

Bayesian counting of photobleaching steps with physical priors

Cite as: J. Chem. Phys. 152, 024110 (2020); doi: 10.1063/1.5132957

Submitted: 21 October 2019 • Accepted: 13 December 2019 •

Published Online: 9 January 2020



View Online



Export Citation



CrossMark

Jon Garry,¹ Yuchong Li,² Brandon Shew,²  Claudiu C. Gradinaru,^{2,a)}  and Andrew D. Rutenberg^{1,b)} 

AFFILIATIONS

¹Department of Physics & Atmospheric Science, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada

²Department of Chemical and Physical Sciences, University of Toronto Mississauga, Mississauga, Ontario L5L 1C6, Canada

^{a)}Electronic mail: claudiu.gradinaru@utoronto.ca

^{b)}Electronic mail: andrew.rutenberg@dal.ca

ABSTRACT

Counting fluorescence photobleaching steps is commonly used to infer the number n_0 of monomeric units of individual oligomeric protein complexes or misfolded protein aggregates. We present a principled Bayesian approach for counting that incorporates the statistics of photobleaching. Our physics-based prior leads to a simple and efficient numerical scheme for maximum *a posteriori* probability (MAP) estimates of the initial fluorophore number \hat{n}_0 . Our focus here is on using a calibration to precisely estimate \hat{n}_0 , though our approach can also be used to calibrate the photophysics. Imaging noise increases with \hat{n}_0 , while bias is often introduced by temporal averaging. We examine the effects of fluorophore number \hat{n}_0 of the oligomer or aggregate, lifetime photon yield μ_{eff} of an individual fluorophore, and exposure time Δt of each image frame in a time-lapse experiment. We find that, in comparison with standard ratiometric approaches or with a “change-point” step-counting algorithm, our MAP approach is both more precise and less biased.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5132957>

I. INTRODUCTION

Photobleaching of fluorophores can be exploited to determine the number of macromolecular subunits within immobilized and isolated complexes or aggregates.^{1–5} After covalently binding one fluorophore to each subunit, counting the steps of fluorescence bleaching for an individual macromolecule then determines its number n_0 of subunits. When all bleach steps are resolved, the number of subunits in, e.g., channels,⁶ membrane receptors,^{7,8} and even soluble proteins⁹ can be determined (see also Ref. 4). Total internal reflection fluorescence (TIRF) microscopy is normally used for data acquisition to minimize the background imaging noise.

Typically, to make a precise estimation of the number of steps n_0 , first the brightness of a single fluorophore v is calibrated, then the calibration is applied to the initial background-subtracted intensity in a ratiometric approach. There are many approaches to calibrating fluorophore brightness, including direct measurement of individual fluorophores,^{10,11} power-spectrum analysis of observed photobleach steps,^{12,13} fluctuation approaches,^{14,15} and Bayesian analysis.¹⁶ In contrast, there has been little systematic consideration

of how to best estimate n_0 , given a precise calibration of fluorophore brightness v .

As n_0 increases, the noise in the measured intensity from a given macromolecule with n_0 fluorescent labels also increases, since the variance increases linearly with n_0 .^{14,17–19} At the same time, the duration of the first bleach step of the macromolecule decreases with n_0 , since if τ is the average time to bleach a single fluorophore then with n_0 fluorophores, the average time until the first step will be τ/n_0 . Both of these effects, increasing noise and initially shorter intervals between bleach steps, work against precise estimation of larger n_0 even with a calibrated v . Furthermore, they are interdependent. Increasing the illumination intensity increases the signal-to-noise ratio (SNR), but shortens the initial bleaching time. The same compromise is also seen when longer imaging intervals Δt are used.

A simple estimate of the total photon budget highlights the problem. Absolute photostability, as characterized by the average number of photons μ emitted by a molecule before photobleaching, is 10^5 – 10^6 for high-performance fluorophores commonly used in single-molecule studies.^{20–22} A typical microscope detection efficiency of 1%–10% then gives a range of the photon

yield, i.e., the number of detected photons per fluorophore, of $\mu_{\text{eff}} \simeq 10^3\text{--}10^5$.

If each imaging frame collects a fraction α of the average lifetime photons of a single unbleached fluorophore, then in one frame the average number of detected photons is $\tilde{\nu} = \alpha\mu_{\text{eff}}$, where the \sim denotes photon units. The initial intensity of n_0 fluorophores is $\tilde{I}_0 \simeq \tilde{\nu}n_0$, while Poisson statistics of photon shot-noise indicates that the standard deviation is $\sigma_{\tilde{I}_0} = \sqrt{\tilde{I}_0}$. This then gives $\sigma_{n_0} = \sigma_{\tilde{I}_0}/\tilde{\nu} = \sqrt{n_0/\tilde{\nu}}$. If we insist on precise counting, we must have $\sigma_{n_0} \lesssim 1$, i.e., $n_0 \lesssim \alpha\mu_{\text{eff}}$, arguing for larger α . However, to resolve the first bleach step, the duration of each imaging frame should be (much) shorter than the first bleach step, i.e., $\alpha\tau \lesssim \tau/n_0$ or $\alpha \lesssim 1/n_0$, arguing for smaller α . Combining these two requirements limits us to $n_0 \lesssim \sqrt{\mu_{\text{eff}}}$, and indicates that more photostable fluorophores will allow larger n_0 to be precisely counted. For a given μ_{eff} , how should we best compromise α between resolving bleach-steps and decreasing noise?

For a given α , we may also ask how best to use the intensity signal and the calibration to determine n_0 . Directly taking the ratio of the background-subtracted intensity to the step-size $(I_0 - a)/\nu$, where a is the background intensity, is a naive but straightforward ratiometric approach. However, this ignores any information after the first image frame and so is unnecessarily noisy.

Some sort of filtering or time-average can reduce the noise in I_0 . Commonly used approaches use step-preserving filters such as the Chung-Kennedy (CK) filter²³ to reduce noise while preserving the step-structure in the signal. Practically, any filtering or local averaging also requires tunable parameters such as the width of the filtering window to achieve reliable results (see, e.g., Refs. 7 and 12). Such tunable parameters must be optimized for every experiment, and so limit throughput. Any local averaging or filtering will also blur out closely separated steps.²⁴ This results in a systematic underestimate (bias) of the total steps with large numbers of steps.²⁵

Since filtering loses information, and so does ignoring data after the first image frame, we could instead try to use all of the available timeseries photobleach data to determine n_0 .²⁶ Since the data are structured, with strong temporal correlations within and between steps, Bayesian inference of n_0 is appropriate.²⁷ With a physically based model and parameters, the likelihood function of the Bayes' theorem is combined with previous knowledge about parameter values or inter-relationships through prior distributions to produce a posterior distribution.²⁸ Recently, Bayesian approaches have been used in fitting photobleach curves, either to account for incomplete labeling²⁹ or to calibrate the size of photobleach steps.¹⁶ Bayesian approaches have also been used to infer state transitions in systems with small numbers of states.³⁰

Ideally, the Bayesian model and priors capture all of the physics that produced the data. In that case, with sufficient data, inferred parameters will converge toward the real physical parameters.²⁷ In practice, simpler *ad hoc* priors are often used to approximate and simplify the posterior.²⁹ For example, Tsekouras *et al.* use an *ad hoc* prior to limit biases of the number of inferred photobleaching steps.¹⁶ Without their prior, they systematically over-estimated the number of bleach steps. What does a principled physical prior look like and how would it behave, both in terms of computational convenience and in terms of bias?

In an alternative approach, directly counting photobleaching steps can be done without any calibration and with no physical priors. For sufficiently small noise, direct counting can be done visually.¹ More generally, algorithms can be used to automatically identify and count steps in the data. These “change-point” (CP) algorithms assume only piecewise constant data (see, e.g., Refs. 31–34) and have been applied to count RNA, DNA, and protein molecules in biological complexes.^{4,8,35} However, accurate counting beyond 10 photobleach steps remains a challenge using CP methods.³⁵

In this paper, we present a Bayesian approach using a principled prior that captures the physics of photobleaching. With our principled prior, we find that maximum *a posteriori* probability (MAP) estimation³⁶ is both mathematically transparent and computationally tractable. We can use our approach to either quantify the initial number of fluorophores (\hat{n}_0) within photobleach traces with known photophysics or calibrate the photophysics with a small number of fluorophores. To illustrate calibration, we calibrate our model parameters with experimental data. We then test our approach by fitting ensembles of simulated photobleach data. We investigate the effects of \hat{n}_0 , camera exposure time, and the photon yield of the fluorophore. We compare our MAP approach with both naive and filtered ratiometric approaches, as well as with a change-point algorithm that does not require calibration.

II. METHODS

A. Intensity model

For a constant camera exposure time $\Delta t = \alpha\tau$, the intensity measured in each frame $i \in \{0, \dots, N-1\}$ is composed of the number of unbleached fluorophores n_i , the brightness factor (step-size) per fluorophore ν , the background illumination intensity a , and a Gaussian noise ξ_i (with zero-mean, $\langle \xi_i \rangle = 0$),

$$I_i = \nu n_i + a + \xi_i. \quad (1)$$

We have approximated the Poisson shot-noise of individual photons¹⁷ with a Gaussian, which is necessary for our subsequent analysis and is reasonable when the number of photons per interval $\tilde{\nu} \gtrsim 20$. The variance in the noise, σ_i^2 , is dependent upon the number of fluorophores present at a given time-step with a shot-noise component σ_1^2 per fluorophore and an additional term accounting for the background fluorescence noise and the detector's dark noise σ_0^2 . Since variances are additive, the total variance is^{17,19}

$$\sigma_i^2 = n_i\sigma_1^2 + \sigma_0^2. \quad (2)$$

Our noise model captures background and dark noise (through σ_0^2) and shot-noise due to fluorophores (through σ_1^2), but does not include pixel-by-pixel camera details (see, e.g., Ref. 37). Random fluctuations of individual fluorophore brightness due to environmental or internal degrees of freedom that are rapid with respect to τ are also captured by σ_1^2 . We have used arbitrary intensity units, rather than photon counts, in Eqs. (1) and (2), so that σ_1 and ν are independent parameters.

B. Simulating bleach traces

For a given initial number of fluorophores \hat{n}_0 and fluorophore lifetime τ , Gillespie's stochastic simulation algorithm (SSA)³⁸ was used to generate bleach times for each simulated trace.

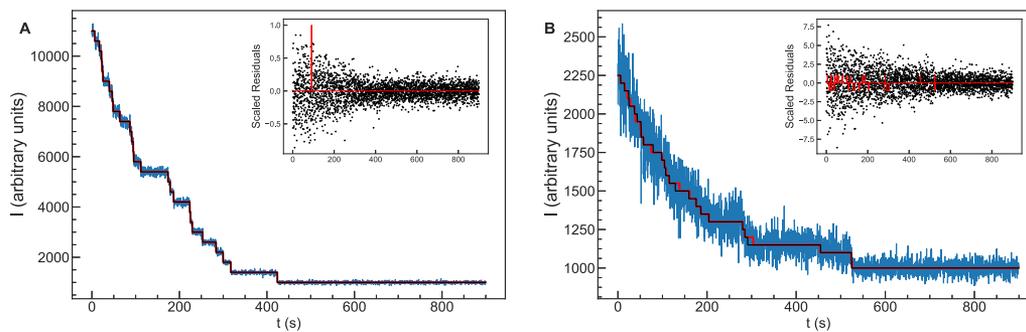


FIG. 1. Illustrative simulated bleach curves (blue lines) each starting with $\hat{n}_0 = 25$ fluorophores. Shown is intensity I vs time t . In each subplot, the thicker red line shows the noise-free steps, the thinner black line our Bayesian MAP fit, and the inset shows the residuals (with respect to the intensity in black points, and with respect to the noise-free intensity in red lines) in units of the fluorophore brightness ν . Each curve has $N = 3000$ data points with $\Delta t = 0.3$ s timesteps. Default parameters are used: background signal $a = 1000$, per fluorophore noise variance $\sigma_1^2 = 900$, background noise variance $\sigma_0^2 = 1600$, and nonbleach probability per time-step $q = 0.998$. (a) Low-noise trace with $\nu = 400$. Bleach steps in the raw intensity are readily identifiable by eye at later times, but not at the earliest times. (b) High-noise trace with $\nu = 50$. Bleach steps in the raw intensity cannot be identified by eye with any certainty.

Using these bleach times, the true set of fluorophore counts $\{\hat{n}_i\} \equiv \{\hat{n}_0, \hat{n}_1, \dots, \hat{n}_{N-1}\}$ was determined. Note that we use $\{\hat{n}_i\}$ to indicate the set of ground-truth values that will be fit with $\{n_i\}$.

Then, for each simulated trace, the set of measured intensities $\{I_i\}$ was constructed from our model [Eq. (1)] with random Gaussian noise according to Eq. (2). Since bleach times did not coincide with exposure intervals, we used time-average fluorophore counts within each exposure and the corresponding variances. Illustrative noise-free steps (red lines) and stochastic traces (blue lines) are shown in Fig. 1.

C. Bayesian model parameters

Bayesian inference uses Bayes' theorem,

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}, \quad (3)$$

to fit probabilistic models to data.²⁷ The posterior distribution $P(\theta|D)$ embodies how well the set of model parameters $\{\theta\}$ describes the measured data $\{D\}$. The likelihood $P(D|\theta)$ —typically used in traditional maximum likelihood estimation—is a measure of how well the data fit the model parameter values. [We ignore the “evidence,” $P(D)$, since it is a constant normalization factor that does not affect our results.] The priors over the model parameters, $P(\theta)$, are commonly used to capture previous measurements or beliefs, but can also capture physical relationships expected between model parameters.

The likelihood follows directly from our intensity model [Eqs. (1) and (2)] because of the independence of the Gaussian noise between the frames. We have simply that

$$P(D|\theta) = \prod_{i=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-(I_i - \nu n_i - a)^2 / (2\sigma_i^2)}, \quad (4)$$

where the product is over the N frames of the photobleach trace, and $\theta = \{\{n_i\}, \nu, a, q, \sigma_1^2, \sigma_0^2\}$ is the set of parameters for that trace. The inferred sets of numbers of fluorophores present in all of the time

steps $\{n_i\} \equiv \{n_0, n_1, \dots, n_{N-1}\}$ are also model parameters—they are inferred from the data $D = \{I_i\}$. However, while the parameters $\{\nu, a, \sigma_0^2, \sigma_1^2\}$, which we call the “photophysics,” are assumed to be the same between the traces, the steps $\{n_i\}$ are random and so will be different in every trace. Since the bleach rate can depend on the local environment of a fluorophore,^{39,40} we also reconsider q for each trace. The survival probability that a given fluorophore has not bleached after Δt is

$$q = e^{-\Delta t/\tau}, \quad (5)$$

where τ is the photobleach lifetime of a single fluorophore.

D. Bayesian prior

An important feature of our approach is our physically motivated prior over the number of fluorophores present at every time-step, i.e., the set of counts $\{n_i\}$ over frames (exposures) indexed by i . From one frame to the next, the number of active fluorophores present within each frame is given by the joint binomial distribution with the probability of not bleaching, q .¹⁴ This presumes that photobleaching of each fluorophore is independent of the others. When all frames are accounted for, we have the prior

$$P(\{n_i\}) = \prod_{i=1}^{N-1} \frac{n_{i-1}!}{n_i!(n_{i-1} - n_i)!} q^{n_i} (1 - q)^{n_{i-1} - n_i}. \quad (6)$$

Since q enters into the prior it is called a “hyperparameter,” but in our case, it is a physical parameter that must be inferred from the data. Our prior naturally accounts for the probability of observing shorter steps with larger number of fluorophores. As part of the prior, we also enforce monotonic decay in the number of fluorophores.

Additional priors can be developed to guide convergence and constrain other parameter values. For example, with known photophysics and background we impose fixed (delta-distributed) values of $\{\nu, a, \sigma_0^2, \sigma_1^2\}$. To infer these parameters from experimental calibration curves, we add uninformed (i.e., flat or uniform) priors to aid numerical convergence.

E. Maximum *a posteriori* probability (MAP) estimation

We found that maximum *a posteriori* probability (MAP) estimation gave simple, reliable, and computationally efficient results. Simply, MAP determines parameter values θ_{MAP} that maximize the posterior for a given data trace,²⁸

$$\theta_{MAP} = \arg \max_{\theta} P(\theta|D).$$

For computational convenience, the log-posterior was used $L \equiv \ln P(\theta|D)$.

F. Determining bleach step locations

Determining the optimal set of bleach steps is nontrivial because $\{n_i\}$ are integer valued and there are many possible discrete configurations. For a given initial number of fluorophores, n_0 , and number of data points, N , the number of possible combinations of bleach step locations are

$$\binom{n_0 + N - 1}{n_0} = \frac{(n_0 + N - 1)!}{n_0!(N - 1)!}. \quad (7)$$

This makes brute-force global optimization computationally impractical for $n_0 \gg 1$.

We also found (see the [supplementary material](#), Fig. S1) that the log-posterior is rough when simultaneously optimizing all of the model parameters when $n_0 > 1$, so that numerical convergence to the global MAP is not guaranteed. There are typically more than 1000 adjustable parameters, depending on the size of the dataset: with the photophysics, background, nonbleach probability, and the number of fluorophores at each time point $\{n_i\}$. A global optimization algorithm is difficult as well as computationally expensive. In practice, we first calibrate (see below) most of the continuous parameters, so that we can take most of the parameter values as sharp priors while determining bleach steps. This results in a smooth log-posterior when $n_0 > 1$, and reliable results.

When $n_0 > 1$, we first determine a range of acceptable n_0 values. The order-of-magnitude range can be determined by knowledge of the experimental system or can be quickly obtained by a naive ratio-metric approach. This range is a uniform prior on n_0 . Within the user defined range of n_0 values, we maximize the log-posterior with a golden section search⁴¹ that is adopted for discrete values. For each value of n_0 considered, an initial set of fluorophore counts $\{n_i\}$ are generated using

$$n_i = \text{round}(n_0 e^{-t_i/\tau}), \quad (8)$$

with the lifetime τ determined from an exponential fit to the curve. This initial array of fluorophore counts serves as a naive guess to start the optimization.

Since the MAP estimation of any continuous parameters (see below) depend on the discrete parameters $\{n_i\}$, effective optimization remains an iterative procedure. At each iteration, the timing of one (randomly chosen) individual bleach event is varied for all possible times between and including the previous and the subsequent bleach events. In principle, this allows the iterative procedure to explore all configurations of $\{n_i\}$. In practice, the largest log-posterior value is chosen at every iteration.

Iteration is continued until the log-posterior cannot be further improved. Since the parametric landscape is smooth when fitting to curves with known continuous parameters, or for calibration with $n_0 = 1$, this iterative procedure efficiently maximizes the posterior distribution.

G. MAP estimation of continuous parameters

The continuous parameters $\{v, a, q, \sigma_1^2, \sigma_0^2\}$ can each be easily determined by maximizing the log-posterior L .

Using $0 = \partial L / \partial q$, we obtain

$$q_{MAP} = \frac{N \langle n \rangle - n_0}{N \langle n \rangle - n_{N-1}}, \quad (9)$$

where $\langle n \rangle$ is the time-average number of fluorophores within the time series $\{n_i\}$ with length N ; n_0 is the number of fluorophores at the start of the series and n_{N-1} is the number of fluorophores at the end. While q_{MAP} requires $\{n_i\}$, we can easily update q_{MAP} at every step of the iterative procedure detailed above for $\{n_i\}$ to get a full MAP estimation.

The MAP estimates of the brightness per fluorophore v_{MAP} and background illumination a_{MAP} are determined by setting $\partial L / \partial v = 0$ and $\partial L / \partial a = 0$ to obtain

$$v_{MAP} = \frac{\sigma_{nI'}^2}{\sigma_{n'}^2} \quad (10)$$

and

$$a_{MAP} = \langle I \rangle' - v_{MAP} \langle n \rangle', \quad (11)$$

where the primes indicate a weighted average using the noise variance (i.e., $\langle I \rangle' \equiv \sum I_i w_i / \sum w_i$ with $w_i = 1/\sigma_i^2$), and $\sigma_{nI'}^2 \equiv \langle nI \rangle' - \langle n \rangle' \langle I \rangle'$.

The MAP estimates of the noise variance parameters σ_0^2 and σ_1^2 cannot be expressed in closed forms but can be determined self-consistently using

$$F \equiv \frac{\langle m_i \rangle}{\langle \sigma_1^2 n_i + \sigma_0^2 \rangle} = 1, \quad (12)$$

$$m_i \equiv (I_i - v_{MAP} n_i - a_{MAP})^2,$$

where the angular brackets denote averages. The bisection method⁴¹ is used to solve $F = 1$ and thus numerically determines the MAP noise parameters for each curve. For each value of σ_0^2 within the range $0 < \sigma_0^2 \leq \langle m_i \rangle$, the solution of F is initially bracketed between

$$\sigma_{1,low}^2 = 0$$

and

$$\sigma_{1,high}^2 = \sum_{n_i \neq 0} \frac{m_i}{n_i} / \left(N - \sum_{n_i=0} m_i / \sigma_0^2 \right). \quad (13)$$

In this paper, these continuous parameters are determined with calibration data with $\hat{n}_0 = 1$, and then applied as priors to general traces with $\hat{n}_0 \geq 1$. Calibration data can be obtained by fitting only the portions of data that contain the last bleach step or it could be done with a dedicated experiment.

H. Calibrating continuous parameters with experimental data

We developed a single-fluorophore ($\hat{n}_0 = 1$) test sample to illustrate calibration. A 28 base-pair single-stranded DNA oligonucleotide was labeled with a single Cy5 dye on the 3'-end, henceforth named A1-Cy5. The complementary strand of A1-Cy5 was modified with 3'-Biotin-TEG (triethylene glycol) for surface immobilization. After hybridizing A1-Cy5 with its complementary strand, the resultant double stranded DNA was immobilized onto the streptavidin-coated surface of a glass coverslip at the bottom of a flow chamber.⁸ The immobilized DNA molecules were suspended in a Tris-EDTA buffer (pH 8.0) with 2.5 mM protocatechuic acid (PCA), 50 nM protocatechuate-3,4-dioxygenase (PCD), and 10 mM Trolox (Sigma-Aldrich) added for photoprotection purposes.⁴² The sample was imaged on a custom-built total internal reflection fluorescence (TIRF) microscope using a 635-nm excitation laser (TECRL-635, WorldStarTech) with an intensity of 60 W/cm² measured at the sample.⁴³ These conditions resulted in negligible fluorophore blinking in our calibration data. The Cy5 fluorescence from the sample was recorded by an EMCCD camera (iXon DU-897BV, Andor) with an exposure time of 0.3 s/frame. A total of 315 molecules were identified in a 43.5 $\mu\text{m} \times 43.5 \mu\text{m}$ region of the sample, and a sequence of 1000 frames was acquired. At the end of the measurement, a significant fraction of the Cy5 molecules within the field of view were photobleached.

We extracted intensity traces from these images using the open-source CellProfiler application.⁴⁴ A CellProfiler pipeline (illustrated in the [supplementary material](#), Figs. S2 and S3) was used to first identify acceptable fluorescent objects of interest, and then to measure the intensities of the accepted objects over the time-separated series of images. Acceptable objects were those that were circular, were within a range of four to eight pixels in diameter, and did not have neighboring objects with any adjacent pixels. These criteria excluded overlapping objects and minimized signal overlap between the objects. We considered $n_0 \in \{0, 1\}$ and obtained MAP estimates of n_0 and all photophysical parameters. We found smooth log-posteriors during calibration. Of the 97 traces initially identified from the image set, a total of 49 remained after discarding those that contained interfering signals or that did not capture a bleach event (with $n_{0, \text{MAP}} = 0$). All 97 traces and the calibration code are provided as the [supplementary material](#).

Histograms of MAP estimates were constructed for the ensemble of traces (shown in the [supplementary material](#), Fig. S4). We estimated the continuous parameters from the means of the v , a , σ_0^2 , and σ_1^2 distributions. The mode of the distribution of q values was used to estimate the nonbleach probability, due to the large skew of the q distribution. We used these estimates (rounded) for modeling with $\hat{n}_0 > 1$: our default parameters are fluorophore brightness $v = 260$, background intensity $a = 1530$, per fluorophore noise variance $\sigma_1^2 = 230$, background noise variance $\sigma_0^2 = 430$, and nonbleach probability $q = 0.9973$. We also use $N = 5000$ data points with $\Delta t = 0.3$ s as a default.

I. Photon counting

Using an intensity unit $v_0 = \sigma_1^2/v$ gives us rescaled parameters with $\hat{v} \equiv v/v_0 = v^2/\sigma_1^2 = \tilde{\sigma}_1^2$. In these units, the variance of the intensity due to a single fluorophore equals its intensity, so these

units count detected photons. The total expected (average) number of detected photons per fluorophore is then easily determined to be

$$\begin{aligned} \mu_{\text{eff}} &= \int_0^\infty e^{-t/\tau} \hat{v} dt / \Delta t \\ &= \frac{\tau v^2}{\sigma_1^2 \Delta t}. \end{aligned} \quad (14)$$

The average fraction of photons detected in a single-timestep is $\alpha = \Delta t/\tau = \hat{v}/\mu_{\text{eff}}$. From our calibration, in photon units, we have $\hat{v} \simeq 294$, $\mu_{\text{eff}} \simeq 1.1 \times 10^5$, $\tau = 111$ s, and $\alpha = \Delta t/\tau \simeq 0.0027$.

J. DNA ladders

For further experimental validation, we used a self-assembled interleaving DNA structure (“ladder”) formed by two building blocks, termed A1 and A2. These are partially complementary 30-bp ssDNA strands which are each labeled with a Cy5 dye at the 5'-end (see below, top panel of [Fig. 5](#)). The initial unit in the ladder contains a biotin at the 3'-end to immobilize the ladder on glass via the pull-down method used for TIRF imaging (see Subsection II H).

We analyzed photobleaching data from a ladder with an average size of 16 ± 5 building blocks, as estimated from the average diffusion coefficient.⁴⁵ We measured the diffusion coefficient of the ladder using fluorescence correlation spectroscopy (FCS);⁴⁶ the FCS curve and the fitting analysis are included in the [supplementary material](#). The DNA photobleaching data were acquired as described in Sec. II H, except the excitation intensity was approximately 6 times lower. We also acquired single-step data from a calibration sample immediately prior to the multi-step sample under exactly the same excitation/detection conditions.

K. Alternative algorithms

The simplest approach used to determine n_0 given a calibrated parameterization is to take the ratio of the initial background-subtracted intensity to the step-size v . This is labeled “ratiometric” in subsequent plots. A significant advantage of this approach is that it is very simple to implement. One can also first filter the data to reduce noise. Commonly, a Chung-Kennedy (CK) edge-preserving filter is used.²³ Here, we use a CK exponent $p = 25$ (chosen for best results) and different widths of the analysis window, w . We only show $w = 4$ and $w = 16$ in subsequent plots for visual clarity. The unfiltered ratiometric approach corresponds to a limiting width $w = 0$.

The analyses of simulated photobleaching traces were also done using a change-point (CP) analysis algorithm that has been described previously,³⁴ implemented with custom software.⁸ A significant advantage of this approach is that no calibration of fluorophore intensity is needed. Briefly, an experimentally measured intensity-time trace $I(t)$ typically exhibits uniform binning, so that data can be represented as a vector $[I(\Delta t), I(2\Delta t), \dots, I(n\Delta t), \dots]$. For CP analysis, a segment of the trace up to the time point $t = n\Delta t$ is considered with $n > 3$. For this truncated trace, a cumulative-sum vector is computed

$$\text{cusum}_n(k) \equiv \sum_{j=1}^k I(j\Delta t), \quad (15)$$

with $k \in \{1, \dots, n\}$. A deviation vector dev_n is then defined as the difference between the cusum_n vector and the cumulative sum of a

virtual trace of constant intensity equal to the mean of all n time points,

$$devi_n(k) \equiv cusum_n(k) - k \frac{\sum_{j=1}^n I(j\Delta t)}{n}. \quad (16)$$

A putative change-point is selected at the time point $k = m$ where $|devi_n(k)|$ is maximum. The candidate change-points are subject to a two-sample t -test screening procedure that requires the intensity values immediately before and after the candidate change-point to be statistically different with a p -value less than 10^{-5} . For more details, see the [supplementary material](#) and Refs. 8 and 34.

III. RESULTS

A. Algorithm performance vs \hat{n}_0

We simulated ensembles of 1000 bleach curves for different \hat{n}_0 values, using the experimentally calibrated default parameters. The computational time required to identify the MAP $\{n_i\}$ scales linearly with the initial number of fluorophores \hat{n}_0 , with the range of n_0 values, and with the number of frames N —and is approximately independent of other parameters. The average computing time needed for one trace with $\hat{n}_0 = 10$ is 1 min, using a one-core (2 GHz CPU) laptop.

We first characterized the width of the MAP distribution of n_0 , σ_{n_0} , which represents the imprecision of our MAP approach. In [Fig. 2\(a\)](#), we plot σ_{n_0} vs \hat{n}_0 with large green circles. The error bars are calculated from bootstrap resampling of the bleach fits. The absolute width remains small for all \hat{n}_0 studied. The approximate asymptotic behavior $\sim \hat{n}_0^{1/2}$ at larger \hat{n}_0 is also indicated (black dashed line). The asymptotic behaviour implies that the fractional error σ_{n_0}/\hat{n}_0 will decrease as \hat{n}_0 increases. In [Fig. 2\(b\)](#), we plot the bias vs \hat{n}_0 . The bias remains small (below one step) for the MAP approach even when $\hat{n}_0 > 10^3$.

Considering ratiometric approaches, we see in [Fig. 2\(a\)](#) that averaging typically *worsens* imprecision with respect to the raw ratiometric estimate by averaging over multiple stochastic steps. We

also see in [Fig. 2\(b\)](#) that averaging also introduces a bias (systematic underestimation) for the same reason. However, the raw ratiometric approach (small dark squares) almost performs as well as the MAP approach in terms of precision. In contrast, the change-point algorithm (CP) does not perform better than either the ratiometric or the MAP approaches, though it does have the advantage of not needing any calibration.

We see that non-CP techniques share the asymptotic imprecision $\sim \hat{n}_0^{1/2}$ dependence. This imprecision arises from the larger noise with \hat{n}_0 , via [Eq. \(2\)](#). While the CP algorithm appears to approach an asymptotically constant imprecision at large \hat{n}_0 , it is at the expense of significant inaccuracy (large bias).

B. Algorithm performance vs exposure time Δt

The exposure times Δt of the camera used in single-molecule photobleach experiments not only determine the temporal resolution but also provide effective averaging for larger Δt . How should Δt be adjusted, either during data acquisition or after, for best algorithmic performance?

The intensity, or photon count, at each time-step is proportional to Δt —and so is the step-size v . Due to the Poisson nature of shot-noise, the noise-variance parameters σ_0^2 and σ_1^2 are also proportional to Δt —and so is the background offset a . However, the nonbleach probability q depends nonlinearly on the bleach rate and exposure time, so from [Eq. \(5\)](#), we have $q = q_0^{\Delta t/\Delta t_0}$. To explore Δt effects, we simulated ensembles of 1000 curves with $n_0 = 10$ using kinetic Monte Carlo and with a time-weighted average over the number of fluorophores within each time-step Δt . Default parameters were used for $\Delta t_0 = 0.3$ s, and they were scaled appropriately with Δt . The same absolute duration t_{max} was used for all Δt , i.e., $N = 5000\Delta t_0/\Delta t$.

In [Fig. 3\(a\)](#), we have scaled Δt by the average time to the first bleach event, $\delta t_1 \equiv \tau/n_0$. For the ratiometric approaches, we see increasing imprecision at both very small and very large Δt . At larger Δt , we are averaging over $O(\Delta t/\delta t_1)$ steps, which provides

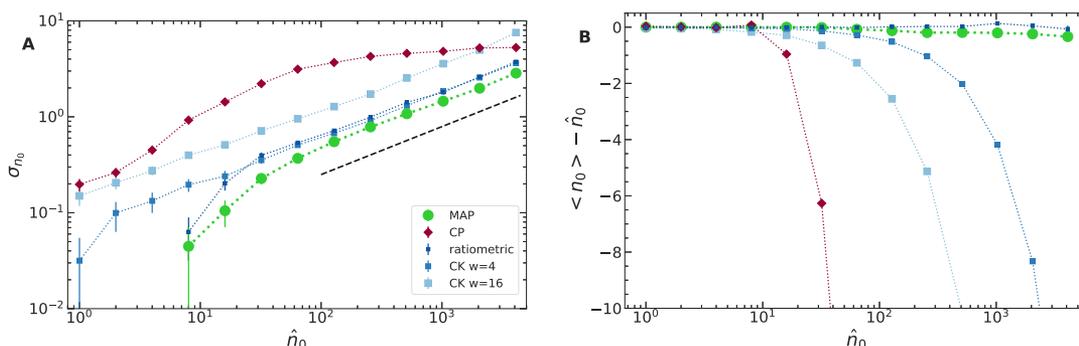


FIG. 2. (a) Imprecision vs \hat{n}_0 . The standard deviation of the estimated number of fluorophores σ_{n_0} vs the number of fluorophores \hat{n}_0 , for various methods as indicated in the legend: MAP algorithm (large green circles), CP (red diamonds), ratiometric without filtering (small dark-blue squares), or with CK filtering ($w = 4$, medium blue squares and $w = 16$, large blue squares). The same legend applies to subsequent figures as well. Averages are taken over 1000 fits of the simulated data, using the default photophysical parameters. Bootstrap error bars are shown. The black dashed line indicates $\sim \hat{n}_0^{1/2}$ asymptotic behavior. For $\hat{n}_0 < 8$, the MAP algorithm returned the exact \hat{n}_0 for all fits, so the estimated $\sigma_{n_0} = 0$. (b) The corresponding bias $\langle n_0 \rangle - \hat{n}_0$ vs the true \hat{n}_0 . Most algorithms have a negative bias, indicating a systematic underestimate of \hat{n}_0 . Bias increases in magnitude with \hat{n}_0 .

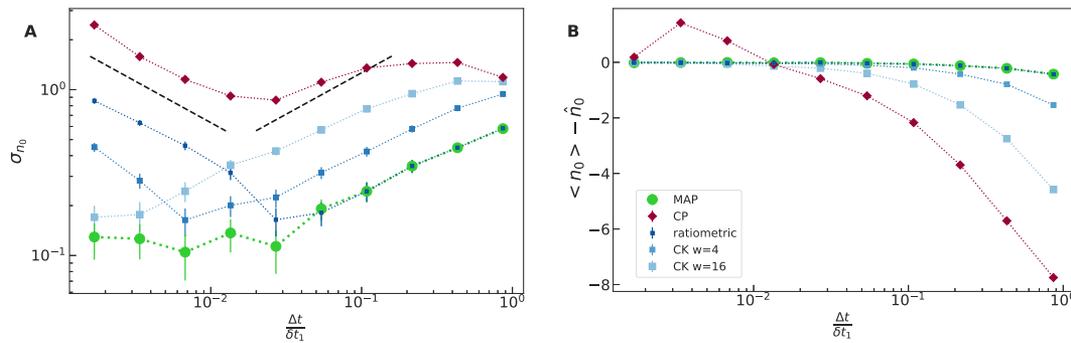


FIG. 3. (a) Imprecision vs scaled exposure time. The standard-deviation of the estimated n_0 , σ_{n_0} vs the scaled exposure time $\Delta t/\delta t_1$, where $\delta t_1 \equiv \tau/\hat{n}_0$ is the average duration of the first bleach step. For each point, an ensemble of 1000 model curves was simulated from default parameters with $\hat{n}_0 = 10$, adjusted for exposure time in the time-series data (see the text), and fit. Imprecision grows with larger Δt as $\sim (\Delta t/\delta t_1)^{1/2}$ (right black dashed line). Imprecision of ratiometric and CP approaches also grows with smaller Δt as $\sim (\Delta t/\delta t_1)^{-1/2}$ (left black dashed line), due to noise in the time-series data. (b) Bias vs $\Delta t/\delta t_1$ for the same data. The increasing bias at larger Δt is due to averaging over initial bleach events.

a standard-deviation of $O(\sqrt{\Delta t/\delta t_1})$ —as indicated by the increasing black dashed line. Supporting this interpretation, in Fig. 3(b), we see that the bias also increases with Δt . For smaller Δt , the decreasing averaging provided by Δt in the ratiometric approaches increases the imprecision by $O(1/\sqrt{\Delta t})$ —as indicated by the decreasing black dashed line. The CP algorithm behaves like the ratiometric approaches, but with a nonmonotonic imprecision at larger Δt .

For the Bayesian MAP approach, imprecision increases at large Δt but *not* at small Δt . This is because all of the data are used in the MAP inference, not just the first data-point as in the ratiometric approaches. For $\Delta t/\delta t_1 \lesssim 0.03$, we obtain a noise-floor that is determined by the finite-brightness of the fluorophores (μ_{eff} , see Sec. III C).

C. More photons

An important part of the experimental design is the choice of fluorophore. From Eq. (14), the total measured photon output of each fluorophore is $\mu_{\text{eff}} = \tau v^2 / (\sigma_1^2 \Delta t)$. As μ_{eff} increases at constant

Δt or α , there are more photons per frame and so a higher signal-to-noise ratio. Intuitively, we expect that larger μ_{eff} should improve the performance of all algorithms.

We varied μ_{eff} by changing the bleach step size v for ensembles of 1000 bleach traces, each with $\hat{n} = 10$ and default parameters. In Fig. 4(a), we plot the imprecision σ_{n_0} vs μ_{eff} . As expected, the precision of the estimate of n_0 generally improves as μ_{eff} increases. The MAP algorithm performs best for all μ_{eff} . We see a characteristic $\sigma_{n_0} \sim 1/\sqrt{\mu_{\text{eff}}}$ asymptotic behavior at larger μ_{eff} : doubling the number of photons effectively doubles the number of “measurements” of \hat{n}_0 .

For ratiometric approaches, the optimal amount of filtering depends on μ_{eff} . Smaller photon yields benefit from more filtering, since they have more noise. However, the filtering leads to increased bias as seen in Fig. 4(b). The bias does not improve with larger μ_{eff} due to the fixed $\Delta t/\delta t_1 = -\hat{n}_0 \ln q \simeq 0.02$. Interestingly, the bias degrades significantly at smaller μ_{eff} —particularly for the MAP and CP approaches. The origin of the bias for the MAP approach appears to be due to using a “point-estimate” of the maximum of the

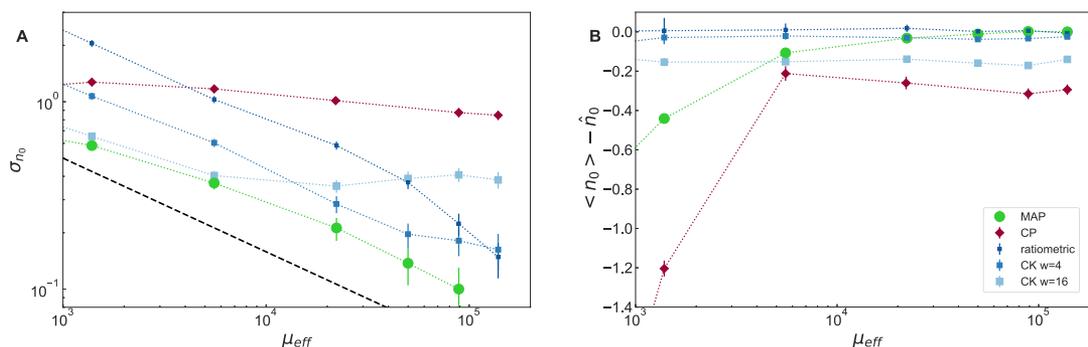


FIG. 4. (a) Imprecision vs photon yield. The standard deviation of the estimated number of fluorophores, σ_{n_0} , vs the total detected photons per fluorophore μ_{eff} . We estimate σ_{n_0} from 1000 simulated traces with $\hat{n}_0 = 10$, $a = 1000$, $q = 0.998$, $\Delta t = 0.3$ s, $\sigma_0^2 = 1600$, and $\sigma_1^2 = 900$. Bootstrap error bars are shown. We vary μ_{eff} by changing the step-size v . For $\mu_{\text{eff}} \gtrsim 10^5$, the MAP algorithm returned the exact \hat{n}_0 for all fits, so the estimated $\sigma_{n_0} = 0$. The black dashed line shows the expected $1/\sqrt{\mu_{\text{eff}}}$ dependence at larger μ_{eff} . (b) Bias vs photon yield. The systematic error of the estimate of n_0 vs μ_{eff} .

posterior, rather than the full (unbiased) posterior distribution. This indicates significant skew asymmetry in the posterior, especially at small μ_{eff} .

D. Multistep ladders

Figures 5(a)–5(d) depict the results of the comparative MAP vs CP analysis of the DNA ladder and the corresponding calibration single-dye sample. Acceptable fluorescent spots in the raw

images [Fig. 5(a)] were identified and intensity traces were extracted using the CellProfiler pipeline (see the [supplementary material](#) for details). A total of 149 intensity traces were analyzed for the DNA ladder and 164 traces for the calibration sample. An example of how a multistep photobleaching trace for a DNA ladder molecule was fit by the two algorithms is shown in Fig. 5(b). Bayes MAP fitting detects 14 steps, while CP fit detects approximately half as many, i.e., 7 steps. The discrepancy arises from the fact that the CP algorithm is not bound to the fixed step intensity imposed in MAP based on

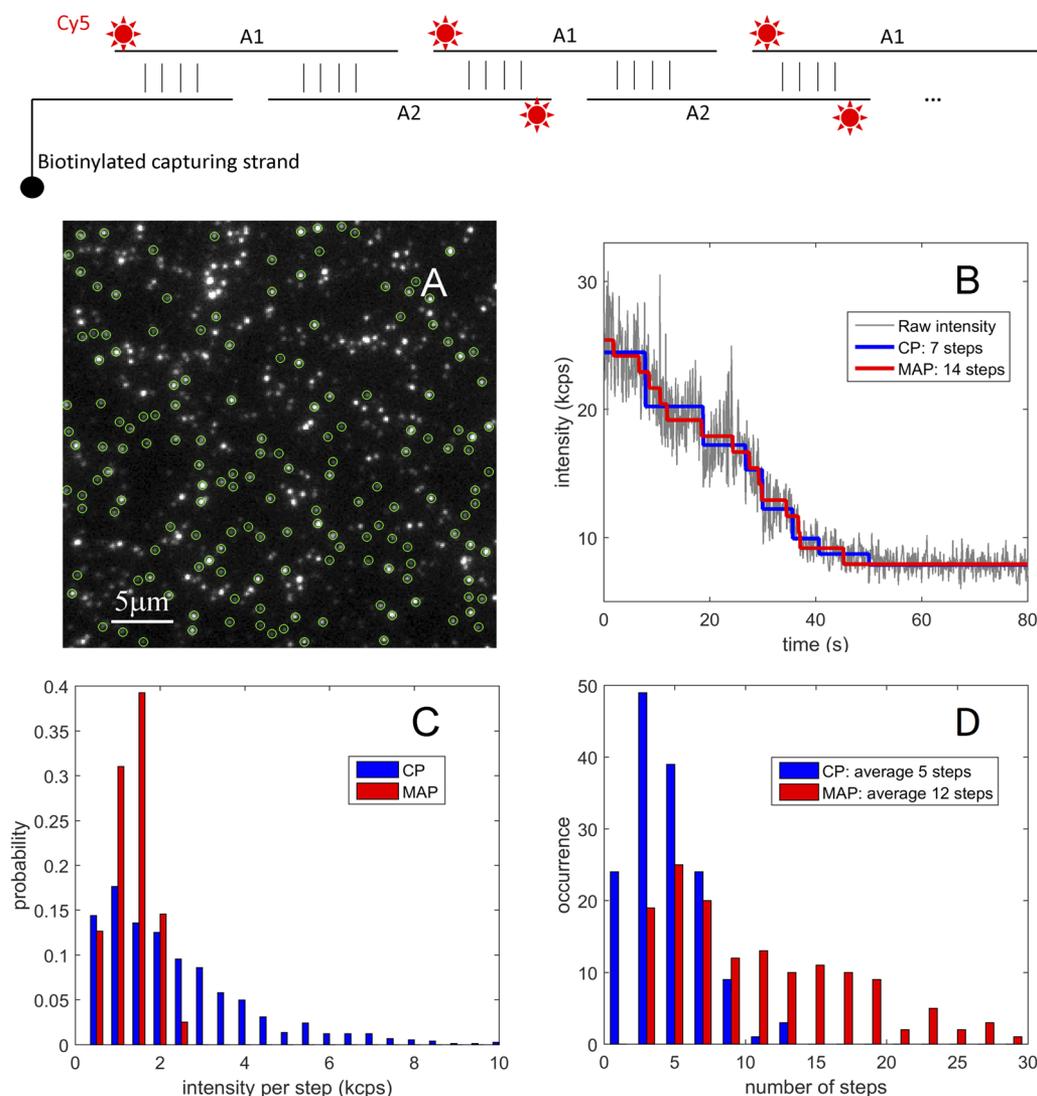


FIG. 5. Comparison of MAP and CP algorithms with experimental multistep photobleaching data. (Top) Scheme of DNA “ladder” of Cy5-labeled, partially complementary ssDNA building blocks, termed A1 and A2; the first subunit is biotinylated for surface immobilization. (a) TIRF image of surface-immobilized DNA ladders, with green circles highlighting the spots selected for photobleaching analysis. (b) Analysis of the intensity trace from a single spot using the CP algorithm detected 7 photobleaching steps (blue), while 14 steps were detected using the MAP algorithm (red). (c) Histograms of the step intensity for CP (blue) and MAP (red) algorithms. The MAP histogram has 164 events from a single-step calibration sample; the average value (1.3 kcps) was used as the calibration step-size. The CP histogram includes 744 step-wise photobleaching events from all DNA ladder molecules. (d) The distribution of the number of steps per DNA ladder molecule reported by CP (blue) and MAP (red) algorithms. A total of 149 ladder molecules from two separate regions in the same sample chamber were analyzed.

a priori, but it can take any value. This is clear in Fig. 5(c), where the step intensity distributions for the control (red) and the ladder (output by CP, blue) are shown. The former is narrow with an average of 1.3 ± 0.2 kcps, which was used by MAP to fit the ladder data, while the latter is broad (average 2.7 ± 1.5 kcps) and extends to 10 kcps. Correspondingly, the distribution of the number of steps (fluorophores) detected per molecule produced by the two methods is significantly different [Fig. 5(d)]. In agreement with the simulation results, CP tends to underestimate the number of fluorophores, probably due to a combination of detection noise and concomitant photobleaching events. On the other hand, MAP analysis gives an average of 12 ± 7 steps per trace, closer to the FCS-determined value (16 ± 5), although the distribution is quite broad, with up to 30 steps per trace. Most likely, this reflects a real size heterogeneity of DNA ladders formed by spontaneous self-assembly of the two DNA building blocks.

IV. DISCUSSION

Many techniques have been developed to count fluorophores in the steplike intensity traces measured from single-molecule photobleaching experiments. Most of these filter out noise from time-traces during analysis, e.g., using the nonlinear edge-preserving Chung-Kennedy (CK) filter.²³ Filtering and other *ad hoc* statistical techniques typically require significant fine-tuning of the analysis pipeline, and can have significant bias in the resulting estimates of step counts.²⁵ Principled model-driven Bayesian approaches have also been proposed, but they have used *ad hoc* priors to avoid biased estimates of step counts.¹⁶

We have presented a Bayesian approach to step-counting with a physics-based prior [Eq. (6)]. Our prior builds in the steplike monotonic decay of molecular photobleaching, with more rapid bleaching at early times when more fluorophores are unbleached. We have shown that a MAP point-estimation (maximizing the posterior distribution) of individual photobleach traces is practical, provides simple formulas for the photophysics parameters, and leads to precise and minimally biased estimates of the true number of steps \hat{n}_0 . We have shown with experimental data that our method lets us calibrate the photophysics, and with simulated data that our method lets us reliably count thousands of fluorophores.

Using simulated data, has allowed us to characterize the imprecision σ_{n_0} and bias $\langle n_0 \rangle - \hat{n}_0$ of the estimate of the number of fluorophores \hat{n}_0 per molecule. The imprecision is the expected random error from one photobleach trace, while the bias is the average systematic error that remains even with many traces. We have explored the roles of the number of fluorophores \hat{n}_0 , the exposure time of individual data points Δt , and the fluorophore stability as parameterized by the average number of photons detected per fluorophore before bleaching μ_{eff} . We have investigated our MAP approach, a change-point (CP) approach, and ratiometric approaches both with and without edge-preserving CK filtering.

For all methods, both the imprecision and bias grew with \hat{n}_0 ; the optimal step size was approximately $\Delta t \approx 0.03 \delta t_1$, i.e., 3% of the first bleach time; and the best results were obtained with the most photostable fluorophores, with large $\mu_{\text{eff}} \gtrsim 10^4$. However, the details vary considerably among the methods. In particular, large bias can be observed, typically by undercounting \hat{n}_0 in the CK or CP approaches. A general conclusion is that analysis of simulated data is

necessary to reliably estimate (and minimize) imprecision and bias for any given step-counting algorithm and for each experimental set-up.

Our Bayesian MAP approach, with a photobleaching prior, has the smallest imprecision of these approaches with minimal bias. Imprecision grows as $\hat{n}_0^{1/2}$ and decreases as $\mu_{\text{eff}}^{1/2}$, both due to intrinsic shot-noise. Furthermore, small $\Delta t/\delta t_1$ does not degrade our principled approach.⁴⁷ This is because the information contained in the trace is not lost by using shorter exposure times since all of the trace is used. In contrast, long exposure times with $\Delta t \gtrsim 0.1\tau/\hat{n}_0$ will provide effectively smooth initial steps. This introduces imprecision, but also bias since our Bayesian model no longer identically represents the data. Nevertheless, the bias in Fig. 3(b) remains small because the MAP approach uses information contained in the noise [Eq. (2)], which reflects \hat{n}_0 even at larger Δt .

Since absolute imprecision grows as $\hat{n}_0^{1/2}$, the relative imprecision gets smaller as \hat{n}_0 increases. However, if we insist on a precise fluorophore count for every trace, with $\sigma_{n_0} \lesssim 0.1$, then Fig. 2 indicates that we will be limited to $n_0 \lesssim 20$ with $\mu_{\text{eff}} \approx 1.1 \times 10^5$. For these small \hat{n}_0 , the bias is insignificant. For a given σ_{n_0} , our MAP approach provides a twofold improvement of the maximal precise \hat{n}_0 over ratiometric approaches and is much better than CP.

The bare ratiometric approach has no bias, and is simple to implement once the photophysics has been calibrated. However, it suffers from increased imprecision due to noise, especially as \hat{n}_0 increases, for smaller exposure times Δt , and for smaller photon yield μ_{eff} . Using CK filtering can reduce imprecision, but it introduces significant bias at larger Δt . As a result, the optimal exposure time Δt will depend on the details of the filtering, and vice versa.

To be more specific, most of our results (Figs. 1–3) reflect the photophysics of the photostable red dye Cy5 with $\mu_{\text{eff}} \approx 1.1 \times 10^5$. Less photostable dyes or fluorophores, such as the green dyes ATTO532 or Alexa532, would have one order of magnitude smaller photon yield,²² i.e., $\mu_{\text{eff}} \approx 10^4$. In that regime, from Fig. 4 we see that significant imprecision is unavoidable with ratiometric approaches. Furthermore, because of the significant noise associated with $\mu_{\text{eff}} \lesssim 10^5$, appropriately filtered ratiometric approaches would be required. With off-the-shelf optics or for live-cell imaging, where oxygen scavenging buffers are not feasible, we may expect $\mu_{\text{eff}} \approx 10^3$ and more pronounced imprecision, and bias. Most applications of internal or external fluorescence standards are ratiometric, and suffer from these effects. That said, errors in the calibration may dominate the effects described here.⁵

A change-point approach has the advantage of requiring no calibration. As such, it may be the easiest approach to analyze only a few photobleach traces with a small number of steps.³⁵ However, it suffers from significant imprecision for all \hat{n}_0 and bias for $\hat{n}_0 \gtrsim 10$ or for smaller Δt or μ_{eff} . It also requires data to extend to $n_i = 0$, i.e., complete bleaching of the molecule.

A principled model-based Bayesian approach to fitting photobleach traces, such as ours, depends on the appropriateness of the physical model. For example, though fluorophore blinking can be minimized experimentally,^{42,48–51} our model does not account for any remaining blinking unless the blinking is fast enough with respect to τ that it can be included in our noise model. We also do not accommodate variation of the photophysical parameters between fluorophores, for example, due to their local molecular or

cellular context—again unless this variation can be treated as a fast random process that can be implicitly included in our noise model. With sufficient experimental guidance, these effects could be explicitly included with a hierarchical Bayesian model.⁵² We expect that the bias and imprecision of any given analysis approach will depend on the details of the system, and will require modeling to properly characterize. Additionally, any prebleaching or inhomogeneous (inefficient) labeling needs to be accounted for after our estimation of n_0 active fluorescent labels.²⁹

We have presented a MAP approach, which is a point-estimate of the Bayesian posterior. We separated photophysics calibration from \hat{n}_0 estimation so that the likelihood of each would remain smooth, and so that we could more easily compare with a variety of other \hat{n}_0 estimation approaches. In principle, calibration can be made arbitrarily accurate with sufficient number of calibration traces. However, a global calibration assumes that the illumination intensity, and other aspects of the photophysics, is uniform across the sample. In principle, efficient Markov chain Monte Carlo (MCMC) approaches should be able to characterize even a rough likelihood for individual photobleach traces without calibration, and would also be able to avoid bias due to skewed posteriors with small μ_{eff} . This is left for future development.

SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for the change-point algorithm details, Figs. S1–S4 about calibration, and *code* and *data* used in the analysis.

ACKNOWLEDGMENTS

We thank the ACGT Corporation (Toronto, Canada) for supplying the DNA ladder construct. This work was supported by the Natural Sciences and Engineering Research Council of Canada (Grant No. RGPIN 2017-06030 to C.C.G. and Grant No. RGPIN 2019-05888 to A.D.R.).

REFERENCES

- M. H. Ulbrich and E. Y. Isacoff, *Nat. Methods* **4**, 319 (2007).
- N. Watanabe and T. J. Mitchison, *Science* **295**, 1083 (2002).
- K. E. Hines, *J. Gen. Physiol.* **141**, 737 (2013).
- R. J. Arant and M. H. Ulbrich, *ChemPhysChem* **15**, 600 (2014).
- V. C. Coffman and J.-Q. Wu, *Mol. Biol. Cell* **25**, 1545 (2014).
- A. Penna, A. Demuro, A. V. Yeromin, S. L. Zhang, O. Safrina, I. Parker, and M. D. Cahalan, *Nature* **456**, 116 (2008).
- H. McGuire, M. R. P. Aourousseau, D. Bowie, and R. Blunck, *J. Biol. Chem.* **287**, 35912 (2012).
- R. V. Shivnaraine, D. D. Fernandes, H. Ji, Y. Li, B. Kelly, Z. Zhang, Y. R. Han, F. Huang, K. S. Sankar, D. N. Dubins, J. V. Rocheleau, J. W. Wells, and C. C. Gradinaru, *J. Am. Chem. Soc.* **138**, 11583 (2016).
- H. Yokota, Y. A. Chujo, and Y. Harada, *Biophys. J.* **104**, 924 (2013).
- Y. Sugiyama, I. Kawabata, K. Sobue, and S. Okabe, *Nat. Methods* **2**, 677 (2005).
- J.-Q. Wu, C. D. McCormick, and T. D. Pollard, *Methods Cell Biol.* **89**, 253 (2008).
- M. C. Leake, J. H. Chandler, G. H. Wadhams, F. Bai, R. M. Berry, and J. P. Armitage, *Nature* **443**, 355 (2006).
- M. C. Leake, N. P. Greene, R. M. Godun, T. Granjon, G. Buchanan, S. Chen, R. M. Berry, T. Palmer, and B. C. Berks, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15376 (2008).
- C. R. Nayak and A. D. Rutenberg, *Biophys. J.* **101**, 2284 (2011).
- N. Rosenfeld, T. J. Perkins, U. Alon, M. B. Elowitz, and P. S. Swain, *Biophys. J.* **91**, 759 (2006).
- K. Tsekouras, T. C. Custer, H. Jashnsaz, N. G. Walter, and S. Pressé, *Mol. Biol. Cell* **27**, 3601 (2016).
- H. Qian and E. L. Elson, *Proc. Natl. Acad. Sci. U. S. A.* **87**, 5479 (1990).
- J. Tellinghuisen and C. W. Wilkerson, *Anal. Chem.* **65**, 1240 (1993).
- M. A. Digman, R. Dalal, A. F. Horwitz, and E. Gratton, *Biophys. J.* **94**, 2320 (2008).
- C. Eggeling, J. Widengren, R. Rigler, and C. Seidel, *Anal. Chem.* **70**, 2651 (1998).
- N. C. Shaner, M. Z. Lin, M. R. McKeown, P. A. Steinbach, K. L. Hazelwood, M. W. Davidson, and R. Y. Tsien, *Proc. SPIE* **7191**, 719105 (2009).
- V. Glembockyte, J. Lin, and G. Cosa, *J. Phys. Chem. B* **120**, 11923 (2016).
- S. H. Chung and R. A. Kennedy, *J. Neurosci. Methods* **40**, 71 (1991).
- B. C. Carter, M. Vershinin, and S. P. Gross, *Biophys. J.* **94**, 306 (2008).
- V. C. Coffman and J.-Q. Wu, *Trends Biochem. Sci.* **37**, 499 (2012).
- In this paper, we ignore the individual photon arrival times and consider only a timeseries of photon counts for each macromolecule that is uniformly binned in time, as characterized by α .
- A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis* (CRC Press, 2014).
- J.-L. Gauvain, *IEEE Trans. Speech Audio Process.* **2**, 291 (1994).
- K. E. Hines, *Biophys. J.* **108**, 2103 (2015).
- J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, and C. H. Wiggins, *Biophys. J.* **97**, 3196 (2009).
- L. P. Watkins and H. Yang, *J. Phys. Chem. B* **109**, 617 (2005).
- B. Kalafut and K. Visscher, *Comput. Phys. Commun.* **179**, 716 (2008).
- B. Shuang, D. Cooper, J. N. Taylor, L. Kislej, J. Chen, W. Wang, C.-B. Li, T. Komatsuzaki, and C. F. Landes, *The J. Phys. Chem. Lett.* **5**, 3157 (2014).
- C. R. Gallistel, S. Fairhurst, and P. Balsam, *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13124 (2004).
- H. Zhang and P. Guo, *Methods* **67**, 169 (2014).
- K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press, 2012).
- M. Hirsch, R. J. Wareham, M. L. Martin-Fernandez, M. P. Hobson, and D. J. Rolfe, *PLoS One* **8**, e53671 (2013).
- D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- A. V. Mamontova, A. M. Bogdanov, and K. A. Lukyanov, *BioTechniques* **58**, 258 (2015).
- A. Acharya, A. M. Bogdanov, B. L. Grigorenko, K. B. Bravaya, A. V. Nemukhin, K. A. Lukyanov, and A. I. Krylov, *Chem. Rev.* **117**, 758 (2017).
- W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, 3rd edition (Cambridge University Press, Cambridge, 2007), Chap. 9.
- C. E. Aitken, R. A. Marshall, and J. D. Puglisi, *Biophys. J.* **94**, 1826 (2008).
- B. Liu, A. Mazouchi, and C. C. Gradinaru, *J. Phys. Chem. B* **114**, 15191 (2010).
- A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, *Genome Biol.* **7**, R100 (2006).
- R. M. Robertson, S. Laib, and D. E. Smith, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7310 (2006).
- Y. Li, R. V. Shivnaraine, F. Huang, J. W. Wells, and C. C. Gradinaru, *Biophys. J.* **115**, 881 (2018).
- We do require that enough photons are delivered per Δt to satisfy our Gaussian approximation of the Poisson photon statistics, i.e. $\alpha\mu_{\text{eff}} \gg 1$.
- Y. Fu, J. Zhang, and J. R. Lakowicz, *Langmuir* **24**, 3429 (2008).
- R. Dave, D. S. Terry, J. B. Munro, and S. C. Blanchard, *Biophys. J.* **96**, 2371 (2009).
- T. Ha and P. Tinnefeld, *Annu. Rev. Phys. Chem.* **63**, 595 (2012).
- Q. Zheng, M. F. Juette, S. Jockusch, M. R. Wasserman, Z. Zhou, R. B. Altman, and S. C. Blanchard, *Chem. Soc. Rev.* **43**, 1044 (2014).
- N. Cressie, C. A. Calder, J. S. Clark, J. M. Ver Hoef, and C. K. Wikle, *Ecol. Appl.* **19**, 553 (2009).