1S Bin specification and normalization for PAID histograms (cf. section 2.1)

When forming a PAID histogram for the large number of photons obtained in fluorescence fluctuation experiments, it is more meaningful to place events in bins than to make a scatter plot. In choosing the size and spacing of bins, one needs to consider that fluorescence fluctuations occur over a large range of time scales. To cover a large range of time scales with a minimum number of histogram bins, log or quasi-log time bins are commonly used in FCS. For the PAID histogram, we choose the bins for the time interval τ axis to be log-spaced, with 10 bins per decade. To use a log scale for the monitor photon axis is more problematic (especially at low photon counts) since the number of monitor photons that arrive is strictly an integer. The clock time resolution Δt can be chosen to be small enough to make the integer nature of the discrete time interval variable $\boldsymbol{\tau} = \lfloor \tau/\Delta t \rfloor$ negligible in the μs regime, but this cannot be done with the number of monitor photons counted. Unless one is willing to use a spacing of bins that is extremely sparse (powers of $2,3,4\ldots$), the discrete spacing of the number of monitor photons will cause log bins to be inconsistently occupied at low $\mathbf{n}$; some bins may not even have an integer in them. So, we use a quasi-log scale adapted from the multiple tau correlation technique [1,2]. The first 16 bins are evenly spaced with increments of 1, $\left(\mathbf{n}_1,\ldots,\mathbf{n}_{16} \right) = \left(0,1,\ldots,15 \right)$; then, with each set of 8 bins, the increment is doubled. The next 8 bins are $\left(\mathbf{n}_{17},\ldots,\mathbf{n}_{24} \right) = \left(16-17,18-19,\ldots,30-31 \right)$ with an increment of 2, followed by $\left(\mathbf{n}_{25},\ldots,\mathbf{n}_{32} \right) = \left(32-35,36-39,\ldots,60-63 \right)$ with an increment of 4, etc. On the large scale these bins are log-spaced, while on the small scale they are linearly spaced. In this way, we can cover a large dynamic range of integers in a consistent manner with a small number of bins. After placing events in the histogram bins, normalization is necessary to obtain $C_{STM}\left(\boldsymbol{\tau}, \mathbf{n} \right)$. First, the histogram is multiplied by $\mathcal{T}/N_S N_T$ in eq 9 in the main text. Second, for a bin that has time interval axis limits $\boldsymbol{\tau}_{low}$ and $\boldsymbol{\tau}_{high}$ and monitor photon axis limits that include the integers $\mathbf{n}_{low}$ through $\mathbf{n}_{high}$, we divide by the size of the bin $\left(\boldsymbol{\tau}_{high} - \boldsymbol{\tau}_{low} \right)\left(\mathbf{n}_{high} - \mathbf{n}_{low} + 1 \right)$. The value for the bin is an average of $C_{STM}\left(\boldsymbol{\tau}, \mathbf{n} \right)$ over the bin limits, rather than an integral over the bin limits. This normalizes the

1

histogram, giving us $C_{STM}(\tau, \mathbf{n})$. Unfortunately, this is a non-ideal representation for a log scale: when plotting a slice of the histogram in the log scale of the monitor photon count axis, for a constant time interval, we want the actual area under the curve to correspond to the value of the correlation $C_{ST}(\tau)$. To do this, we approximate the photon monitor variable $\mathbf{n}$ as a continuous variable n, then convert to a log scale using the expression $\zeta = \log_{10} n$. We want to keep the relation in eq 10 in the main text valid in the new variable. We approximate the sums over $\mathbf{n}$ as integrals over a continuous variable n, and convert to the log variable ζ:

$$
\begin{aligned}
C_{ST}(\tau) &= \sum_{\mathbf{n}=0}^{\infty} C_{STM}(\tau, \mathbf{n}) \\
&= \int_{0}^{\infty} dn\, C_{STM}(\tau, [n]) \\
&= \int_{-\infty}^{\infty} d\zeta\, (n \ln 10)\, C_{STM}(\tau, [n]) \\
&\equiv \int_{-\infty}^{\infty} d\zeta\, C'_{STM}(\tau, \zeta)
\end{aligned}
\tag{1S}
$$

By looking at the differentials, we find that by changing variables from n to ζ we change amplitude of the PAID by a factor $(\ln 10)n$. So, the histogram bins are in addition weighted by a factor of $(\mathbf{n}+0.5)\ln 10$. We add 0.5 because we consider each bin in n as covering a range between $\mathbf{n}$ and $\mathbf{n}+1$, and the average over this range is $\mathbf{n}+0.5$. This only makes a difference at low n, and causes the $n=0$ bin to be weighted by the factor 0.5 rather than 0. We emphasize that this rescaling is made to ease the visual interpretation of the histogram, and in no way affects the statistical characteristics of the histogram.

2S Algorithms for constructing PAID histograms (cf. section 2.1)

A direct algorithm to form PAID histograms from photon streams compares each photon from the start channel with each photon in the stop channel, then calculates the number of monitor photons that fall between each start and stop photon. If there are N_S photons in the start channel, and N_T photons in the stop channel, there are $\sim N_S N_T$ entries into the histogram, making the algorithm $O(N^2)$, and limiting the dynamic range of the PAID histogram. A more efficient algorithm exploits the log spacing of the time interval bins $\boldsymbol{\tau}_a$ and the quasi-log spacing of the monitor photon count bins $\mathbf{n}_b$ (Fig. 1S-A; filled square placed at the integer-valued time of arrival of each detected photon, $\mathbf{t}_i^S$; time interval and monitor photon count bins corresponding to the first start photon are denoted by the arrow at the top). The time interval bins, log-spaced with integer time intervals 2,4,8,16,… are shown below the S photon stream. The monitor photon count bins, log-spaced with monitor photon counts 1,2,4,… are shown below the M photon stream. The time interval and monitor photon count bins, combined to form the 2D-histogram bins, are shown below S. The number of stop photon inside each bin is shown. These values are transferred to the 2D-histogram in Fig. 1S-B. The algorithm: (1) consider each start photon arrival time $\mathbf{t}_i^S$, and search for photons in T and M that are closest to this time; (2) set the current time interval (τ) bin to $a_{curr} = 1$, and the current monitor photon count $(\mathbf{n})$ bin to $b_{curr} = 1$; (3) calculate the time interval τ_M at which M switches to the next monitor count bin $b_{curr} + 1$; (4) if τ_M is less than the time interval of the $(a_{curr} + 1)^{th}$ τ bin, perform a binary search on T to find the photon arriving just after τ_M, to determine how many stop photons arrive in the current bin and add them to the (a_{curr}, b_{curr}) bin of the histogram; advance $b_{curr} = b_{curr} + 1$, and go to step 3; (5) otherwise, perform a binary search on T to find how many stop photons arrive up to the time of the $(a_{curr} + 1)^{th}$ τ bin, and add them to the (a_{curr}, b_{curr}) bin of the histogram and advance $a_{curr} = a_{curr} + 1$; go to step 3, till there are no more stop photons or (a_{curr}, b_{curr}) is outside of the histogram; (6) go to step 1 till there are no more start photons.
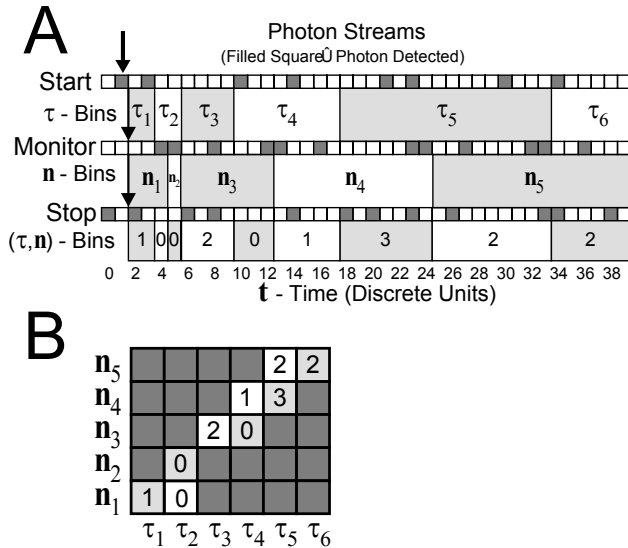
Figure 1S. Generation of 2D-PAID histograms with log time-interval axis and log or quasi-log monitor photon count axis. A, photons detected in the start, monitor, and stop channels (filled squares) in the time axis (shown with discrete clock units, figure 3). The contribution to the PAID histogram for the first start photon, marked with the top black arrow, is shown in B. The time interval bins are shown as gray and white bands between the start channel and the monitor channel. The bins are log-spaced, with 2,4,8,16,…clock units. The monitor photon count bins are shown as gray and white bands between the monitor channel and the stop channel. These bins are also log-spaced, with 1,2,4,8,… photon counts. The gray and white bands below the stop photon channel show how the time interval and monitor photon count bins combine to form the two-dimensional histogram bins. Each bin is labeled with the number of stop photons counted within the bin, subsequently transferred to the histogram in B.

This algorithm uses the fact that the start, stop, and monitor channels are ordered lists (each successive photon is at a later time) by performing binary searches. A modified search algorithm that uses increments of increasing size from the initial search index to bracket the desired value before performing a standard binary search was found to be most effective (see description of **hunt** in ref 3.) Also, because of the log spacing on both axes, a small number of binary searches can cover a large dynamic range. The algorithm is extendable to multiple monitor channels. On a 1.2 GHz Pentium 3

4

based PC, the algorithm shown here is able to form the PAID histogram of a 10 sec data set with 3,750,000 photons in 2 minutes and the PAID histogram of a 30 sec data set with 150,000 photons in 4.4 seconds, scaling nearly linearly with the number of photons in the data set.

3S PAID Function Model (cf. section 2.3)

We develop a model of the PAID function for several diffusing species with a Poisson background in a tightly focused laser excitation volume. The ensemble averages in the PAID function in eq 8 in the main text are simplified by averaging over the "shot noise" contribution (intrinsic to photon counting experiments). We then split the photon emission rates into independent contributions from each molecule and background (assuming that all species diffuse independently, with the time between association and dissociation much longer than the diffusion time). The PAID function for all species is expressed as a sum of convolutions of distributions for each species, which are approximated by Monte Carlo simulation of possible diffusion paths through V_{eff} [4]. These path integrals need only be approximated once; changes in the diffusion time and brightness parameters are accounted for by a scaling law. This means that, although the model is not expressed in closed analytical form, it can still be used practically in a fitting routine using a pre-calculated kernel.

In modeling the light emission and detection processes, we assume that the lifetime of each fluorophore is negligible compared to both the time scale under consideration (>1 μs) and the inverse of the excitation rate (far from singlet-state saturation). This assumption means that the photon emission statistics are determined by a Poisson process with a single fluctuating rate $k_A(t)$ for each channel A (fluctuations can be due to diffusion, intersystem crossing, etc.). Additionally, we assume that the fluorophores only undergo fluorescence transitions, ignoring photophysical processes such as intersystem crossing and photobleaching. With these assumptions, the photon emission rate $k_A(t)$ equals the laser excitation intensity multiplied by the absorption coefficient and the quantum efficiency. Using this rate, the probability for the number of photons counted up to time t follows the series of differential equations $\frac{\partial P_A(t,\mathbf{n})}{\partial t} = k_A(t)P_A(t,\mathbf{n}-1) - k_A(t)P_A(t,\mathbf{n})$; the probability to have $\mathbf{n}$ counted photons is increased by transitions from the $\mathbf{n}-1$ photons state, and decreased by transitions out of the $\mathbf{n}$ counted photons state. Solving these equations, we obtain the Poisson probability distribution[5,6],

$$P_A\left(t,\mathbf{n}|t_0\right)=\exp\left[-K_A\left(t_0,t\right)\right]\frac{\left[K_A\left(t_0,t\right)\right]^{\mathbf{n}}}{\mathbf{n}!}\equiv\mathrm{Poi}\left(K_A\left(t_0,t\right),\mathbf{n}\right) \qquad (2S)$$

where we define $K_A\left(t_0,t\right)=\int_{t_0}^{t}dt'\,k_A\left(t'\right)$ as the cumulative intensity. This expression only accounts for the "shot noise" intrinsic to the photon counting process; any random fluctuations in $k_A\left(t\right)$ widen this distribution further.

We now apply these expressions to the PAID function given in eq 8 in the main text , where the ensemble averages are performed first to account for the shot noise contribution, then for the rate fluctuation contribution $\langle\cdots\rangle=\left\langle\left\langle\cdots\right\rangle_{\mathrm{shot}}\right\rangle_{\mathrm{rate}}$. $k\left(t\right)$ and its integral $K_A\left(t_0,t\right)$ are functions of continuous time, and will not be written using discrete time variables (we will use $t=\mathbf{t}\Delta t$ and $\tau=\boldsymbol{\tau}\Delta t$). We assume that the stochastic processes are stationary so that ensemble averages and time averages are equivalent, and the variable $\mathbf{t}$ can then be dropped. For $\left\langle I_S\left(\mathbf{t}\right)\right\rangle$ in the denominator, stationarity implies $\left\langle I_S\left(\mathbf{t}\right)\right\rangle=\left\langle I_S\left(0\right)\right\rangle$; the shot noise contribution is found by averaging the number of photon counts $I_S\left(\mathbf{t}\right)$ against the probability distribution in eq 2S (the continuous time interval corresponding to the discrete time $\mathbf{t}$ is of duration Δt),

$$\left\langle I_S\left(\mathbf{t}\right)\right\rangle=\left\langle\sum_{I_S=0}^{\infty}I_S\mathrm{Poi}\left[K_S\left(0,\Delta t\right),I_S\right]\right\rangle_{\mathrm{rate}} \qquad (3S)$$
$$=\left\langle K_S\left(0,\Delta t\right)\right\rangle_{\mathrm{rate}}\approx\left\langle k_S\left(0\right)\Delta t\right\rangle_{\mathrm{rate}}$$

The final approximation assumes that Δt is much smaller than the fluctuation time scale of $k\left(t\right)$. Defining $\overline{k}_S$ to be the average $\overline{k}_S=\left\langle k_S\left(t\right)\right\rangle_{\mathrm{rate}}$, we can write $\left\langle I_S\left(\mathbf{t}\right)\right\rangle=\overline{k}_S\Delta t$ (similarly, $\left\langle I_T\left(\mathbf{t}\right)\right\rangle=\overline{k}_T\Delta t$). For the numerator in eq 8, the shot noise contribution is independent for different time windows, allowing the shot noise averages to be evaluated separately for each factor. We obtain $\left\langle I_S\left(\mathbf{t}\right)\right\rangle_{\mathrm{shot}}=k_S\left(0\right)\Delta t$ and $\left\langle I_T\left(\mathbf{t}+\boldsymbol{\tau}\right)\right\rangle_{\mathrm{shot}}=k_T\left(\tau\right)\Delta t$ for the first two factors. The third factor is evaluated as follows,

7

$$\left\langle \mathbf{\delta}\left[\mathbf{n},\mathbf{n}_M\left(\mathbf{t},\mathbf{t}+\boldsymbol{\tau}\right)\right]\right\rangle_{\text{shot}} = \sum_{\mathbf{n}_M=0}^{\infty} \mathbf{\delta}\left(\mathbf{n},\mathbf{n}_M\right)\text{Poi}\left[K_M\left(\Delta t,\tau\right),\mathbf{n}_M\right]$$

$$= \text{Poi}\left[K_M\left(\Delta t,\tau\right),\mathbf{n}\right] \tag{4S}$$

Equation 8 is rewritten in terms of rates,

$$C_{STM}\left(\boldsymbol{\tau},\mathbf{n}\right) = \left\langle k_S\left(0\right)k_T\left(\tau\right)\text{Poi}\left[K_M\left(\Delta t,\tau\right),\mathbf{n}\right]\right\rangle_{\text{rate}} / \overline{k}_S \overline{k}_T, \tag{5S}$$

We then perform ensemble averaging over the "rate fluctuations." To do this, each rate $k_A(t)$ is split into contributions from background and each of $\mathcal{N}_\alpha$ independently diffusing and emitting molecules from each species $\alpha = 1\ldots F$ in the sample volume $\mathcal{V}_{\text{sample}}$ (concentration is $[\alpha] = \mathcal{N}_\alpha / \mathcal{V}_{\text{sample}} = c_\alpha / V_{\text{eff}}$),

$$k_A\left(t\right) = \sum_{\alpha=0}^{F}\sum_{i=1}^{\mathcal{N}_\alpha} k_{\alpha iA}\left(t\right) \equiv \sum_{(\alpha,i)} k_{\alpha iA}\left(t\right) \tag{6S}$$

For the constant background $\left(\alpha = 0\right)$, we set $\mathcal{N}_0 = 1$ and $k_{01A}\left(t\right) = k_{0A}$. Since the probability distribution for a sum of independent random variables is the convolution of the probability distributions for each of the independent random variables, and $K_M\left(\Delta t,\tau\right) = \sum_{(\alpha,i)} K_{\alpha iM}\left(\Delta t,\tau\right)$ (a direct extension of eq 6S), eq 4S can be rewritten

$$\text{Poi}\left[K_M\left(\Delta t,\tau\right),\mathbf{n}\right] = \prod_{(\alpha,i)}\left(\text{Poi}\left[K_{\alpha iM}\left(\Delta t,\tau\right),\mathbf{n}\right]*\right). \tag{7S}$$

The star inside the parantheses indicates repeated convolution in the variable $\mathbf{n}$ rather than multiplication. For later convenience, we define $S_{\alpha iM}\left(\tau,\mathbf{n}\right) \equiv \text{Poi}\left[K_{\alpha iM}\left(\Delta t,\tau\right),\mathbf{n}\right]$ (S is for "shot noise"). Using the relation $\langle xy \rangle = \langle x \rangle \langle y \rangle$ which is true for any independent random variables x and y, along with eqs 5S, 6S, and 7S, and grouping direct multiplication terms separately from cross terms, we obtain

8

$$C_{STM}\left(\boldsymbol{\tau},\mathbf{n}\right)=\sum_{(\alpha,i)}\left\langle k_{\alpha iS}\left(0\right)k_{\alpha iT}\left(\tau\right)S_{\alpha iM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}$$

$$*\prod_{(\gamma,l)\neq(\alpha,i)}\left(\left\langle S_{\gamma lM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}*\right)\Big/\overline{k}_S\overline{k}_T+$$

$$\sum_{(\alpha,i)\neq(\beta,j)}\left\langle k_{\alpha iS}\left(0\right)S_{\alpha iM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}*\left\langle k_{\beta jT}\left(\tau\right)S_{\beta jM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate} \tag{8S}$$

$$*\prod_{(\gamma,l)\neq\{(\alpha,i);(\beta,j)\}}\left(\left\langle S_{\gamma lM}\left(\tau,\mathbf{n}\right)\right\rangle *\right)_{rate}\Big/\overline{k}_S\overline{k}_T$$

Equation 8S breaks $C_{STM}\left(\boldsymbol{\tau},\mathbf{n}\right)$ into combinations of four single-molecule expressions. The first is

$\left\langle k_{\alpha iS}\left(0\right)k_{\alpha iT}\left(\tau\right)S_{\alpha iM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}$, which is the numerator for the PAID function of the single molecule i of

species α. To evaluate this ensemble average, there are only three quantities, $k_{\alpha iS}\left(0\right)$, $k_{\alpha iT}\left(\tau\right)$, and

$K_{\alpha iM}\left(\Delta t,\tau\right)$, required for each set of rate functions $k_{\alpha iS}\left(t\right)$, $k_{\alpha iT}\left(t\right)$, and $k_{\alpha iM}\left(t\right)$. Therefore, the rate

ensemble averages can be written as integrals over a probability distribution $P_{\alpha STM}\left(k_S,k_T,K_M\,|\tau\right)$.

Since all molecules i of a species α share the same properties, this probability distribution is the same

for all i. Therefore, the index i is dropped.

$$\left\langle k_{\alpha iS}\left(0\right)k_{\alpha iT}\left(\tau\right)S_{\alpha iM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}=\int_0^\infty dk_S\int_0^\infty dk_T\int_0^\infty dK_M P_{\alpha STM}\left(k_S,k_T,K_M\,|\tau\right)k_S k_T \mathrm{Poi}\left(K_M,\mathbf{n}\right)$$
$$=\int_0^\infty dK_M f_{\alpha STM}^{(1,1)}\left(K_M\,|\tau\right)\mathrm{Poi}\left(K_M,\mathbf{n}\right) \tag{9S}$$

k_S, k_T, and K_M are integration variables that take the place of $k_{\alpha iS}\left(0\right)$, $k_{\alpha iT}\left(\tau\right)$, and

$K_{\alpha iM}\left(\Delta t,\tau\right)$ respectively. The final expression, $f_{\alpha STM}^{(1,1)}\left(K_M\,|\tau\right)=\int_0^\infty dk_S\int_0^\infty dk_T P_{\alpha STM}\left(k_S,k_T,K_M\,|\tau\right)k_S k_T$, is

obtained by integrating over the variables k_S and k_T; we retain the integration over K_M.

The remaining single molecule expressions in eq 8S are written as follows,

$$\left\langle k_{\alpha iS}\left(0\right)S_{\alpha iM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}=\int_0^\infty dK_M f_{\alpha STM}^{(1,0)}\left(K_M\,|\tau\right)\mathrm{Poi}\left(K_M,\mathbf{n}\right)$$
$$\left\langle k_{\beta jT}\left(\tau\right)S_{\beta jM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}=\int_0^\infty dK_M f_{\beta STM}^{(0,1)}\left(K_M\,|\tau\right)\mathrm{Poi}\left(K_M,\mathbf{n}\right) \tag{10S}$$
$$\left\langle S_{\gamma lM}\left(\tau,\mathbf{n}\right)\right\rangle_{rate}=\int_0^\infty dK_M f_{\gamma STM}^{(0,0)}\left(K_M\,|\tau\right)\mathrm{Poi}\left(K_M,\mathbf{n}\right)$$

where $f_{\alpha STM}^{(p,q)}\left(K_M\,|\tau\right)=\int_0^\infty dk_S\int_0^\infty dk_T P_{\alpha STM}\left(k_S,k_T,K_M\,|\tau\right)k_S^p k_T^q$. eq 8S can now be simplified to read

$$C_{STM}(\boldsymbol{\tau},\mathbf{n}) = \int_0^\infty dK_M \mathrm{Poi}(K_M,\mathbf{n})\left\{\sum_\alpha \mathcal{N}_\alpha f_{\alpha STM}^{(1,1)}(K_M|\tau) * \prod_\gamma \left[f_{\gamma STM}^{(0,0)}(K_M|\tau)*\right]^{\mathcal{N}_\gamma} + \right.$$

$$\left. \sum_{\alpha,\beta} \mathcal{N}_\alpha \mathcal{N}_\beta f_{\alpha STM}^{(1,0)}(K_M|\tau) * f_{\beta STM}^{(0,1)}(K_M|\tau) * \prod_\gamma \left[f_{\gamma STM}^{(0,0)}(K_M|\tau)*\right]^{\mathcal{N}_\gamma} \right\} \bigg/ \overline{k}_S \overline{k}_T \qquad (11S)$$

where we have used the approximation $\mathcal{N}_\alpha \gg 1$; $\mathcal{V}_{\text{sample}}$ and $\mathcal{N}_\alpha$ are taken to be very large, $\mathcal{N}_\alpha \gg 1$ and $\mathcal{V}_{\text{sample}} \gg V_{\text{eff}}$. Also, as a simplification, the integration over K is performed only once, after all of the convolutions have been performed in K space rather than $\mathbf{n}$ space (see section 3.1S for proof of equivalence). Note that the powers are multiple convolutions, not multiplications.

We subsequently compute the functions $f^{(p,q)}$. We use Monte Carlo sampling of diffusion paths through the detection volume (Gaussian or numerically computed volume) to estimate them. A diffusion path L is defined by its position as a function of time $\vec{x}(t|L)$. The probability for each diffusion path L can be found by dividing the path into a series of $\mathbf{X}$ positions evenly spaced in time; each position is separated by the time interval $\delta\tau$ (distinct from Δt), so that $\mathbf{X} = \tau/\delta\tau$ [7]. By taking the limit $\delta\tau \to 0$, we obtain a probability density for L suitable for path integration.

$$P_\alpha\left(L|\vec{x}(0),\vec{x}(\tau)\right) = \lim_{\delta\tau\to 0}\left(\prod_{\mathbf{h}=1}^{\mathbf{X}-1}\frac{d\vec{x}_{\mathbf{h}}}{(4\pi D_\alpha \delta\tau)^{3/2}}\right)\exp\left[\sum_{\mathbf{h}=0}^{\mathbf{X}-1} -\frac{(\vec{x}_{\mathbf{h}+1}-\vec{x}_{\mathbf{h}})^2}{4 D_\alpha \delta\tau}\right] \qquad (12S)$$

The exponential factor comes from the probability for a particle to diffuse from position $\vec{x}_{\mathbf{h}}$ to position $\vec{x}_{\mathbf{h}+1}$ after a time $\delta\tau$; the probability of each diffusion path depends on the diffusion constant D_α. For each path L, there is a corresponding rate for each channel ($A = S,T,M$), $k_{\alpha A}(\mathbf{h}\delta\tau|L) = q_{\alpha A}\phi(\vec{x}_{\mathbf{h}})$, where $q_{\alpha A}$ is the brightness and ϕ is the detectivity, both defined earlier. From these rates we calculate $k_{\alpha i S}(t)$, $k_{\alpha i T}(t)$, and $K_{\alpha i M}(t_0,t)$. The relationship between the diffusion time (where the correlation function decays by a factor of 2) and the diffusion constant $\tau_\alpha^D = \eta/D_\alpha$ (η is a constant of proportionality) depends on the detectivity, and needs to be determined for each detection volume.

To calculate $f^{(p,q)}$, we first find $P_{\beta STM}\left(k_S, k_T, K_M \mid \boldsymbol{\tau}\right)$. To calculate this probability, we average over the initial, final, and all intermediate positions, restricting the values of the following variables:

$k_S = q_{\alpha S}\phi\left(\vec{x}_0\right)$, $k_T = q_{\alpha T}\phi\left(\vec{x}_{\mathbf{X}}\right)$, and $K_M = \sum_{\mathbf{h}=1}^{\mathbf{X}-1} q_{\alpha M}\phi\left(\vec{x}_{\mathbf{h}}\right)\delta\tau$. We also need the probability distribution for initial positions, which we take as being equal throughout the sample volume, $P\left(\vec{x}_0\right) = dV_0/\mathcal{V}_{\text{sample}}$. $\mathcal{V}_{\text{sample}}$ is finite, but large enough so that converging integrals may be extended to infinity with negligible error. Integrating over all positions with the restrictions just mentioned, we obtain

$$
\begin{aligned}
P_{\beta STM}\left(k_S, k_T, K_M \mid \boldsymbol{\tau}\right) dk_S dk_T dK_M = \lim_{\delta\tau \to 0} & \left\{ \int \frac{dV_S}{\mathcal{V}_{\text{sample}}} \int dV_T\, \delta\left[k_S - q_{\alpha S}\phi\left(\vec{x}_0\right)\right]\delta\left[k_T - q_{\alpha T}\phi\left(\vec{x}_{\mathbf{X}}\right)\right] \right. \\
& \left. \left(\prod_{\mathbf{h}=1}^{\mathbf{X}-1}\int \frac{d\vec{x}_{\mathbf{h}}}{\left(4\pi D_\alpha \delta\tau\right)^{3/2}}\right)\exp\left[\sum_{\mathbf{h}=0}^{\mathbf{X}-1} -\frac{\left(\vec{x}_{\mathbf{h}+1}-\vec{x}_{\mathbf{h}}\right)^2}{4D_\alpha \delta\tau}\right]\delta\left[K_M - \sum_{\mathbf{h}=1}^{\mathbf{X}-1} q_{\alpha M}\phi\left(\vec{x}_{\mathbf{h}}\right)\delta\tau\right]\right\} dk_S dk_T dK_M
\end{aligned}
\tag{13S}
$$

Now, integrating over k_S and k_T to obtain $f_{\alpha STM}^{(p,q)}\left(K_M \mid \tau\right)$,

$$
\begin{aligned}
f_{\alpha STM}^{(p,q)}\left(K_M \mid \tau\right) dK_M = \lim_{\delta\tau \to 0} & \left\{ \int \frac{dV_S}{\mathcal{V}_{\text{sample}}} \int dV_T\, \left[q_{\alpha S}\phi\left(\vec{x}_0\right)\right]^p \left[q_{\alpha T}\phi\left(\vec{x}_{\mathbf{X}}\right)\right]^q \right. \\
& \left. \left(\prod_{\mathbf{h}=1}^{\mathbf{X}-1}\int \frac{d\vec{x}_{\mathbf{h}}}{\left(4\pi D_\alpha \delta\tau\right)^{3/2}}\right)\exp\left[\sum_{\mathbf{h}=0}^{\mathbf{X}-1} -\frac{\left(\vec{x}_{\mathbf{h}+1}-\vec{x}_{\mathbf{h}}\right)^2}{4D_\alpha \delta\tau}\right]\delta\left[K_M - \sum_{\mathbf{h}=1}^{\mathbf{X}-1} q_{\alpha M}\phi\left(\vec{x}_{\mathbf{h}}\right)\delta\tau\right]\right\} dK_M
\end{aligned}
\tag{14S}
$$

An extremely important property of $f_{\alpha STM}^{(p,q)}\left(K_M \mid \tau\right)$ is obtained if we change variables in the expression, now using $\delta\tau^0 = \delta\tau/\tau_\alpha^D$, $K_M^0 = K_M/q_{\alpha M}\tau_\alpha^D$ and $\tau^0 = \tau/\tau_\alpha^D$,

$$
f_{\alpha STM}^{(p,q)}\left(K_M \mid \tau\right) dK_M = q_{\alpha S}^p q_{\alpha T}^q f_{0STM}^{(p,q)}\left(K_M^0 \mid \tau^0\right) dK_M^0
\tag{15S}
$$

Once the function $f_{0STM}^{(p,q)}\left(K_M^0 \mid \tau^0\right)$ (where $q_{\alpha S}^0 = q_{\alpha T}^0 = q_{\alpha M}^0 = \tau_\alpha^{D0} = 1$) is known, the function for any other parameter values can be obtained by scaling the variables according to the definitions of τ^0 and K_M^0 and eq 15S, making the use of Monte Carlo simulation of diffusion paths practical for model calculations (the kernel simulation only needs to be performed once).

Using eq 14S without the limit, a series of diffusion paths is simulated, and the function $f_{0STM}^{(p,q)}\left(K_M^0\,\middle|\,\tau^0\right)$ is estimated by averaging over 10^4-10^7 diffusion paths. $\delta\tau$ is a small time window (100 times smaller than $\tau_\alpha^D = \eta/D_\alpha$) where the position of the molecule is assumed constant. There is a useful time symmetry in eq 14S: the reverse of each path is equally likely. This implies that $f_{0STM}^{(1,0)}\left(K_M^0\,\middle|\,\tau^0\right) = f_{0STM}^{(0,1)}\left(K_M^0\,\middle|\,\tau^0\right)$. For $f_{0STM}^{(1,1)}$, $f_{0STM}^{(1,0)}$, and $f_{0STM}^{(0,1)}$, the factors $q_{\alpha S}^0\phi(\vec{x}_0)$ and $q_{\alpha T}^0\phi(\vec{x}_X)$ act to emphasize paths that emit more strongly at the endpoints, when molecules are in the detection volume. Therefore, when estimating these functions with Monte Carlo simulation, the probability distribution for the starting points of the diffusion paths are chosen to match the detectivity $\phi(\vec{x})$. Because of this emphasis on strongly emitting paths, the sample volume is allowed to go to infinity, $\mathcal{V}_{\text{sample}} \to \infty$. This is not possible for the function $f_{0STM}^{(0,0)}$, since the factors $q_{\alpha S}^0\phi(\vec{x}_0)$ and $q_{\alpha T}^0\phi(\vec{x}_X)$ are not present in the expression; molecular paths with initial and final points anywhere in the sample volume contribute equally. To calculate the PAID function in eq 11S, it is necessary to perform many successive convolutions on $f_{0STM}^{(0,0)}$. In order to limit this number, while providing accurate results, a simulation volume $\mathcal{V}_{\text{sample}}$ is chosen that is 100 to 1000 times as large as the detection volume V_{eff}. To obtain an occupancy of c_α for species α, the number of convolutions $\mathcal{N}_\alpha = c_\alpha \mathcal{V}_{\text{sample}}/V_{\text{eff}}$ is used in eq 11S to simulate the presence of many molecules.

Depending on the accuracy desired, $\phi(\vec{x})$ is set as an analytical Gaussian detection volume, as a numerical approximation, or as an experimentally-measured detection volume, allowing the direct application of the expected detection volume to the model. In contrast to either FIDA or FIMDA (where for simplicity, only the volume density for a given brightness value is required), **our method directly models the possible diffusion paths through V_{eff}** [8,9]. To obtain expressions for the PAID function through combining single-molecule expressions, it is necessary to compute many convolutions, producing the primary computational task in the model calculation [similar to PCH [10] and FIDA[9]]. Because of the wide temporal and dynamic ranges over which fluorescence fluctuations occur, log axes

12

are used. For the most efficient calculation, the convolution method used must work in a log domain. However, pure Fast Fourier Transform (FFT) methods demand linearly-spaced data, which quickly produces huge arrays. We have developed a novel method that combines the use of the FFT with a quasi-log scale, making the model calculation practical (sections 3.2S).

The computation of the PAID function begins by scaling the pre-computed kernels to match the brightness in each channel $q_{\alpha A}$ and diffusion time τ_{α}^{D} for each species α, obtaining $f_{\alpha STM}^{(p,q)}\left(K_{M}\,|\,\tau\right)$, re-binning the functions in the quasi-log bins used for the convolution method (section 3.2S). Using the relation $\mathcal{N}_{\alpha}=c_{\alpha}\mathcal{V}_{\text{sample}}/V_{\text{eff}}$, the presence of multiple molecules is accounted for using the convolutions in eq 11S. The final step is to implement the final integration as a matrix multiplication, described in section 3.3S.

3.1S Equivalence of convolutions in n-space and K-space

One property that needs to be established is the equivalence of performing convolutions in $\mathbf{n}$ space and K space. Consider three distributions $f(\mathbf{n})$, $g(\mathbf{n})$, and $h(\mathbf{n})$, such that h is the convolution of f and g:

$$h(\mathbf{n})=\sum_{\mathbf{n}'=0}^{\mathbf{n}}f(\mathbf{n}')g(\mathbf{n}-\mathbf{n}')\equiv f(\mathbf{n})*g(\mathbf{n}) \tag{16S}$$

We can write each of these distributions in $\mathbf{n}$ as a Poisson transformation: $f(\mathbf{n})=\int_{0}^{\infty}\tilde{f}(K)\text{Poi}(K,\mathbf{n})dK$,

$g(\mathbf{n})=\int_{0}^{\infty}\tilde{g}(K)\text{Poi}(K,\mathbf{n})dK$, and $h(\mathbf{n})=\int_{0}^{\infty}\tilde{h}(K)\text{Poi}(K,\mathbf{n})dK$. By substituting these expressions for $f(\mathbf{n})$ and $g(\mathbf{n})$ into eq 16S, we get

$$\begin{aligned}h(\mathbf{n})&=\int_{0}^{\infty}dK_{\text{f}}\,\tilde{f}(K_{\text{f}})\int_{0}^{\infty}dK_{\text{g}}\,\tilde{g}(K_{\text{g}})\sum_{\mathbf{n}'=0}^{\mathbf{n}}\text{Poi}(K_{\text{f}},\mathbf{n}')\text{Poi}(K_{\text{g}},\mathbf{n}-\mathbf{n}')\\[6pt]&=\int_{0}^{\infty}dK_{\text{f}}\,\tilde{f}(K_{\text{f}})\int_{0}^{\infty}dK_{\text{g}}\,\tilde{g}(K_{\text{g}})\text{Poi}(K_{\text{f}}+K_{\text{g}},\mathbf{n})\\[6pt]&=\int_{0}^{\infty}dK\left[\int_{0}^{K}dK'\tilde{f}(K')\tilde{g}(K-K')\right]\text{Poi}(K,\mathbf{n})\end{aligned} \tag{17S}$$

Since we also know that $h(\mathbf{n}) = \int_0^\infty dK\, \tilde{h}(K)\,\text{Poi}(K, \mathbf{n})$, we find

$$\tilde{h}(K) = \int_0^K dK'\, \tilde{f}(K')\,\tilde{g}(K - K') = \tilde{f}(K) * \tilde{g}(K) \qquad (18S)$$

This shows that convolutions can equivalently be performed either in $\mathbf{n}$-space or K-space.

3.2S Efficient calculation of convolutions on a quasi-logarithmic scale

Because of the large number of convolutions necessary to calculate the PAID histogram, it is necessary to have an efficient algorithm for convolutions. The standard method is to Fast Fourier Transform (FFT) the data, multiply the data in the complex transform space, and then FFT back to real space (See for example Ch. 13 of ref 3). The FFT, however, requires evenly spaced bins, which can produce huge arrays for a large dynamic range. We use a convolution method inspired by the multiple-tau correlation technique[1,2] that uses the FFT for small convolutions. For each vector ($\mathbf{a}$ and $\mathbf{b}$) to be convolved, we produce a series of linearly spaced arrays of length 64, where the resolution of each is reduced by a factor of 2 compared to its predecessor. The structure of these vectors is shown in Fig. 2S-A. In the figure, the arrays are of length 8 (rather than 64) for each factor of 2 in resolution. The array with the smallest spacing is shown at the left. In Fig. 2S-A, the first array has a spacing of $\Delta K = 0.1$. The second array has double the spacing, $\Delta K = 0.2$. The first four elements are each made up of a sum of two elements from the first array. The third array has a spacing that is doubled again, $\Delta K = 0.4$. Again, the first four elements are each made up of a sum of two elements from the second array. Continuing this structure to successively lower resolution, one obtains a series of linearly spaced arrays that can efficiently span a large dynamic range.

In Figs. 2S-B and 2S-C, we illustrate how the convolution of two series of such arrays is performed. One array is labeled with lower-case letters, and the other is labeled with upper-case letters. To calculate the convolution of two vectors, $\mathbf{a} * \mathbf{b}$, the arrays with the smallest spacing ($\Delta K = 0.1$) are convolved with each other first, as shown in Fig. 2S-B. The FFT procedure described above is used, with the arrays are zero-padded to twice the original length. We use the FFTW (Fastest Fourier

14

Transform in the West) set of routines, developed by M. Frigo and S. G. Johnson. The lower half of the array becomes the new array for the smallest spacing. The whole array is rebinned to the next largest spacing, and serves as a contribution to the convolution with spacing $\Delta K = 0.2$.

In Fig. 2S-C, we show how the convolution for the larger spacing $\Delta K = 0.2$ is performed. Because we already performed part of the convolution with that spacing, we need to exclude that contribution in subsequent calculations. This is done by setting (or "clipping") the lower half of one sub-array (*lower-case*) to 0, and convolving it with the other sub-array (*upper-case*). The clipped version from **a** is convolved with the full version from **b**, and vice versa. By adding these results with the array obtained by rebinning the result with spacing $\Delta K = 0.1$, we obtain the final result for $\Delta K = 0.2$. This allows the use of information from higher resolution arrays to contribute to the convolutions with lower resolution arrays.

In the quasi-logarithmically spaced arrays, we approximate the function to be convolved by a series of steps; over the range of each bin, the function is assumed to be constant. When two such steps are convolved, we get a triangle, shown in Fig. 2-SD. The convolution of the two steps leaks into the next bin. The discrete convolutions shown in Figs. 2S-B and 2S-C do not account for this because they assume that the function is nonzero only at discrete values. In order to approximate the continuous functions with the method presented, the convolution is shifted one-half bin up after each FFT-based convolution in Figs. 2S-B and 2S-C.
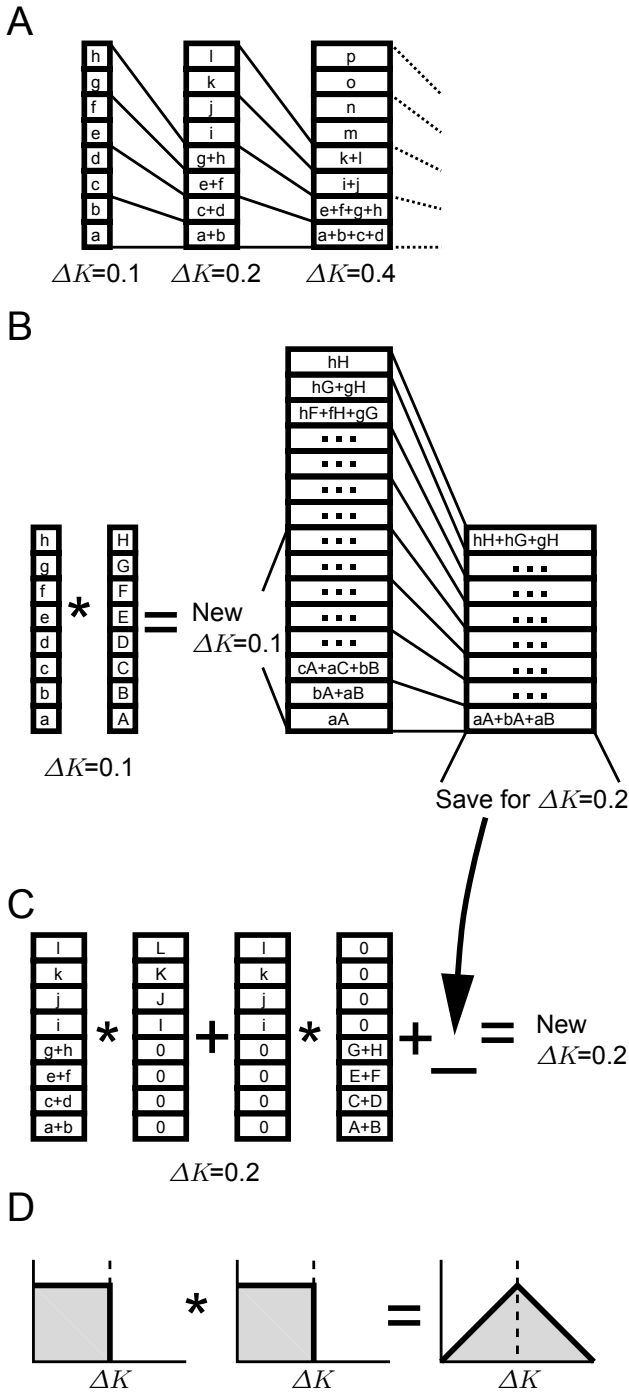
A

h	l	p
g	k	o
f	j	n
e	i	m
d	g+h	k+l
c	e+f	i+j
b	c+d	e+f+g+h
a	a+b	a+b+c+d

ΔK=0.1 ΔK=0.2 ΔK=0.4

B

ΔK=0.1

C

ΔK=0.2

D

ΔK ΔK ΔK

Fig. 2S: In A, the structure of the arrays with quasi-logarithmic spacing used for fast convolutions is shown. In B and C, we illustrate how the convolution of two series of such arrays is performed. In D, the convolution of two step functions is shown.

In eq 11S, there are successive convolutions performed on the same array which have the form $\left[f_{\gamma STM}^{(0,0)} \left(K_M \,|\, \tau \right) * \right]^{\mathcal{N}_\gamma}$. In order to compute $\mathcal{N}_\gamma$ successive convolutions on the original distribution

16

$f_{\gamma STM}^{(0,0)}\left(K_M|\tau\right)$, we can use a trick to prevent performing $\mathcal{N}_\gamma$ convolutions directly. We express the number $\mathcal{N}_\gamma$ in binary. By recursively convolving the original array, we obtain a series of vectors for each convolved power of 2: 1,2,4,8,… By convolving only those convolved powers of 2 in the binary representation of $\mathcal{N}_\gamma$, we obtain $\left[f_{\gamma STM}^{(0,0)}\left(K_M|\tau\right)*\right]^{\mathcal{N}_\gamma}$. This allows the number of convolutions performed to increase logarithmically with $\mathcal{N}_\gamma$.

3.3S Conversion of PAID function from K-space to n-space

The PAID function is expressed in eq 11S as an integral of the Poisson distribution over a function of K (defining $\tilde{C}_{STM}\left(\boldsymbol{\tau},K\right)$),

$$C_{STM}\left(\boldsymbol{\tau},\mathbf{n}\right)=\int_0^\infty dK\,\mathrm{Poi}\left(K,\mathbf{n}\right)\tilde{C}_{STM}\left(\boldsymbol{\tau},K\right) \tag{19S}$$

We have dropped the subscript M for simplicity.

The kernels used as the basis for the model have finite sized bins in K, which are indexed by b and have the range $\left[K_{min}^b,K_{max}^b\right)$. Because of the finite bin size, what is really calculated is the average over a bin $\int_{K_{min}^b}^{K_{max}^b}dK'\tilde{C}_{STM}\left(\tau,K'\right)\Big/K_{max}^b-K_{min}^b$. We assume that the amplitude is constant across each bin, so that $\tilde{C}_{STM}^b\left(\tau,K\right)\approx\int_{K_{min}^b}^{K_{max}^b}dK'\tilde{C}_{STM}\left(\tau,K'\right)\Big/K_{max}^b-K_{min}^b$ for K in the range $\left[K_{min}^b,K_{max}^b\right)$.

The expression for $C_{STM}\left(\boldsymbol{\tau},\mathbf{n}\right)$ as calculated until now gives the instantaneous rates at a particular τ. However, the time interval bins for the photon counting data have finite extent, and so $C_{STM}\left(\boldsymbol{\tau},\mathbf{n}\right)$ must be averaged over the range of the time bin. To calculate $\tilde{C}_{STM}\left(\tau,K\right)$ within the time interval τ_{min} and τ_{max}, we interpolate between $\tilde{C}_{STM}\left(\tau_{min},K\right)$ and $\tilde{C}_{STM}\left(\tau_{max},K\right)$. We project $\tilde{C}_{STM}\left(\tau_{min},K\right)$ forward in time from τ_{min}, noting that the limits of a bin in cumulative intensity scales with time interval, $K_{min}'^b\left(\tau\right)=K_{min}^b\,\tau/\tau_{min}$ and $K_{max}'^b\left(\tau\right)=K_{max}^b\,\tau/\tau_{min}$. The value interpolated forward from τ_{min}

is $^{\mathrm{frwd}}\tilde{C}^b_{STM}(\tau,K) \approx \int_{K^b_{\min}}^{K^b_{\max}} dK'\tilde{C}_{STM}(\tau_{\min},K') \Big/ \left(K^b_{\max}-K^b_{\min}\right)\tau/\tau_{\min}$ for K in the range $\left[K'^b_{\min}(\tau), K'^b_{\max}(\tau)\right)$.

$^{\mathrm{frwd}}\tilde{C}^b_{STM}(\tau,K)$ is 0 outside of this range. Similarly, the value interpolated backward from $\tau_{\max}$ (with

$K''^b_{\min}(\tau) = K^b_{\min}\,\tau/\tau_{\max}$ and $K''^b_{\max}(\tau) = K^b_{\max}\,\tau/\tau_{\max}$)

is $^{\mathrm{bkwd}}\tilde{C}^b_{STM}(\tau,K) \approx \int_{K^b_{\min}}^{K^b_{\max}} dK'\tilde{C}_{STM}(\tau_{\max},K') \Big/ \left(K^b_{\max}-K^b_{\min}\right)\tau/\tau_{\max}$ for K in the range $\left[K''^b_{\min}(\tau), K''^b_{\max}(\tau)\right)$.

$^{\mathrm{bkwd}}\tilde{C}^b_{STM}(\tau,K)$ is 0 outside of this range. A weighted average over $\tau \in [\tau_{\min},\tau_{\max})$ of the forward and backward projections is performed, with the results rebinned into the original spacing,

$$^{\mathrm{avgd}}\tilde{C}^b_{STM}(\tau,K) \approx \sum_{b'} \int_{\tau_{\min}}^{\tau_{\max}} d\tau' \,^{\mathrm{frwd}}\tilde{C}^{b'}_{STM}(\tau,K)\frac{\tau_{\max}-\tau'}{\tau_{\max}-\tau_{\min}} + \,^{\mathrm{bkwd}}\tilde{C}^{b'}_{STM}(\tau,K)\frac{\tau'-\tau_{\min}}{\tau_{\max}-\tau_{\min}} \Big/ \tau_{\max}-\tau_{\min} \quad (20S)$$

for K in the range $\left[K^b_{\min}, K^b_{\max}\right)$, and τ in the range $[\tau_{\min},\tau_{\max})$.

The final integration over K shown in eq 19S is implemented as a matrix multiplication. The approximation that the value of $\tilde{C}_{STM}(\tau,K)$ is constant over a bin with limits $K^b_{\min}$, and $K^b_{\max}$ leads to the expression,

$$C_{STM}(\tau,\mathbf{n}) = \sum_b \,^{\mathrm{avgd}}\tilde{C}^b_{STM}(\tau,K)\int_{K^b_{\min}}^{K^b_{\max}} dK'\mathrm{Poi}(K',\mathbf{n}) \quad (21S)$$

The integral on the right can be expressed in terms of the incomplete gamma

function $\gamma(a,x) = \dfrac{1}{\Gamma(a)}\int_0^x e^{-t}t^{a-1}dt$,

$$\int_{K^b_{\min}}^{K^b_{\max}} \mathrm{Poi}(K,\mathbf{n})dK = \gamma\left(\mathbf{n}+1,K^b_{\max}\right) - \gamma\left(\mathbf{n}+1,K^b_{\min}\right) \quad (22S)$$

We keep the bin spacing fixed, so the integrals in eq22S need to be performed only once to create the matrix.

4S Figure of numerically approximated detection volume (cf. section 3.5)
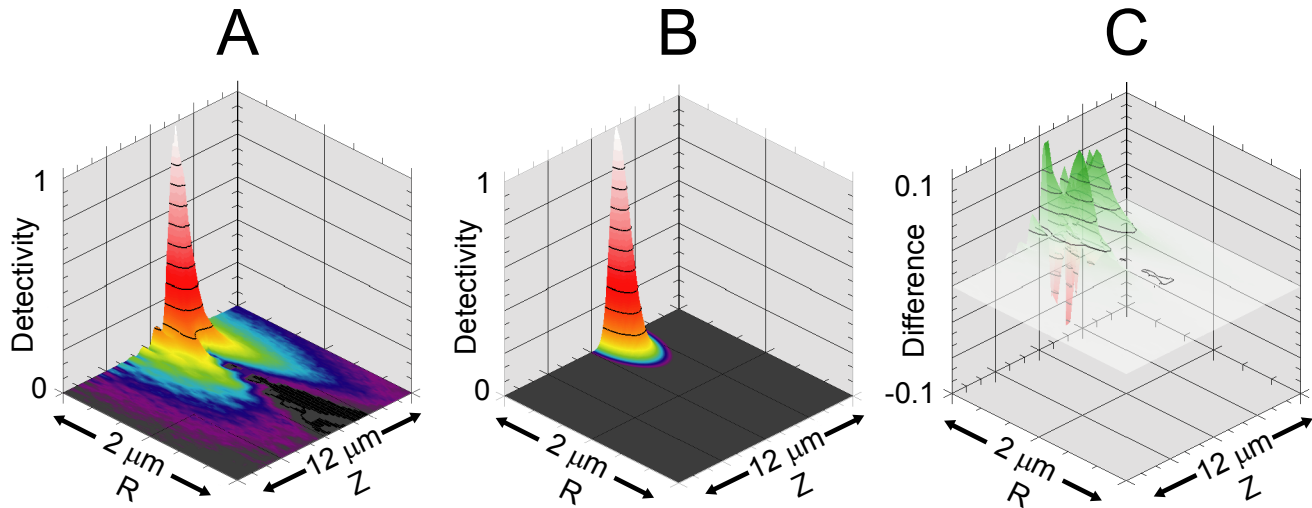

Figure 3S: Comparison of the numerically approximated detection volume to a Gaussian detection volume (see section 3.5 for description of calculation). A, Numerically approximated detection volume. B, Gaussian detection volume fitted to detection volume in A (lateral width = 0.27 mm; longitudinal width = 0.84 mm). C, Difference between detection volumes. Although the central peak is well-fitted, the wings of the detection volume are poorly fitted.

5S Single-channel, intermediate-occupancy simulations; modified FIMDA (cf. section 4.2)

5.1S Expanded tables including intermediate occupancy simulations

Table 1S. Parameters Extracted using PAID, modified FIMDA, and FIMDA for single-channel, one-component simulations in a Gaussian detection volume

parameters	Simulation	PAID	modified FIMDA	FIMDA	FCS	FIDA
Averaged Fits for 10 Simulations (30 s each): Low Occupancy
χ^2	-	1.2 ± 0.1	1.1 ± 0.1	4.7 ± 0.2	1.3 ± 0.1	0.6 ± 0.1
k_{01} (kHz)	0.0	0.01 ± 0.01	0.002 ± 0.001	0.01 ± 0.01	N/A	0.02 ± 0.01
c_1 (mol)	0.1	0.100±0.001	0.099±0.001	0.100±0.001	0.099±0.001	0.105±0.001
τ_1^D (µs)	100.0	98 ± 1	99 ± 1	99 ± 1	100 ± 1	N/A
q_{11} (kHz)	50.0	48.9 ± 0.3	49.2 ± 0.3	49.7 ± 0.3	N/A	47.3 ± 0.3
Averaged Fits for 10 Simulations (10 s each): Intermediate Occupancy
χ^2	-	1.0 ± 0.1	0.8 ± 0.1	1.0 ± 0.1	1.0 ± 0.1	0.7 ± 0.1
k_{01} (kHz)	0.0	0.03 ± 0.01	0.06 ± 0.01	0.13 ± 0.03	N/A	0.3 ± 0.1
c_1 (mol)	1.0	0.99 ± 0.01	1.00 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	1.05 ± 0.01
τ_1^D (µs)	100.0	99 ± 1	101 ± 1	94 ± 1	100 ± 1	N/A
q_{11} (kHz)	50.0	49.5 ± 0.2	48.6 ± 0.2	50.8 ± 0.2	N/A	47.6 ± 0.2

Table 2S. Parameters Extracted using PAID, modified FIMDA, and FIMDA for single-channel, two-component simulations in a Gaussian detection volume

parameters	simulation	PAID	modified FIMDA	FIMDA	FCS	FIDA
Averaged Fits for 10 Simulations (30 s each): Low Occupancy
χ^2	-	0.77 ± 0.04	0.75±0.04	1.1 ± 0.3	1.3 ± 0.2	0.5 ± 0.1
k_{01} (kHz)	2.0	2.00 ± 0.01	2.00 ± 0.01	1.90 ± 0.01	2.0	1.99 ± 0.03
c_1 (mol)	0.05	0.048±0.001	0.048±0.001	0.064±0.001	0.05	0.059±0.002
τ_1^D (µs)	100.0	100 ± 2	99 ± 2	117 ± 4	105 ± 7	N/A
q_{11} (kHz)	50.0	50 ± 1	50 ± 1	51 ± 2	50.0	48 ± 2
c_2 (mol)	0.05	0.052±0.001	0.052±0.001	0.042±0.002	0.05	0.048±0.003

τ_2^D (μs)	*150.0*	145 ± 2	146 ± 2	146 ± 3	142 ± 3	N/A
q_{21} (kHzl)	*100.0*	98 ± 1	98 ± 1	106 ± 1	*100.0*	100 ± 2
Averaged Fits for 10 Simulations (10 s each): Intermediate Occupancy						
χ^2	-	0.9 ± 0.1	0.58 ± 0.02	0.55 ± 0.02	1.4± 0.2	0.7 ± 0.1
k_{01} (kHz)	*2.0*	2.03 ± 0.02	2.05 ± 0.01	1.6 ± 0.2	*2.0*	1.6 ± 0.3
c_1 (mol)	*0.5*	0.45 ± 0.03	0.43 ± 0.04	0.51 ± 0.04	*0.5*	0.57 ± 0.03
τ_1^D (μs)	*100.0*	88 ± 6	95 ± 8	92 ± 5	120 ± 7	N/A
q_{11} (kHz)	*50.0*	53 ± 1	50 ± 2	52 ± 2	*50.0*	46 ± 4
c_2 (mol)	*0.5*	0.53 ± 0.02	0.56 ± 0.04	0.50 ± 0.04	*0.5*	0.52 ± 0.04
τ_2^D (μs)	*150.0*	153 ± 1	148 ± 3	144 ± 4	141 ± 3	N/A
q_{21} (kHz)	*100.0*	96 ± 1	94 ± 2	101 ± 2	*100.0*	97 ± 2

5.2S Modified FIMDA model

A new model for FIMDA can be developed using the methods in the section 3S. A modification of Eq. 11S may be used to obtain the photon count distribution for a bin width τ,

$$P_{STM}(\tau, \mathbf{n}) = \int_0^\infty dK_M \mathrm{Poi}(K_M, \mathbf{n}) \prod_\gamma \left[f_{\gamma STM}^{(0,0)}(K_M | \tau) * \right]^{N_\gamma} \qquad (23S)$$

The scaling laws described with the PAID function model are used here as well, and the convolutions are performed as described in section 3.2S. We refer to this model as "modified FIMDA".

The columns in Tables 1S and 2S labeled "modified FIMDA" use our model to extract the parameters from the histograms, and the columns labeled "FIMDA" use the original FIMDA model[8]. The simulations for Tables 1S and 2S are the same as those for Tables 1 and 2 in the main text. Some differences were found between fitting the FIMDA histograms with the modified FIMDA model and the original model in ref 8. For the single-species simulations at low occupancy, the accuracy of the extracted values is similar to the accuracy found using PAID and modified FIMDA, but χ^2 is significantly higher than 1. This comes from the region where time delays are larger than the diffusion time. The fits for one modified FIMDA histogram using the modified FIMDA model and the original FIMDA model are shown in Fig. 4S. It is possible to see in fig. 4S-C and 4S-D the deviation of the FIMDA model (*green line*) from the data (*black line*).
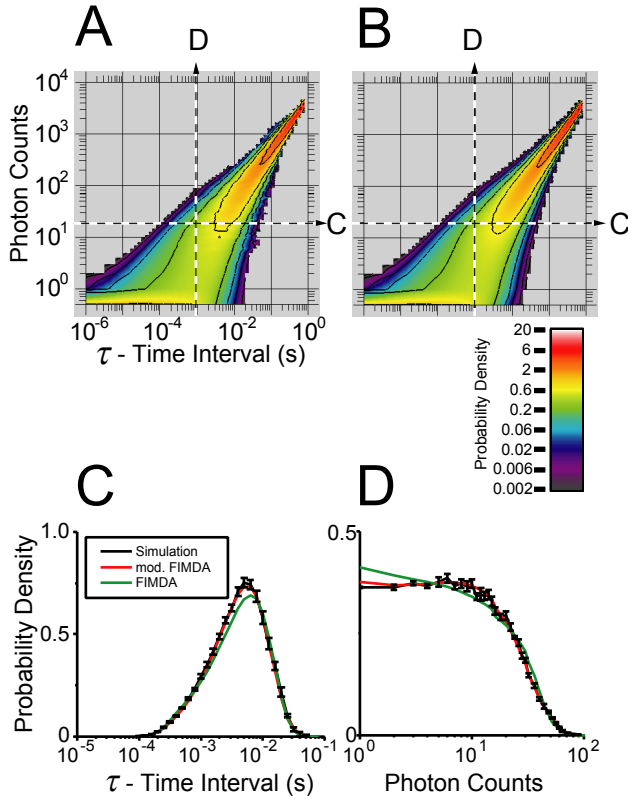
Fig. 4S: One of the fits in Table 1S using FIMDA. The occupancy is $c_1 = 0.1$, the diffusion time is $\tau_1^D = 100 \ \mu s$, the brightness is $q_{11} = 50 \ kHz$, and the background count rate is $q_{01} = 0 \ kHz$. The FIMDA histogram for the simulation is in A (represented using the normalization described in this section), and the fit using modified FIMDA is in B. Horizontal slices of both are shown in C, and vertical slices are shown in D. The slices of the simulation are shown in black with error bars, the slices of the fit using modified FIMDA are shown in red, and the slices of the fit using the FIMDA model, specialized to the Gaussian detection volume with no triplet state, are shown in green.

The difference between the models is less pronounced at higher occupancies. For the single-species simulations at intermediate occupancy, the χ^2 is near 1, and the extracted parameters are close to the simulation values, except for a 5% downward bias in the diffusion time using the FIMDA model.

The parameters extracted using the FIMDA model for the simulations of the low occupancy mixtures had larger biases than those extracted with the modified FIMDA model (Table 2S). For the low

occupancy mixture, the biases in the occupancies were close to 20%. It performed better with the intermediate occupancy simulations (with smaller biases in the occupancy values), except that the background count rate values extracted were extracted more accurately with PAID and modified FIMDA.

We attribute the increased χ^2 and biases found using the original FIMDA model to the assumption made that the functional form of the photon counting histogram changes only through changes in effective brightness and effective occupancy. In the view of the original model, at time bins larger than the diffusion time, the effective occupancy is increased while the effective brightness is reduced; the detection volume increases in size as time passes, but retains the same shape. This view does not account for the dynamic picture, neglecting for example the chance that a molecule re-crosses the detection volume.

6S Single-channel, intermediate-occupancy experiments; initial experiments using Cy3 (cf. section 4.3)

6.1S Tables for single-channel, intermediate-occupancy experiments

Table 3S. Parameters extracted using single-component PAID fits for single-channel, single-species experiments

parameters	DNA^{Cy3B,1T}	DNA$^{Cy3B,1T/Cy3B,65B}$
Averaged Fits for 10 Measurements (30 s each): Intermediate Occupancy		
χ^2	1.5 ± 0.1	1.2 ± 0.1
k_{01} (kHz)	0.23 ± 0.04	2.2 ± 0.3
c_1 (mol)	2.21 ± 0.02	2.32 ± 0.04
τ_1^D (µs)	620 ± 10	650 ± 10
q_{11} (kHz)	8.7 ± 0.1	17.9 ± 0.2

Buffer-only measurements - $k_{01} = 0.24 \pm 0.01$ kHz

Table 4S. Parameters Extracted using two-component PAID fits for single-channel, one- and two-species experiments

Parameters	DNA^{Cy3B,1T} fixed ratio fit[a]	DNA^{Cy3B,1T} restricted fit[b]	DNA$^{Cy3B,1T}_{/Cy3B,65B}$ fixed ratio fit	DNA$^{Cy3B,1T}_{/Cy3B,65B}$ restricted fit	DNA^{Cy3B,1T} DNA$^{Cy3B,1T}_{/Cy3B,65B}$ fixed ratio fit	DNA^{Cy3B,1T} DNA$^{Cy3B,1T}_{/Cy3B,65B}$ restricted fit
Averaged Fits for 10 Measurements (30 s each): Intermediate Occupancy						
χ^2	1.3 ± 0.1	1.6 ± 0.1	1.0 ± 0.1	1.3 ± 0.1	2.5 ± 0.1	2.8 ± 0.1
k_{01} (kHz)	0.7 ± 0.1	*0.24*	0.5 ± 0.1	*0.24*	0.9 ± 0.1	*0.24*
c_1 (mol)	2.04 ± 0.04	2.19 ± 0.01	0.5 ± 0.1	0.2 ± 0.1	1.4 ± 0.1	1.99 ± 0.04
τ_1^D (µs)	640 ± 10	*620*	630 ± 10	*620*	680 ± 10	*620*
q_{11} (kHz)	9.0 ± 0.2	*8.7*	9.2 ± 0.1	*8.7*	7.4 ± 0.1	*8.7*
c_2 (mol)	0.10 ± 0.03	0.005 ± 0.002	2.1 ± 0.1	2.41 ± 0.04	1.0 ± 0.1	0.49 ± 0.03
τ_2^D (µs)	640 ± 10	*620*	630 ± 10	*620*	680 ± 10	*620*
q_{21} (kHz)	4.5 ± 0.1	*17.4*	18.5 ± 0.3	*17.4*	14.9 ± 0.3	*17.4*

[a]The ratio between the brightness of both components fixed to a factor of 2, and the diffusion times of the two components are linked.

6.2S Initial single-channel PAID experiments using Cy3

Initially, our single-channel PAID experiments were performed using Cy3 (rather than Cy3B); in this case, the ratio in brightness between DNA$^{Cy3,1T/Cy3,65B}$ and DNA^{Cy3,1T} was measured to be ~1.5 (data not shown). Even at low excitation intensities (<10 kW/cm^2), the factor of ~1.5 persisted, excluding triplet state saturation and photobleaching from the possible sources of the discrepancy. Using FCS, we observed a fluctuation that became faster with increasing excitation intensity without a corresponding change in the fluctuation amplitude (unlike triplet fluctuations). A similar effect was observed previously for Cy5, and was identified as photo-induced isomerization[11]. Since isomerization of two distinct fluorophores is uncorrelated, the amplitude of the fluctuation is reduced by a factor of 2 in the DNA$^{Cy3,1T/Cy3,65B}$, increasing the apparent concentration of the double-labeled species as compared to the single-labeled DNA^{Cy3,1T}, resulting in only a factor of ~1.5 difference in brightness. Using Cy3B, a conformationally-constrained analog of Cy3 which prevents isomerization, the fluctuation is absent. This emphasizes the importance of the choice of fluorophore when attempting to use brightness as a measure of stoichiometry. Dual-channel methods (discussed in sections 4.4-4.6) are less sensitive to such effects.

7S Simulations–quantitative analysis of binding using dual-channel PAID; PAID histograms with two monitor channels (cf. section 4.4)

7.1S Simulations–quantitative analysis of binding using dual-channel PAID

PAID and FCCS in dual-channel applications with two spectrally-separable fluorophores were investigated using multiple species simulations. As in the single-channel case, low- and intermediate-occupancy regimes were studied (Tables 5S and 6S). The parameters used for the simulations are listed in the 2nd column of tables 5S and 6S.

For each simulation, three fits were performed. The 1st fit (Tables 5S and 6S, 3rd column) uses PAID with all parameters unrestricted. The 2nd fit (Tables 5S and 6S, 4th column) also uses PAID, with all parameters of the free components except for the occupancies being fixed. For the 1st and 2nd fits, all combinations of the dual-channel PAID histogram are fitted simultaneously; since each histogram emphasizes different species, parameters for all species can be extracted. The 3rd fit (Tables 5S and 6S, 5th column) uses FCCS to simultaneously fit the autocorrelations of the red and yellow channels and the two cross-correlations, with all brightness and background values fixed.

Table 5S. Parameters Extracted using PAID and FCCS fits for two-channel, three-component simulations in a Gaussian detection volume at low occupancy

parameters	Simulation	PAID- unrestricted fit	PAID- restricted fit	FCCS
Averaged Fits for 10 Simulations (30 s each): Low Occupancy				
χ^2	-	0.72 ± 0.02	0.79 ± 0.02	0.95 ± 0.04
k_{0R} (kHz)	2.0	2.00 ± 0.01	2.0	2.0
k_{0Y} (kHz)	2.0	2.00 ± 0.01	2.0	2.0
c_1 (mol)	0.05	0.049 ± 0.001	0.049 ± 0.001	0.050 ± 0.001
τ_1^D (µs)	300.0	287 ± 7	300.0	311 ± 4
q_{1R} (kHz)	50.0	50.3 ± 0.4	50.0	50.0
q_{1Y} (kHz)	0.0	0.05 ± 0.01	0.0	0.0
c_2 (mol)	0.05	0.049 ± 0.001	0.050 ± 0.001	0.050 ± 0.001
τ_2^D (µs)	300.0	294 ± 7	300.0	306 ± 4

q_{2R} (kHz)	5.0	4.8 ± 0.1	5.0	5.0
q_{2Y} (kHz)	45.0	44.8 ± 0.3	45.0	45.0
c_3 (mol)	0.05	0.050 ± 0.001	0.050 ± 0.001	0.051 ± 0.001
τ_3^D (µs)	400.0	394 ± 6	389 ± 7	404 ± 8
q_{3R} (kHz)	55.0	54.9 ± 0.3	55.2 ± 0.1	55.0
q_{3Y} (kHz)	45.0	44.7 ± 0.2	44.8 ± 0.1	45.0

Values that are fixed are shown in italics, with no errors listed.

Table 6S. Parameters Extracted using PAID and FCCS fits for two-channel, three-component simulations in a Gaussian detection volume at intermediate occupancy

parameters	Simulation	PAID-unrestricted fit	PAID - restricted fit	FCCS
Averaged Fits for 10 Simulations (10 s each): Intermediate Occupancy				
χ^2	-	0.81 ± 0.02	0.90 ± 0.04	0.76 ± 0.09
k_{0R} (kHz)	2.0	2.02 ± 0.02	2.0	2.0
k_{0Y} (kHz)	2.0	2.00 ± 0.02	2.0	2.0
c_1 (mol)	0.5	0.50 ± 0.01	0.49 ± 0.01	0.050 ± 0.01
τ_1^D (µs)	300.0	320 ± 9	300.0	313 ± 3
q_{1R} (kHz)	50.0	49.1 ± 0.2	50.0	50.0
q_{1Y} (kHz)	0.0	0.06 ± 0.02	0.0	0.0
c_2 (mol)	0.5	0.49 ± 0.01	0.49 ± 0.01	0.50 ± 0.01
τ_2^D (µs)	300.0	306 ± 11	300.0	303 ± 7
q_{2R} (kHz)	5.0	4.9 ± 0.2	5.0	5.0
q_{2Y} (kHz)	45.0	44.5 ± 0.4	45.0	45.0
c_3 (mol)	0.5	0.50 ± 0.01	0.49 ± 0.01	0.49 ± 0.01
τ_3^D (µs)	400.0	395 ± 4	390 ± 6	407 ± 5
q_{3R} (kHz)	55.0	54.3 ± 0.4	55.0 ± 0.2	55.0
q_{3Y} (kHz)	45.0	44.5 ± 0.5	45.1 ± 0.2	45.0

The unrestricted fit extracted reliable values, within 1-10% for all parameters in both sets of simulations. Unexpectedly, fixing the brightness of the free components and the background resulted only in a modest improvement for the brightness values for the complex and no improvement for the diffusion time or for the occupancies. The parameters extracted using FCCS had similar statistical accuracy (~1% range) to those found using PAID. The extracted diffusion times are somewhat better

with FCCS (the error bars were ~20% smaller), mainly due to the smaller numbers of fitted parameters (brightness and background were fixed to their simulation values; for experiments, it would be necessary to measure the brightness and background values using a method different than FCS.)

7.2S PAID histograms with two monitor channels

Higher sensitivity could be achieved if two monitor channels were used for the two-channel experiments, allowing better use of ratiometric information available in dual-channel experiments. Observables in single-molecule diffusion studies that depend on the ratio of two channels are more sensitive than brightness. A PAID histogram with two monitor channels has three axes: the time interval axis τ, the number of photons in the red channel $\mathbf{n}_R$, and the number of photons counted in the yellow channel $\mathbf{n}_Y$. Fig. 5S-A shows three slices of this histogram at different time intervals when A^y, B^r, and $A^y B^r$ are present, and Fig. 5S-B shows the same three slices when $A^y B^r$ is absent. The histograms shown are for the PAID histogram with the red channel assigned as the start channel ($S = R$), the yellow channel assigned as the stop channel ($T = Y$), and a monitor channel for each channel ($M_1 = R$ and $M_2 = Y$). A cartoon of a particular type of molecule is placed in the $\tau = 1$ ms slice at the approximate location where it contributes. The histograms in Figs. 6C and 6G in the main text are the collapse of Figs. 5S-A and 5S-B, respectively, summing along the yellow monitor channel axis, while keeping the time interval and red monitor channels axes. For an isolated burst, the number of photons counted depends on the diffusion path taken through the detection volume, whereas the ratio between two channels does not. For example, the correlation density peak in Fig. 5S-A at $\tau = 1$ ms corresponding the complex is wider in the direction of the diagonal $\mathbf{n}_R = \mathbf{n}_Y$ than in the perpendicular direction. The one monitor channel PAID histogram is a collapse of the two monitor channel histogram onto the Y or R axis. This collapse smears the central peak; the subpopulations of complex and free molecules are not clearly separated in the one monitor channel case, but are separated with two monitor channels (compare the $\tau = 1$ ms slice in Fig. 5S-A with the corresponding single

monitor channel histogram in Fig. 6C.) Although the benefits are clear, the fitting model has not yet been extended to account for two monitor channels.
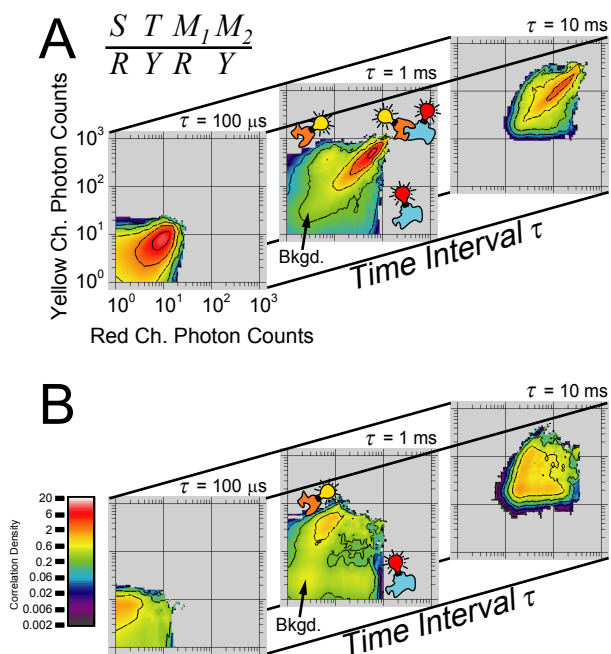

Figure 5S. Two-channel, two monitor channel PAID histograms for the same simulations as in fig 5 in the main text. The red channel is assigned as the start channel and the first monitor channel along the horizontal axis. The yellow channel is assigned as the stop channel and the second monitor channel along the vertical axis. Because the start and stop channels are assigned to different channels, these histograms emphasize the species that emits in both channels, the complex.   A two-monitor channel PAID histogram is three-dimensional, so three two-dimensional slices at $\tau = 100\ \mu s$, $\tau = 1\ ms$, and $\tau = 10\ ms$ are shown for each histogram. In A, both free species and complex are present. In the $\tau = 1\ ms$ slice, a cartoon of each diffusing species is placed next to the contribution from that species. In B, the complex is absent.

8S Correlations for dual-channel, low-occupancy experiments; PAID and FCS analysis of dual-channel, intermediate-occupancy experiments (cf. section 4.5)

8.1S Figure of correlations for dual-channel, low occupancy experiments
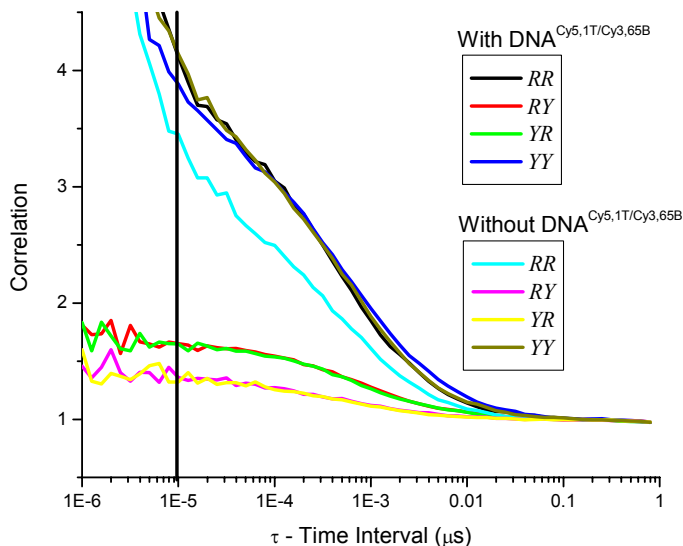

Fig. 6S: Correlations for dual-channel, low-occupancy mixture samples discussed in section 4.6. These correlations are collapses of the PAID histograms shown in fig 7 in the main text. *RR* and *YY* are the autocorrelations of channels *R* and *Y*, respectively; RY and YR are cross-correlations. When these data are fitted using the PAID model, time intervals < 10 μs are excluded from the fit (*black line*). The dramatic increase in the autocorrelations at short time is from afterpulsing effects from the APDs. Note that the cross-correlation amplitudes increase when DNA$^{Cy5,1T/Cy3,65B}$ is present (the residual cross-correlation without DNA$^{Cy5,1T/Cy3,65B}$ is due to leakage of the fluorophore *y* into channel *R*).

8.2S PAID analysis of dual-channel, intermediate occupancy experiments

In addition to the low-occupancy dual-channel experiments discussed in section 4.6, we also performed intermediate-occupancy dual-channel experiments. The occupancies for the intermediate-occupancy samples were expected to be 0.85 ± 0.09 for DNA^{Cy5,1T}, 1.15 ± 0.05 for DNA^{Cy3,65B}, and 0.84

± 0.03 for DNA$^{\text{Cy5,1T/Cy3,65B}}$. As in the main text, the fragments were prepared as free components and in the following mixtures: DNA$^{\text{Cy5,1T}}$/DNA$^{\text{Cy3,65B}}$ (to simulate non-interacting species); and DNA$^{\text{Cy5,1T/Cy3,65B}}$ /DNA$^{\text{Cy5,1T}}$/DNA$^{\text{Cy3,65B}}$ (to simulate interacting species).

   Using PAID, we extracted occupancy, diffusion time, and brightness from dual-channel experiments performed on single-species samples (Table 7S), then from dual-channel experiments performed on mixtures of those species (Table 8S). As in Tables 5 and 6 in the main text, we obtained fits with χ^2 in the range $10 > \chi^2 > 1$ (ideally $\chi^2 \sim 1$). Consistency between the values extracted from fits of data from single species samples and values extracted from multiple-species fits of mixtures demonstrates the ability of PAID to analyze mixtures of species.

Table 7S. Parameters extracted using PAID fits for two-channel, single-species, intermediate occupancy experiments

parameters	DNA$^{\text{Cy5,1T}}$	DNA$^{\text{Cy3,65B}}$	DNA$^{\text{Cy5,1T/Cy3,65B}}$
Averaged Fits for 10 Measurements (30 s): Intermediate occupancy			
χ^2	4.7 ± 0.2	2.2 ± 0.1	1.4 ± 0.1
k_{0R} (kHz)	1.30 ± 0.01	0.99 ± 0.01	0.92 ± 0.01
k_{0Y} (kHz)	1.33 ± 0.01	1.44 ± 0.01	1.29 ± 0.01
DNA Fragment 1	*DNA$^{\text{Cy5,1T}}$*	*DNA$^{\text{Cy3,65B}}$*	*DNA$^{\text{Cy5dark,1T/Cy3,65B}}$*
c_1 (mol)	0.68 ± 0.01	1.08 ± 0.01	0.32 ± 0.01
τ_1^{D} (µs)	390 ± 10	550 ± 20	700 ± 10
q_{1R} (kHz)	9.4 ± 0.2	0.82 ± 0.01	0.70 ± 0.03
q_{1Y} (kHz)	0.02 ± 0.01	9.0 ± 0.1	10.7 ± 0.2
DNA Fragment 2	*None*	*None*	*DNA$^{\text{Cy5,1T/Cy3,65B}}$*
c_2 (mol)			0.24 ± 0.01
τ_2^{D} (µs)			580 ± 10
q_{2R} (kHz)			7.6 ± 0.1
q_{2Y} (kHz)			7.6 ± 0.2

   Table 7S show fitted values for the samples containing only one species of labeled DNA, where the background rates, occupancies, diffusion times, and brightness in each channel were fitted parameters.

For the single-species, low-occupancy data (Table 6), the extracted occupancies were 35%-70% lower than expected from higher-concentration FCS experiments, whereas for the single-species, intermediate-occupancy data (Table 7S), they are 4%-20% lower than expected. The fact that the difference is smaller for the intermediate occupancy supports the suggestion in the main text that the difference between extracted and estimated occupancies can be attributed to loss of DNA on surfaces during handling (this problem is less severe at higher concentrations). The diffusion times extracted for $DNA^{Cy5,1T/Cy3,65B}$ and $DNA^{Cy3,65B}$ are similar to those found for low occupancy theoretical ones (550-700 μs); diffusion times for $DNA^{Cy5,1T}$ are similar as well (~400 μs).

Table 8S. Parameters extracted using PAID fits for two-channel, multiple-species, intermediate occupancy experiments

Parameters	$DNA^{Cy3,65B}$, $DNA^{Cy5,1T}$ unrestricted 2 component fit	$DNA^{Cy3,65B}$, $DNA^{Cy5,1T}$ restricted 3 component fit	$DNA^{Cy3,65B}$, $DNA^{Cy5,1T}$, $DNA^{Cy5,1T/Cy3,65B}$ unrestricted 3 component fit	$DNA^{Cy3,65B}$, $DNA^{Cy5,1T}$, $DNA^{Cy5,1T/Cy3,65B}$ restricted 3 component fit
Averaged Fits for 10 Measurements (30 s each): Intermediate occupancy				
χ^2	1.8 ± 0.1	3.0 ± 0.2	3.0 ± 0.1	4.8 ± 0.2
k_{0R} (kHz)	1.4 ± 0.1	*0.71*	1.66 ± 0.01	*0.71*
k_{0Y} (kHz)	1.44 ± 0.01	*1.15*	1.3 ± 0.1	*1.15*
DNA Frag. 1	$DNA^{Cy5,1T}$	$DNA^{Cy5,1T}$	$DNA^{Cy5,1T}$	$DNA^{Cy5,1T}$
c_1 (mol)	0.43 ± 0.01	0.57 ± 0.01	0.37 ± 0.01	0.74 ± 0.01
τ_1^D (μs)	350 ± 10	*389*	360 ± 10	*389*
q_{1R} (kHz)	11.0 ± 0.1	*9.4*	14.5 ± 0.4	*9.4*
q_{1Y} (kHz)	0.20 ± 0.03	*0.02*	0.3 ± 0.1	*0.02*
DNA Frag. 2	$DNA^{Cy3,65B}$	$DNA^{Cy3,65B}$	$^aDNA^{Cy3,65B}$	$^aDNA^{Cy3,65B}$
c_2 (mol)	0.86 ± 0.01	0.96 ± 0.01	1.0 ± 0.1	1.56 ± 0.01
τ_2^D (μs)	560 ± 10	*554*	500 ± 10	*554*
q_{2R} (kHz)	0.75 ± 0.02	*0.82*	0.5 ± 0.1	*0.82*
q_{2Y} (kHz)	9.8 ± 0.1	*9.0*	11.5 ± 0.3	*9.0*
DNA Frag. 3	*None*	$DNA^{Cy5,1T/Cy3,65B}$	$DNA^{Cy5,1T/Cy3,65B}$	$DNA^{Cy5,1T/Cy3,65B}$
c_3 (mol)		0.01 ± 0.01	0.58 ± 0.05	0.22 ± 0.01
τ_3^D (μs)		*575*	450 ± 10	*575*
q_{3R} (kHz)		*7.6*	5.3 ± 0.4	*7.6*

q_{3Y} (kHz)	7.6	6.4 ± 0.1	7.6

^{a} Includes contributions from DNA$^{Cy5dark,1T/Cy3,65B}$

As in table 6 in the main text, we performed two fits for mixtures of DNA fragments simulating non-interacting and interacting species (Table 8S). The 1st fit assumes the correct number of species, but allows all parameters to freely vary; the 2nd fit uses the single-species parameters already extracted to restrict the parameters for the free components, except for occupancy. These fits show that a sample with two species can be distinguished from a sample with three species, as is necessary for measuring interactions. PAID performed extremely well for all fits except the unrestricted fit of the three-component mixture. Although the three-component, unrestricted fit at intermediate occupancy identified three species, it produced values less consistent with values obtained with other fits.

The occupancies extracted from the mixtures using restricted fits were consistent with the occupancies extracted from the single-species samples (within 15%, except for DNA^{Cy5,1T} at low occupancy, which is 27% lower).

For the unrestricted intermediate-occupancy fits in Table 8S, the occupancies for DNA^{Cy5,1T} were smaller (25%-50%), with compensating increases in the brightness in R. The occupancies extracted using the unrestricted fit from DNA^{Cy3,65B} and DNA$^{Cy5,1T/Cy3,65B}$ in the three-fragment mixture are different from those obtained with the restricted fit (a 33% decrease and a 160% increase). With the unrestricted fits, the diffusion times are similar to those extracted for the single-species samples. The consistency found for brightness is excellent for all fits (typically ~10%).

These intermediate-occupancy experiments demonstrate that PAID can work in a variety of concentration regimes, and give additional support to the findings in the main text.

8.3S FCS analysis of dual-channel, intermediate occupancy experiments

The effects of triplet state fluctuations and fluorescence saturation[12], photobleaching[13,14], and photo-induced isomerization[11] are well-characterized in FCS; many of these features are evident in the data analyzed here. To compare with results using PAID, the data from Tables 7S and 8S were analyzed using FCS over the same range of time intervals, modeling only diffusion. We used simulated diffusion

paths through the non-Gaussian detection volume in section 3.5 to calculate the expected correlation function $C_0(\tau)$ for unit occupancy and unit diffusion time. This function was scaled to provide a functional form for fitting (τ^D is diffusion time and c is occupancy): $C(\tau) = \left[C_0\left(\tau/\tau^D\right) - 1 \right]\big/c + 1$.

Table 9S. Parameters extracted using FCS fits for only one two-channel correlation function in single-species, intermediate occupancy experiments – modeling diffusion time only

parameters	DNA^{Cy5,1T} Cy5 autocorr.	DNA^{Cy3,65B} Cy3 autocorr.	DNA$^{Cy5,1T/Cy3,65B}$ Cy5 autocorr.	DNA$^{Cy5,1T/Cy3,65B}$ Cy3 autocorr.	DNA$^{Cy5,1T/Cy3,65B}$ Cy3-Cy5 cross-corr.
Averaged Fits for 10 Measurements (30 s): Intermediate occupancy					
χ^2	1.7 ± 0.1	1.6 ± 0.1	1.0 ± 0.1	1.0 ± 0.1	0.6 ± 0.1
c_1 (mol)	1.03 ± 0.01	1.48 ± 0.03	0.59 ± 0.01	0.86 ± 0.01	1.31 ± 0.02
τ_1^D (µs)	510 ± 10	710 ± 10	580 ± 10	750 ± 10	900 ± 10

Table 9S shows parameters extracted using FCS fits of single correlation functions from the single-species dual-channel experiments used in Table 7S. The cross-correlation of the DNA$^{Cy5,1T/Cy3,65B}$ sample provides the longest, least biased diffusion time: photophysics of the Cy3 and Cy5 fluorophores are uncorrelated with each other, so the cross-correlation is affected only by diffusion. The remaining autocorrelations are shorter due to photophysical properties of the fluorophores.

Table 10S. Parameters extracted using FCS fits for only one two-channel correlation function in single-species, intermediate occupancy experiments – modeling diffusion time and a short time scale photophysical process (most likely photo-induced isomerization)

Parameters	DNA^{Cy5,1T} Cy5 autocorr.	DNA^{Cy3,65B} Cy3 autocorr.	DNA$^{Cy5,1T/Cy3,65B}$ Cy5 autocorr.	DNA$^{Cy5,1T/Cy3,65B}$ Cy3 autocorr.	DNA$^{Cy5,1T/Cy3,65B}$ Cy3-Cy5 cross-corr.
Averaged Fits for 10 Measurements (30 s): Intermediate occupancy					
χ^2	0.8 ± 0.1	0.7 ± 0.1	0.8 ± 0.1	0.63 ± 0.04	0.6 ± 0.1
c_1 (mol)	1.13 ± 0.01	1.60 ± 0.02	0.64 ± 0.01	0.94 ± 0.01	1.31 ± 0.02
τ_1^D (µs)	590 ± 10	840 ± 30	650 ± 10	860 ± 10	900 ± 10
F	0.26 ± 0.01	0.18 ± 0.01	0.26 ± 0.04	0.18 ± 0.01	-
$\tau_1^{Photophysics}$ (µs)	25 ± 2	60 ± 10	40 ± 10	54 ± 3	-

Table 10S uses the same data, but now fits an additional "photophysical factor", which can be due to either triplet state related fluctuations or photo-induced isomerization (the latter is the primary source of fluctuations for the Cy3 and Cy5 fluorophores[11]). Accounting for this, the diffusion times for the Cy3 autocorrelations are less than 10% smaller than the cross-correlation value. However, the diffusion times for the Cy5 autocorrelations are still significantly smaller, indicating that another effect such as photobleaching within the detection volume may play a role in the shortened diffusion time.

The previous fits show how photophysical properties of the fluorophores can give rise to variations in apparent diffusion time. However, the fitted diffusion times for the PAID histograms combine information from autocorrelations and cross-correlations. Therefore, the precise diffusion time values extracted using PAID for different DNA species depends on the contribution each species makes to the $S=T$ PAID histograms (analogous to autocorrelation in FCS) and $S{\neq}T$ PAID histograms (analogous to cross-correlation in FCS), as well as the presence of other species. For example, the value for $DNA^{Cy5,1T/Cy3,65B}$ is affected by all of the histograms, since it emits equally in both channels, whereas $DNA^{Cy5,1T}$ is affected primarily by the $S=T=Cy5$ PAID histograms.

To compare FCS-extracted and PAID-extracted diffusion times, the Cy3 and Cy5 autocorrelations and the Cy3-Cy5 cross-correlations are simultaneously fitted for the data sets used in Tables 7S and 8S (the fit is similar to the FCCS fits performed in Tables 5S and 6S). The background levels, brightnesses, and occupancies are all fixed to the values from Table 7S or the unrestricted fit values from Table 8S. Only the diffusion times are allowed to vary. The results of these fits are in Tables 11S and 12S. The fitted diffusion times were around 15% longer than extracted using PAID. Spreading the correlation function over the monitor photon axis likely enhances the effects of short time scale fluctuations. Most importantly, the variations in diffusion time extracted using FCS correlated with those extracted using PAID in Table 7S and 8S (see Figure 7S), indicating a common source for the variations.

Table 11S. Parameters extracted using simultaneous FCS fits for all two-channel autocorrelations and cross-correlations in single species, intermediate occupancy experiments – modeling diffusion time only

parameters	$DNA^{Cy5,1T}$	$DNA^{Cy3,65B}$	$DNA^{Cy5,1T/Cy3,65B}$
Averaged Fits for 10 Measurements (30 s): Intermediate occupancy			
χ^2	1.8 ± 0.2	1.8 ± 0.1	1.3 ± 0.1
k_{0R} (kHz)	*1.30*	*0.99*	*0.92*
k_{0Y} (kHz)	*1.33*	*1.44*	*1.29*
DNA Fragment 1	$DNA^{Cy5,1T}$	$DNA^{Cy3,65B}$	$DNA^{Cy5dark,1T/Cy3,65B}$
c_1 (mol)	*0.68*	*1.08*	*0.32*
τ_1^D (µs)	450 ± 10	630 ± 10	770 ± 30
q_{1R} (kHz/mol)	*9.4*	*0.82*	*0.70*
q_{1Y} (kHz/mol)	*0.02*	*9.0*	*10.7*
DNA Fragment 2	*None*	*None*	$DNA^{Cy5,1T/Cy3,65B}$
c_2 (mol)			*0.24*
τ_2^D (µs)			670 ± 20
q_{2R} (kHz/mol)			*7.6*
q_{2Y} (kHz/mol)			*7.6*

Table 12S. Parameters extracted using simultaneous FCS fits for all two-channel autocorrelations and cross-correlations in multiple species, intermediate occupancy experiments – modeling diffusion time only

parameters	$DNA^{Cy3,65B}$, $DNA^{Cy5,1T}$ 2 component fit	$DNA^{Cy3,65B}$, $DNA^{Cy5,1T}$, $DNA^{Cy5,1T/Cy3,65B}$ 3 component fit
Averaged Fits for 10 Measurements (30 s each): Intermediate occupancy		
χ^2	1.2 ± 0.1	1.5 ± 0.1
k_{0R} (kHz)	*1.4*	*1.66*
k_{0Y} (kHz)	*1.44*	*1.3*
DNA Fragment 1	$DNA^{Cy5,1T}$	$DNA^{Cy5,1T}$
c_1 (mol)	*0.43*	*0.37*
τ_1^D (µs)	420 ± 20	380 ± 20

q_{1R} (kHz/mol)	*11.0*	*14.5*
q_{1Y} (kHz/mol)	*0.20*	*0.3*
DNA Fragment 2	***DNA**^{Cy3,65B}*	*^{a}**DNA**^{Cy3,65B}*
c_2 (mol)	*0.86*	*1.0*
τ_2^D (µs)	650 ± 20	560 ± 10
q_{2R} (kHz/mol)	*0.75*	*0.5*
q_{2Y} (kHz/mol)	*9.8*	*11.5*
DNA Fragment 3	***None***	***DNA**$^{Cy5,1T/Cy3,65B}$*
c_3 (mol)		*0.58*
τ_3^D (µs)		630 ± 50
q_{3R} (kHz/mol)		*5.3*
q_{3Y} (kHz/mol)		*6.4*

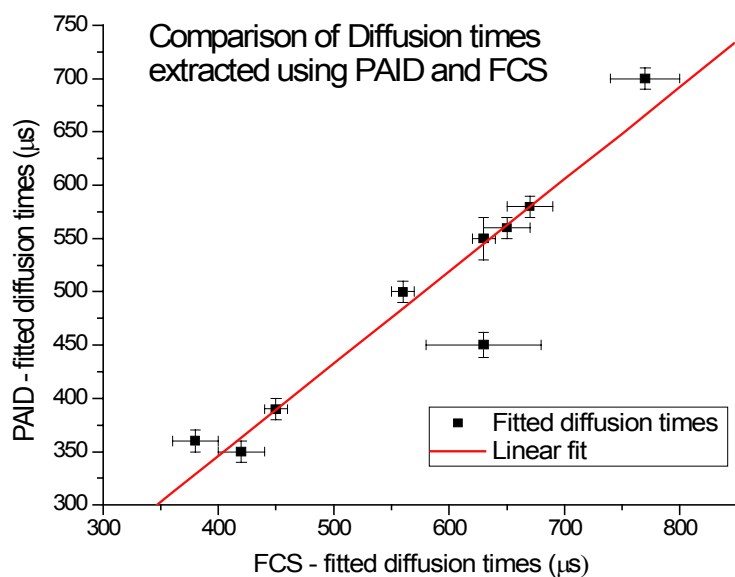^{a}Includes contributions from DNA$^{Cy5dark,1T/Cy3,65B}$

Fig. 7S: Variations in diffusion times fitted using PAID are correlated with variations in diffusion times fitted using FCS when only diffusion time is modeled. The PAID values are taken from tables 7S and 8S, and the FCS values are taken from tables 11S and 12S . The correlation coefficient is 0.96, and the linear fit using $\tau_{PAID}^D = \alpha \tau_{FCS}^D$ gives $\alpha = 0.86 \pm 0.02$.

References.

(1)	Schatzel, K. Inst. Phys. Conf. Ser. No. 77: session 4 1985, No. 77: session 4, 175.

(2)	Schatzel, K.; Peters, R. SPIE vol. 1430 Photon Correlation Spectroscopy: Multicomponent Systems 1991, 1430, 109.

(3)	Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery Numerical recipes in C : the art of scientific computing, 2nd ed.; Cambridge University Press: Cambridge, U.K., 1992.

(4)	Enderlein, J., David L. Robbins, W. Patrick Ambrose, Peter M. Goodwin, and Richard A. Keller. J. Phys. Chem. B 1997, 101, 3626.

(5)	Mandel, L. Proc. Phys. Soc. 1958, 72, 1037.

(6)	Mandel, L. Proc. Phys. Soc. 1959, 74, 233.

(7)	Enderlein, J. Phys. Lett. A 1996, 221, 427.

(8)	Palo, K.; Mets, U.; Jäger, S.; Kask, P.; Gall, K. Biophys. J. 2000, 79, 2858.

(9)	Kask, P.; Palo, K.; Ullmann, D.; Gall, K. Proc. Natl. Acad. Sci. U.S.A. 1999, 96, 13756.

(10)	Chen, Y.; Müller, J. D.; So, P. T.; Gratton, E. Biophys. J. 1999, 77, 553.

(11)	Widengren, J.; Schwille, P. J. Phys. Chem. A. 2000, 104, 6416.

(12)	Widengren, J.; Mets, U.; Rigler, R. J. Phys. Chem. 1995, 99, 13368.

(13)	Widengren, J.; Rigler, R. Bioimaging 1996, 4, 149.

(14)	Dittrich, P. S.; Schwille, P. Applied Physics B-Lasers and Optics 2001, 73, 829.