

Measuring localization performance of super-resolution algorithms on very active samples

Steve Wolter,¹ Ulrike Endesfelder,^{1,2} Sebastian van de Linde,¹
Mike Heilemann,^{1,2} and Markus Sauer^{1,*}

¹ *Biotechnology and Biophysics, Julius-Maximilians-Universität Würzburg, Germany*

² *Angewandte Laserphysik & Laserspektroskopie, Bielefeld University, Germany*

*m.sauer@uni-wuerzburg.de

<http://super-resolution.de>

Abstract: Super-resolution fluorescence imaging based on single-molecule localization relies critically on the availability of efficient processing algorithms to distinguish, identify, and localize emissions of single fluorophores. In multiple current applications, such as three-dimensional, time-resolved or cluster imaging, high densities of fluorophore emissions are common. Here, we provide an analytic tool to test the performance and quality of localization microscopy algorithms and demonstrate that common algorithms encounter difficulties for samples with high fluorophore density. We demonstrate that, for typical single-molecule localization microscopy methods such as *d*STORM and the commonly used rapid*d*STORM scheme, computational precision limits the acceptable density of concurrently active fluorophores to 0.6 per square micrometer and that the number of successfully localized fluorophores per frame is limited to 0.2 per square micrometer.

© 2011 Optical Society of America

OCIS codes: (100.6640) Superresolution; (180.2520) Fluorescence microscopy; (350.4800) Optical standards and testing.

References and links

1. S. W. Hell, "Far-Field Optical Nanoscopy," *Science* **316**, 1153–1158 (2007).
2. P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, and M. G. L. Gustafsson, "Super-resolution video microscopy of live cells by structured illumination," *Nat. Methods* **6**, 339–342 (2009).
3. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution," *Science* **313**, 1642–1645 (2006).
4. S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophys. J.* **91**, 4258–4272 (2006).
5. B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science* **319**, 810–813 (2008).
6. M. Heilemann, S. van de Linde, M. Schüttelpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer, "Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes," *Angew. Chem. Int. Ed.* **47**, 6172–6176 (2008).
7. J. Vogelsang, T. Cordes, C. Forthmann, C. Steinhauer, and P. Tinnefeld, "Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy," *Proc. Nat. Acad. Sci. U.S.A.* **106**, 8107–8112 (2009).
8. R. Wombacher, M. Heidbreder, S. van de Linde, M. P. Sheetz, M. Heilemann, V. W. Cornish, and M. Sauer, "Live-cell super-resolution imaging with trimethoprim conjugates," *Nat. Methods* **7**, 717–719 (2010).

9. H. Shroff, C. G. Galbraith, J. A. Galbraith, and E. Betzig, "Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics," *Nat. Methods* **5**, 417–423 (2008).
10. T. Klein, A. Löschberger, S. Proppert, S. Wolter, S. van de Linde, and M. Sauer, "Live-cell dstorm with snap-tag fusion proteins," *Nat. Methods* **8**, 7–9 (2011).
11. N. Bobroff, "Position measurement with a resolution and noise-limited instrument," *Rev. Sci. Instrum.* **57**, 1152–1157 (1986).
12. M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophys. J.* **81**, 2378–2388 (2001).
13. R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise Nanometer Localization Analysis for Individual Fluorescent Probes," *Biophys. J.* **82**, 2775–2783 (2002).
14. K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg, "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nat. Methods* **7**, 377–381 (2010).
15. C. Shannon, "Communication in the Presence of Noise (reprinted)," *Proc. IEEE* **72**, 1192–1201 (1984).
16. S. van de Linde, S. Wolter, M. Heilemann, and M. Sauer, "The effect of photoswitching kinetics and labeling densities on super-resolution fluorescence imaging," *J. Biotechnol.* **149**, 260–266 (2010).
17. T. Cordes, M. Strackharn, S. W. Stahl, W. Summerer, C. Steinhauer, C. Forthmann, E. M. Puchner, J. Vogel-sang, H. E. Gaub, and P. Tinnefeld, "Resolving single-molecule assembled patterns with superresolution blink-microscopy," *Nano Lett.* **10**, 645–651 (2010). PMID: 20017533.
18. M. F. Juette, T. J. Gould, M. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Be-wersdorf, "Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples," *Nat. Methods* **5**, 527–529 (2008).
19. S. Manley, J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz, "High-density mapping of single-molecule trajectories with photoactivated localization microscopy," *Nat. Methods* **5**, 155–157 (2008).
20. M. B. J. Roeffaers, B. F. Sels, H. Uji-i, F. C. De Schryver, P. A. Jacobs, D. E. De Vos, and J. Hofkens, "Spa-tially resolved observation of crystal-face-dependent catalysis by single turnover counting," *Nature* **439**, 572–575 (2006).
21. M. B. J. Roeffaers, G. De Cremer, J. Libeert, R. Ameloot, P. Dedecker, A.-J. Bons, M. Bückins, J. A. Martens, B. F. Sels, D. E. De Vos, and J. Hofkens, "Super-resolution reactivity mapping of nanostructured catalyst parti-cles," *Angew. Chem. Int. Ed.* **48**, 9285–9289 (2009).
22. H. Bornfleth, K. Sätzler, R. Eils, and C. Cremer, "High-precision distance measurements and volume-conserving segmentation of objects near and below the resolution limit in three-dimensional confocal fluorescence micros-copy," *Journal of Microscopy* **189**, 118–136 (1998).
23. S. Holden, S. Uphoff, and A. Kapanidis, "Crowded-field photometry increases maximum super-resolution imag-ing density by an order of magnitude," *Nat. Methods* (2010). Manuscript submitted.
24. S. Wolter, M. Schüttelpeiz, M. Tscherepanow, S. van de Linde, M. Heilemann, and M. Sauer, "Real-time compu-tation of subdiffraction-resolution fluorescence images," *J. Microsc.* **237**, 12–22 (2010).
25. J. Tang, J. Akerboom, A. Vaziri, L. L. Looger, and C. V. Shank, "Near-isotropic 3D optical nanoscopy with photon-limited chromophores," *Proc. Nat. Acad. Sci. U.S.A.* **107**, 10068–10073 (2010).
26. T. Quan, P. Li, F. Long, S. Zeng, Q. Luo, P. N. Hedde, G. U. Nienhaus, and Z.-L. Huang, "Ultra-fast, high-precision image analysis for localization-based super resolution microscopy," *Opt. Express* **18**, 11867–11876 (2010).
27. D. A. Grossman and O. Frieder, *Information Retrieval: Algorithms and Heuristics*, The Kluwer International Series of Information Retrieval (Springer, P.O.Box 17, 3300 AA Dordrecht, The Netherlands, 2004), 2nd ed.
28. A. Technology, *iXon camera manual* (Andor Technology, 7 Millennium Way, Springvale Business Park, Belfast, BT12 7AL, NORTHERN IRELAND, 2008).
29. M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, M. Booth, and F. Rossi, *Gnu Scientific Library: Refer-ence Manual* (Network Theory Ltd., 2003).
30. M. Matsumoto and T. Nishimura, "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Trans. Model. Comput. Simul.* **8**, 3–30 (1998).
31. S. Ehrlich, "Error bounds for Gauss–Kronrod quadrature formulae," *Math. Comp.* **62**, 295–304 (1994).
32. D. M. Thomann, "Algorithms for detection and tracking of objects with super-resolution in 3d fluorescence microscopy," Ph.D. thesis, ETH Zürich (2003).
33. R. Henriques, M. Lelek, E. F. Fornasiero, F. Valtorta, C. Zimmer, and M. M. Mhlanga, "QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ," *Nat. Methods* **7**, 339–340 (2010).
34. T. A. Laurence and B. A. Chromy, "Efficient maximum likelihood estimator fitting of histograms," *Nat Meth* **7**, 338–339 (2010).
35. N. A. Frost, H. Shroff, H. Kong, E. Betzig, and T. A. Blanpied, "Single-molecule discrimination of discrete perisynaptic and distributed sites of actin filament assembly within dendritic spines," *Neuron* **67**, 86 – 99 (2010).
36. U. Endesfelder, S. van de Linde, S. Wolter, M. Sauer, and M. Heilemann, "Subdiffraction-resolution fluorescence microscopy of myosin-actin motility," *Phys. Chem. Chem. Phys.* **11**, 836–840 (2010).

1. Introduction

In recent years, subdiffraction-resolution far-field fluorescence microscopy methods attracted considerable interest because they allow the noninvasive observation of cellular processes with almost molecular resolution [1–8].

Single-molecule based localization microscopy methods such as photoactivated localization microscopy (PALM, [3]), fluorescence photoactivation localization microscopy (FPALM, [4]), stochastic optical reconstruction microscopy (STORM, [5]), and direct stochastic optical reconstruction microscopy (*d*STORM, [4, 8]) are very promising because they are comparably simple, require only moderate irradiation intensities which makes them ideally suited for live-cell imaging [8–10], and can be implemented in standard wide-field fluorescence microscopes. In localization microscopy target structures are labeled with fluorophores that can be stochastically switched between a fluorescent on state and a non-fluorescent off state upon irradiation with light of appropriate wavelength. Ideally, only a small subset of fluorophores is active at any time of the experiment generating isolated images of the point spread function (PSF), so-called *spots*, on the detector that are well resolvable within the optical resolution limit. A model function is approximated to the PSF [11–14] to determine the nanometer-precise fluorophore position called *localization*. A cycle of activation, detection, and photobleaching or transfer to a reversible off state, respectively, is repeated several thousand times to collect a sufficient number of localizations (typically several ten thousands to millions of localizations) to reconstruct a super-resolved image.

When considering the resolution of localization microscopy, two quantities have to be distinguished: The optical resolution, i.e. the shortest distance at which two point emitters can be distinguished, and the structural resolution, i.e. the finest resolvable level of detail in a continuous structure. Contrary to many classical microscopy methods, the structural resolution is potentially much lower than the optical resolution due to the nonlinear computational processing.

Achieving high structural resolution with localization microscopy has three prerequisites: The fluorophore density must be sufficiently high according to the Nyquist-Shannon sampling theorem [15], the imaging speed must surpass the sample dynamics, and the density of generated spots must be small enough to be optically resolvable to avoid imaging artifacts and false localizations [9, 16, 17]. While these prerequisites are easily fulfilled as long as only extended filaments, e.g. microtubulin and actin, are imaged, super-resolution imaging of complex and densely labeled structures necessitates the use of photoswitchable fluorophores with highly stable off states and appropriately set photoswitching rates [16, 17]. Three-dimensional and dynamic super-resolution imaging with high spatiotemporal resolution [5, 8, 9, 18, 19] or single turnover counting for spatially resolved observation of catalysis [20, 21] are even more challenging. In these cases, precise photophysical control of the number of active fluorophores is difficult and high spot densities are often unavoidable.

While spot density issues have been researched for localization microscopy of fluorescent particles [22], only recently [23] an algorithm was implemented for single-molecule super-resolution data analysis based on simultaneously fitting the data of overlapping PSF with multiple kernels to increase the allowed number of simultaneously active fluorophores. Unfortunately, current algorithms used for single-molecule localization microscopy mainly focus on spatial localization precision and computational speed and rely on well-separated input spots [24–26]. The lack of efficient algorithms results, in our opinion, primarily from the difficulties to check the performance of the algorithm on experimental data obtained from densely labeled samples. While it has been shown previously [24] that stochastic simulations can generate data sets sufficiently close to experimental localization microscopy data, direct stochastic analysis of suchlike simulations is difficult because the density of active fluorophores prevents

direct mapping of active fluorophores and localizations (Fig. 1a).

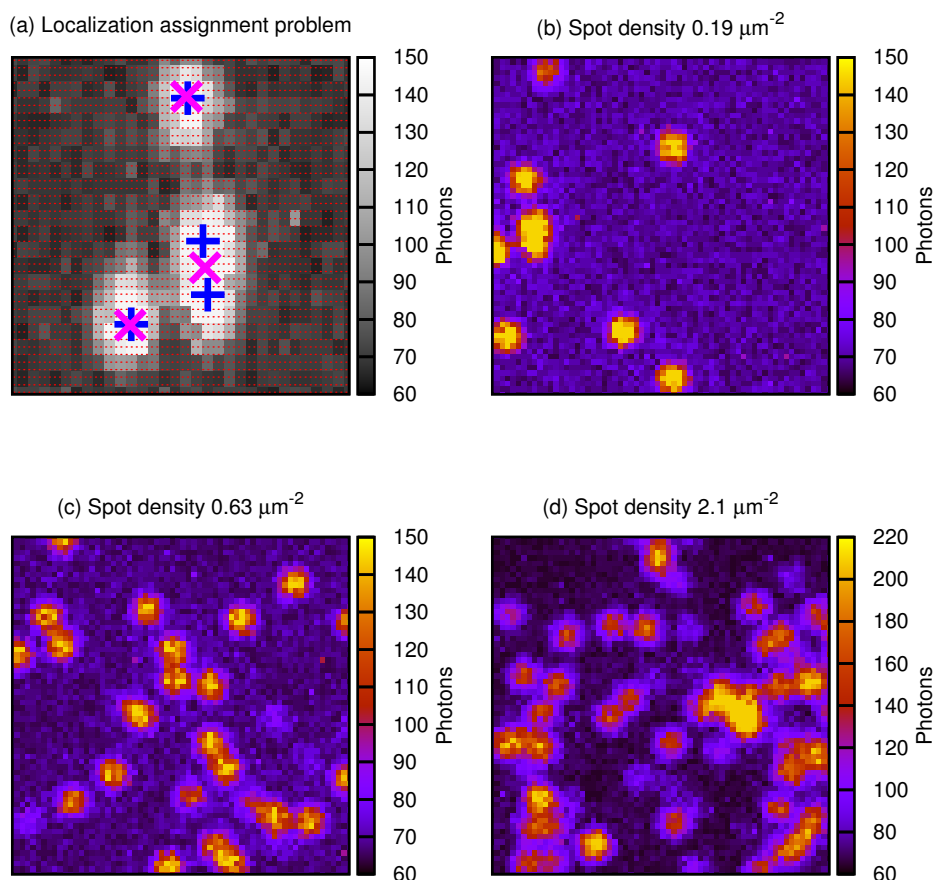


Fig. 1. Illustration of localization assignment problem and typical input images at different spot densities. (a) Example of typical multi-spot error. Red dots mark simulated fluorophores, with blue plusses marking active fluorophores. The resulting signal is indicated with grey values in the background, and a possible set of localizations is displayed with purple crosses. The localization for the multi-spot event is clearly false and would bias the localization accuracy for both close blue fluorophores if it was assigned to either localization and thus must be counted as false positive localization, but distinguishing between such multi-spot localizations and correct single-spot localizations is not trivial. (b-d) Examples for generated input images at different photon counts. Note that the photon count scale has an unknown offset since the background noise was determined experimentally.

Here, we introduce a method to evaluate the performance of localization microscopy imaging algorithms on samples with high fluorophore density using simulated fluorophore lattices. Our method is capable of quantifying the performance of a localization microscopy algorithm with three standard characteristics: *stochastic precision*, *recall*, and *spatial precision*. The stochastic precision is defined as the quotient of true positive localizations to all localizations found, and the recall as the quotient of the number of true positive localizations to the total

number of spots that should have been localized [27]. The spatial precision gives the uncertainty in fluorophore localization, that is, the spatial difference between the exact position of an emitting fluorophore and the determined localization. (Stochastic precision and recall are often referred to as false positive/negative rate or miss/hit probability. Spatial precision is also known as localization precision, precision or optical resolution, renamed here to avoid confusion with stochastic precision.) To demonstrate its performance, we applied our method to the rapidSTORM algorithm [24], an algorithm in the important class of Gaussian PSF least-squares fitters to determine its performance on typical dSTORM data.

2. Material and methods

We measured the impact of high spot densities on algorithmic localization performance by simulating fluorophores located on a dense lattice cycling reversibly between a fluorescent on and a non-fluorescent off state. While this method can generate signals sufficiently close to experimental data, straightforward comparison of the simulation ground truth, i.e. the known position of simulated emissions, and the computational result, i.e. the position of algorithmically recognized localization positions, has proven difficult under dense-spot conditions since different simulated fluorophores can cause any given localization. Our methodical solution to this problem is an analysis based on a localization histogram from a large number of analyzed frames with all lattice points averaged into a mean lattice interval. In other terms, we (i) simulated a long localization microscopy image stack, (ii) employed the algorithm under scrutiny to find localizations in the stack, (iii) subtracted the position of the nearest fluorophore from each localization, regardless of this fluorophore's state, (iv) fitted the histogram of these localization offsets with a sum of Gaussian functions, and (v) extracted the observables of interest from the obtained parameters.

2.1. Input data simulation

We modelled localization microscopy input data as given in Eq. 1. $PSF(f, p)$ models the point spread function for a simulated fluorophore f at the pixel p , i.e. the probability that a photon emitted by the fluorophore f (assumed to be point-like) is detected in the pixel p . The photon rate N_p gives the number of photons emitted per time unit while residing in the on state, and this rate was assumed to be constant. $t_{on}(f, t)$ denotes the time the fluorophore f resided in the on state during a time period t . The product of these three quantities, assumed to be the number of photons detected by one pixel of a simulated CCD camera for one image and one fluorophore, was varied by Poisson statistics (denoted by Pois). No additional camera properties were considered because modern scientific CCD cameras come very close to linear response [28]. The contributions were summed over a set of fluorophores F and additive background noise was modelled by randomly choosing a value G_r out of a set of likely background noise values G .

$$S(p, t) = G_r + \sum_{f \in F} \text{Pois}[t_{on}(f, t)N_p PSF(f, p)] \quad (1)$$

G was generated by selecting all pixels further than 10 pixels from all localizations from a real dSTORM acquisition. As for other random numbers in this article, the randomness was drawn from the GSL Mersenne Twister implementation [29, 30].

F was generated by placing one fluorophore on each junction of a 40 nm rectangular lattice. Fluorophore behavior was modeled as a time-continuous Markov process between a dark and a bright state with lifetimes of τ_{off} and τ_{on} , respectively. In other terms, the time a fluorophore spends in each state follows an exponential distribution with mean τ_{off} and τ_{on} , respectively, and each bright state phase is followed by a dark state phase and vice versa. The simulated

point spread function was computed by assuming perfect focusing and sample planarity, i.e. an optical point spread function equal to the Besselian of first kind and first order J_1 , scaled by a factor κ to match common experimental numerical apertures, and integrated over the camera pixel size numerically using 87-point Gauss-Kronrod [31] integration (Eq. 2). \vec{x}_f denotes the subpixel-precise fluorophore position in this equation, and α a scale factor chosen such that $PSF(f, p) = 1$. While this formulation assumes point-like fluorophores, extension to two-dimensional or three-dimensional objects is straightforward.

$$PSF(f, p) = \alpha \int_{x_p \in p} \frac{J_1(\kappa \|\vec{x}_p - \vec{x}_f\|)}{\|\vec{x}_p - \vec{x}_f\|^2 \kappa^2} dx_p \quad (2)$$

The average lifetime of the on-state was chosen to three times the simulated acquisition time for a single image, but not synchronized with the simulated acquisition time intervals, to produce simulated images with a broad spectrum of spots with different photon counts. This mainly serves to emulate the broad distribution of spot photon counts experienced in real TIRF experiments. We simulated fluorophores spaced on a 40 nm lattice with a detection pixel raster of 85 nm, causing many different detection raster/fluorophore lattice orientations to occur. (The raster and lattice were chosen with a small lowest common multiple to ensure computational tractability. This way, only a limited number of PSFs had to be computed.) The Besselian PSFs was scaled with $\kappa = 1.37$, equivalent to a spot full width at half maximum (FWHM) of 370 nm.

2.2. Simulation parameters

Three parameters were varied to identify influences in spot detection rate: The density of spots on the camera, the signal-to-noise ratio and the sampling raster width. The spot density was varied by prolonging the simulated lifetime of the dark state while keeping the on state lifetime and fluorophore density constant; the signal-to-noise ratio was varied by changing the simulated photon emission rate of fluorophores while keeping the background noise constant; and the sampling raster width was simulated by changing the simulated pixel raster while keeping the PSF size constant.

By default, we used physical parameters similar to typical *d*STORM experiments: integration time 0.1 s, $\tau_{\text{on}} = 0.3$ s, $N_p = 10$ kHz and a pixel size ρ_p of 0.24 PSF FWHM. Each photon was counted as 16 A/D counts in the linear part of the camera response.

Typical images generated with these parameters are shown in Fig. 1 and in the boxes in Fig. 4.

2.3. Algorithmic molecule localization

We processed the simulated images using our previously published *rapid*dSTORM algorithm [24]. The *rapid*dSTORM algorithm is optimized for comparably noisy input images as commonly recorded in widefield single-molecule localization microscopy, and processes input images in two steps. In the first step, likely positions of bright-state fluorophores are pre-selected by applying a suitable smoothing algorithm, such as a *Spalttiefpass* filter, to the input images and selecting the local maxima of the resulting image. In the second step, a Gaussian function with fixed covariance matrix (i.e., the widths σ_x and σ_y and the X-Y-correlation are either estimated manually or automatically prior to the fitting) is fitted to the pixels around the strongest local maximums. By thresholding the amplitude of the fitting Gaussian, a distinction is made between random background noise and a real fluorophore emission. Strong local maxima are fitted in decreasing intensity until a predefined number of successive maxima were fitted with an amplitude below the threshold. We used an amplitude threshold of 180 times the noise standard deviation and filtered spots containing emissions from multiple fluorophores (*multi-spots*)

by fitting the data from the first spot with a sum of two Gaussian kernels [32]. The start positions of the centers were chosen 1 pixel apart along a line connecting the one-kernel center and the highest residues. The start amplitudes were set to half of the one-kernel amplitude. All localizations were tagged with the quotient of the sum of squared residues of the two-kernel and the one-kernel fit, termed their suspectedness. If the two centers found in two-kernel analysis differed by more than a threshold θ_{dist} , the two-kernel fit was instead discarded and the suspectedness set to 0. After computing all results, we discarded localizations that surpassed a suspectedness threshold of θ_{fishy} .

Three parameters were varied in the data processing: First, we used two different spot finding methods (*Spalttieypass* and Gaussian smoothing, see [24]) to check whether the used spot detection influences the results. Second, we changed the distance threshold θ_{dist} and the suspectedness threshold θ_{fishy} to increase the probability to generate at least one reasonably good set of parameters.

By default, we used a *Spalttieypass* smoother, $\theta_{\text{dist}} = 0.5 \mu\text{m}$ and $\theta_{\text{fishy}} = 0.1$.

To control our measurements with an independently implemented algorithm and to show the easy integration of our method with other algorithms, we also computed the images generated with default photophysical parameters using QuickPALM [33] version 1.1 with its default settings.

2.4. Statistical characterization of localizations

The spot density was characterized by calculating the area within a circle with a diameter of one PSF FWHM. The average density of simulated fluorophores, the average fraction of the acquisition time each fluorophore spent in the on state and the average density of spots was calculated to determine the average number of spots per area within a circle of one FWHM diameter by linear arithmetics. We did not correct for integration time.

We analyzed the localization distribution generated by the *rapidSTORM* algorithm by histogramming the offsets of localizations relative to the known fluorophore positions at the lattice points, excluding a border of 5 pixels resulting in a two-dimensional point distribution representing a mean lattice interval. We fitted the histogram with a sum of a 5 by 5 lattice of symmetrical, identical two-dimensional Gaussian functions centered at the theoretical lattice points with a width of σ plus a constant offset B to the data, resulting in the model function given by Eq. 4. The histogram was fitted with the Levenberg-Marquardt maximum likelihood estimator published by Laurence et al [34].

$$K(\vec{x}, \vec{x}_0) = \frac{A}{2\pi\sigma^2} \exp\left(-\frac{\|\vec{x} - \vec{x}_0\|^2}{2\sigma^2}\right) \quad (3)$$

$$H(\vec{x}) = B + \sum_{x_c=-2}^2 \sum_{y_c=-2}^2 K\left(\vec{x}, \begin{pmatrix} x_c \\ y_c \end{pmatrix}\right) \quad (4)$$

The width of the Gaussian functions directly gives the spatial precision. The localizations explained by the Gaussian functions give the number of true positive localizations. On the other hand, the number of localizations explained by the shift gives the number of unspecific localizations (false positives), which include erroneously fitted background noise and localizations stemming from multi-spots. Both of these sources of false localizations can be expected to produce localizations with a very broad distribution, thus appearing identically distributed on the mean lattice interval. The total number of spots that should have been detected was determined from the number of spots in the simulation that contained enough photons above the background threshold. From these values, we computed stochastic precision and recall accordingly.

Due to multiple stochastic simplifications made, we computed each stochastic simulation 5 times with random number generation seeds 41–45 to gain information about data point validity. We performed these computations on an Intel(R) Core(TM) i5 CPU 650 clocked at 3.20 GHz with four cores, using close to a week of computer time for the whole set of data found on our website. The program code and additional scripts we used for generating and evaluating data sets can be found on our website.

3. Results and discussion

The most researched characteristic for super-resolution microscopy is its spatial precision, i.e. for localization microscopy the stochastic uncertainty in each localization. By applying our evaluation method, we found the spatial precision of rapidSTORM to be very stable and decreasing only by a few nanometers for spot densities up to 5 spots per μm^2 (Fig. 2), corresponding to an average distance to the nearest neighbouring molecule of 0.7 FWHM of a PSF. The deterioration of spatial precision for higher spot densities cannot be accurately determined due to the very small stochastic precision. The decrease in spatial precision is consistent across a wide range of parameter variations, including photon count rate, pixel size, smoothing algorithm, and multi-spot search thresholds.

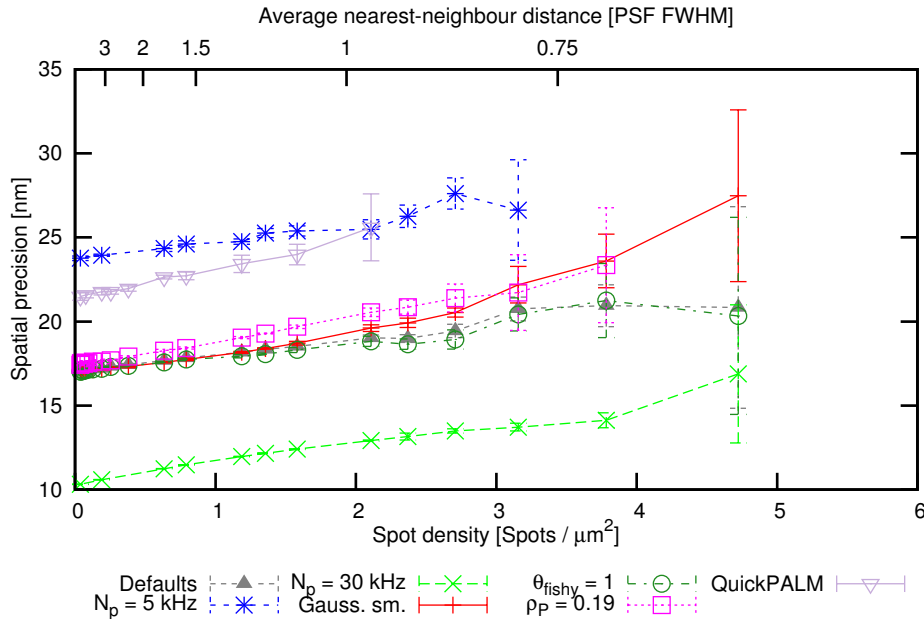


Fig. 2. Spatial precision decrease versus increasing spot density. Each displayed curve differs from the default settings in the indicated parameter: photon emission rate (N_p), smoothing algorithm choice, multi-spot suspectedness thresholds (θ_{fishy}), multi-spot distance threshold (θ_{dist}), spot density (ρ_s) in spots per μm^2 and pixel sizes (ρ_p) in PSF full-widths at half-maximum. At all settings, a significant decrease in spatial precision is observed with higher spot densities, but the decrease is small in comparison to other sources of spatial uncertainty. The error bars indicate the standard deviation within 5 simulation runs differing only by random seed. Points with standard deviations greater than their mean were discarded.

The impact of stochastic precision and recall was investigated in two steps, first fixing optimal parameters for multi-spot analysis and then analyzing variations along the remaining pa-

rameters. The stochastic precision-recall diagram (Fig. 3), a parametric diagram which shows the trajectory in precision-recall-space caused by variation of the spot search threshold, was used to identify critical points in multi-spot search parameters. Readers unfamiliar with this kind of diagram should note that the points are not a function of stochastic precision, but rather of θ_{fishy} , and that both axes in the diagram show measured quantities. In general, points close to the upper right corner of a precision-recall diagram are considered optimal, and curves will tend to run from top left (good recall, i.e. many true positives, and bad precision, i.e. many false positives) to bottom right (few true positives and few false positives), with curves running from bottom left to top right indicating suboptimal parameters (e.g. for very low θ_{fishy} both stochastic precision and recall decrease, showing $\theta_{\text{fishy}} = 0$ to be an absolutely inferior choice compared to $\theta_{\text{fishy}} = 0.01$). The consistent shape of the curve demonstrates that the combined optimum of stochastic precision and recall consistently occurs at a residual quotient $\theta_{\text{fishy}} = 0.1$ and a high distance threshold θ_{dist} should be selected. Therefore, we fixed these settings to $\theta_{\text{fishy}} = 0.1$ and $\theta_{\text{dist}} = 0.5 \mu\text{m}$ in the following analysis. It should be noted that the overall low values of stochastic precision and recall stem primarily from the high default spot density of $0.64 \text{ spots } \mu\text{m}^{-2}$ (1.74 PSF FWHMs), which was chosen to cause many multi-spot events and thus test the effectivity of multi-spot search.

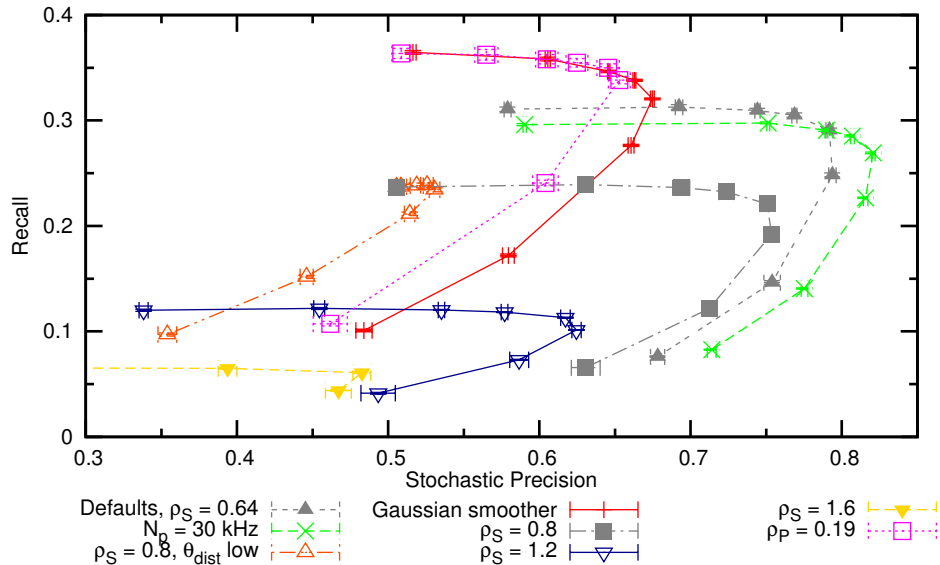


Fig. 3. Stochastic precision-recall-diagram. Each displayed curve differs from the default settings in the indicated parameters: photon emission rate (N_p), smoothing algorithm choice, multi-spot distance threshold (θ_{dist}), spot density (ρ_S) in spots per μm^2 and pixel sizes (ρ_P) in PSF full-widths at half-maximum. The plot is parametric with points along each curve varying in double spot search aggressiveness (θ_{fishy}) with 1 being at the upper left edge of each curve and 0.5, 0.3, 0.2, 0.1, 0.05, 0.01 and 0 following. While almost all curves differ from each other, indicating sensitivity of stochastic precision and recall to all parameters, most reach their optimum at or close to the fifth point, i.e. $\theta_{\text{fishy}} = 0.1$, indicating an optimal value for θ_{fishy} .

3.1. Stochastic precision and recall

Despite applying multi-spot analysis, we found that the *rapidSTORM* algorithm encounters problems with accurate spot identification when the number of simultaneously active fluorophores increases. Both recall and stochastic precision show a distinct decrease with increasing spot density ρ_S (Figs. 4a and b). In other words, the number of false positive localizations increases in the super-resolved reconstructed image with increasing ρ_S . An exponential decay in the recall emerges for all curves, even though the actual values differ by up to a factor of two. Since the number of true localizations is given by the camera area multiplied by the spot density and the recall rate, the density of true localizations scales with $\rho_S \exp(-\frac{\rho_S}{k})$ with k being algorithm-dependent (e.g. $0.6 \mu\text{m}^{-2}$ for default settings) and has a maximum at $\rho_S = k$ (Fig. 4c). We will refer to the density of true localizations per frame as throughput. Using default settings, the maximum occurs at $0.6 \mu\text{m}^{-2}$ (1.8 PSF FWHMs) and allows a throughput of 0.17 localizations per frame and μm^2 (Fig. 4c), corresponding to a mean nearest-neighbour distance of correctly identified localizations of 3.4 PSF FWHMs. Consistently, algorithmic alternative parameters that employ less smoothing in the preprocessing stage offer better throughput. Note, however, that the maximum throughput is offset by a considerable amount of false localizations, i.e. noise or artifacts impair the reconstructed image.

This tradeoff can be analyzed and visualized by using a stochastic precision-throughput diagram (Fig. 5), i.e. a diagram analogous to a stochastic precision-recall diagram with the ordinate scaled by the spot density. When plotting throughput versus the stochastic precision, the slope curves along which points differ by spot density, the slope in the stochastic precision-throughput diagram characterizes both the existence and the sharpness of the tradeoff between stochastic precision and throughput. Positive slopes in the diagram show that stochastic precision and throughput can be optimized concurrently when changing spot density and negative slopes close to 0 or $-\infty$ show that stochastic precision or throughput, respectively, can be gained at little cost.

3.2. Influence on different types of localization microscopy

The interdependence of throughput and stochastic precision is influenced by experimental constraints that can be divided into three categories: (i) the fluorophore-limited case defined by unfavorable photostability, including experiments with photoactivatable fluorescent proteins (e.g. PALM and FPALM, [3, 4]) and necessitating maximum recall rates, (ii) the ratio-limited case defined by the photoswitching rates or lifetimes of the on and off state (e.g. STORM and *d*STORM, [16]), and (iii) the throughput-limited case defined by short acquisition times caused by experimental instability or the need to acquire many localizations in a short time (e.g. dynamic super-resolution imaging in living cells, [8, 9]). This need is exemplified by the low spot densities ($0.01 - 0.03 \mu\text{m}^{-2}$) used recently by Frost et al. [35].

For the fluorophore-limited case, low spot densities should be adjusted to guarantee good recall because in most cases each fluorophore only produces few spots before photobleaching. The density-stochastic precision and density-recall diagram (Figs. 4a and b) demonstrate that both recall and stochastic precision approach 100% for very low spot densities highlighting the stability of the *rapidSTORM* algorithm against background noise [24].

The ratio-limited case occurs when reversible photoswitching allows repeated detection of a fluorophore, making recall less relevant. Here, stochastic precision is the relevant factor and mainly limited by the photoswitching rates, i.e. the ratio of the lifetime of the off and on state, $r = \frac{\tau_{\text{off}}}{\tau_{\text{on}}}$ [16, 17]. Thus, the ratio determines the acceptable fluorophore and spot density, and finally the achievable structural resolution [16]. Without multi-spot analysis the stochastic precision decreases exponentially with increasing spot density. On the other hand, applying multi-spot analysis the fraction of true positive localizations increases considerably enabling a sto-

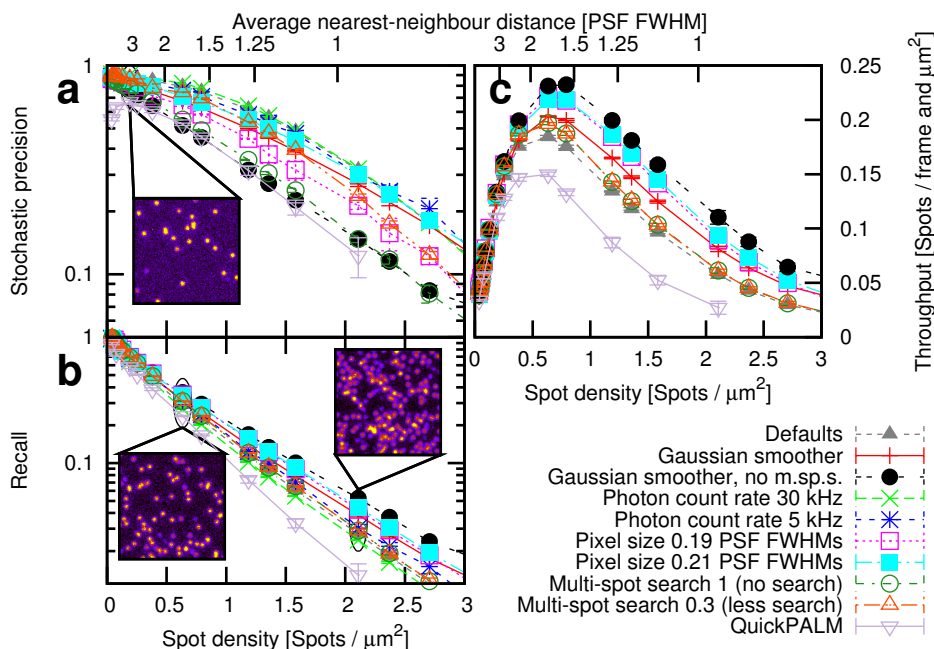


Fig. 4. Recall, stochastic precision and throughput as a function of spot density. The curves differ in photophysical and algorithmic properties. A photon count rate of 1 kHz corresponds to up to 100 photons per spot, and m.sp.s. abbreviates multi-spot search. (a,b) All recall curves and the stochastic precision curves for some settings, including those without multi-spot search, show exponential decay with slope varying between algorithms. The exponential behavior implies the existence of a maximum for the number of detected spots per time, located at ~ 0.7 spots per μm^2 . Double-spot search improves the stochastic precision by a factor of up to 2. (c) Precision-throughput-diagram with parametric curves. The points along each curve vary in spot density, with abscissa and ordinate showing the achieved stochastic precision and throughput. This diagram shows how different algorithmic approaches offer different trade-offs, with our default settings offering high stochastic precision and several low-precision, high-throughput alternatives that might be useful for time-resolved measurements.

chastic precision of $> 80\%$ for spot densities of $0.6 \mu\text{m}^{-2}$ or, in other words, more than 80% of all localizations composing the reconstructed image are accurate localizations and contain sample information (Fig. 4a). The exponential behavior is also shown by the QuickPALM algorithm, indicating a broad applicability of a simple exponential decay model for stochastic precision and recall.

In the throughput-limited case when the acquisition speed is the crucial parameter, the stochastic precision-throughput diagram (Fig. 5) is most useful. By plotting the product of spot density and recall against the stochastic precision several optimal zones can be identified. Already with default settings the *rapid*STORM algorithm allows a throughput of almost 0.2 localizations per frame and μm^2 at a stochastic precision of nearly 80% which can be further optimized by changing the pixel size and other algorithmic input parameters (Fig. 5).

These results show that effort is necessary to optimize the performance of the *rapid*STORM algorithm and probably other, similar algorithms for localization microscopy. While a full inquiry into the causes of failure is outside the scope of this article, several studies [9,22] indicate

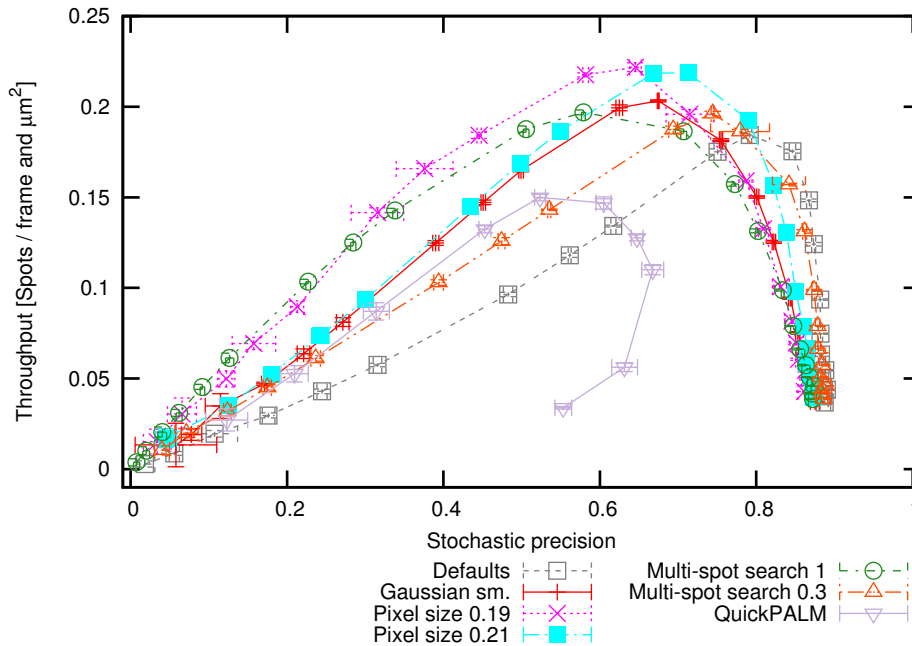


Fig. 5. Parametric diagram of stochastic precision and throughput with changing spot density. The points along each curve vary in spot density, with abscissa and ordinate showing the achieved stochastic precision and throughput given in spots per frame and μm^2 . For example, the curve showing default settings (*hollow squares*) starts at the lower right with high precision and low throughput, and shows loss of precision (that is, towards lower x-values) as well as rise and subsequent fall of throughput along the curve. By evaluating the slope of the curve, it can be determined how much precision is lost for gains in throughput. The diagram also shows how different algorithmic approaches offer different trade-offs, with our default settings offering high precision and several low-precision, high-throughput alternatives that might be useful for time-resolved measurements.

that higher performance under high spot density conditions should be possible. However, these studies did not provide a false positive rate. It has to be pointed out that the investigation of recall or localization precision alone is not adequate to judge the quality of algorithmic processing. Since the distribution of false positives is both different from that of true positives and can be expected to be highly dependent on the spatial configuration of simulated fluorophores, localization precision, and stochastic precision must be measured independently. At the same time, both recall and stochastic precision are necessary for a meaningful stochastic analysis. In other words, performance at high spot density can only be compared with previously published results (reporting e.g. high recall values also for higher spot densities [9, 22]) and new algorithms evaluated only if all three parameters stochastic precision, recall, and localization precision are considered, underlining the importance of the proposed measurement method.

3.3. Information throughput and necessary acquisition time

The maximal throughput can be used to predict the minimal acquisition time necessary to achieve a desired experimental Nyquist-Shannon-limited spatial image resolution. To resolve a structure with a structural resolution of 20 nm, i.e. reliably detect irregularities of at least 20 nm in size, one data point has to be recorded every 10 nm and therefore up to 10,000 fluoro-

phores μm^{-2} are necessary. At the maximum throughput we measured, at least 50,000 images have to be acquired to reconstruct a super-resolved image containing a sufficient number of true localizations. For example, at a frame rate of 100 Hz the acquisition time sums up to ~ 8 minutes and prevents dynamic super-resolution imaging of highly dynamic samples. However, this number should be interpreted with two caveats: firstly, less complex structures such as filaments or small multi-protein complexes require much lower labeling density and allow us to perform super-resolution imaging in much shorter acquisition times [36]. Secondly, we ignored the effects of localization redundancy, i.e. more than one localization per fluorophore, and localization precision, both of which necessitate even longer acquisition times for resolution.

The throughput limit we found also implies an information limit for localization microscopy applicable to all localization microscopy applications, surpassing structural resolution considerations. Since the number of true localizations determined per time unit is limited and each localization is determined with an inherent uncertainty given by the localization precision, localization microscopy must be limited by the Shannon-Hartley theorem [15], with the localization throughput representing the bandwidth of a classical communication channel and the acquisition area size and the localization precision determining the signal-to-noise ratio. Thereby, our results give an information throughput measured in bits per image for localization microscopy algorithms; if the results of the experiment can be easily treated in information theory terms, this result allows an estimation of the necessary acquisition time.

3.4. Applicability

We consider our simulation-derived results to be applicable quantitatively to real *d*STORM measurements since we adapted realistic noise measurements and photophysical parameters from *d*STORM measurements. The largest incongruence with reality is the distribution of localization amplitudes, i.e. the estimated number of photons per spot: while computations on real measurements show a broad distribution of localization amplitudes, the simulations generate many localizations with small variation around the photon rate times the integration time. This incongruence might be due to different excitation intensities of differently located fluorophores induced e.g. by total internal reflection. However, with the simulated spots already showing significant variation towards low amplitudes due to the chosen simulated integration time and the combination of many spots probably overshadowing the photon statistics of each single spot, we deem the variance in spot strength in our simulations sufficient and did not try to enhance the model to match a spot strength histogram of real data more closely.

While the above results have been obtained with simulation parameters typical for the *d*STORM method and the *rapid**d*STORM algorithm, the proposed method for measuring dense-spot performance is easily applicable to all current localization microscopy methods and algorithms. We provide evaluation software on our website and can provide the stochastically generated input image stacks on request. We expect our results to hold quantitatively for many current localization microscopy algorithms, which generally follow the the same pattern as *rapid**d*STORM of denoising, identifying spot positions by maximum search or thresholding, and non-linearly fitting functional approximations of the PSF to these positions. However, it should be stressed that photophysical parameters were assigned to match typical *d*STORM data and were measured on *rapid**d*STORM, and thus our numerical results can not be applied directly to other localization microscopy algorithms that operate with different premises, under different conditions (e.g. molecule brightness and background noise) or with different computational approaches. For these algorithms, we suggest using the demonstrated method of simulated fluorophore lattices to obtain their statistical properties, which is eased by the supplied software and, if photophysical parameters allow, our generated input image stacks. In general, our method should be considered as a practical proof and testing procedure for localization microscopy

algorithms, and does not indicate the limits of localization microscopy.

4. Conclusions and outlook

Our results demonstrate that we have a powerful method for automatic and reliable testing of localization microscopy algorithms under high spot density conditions. The method relies on computing localization precision, stochastic precision and recall from the parameters of a sum of Gaussian functions fitted to an average raster interval histogram. The usefulness of stochastic precision recall, spot density recall, spot density precision, and stochastic precision throughput diagrams has been demonstrated by identifying the best algorithmic parameters and by predicting ideal physical parameters for algorithmic performance. We determined optimal spot densities on the basis of the *rapid*STORM algorithm demonstrating that fluorophore-limited experiments (PALM, FPALM) should be performed well below 0.5 spots per μm^2 , ratio-limited experiments with reversibly photoswitchable fluorophores (STORM, *d*STORM) at 0.6 spots per μm^2 , and throughput-limited dynamic super-resolution imaging experiments are limited to 0.2 achieved localizations per frame and μm^2 . Our results highlight the complex interrelation of spot density, photophysical fluorophore parameters, and acquisition speed expressed as throughput (spots per frame and μm^2). They demonstrate that high labeling densities are prone to generate false and artificial localizations unless experimental parameters such as photoswitching and acquisition rates are set appropriately and that very long acquisition times are necessary when localization throughput is limited.

Our quantitative characterization of a localization microscopy algorithm is an important step towards a refined understanding of the resolution and quantification capability of single-molecule based localization microscopy methods. Furthermore, we expect the proposed lattice histogram method used to evaluate evaluating dense-spot performance to be very useful in designing and testing algorithms that extend the capabilities of standard localization microscopy methods.

Acknowledgement

This work was supported by the Biophotonics and the Systems Biology Initiative (FORSYS) of the German Ministry of Research and Education (BMBF). This publication was funded by the German Research Foundation (DFG) in the funding programme Open Access Publishing.