SUPPORTING MATERIAL

Empirical Bayes Method Using Surrounding Pixel Information for Number and Brightness Analysis

R. Fukushima, J. Yamamoto, M. Kinjo

Herein, we provide supplemental information to describe the theoretical derivations, materials and methods, and supplemental data in detail. First, we briefly describe the derivation of the equations in the main text. Then, we describe the simulation method, preparation of EGFP and cells, and measurement and analysis methods. Finally, in the supplemental data, we present additional results that complement the main text and figures.

THEORETICAL DERIVATIONS

Derivation of the recursive formula describing the probability distribution for photon count detected by double detectors

The probability distribution for observing X and Y photons by double detectors is given in Eq. 6. The equation is a summation of W time Bernoulli processes, and W is an independent random variable determined by a Neyman type A distribution. The probability-generating function is given by

$$\begin{aligned} G(z_1, z_2) &= \sum_{X=0}^{\infty}\sum_{Y=0}^{\infty} P(X,Y) {z_1}^X {z_2}^Y \\ &= \exp[\nu(\exp[\varepsilon(pz_1 + qz_2 - 1)] - 1)], \end{aligned} \tag{S1}$$

where $p + q = 1$. The a-th order derivative of the probability-generating function with respect to z_1 is given as follows from mathematical induction ($a \geq 1$):

$$\frac{\partial^a G}{\partial {z_1}^a}(z_1, z_2) = \sum_{l=0}^{a-1} \binom{a-1}{l} \nu(\varepsilon p)^{a-l} \exp[\varepsilon(pz_1 + qz_2 - 1)] \frac{\partial^l G}{\partial {z_1}^l}(z_1, z_2). \tag{S2}$$

The b-th order derivative of Eq. S2 with respect to z_2 is given as follows from mathematical induction ($b \geq 0$):

$$\begin{aligned} &\frac{\partial^{a+b} G}{\partial {z_1}^a \partial {z_2}^b}(z_1, z_2) \\ &= \sum_{m=0}^{b} \binom{b}{m} \sum_{l=0}^{a-1} \binom{a-1}{l} \nu \varepsilon^{a-l+b-m} p^{a-l} q^{b-m} \exp[\varepsilon(pz_1 + qz_2 - 1)] \frac{\partial^{l+m} G}{\partial {z_1}^l \partial {z_2}^m}(z_1, z_2). \end{aligned} \tag{S3}$$

Substituting $z_1 = z_2 = 0$ into the definition of the probability-generating function yields

$$\frac{\partial^{a+b} G}{\partial {z_1}^a \partial {z_2}^b}(0,0) = a!\, b!\, P(a,b). \tag{S4}$$

Substituting $z_1 = z_2 = 0$ and Eq. S4 into Eq. S3 gives

$$P(a,b) = \frac{1}{a} \nu \varepsilon p \exp[-\varepsilon] \sum_{m=0}^{b} \frac{(\varepsilon q)^{b-m}}{(b-m)!} \sum_{l=0}^{a-1} \frac{(\varepsilon p)^{a-l-1}}{(a-l-1)!} P(l,m). \tag{S5}$$

Derivation of a photon-counting model for a dead-time-affected single detector

Bédard described the probability for dead-time-affected photon count (1). The probability distribution for observing W photons on a dead-time-affected detector is given as follows for $W < 1/\delta$:

$$\begin{aligned} P(W|\eta,\delta) &= \int_0^\infty P(W|\eta, I_0, \delta)P(I_0)\mathrm{d}I_0 \\ &= \int_0^\infty \{F(W-1|\eta, I_0, \delta) - F(W|\eta, I_0, \delta)\}P(I_0)\mathrm{d}I_0, \end{aligned} \tag{S6}$$

where the probability distribution F is given by

$$F(W|\eta, I_0, \delta) = \frac{\gamma\{W+1, \eta I_0(1-W\delta)\}}{W!}, \tag{S7}$$

where $\gamma(W+1,\xi)$ is an incomplete gamma function, and $\gamma(W+1,\xi) = \int_0^\xi u^W \exp(-u)\,\mathrm{d}u$. δ is obtained by dividing dead time τ by the sampling time ($\delta = \tau/T$). O'Donnell derived a Taylor expansion of the probability distribution from the first to third terms, and Ackermann and Hogreve derived that from the fourth to fifth terms (2). Hillesheim and Müller corrected the dead-time effect with the first term of the expansion (3, 4). We generalize this formulation and present it in a different form for our purpose of approximation. The Taylor expansion is given by

$$\begin{aligned} P(W|\eta, I_0, \delta) &= \sum_{k=0}^{\infty} \frac{\delta^k}{k!}\frac{\partial^k P}{\partial \delta^k}(W|\eta, I_0, 0) \\ &= \sum_{k=0}^{\infty} \frac{\delta^k}{k!}\left\{\frac{\partial^k F}{\partial \delta^k}(W-1|\eta, I_0, 0) - \frac{\partial^k F}{\partial \delta^k}(W|\eta, I_0, 0)\right\}, \end{aligned} \tag{S8}$$

where the derivative of the distribution F is given as follows from mathematical induction ($k = 1, 2, 3, \ldots$):

$$\frac{\partial^k F}{\partial \delta^k}(W|\eta, I_0, 0) = (W)^k \mathrm{Poi}(W|\eta I_0) \sum_{l=0}^{k-1}\left[(-1)^{l+1}\binom{k-1}{l}(\eta I_0)^{k-l}\frac{W!}{(W-l)!}\right]. \tag{S9}$$

$\binom{a}{b}$ is a binomial coefficient given by $\binom{a}{b} = \frac{a!}{b!(a-b)!}$. Each term in the infinite series of Eq. S8 contains a Poisson distribution multiplied by $(\eta I_0)^k$: $\mathrm{Poi}(W|\eta I_0)$, $(\eta I_0)\mathrm{Poi}(W|\eta I_0)$, …or $(\eta I_0)^k \mathrm{Poi}(W|\eta I_0)$. Such a Poisson distribution can be rewritten by using $(\eta I_0)^k \mathrm{Poi}(W|\eta I_0) = (W+k)!/W!\,\mathrm{Poi}(W+k|\eta I_0)$. The infinite series can be separated into multiple terms having $\mathrm{Poi}(W+m|\eta I_0)$ and function $A_m(W,\delta)$ with $m = 0,1,2,\ldots$ as follows:

$$P(W|\eta, I_0, \delta) = \sum_{m=0}^{\infty} A_m(W, \delta)\mathrm{Poi}(W + m|\eta I_0). \tag{S10}$$

Under our assumption, the probability distribution for observing W photons on a dead-time-affected detector is

$$P(W|\nu, \varepsilon, \delta) = \sum_{m=0}^{\infty} A_m(W, \delta)\mathrm{Ney}(W + m|\nu, \varepsilon). \tag{S11}$$

Derivation of a photon-counting model for dead-time-affected double detectors

For a double-detector system, the joint probability distribution for observing X and Y photons on dead-time-affected detectors is given as follows:

$$P(X, Y|\eta_1, \eta_2, \delta_1, \delta_2) = \int_0^{\infty}\int_0^{\infty} P(X|\eta_1, I_1, \delta_1)P(Y|\eta_2, I_2, \delta_2)P(I_1, I_2)\,\mathrm{d}I_1\,\mathrm{d}I_2, \tag{S12}$$

where δ_1 and δ_2 represent the dead time of each detector divided by the sampling time. Under our assumption, the joint probability is rewritten as follows for $X < 1/\delta_1$ and $Y < 1/\delta_2$ with $f = 0,1,2,\ldots$ and $g = 0,1,2,\ldots$:

$$\begin{aligned} P(X, Y|\nu, \varepsilon, p, \delta_1, \delta_2) &= \sum_{m=0}^{\infty}\sum_{h=0}^{m} A_h(X, \delta_1)A_{m-h}(Y, \delta_2)P(X + h, Y + m - h|\nu, \varepsilon, p) \\ &= A_0(X, \delta_1)A_0(Y, \delta_2)P(X, Y|\nu, \varepsilon, p) + A_1(X, \delta_1)A_0(Y, \delta_2)P(X + 1, Y|\nu, \varepsilon, p) \\ &+ A_0(X, \delta_1)A_1(Y, \delta_2)P(X, Y + 1|\nu, \varepsilon, p) + A_2(X, \delta_1)A_0(Y, \delta_2)P(X + 2, Y|\nu, \varepsilon, p) \\ &+ A_1(X, \delta_1)A_1(Y, \delta_2)P(X + 1, Y + 1|\nu, \varepsilon, p) + A_0(X, \delta_1)A_2(Y, \delta_2)P(X, Y + 2|\nu, \varepsilon, p) \\ &+ \cdots + A_f(X, \delta_1)A_g(Y, \delta_2)P(X + f, Y + g|\nu, \varepsilon, p). \end{aligned} \tag{S13}$$

Derivation of $\mathcal{Q}$ function on expectation maximization algorithm

We describe the derivation of the $\mathcal{Q}$ function on lognormal prior with gamma hyperprior. Bishop described the details of the derivation generally (5) (Section 9.4, the EM Algorithm in General). The $\mathcal{Q}$ function on lognormal prior is derived similarly. The marginal likelihood is given in Eq. 48. The natural logarithm of the marginal likelihood is given by

$$\log \mathcal{M}\left(\boldsymbol{x}_{1:J}, \boldsymbol{y}_{1:J} | \hat{p}, \mu, \sigma, \alpha, \beta\right) = \sum_{j=1}^{J}\left[\log \int_0^\infty \int_0^\infty \mathcal{L}\left(\boldsymbol{x}_j, \boldsymbol{y}_j | \nu_j, \varepsilon_j, \hat{p}\right) \mathrm{LN}\left(\nu_j | \mu, \sigma\right) \mathrm{Uni}\left(\varepsilon_j\right) \mathrm{d}\nu_j \, \mathrm{d}\varepsilon_j\right] + \log \mathrm{Gam}(\sigma | \alpha, \beta). \tag{S14}$$

We introduce a probability distribution $R\left(\nu_j, \varepsilon_j\right)$ and use Jensen's inequality as follows:

$$\begin{aligned}
&\log \mathcal{M}\left(\boldsymbol{x}_{1:J}, \boldsymbol{y}_{1:J} | \hat{p}, \mu, \sigma, \alpha, \beta\right) \\
&= \sum_{j=1}^{J}\left[\log \int_0^\infty \int_0^\infty R\left(\nu_j, \varepsilon_j\right) \frac{\mathcal{L}\left(\boldsymbol{x}_j, \boldsymbol{y}_j | \nu_j, \varepsilon_j, \hat{p}\right) \mathrm{LN}\left(\nu_j | \mu, \sigma\right) \mathrm{Uni}\left(\varepsilon_j\right)}{R\left(\nu_j, \varepsilon_j\right)} \mathrm{d}\nu_j \, \mathrm{d}\varepsilon_j\right] + \log \mathrm{Gam}(\sigma | \alpha, \beta) \\
&\geq \sum_{j=1}^{J}\left[\int_0^\infty \int_0^\infty R\left(\nu_j, \varepsilon_j\right) \log \frac{\mathcal{L}\left(\boldsymbol{x}_j, \boldsymbol{y}_j | \nu_j, \varepsilon_j, \hat{p}\right) \mathrm{LN}\left(\nu_j | \mu, \sigma\right) \mathrm{Uni}\left(\varepsilon_j\right)}{R\left(\nu_j, \varepsilon_j\right)} \mathrm{d}\nu_j \, \mathrm{d}\varepsilon_j\right] + \log \mathrm{Gam}(\sigma | \alpha, \beta) \\
&= \sum_{j=1}^{J}\left[\int_0^\infty \int_0^\infty R\left(\nu_j, \varepsilon_j\right) \log \mathrm{LN}\left(\nu_j | \mu, \sigma\right) \mathrm{d}\nu_j \, \mathrm{d}\varepsilon_j\right] + \log \mathrm{Gam}(\sigma | \alpha, \beta) + \mathrm{const.}
\end{aligned} \tag{S15}$$

The right-hand side of the last equation in Eq. S15 is a lower bound on the log-marginal likelihood. The lower bound is maximized with a fixed μ_{old} and σ_{old} when $R\left(\nu_j, \varepsilon_j\right)$ is equal to $P\left(\nu_j, \varepsilon_j, | \boldsymbol{x}_j, \boldsymbol{y}_j, \hat{p}, \mu_{\mathrm{old}}, \sigma_{\mathrm{old}}, \alpha, \beta\right)$, and the maximized lower bound is a $\mathcal{Q}$ function.

Derivation of overlap fraction

We simply consider the overlap fraction during scanning in the *x*–*y* plane. We assume that the confocal volume has a radius r and the microscope scans a pixel size s within a sampling time. The illuminated area within a sampling time is the trajectory that is given by moving the circular confocal volume along the line of scanning. The illuminated area is the summation of the areas of the circle with a radius r and the rectangle with lengths of $2r$ and s for its two sides: $\pi r^2 + 2sr$. The overlapped area between the adjacent pixels is the circular area, and is thus given by πr^2. Therefore, the overlap fraction is given by $\pi r^2/(\pi r^2 + 2sr)$.

Relation between average of particle number, particle brightness, and photon count average

The product of the particle number and particle brightness is equal to the temporal average of the photon counts (Eq. 18) for MoM and ML. Either the particle number or particle brightness is underestimated if the other is overestimated. One could consider that these relations are the same for the average of the estimates as well, and that the results presented in Fig. 2 are doubtful. Fig. 2G shows an overestimation of the particle number, but Fig. 2H shows an accurate estimation of the particle brightness in MoM and ML. However, these relations are not the same for the average of the estimates. We present a counter-example below, wherein we assume that the spatial distribution of the particle number and the temporal average of the photon counts follow gamma distribution and delta distribution, respectively:

$$\mathrm{Gam}(\nu|\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\,\nu^{\alpha-1}\exp(-\beta\nu), \tag{S16}$$

$$\mathrm{Del}(\lambda|d) = \begin{cases}\infty, & \lambda = d,\\ 0, & \lambda \neq d,\end{cases} \tag{S17}$$

where λ is a random variable for the temporal average of photon counts. The assumption of delta distribution means that the temporal average of the photon counts is a constant d at any spatial position. The distribution of $1/\nu$ is an inverse gamma distribution, and $\varepsilon = \lambda/\nu$. The distribution of particle brightness is given by

$$\mathrm{P}(\varepsilon|\alpha,\beta,d) = \frac{1}{d}\frac{\beta^{\alpha}}{\Gamma(\alpha)}\left(\frac{\varepsilon}{d}\right)^{-\alpha-1}\exp\left(-\frac{\beta d}{\varepsilon}\right). \tag{S18}$$

The spatial average of the particle number and particle brightness is $\mathbb{E}[\nu] = \alpha/\beta$ and $\mathbb{E}[\varepsilon] = d\beta/(\alpha-1)$, respectively. Therefore, the product of the averages is $\mathbb{E}[\nu]\mathbb{E}[\varepsilon] = \alpha d/(\alpha-1)$. In contrast, the spatial average of photon count averages is $\mathbb{E}[\lambda] = d$. The product of the averages is not equal to the average of the photon count averages, i.e., $\mathbb{E}[\nu]\mathbb{E}[\varepsilon] > \mathbb{E}[\lambda]$. For a sufficiently large α, the gamma distribution converges to normal distribution, and the product of averages is close to d. This discussion reveals that the results presented in Fig. 2 are not doubtful.

Appearance of negative particle number and particle brightness in MoM

In MoM, a negative value of the particle number and brightness is occasionally observed. The cause for this is an insufficient number of observations for covariance estimation. Let X_i and Y_i be a photon count at the ith observation ($i = 1, 2, \ldots, I$) in detectors 1 and 2, respectively. The estimators for the particle number and brightness in TD–N&B are respectively as follows (6):

$$\hat{\nu} = \frac{\bar{X}\bar{Y}}{\bar{C}}, \tag{S19}$$

$$\hat{\varepsilon} = \frac{(\bar{X} + \bar{Y})\bar{C}}{\bar{X}\bar{Y}}. \tag{S20}$$

where $\bar{X}$ and $\bar{Y}$ are the sample average of X_i and Y_i, respectively, and $\bar{C}$ is sample covariance. $\bar{X}$, $\bar{Y}$, and $\bar{C}$ are given as follows:

$$\bar{X} = \frac{1}{I}\sum_{i=1}^{I} X_i, \tag{S21}$$

$$\bar{Y} = \frac{1}{I}\sum_{i=1}^{I} Y_i, \tag{S22}$$

$$\bar{C} = \frac{1}{I}\sum_{i=1}^{I} (X_i - \bar{X})(Y_i - \bar{Y}). \tag{S23}$$

The expected values of the sample average and covariance are given from the probability-generating function (Eq. S1) as follows:

$$\mathbb{E}[\bar{X}] = p\nu\varepsilon, \tag{S24}$$

$$\mathbb{E}[\bar{Y}] = (1-p)\nu\varepsilon, \tag{S25}$$

$$\mathbb{E}[\bar{C}] = p(1-p)\nu\varepsilon^2. \tag{S26}$$

If the number of observations is infinite, the estimators for particle number and brightness are equal to the true particle number and brightness, respectively; thus, the estimates are positive. However, there is no guarantee that the estimators are equal to the true ones if the number of observations is finite. Under finite observations, estimates are close to the true values, and deviate from the true value if the number of observations is insufficient. The sample covariance could become negative if the number of observations is insufficient. The negative value of the sample covariance results in negative values of particle number and brightness.

MATERIALS AND METHODS

Simulation

Photon count images were numerically simulated using Microsoft Visual Studio Community 2017, version 15.9.10 (Microsoft, USA). All programs were written in Visual C++. A typical simulation of FCS must simulate the trajectories of diffusing particles, but the temporal fluctuation of the fluorescence intensity is independent in number and brightness (N&B) analysis; thus, we simulated the temporal fluctuation with independent generation of random numbers. We generated the random numbers by using functions implemented in the header "random." The seeds for random numbers were generated by the function "std::random_device" and pseudorandom numbers were generated by the function "std::default_random_engine" and the seed. The pseudorandom numbers were transformed into distributions representing photon counts by using the functions "std::poisson_distribution" and "std::binomial_distribution." The separation probability on binomial distribution was assumed to be 0.5665 (this value is the same as the experimentally estimated value). We did not add any noise and background, to simplify our model. The photon counts were simulated with a sampling time of 10 μs. The frame size of the simulated images was 18×18, and the pixels excluding the edge were analyzed. As a result, the particle number and brightness with the size of 16×16 were obtained. One thousand images were simulated with a fixed particle brightness of 10 kHz and different particle numbers (ν = 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100) (Fig. 2). Further, the frame was separated into 2×2 grids, and 1000 images were simulated with a fixed particle number of 50 or 100 and different particle brightness (ε = 0.025, 0.050, 0.075, 0.100, 0.125, 0.150, 0.175, and 0.200) (Fig. 3). These simulations were performed independently on each pixel.

Purification of EGFP

The EGFP was purified using the His-tag method, which has been described previously (6).

Preparation of EGFP tandem oligomer-transfected HeLa cells

HeLa cells were maintained at 37 ℃ under a humidified atmosphere of 5% CO_2 in a Dulbecco's Modified Eagle's Medium (DMEM; Sigma-Aldrich, USA) supplemented with 100 μg/mL penicillin G (Sigma-Aldrich, USA) and 100 μg/mL streptomycin (Sigma-Aldrich, USA). The HeLa cells were seeded into an 8-well chambered coverglass (#155411, Thermo Fisher Scientific, USA) and incubated for 24 h. The cells (1.5×10^4 cells) were then transfected with 0.02 μg of EGFP-C1 plasmids encoding EGFP monomer, or plasmids encoding flexible linker (FL) linked EGFP tandem-oligomers (2–3 mer) (7, 8) using ViaFect (Promega, USA). The pCAGGS plasmid was added as necessary to increase the total amount of plasmids to 0.1 μg. The cells were incubated for 24 h. The medium was then renewed, and incubation was continued for 24 h. The medium was renewed again before the measurements were performed.

Image acquisition for N&B analysis

Confocal fluorescence microscopy images were obtained using an LSM 510 META ConfoCor3 system (Carl Zeiss, Germany) with a C-Apochromat 40×/1.2W Corr objective (Carl Zeiss, Germany). The EGFP was excited at a wavelength of 488 nm. EGFP fluorescence was split into two channels by a half mirror, filtered using a long-pass filter (LP505), and detected by two avalanche photodiodes, which has been described previously (6). The pinhole

size was 94 μm (1.32 airy unit). The zoom factor was ×6, and the pixel size was 0.146 μm × 0.146 μm. The X- and Y-scanning sizes were 256×64 pixels. The pixel dwell time was 6.39 μs, and the scanning was raster scan. Two thousand images were sequentially obtained. The laser had an output of 0.30 μW at the objective for 1.0% output (see Fig. S1.3 for details). For Fig. 4, the images were obtained with different dilution ratios of EGFP with a laser output of 1.0%. For Fig. 5, the images were obtained with different laser outputs (0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0%). For Fig. 6, the images of the cells were obtained with a laser output of 0.4%. To minimize the background intensity, we selected the cell having a fluorescence intensity greater than 400 kHz on average in the cytoplasm. The fluorescence intensity of the non-transfected cell was less than 20 kHz in an area where the fluorescence intensity was stationary. Thus, the fraction of the background intensity would be less than 0.05 in cytoplasm. On the other hand, the fluorescence intensity in the nucleus was lower than that in the cytoplasm for EGFP dimer and trimer. This could be because of the larger molecular size than an EGFP monomer; EGFP dimer and trimer are difficult to pass through nuclear pores. The fluorescence intensity in the nucleus was greater than 150 kHz on average. Thus, a fraction of the background intensity would be less than 0.13.

Analysis conditions

For analysis based on method of moments (MoM), we adopted two-detector number and brightness analysis (TD–N&B), which has been described previously (6). A method based on covariance estimation of photon count sequence and correction for dead time are not required. In this study, linear regression, boxcar filtering, and median filtering were not implemented, unlike in a previous study (6).

For maximum likelihood (ML) estimation and empirical Bayes–maximum a posteriori (EB–MAP) estimation, a statistical model without correction for dead time was applied to the simulated images. In contrast, a statistical model with correction for dead time was applied to the experimentally obtained images. We assumed the same dead time of 50 ns (performance data, EG&G, Canada) on both detectors ($\delta_1 = \delta_2 = 7.82 \times 10^{-3}$), and a separation probability $\hat{p}$ of 0.5665. The estimated value and standard error for $\hat{p}$ was 0.566565 ± 0.000031. In the calculation of A_0, A_1, and A_2, we approximated the summation in Eq. 10 by partial summation from 0 to 5, 1 to 5, and 2 to 5, respectively. Although the upper limit of the summation is $W + m$, the difference only slightly affects our result. This is because the difference between the total and partial summations is negligibly small. We plan to calculate the total summation in future experiments.

In the Newton–Raphson method, the initial pairs of $\nu^{(0)}$ and $\varepsilon^{(0)}$ were determined by choosing the pair giving the maximum log-likelihood between candidates. The candidates were chosen arbitrarily within the interval of the parameters. The iteration process of the Newton–Raphson method was terminated when the relative error of the parameters was less than 1.0×10^{-6} for each parameter. The relative error was estimated as follows for $\nu^{(k)}$ and $\nu^{(k+1)}$: $\nu^{(k+1)}/\nu^{(k)} - 1$. The relative error for $\varepsilon^{(k)}$ was similarly estimated. In ML estimation, log-likelihood sometimes gives maximum on the boundary of the parameter space. In the iteration process, we considered the estimates as boundary estimates, $\hat{\nu}$ of infinity, and $\hat{\varepsilon}$ of 0, when the estimates did not satisfy the above termination condition after 100 iterations. To solve the simultaneous linear equations in Eq. 31, we used Boost uBLAS (http://www.boost.org/libs/numeric/ublas/) implemented in Boost C++ Libraries, version 1.72.0.

In the EM algorithm, the choice of the initial pair of μ_{old} and σ_{old} and the termination of the iteration process

were the same as those in the Newton–Raphson method. EB–MAP estimation was not performed on the edge of the image because the number of surrounding pixels is less than 8. For numerical integration, we used a double exponential formula (9) implemented with the transformation $\exp(\pi/2 \sinh t)$. We calculated the double integral as iterated integral, and applied the transformation formula to the first and second integrals. The numerical integration was iteratively performed by doubling the number of nodes. The iteration was terminated when the relative error of the integral was less than 1.0×10^{-6}. The EM algorithm requires an iterative calculation of the double integral with different μ and σ. In the first integral with respect to ε, the integration does not depend on μ and σ. To reduce the computational costs, we recorded the integrated value of the first integral, and calculated the second integral iteratively using the recorded first integral. For the first integral, we truncated the infinite summation when the truncation error was less than 1.0×10^{-6}. For the second integral, we fixed the interval of t on the second integration as -2 to 3 ($3.36 \times 10^{-3} < \nu < 6.82 \times 10^{6}$), and confirmed that the truncation error is less than 1.0×10^{-6} at the endpoints. In this numerical integration, arithmetic computations were performed in a logarithmic scale to avoid underflow and overflow. For the analysis in Fig. 2, the typical computation time was 30 ms, 6 s, and 1400 s for MoM, ML, and EB–MAP, respectively.

Evaluation of analysis results

The fluorescence intensity was calculated by dividing the temporal average of the photon count by the sampling time. After estimations by MoM, ML, and EB–MAP, the variation of the estimates was shown by a box plot, and the sample average and sample standard deviation were estimated within the regions of interest (ROIs). In Fig. 2 and 3, the normalized bias for the particle number was estimated by $\bar{\nu}/\nu_{\mathrm{true}} - 1$, where $\bar{\nu}$ is the sample average of the particle number and ν_{true} is the true particle number. The normalized bias for particle brightness was similarly estimated. Normalized standard deviation was estimated from the sample standard deviation divided by true parameters. In the calculation of the sample average and sample standard deviation, infinite values were excluded, but negative values and 0 were included. Note that the fraction of the negative values in the particle number and particle brightness estimated by MoM is the same because the estimates of MoM satisfy Eq. 18. Further, the weighted least squares method was applied with the weight of sample variance (Fig. 4 and 5). One could consider that the negative values of particle number and brightness are non-physical and should be excluded when calculating the sample average. However, we do not recommend excluding the negative values. Fig. S3.5 shows the influence of negative values on the calculation of the normalized bias. Excluding negative values resulted in an overestimation of the normalized bias. Thus, we do not recommend excluding the negative values when calculating the sample average. Readers interested in why negative values appear may refer to the section above, “Appearance of negative particle number and particle brightness in MoM.”

For experiments with EGFP in solution, an ROI of size 16×16 was placed at the center of a 256×64 image of the estimates. For experiments with EGFP monomer, dimer, and trimer in a cell, the position of the ROIs of size 16×16 was chosen carefully within the nucleus and cytoplasm. Our theory assumed independent and identical distribution of the photon count sequence. Cell migration and photobleaching could be the causes of artifacts, owing to nonstationary changes in the sequence. We confirmed that the sequences are stationary on each pixel in the ROI. Strong distortion of the sequence would be easily identified, but weak distortion would be missed if the sequence is

checked only pixel-by-pixel. To avoid missing the weak distortion, we additionally confirmed that the time sequence made by averaging the spatial photon counts in the ROIs is stationary.

FCS measurements

FCS measurements were performed using an LSM 510 META ConfoCor3 system (Carl Zeiss, Germany) with a C-Apochromat 40×/1.2W Corr objective (Carl Zeiss, Germany). The EGFP was excited at a wavelength of 488 nm. EGFP fluorescence was split into two channels by a half mirror and filtered using a long-pass filter (LP505). The pinhole size was 66 μm. The measurement duration was 10 s × 20 times. In Fig. S5.3, the measurement duration was 20 s × 20 times for the lowest and second-lowest laser power. A single-component diffusion model with one triplet state was used for curve-fitting:

$$G(\tau_{\mathrm{cor}}) = \frac{1}{\nu_{\mathrm{FCS}}}\left[1 + f_{\mathrm{tri}} \exp\left(-\frac{\tau_{\mathrm{cor}}}{\tau_{\mathrm{tri}}}\right)\right]\left[1 + \frac{\tau_{\mathrm{cor}}}{\tau_{\mathrm{dif}}}\right]^{-1}\left[1 + \frac{\tau_{\mathrm{cor}}}{\varphi^2 \tau_{\mathrm{dif}}}\right]^{-\frac{1}{2}} + 1, \tag{S27}$$

where $G(\tau_{\mathrm{cor}})$ is a temporal correlation function, f_{tri} is a triplet fraction, τ_{tri} denotes the relaxation time of triplet state, ν_{FCS} denotes the average number of fluorescent particles in the confocal volume, τ_{dif} is the diffusion time, and φ is a structure parameter representing the ratio of beam waist to the axial radius. After pinhole adjustment, the structure parameter was determined using a 10^{-7} M standard solution of rhodamine 6G, whose diffusion coefficient is known, i.e., D_{R6G} = 414 μm^2/s (10, 11). The radius of confocal volume was calculated by $(4\tau_{\mathrm{dif}} D_{\mathrm{R6G}})^{0.5}$ theoretically. The particle brightness was calculated as a summation of count rate divided by the number of particles determined by curve-fitting with a cross correlation function. For Fig. 4, the measurements were performed with different dilution ratios of EGFP with a laser output of 1.0%. For Fig. 5, the measurements were performed with different laser outputs (0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0%). For Fig. 6, the measurements on cell were performed with a laser output of 0.2%.

SUPPLEMENTAL DATA

Validation of construction for prior distribution with lognormally distributed particle number

In a single pixel analysis, MoM and ML estimations are conducted at each pixel on an image, and the neighboring pixels of the analyzed pixel play no role in this. In contrast, the EB–MAP estimation utilizes the information of the surrounding pixels. To verify the EB method, we simulated a lognormally distributed particle number and determined the hyperparameter using two models, namely, lognormal prior and lognormal prior with gamma hyperprior. The purpose is to choose the appropriate hyperparameters for prediction; hyperparameters representing ambiguous prediction are acceptable. Fig. S1.1A and S1.1B show the spatial distribution of the particle number. Fig. S1.1C and S1.1D show the estimated hyperparameters and the Q_{25}–Q_{75} interval, respectively, determined by lognormal prior. In Fig. S1.1C, the estimates are scattered around the intersection of the true values. In Fig. S1.1D, the estimated Q_{25}–Q_{75} interval is sometimes shorter than the true one. These results suggest that underprediction

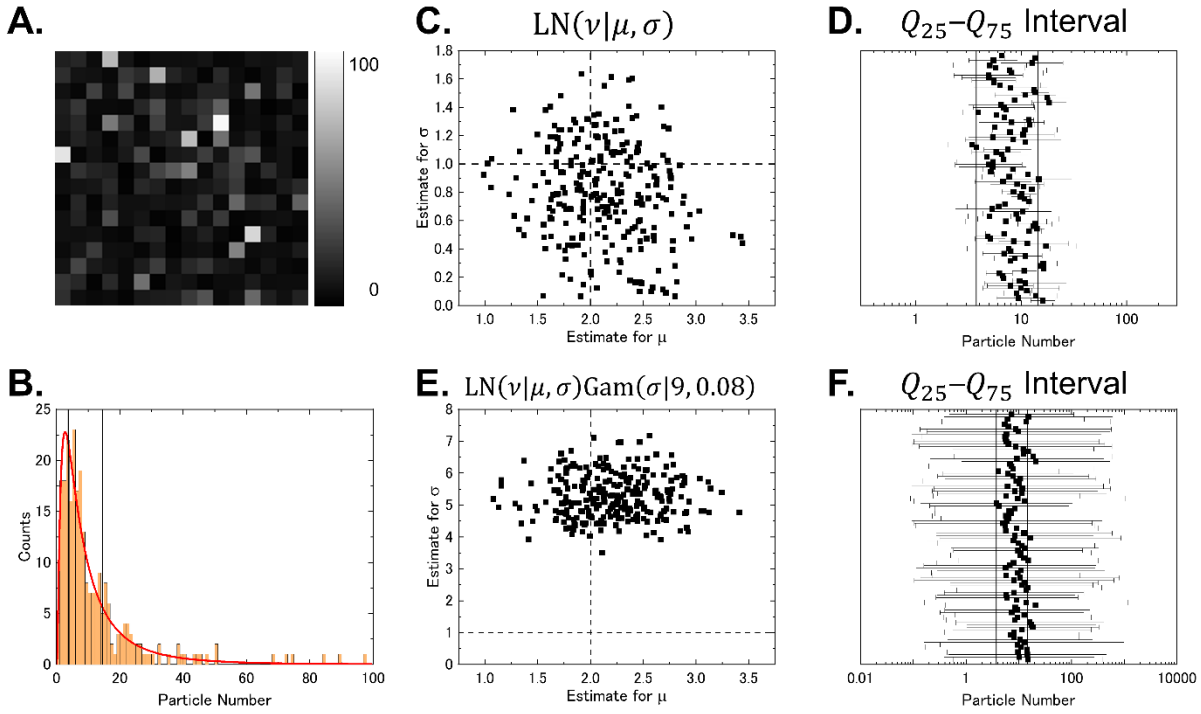

FIGURE S1.1: Comparison between prior models with lognormally distributed particle number

One thousand images were simulated with particle number and particle brightness of 10 kHz, and hyperparameters were estimated by the EB method with different models. The particle number was lognormally distributed independently ($\mu = 2$, $\sigma = 1$). (A) Spatial distribution of particle number. (B) Histogram of the particle number in A. The red curved line is a theoretical line of lognormal distribution. The black solid line shows the 25th and 75th percentiles of the distribution. (C, E) Scatter plot of the hyperparameters estimated by lognormal prior (C) and lognormal prior with gamma hyperprior (E). The dashed lines are true hyperparameters. (D, F) Q_{25}–Q_{75} interval of the lognormal distribution for which the estimated hyperparameters are substituted. The symbols show median, the error bars show the 25th and 75th percentiles of the lognormal distribution. The black solid lines show the 25th and 75th percentiles of the true distribution.

of the Q_{25}–Q_{75} interval occurs sometimes, which would cause bias. Thus, the lognormal prior model is not appropriate for the prediction of particle number. Fig. S1.1E and S1.1F show the estimated hyperparameters and the Q_{25}–Q_{75} interval, respectively, determined by lognormal prior with gamma hyperprior. Fig. S1.1E shows that the estimates for σ increased, but those for μ did not change markedly as before. Fig. S1.1F shows that the estimated Q_{25}–Q_{75} interval is wider than the true one. These results suggest that the prediction is ambiguous and it is difficult to miss the prediction, although the estimation of the hyperparameters is not accurate. Thus, the lognormal prior with gamma hyperprior can be applied to predict the particle number.

Validation of construction for prior distribution with grid-patterned particle number

In the lognormal prior model, a problem would arise if the distribution of the particle number on the surrounding pixels does not follow lognormal distribution. However, we assumed gamma distribution in addition to the lognormal distribution. The gamma distribution widens the lognormal distribution by increasing $\hat{\sigma}$. The wide prior distribution weakly constrains the estimation. It can be applied to predict distributions other than lognormal distribution. We simulated 2×2 grid-patterned images of the particle number to confirm that the presence of a sharp change in the surrounding pixels does not affect the estimation. The grid-patterned image has two types of area, each of which has the same particle number. The pixels have the particle number of 50 in one area, and 100 in the other.

Fig. S1.2A shows the distribution of the estimated $\hat{\mu}$ and $\hat{\sigma}$ for different particle brightness. At a low particle brightness, $\hat{\sigma}$ was estimated highly. Fig. S1.2B and S1.2C show the $Q_{25}-Q_{75}$ interval of the lognormal

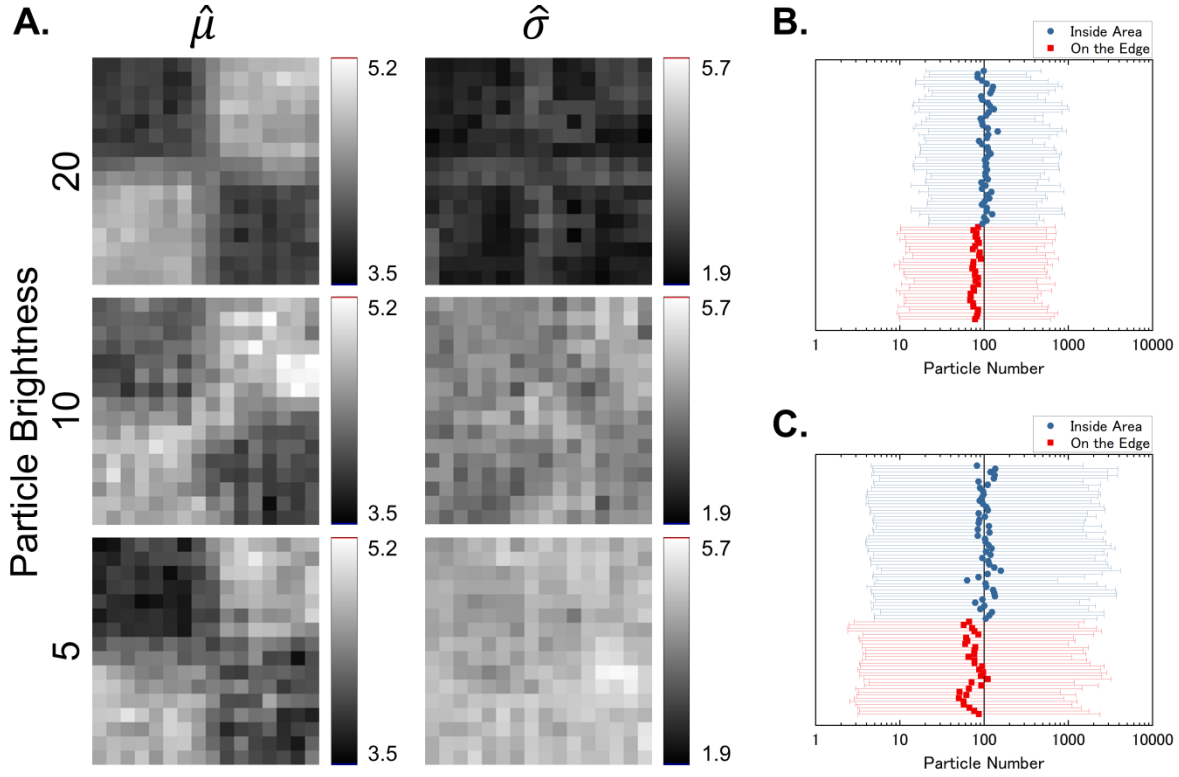

FIGURE S1.2: Optimization of hyperparameters by lognormal prior with gamma hyperprior on grid-patterned particle number

Simulation condition is the same as that in Fig. 3. (A) Spatial distribution of estimated hyperparameters with different particle brightness [kHz]. (B, C) Q_{25}–Q_{75} interval of the lognormal distribution for which the estimated hyperparameters are substituted. The intervals are shown for a particle brightness of 20 kHz (B) and 5 kHz (C). The symbols represent median, and error bars represent the 25th and 75th percentiles of the lognormal distribution. All plots belong to the area with the true particle number of 100. The results on the edge of the area are colored in red, and those inside the area are colored in blue. The black solid lines denote the particle number of 100.

distribution for which $\hat{\mu}$ and $\hat{\sigma}$ were substituted. Fig. S1.2B shows the intervals for a particle brightness of 20 kHz.

$\hat{\mu}$ inside the grid-patterned area having a particle number of 100 was close to the true value, and $\hat{\mu}$ on the edge of the grid-patterned area was smaller than the true value. The smaller $\hat{\mu}$ could have been caused by the adjacent pixels having a particle number of 50. The Q_{25}–Q_{75} intervals contain the true value both on the edge and in the inner area, and they would be sufficiently wide to predict the particle number. Fig. S1.2C shows the intervals for a particle brightness of 5 kHz. The Q_{25}–Q_{75} intervals were wider than those in Fig. S1.2B, and contain the true value. The lognormal prior with gamma hyperprior can be applied if the particle number changes sharply between the surrounding pixels.

Relation between laser power setting and actual output

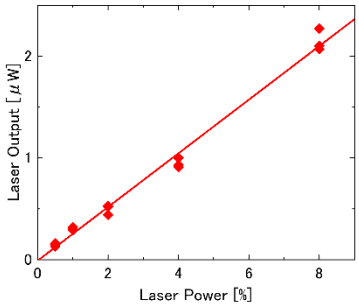

FIGURE S1.3: Dependence of actual laser output on laser power setting

The solid line is a fitted line obtained by the least square method (see Table S1.1 for details on parameters of the line).

Slope	Intercept
0.2636±0.0074	(-0.9±3.1)×10^{-2}

TABLE S1.1: Slope and intercept for the fitted line in Fig. S1.3

Estimated value and standard error

Supplemental data for Fig. 2

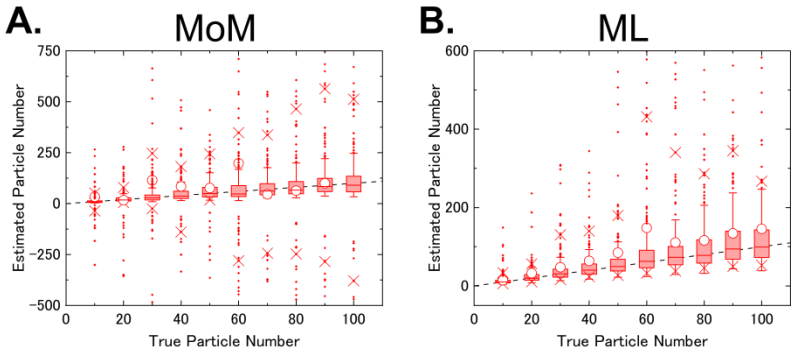

FIGURE S2.1: Broad view of estimated particle number in Fig. 2A and 2B

The view is adjusted to show the 5th and 95th percentiles of the particle number. The plots are shown similarly in Fig. 2. The minimum and maximum values are shown in Table S2.1 and S2.2. (A) Broad view of Fig. 2A. (B) Broad view of Fig. 2B.

True Particle Number	Minimum	Maximum
10	-1.439×10^{3}	7.999×10^{3}
20	-6.468×10^{2}	2.782×10^{2}
30	-1.251×10^{3}	1.020×10^{4}
40	-4.913×10^{3}	2.217×10^{4}
50	-8.467×10^{2}	1.795×10^{3}
60	-2.181×10^{3}	3.442×10^{4}
70	-3.832×10^{3}	9.617×10^{2}
80	-1.168×10^{4}	1.118×10^{4}
90	-4.259×10^{3}	3.290×10^{3}
100	-7.638×10^{5}	2.478×10^{3}

TABLE S2.1: Minimum and maximum values of estimated particle number in Fig. 2A

True Particle Number	Minimum	Maximum
10	4.493	1.488×10^{2}
20	7.980	1.102×10^{3}
30	1.372×10	1.258×10^{3}
40	1.717×10	1.798×10^{3}
50	2.256×10	3.578×10^{3}
60	2.372×10	8.799×10^{3}
70	2.918×10	1.661×10^{3}
80	3.196×10	1.612×10^{3}
90	4.362×10	1.675×10^{3}
100	3.847×10	5.089×10^{3}

TABLE S2.2: Minimum and maximum values of estimated particle number in Fig. 2B

Supplemental data for Fig. 2

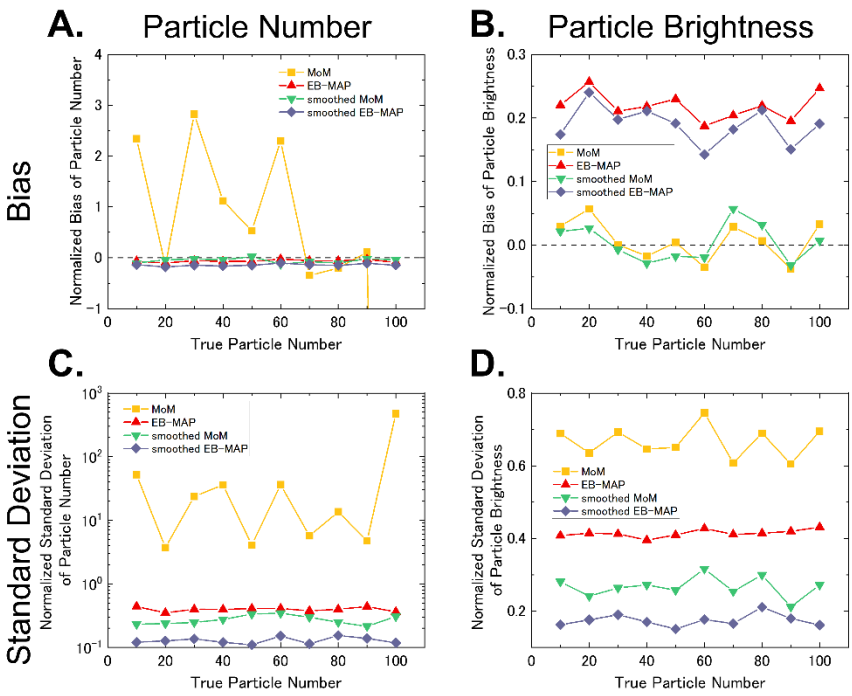

FIGURE S2.2: Noise smoothing by median filter

For noise smoothing, a median filter of size 3×3 was applied to the result of MoM and EB–MAP in Fig. 2. Normalized bias and standard deviation are shown for particle numbers (A, C) and particle brightness (B, D).

Supplemental data for Fig. 3

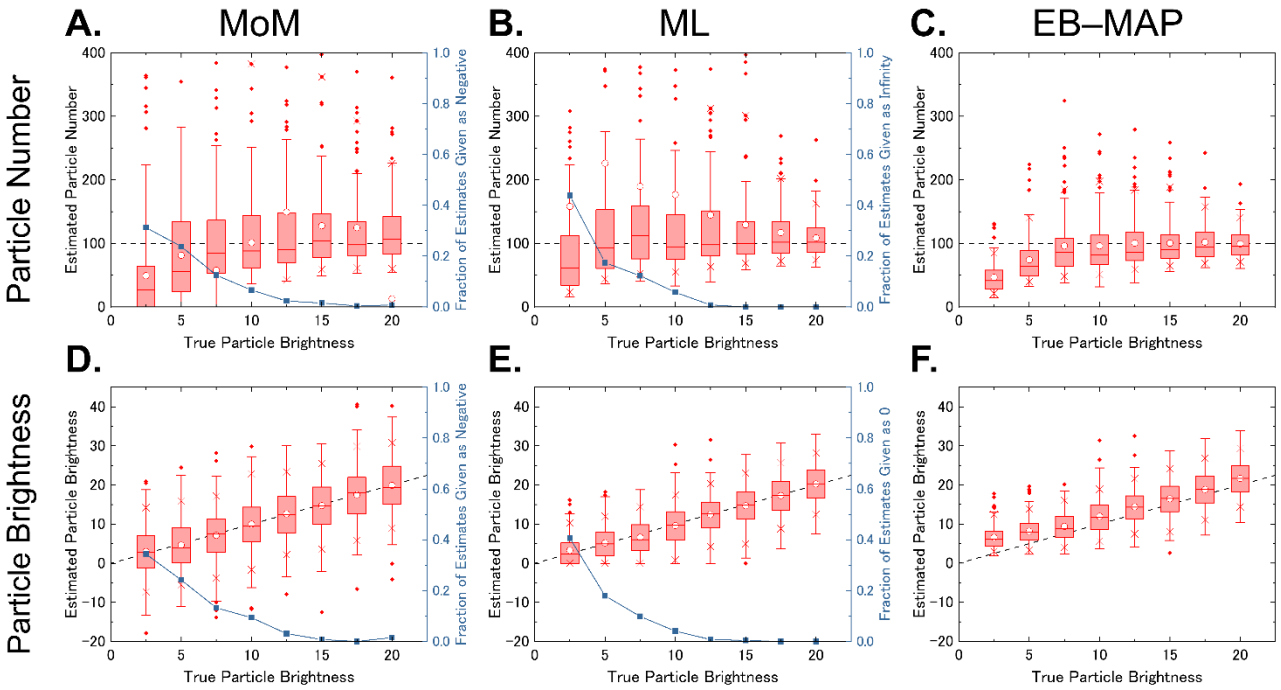

FIGURE S3.1: Variation in estimates for different particle brightness

One thousand images were simulated with a fixed particle number and different particle brightness, and analyzed by MoM (A, D), ML (B, E), and EB–MAP (C, F). (A, B, C) Dependence of estimated particle number on true particle brightness [kHz]. The particle number is shown for the area having a true particle number of 100. Figures S3.1A and S3.1B are shown in enlarged view (see Fig. S3.2 for broad view). (D, E, F) Dependence of estimated particle brightness [kHz] on true particle brightness. All figures are shown similarly in Fig. 2.

Supplemental data for Fig. 3

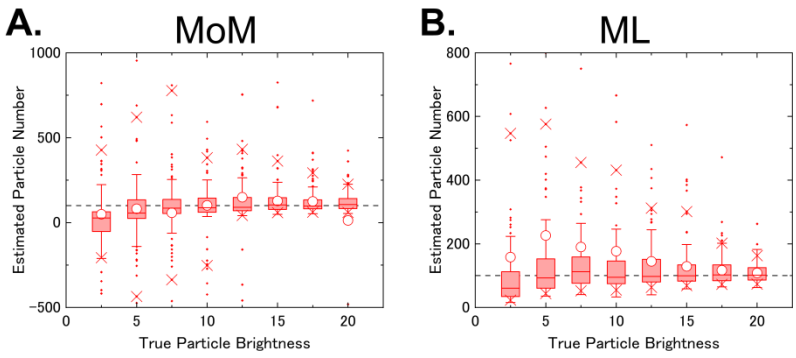

FIGURE S3.2: Broad view of estimated particle number in Fig. S3.1A and S3.1B

The view is adjusted to show the 5th and 95th percentiles of the particle number. The minimum and maximum values are listed in Table S3.1 and S3.2. (A) Broad view of Fig. S3.1A. (B) Broad view of Fig. S3.1B.

True Particle Brightness	Minimum	Maximum
2.5	-1.331×10^{4}	4.244×10^{2}
5.0	5.244×10	7.188×10^{2}
7.5	-4.770×10^{3}	3.087×10^{3}
10.0	-1.656×10^{3}	2.922×10^{3}
12.5	-1.140×10^{3}	2.357×10^{3}
15.0	-7.794×10^{3}	2.485×10^{3}
17.5	-5.578×10^{3}	5.790×10^{3}
20.0	-5.123×10^{2}	1.851×10^{3}

TABLE S3.1: Minimum and maximum values of the estimated particle number in Fig. S3.1A

True Particle Brightness	Minimum	Maximum
2.5	6.250×10	2.629×10^{2}
5.0	6.454×10	4.720×10^{2}
7.5	5.863×10	1.092×10^{3}
10.0	3.974×10	1.380×10^{3}
12.5	3.290×10	2.492×10^{3}
15.0	4.023×10	2.583×10^{3}
17.5	3.595×10	8.441×10^{3}
20.0	1.561×10	3.534×10^{3}

TABLE S3.2: Minimum and maximum values of the estimated particle number in Fig. S3.1B

Supplemental data for Fig. 3

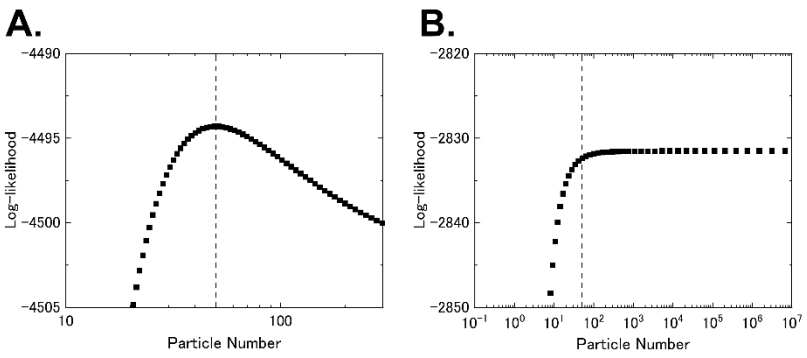

FIGURE S3.3: Shape of likelihood function

Typical results obtained by simulation in Fig. 3. Shape of the likelihood function when the ML estimate is (A) not on the boundary and (B) on the boundary. The dashed lines show the true value of particle number. The images were simulated with a particle brightness of (A) 20 kHz and (B) 5 kHz.

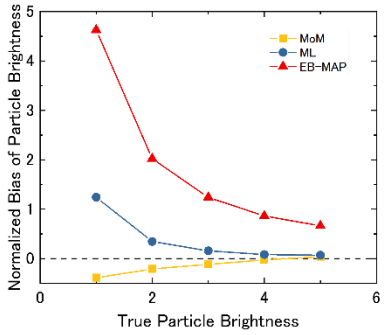

FIGURE S3.4: Normalized Bias of particle brightness

One thousand images were simulated with a fixed particle number of 100 and different particle brightness, and analyzed by MoM, ML, and EB–MAP. The size of the simulated images is 34×34, and the bias was estimated by excluding the edge of the image (32×32 pixels).

Supplemental data for Fig. 3

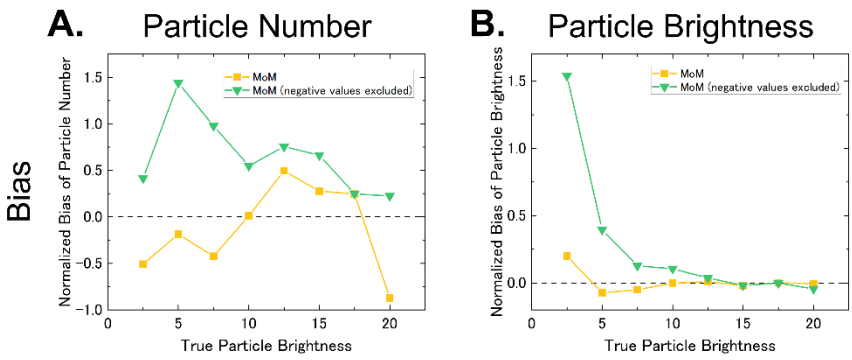

FIGURE S3.5: Influence of negative values on normalized bias

Normalized bias was estimated by excluding negative values in the results of Fig. 3. (A, B) Normalized bias for true particle number of 100 particles (A) and particle brightness (B).

Supplemental data for Fig. 4

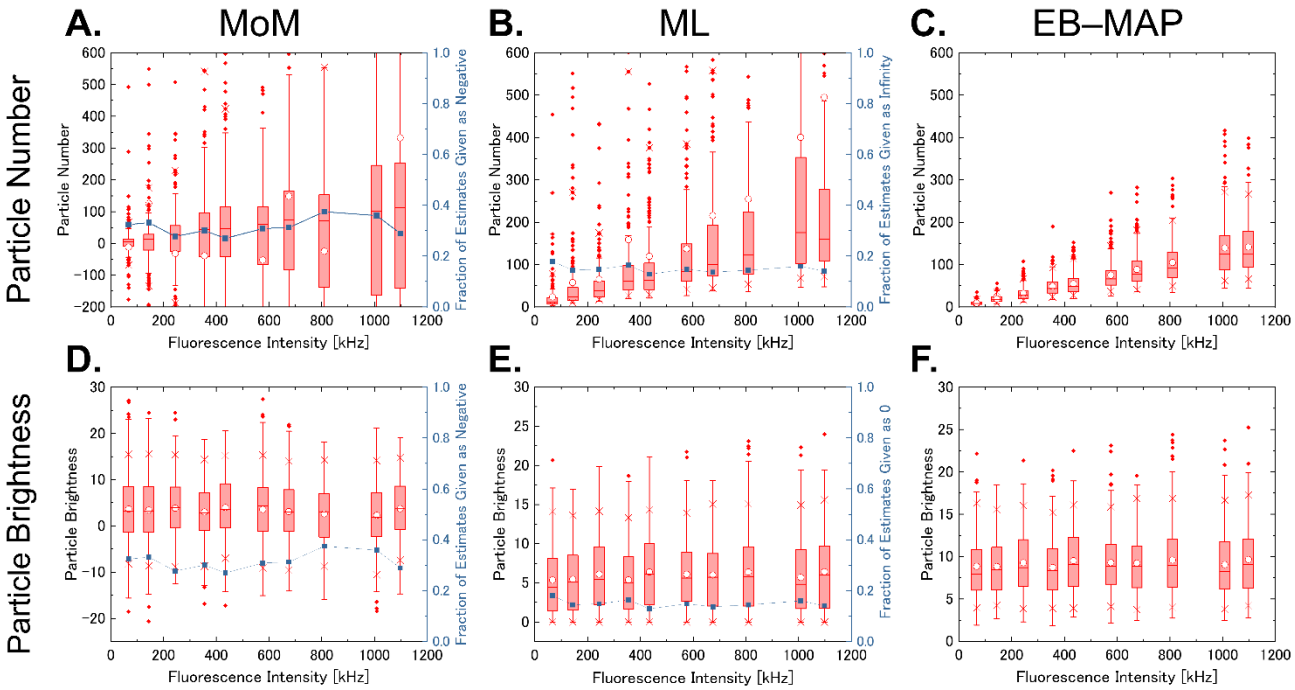

FIGURE S4.1: Variation in estimates for different concentrations

Two thousand images were obtained with different concentrations, and analyzed by MoM (A, D), ML (B, E), and EB–MAP (C, F). (A, B, C) Dependence of the estimated particle number on fluorescence intensity [kHz]. Figure S4.1A and S4.1B are shown in enlarged view (see Fig. S4.2 for broad view). (D, E, F) Dependence of the estimated particle brightness [kHz] on fluorescence intensity [kHz]. All figures are shown similarly in Fig. 2.

Supplemental data for Fig. 4

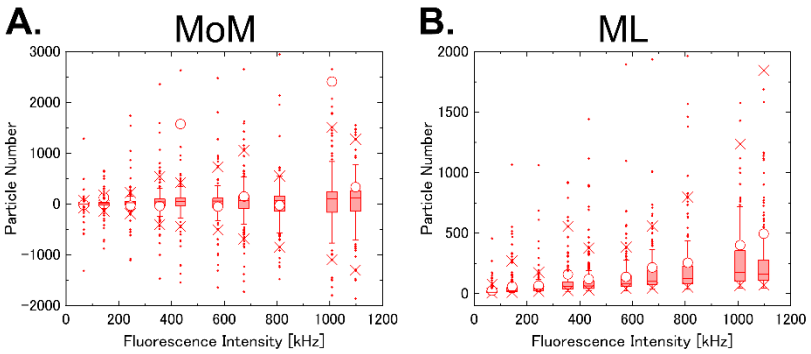

FIGURE S4.2: Broad view of estimated particle number in Fig. S4.1A and S4.1B

The view is adjusted to show the 5th and 95th percentiles of the particle number. The minimum and maximum values are listed in Table S4.1 and S4.2. (A) Broad view of Fig. S4.1A. (B) Broad view of Fig. S4.1B.

Fluorescence Intensity [kHz]	Minimum	Maximum
1.098×10^{3}	-5.550×10^{3}	3.751×10^{4}
1.009×10^{3}	-9.695×10^{3}	5.607×10^{5}
8.100×10^{2}	-1.030×10^{4}	7.708×10^{3}
6.750×10^{2}	-2.347×10^{3}	1.297×10^{4}
5.762×10^{2}	-4.709×10^{4}	1.948×10^{4}
4.344×10^{2}	-1.030×10^{4}	3.871×10^{5}
3.554×10^{2}	-1.882×10^{4}	8.033×10^{3}
2.442×10^{2}	-1.589×10^{4}	1.738×10^{3}
1.438×10^{2}	-3.185×10^{3}	3.144×10^{4}
6.725×10	-2.501×10^{3}	1.287×10^{3}

TABLE S4.1: Minimum and maximum values of the estimated particle number in Fig. S4.1A

Fluorescence Intensity [kHz]	Minimum	Maximum
1.098×10^{3}	4.760×10	9.735×10^{3}
1.009×10^{3}	4.654×10	8.153×10^{3}
8.100×10^{2}	3.524×10	7.532×10^{3}
6.750×10^{2}	3.879×10	5.165×10^{3}
5.762×10^{2}	2.668×10	1.895×10^{3}
4.344×10^{2}	2.104×10	1.442×10^{3}
3.554×10^{2}	1.942×10	3.386×10^{3}
2.442×10^{2}	1.278×10	1.061×10^{3}
1.438×10^{2}	8.527	1.066×10^{3}
6.725×10	3.227	4.544×10^{2}

TABLE S4.2: Minimum and maximum values of the estimated particle number in Fig. S4.1B

Supplemental data for Fig. 4

	Slope	Intercept
MoM	$(1.24\pm0.74)\times10^{-1}$	-21±16
ML	$(2.73\pm0.27)\times10^{-1}$	5.9±5.7
EB-MAP	$(1.307\pm0.016)\times10^{-1}$	0.41±0.30

TABLE S4.3: Slope and intercept for the fitted line in Fig. 4A

Estimated value and standard error

	Slope	Intercept
MoM	$(-6.9\pm5.0)\times10^{-4}$	3.71±0.32
ML	$(7.0\pm3.6)\times10^{-4}$	5.54±0.22
EB-MAP	$(5.8\pm2.7)\times10^{-4}$	8.89±0.17

TABLE S4.4: Slope and intercept for the fitted line in Fig. 4B

Estimated value and standard error

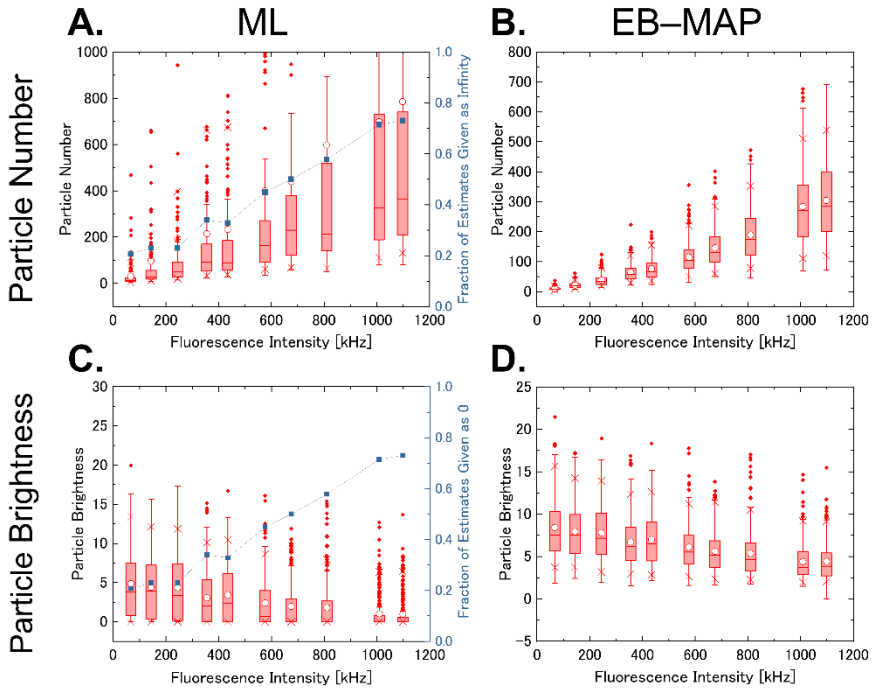

FIGURE S4.3: Variation in estimates without correction for dead time

The images for analysis are the same as those in Fig. 4 and Fig. S4.1, and were analyzed by ML (A, C) and EB–MAP (B, D) without correction for dead time. (A, B) Dependence of the estimated particle number on fluorescence intensity. (C, D) Dependence of the estimated particle brightness [kHz] on fluorescence intensity. All figures are shown similarly in Fig. 2.

Supplemental data for Fig. 4

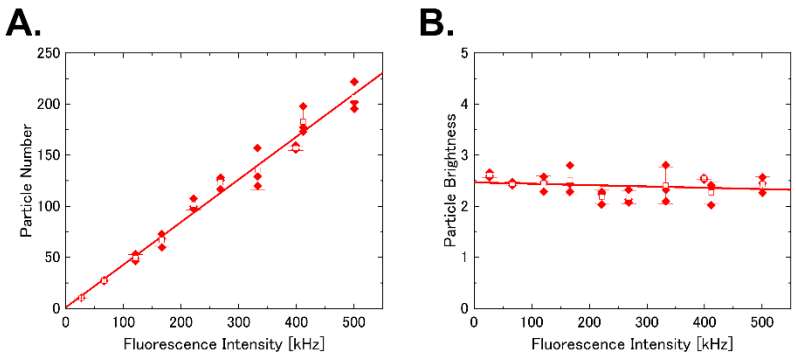

FIGURE S4.4: Estimation by FCS with different concentration

FCS measurements on EGFP solution with different concentrations. (A) Dependence of the estimated particle number on fluorescence intensity. (B) Dependence of the estimated particle brightness [kHz] on fluorescence intensity. In the plot, diamonds: results obtained from a single measurement, open squares: averages, and error bars: standard deviations. The solid lines are fitted lines obtained by the least squares method (see Table S4.5 for details on parameters of the lines).

	Slope	Intercept
Particle Number	0.417±0.013	1.2±3.7
Particle Brightness	$(-2.6\pm2.5)\times10^{-4}$	2.466±0.074

TABLE S4.5: Slope and intercept for the fitted line in Fig. S4.4A and S4.4B

Estimated value and standard error

Supplemental data for Fig. 5

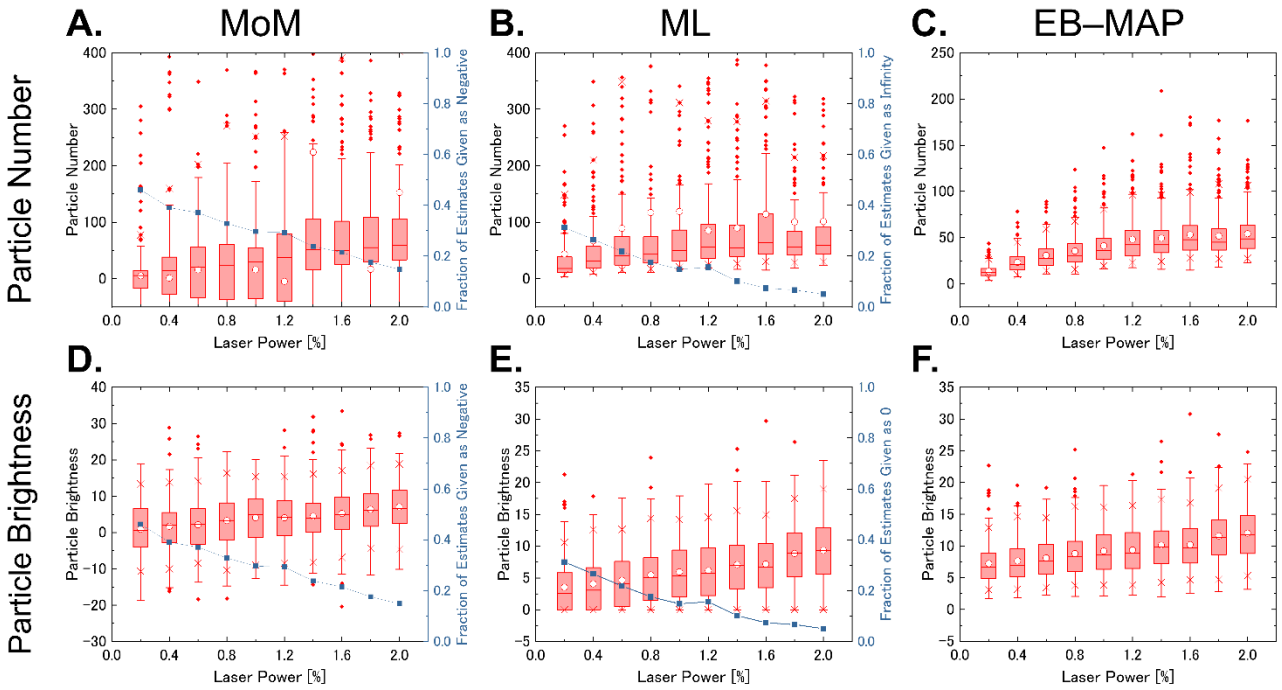

FIGURE S5.1: Variation in estimates for different laser powers

Two thousand images were obtained with different laser powers, and analyzed by MoM (A, D), ML (B, E), and EB–MAP (C, F). (A, B, C) Dependence of the estimated particle number on laser power. Figure S5.1A and S5.1B are shown in enlarged view (see Fig. S5.2 for broad view). (D, E, F) Dependence of the estimated particle brightness [kHz] on laser power. All figures are shown similarly in Fig. 2.

Supplemental data for Fig. 5

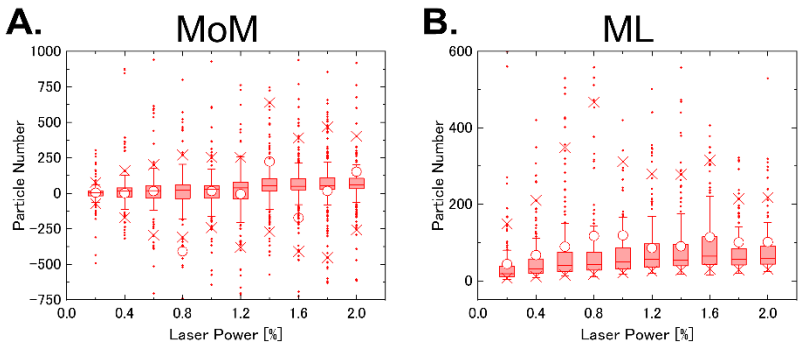

FIGURE S5.2: Broad view of estimated particle number in Fig. S5.1A and S5.1B

The view is adjusted to show the 5th and 95th percentiles of the particle number. The minimum and maximum values are listed in Table S5.1 and S5.2. (A) Broad view of Fig. S5.1A. (B) Broad view of Fig. S5.1B.

Laser Power [%]	Minimum	Maximum
0.2	-4.907×10^{2}	1.452×10^{3}
0.4	-2.592×10^{3}	2.023×10^{3}
0.6	-3.197×10^{3}	2.639×10^{3}
0.8	-1.090×10^{5}	1.831×10^{3}
1.0	-3.583×10^{3}	4.386×10^{3}
1.2	-6.125×10^{3}	4.618×10^{3}
1.4	-4.432×10^{3}	2.019×10^{4}
1.6	-4.189×10^{4}	5.072×10^{3}
1.8	-4.242×10^{3}	1.320×10^{3}
2.0	-1.317×10^{4}	3.506×10^{4}

TABLE S5.1: Minimum and maximum values of the estimated particle number in Fig. S5.1A

Laser Power [%]	Minimum	Maximum
0.2	3.666	6.474×10^{2}
0.4	8.691	1.153×10^{3}
0.6	1.180×10	1.303×10^{3}
0.8	1.082×10	2.768×10^{3}
1.0	1.749×10	3.255×10^{3}
1.2	1.884×10	5.013×10^{2}
1.4	1.657×10	9.492×10^{2}
1.6	1.538×10	2.017×10^{3}
1.8	1.920×10	2.701×10^{3}
2.0	2.423×10	3.165×10^{3}

TABLE S5.2: Minimum and maximum values of the estimated particle number in Fig. S5.1B

Supplemental data for Fig. 5

	Slope	Intercept
MoM	3.25 ± 0.17	0.37 ± 0.22
ML	3.11 ± 0.17	2.80 ± 0.19
EB-MAP	2.55 ± 0.14	6.62 ± 0.17

TABLE S5.3: Slope and intercept for the fitted line in Fig. 5B

Estimated value and standard error

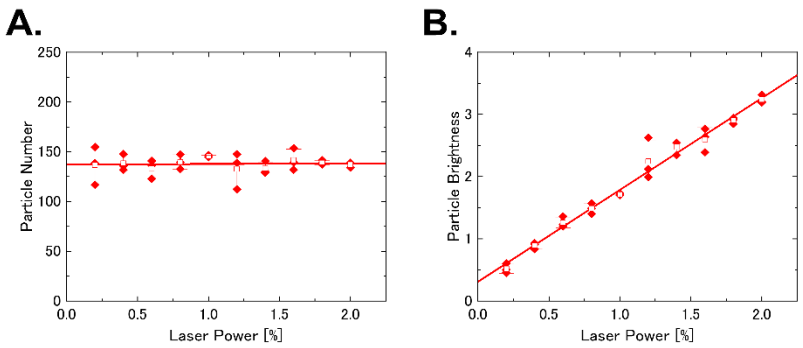

FIGURE S5.3: FCS measurements on EGFP solution with different laser powers

(A) Dependence of the estimated particle number on laser power. (B) Dependence of the estimated particle brightness [kHz] on laser power. In the plot, diamonds: results obtained from a single measurement, open squares: averages, and error bars: standard deviations. The solid lines are fitted lines obtained by the least squares method (see Table S5.4 for details on parameters of the lines).

	Slope	Intercept
Particle Number	0.6±3.1	137.2±3.8
Particle Brightness	1.477±0.046	0.310±0.057

TABLE S5.4: Slope and intercept for the fitted line in Fig. S5.3A and S5.3B

Estimated value and standard error

Supplemental data for Fig. 5

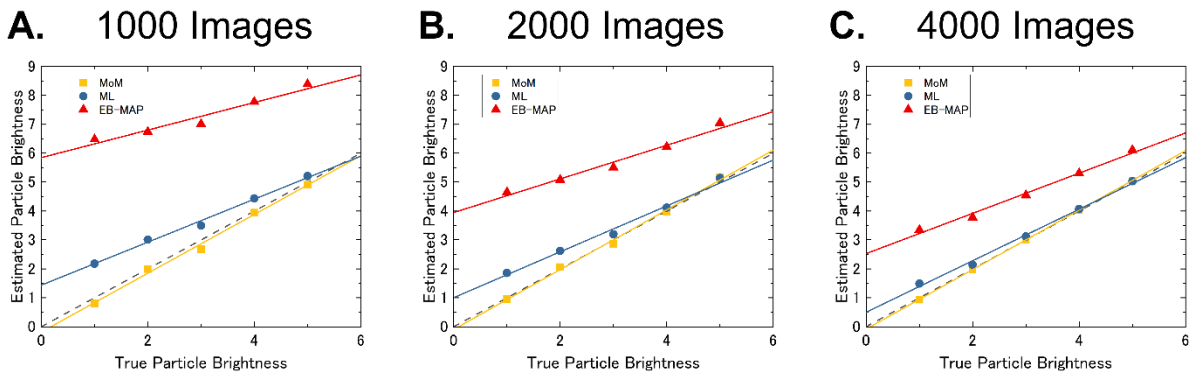

FIGURE S5.4: Simulations with different numbers of images

1000 (A), 2000 (B), and 3000 images (C) of size 34×34 were simulated with a fixed particle number of 10 and different particle brightness. (A, B, C) Dependence of the estimated particle brightness on true particle brightness. The symbols show the average of particle brightness. The dashed lines represent the true value. The solid lines are fitted lines obtained by the weighted least squares method (see Table S5.5 for details on parameters of the lines).

		Slope	Intercept
S5.4A	MoM	1.016±0.043	-0.18±0.14
	ML	0.744±0.036	1.44±0.11
	EB–MAP	0.477±0.065	5.85±0.21
S5.4B	MoM	1.034±0.038	-0.09±0.13
	ML	0.793±0.048	1.00±0.14
	EB–MAP	0.582±0.056	3.95±0.17
S5.4C	MoM	1.0292±0.0048	-0.084±0.016
	ML	0.891±0.038	0.50±0.11
	EB–MAP	0.696±0.044	2.53±0.13

TABLE S5.5: Slope and intercept for the fitted line in Fig. S5.4

Estimated value and standard error

Supplemental data for Fig. 6

Fig. S6.1 and S6.2 show the particle number and particle brightness in the experiments with EGFP tandem-oligomer in a living cell. The fluorescence intensity in the nucleus is lower than that in the cytoplasm for the EGFP dimer and trimer. This could have been caused by the higher molecular weight of the EGFP dimer and trimer than the monomer. The higher molecular weight species would be difficult to pass through nuclear pores. At the boundary between the nucleus and cytoplasm, the estimation of the particle number was markedly lower, and that of particle brightness was markedly higher than those in the nucleus and cytoplasm; this is remarkable in the EGFP trimer (Fig. S6.2D and S6.2F). This occurs due to the non-stationary change in fluorescence intensity by the movement of the nuclear membrane within the field of imaging. In actuality, the total measurement time is approximately 8 min to obtain 2000 images, and the cells migrate during the entire time period. We have previously reported that non-stationary changes cause artifacts, underestimation of particle number, and overestimation of particle brightness (6). Thus, we carefully placed the ROIs by avoiding the areas on which the pixel values would be biased by the artifact. Performing corrections for the non-stationary changes would improve the estimations.

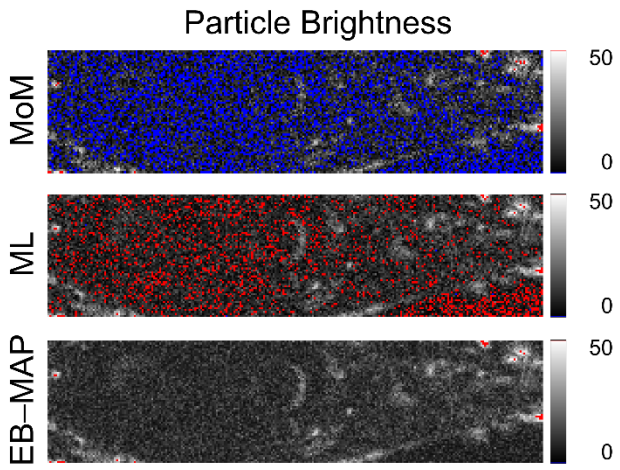

FIGURE S6.1: Particle brightness images of EGFP monomer

Estimated images of particle brightness [kHz]. Fluorescence intensity and particle number are shown in Fig. 6. The pixels below the range are colored in blue, and those beyond the range are colored in red. (Please view the digital version of this figure because the printed version could appear different.)

Supplemental data for Fig. 6

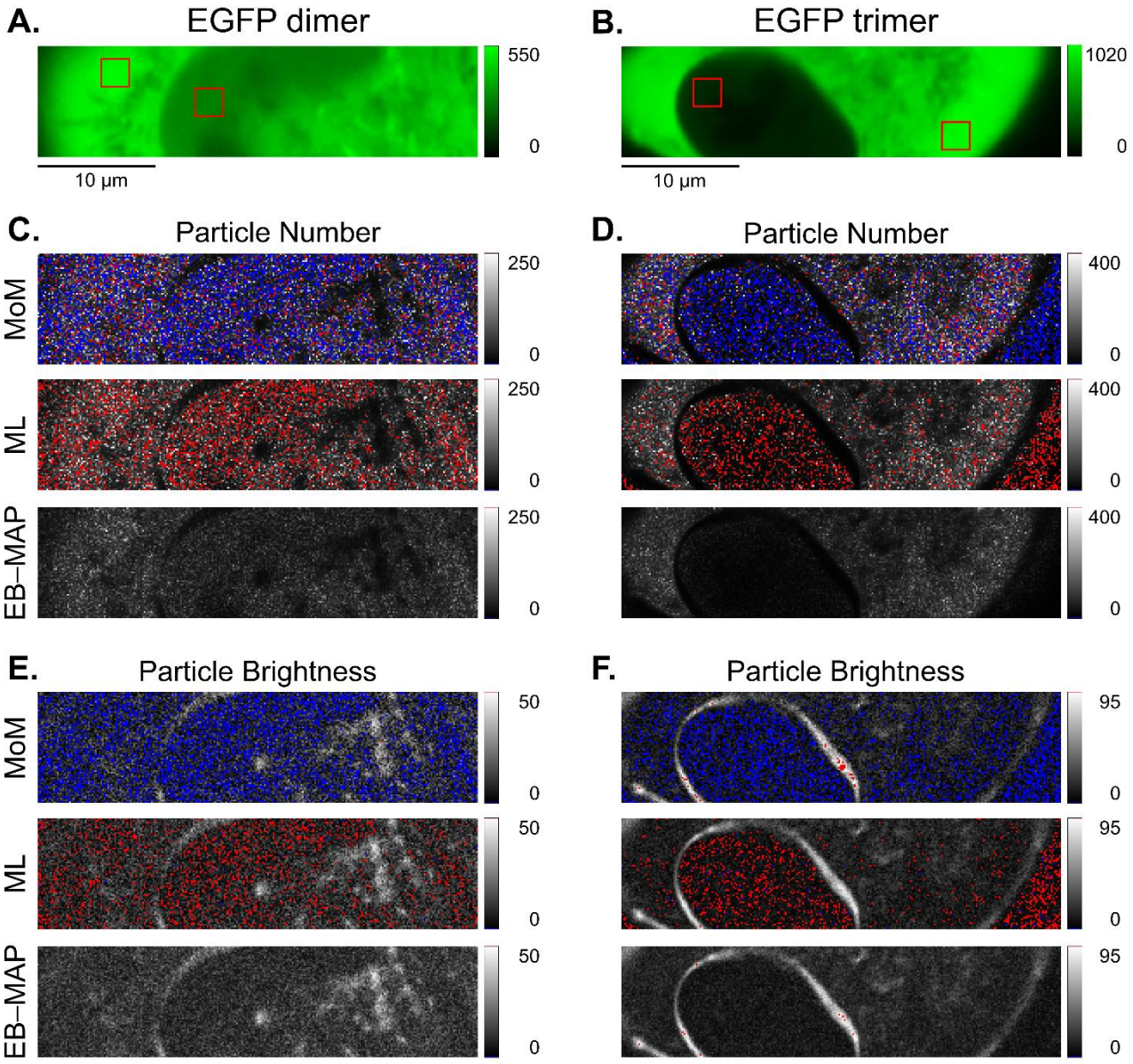

FIGURE S6.2: Estimated images for EGFP dimer and trimer

Experiments on EGFP dimer and trimer. Images of fluorescence intensity [kHz] for EGFP dimer (A) and trimer (B). Red squares represent the ROIs placed in the nucleus and cytoplasm. Estimated images of particle number for dimer (C) and trimer (D), and that of particle brightness [kHz] for dimer (E) and trimer (F). The pixels below the range are colored in blue, and those beyond the range are colored in red. (Please view the digital version of this figure because the printed version could appear different.)

Supplemental data for Fig. 6

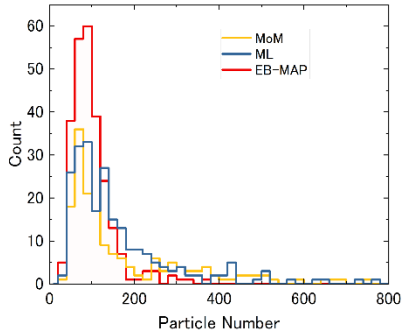

FIGURE S6.3: Histogram of particle number

The histogram was created from the particle number for EGFP monomer in the cytoplasm (ROI is shown in Fig. 6A).

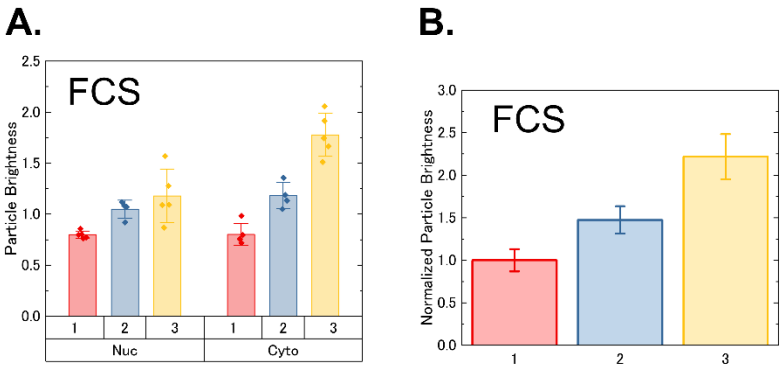

FIGURE S6.4: FCS measurements on EGFP tandem oligomers

(A) Dependence of the particle brightness [kHz] on the number of EGFP units in the nucleus (Nuc) and cytoplasm (Cyto). Dots: particle brightness, bar: sample average of particle brightness, and error bar: sample standard deviation of particle brightness. (B) Normalized particle brightness with respect to the particle brightness of the monomer in the cytoplasm. Bar: fold change in particle brightness (sample average of normalized particle brightness), error bar: sample standard deviation of normalized particle brightness.

Supplemental data for Fig. 6C–E and Fig. S6.4A

Verification of linearity in particle brightness of EGFP tandem oligomers

In experiments with EGFP tandem oligomers, particle brightness is expected to proportionally increase (with intercept of 0) for the monomer, dimer, and trimer. Even if a non-fluorescent state of EGFP appears, the particle brightness linearly increases (with intercept of any value). In Fig. S6.4A, the average particle brightness seems to increase linearly in the cytoplasm for FCS. However, in Fig. 6C–E, the average of the average particle brightnesses does not seem to increase linearly in the cytoplasm for MoM, ML, and EB–MAP. One could consider that particle brightness of the trimer is positively biased.

We assume that the particle brightness for the monomer and dimer was estimated accurately, but that for the trimer was overestimated. To evaluate the uncertainty, we estimated credible intervals for the difference between actual and predicted particle brightness of trimer. The particle brightness is predicted by assuming linearity between particle brightness of the monomer and dimer. The aim of this evaluation is not to specify the cause resulting in the non-linear change of particle brightness in our experiments, but to provide a method to evaluate the uncertainty of results in readers' experiments. In a strict sense, we cannot specify the cause that resulted in the non-linear change of particle brightness because our experiment is conducted on three different oligomeric states. Other combinations of assumptions are possible: The particle brightness for the dimer and trimer were estimated accurately, but the particle brightness for the monomer was overestimated; the particle brightness for the monomer and trimer were estimated accurately, but the particle brightness for the dimer was underestimated. Naturally, the linearity of the particle brightness between the monomer, dimer, and trimer is lost if two of the three particle brightness are not estimated accurately. The following analyses do not determine whether a scientific assumption is true. In the interpretation of the following results, we referred to papers (12, 13).

Statistical Model

First, we predict the particle brightness of the trimer by assuming linearity between the particle brightness of the monomer and dimer. Let X be the EGFP subunit number for the monomer, dimer, and trimer ($X = 1, 2, 3$) and Y be the particle brightness. Note that variables used in this section are different from those used in the main text. We assume that particle brightness plotted as dots in Figs. S6.4A and 6C–E follow an independent normal distribution with average μ and standard deviation σ, and we assume that the average of the normal distribution is equal to $aX + b$ for the monomer and dimer, where a and b are the regression coefficient and intercept, respectively. In addition, we assume that the prior distribution of a and b follow a normal distribution, and the prior distribution of σ follows a half-Cauchy distribution. The joint probability distribution is given by

$$P(\boldsymbol{Y}_{\mathrm{m}}, \sigma_{\mathrm{m}}, \boldsymbol{Y}_{\mathrm{d}}, \sigma_{\mathrm{d}}, a, b) = P(\boldsymbol{Y}_{\mathrm{m}}|a, b, \sigma_{\mathrm{m}})P(\boldsymbol{Y}_{\mathrm{d}}|a, b, \sigma_{\mathrm{d}})P(a, b, \sigma_{\mathrm{m}}, \sigma_{\mathrm{d}}), \tag{S28}$$

where

$$P(\boldsymbol{Y}_{\mathrm{m}}|a,b,\sigma_{\mathrm{m}}) = \prod_{i=1}^{I_{\mathrm{m}}} \mathrm{Nor}\left(Y_{\mathrm{m}}^{(i)}|a+b,\sigma_{\mathrm{m}}\right), \tag{S29}$$

$$P(\boldsymbol{Y}_{\mathrm{d}}|a,b,\sigma_{\mathrm{d}}) = \prod_{i=1}^{I_{\mathrm{d}}} \mathrm{Nor}\left(Y_{\mathrm{d}}^{(i)}|2a+b,\sigma_{\mathrm{d}}\right), \tag{S30}$$

$$P(a,b,\sigma_{\mathrm{m}},\sigma_{\mathrm{d}}) = \mathrm{Nor}(a|\mu_1,\sigma_1)\mathrm{Nor}(b|\mu_2,\sigma_2)\mathrm{Cau}(\sigma_{\mathrm{m}}|\varphi_1,\psi_1)\mathrm{Cau}(\sigma_{\mathrm{d}}|\varphi_2,\psi_2). \tag{S31}$$

The subscripts m and d represent variables for the monomer and dimer, respectively. The superscript (i) represents the ordinal number of observations. For I_{m} times observation, $\boldsymbol{Y}_{\mathrm{m}} = \left(Y_{\mathrm{m}}^{(1)}, Y_{\mathrm{m}}^{(2)}, \dots, Y_{\mathrm{m}}^{(I_{\mathrm{m}})}\right)$. Nor and Cau are normal and half-Cauchy distribution, respectively:

$$\mathrm{Nor}(Y|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(Y-\mu)^2}{2\sigma^2}\right), \tag{S32}$$

$$\mathrm{Cau}(\sigma|\varphi,\psi) = \frac{2}{\pi\psi}\left[\frac{\psi^2}{\psi^2+(\sigma-\varphi)^2}\right]. \tag{S33}$$

The hyperparameters are fixed as follows: $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 10^3$, $\varphi_1 = \varphi_2 = 0$, and $\psi_1 = \psi_2 = 25$. We believe that prior distributions with the hyperparameters would not be informative. In experiments with EGFP in the cells, the plausible values of particle brightness are in the range from 0 to 20 kHz, and a particle brightness of 40 kHz has a less likely yield than 20 kHz. For the hyperparameters, the prior distributions for $a+b$ and $2a+b$ are given by $\mathrm{Nor}\left(a+b|0,\sqrt{2}\times10^3\right)$ and $\mathrm{Nor}\left(2a+b|0,\sqrt{5}\times10^3\right)$, respectively. $\mathrm{Nor}\left(a+b|0,\sqrt{2}\times10^3\right)$ is flat in the range from 0 to 100 kHz, while the probability density is 2.82×10^{-4} at $a+b=0$ and 2.81×10^{-4} at $a+b=100$. $\mathrm{Nor}\left(2a+b|0,\sqrt{5}\times10^3\right)$ is flat in the range from 0 to 100 kHz. A plausible value of the standard deviation for particle brightness is less than 10 kHz. Particle brightness is a universal parameter under the same experimental conditions, where the change of particle brightness in different cells is small. $\mathrm{Cau}(\sigma_{\mathrm{m}}|0,25)$ and $\mathrm{Cau}(\sigma_{\mathrm{d}}|0,25)$ are weakly-informative (14–16). The posterior distribution of predicted particle brightness for the trimer is $\mu_{\mathrm{t}}^{\mathrm{pred}} = 3a+b$ by the posterior distribution for a and b, where the subscript t represents variables for the trimer. Next, we estimate the average particle brightness of the trimer. We assume that the prior distribution of μ_{t} and σ_{t} follow normal and half-Cauchy distributions, respectively. The joint probability distribution is given as follows, with $\boldsymbol{Y}_{\mathrm{t}} = \left(Y_{\mathrm{t}}^{(1)}, Y_{\mathrm{t}}^{(2)}, \dots, Y_{\mathrm{t}}^{(I_{\mathrm{t}})}\right)$:

$$P(\boldsymbol{Y}_{\mathrm{t}},\mu_{\mathrm{t}},\sigma_{\mathrm{t}}) = \prod_{i=1}^{I_{\mathrm{t}}}\left[\mathrm{Nor}\left(Y_{\mathrm{t}}^{(i)}|\mu_{\mathrm{t}},\sigma_{\mathrm{t}}\right)\right]\mathrm{Nor}(\mu_{\mathrm{t}}|\mu_3,\sigma_3)\mathrm{Cau}(\sigma_{\mathrm{t}}|\varphi_3,\psi_3). \tag{S34}$$

The hyperparameters are fixed similarly as follows: $\mu_3 = 0$, $\sigma_3 = 10^3$, $\varphi_3 = 0$, and $\psi_3 = 25$. $\mathrm{Nor}(\mu_\mathrm{t}|0, 10^3)$ is flat in the range from 0 to 100 kHz; the probability density is 3.99×10^{-4} at $\mu_\mathrm{t} = 0$ and 3.97×10^{-4} at $\mu_\mathrm{t} = 100$. The difference between the actual and predicted particle brightness of the trimer is given by the posterior distribution of μ_t and $\mu_\mathrm{t}^{\mathrm{pred}}$ as follows:

$$\delta_\mu = \mu_\mathrm{t} - \mu_\mathrm{t}^{\mathrm{pred}}. \tag{S35}$$

Results

We evaluated the difference between the actual and predicted particle brightness of the trimer, δ_μ, using different methods: FCS, MoM, ML, and EB–MAP. We used Markov chain Monte Carlo (MCMC) to simulate random numbers drawn from the posterior distribution of δ_μ. The random numbers are a better approximation of the posterior distribution at the convergence to stationary distribution. We assessed the convergence using $\hat{n}_{\mathrm{eff}}/(mn)$ and $\hat{R}$ (17), where $\hat{n}_{\mathrm{eff}}$ is the effective number of the simulated draws, m is the number of multiple simulated sequence, n is the actual number of the simulated draws in each simulated sequence, and $\hat{R}$ is the potential scale reduction factor. Fig. S7.1 shows the posterior distribution of δ_μ, and Table S7.1 shows the posterior mean and percentiles, assessment factors for convergence, and the probability for differences less and greater than 0. In Table S7.1, $\hat{n}_{\mathrm{eff}}/(mn)$ for all methods ranged from 0.86 to 0.99, indicating minimal autocorrelation between the simulated draws, and $\hat{R}$ for all methods ranged from 0.9999 to 1.0000, indicating the convergence of the simulated draws. In Fig. S7.1 and Table S7.1, FCS results suggest that the actual particle brightness of the trimer is 0.22 kHz higher than the predicted particle brightness on average. On the contrary, the 95% credible interval suggests that the difference in the particle brightness ranging from −0.58 to 1.02 kHz is compatible with our data, given our assumptions. From the credible interval, we infer the following: The difference of −0.58 indicates that the actual particle brightness of the trimer is 0.58 kHz lower than the predicted line, the difference of 0 indicates that the actual particle brightness of the trimer is on the predicted line, and the difference of 1.02 indicates that the actual particle brightness of the trimer is 1.02 kHz higher than the predicted line. In Table S7.1, the posterior probability for δ_μ less and greater than 0 are 0.22 and 0.78, respectively, indicating that the probability for δ_μ greater than 0 is high compared with the probability for less than 0. In Fig. S7.1 and Table S7.1, MoM, ML, and EB–MAP showed a wider range of 95% credible interval, and the posterior mean was higher than FCS. The posterior probability for $\delta_\mu > 0$ increased to 0.95, 0.92, and 0.92 for MoM, ML, and EB–MAP, respectively. To show how suitable our statistical model for characterizing observed samples are, we calculated the posterior predictive p-value (17) (Chapter 6, Model checking, p. 141), using sample mean and standard deviation as test quantities. In Table S7.2, the posterior predictive p-value of the sample average for all methods ranged from 0.49 to 0.50, indicating the sample averages are similar for replicated and observed samples. In Table S7.3, the posterior predictive p-value of sample standard deviations ranged from 0.61 to 0.64, indicating the sample standard deviations of replicated samples are occasionally higher than that of the observed samples.

Supplemental data for Fig. 6C–E and Fig. S6.4A

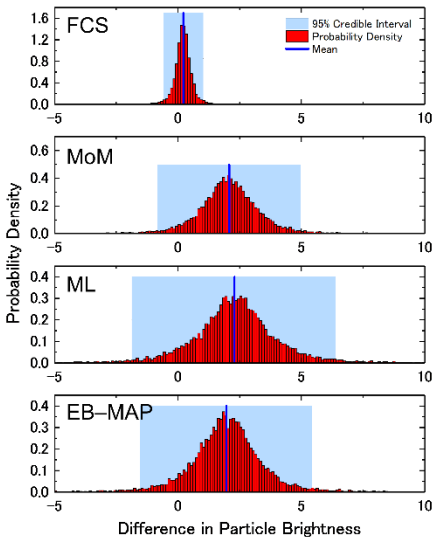

FIGURE S7.1: Uncertainty in linearity of particle brightness for EGFP trimer

Posterior distribution for the difference between actual and predicted particle brightness [kHz] of the trimer in the cytoplasm. The light blue shaded area represents 95% credible interval (interval between 2.5th and 97.5th percentiles), red histogram the posterior probability density, and solid blue line the posterior mean.

	Posterior Mean and Percentiles				Convergence Assessment		Posterior Probability	
	Mean and SE	Median	$Q_{2.5}$	$Q_{97.5}$	$\hat{n}_{\mathrm{eff}}/(mn)$	$\hat{R}$	$P(\delta_\mu < 0)$	$P(\delta_\mu > 0)$
FCS	0.2181 ± 0.0024	0.2177	-0.5772	1.0243	0.8646	0.9999	0.217	0.783
MoM	2.0824 ± 0.0076	2.0780	-0.8228	4.9718	0.9027	1.0000	0.054	0.946
ML	2.2971 ± 0.0108	2.3213	-1.8680	6.3851	0.9880	0.9999	0.082	0.918
EB–MAP	1.9731 ± 0.0089	1.9744	-1.5325	5.4337	0.9798	1.0000	0.080	0.920

TABLE S7.1: Summary of posterior distribution in Fig. S7.1

Supplemental data for Fig. 6C–E and Fig. S6.4A

		Test Quantity	Convergence Assessment	
		Sample Mean	$\hat{n}_{\text{eff}}/(mn)$	$\hat{R}$
FCS	Y_{m}	0.5026 ± 0.0022	1.0143	0.9999
	Y_{d}	0.4983 ± 0.0022	0.9997	1.0000
	Y_{t}	0.4995 ± 0.0023	0.9420	1.0000
MoM	Y_{m}	0.5000 ± 0.0022	1.0061	1.0000
	Y_{d}	0.4942 ± 0.0022	0.9928	1.0000
	Y_{t}	0.4976 ± 0.0022	0.9882	1.0000
ML	Y_{m}	0.5000 ± 0.0022	1.0048	1.0000
	Y_{d}	0.5012 ± 0.0023	0.9739	1.0000
	Y_{t}	0.5012 ± 0.0022	0.9934	1.0001
EB–MAP	Y_{m}	0.4980 ± 0.0023	0.9850	1.0001
	Y_{d}	0.4993 ± 0.0022	1.0092	1.0000
	Y_{t}	0.4995 ± 0.0022	0.9880	1.0000

TABLE S7.2: Summary of posterior predictive p-values estimated by sample mean for Fig. S7.1

		Test Quantity	Convergence Assessment	
		Sample Standard Deviation	$\hat{n}_{\text{eff}}/(mn)$	$\hat{R}$
FCS	Y_{m}	0.6138 ± 0.0022	1.0042	1.0000
	Y_{d}	0.6448 ± 0.0022	0.9885	0.9999
	Y_{t}	0.6194 ± 0.0022	1.0150	1.0000
MoM	Y_{m}	0.6168 ± 0.0022	0.9987	1.0000
	Y_{d}	0.6443 ± 0.0021	1.0058	1.0000
	Y_{t}	0.6157 ± 0.0022	0.9705	1.0000
ML	Y_{m}	0.6214 ± 0.0022	0.9789	1.0000
	Y_{d}	0.6437 ± 0.0021	0.9996	1.0000
	Y_{t}	0.6194 ± 0.0022	0.9614	1.0001
EB–MAP	Y_{m}	0.6131 ± 0.0022	1.0072	1.0001
	Y_{d}	0.6443 ± 0.0022	0.9556	0.9999
	Y_{t}	0.6160 ± 0.0022	0.9891	1.0000

TABLE S7.3: Summary of posterior predictive p-values estimated by sample standard deviation for Fig. S7.1

Discussions and Comments

We predicted the particle brightness of the EGFP trimer by assuming a linearity between the particle brightness of the monomer and dimer. The 95% credible intervals in Fig. S7.1 indicate a wide range of uncertainty; thus, we cannot make strong implications, such as whether the actual particle brightness of the trimer is higher than the predicted line or whether the actual particle brightness is on the predicted line, for all methods. We need to increase the sample sizes to distinguish whether the actual particle brightness is on the predicted line. The 95% credible interval in FCS has an important sense practically; the difference of −0.58 and 1.02 kHz between the actual and predicted particle brightness is not negligible, compared with the sample average of 1.78 kHz for the particle brightness of the trimer (Fig. S6.4A). The uncertainty of the difference between the particle brightnesses are also not negligible for MoM, ML, and EB–MAP. According to statistical hypothesis testing with a significance level of 0.05, if the 95% credible interval does not contain a difference of 0 between the actual and predicted particle brightness, it is possible to mechanically make the decision that the actual and predicted particle brightness is significantly different. Note that we obtained these results using small sample sizes and our assumptions, and we should not place too much confidence on the exact numerical values of these results. This is because our statistical model is not perfectly true. Although we constructed our statistical model using normal and half-Cauchy distributions, we have no evidence that the model is true. However, we can show the suitableness of our statistical model for characterizing observed samples through the posterior predictive p-value. If the statistical model is suitable to characterize the observed samples, replicated samples produced by posterior predictive distribution and observed samples would be similar. The results in Table S7.2 suggest that our statistical model is suitable to estimate the mean, but the results in Table S7.3 suggest that our statistical model tends to overestimate the standard deviation. However, we believe that the use of our statistical model for the observed samples do not cause serious problems. The posterior predictive p-value of approximately 0.6 indicates overestimated standard deviation in 60% of replicated samples among the multiple sets of replicated samples. We believe that the deviation between 0.5 and 0.6 is not significantly large.

Methods

Markov chain Monte Carlo (MCMC)

Posterior draws were simulated using R, version 4.0.3 (18), and Stan, version 2.21.2 (19). Five sequences of draws are simulated in parallel to assess convergence of the simulated draws. The total number of simulated draws for each sequence is 305,000. The first 5,000 draws were discarded, and the remainder of the sequence was thinned by keeping every 30th draw, resulting in 10,000 draws for each sequence.

Supplemental data for Fig. 6F and Fig. S6.4B

Credible interval of fold change in particle brightness

In experiments with EGFP tandem oligomers, Fig. 6F shows the normalized particle brightness, where the average of normalized particle brightness is the fold change. We estimated the credible interval of the fold change for FCS, MoM, ML, and EB–MAP. The credible interval is useful to compare results obtained by different methods and verify the compatibility of the results. Using the same notation from the previous section, we assume that the joint probability distributions are given by

$$P(\boldsymbol{Y}_{\mathrm{m}}, \mu_{\mathrm{m}}, \sigma_{\mathrm{m}}) = \prod_{i=1}^{I_{\mathrm{m}}} \left[\mathrm{Nor}\left(Y_{\mathrm{m}}^{(i)} | \mu_{\mathrm{m}}, \sigma_{\mathrm{m}}\right)\right] \mathrm{Nor}(\mu_{\mathrm{m}} | \mu_4, \sigma_4) \mathrm{Cau}(\sigma_{\mathrm{m}} | \varphi_4, \psi_4), \tag{S36}$$

$$P(\boldsymbol{Y}_{\mathrm{d}}, \mu_{\mathrm{d}}, \sigma_{\mathrm{d}}) = \prod_{i=1}^{I_{\mathrm{d}}} \left[\mathrm{Nor}\left(Y_{\mathrm{d}}^{(i)} | \mu_{\mathrm{d}}, \sigma_{\mathrm{d}}\right)\right] \mathrm{Nor}(\mu_{\mathrm{d}} | \mu_5, \sigma_5) \mathrm{Cau}(\sigma_{\mathrm{d}} | \varphi_5, \psi_5), \tag{S37}$$

$$P(\boldsymbol{Y}_{\mathrm{t}}, \mu_{\mathrm{t}}, \sigma_{\mathrm{t}}) = \prod_{i=1}^{I_{\mathrm{t}}} \left[\mathrm{Nor}\left(Y_{\mathrm{t}}^{(i)} | \mu_{\mathrm{t}}, \sigma_{\mathrm{t}}\right)\right] \mathrm{Nor}(\mu_{\mathrm{t}} | \mu_6, \sigma_6) \mathrm{Cau}(\sigma_{\mathrm{t}} | \varphi_6, \psi_6). \tag{S38}$$

We fixed the hyperparameters as follows: $\mu_4 = \mu_5 = \mu_6 = 0$, $\sigma_4 = \sigma_5 = \sigma_6 = 10^3$, $\varphi_4 = \varphi_5 = \varphi_6 = 0$, and $\psi_4 = \psi_5 = \psi_6 = 25$. The fold changes in the particle brightness for the EGFP dimer and trimer are given as follows using the posterior distributions of μ_{m}, μ_{d}, and μ_{t}:

$$r_{\mathrm{d}} = \frac{\mu_{\mathrm{d}}}{\mu_{\mathrm{m}}}, \tag{S39}$$

$$r_{\mathrm{t}} = \frac{\mu_{\mathrm{t}}}{\mu_{\mathrm{m}}}. \tag{S40}$$

Results

Fig. S8.1 shows the posterior distributions of r_{d} and r_{t} for FCS, MoM, ML (+), and EB–MAP (+). We used MCMC to simulate the draws from the posterior distribution in the same way as the previous section. Table S8.1 summarizes the posterior mean, posterior percentiles, and the assessment factors for convergence. $\hat{n}_{\mathrm{eff}}/(mn)$ for all methods ranged from 0.94 to 1.00, indicating minimal autocorrelation between the simulated draws, and $\hat{R}$ for all methods ranged from 1.0000 to 1.0001, indicating the convergence of the simulated draws. In Fig. S8.1 and Table S8.1, FCS result for the dimer shows that the 95% credible interval ranged from 0.98—indicating that the particle brightness of the dimer is equal to that of the monomer—to 2.12—indicating that the particle brightness of the dimer is approximately two times higher than that of the monomer. The 95% credible intervals of the fold change for the dimer in each method are within the range of 0.6 to 2.2 and overlapped each other, indicating that

the results are compatible with each other. Similarly, the 95% credible intervals of the fold change for the trimer are compatible with each other. Tables S8.2 and S8.3 summarize the posterior predictive p-value calculated by the sample mean and standard deviation, respectively. In Table S8.2, the posterior predictive p-value of the sample averages for all methods ranged from 0.49 to 0.50, indicating that they were similar between replicated and observed samples. In Table S8.3, the posterior predictive p-value of sample standard deviations ranged from 0.62 to 0.65, indicating the sample standard deviations of the replicated samples are occasionally higher than that of the observed samples.

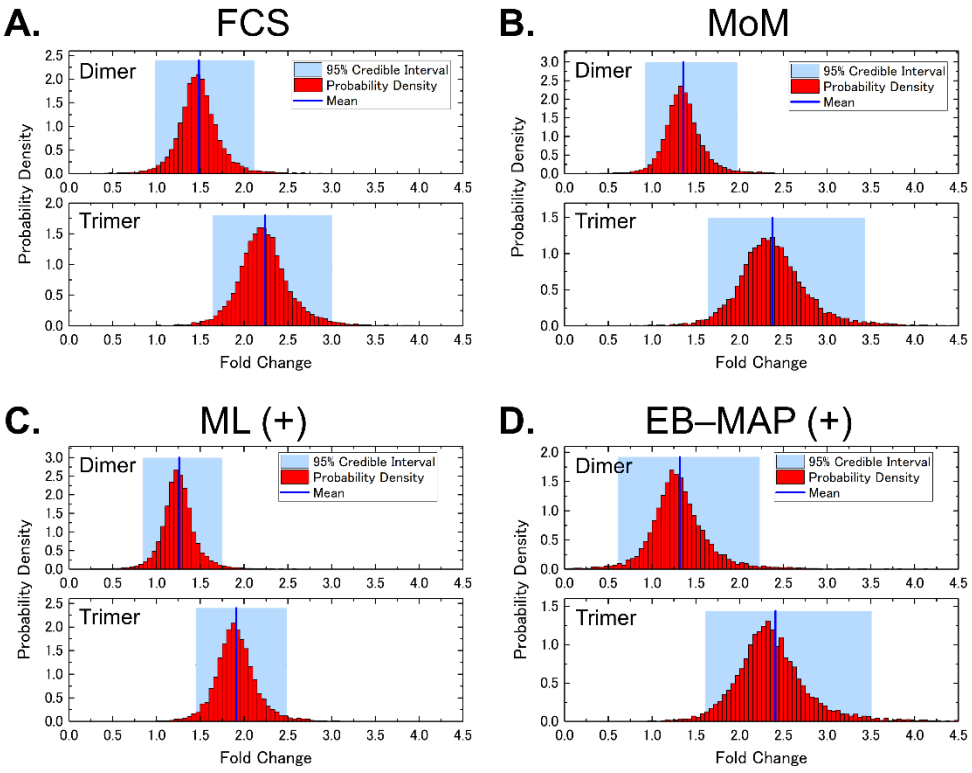

FIGURE S8.1: Uncertainty in fold change of particle brightness for EGFP dimer and trimer with respect to monomer

Posterior distribution for fold change in particle brightness in cytoplasm. The light blue shaded area represents 95% credible interval: interval between 2.5th and 97.5th percentiles, red histogram: posterior probability density, and blue solid line: posterior mean.

Discussions and Comments

In Fig. S8.1, the 95% credible intervals for FCS, MoM, ML (+), and EB–MAP (+) showed a wide range of uncertainty that was non-negligible. In Fig. 6F, particle brightness is divided by the sample average of particle brightness for monomer, which is the normalized particle brightness, and thus, the uncertainty in the population average of the particle brightness for the monomer is neglected. The credible interval is more useful than the standard deviation of the normalized particle brightness to estimate uncertainty. However, while useful to translate the uncertainty into numerical values, credible intervals must be interpreted with care. This is because the statistical

model is not true, similar to the previous section. Tables S8.2 and S8.3 summarize the similarity between the observed and replicated samples. The statistical model introduced in this section is not the only model to estimate uncertainty, and other models can be used.

	Tandem Oligomer	Posterior Mean and Percentiles				Convergence Assessment	
		Mean and SE	Median	$Q_{2.5}$	$Q_{97.5}$	$\hat{n}_{\text{eff}}/(mn)$	$\hat{R}$
FCS	Dimer	1.4854 ± 0.0033	1.4728	0.9779	2.1153	0.9421	1.0000
	Trimer	2.2404 ± 0.0045	2.2164	1.6432	3.0051	0.9997	1.0000
MoM	Dimer	1.3544 ± 0.0118	1.3431	0.9200	1.9735	1.0005	1.0001
	Trimer	2.3773 ± 0.0200	2.3538	1.6383	3.4320	0.9979	1.0000
ML (+)	Dimer	1.2584 ± 0.0025	1.2497	0.8410	1.7530	0.9903	1.0001
	Trimer	1.9120 ± 0.0033	1.8971	1.4467	2.4853	0.9960	1.0000
EB–MAP (+)	Dimer	1.3184 ± 0.0106	1.2832	0.6109	2.2314	0.9934	1.0000
	Trimer	2.4101 ± 0.0277	2.3326	1.6109	3.5076	0.9989	1.0000

TABLE S8.1: Summary of posterior distribution in Fig. S8.1

		Test Quantity	Convergence Assessment	
		Sample Mean	$\hat{n}_{\text{eff}}/(mn)$	$\hat{R}$
FCS	Y_{m}	0.5027 ± 0.0022	0.9924	1.0000
	Y_{d}	0.4972 ± 0.0022	0.9907	1.0001
	Y_{t}	0.4974 ± 0.0022	0.9951	1.0000
MoM	Y_{m}	0.4971 ± 0.0022	0.9939	1.0000
	Y_{d}	0.5042 ± 0.0022	1.0091	1.0000
	Y_{t}	0.4992 ± 0.0022	1.0161	1.0000
ML (+)	Y_{m}	0.5001 ± 0.0022	0.9926	1.0000
	Y_{d}	0.4975 ± 0.0023	0.9812	1.0000
	Y_{t}	0.5018 ± 0.0023	0.9832	1.0000
EB–MAP (+)	Y_{m}	0.4993 ± 0.0022	0.9947	1.0001
	Y_{d}	0.5006 ± 0.0022	1.0016	0.9999
	Y_{t}	0.5027 ± 0.0022	1.0197	1.0000

TABLE S8.2: Summary of posterior predictive p-values estimated by sample mean for Fig. S8.1

Supplemental data for Fig. 6F and Fig. S6.4B

		Test Quantity	Convergence Assessment	
		Sample Standard Deviation	$\hat{n}_{\text{eff}}/(mn)$	$\hat{R}$
FCS	Y_{m}	0.6183 ± 0.0022	0.9908	1.0000
	Y_{d}	0.6471 ± 0.0021	0.9903	0.9999
	Y_{t}	0.6194 ± 0.0022	0.9863	1.0000
MoM	Y_{m}	0.6185 ± 0.0022	0.9683	0.9999
	Y_{d}	0.6451 ± 0.0022	0.9947	1.0000
	Y_{t}	0.6181 ± 0.0022	0.9723	0.9999
ML (+)	Y_{m}	0.6224 ± 0.0022	1.0032	0.9999
	Y_{d}	0.6481 ± 0.0021	0.9873	1.0000
	Y_{t}	0.6160 ± 0.0022	0.9907	1.0000
EB–MAP (+)	Y_{m}	0.6219 ± 0.0022	0.9770	1.0000
	Y_{d}	0.6480 ± 0.0021	1.0018	1.0000
	Y_{t}	0.6178 ± 0.0022	0.9986	1.0000

TABLE S8.3: Summary of posterior predictive p-values estimated by sample standard deviation for Fig. S8.1

SUPPORTING REFERENCES

1. Bédard, G. 1967. Dead-time corrections to the statistical distribution of photoelectrons. *Proc. Phys. Soc.* 90.
2. Ackermann, J., and H. Hogreve. 2010. Small dead-time expansion in counting distributions and moments. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* 614:297–302.
3. Hillesheim, L.N., and J.D. Müller. 2003. The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy with Non-Ideal Photodetectors. *Biophys. J.* 85:1948–1958.
4. Hillesheim, L.N., and J.D. Müller. 2005. The dual-color photon counting histogram with non-ideal photodetectors. *Biophys. J.* 89:3491–3507.
5. Bishop, C.M. 2006. Pattern Recognition and Machine Learning. Springer, Germany.
6. Fukushima, R., J. Yamamoto, H. Ishikawa, and M. Kinjo. 2018. Two-detector number and brightness analysis reveals spatio-temporal oligomerization of proteins in living cells. *Methods*. 140–141:161–171.
7. Oura, M., J. Yamamoto, H. Ishikawa, S. Mikuni, R. Fukushima, and M. Kinjo. 2016. Polarization-dependent fluorescence correlation spectroscopy for studying structural properties of proteins in living cell. *Sci. Rep.* 6:31091.
8. Pack, C., K. Saito, M. Tamura, and M. Kinjo. 2006. Microenvironment and Effect of Energy Depletion in the Nucleus Analyzed by Mobility of Multiple Oligomeric EGFPs. *Biophys. J.* 91:3921–3936.
9. Takahasi, H., and M. Mori. 1973. Double exponential formulas for numerical integration. *Publ. Res. Inst. Math. Sci.* 9:721–741.
10. Müller, C.B., A. Loman, V. Pacheco, F. Koberling, D. Willbold, W. Richtering, and J. Enderlein. 2008. Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. *Europhys. Lett.* 83:46001.
11. Culbertson, C.T., S.C. Jacobson, and J. Michael Ramsey. 2002. Diffusion coefficient measurements in microfluidic devices. *Talanta*. 56:365–73.
12. Amrhein, V., S. Greenland, and B. McShane. 2019. Scientists rise up against statistical significance. *Nature*. 567:305–307.
13. Johnson, D.H. 1999. The Insignificance of Statistical Significance Testing. *J. Wildl. Manage.* 63:763.
14. Lemoine, N.P. 2019. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos*. 128:912–928.
15. Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* 1:515–534.
16. Polson, N.G., and J.G. Scott. 2012. On the Half-Cauchy Prior for a Global Scale Parameter. *Bayesian Anal.* 7:887–902.
17. Gelman, A., J. Carlin, H. Stern, and D. Rubin. 2004. Bayesian data analysis. 3rd ed. Chapman & Hall/CRC, UK.
18. R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.
19. Stan Development Team. 2020. RStan: the R interface to Stan. http://mc-stan.org/.