errors $\varepsilon_{il}$ of variance $\sigma^2$ which are uncorrelated for each fixed $i$, where

$$D_{il} = \delta_i - \delta_l,$$

and thus

$$\Delta_{il} = \delta_i - \delta_l + \varepsilon_{il}.$$  (5)

The $\Delta_{il}$ depend linearly on the $(n-1)$ unknown drift values $\delta_i$, $i = 2, ..., n$, plus noise, suggesting to estimate the drift by least-squares. It is straight-forward to show that the least-squares estimator for $\delta_i$ is given by averaging the inferior estimates $\Delta_{il}$ over $l$, see appendix; it can thus be computed very efficiently.

We stress that such a least-squares estimate makes no assumption on the shape of the drift; we therefore call it “model-free”. It is important that by the very nature of this approach we can only hope to take into account uncorrelated noise of the estimates. In other words, the estimate for $\delta_i$ will not describe the real physical drift but contains possibly correlated errors, i.e. apparent shifts of whole frames relative to all others which are clearly indistinguishable from an actual physical shift without additional information.
It is instructive to remark that, assuming the model from Eq. (5), method “first” results in a constant variance of $2\sigma^2$, whereas for independent errors “pre” would result in a random walk of the estimate’s error and thus increasing variance. The “model-free” least-squares approach, however, results in a variance proportional to $1/(n-1)$. Without any further assumptions, this clearly is the best we can do: reconstruct $D_i$ in Eq. (3) fulfilling the relation in Eq. (4) for some $\delta$.

2.2.3. Mixed effects model

Often it is reasonable to assume more, e.g. that the drift is well described by a linear or quadratic function, or by a spline. More generally, let us consider a drift given by some linear model equation

$$\delta = X\beta$$  \hspace{1cm} (6)

with $\beta$ being the (unknown) parameter vector of dimension $\leq n$ and $X$ the full-rank design matrix. We can then rewrite Eq. (4) as

$$D_{ij} = d_i - d_j$$ \hspace{1cm} (7)

with noisy drifts

$$d = \delta + \eta = X\beta + \eta,$$ \hspace{1cm} (8)

where the errors $\eta_i$ are independent, zero-mean noise with inter-frame variance $\tau^2$, independent of the errors $\epsilon_{ij}$ which now can be taken to be independent as well. Putting Eq. (3), Eq. (7) and Eq. (8) together, we obtain

$$\Delta_{ij} = \delta_i - \delta_j + \eta_i - \eta_j + \epsilon_{ij} = (X\beta)_i - (X\beta)_j + \eta_i - \eta_j + \epsilon_{ij}.$$ \hspace{1cm} (9)

This gives the appropriate covariance structure of the pair-wise drift estimates: $\Delta_{ij}$ and $\Delta_{i'j'}$ are independent if the pairs of images are disjoint; they are identical if they are based on the same pairs; and they feature some correlation $\pm \rho$ if they have exactly one image in common. Importantly, $\rho$ is a constant if the experimental conditions do not change over time; in model Eq. (9) we get $\rho = \tau^2 / (2\tau^2 + \sigma^2)$. Again we stress that $\delta$ and $\eta$ in Eq. (8) can only be distinguished if we make assumptions about the drift, otherwise we can only hope to reconstruct the drift including these apparent shifts of the whole frame. As the errors $\eta_i - \eta_j + \epsilon_{ij}$ in Eq. (9) are correlated, this is a linear mixed effects model which can be fitted by standard techniques, cf. e.g [31]. Please note that it is still valid to use the “model-free” least-squares estimate under this modified noise model by substituting $\epsilon_{ij}$ in Eq. (5) with $\eta_i - \eta_j + \epsilon_{ij}$, as these are indeed uncorrelated for fixed $i$. Assuming that the drift can indeed be described by Eq. (6), this would however result in an inferior estimator as it ignores available a-priori information.

How to model the drift in Eq. (9) will depend on the problem at hand; any model that is linear in some parameter $\beta$ may be used – many nonparametric curve estimators are of this form, e.g. splines, or local polynomials, see e.g [32]. Usually, one then has to select a bandwidth or some other smoothing parameter; this is a standard problem in non-parametric curve estimation, hence will not be discussed here. We emphasize that the “model-free” approach results in drift estimates that correspond to the raw data in the standard model of non-parametric curve estimation, with the only difference that the errors here are correlated.

2.3 Illustrative 1D simulation

In order to exemplify the points discussed above, we show a small and simplified simulation study of a one-dimensional drift with periodic boundary conditions, i.e. on the circle, represented by $[0, 2\pi)$. The true “image” comprises a sum of three shifted Gaussians on a small constant background, mimicking the event distribution of three features in a sample (Fig. 1(a)). From this, we draw Poisson observations $Z_{ij}$ according to Eq. (2). We use 50
pixels and 15 time bins covering [0, 1]; the maximal signal-to-noise ratio was less than 1.6. The simulated drift was given as \( \sin(1.5 \pi t_i) \). The \( \Delta_i \) were obtained as the points of maximal cross-correlation, based on which estimates for the drift were determined based on methods “first”, “pre”, “model-free”, and a mixed-effects model with a natural cubic spline having 5 degrees of freedom.

To measure the precision of the drift, we determined the integrated squared error (ISE). As discussed above, the drift’s offset is not identifiable; we therefore define the (adapted) ISE by

\[
\min_{\omega} \frac{1}{n} \sum_{i=1}^{n} |\delta_i - \omega|^2,
\]

The minimization removes the ambiguity when measuring the quality of estimation (or by how much two estimations of the drift differ) by means of the ISE. In the following, ISE is therefore always to be understood in this adapted sense defined above. Clearly the minimizing value of \( \omega \) is given by the average of \( (\delta^* - \delta_i) \); thus it suffices to demean the (estimated) drifts prior to comparison.

We repeated this analysis 10,000 times, both with and without applying the Anscombe-like transformation prior to the pair-wise cross-correlation analysis. We then obtained the mean ISE (MISE) by averaging over all simulation runs. The table reports square roots of the MISE (rMISE), i.e. the precision of the drift estimates (Fig. 1(b)). Evidently, the Anscombe-like transformation has a beneficial effect, halving the rMISE. As expected, “pre” resulted in the largest rMISE, roughly 2.5 times larger than with “first”. The “model-free” approach is able to reduce the rMISE by an order of magnitude, as it is able to eliminate the errors \( \epsilon_i \) with increasing \( n \). Finally, the mixed-effects spline model for the drift can also address the inter-class errors \( \eta_i \), resulting in a drift estimate that is visually indistinguishable from the true drift, even with this small sample of very noisy data.

3. Experimental results

3.1. Setup and acquisition protocol

For experimental validation of the estimation methods proposed above, we used a standard SMS-setup (modified from [7, 11, 33]). The excitation laser light was provided by a diode pumped solid-state laser (LC-532-200, Shanghai Laser Century Technology Co., Shanghai, China) at 532 nm. The laser light which was delivered to the sample was intensity-controlled by an acousto-optical tunable filter (AOTFnc-VIS-TN, AA Optoelectronic, Orsay cedex,
France), expanded by a telescope and spectrally cleaned up (Z532/10x, Chroma Technology, Rockingham, VT, USA). For switching the fluorophores, a 375 nm (Cube 375-16C, Coherent Inc., Santa Clara, CA, USA) was used. The green excitation laser beam and the UV laser light were combined at a dichroic mirror (Z415 RDC, AHF Analysentechnik, Tübingen, Germany) and converted from linear to circular polarization by quarter wave plates. Subsequently, the beams were directed into the side port of a commercial wide-field microscope (DMIRE 2, Leica Microsystems, Wetzlar, Germany). Uniform epiillumination of a field of view of approximately 10 µm in diameter was achieved by under-illuminating the back-aperture of the objective lens (HCX APO 100x/1.30 Oil U-V-I 0.17/D, Leica Microsystems, Germany). Moving the sample in the axial direction was done by a single axis nano-positioning system (Nano Z200, Mad City Labs Inc., Madison, WI, USA) which was integrated into a home-built stable scanning block and sample holder. Lateral movement of the sample was done with the help of two piezo-electric motors (Piezo LEGS, PiezoMotor, Uppsala, Sweden; Motion Commander, NANOS-Instruments GmbH, Hamburg, Germany). Further, fine uni-directional movement of the sample in one lateral direction was performed by a piezo-electric actuator (Physik Instrumente, Karlsruhe/Palmbach, Germany). The fluorescence emitted by the sample was collected by the same lens and decoupled from the excitation beam path by a dichroic mirror (Z532/NIR RPC, AHF Analysentechnik, Tübingen, Germany). It was further cleaned up from residual laser light and background light outside the fluorophore’s emission spectrum by a notch filter (NF01-532U-25, Semrock, Rochester, NY, USA) and a band-pass filter (FF01-582/75-25, Semrock, Rochester, NY, USA) and subsequently imaged onto an EMCCD camera (Ixon plus DU-860, Andor Technology, Belfast, Northern Ireland).

The microtubule network (β-tubulin) was immuno-labelled with the photochromic rhodamine RhS [34] in fixed PtK2-cells. For sample preparation, a drop of a water-based solution with 1% (w/v) polyvinyl alcohol (PVA) (88 mol% hydrolyzed, Polysciences Europe, Eppelheim, Germany) and fluospheres (FluoSpheres® carboxylate-modified microspheres, 0.2 µm, crimson fluorescent (625/645) *2% solids, Invitrogen, Carlsbad, CA, USA) was pipetted onto the cells adherent to the coverslip which was then spin-coated for 20 s at 3000 rpm (KW-4A, Chemat Technology, Northridge, CA, USA).

The acquisition protocol was the following. A series of typically 40,000 frames was taken with a frame exposure time of 10 ms and an electron multiplying gain of 60. The excitation intensity was in the range of 10 kW/cm². Switching light was applied when necessary and increased in a stepwise manner from 0 to a few 10 W/cm². Optionally, an experimental drift was applied. Because the sample holder was not guided, the uni-directional piezo movement resulted in movement of the sample with respect to the objective lens with components in both lateral directions, though a dominant component was in the direction of the piezo movement.

The bright fluorescent beads were incorporated into the sample as fiducial markers for retracing the lateral movement during the measurement. For this purpose, their positions were registered during the analysis process. The bead position information can either be used to correct the drift in the determined positions of individual fluorophores during the post-processing analysis or it can be used as a reference value for the results obtained by the proposed methods. More precisely, the bead position in a frame was determined by averaging the bead positions over a span of 1000 frames centered at the respective frame. Usually, one or two beads could be identified per measurement. Because of photobleaching during the measurement, the bead localization precision gradually decreased; nevertheless, the detected photon number was well above 1000 photons throughout the entire measurement. Because of the excess noise factor due to the electron multiplication within the EMCCD, this photon number is effectively halved. According to Thompson et al. [25], the position precision is therefore well below 13 nm full width half maximum (FWHM) (neglecting pixelation effects and background noise). Averaging over a span of 1000 frames theoretically improves this localization precision to well below 0.5 nm. However, one must keep in mind that there are systematic effects, among others the ones which originate in imperfections of the detector, which limit the localization precision. According to Pertsinidis et al., the photo response non-
uniformity between different pixels or within one single pixel, is a major contributor to this
limit and results in distortions of some nanometers over scales of 3-5 pixels and of 0.5 nm
standard deviation (which corresponds to 1.2 nm FWHM) over a 1 pixel scale [35]. This
means that the precision of the localized position of the fiducial marker, when its real position
moves along the CCD-pixel, is limited by a systematic error which does not improve upon
averaging.

3.2. Drift estimation for experimental data in 2D

We will now apply the methodology described in Section 2 to the data thus obtained. Along
the way, we also discuss practical issues.

3.2.1. Choice of parameters and validation

For the experimental data, the quality of the drift correction will also depend on the spatial
and temporal binning of the 2D histograms. First, if the spatial binning is too coarse, i.e. if
the pixel size of the 2D histogram is too large, smaller features of the marker distribution are
lost. If it is too fine, the number of positioned markers per pixel again gets too small, such that
the position histogram no longer represents the overall marker distribution in the sample,
again resulting in outliers. It seems reasonable to apply similar criteria to the pixel size in
standard microscopy where the sampling is usually chosen about twice finer than the standard
deviation of the minimum effective structure size of interest. Often the Nyquist Shannon
sampling theorem [36] is cited as a motivation for this particular choice. Here, the effective
structure is determined by the marker distribution’s $\Sigma_S$ and by the localization precision $\Sigma_A$.
Because these are independent, the total $\Sigma_T$ is given by $\sqrt{\Sigma_S^2 + \Sigma_A^2}$. In our experiments, the
expectation value of the exponential fit of the photon distribution per event is 340 photons.
The resulting average localization precision is 22 nm (FWHM) for the detection PSF of 285
nm (FWHM). The diameter of the microtubules is typically 25 nm, which appears increased
because of the immunolabeling to up to 65 nm. Consequently, in accordance with the above
argument we chose a spatial binning of 30 nm for the histogram.

Second, if the time bins are too large, the temporal and spatial drift variations occurring
within one interval will be smoothed out, i.e. the drift cannot be sampled sufficiently,
resulting in a bias. On the other hand, if the time bins are too small, the events collected in the
respective time interval will not give a meaningful representation of the marker distribution
and therefore, the cross-correlation will return maxima which may be arbitrarily far from the
true drift, i.e. the $\Delta威尔$ contain many outliers. Smaller time bins will also lead to larger inter-
frame variance $\tau^2$ which is ‘invisible’ in the “model-free” approach. Here the time-bin size
takes the role of the smoothing or bandwidth parameter in the mixed-effects model where the
increasing $\tau^2$ is compensated by the increasing number of images and thus pair-wise drift
estimates $\Delta威尔$ when decreasing the time-bin size. Especially for the model-free approach the
time bins should be chosen as small as possible but large enough such that the cross-
correlation analysis does not yet produce obvious (single) outliers. In order to get an estimate
of the sampling rate, we calculated the average number of events per pixel within the imaged
structure. For this purpose, we assigned all pixels with less than 2 events to the background
and divided the sum of the events in the remaining pixels by the number of the remaining
pixels and obtained an average event density of 10 for the measurements shown in the next
paragraphs. When binning 1000 frames together these events are split into 40 time bins, i.e. at
each time point the structure is sampled by one event in every fourth structure pixel. Since the
histogram is spatially binned by pixels of approximately half the smallest effective structure
size, even the smallest structures typically cover an area containing at least a square of 4
pixels thus containing an average of 1 event per time bin. It is therefore not surprising that the
cross-correlation analysis still delivers good results.

These parameter settings (time bin: 1000 frames, space bin: 30 nm) were applied for drift
evaluation of a 40,000 frame long measurement with an experimentally applied drift over
approximately 500 nm (Fig. 2). Here, the sample structure is comprised of 4,070,054
localized events and stretches over 384,869 pixels, resulting in a event density of 10.6 events
per pixel. First, the drift was calculated by using the model-free approach. Here, the localized events close to the bead position had been removed from the data. The determined drift was assigned to all frames in the respective bin. For comparison, the drift was deduced from the bead position. Specifically, the average of the bead positions over a span of 1000 frames was subtracted from all marker positions registered in the respective frame. The drift of the

Fig. 2. Drift estimation methods reconstruct a considerably blurred measurement of RhS-labeled β-tubulin network in a fixed PtK2-cell. (a) Histogram of the raw SMS-data. (b) Histogram of the model-free corrected SMS-data. (c) Histogram of the bead tracking corrected SMS-data. The gray rectangle marks the region where the localized events close to the bead were cut out. (d-f) Close ups of the regions marked by the white rectangles. (g) Profiles in x-direction integrated over y within the areas marked in (d-f). (h) Close up of (g), showing data points (markers) and Gaussian fits (lines) for model-free, bead tracking and for the set of curves “first”, “second”, …, “last”.

boundary frames for which the span of 1000 frames exceeded the first frame or the last frame, respectively, was set to the nearest value.

While the lateral drift of ~500 nm severely blurs the histogram of the raw SMS-data (Fig. 2(a)) the histograms of the model-free (Fig. 2(b)) and the bead tracking (Fig. 2(c)) corrected SMS-data look much sharper. The close ups shown in Fig. 2(d)-2(f) reveal that the quality of the correction which is based on the SMS-data itself is at least as good as the correction based on bead-tracking. From the profiles in Fig. 2(g) it can be seen that both methods reconstruct the shape of the microtubule-filament evenly well and improve its FWHM (drawn from the
Gaussian fit) to about 70 nm. The zoom in on the x-axis of Fig. 2(g) in Fig. 2(h) reveals that the data points and the Gaussian fit of the data which was corrected from pair-wise drift estimates with only one frame ("first", "second", ..., "last") exhibits considerable variance. As expected from the theoretical results in Sections 2.2, the model-free method lies in the middle of the set of curves and is not susceptible to the uncertainty in the drift estimates when relying just on one arbitrary reference bin [24] (e.g. the first one in “first”). A slight difference between the model-free and the bead-tracking corrected curves is also cognoscible. This effect is most likely due to small sample alterations as the model-free approach determines an average drift over the whole field of view, while bead-tracking is a local method. This example demonstrates the applicability of the model-free approach to perform drift compensation of even strongly motion-blurred experimental data. Similar results were obtained for 14 measurements (data not shown). This means that apparently spoilt data can be recovered without the necessity to use artificially inserted fiducial markers.

3.2.2 Comparing different bin sizes and quality checks

Having shown the method’s general applicability, the above somewhat crude parameter choice was then re-examined. To this end, we varied the spatial bin size from 10 nm to 120 nm and the time bin size from 250 to 6000 frames, determining the drift based on the model-free approach for each pair of parameter values. These analyses were undertaken on a set of 14 similar experimental data sets (data not shown) and they support the choice of the above motivated spatial binning of 30 nm. More precisely, the drift estimation seems to be robust against a reasonable change of the spatial binning within ± 50%, while simultaneously keeping the temporal bin size constant. Further, there is a tendency to produce outliers in the cross-correlation for spatial or temporal bin size much smaller than our settings. Here, this behaviour is exemplified by the comparison of two parameter sets for the experimental data shown in section 3.2.4: Fig. 3 shows the drift estimates drawn from the model-free approach

![Image](image_url)

Fig. 3. Effects of different temporal and spatial binning on the drift estimates of real data. (a) Comparison of drift estimates $\delta$ for the y-direction (for the data shown in section 3.2.4) derived from the model-free approach. (black) time bin = 1000 frames, space bin = 30 nm; (red) time bin = 250 frames, space bin = 20 nm. (b) quantile plot for the data in (a). The inset shows the whole data range. Note that the right y-axis covers a range from -7000 to +7000 nm.

with the above motivated parameter set (time bin: 1000 frames, spatial bin: 30 nm) (black line) and the drift estimates for reduced time and space bin sizes (time bin: 250 frames, space bin: 20 nm) (red line). The sample structure is comprised of 1,270,025 localized events and stretches over 139,779 pixels, resulting in a event density of 9.1 events per pixel. Clearly, too few events are binned together in space and time for the second parameter set which leads to the appearance of significant single-frame outliers in the drift estimate $\delta$ plotted against time (Fig. 3(a)). The amount of such outliers is therefore a reasonable control to check whether valid parameters where chosen.
While systematically analyzing data with different parameter settings is a perfectly valid pathway to an optimized drift correction it is important to know that there is a way to detect the presence of outliers and thus an invalid choice of parameters directly, without this extra effort, by standard techniques for residual diagnostics, cf. e.g [37]. In particular, quantile plots – a.k.a. q-q-plots, or probability plots [38] – can usefully be employed to detect outliers by plotting the empirical quantiles against the theoretical quantiles of a standard normal distribution (Fig. 3(b)). We observe that the outliers produced by the too small choice for the time and space bin size (250 frames and 20 nm respectively) are immediately visible in such a plot from the pronounced sigmoidal shape which is absent for the more moderate parameter choice (1000 frames, 30 nm). The most extreme points in these plots are very volatile, so an indicator for poor parameter choice is a sigmoidal shape in the center of the plot. However, the zoomed-out graph shown in the inset of Fig. 3(b) is useful to illustrate that some pair-wise drift estimates can deviate from the mean drift estimate by several micrometers for the too small parameter set (Fig. 3(b), inset).

3.2.3 The method’s performance: smoothing and discontinuities

Until now, we have not made any assumptions on the type of drift. However, as the typical drift in a microscope setup – e.g. caused by thermal expansion or by a continuous drift of moving parts – is smooth, we can take advantage of this. We use the linear mixed effects model (lme) from Eq. (9) using cubic splines, resulting in a nonparametric curve estimator, which also addresses the inter-image errors \( \eta \). Figure 4(a) compares the performance of this lme model with the model-free approach. For further evaluation, we then plotted the positions of the two beads which were incorporated in the sample. The beads’ signals started with 25,000 and 15,000 photons and decreased exponentially to 1200 and 1000 photons at the end of the 40,000 frames. Consequently, the localization precision is better than 13 nm for all frames. When averaging over 1000 frames and accounting for the photo response non-uniformity of the CCD-camera on the scale of 1 pixel, which is presumably existent for most CCD-cameras [35], this precision is likely improved to the range of 1-2 nm. However, when comparing the two drift curves extracted from the two beads (Fig. 4(a)), it becomes obvious that the bead position must not be taken as the absolute reference for the samples drift. Indeed, the square root of ISE (rISE) of the drift estimates calculated from the two beads is 11.8 nm which is much larger than the expected localization precision. This finding is further supported by analyses of 4 additional data sets (data not shown). The results of the model-free approach are
equally far away or even closer to each bead than the beads are to each other and the drift determination obtained by the linear mixed effects model is equally good. This explains why the model-free and the bead tracking approaches do not agree exactly in Fig. 2(g), 2(h). This suggests that our data driven method determines the drift at least as well as the bead tracking for this type of measurements. Figure 4(a) also shows the drift estimates obtained from estimating the drift from pair-wise drift estimates with only one frame ("first", "second", "third", ...) which clearly exhibit a much larger variance.

For this example, the presented drift estimation algorithm works thus reliably for continuous drift functions which were modeled by a cubic spline. However, in the experimental routine, discontinuous perturbations (i.e. jumps of the sample position) might occur as well, e.g. caused by accidental physical contact with the setup.

In order to mimic this situation, an artificial jump of 100 nm in the x-direction was introduced into the raw data. Both drift estimation algorithms (model-free and lme) were applied again as described above and the resulting drift estimates are shown in Fig. 4(b). The position of the jump is clearly identifiable with both methods. Whereas the model-free approach tracks the shape of the jump very well, the spline-based linear mixed effects model tends to smooth the edges on account of its assumption of continuity, as expected. The drift estimation can thus be used to either correct the drift in the whole range of frames or to identify the frames near the jump and discard the corresponding events.

3.2.4 Drift compensation is essential for quantitative analysis

The advantage of drift correction in super-resolution imaging in case of large drifts, as exemplified in Fig. 2, is obvious. However, even for small drifts on the scale of some tens of nanometers or even less, drift correction is beneficial. It is even essential when it comes to quantitative image analysis. Figure 5 presents the 2D histogram of “raw”, model free corrected and bead tracking corrected superresolution data. Here, the drift was about 90 nm, which is only ~1.5 times the smallest effective structure size within the images. Although, the presence of drift is not obvious at first glance (Fig. 5(a)), it has a considerable effect during quantitative analysis. For example, the doughnut like shape shown in Fig. 5(b) gets only unveiled after drift compensation. Again, the data driven and the bead tracking method agree and yield a similar result, which is also supported by the profiles. The reason for this effect is the temporal appearance of the individual events which constitute the nanoscale feature: in this example they were registered over an extended period of time and therefore their position was spread by the drift. The same effect can also lead to an increase in apparent FWHM of individual filaments. However, if the events constituting two different features stem from rather small but distant time windows, for example one from the beginning and the other from the end of the measurement, the local resolution is not considerably affected but the drift will change the relative position in the images. This is exemplified in Fig. 5(c) where the distance between two filaments seems to be 512 nm in the “raw” image, while the drift corrected images depict that the true distance is about 40 nm less (469 nm and 472 nm for model-free and bead tracking correction respectively).

4. Conclusion

Movement of the sample with respect to the objective lens, even if slight, is an important issue in optical microscopy. Especially when the local image acquisition time is long and the resolution is high, as it is the case in SMS based nanoscopy schemes. If the stability of the experimental setup does not limit mechanical drifts to less than about half the resolution, drift detection is indispensable, in order to allow for the correction of movement induced image...
Fig. 5. Drift compensation is essential for quantitative analysis. (a) Position histograms generated from the raw SMS data (top), drift-corrected by the model-free approach (middle) and by bead-tracking (bottom). The gray rectangle marks the region where the localized events close to the bead were cut out. (b) Close-ups and corresponding profiles of the region marked by the small white rectangle in (a). (c) Close-ups and corresponding profiles of the region marked by the large white rectangle in (a). The positions of the localized events shown in (b) and (c) were rotated by 40° before plotting the histograms.

degradations and distortions. Importantly, sample movement can also occur if the setup itself is perfectly stable e.g. by cells moving relative to the cover slip making data-based detection of motion indispensable.

In this study, we have therefore analyzed the problem of drift detection in SMS based nanoscopy schemes, and propose a versatile method to deduce the drift from the SMS-data itself. This renders the use of a detection scheme for lateral drift obsolete and thus reduces the overall experimental complexity. For example, fiducial markers do not have to be incorporated into the sample, image acquisition is not confined to regions close to a fiducial and acquisition parameters can be freely tuned, as they do not have to be matched to the marker’s brightness and spectrum (and vice versa). Our study further suggests that for certain scenarios using the data itself could well turn out to be more reliable than the use of fiducials which may move or fade during data acquisition.

While we used a standard cross-correlation analysis to determine the pairwise shifts between time points, other techniques like Fourier methods together with techniques from M-estimation [39] could also have been combined with our estimation scheme. In particular, our specific assumptions on the raw data itself were necessary only for our specific way to determine the pairwise shifts; for the combination of these estimates we only have to assume that they were derived from independent frames, and have constant variance. It would be worthwhile to study in detail the specific strengths and weaknesses in terms of spatial and...
temporal precision of different pairwise shift determination methods for SMS data, which may well depend on the structure being imaged. Irrespective of the method chosen the important result of this study is that incorporating all possible pairwise shifts into the estimation procedure is by far superior to comparing each time point with a single reference point. This should allow application of our approach even to structures for which drift correction hitherto seemed impossible due to the large errors of the pairwise estimates. Importantly in this context, our strategy allows for calculating not only the drift estimate itself but also delivers a quality measure for its precision, which is important when it comes to the question whether a specific data set can be drift corrected, or should be rejected or when different pairwise determination methods are to be compared and tested for applicability. Furthermore, the specific choice of parameters (spatial and temporal bin size in case of cross correlation analysis) can reliably be tested by q-q-plot diagnostics, which may well enable the design of an automatic drift correction system.

Under the experimental conditions used for this study, our method can determine sample drift at least with the precision which can be reached by tracking fluorescent particles. In order to investigate whether the precision of our method is even higher, our ranking method has to be refined. If a priori knowledge about the type of drift is available, for example smoothness, the precision can be further increased by fitting a specific drift curve to the pairwise data. Even discontinuity detection is possible.

We have restricted ourselves to the drift detection in 2D, however, our method is readily expandable to 3D in case such SMS-based nanoscopy data is present. For example one could project the 3D data along the x, y and z-direction respectively [24] and determine the pairwise shifts from these 2D histograms. Another straightforward possibility is to determine the pairwise drifts directly from the 3D data itself and to expand our model to 3D even though this strategy is much likely to require a higher number of localized events than the projection approach.

In future our methodology might be also applicable to cases where morphological deformations change the object during the measurement process. Finally, we want to mention that the method presented here will offer new calibration strategies to other fields of imaging, where only partial information of the probe is available in each acquisition step and the probe changes unexpectedly during the entire measurement process.

Appendix: derivation of the “model-free” estimator

The “model-free” estimator by definition is the one which minimizes the sum of least squares,

\[ \sum_{i,l} (\Delta_{il} - \delta_i + \delta_l)^2, \]

where indices \(i\) and \(l\) in the sum run from 1 to \(n\). This expression clearly is invariant under the addition of some offset \(\omega\) to the drift \(\delta\), so its minimum remains unchanged however we choose the offset. Fixing the offset such that the mean drift is 0, cf. Equation (10), thereby is equivalent to adding twice the squared sum of the drift estimates to the functional, i.e. we minimize the functional

\[ \sum_{i,l} (\Delta_{il} - \delta_i + \delta_l)^2 + 2 \sum_i \delta_i \]

Differentiating with respect to \(\delta_i\) immediately gives

\[ 2 \sum_l (\Delta_{il} - \Delta_{il} - 2\delta_i + 2\delta_l) + 4 \sum_l \delta_l = 4n\delta_i - 4 \sum_l \Delta_{il}, \]

where we used the antisymmetry of \(\Delta\). The Hessian of the functional is therefore \(4n\) times the identity matrix, showing that the estimator is uniquely determined; setting the last expression to zero, we finally obtain
\[ \delta_l = \frac{1}{n} \sum \Delta_l, \]

which shows that the “model-free” estimator is indeed given by averaging the inferior estimates \( \Delta_l \) over \( l \) as claimed in Section 2.2.

Acknowledgment

We thank R. Pick for aid in the design of the mechanics, E. Rothermel and T. Gilat for technical assistance, and V. Belov for providing us with Rhodamine S. We acknowledge the contribution of L. Wang to the data acquisition. T.H., A.M., S.W.H., A.S. acknowledge support of BMBF INVERS - 03MUPAH6 - Health and Medical Technology. This work was supported by a grant of the Deutsche Forschungsgemeinschaft to A.E. and A.M. (SFB 755).