Single-Photon Single-Particle Tracking

Lance W.Q. Xu (徐伟青)[1,2], Nathan Ronceray[3], Marianna Fanouria Mitsioni[3,4], Aleksandra Radenovic[3], and Steve Pressé[1,2,5,*]

[1]Center for Biological Physics, Arizona State University, Tempe, AZ, USA
[2]Department of Physics, Arizona State University, Tempe, AZ, USA
[3]Institute of Bioengineering, School of Engineering, EPFL, Lausanne, Switzerland
[4]Institute of Bioengineering, School of Life Sciences, EPFL, Lausanne, Switzerland
[5]School of Molecular Sciences, Arizona State University, Tempe, AZ, USA
[*]Corresponding author (email: spresse@asu.edu)

Abstract

Multiple biological "particles", from biomolecules to viral capsids, diffuse with rates exceeding $1\,\mu\text{m}^2/\text{s}$, introducing blurring so significant when imaged with traditional integrative detectors such as sCMOSs and EMCCDs that one particle may appear as many. While single-photon detector arrays hold promise to strongly mitigate blurring artefacts by allowing for tenfold temporal resolution improvement, tracking from arrays with binary 1-bit output (photon detected or not) falls outside the existing tracking paradigm. For instance, the core of existing tracking algorithms relies on intensity-based information per pixel to achieve single-particle localization. Yet this is not possible with a 1-bit output. Here, we circumvent the localization paradigm and directly analyze single-photon detector array output to leverage the temporal resolution afforded by short single-photon detector array exposures. To achieve this, we propose a computationally feasible strategy to propose particle numbers and associated trajectories consistent with the entire 1-bit per pixel image stack. Our method, single-photon single-particle tracking (SP^2T), deals with both hot and cold pixels to which single-photon detector array data interpretation is highly sensitive. We consider multiple examples, including controlled experiments with fluorescent beads following pre-programmed trajectories, where ground truth is partially available, and the diffusion of aerolysin monomers on supported lipid bilayers. These results showcase SP^2T as a powerful tool capable of tenfold temporal resolution improvement for advancing the resolution of dynamical events from single-photon detector array architectures.

Introduction

Advances in fluorophore photophysics have allowed us to surpass the diffraction limit of light, enabling fluorescent localization of molecular positions with remarkable precision, down to approximately 70 nm [1–14]. More recent techniques have pushed this precision even further, with localization down to 6 nm achieved through excitation beam modulation [15], with subsequent *in vitro* techniques reaching sub-nanometer precision [16].With these advancements we can now directly observe a number of biomedical structures [17], nanoparticle endocytosis and targeted drug delivery to cancer sites [18, 19], as well as viral particle activity within living systems [20, 21].

Despite the impressive spatial resolution achieved, temporal resolution remains constrained to the range above a few milliseconds by the acquisition rates of commonly used detectors such as sCMOS and EMCCD cameras [22–24]. At these slower acquisition rates, blurring artefacts become manifest, even for relatively slow diffusing particles ($\approx 1\,\mu\mathrm{m}^2/\mathrm{s}$), see Fig. 1a.

Yet, whereas blurring artefacts are already manifest at $1\,\mu\mathrm{m}^2/\mathrm{s}$, a SARS-CoV-2 virion diffuses with a coefficient greater than $5\,\mu\mathrm{m}^2/\mathrm{s}$ [25] while that of a cytoplasmic proteins approach $10\,\mu\mathrm{m}^2/\mathrm{s}$ [26]. Here, motion over an exposure period introduces challenges in distinguishing between blurring artefacts from neighboring biomolecules. This is only further exacerbated by small biomolecules appearing as larger patterns ($\approx$250 nm) due to the diffraction-limited point spread function (PSF) [27]. While custom modifications to detectors can enhance frame rates of traditional integrative detectors [23], these improvements often come at the expense of a reduced FOV, limiting the ability to achieve both high temporal resolution and wide spatial coverage. Furthermore, shorter exposure times, while nominally providing data on faster timescales, reduce the signal-to-noise ratio and, therefore, complicate image interpretation and analysis.

Overcoming the trade-off between FOV and data acquisition rate requires using single-photon detectors (SPDs) offering high sensitivity at fast acquisition speeds. Most commercial SPDs are solid-state devices made from either semiconductor or superconducting materials, falling into four primary families: photomultiplier tubes (PMTs), single-photon avalanche diodes (SPADs), superconducting nanowire single-photon detectors (SNSPDs), and transition-edge sensors (TES) [28]. These SPDs have been used across various fields such as quantum computing [28, 29], long-range sensing [30], and various microscopy methods [31–36] yet they remain relatively under-utilized in biological applications.

Here, we focus on SPAD arrays, specifically the SPAD512^2 from Pi Imaging, which offers a large field of view (512×512 pixels) and a fast data acquisition rate of 97.7 kHz. These arrays operate reliably at room temperature or slightly below, and their compact design, low operating voltage, and high performance make them well-suited for a variety of applications [37]. SPAD arrays have already proven valuable in fields such as fluorescence microscopy [31–36], spectroscopy [38–40], and nuclear medicine [38], where they serve as powerful diagnostic tools.

Despite the fast data acquisition rate of SPAD arrays, no current method can directly monitor dynamical events from raw 1-bit (single-photon) outputs. The core challenge lies

in the need for superresolution while tracking, which relies on the accurate localization of emitting particles. Most localization methods require sufficient detected photons (and thus intensities per pixel) to achieve high localization accuracy [41]. This approach inherently conflicts with the need to localize without intensities from single-photon detection. Indeed, recent studies using SPDs, including SPAD arrays, still depend on binning 1-bit frames for analysis [42, 43], a process that diminishes the exceptional temporal resolution SPDs can offer, as illustrated in Fig. 1. While binning allows us to localize particles using traditional tracking methods, binning frames do not add information to the dataset; instead, they average out temporal information. This suggests that a method avoiding binning may leverage the information available from 1-bit frames.

Here, we introduce single-photon single-particle tracking (SP^2T), a framework that enables direct particle tracking from 1-bit frames. We fundamentally overcome the localization paradigm by proposing the number of particles and their associated trajectories consistent over the entire 1-bit-per-pixel image stack.

Unlike the analysis for integrative detectors, an accurate SPAD array algorithm requires careful calibration of each pixel. Hot pixels can mimic additional particles by consistently registering false-positive detections, while cold pixels may create areas that appear to repel particles, reducing detection probability in those regions. Thus, as we will see in SP^2T, we account for realistic complications in SPAD array detectors, including hot and cold pixels.

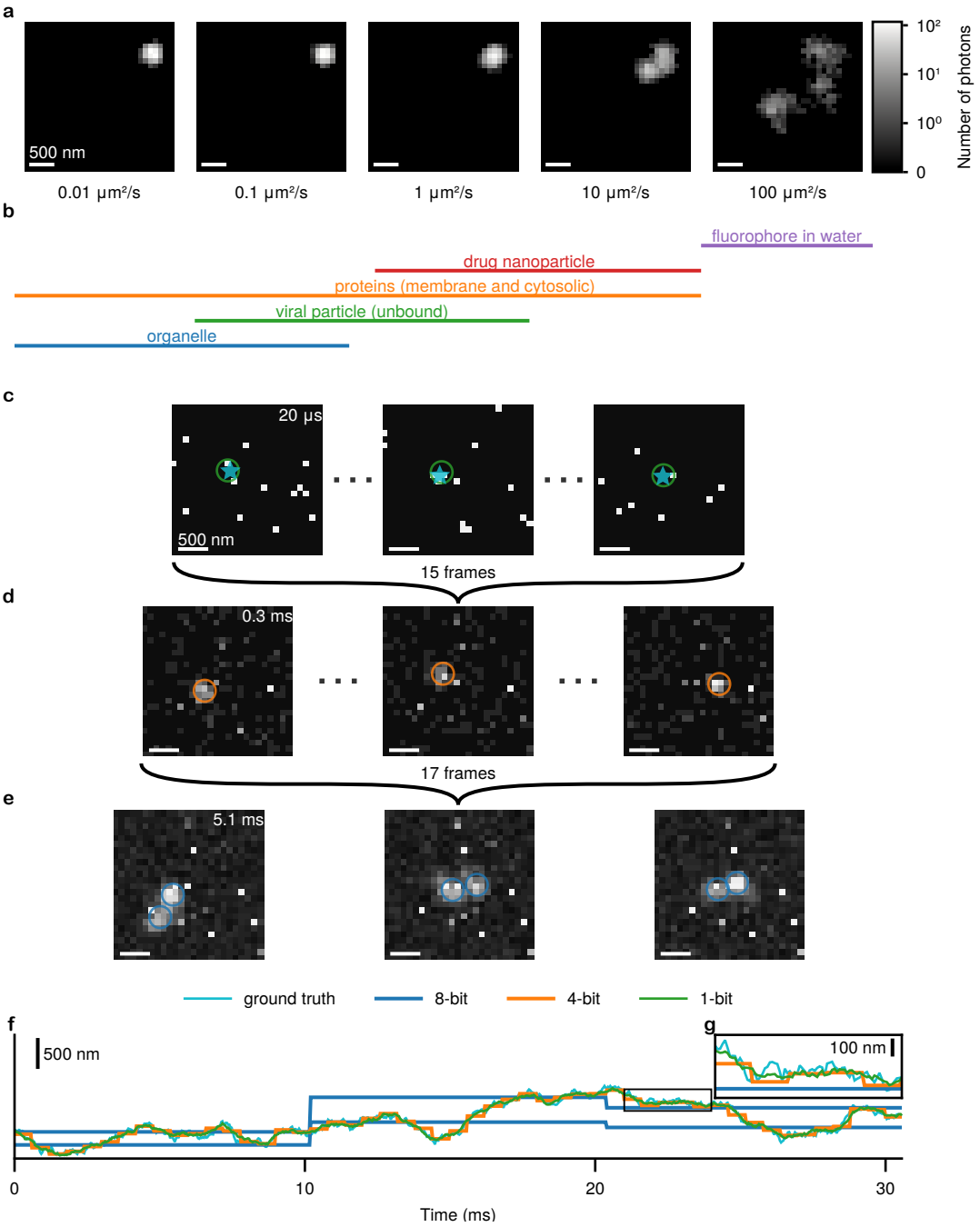

Figure 1: Advantages of monitoring molecular targets with 1-bit frames. **a**, Increasing diffusion coefficients from left to right. The frame contrast with a higher diffusion coefficient is enhanced for clearer visualization. The total number of photon detections per frame is constant. Interestingly, at higher diffusion coefficients, not only is the smearing of the diffraction-limited spot apparent, but a single molecule may even appear and be detected as many by standard tracking tools, as we will see shortly.

Figure 1: **b**, Typical diffusion coefficient ranges for different biological systems [44]. **c**, Example 1-bit frames with molecule positions (with ground truth shown as a star) extracted by our proposed framework here, indicated by green circles. **d**, 15 1-bit frames are summed to produce a 4-bit frame. Example 4-bit frames with target positions localized using TrackMate [45], indicated by orange circles. **e**, Seventeen 4-bit frames are summed to create an 8-bit frame. Example 8-bit frames, where blue circles show particle positions detected by TrackMate. Due to severe blurring artefacts, TrackMate incorrectly identifies two instead of one target. For visualization purposes, the contrast in frames of lower bit-depth is enhanced in **c**–**e**. **f**, Comparison of tracks obtained from frames with different bit-depths: 1-bit tracks are extracted using the proposed framework, while 4-bit and 8-bit tracks are obtained using TrackMate. **g**, Zoom in onto the boxed region in **f**.

Results

To validate that SP^{2}T can reliably track particles using 1-bit frames, we begin with controlled experiments involving fluorescent beads fixed to a piezoelectric (piezo) stage programmed to follow a known trajectory. This setup directly compares the programmed trajectory and SP^{2}T's tracking results. We then apply SP^{2}T to a biologically relevant dataset, tracking aerolysin molecules diffusing on a supported lipid bilayer. Unlike the fluorescent beads, aerolysin molecules are labeled with Cy3B, which reduces their brightness by approximately half, presenting additional challenges. Finally, we assess the robustness of SP^{2}T using synthetic datasets across various parameter regimes, including emitter separation, diffusion coefficients, emission brightness, and dark count rates, to ensure its adaptability to diverse experimental conditions.

As SP^{2}T is based on a probabilistic framework within the Bayesian paradigm, it predicts full distributions over particle numbers and associated tracks obtained by drawing samples from the posterior probability distribution, not just point estimates over particle numbers and their most probable tracks. However, to plot our results, we typically plot point estimates obtained as the *maximum a posteriori* (MAP) sample representing the particle numbers and their associated tracks, maximizing the posterior. We generate multiple posterior samples to quantify uncertainty, typically around 500 for each track.

Fluorescent beads moving with a piezo stage

The first control experiments we present involve fluorescent beads fixed to a piezo stage following a pre-programmed trajectory. This trajectory is designed by holding the stage stationary for a constant dwell time (40 ms), then moving by a predetermined step size (200 nm) in a randomly chosen direction: left, right, up, or down.

This setup enables us to 1) directly compare our tracking results with the known programmed trajectories; 2) assess the consistency of the tracks across different beads, confirming the robustness of our framework; and 3) benchmark our results against tracks obtained

using existing tracking tools and binned frames, demonstrating that SP^2T achieves comparable spatial resolution while delivering temporal resolution at least an order of magnitude faster.

As illustrated in Fig. 2a, the dataset spanning approximately 115 ms clearly shows three shifts of around 200 nm in the particle positions. These shifts occur at approximately 15 ms, 55 ms, and 95 ms, closely aligning with the expected dwell time of 40 ms. Additionally, the directions of these shifts are consistent with the experimental settings: the first shift is to the left, the second is upward, and the third shift is downward. In Fig. 2b, we align the three tracks for spatially separated particles for direct comparison. Remarkably, they are superresolved—agreeing within a range of approximately 50 nm, well below the diffraction limit of light (around 250 nm) despite being tracked from 1-bit per pixel data. We calculated the Euclidean distances between the shifted tracks to quantify the agreement between the particle tracks. The maximum distance between any two beads in any frame is less than 65 nm, and the average distance between any two beads over all frames is about 21 nm.

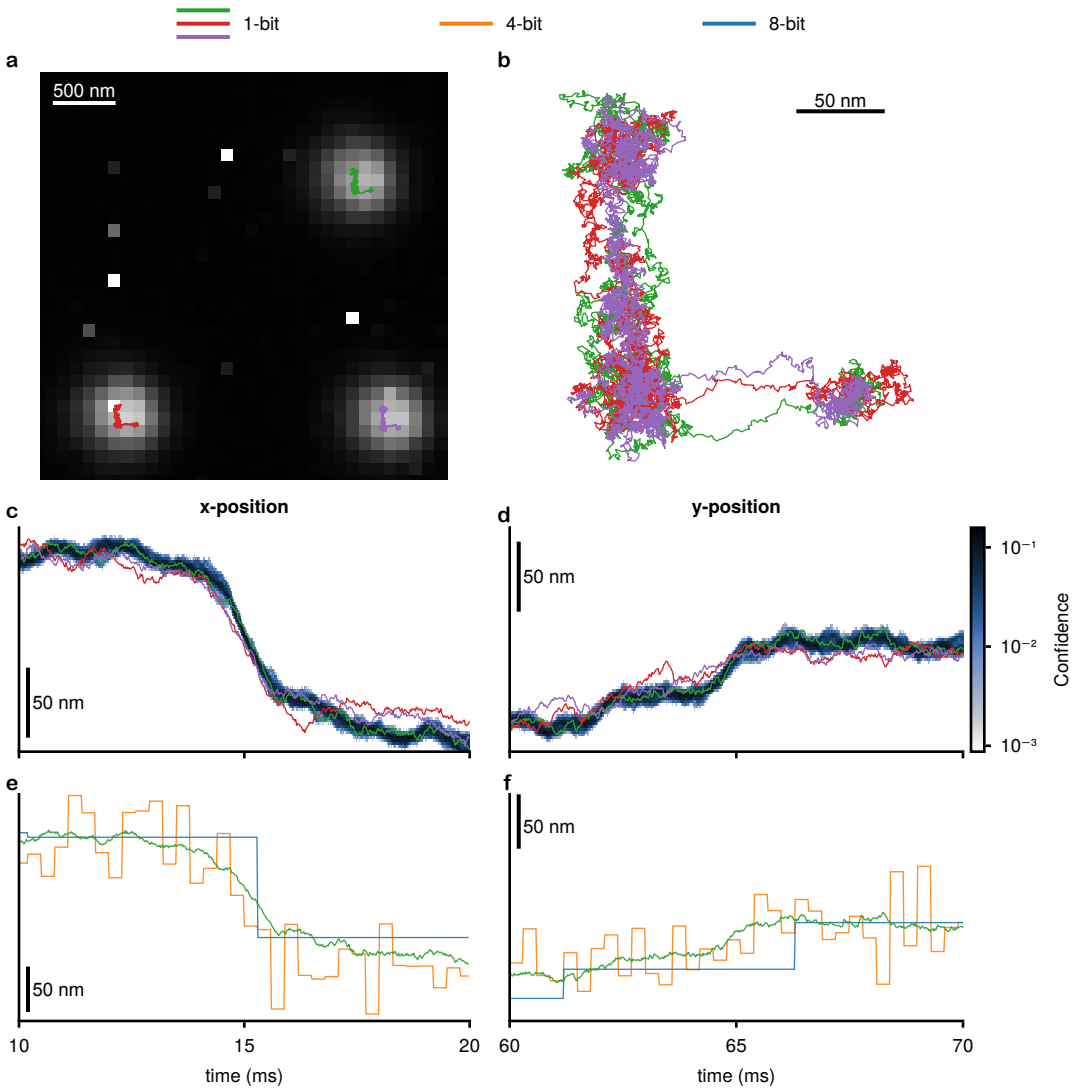

Figure 2: Results of fluorescent beads moving with a piezo stage with 40 ms dwell time and 200 nm step size. **a**, An example frame generated by summing all 5,750 45×45 1-bit frames. The frame is cropped and contrast-adjusted for illustrative purposes. **b**, The bead tracks inferred by our framework are overlaid on top of each other to assess consistency. **c**, Confidence of bead one's x-track plotted as a shaded region in a time interval (10 ms to 20 ms); darker colors indicate higher confidence. **d**, The same layout as **c** but for bead one's y-track **e**, Direct comparison of bead one's 1-bit track with its 4-bit and 8-bit tracks during the same time interval as in **c**. **f**, Direct comparison of bead one's 1-bit track with its 4-bit and 8-bit tracks during the same time interval as in **d**.

In addition, since the beads remain stationary during the dwell time, the spread of

their inferred tracks during these periods provides a metric for quantifying tracking errors. Figures 2c and 2d show that the beads remain fixed at the three vertices of the "L"-shaped trajectories, and the track spreads at these vertices are consistently below 50 nm. This result demonstrates that our proposed framework achieves superresolved tracking directly from 1-bit frames. We highlight that no method can currently achieve what is shown in this figure.

Furthermore, since the fluorescent beads are known to remain stationary during the dwell time, it is feasible to sum the binary frames over time (binning) to perform conventional tracking. For this purpose, we binned 125 binary frames and used TrackMate [45] to analyze the resulting frames. The tracks generated by TrackMate are also plotted in Figs. 2e and 2f, showing a high degree of alignment with the SP^{2}T results. This strong correlation indicates that SP^{2}T achieves a similar spatial resolution while offering a temporal resolution that is two orders of magnitude faster.

As for traditional (intensity-based) tracking tools, while selecting a smaller bin size for frames can improve the temporal resolution for TrackMate, this comes at the cost of reduced spatial resolution. As the bin size decreases, the limited information in each frame results in poorer localization accuracy, according to the Cramér–Rao lower bound [41, 46].

Next, we compare the SP^{2}T results with each other. As illustrated in Figs. 2a and 2b, even without assuming any correlation between the particle tracks, all three SP^{2}T tracks align with each other within a range of 50 nm after applying a constant translational shift. This precision is well below the diffraction limit of approximately 230 nm in this dataset, demonstrating the consistency and accuracy of SP^{2}T in tracking particles.

Aerolysin diffusing on an SLB

In the following example, we shift to a more biologically relevant scenario: aerolysin diffusing on a supported lipid bilayer. Unlike fluorescent beads used in the previous dataset, aerolysin are labeled with Cy3B, reducing their brightness twofold; see Fig. 3a.

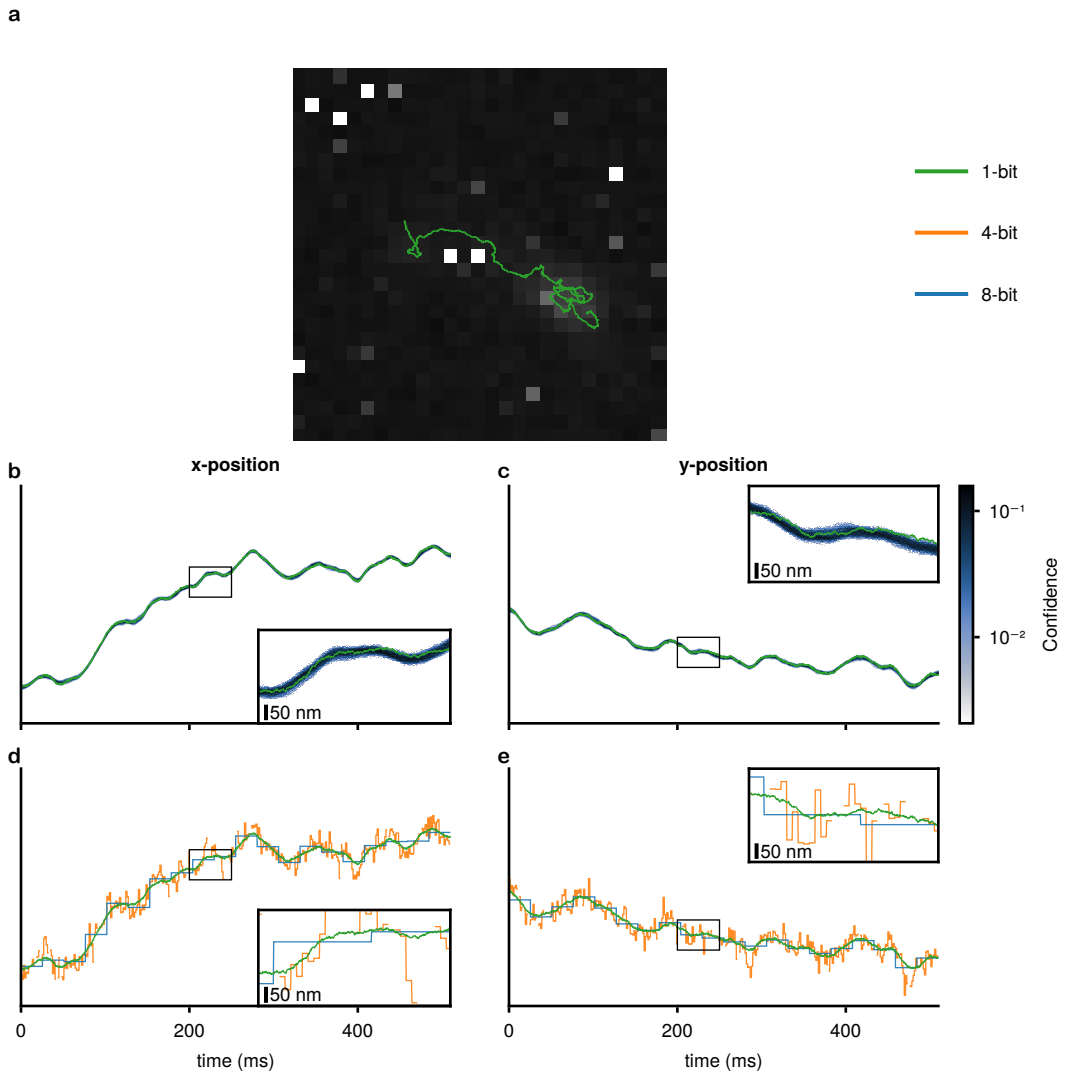

Figure 3: Results of aerolysin labeled with Cy3B diffusing on a supported lipid bilayer. **a**, An example frame generated by summing all 5,100 40×40 1-bit frames. The frame is contrast-adjusted for illustrative purposes. The aerolysin track inferred by our framework is overlaid on top. **b**, Confidence of this aerolysin's x-track plotted as a shaded region; darker colors indicate higher confidence. **c**, For the sake of illustration, we zoom in to the region enclosed by the box. **d**, As we do not have ground truth for a real experiment, as an approximate control, we directly compare aerolysin's 1-bit track with its 4- and 8-bit tracks obtained using TrackMate [45] **e**, For illustration, we zoom in to the region enclosed by the box. A comparison in the y-direction is largely the same and not shown.

This reduced brightness presents a robustness test to our framework. Despite this, as

depicted in Figs. 3b and 3c, our framework achieves superresolution, *i.e.*, a tracking error of approximately 50 nm using 1-bit frames, consistent with the results obtained from the bead dataset. In Figs. 3d and 3e, while we have no ground truth for the 1-bit track, we present a direct comparison between our tracks and those generated by TrackMate [45]. In the zoomed-in region, it is evident that the 4-bit tracks exhibit missing segments, attributed to the lower brightness of the aerolysin labels. This further highlights the limitations of the existing intensity-based paradigm when dealing with low-intensity signals and underscores the robustness of our approach.

Robustness tests with synthetic data

In this section, we focus on using synthetic data to evaluate SP^{2}T's robustness in different parameter regimes (separation between emitters in Fig. 4, diffusion coefficient in Fig. S.1, emitter brightness in Fig. S.2, and dark counts in Fig. S.3).

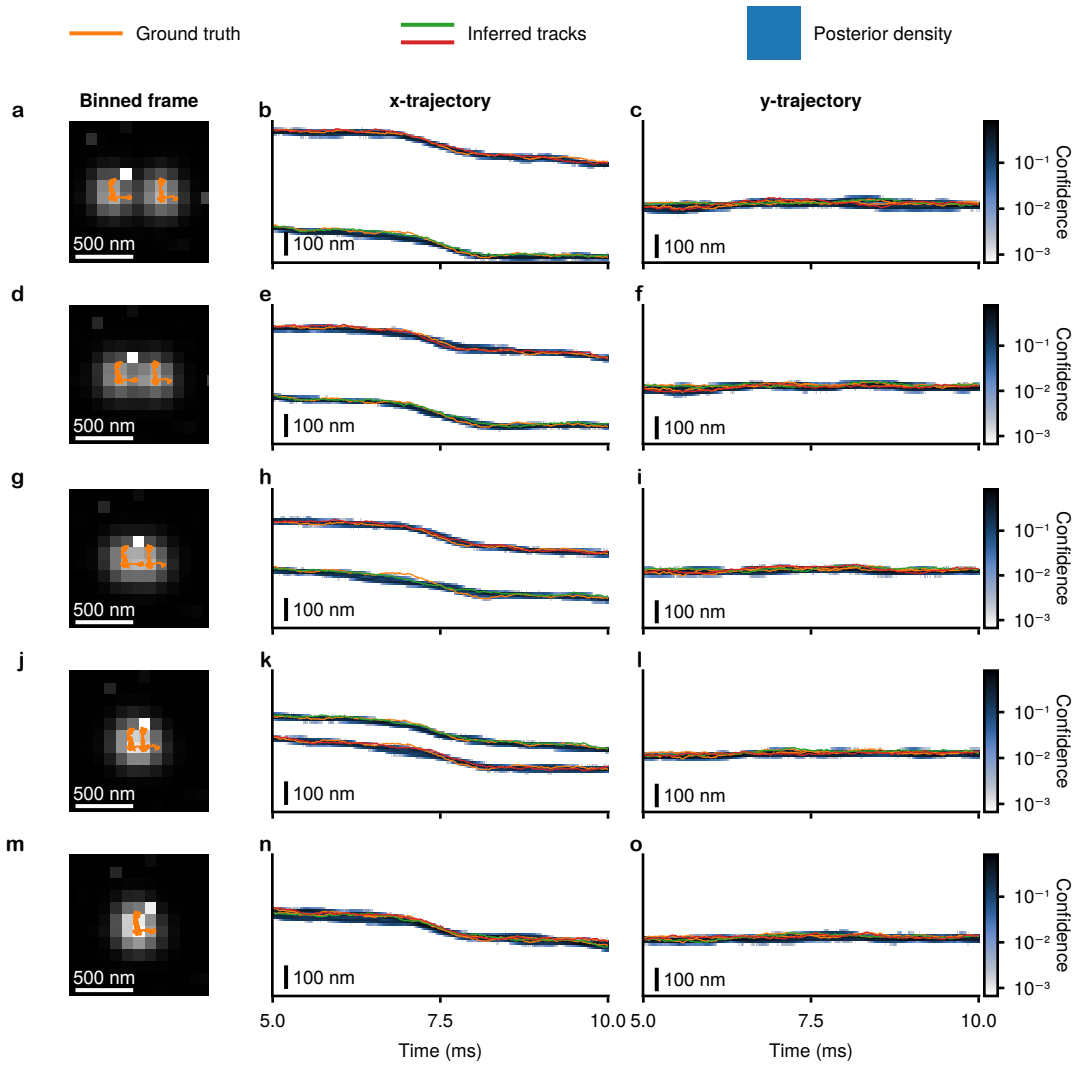

Figure 4: Performance of SP^{2}T under variable separation between emitter trajectories. **a** Binned frames and the ground truth trajectories, the spatial separation here is about 400 nm. **b**–**c** Zoomed-in tracking results of SP^{2}T for x and y trajectories, corresponding to Supplementary Data 3 . The ground truth tracks are orange, while SP^{2}T's MAP tracks are green and red. Shading indicates confidence (the probability of the emitter being in a 10 nm bin), with darker colors representing higher confidence. **d**–**f** Same layout as **a**–**c** but for Supplementary Data 4 with a separation of 300 nm. **g**–**i** Same layout as **a**–**c** but for Supplementary Data 5 with a separation of 200 nm. **j**–**l** Same layout as **a**–**c** but for Supplementary Data 6 with a separation of 100 nm. **m**–**o** Same layout as **a**–**c** but for Supplementary Data 7 with a separation of 0 nm.

As illustrated in Fig. 4a, we used two estimated bead trajectories from Fig. 2as ground truth, placing them side-by-side with a separation of approximately 400 nm. We then simulated 1-bit frames using real dark count calibrations (from Fig. 2) and the model described in the Methods. In Figs. 4b and 4c, we present SP^{2}T's results in the same format as Figs. 2c, 2d, 3b and 3c. Consistent with previous findings, SP^{2}T found the correct number of emitters and the inferred tracks align with the ground truth trajectories within a range of around 50 nm.

In the lower panels of Fig. 4, we gradually decreased the spatial separation between the two tracks in increments of 100 nm until the tracks largely overlapped in most frames. As shown in Figs. 4d to 4o, SP^{2}T maintained its accuracy even as the emitters overlapped almost entirely. This robustness is due to SP^{2}T's use of brightness calibration, allowing it to reliably infer the number of emitters even in challenging scenarios with close proximity.

Discussion

In this study, we have demonstrated the robust performance of SP^{2}T, a tracking method designed to work directly with binary frames from single-photon detection. SP^{2}T consistently identified the correct number of particles and produced accurate tracks across various datasets, including both controlled experiments with fluorescent beads on a piezoelectric stage and more complex biological scenarios involving Cy3B-labeled aerolysin monomers diffusing on a supported lipid bilayer.

We have also compared SP^{2}T to the widely used SPT tool, TrackMate [45], across all datasets in this study. While TrackMate and other tracking tools [47] rely on long-exposure frames to achieve high spatial resolution, they do so at the cost of reduced temporal resolution. In contrast, SP^{2}T provides tracks with comparable spatial resolution while achieving over ten times the temporal resolution, eliminating blurring artefacts and enabling us to analyze both slowly diffusing organelles [44] and faster-diffusing drug nanoparticles, as illustrated in Fig. 1b. This improvement holds promise in capturing faster dynamic processes, including vesicle transport and drug nanoparticle delivery, with a level of previously inaccessible temporal resolution.

As discussed, SPDs hold tremendous promise for the future of high-resolution imaging, with various architectures suited to diverse applications. Although SPDs are based on different physical principles and tailored to specific research scenarios, analyzing single-photon datasets requires a fundamental departure from traditional localization paradigms—something our framework accomplishes by tracking directly from single-photon data.

Our work has primarily focused on pixel-based SPDs, such as SPAD arrays, but the framework can be adapted to non-array-based SPDs like LINCam [48] and NCam [49–51]. These detectors provide an exciting opportunity to enhance spatial resolution even further by eliminating the need for spatial binning, just as we avoid temporal binning in our current approach. This adaptability might enable truly bin-free imaging in both spatial and temporal dimensions, optimizing resolution across all domains.

Additionally, without binning constraints, we could potentially distinguish between mul-

tiple molecular species by leveraging fluorescence lifetimes. Detectors like the LINCam [48] and TCSPC-capable SPAD arrays [32], coupled with pulsed laser sources, already capture lifetime information, enabling discrimination between different fluorophores based on their unique decay profiles [52]. This capability opens the door to achieving higher resolution and multi-species tracking in complex environments, making SPDs a powerful platform for future imaging applications.

While SP^2T has demonstrated significant advantages in tracking performance, there are areas of improvement. Specifically, because SP^2T works directly with binary frames, the number of frames it analyzes is substantially higher—up to 255 times more—than that of a typical 8-bit grayscale image. This dramatic increase in the number of frames raises potential challenges, particularly regarding computational feasibility. Performing an analysis that jointly considers all binary frames could become computationally intensive and impractical for large datasets, potentially limiting the method's application in scenarios where computational resources are constrained. We have implemented two primary strategies to address the computational challenges posed by the large number of binary frames. First, as detailed in the Methods section, the most computationally intensive task in SP^2T—likelihood calculation—is parallelized in a two-layer structure and executed on a GPU. This parallelization significantly accelerates processing, achieving up to a 50-fold speedup on a GTX 1060 GPU compared to a conventional CPU setup like the i7-7700k.

The second approach to reducing SP^2T's computational demands is to offer users the flexibility to operate in either parametric or nonparametric mode. In parametric mode, users specify the number of particles, allowing SP^2T to bypass the computationally intensive process of particle count estimation. Users may initially run in nonparametric mode, then switch to parametric mode for a five-fold performance boost. This adaptability enables users to tailor the analysis to their needs, balancing computational efficiency with the desired analytical depth for their application. Additionally, we recommend using tracks obtained from long-exposure frames as input to SP^2T, which can yield a further two-fold speed increase.

With these computational optimizations, SP^2T can process over 5,000 50×50 frames in just a couple of hours using an NVIDIA GTX 1060. Additionally, its computational cost remains stable, regardless of the number of frames or particles, until the GPU cores reach saturation.

Thanks to their versatility, SPDs can substitute conventional integrative detectors in a range of applications, including fluorescence lifetime microscopy [34, 35], image-scanning microscopy [32, 36], and Förster resonance energy transfer [33]. Similarly, as a mathematical framework designed to work directly with single-photon data, SP^2T can be adapted to various experimental setups and purposes. Here are a few examples: 1) SP^2T can be extended for 3D or specialized PSFs by adjusting the integrand in equation (4), enabling direct tracking in three dimensions; 2) equation (5) could be replaced by a Poisson distribution if a detector can count multiple photons arriving at a pixel within a frame period; and 3) for SPDs that capture single-photon arrival times, we could couple equation (5) with an arrival time distribution, allowing SP^2T to perform tracking and species discrimination based on fluorescence lifetimes simultaneously.

References

1. Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313,** 1642 (2006).
2. Rust, M. J., Bates, M. & Zhuang, X. Sub-Diffraction-Limit Imaging by Stochastic Optical Reconstruction Microscopy (STORM). *Nat. Methods* **3,** 793 (2006).
3. Hess, S., Girirajan, T. & Mason, M. Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. *Biophys. J.* **91,** 4258 (2006).
4. Lee, A., Tsekouras, K., Calderon, C., Bustamante, C. & Pressé, S. Unraveling the Thousand Word Picture: An Introduction to Super-Resolution Data Analysis. *Chem. Rev.* **117,** 7276 (2017).
5. Tavakoli, M., Taylor, J. N., Li, C. B., Komatsuzaki, T. & Pressé, S. English (US). in *Advances in Chemical Physics* Publisher Copyright: © 2017 John Wiley & Sons, Inc. All rights reserved., 205 (Wiley, Mar. 2017). ISBN: 9781119324577.
6. Kervrann, C., Sánchez-Sorzano, C., Acton, S., Olivo-Marin, J. & Unser, M. A Guided Tour of Selected Image Processing and Analysis Methods for Fluorescence and Electron Microscopy. *IEEE J. Sel. Top. Signal Process.* **10,** 6 (2016).
7. Smal, I., Loog, M., Niessen, W. & Meijering, E. Quantitative Comparison of Spot Detection Methods in Fluorescence Microscopy. *IEEE Trans. Med. Imaging* **29,** 282 (2010).
8. Sage, D. *et al.* Quantitative Evaluation of Software Packages for Single-Molecule Localization Microscopy. *Nat. Methods* **12,** 717 (2015).
9. Von Diezmann, A., Shechtman, Y. & Moerner, W. Three-Dimensional Localization of Single Molecules for Super-Resolution Imaging and Single-Particle Tracking. *Chem. Rev.* **117,** 7244 (2017).
10. Beghin, A. *et al.* Localization-Based Super-Resolution Imaging Meets High-Content Screening. *Nat. Methods* **14,** 1184 (2017).
11. Patterson, G., Davidson, M., Manley, S. & Lippincott-Schwartz, J. Superresolution Imaging Using Single-Molecule Localization. *Ann. Rev. Phys. Chem.* **61,** 345 (2010).
12. Egner, A. *et al.* Fluorescence Nanoscopy in Whole Cells by Asynchronous Localization of Photoswitching Emitters. *Biophys. J.* **93,** 3285 (2007).
13. Thompson, R., Larson, D. & Webb, W. Precise Nanometer Localization Analysis for Individual Fluorescent Probes. *Biophys. J.* **82,** 2775 (2002).
14. Holden, S., Uphoff, S. & Kapanidis, A. DAOSTORM: An Algorithm for High-Density Super-Resolution Microscopy. *Nat. Methods* **8,** 279 (2011).
15. Balzarotti, F. *et al.* Nanometer Resolution Imaging and Tracking of Fluorescent Molecules with Minimal Photon Fluxes. *Science* **355,** 606 (2017).

16. Reinhardt, S. C. *et al.* Ångström-Resolution Fluorescence Microscopy. *Nature* **617,** 711 (2023).

17. Padmanabhan, P. *et al.* Single-Molecule Imaging Reveals Tau Trapping at Nanometer-Sized Dynamic Hot Spots Near the Plasma Membrane That Persists after Microtubule Perturbation and Cholesterol Depletion. en. *EMBO J.* **41,** e111265 (Oct. 2022).

18. Yao, Y. *et al.* Nanoparticle-Based Drug Delivery in Cancer Therapy and Its Role in Overcoming Drug Resistance. en. *Front. Mol. Biosci.* **7,** 193 (Aug. 2020).

19. Patel, S. *et al.* Brief Update on Endocytosis of Nanomedicines. *Adv. Drug Deliv. Rev.* **144,** 90 (2019).

20. Zhang, Q. *et al.* Entry Dynamics of Single Ebola Virus Revealed by Force Tracing. *ACS Nano* **14,** 7046 (2020).

21. Arista-Romero, M., Pujals, S. & Albertazzi, L. Towards a Quantitative Single Particle Characterization by Super Resolution Microscopy: From Virus Structures to Antivirals Design. en. *Front. Bioeng. Biotechnol.* **9,** 647874 (Mar. 2021).

22. Do Ó, C. R. *et al. GPI 2.0: Performance Evaluation of the Wavefront Sensor's EMCCD* 2023.

23. Mandracchia, B. *et al.* High-Speed Optical Imaging with sCMOS Pixel Reassignment. *Nat. Commun.* **15,** 4598 (2024).

24. Fazel, M. *et al.* Fluorescence Microscopy: A Statistics-Optics Perspective. *Rev. Mod. Phys.* **96,** 025003 (2 2024).

25. Adamczyk, Z., Batys, P. & Barbasz, J. SARS-CoV-2 Virion Physicochemical Characteristics Pertinent to Abiotic Substrate Attachment. *Curr. Opin. Colloid Interface Sci.* **55,** 101466 (2021).

26. Schavemaker, P. E., Boersma, A. J. & Poolman, B. How Important is Protein Diffusion in Prokaryotes? en. *Front. Mol. Biosci.* **5,** 93 (Nov. 2018).

27. Sgouralis, I. *et al.* BNP-Track: a framework for superresolved tracking. *Nat. Methods* **21,** 1716. ISSN: 1548-7105 (Sept. 2024).

28. Couteau, C. *et al.* Applications of single photons to quantum communication and computing. *Nature Reviews Physics* **5,** 326. ISSN: 2522-5820. `https://doi.org/10.1038/s42254-023-00583-2` (June 2023).

29. You, L. *Nanophotonics* **9,** 2673. `https://doi.org/10.1515/nanoph-2020-0186` (2020).

30. Hadfield, R. H. *et al.* Single-photon detection for long-range imaging and sensing. *Optica* **10,** 1124. `https://opg.optica.org/optica/abstract.cfm?URI=optica-10-9-1124` (Sept. 2023).

31. Bruschini, C., Homulle, H., Antolovic, I. M., Burri, S. & Charbon, E. Single-photon avalanche diode imagers in biophotonics: review and outlook. *Light: Science & Applications* **8,** 87 (2019).

32. Buttafava, M. *et al.* SPAD-based asynchronous-readout array detectors for image-scanning microscopy. *Optica* **7,** 755. https://opg.optica.org/optica/abstract.cfm?URI=optica-7-7-755 (July 2020).

33. Smith, J. T. *et al.* In vitro and in vivo NIR fluorescence lifetime imaging with a time-gated SPAD camera. *Optica* **9,** 532. https://opg.optica.org/optica/abstract.cfm?URI=optica-9-5-532 (May 2022).

34. Smith, J. T. *et al.* In vitro and in vivo NIR Fluorescence Lifetime Imaging with a Time-Gated SPAD Camera. *Optica* **9,** 532 (2022).

35. Ulku, A. C. *et al.* A 512 × 512 SPAD Image Sensor with Integrated Gating for Widefield FLIM. *IEEE J. Sel. Top. Quantum Electron.* **25,** 1 (2019).

36. Castello, M. *et al.* A robust and versatile platform for image scanning microscopy enabling super-resolution FLIM. *Nature Methods* **16,** 175. ISSN: 1548-7105. https://doi.org/10.1038/s41592-018-0291-9 (Feb. 2019).

37. Ceccarelli, F. *et al.* Recent Advances and Future Perspectives of Single-Photon Avalanche Diodes for Quantum Photonics Applications. *Advanced Quantum Technologies* **4,** 2000102. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qute.202000102. https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.202000102 (2021).

38. Cusini, I. *et al.* Historical Perspectives, State of Art and Research Trends of SPAD Arrays and Their Applications (Part II: SPAD Arrays). *Frontiers in Physics* **10.** ISSN: 2296-424X. https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2022.906671 (2022).

39. Slenders, E. *et al.* Confocal-based fluorescence fluctuation spectroscopy with a SPAD array detector. *Light: Science & Applications* **10,** 31. ISSN: 2047-7538. https://doi.org/10.1038/s41377-021-00475-z (Feb. 2021).

40. Madonini, F. & Villa, F. Single Photon Avalanche Diode Arrays for Time-Resolved Raman Spectroscopy. *Sensors* **21.** ISSN: 1424-8220. https://www.mdpi.com/1424-8220/21/13/4287 (2021).

41. Lelek, M. *et al.* Single-molecule localization microscopy. *Nature Reviews Methods Primers* **1,** 39. ISSN: 2662-8449. https://doi.org/10.1038/s43586-021-00038-x (June 2021).

42. Bucci, A. *et al.* 4D Single-Particle Tracking with Asynchronous Read-Out Single-Photon Avalanche Diode Array Detector. *Nat. Commun.* **15,** 6188 (2024).

43. Gyongy, I. *et al. High-Speed Particle Tracking in Microscopy Using SPAD Image Sensors* in *High-Speed Biomedical Imaging and Spectroscopy III: Toward Big Data Instrumentation and Management* (eds Tsia, K. K. & Goda, K.) **10505** (SPIE, 2018), 105050A.

44. Phillips, R., Kondev, J., Theriot, J. & Garcia, H. *Physical Biology of the Cell* (Garland Science, 2012).

45. Tinevez, J.-Y. *et al.* TrackMate: An Open and Extensible Platform for Single-Particle Tracking. *Methods* **115,** 80. ISSN: 1046-2023 (2017).

46. Kay, S. M. *Fundamentals of Statistical Signal Processing* (Prentice-Hall, 1993).

47. Chenouard, N. *et al.* Objective Comparison of Particle Tracking Methods. *Nat. Methods* **11,** 281 (2014).

48. Prokazov, Y., Turbin, E., Weber, A., Hartig, R. & Zuschratter, W. Position Sensitive Detector for Fluorescence Lifetime Imaging. *J. Instrum.* **9,** C12015 (2014).

49. Stonehill, L. C. *et al. Cross-Strip Anodes for High-Rate Single-Photon Imaging* in *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)* (2009), 1417.

50. Cook, J. M. *et al. A Photon-Counting Camera System Developed from a Crossed-Strip Detector* in *Biosensing and Nanomedicine V* **8460** (SPIE, 2012), 170.

51. Siegmund, O. H. W. *et al. Cross Strip Microchannel Plate Imaging Photon Counters with High Time Resolution* in *Advanced Photon Counting Techniques IV* **7681** (SPIE, 2010), 54.

52. Fazel, M. *et al.* Building Fluorescence Lifetime Maps Photon-by-Photon by Leveraging Spatial Correlations. *ACS Photonics* **10,** 3558 (2023).

53. Goudie, R. J. B., Turner, R. M., De Angelis, D. & Thomas, A. MultiBUGS: A Parallel Implementation of the BUGS Modeling Framework for Faster Bayesian Inference. *J. Stat. Softw.* **95,** 1 (2020).

54. Luengo, D., Martino, L., Bugallo, M., Elvira, V. & Särkkä, S. A Survey of Monte Carlo Methods for Parameter Estimation. *EURASIP J. Adv. Signal Process.* **2020,** 25 (2020).

55. Wang, X., Jin, Y., Schmitt, S. & Olhofer, M. Recent Advances in Bayesian Optimization. *ACM Comput. Surv.* **55,** 1 (2023).

56. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: The MCMC Hammer. *Publ. Astron. Soc. Pac.* **125,** 306 (2013).

57. Raffenetti, K. *et al. Why is MPI So Slow? Analyzing the Fundamental Limits in Implementing MPI-3.1* in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (Association for Computing Machinery, Denver, Colorado, 2017). ISBN: 9781450351140.

58. Ansel, J. *et al. PyTorch 2: Faster Machine Learning through Dynamic Python Bytecode Transformation and Graph Compilation* in *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2* (Association for Computing Machinery, La Jolla, CA, USA, 2024), 929. ISBN: 9798400703850.

59. Martín Abadi *et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems* Software available from tensorflow.org. 2015.

60. Innes, M. Flux: Elegant Machine Learning with Julia. *Journal of Open Source Software* **3,** 602 (2018).

61. Innes, M. *et al. Fashionable Modelling with Flux* 2018.

62. Ronceray, N. *et al.* Liquid-Activated Quantum Emission from Pristine Hexagonal Boron Nitride for Nanofluidic Sensing. *Nat. Mater.* **22,** 1236 (2023).

63. Cirauqui, N., Abriata, L. A., Van Der Goot, F. G. & Dal Peraro, M. Structural, Physicochemical and Dynamic Features Conserved within the Aerolysin Pore-Forming Toxin Family. *Sci. Rep.* **7,** 13932 (2017).

64. Mayer, S. F. *et al.* Lumen charge governs ion transport in β-barrel nanopores. *bioRxiv* (2024).

65. Menestrina, G. Ionic Channels Formed by Staphylococcus aureus alpha-toxin: Voltage-Dependent Inhibition by Divalent and Trivalent Cations. *J. Membr. Biol.* **90,** 177 (1986).

66. Schindler, H. & Rosenbusch, J. P. Matrix Protein from Escherichia coli Outer Membranes Forms Voltage-Controlled Channels in Lipid Bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **75,** 3751 (1978).

67. Nestorovich, E. M. & Bezrukov, S. M. Beta-Barrel Channel Response to High Electric Fields: Functional Gating or Reversible Denaturation? *Int. J. Mol. Sci.* **24,** 16655 (2023).

68. Ronceray, N. *et al.* Multiplexed Lifetime Imaging of Single Molecules with a Gated Single-Photon Camera. *bioRxiv* (2024).

Methods

Theoretical framework

This section outlines the mathematical framework of SP^{2}T. For notational simplicity, we first define the following terms: N represents the number of frames, P denotes the number of pixels per frame, w_n^p refers to the data collected at pixel p on frame n, M indicates the number of particles, $\mathbf{x}_n^m$ represents the m-th particle's position at frame n, and $\mathcal{D}$ is the transition parameter which represents mean squared displacement, as defined in equation (9). With these definitions, we can express all collected data as $w_{1:N}^{1:P} = \{w_n^p | n = 1, \ldots, N, p = 1, \ldots, P\}$. Similarly, the collection of all particle tracks is denoted as $\mathbf{x}_{1:N}^{1:M}$.

As discussed in the sections above, the goal of SP^{2}T is to determine all variables of interest (including the number of emitting particles, particle tracks, and the transition parameter) along with the corresponding credible intervals. We achieve this by constructing the full probability distribution, known as the posterior, $\mathbb{P}\left(M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D} \middle| w_{1:N}^{1:P}\right)$, representing the probability distribution of all variables of interest given the collected data. Following Bayes' theorem, this posterior can be expressed as

$$\mathbb{P}\left(M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D} \middle| w_{1:N}^{1:P}\right) = \frac{\mathbb{P}\left(w_{1:N}^{1:P} \middle| M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D}\right) \mathbb{P}\left(M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D}\right)}{\mathbb{P}\left(w_{1:N}^{1:P}\right)}. \tag{1}$$

The first term in the numerator, $\mathbb{P}\left(w_{1:N}^{1:P} \middle| M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D}\right)$, represents the likelihood, which is the probability distribution of the data given the variables of interest. This term usually

describes the system's physics (the model for photon detection in this context) and realistic complications such as hot pixels. The second term, $\mathbb{P}\left(M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D}\right)$, known as the prior, imposes additional constraints on the variables of interest, such as a particle's motion model and ensuring the transition parameter $\mathcal{D}$ remains non-negative. The denominator, $\mathbb{P}\left(w_{1:N}^{1:P}\right)$, is referred to as the evidence and is generally treated as a normalization constant, as it does not depend on any variable of interest.

In the following sections, we will establish the base likelihood using only the photon detection model. Next, we will generalize this base likelihood to account for the aforementioned realistic complications and then discuss the prior we place on the variables of interest. Finally, we will briefly introduce our approach to constructing and sampling the posterior.

Base likelihood

As introduced above, the likelihood $\mathbb{P}\left(w_{1:N}^{1:P} \middle| M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D}\right)$ represents the probability distribution of the data given the number of particles, the particle tracks, and the transition parameter. To express this likelihood mathematically, we first observe that once all particle tracks are known, neither the number of particles nor the transition parameter affects photon detection. This observation allows us to simplify $\mathbb{P}\left(w_{1:N}^{1:P} \middle| M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D}\right)$ to $\mathbb{P}\left(w_{1:N}^{1:P} \middle| \mathbf{x}_{1:N}^{1:M}\right)$.

Furthermore, considering that we are dealing with fluorescent emitters whose lifetimes (on the order of nanoseconds) are significantly shorter than the frame period (at least $10\,\mu$s), the probability that an emitting particle excited during the exposure period of one frame will result in a photon detection in the next frame is negligible (less than 10^{-200}, assuming a 10 ns lifetime and excitation occurring in the middle of an exposure). Therefore, it is reasonable to assume that particle positions at frame n do not influence other frames. This assumption allows us to further simplify the likelihood as a product:

$$\mathbb{P}\left(w_{1:N}^{1:P} \middle| \mathbf{x}_{1:N}^{1:M}\right) = \prod_{n=1}^{N} \mathbb{P}\left(w_n^{1:P} \middle| \mathbf{x}_n^{1:M}\right). \tag{2}$$

Similarly, since the pixels of a detector usually operate independently, we can express the per-frame likelihood $\mathbb{P}\left(w_n^{1:P} \middle| \mathbf{x}_n^{1:M}\right)$ as a product over the individual pixel likelihoods:

$$\mathbb{P}\left(w_n^{1:P} \middle| \mathbf{x}_n^{1:M}\right) = \prod_{p=1}^{P} \mathbb{P}\left(w_n^{p} \middle| \mathbf{x}_n^{1:M}\right). \tag{3}$$

The per-pixel likelihood $\mathbb{P}\left(w_n^{p} \middle| \mathbf{x}_n^{1:M}\right)$ is influenced by four key factors: the point spread function (PSF) of the optical system, the pixel area A, the brightness of the particles h, and the number of particles M. The PSF characterizes how light from a point source is distributed across the detector [24]. At the same time, brightness represents the expected number of photons arriving at a pixel perfectly aligned with the PSF of a particle located at the focal plane. The total number of photons received by pixel p from all emitting particles, denoted as u_n^p, can be expressed as:

$$u_n^p = h \sum_{m=1}^{M} \iint_{A^p} \mathrm{PSF}\left(x, y; \mathbf{x}_n^m\right) \mathrm{d}x\, \mathrm{d}y. \tag{4}$$

Since a single-photon detector can only distinguish between the absence of photons (false) and the presence of at least one photon (true), the measurement w_n^p must be sampled from a Bernoulli distribution. The parameter of this Bernoulli distribution is determined by the probability of detecting at least one photon, which is given by $1 - \text{Poisson}(0; u_n^p)$. Here, $\text{Poisson}(0; u_n^p) = \exp(-u_n^p)$ represents the probability of detecting zero photons. Consequently, the probability of detecting at least one photon is $1 - \exp(-u_n^p)$, and the corresponding mathematical expression is:

$$w_n^p \sim \text{Bernoulli}\left(1 - \exp\left(-u_n^p\right)\right). \tag{5}$$

Thus, the overall likelihood can be obtained by combining equations (2) to (5).

Dark counts

Single-photon detectors are known to exhibit dark counts [32], which are false detections of photons that occur even in the absence of light. These dark counts, specific to each pixel, arise due to thermal noise and other intrinsic factors within the detector. As a result, the likelihood model must be adjusted to account for these dark counts, which contribute to an additional noise source and affect the accuracy of photon detection. To do so, we introduce a quantity that represents the expected number of dark counts per frame exposure, DC^p, and we modify equation (4) as follows

$$u_n^p = \text{DC}^p + h \sum_{m=1}^{M} \iint_{A^p} \text{PSF}\left(x, y; \mathbf{x}_n^m\right) \mathrm{d}x\, \mathrm{d}y. \tag{6}$$

Other pixel defects

Similar to dark counts, pixels of a single-photon detector can also carry other defects. For example, some pixels may be imperfectly connected to the circuit, therefore, may not produce any signal during an experiment. As these defects can originate from numerous possibly unclear causes, we allow users to set a boolean label, F^p, which is zero when the corresponding pixel has a defect and one otherwise, to determine whether a pixel should be excluded from the likelihood calculation. Specifically, we change equation (3) to

$$\mathbb{P}\left(w_n^{1:P} \middle| \mathbf{x}_n^{1:M}\right) = \prod_{p=1}^{P} \left[\mathbb{P}\left(w_n^p \middle| \mathbf{x}_n^{1:M}\right)\right]^{F^p}. \tag{7}$$

Full likelihood

To summarize the sections above, we present the mathematical form of the full likelihood here:

$$\mathbb{P}\left(w_{1:N}^{1:P} \middle| \mathbf{x}_{1:N}^{1:M}\right) = \prod_{n=1}^{N} \prod_{p=1}^{P} \left[\text{Bernoulli}\left(w_n^p; 1 - \exp\left(-u_n^p\right)\right)\right]^{F^p}, \tag{8}$$

where the expected photon count u_n^p is given by equation (6).

Motion model prior

Now that we have established the full likelihood $\mathbb{P}\left(w_{1:N}^{1:P} \middle| \mathbf{x}_{1:N}^{1:M}\right)$, we briefly introduce the motion model prior so that the full posterior can be constructed. As mentioned earlier, we characterize the motion of a particle using transition parameter $\mathcal{D}$. Specifically, we choose

$$\mathbf{x}_n^m \sim \text{Normal}_3\left(\mathbf{x}_{n-1}^m, \mathcal{D}\right). \tag{9}$$

Posterior

According to equation (1), we can now write the full posterior as

$$\begin{aligned}
&\mathbb{P}\left(M, \mathbf{x}_{1:N}^{1:M}, \mathcal{D} \middle| w_{1:N}^{1:P}\right) \\
&\propto \left[\prod_{n=1}^{N}\prod_{p=1}^{P}\left[\text{Bernoulli}\left(w_n^p; 1-\exp\left(-u_n^p\right)\right)\right]^{F^p}\right]\left[\prod_{n=2}^{N}\text{Normal}_3\left(\mathbf{x}_n^m; \mathbf{x}_{n-1}^m, 2\mathcal{D}\right)\right] \\
&\times \mathbb{P}\left(\mathbf{x}_1^{1:M}\right)\mathbb{P}\left(M\right)\mathbb{P}\left(\mathcal{D}\right).
\end{aligned} \tag{10}$$

Here, the last three priors $\mathbb{P}\left(\mathbf{x}_1^{1:M}\right)$, $\mathbb{P}\left(M\right)$, and $\mathbb{P}\left(\mathcal{D}\right)$ are chosen for computational convenience. As they primarily define the support of each variable and do not otherwise integrate critical physics, we introduce them in the Supplementary Information.

Parallelization

Here, we anticipate being able to parallelize 1) the likelihood calculation at each frame and 2) within each frame. Conceptually, this parallelization scheme stems from the notion that the full likelihood can be expressed as the product of per-pixel likelihoods across all pixels and frames. As an important technical point, these per-pixel likelihoods are independent when conditioned on θ as long as SPD pixels are uncorrelated. This independence allows us to parallelize calculations efficiently. Inspired by parallelization schemes across Bayesian applications [53–56], we specifically tailor our framework to SPDs to achieve the best possible computational efficiency.

Significant communication between cores is required as all per-pixel likelihoods must be combined to compute the full likelihood across all pixels. This frequent communication poses a challenge for distributed computing frameworks, where high communication overhead can become a bottleneck [57]. Unlike parallelization schemes assuming infrequent communication [53–56], our proposed framework leverages shared-memory architectures, specifically desktop GPUs for ease of use better suited to manage continuous communication demands efficiently. Moreover, our parallelization scheme is built on top of well-optimized numerical routines, such as batched matrix multiplication, already widely adopted in the DL community [58–61]. These routines automatically leverage lower-level hardware-specific numerical libraries, ensuring our framework achieves optimal performance across hardware platforms.

Experimental methods

SPAD array

Experiments were performed on a custom-built widefield single-molecule fluorescence microscope equipped with a time-gated large 512×512 SPD array, SPAD512^2 camera from Pi-Imaging with each pixel capable of detecting single photons. This commercial version of the SwissSPAD^2 camera was developed by the AQUA group (of Prof. Charbon) at EPFL. The camera captures binary images at an exceptional frame rate of up to 97,700 frames per second. A supercontinuum pulsed laser source was used for illumination, which could output an arbitrary wavelength in the visible range. We used either epi-illumination or total internal reflection illumination, allowing monitoring thousands of individual molecules simultaneously within a $\approx 50 \times 50\,\mu\text{m}^2$ FOV. The detection path can be swapped for instrument comparison with a conventional EMCCD camera.

Beads and piezo stage

Widefield imaging was performed using a custom fluorescence microscope, as described in Ref. [62]. Briefly, we imaged fluorescent beads while following a programmed trajectory set by the motion of the sample placed on a piezoelectric stage (Mad City Labs).

Aerolysin

Full-length wild-type aerolysin protein in the pET22b vector with a C-terminal hexa-histidine-tag was used. Cysteines were introduced into specific pore regions by mutating amino acids like G68, enabling labeling and tracking conformational changes during pore formation [63]. The mutants were generated by Genscript using the wild-type plasmid as a template, as was previously done for the production of 26 other aerolysin mutants [64]. By labeling specific aerolysin sites, we could leverage these mutants in combination with our imaging and tracking techniques to address fundamental questions regarding the oligomerization process or the voltage-induced closure of pores termed gating [65–67]. Aerolysin production, purification, labeling protocols, and the preparation of supported lipid bilayers where pore activity could be observed had already been established in the Radenovic lab [68].