Biophysical Journal, Volume 98

**Supporting Material**

# Intramolecular Fluorescence Correlation Spectroscopy in a Feedback Tracking Microscope

Kevin McHale and Hideo Mabuchi

Biophysical Journal, Volume 98

**Supporting Material**

# Intramolecular fluorescence correlation spectroscopy in a feedback tracking microscope

Kevin McHale and Hideo Mabuchi

# Supporting information for "Intramolecular Fluorescence Correlation Spectroscopy in a Feedback Tracking Microscope"

Kevin McHale and Hideo Mabuchi

March 1, 2010

## 1 Derivation of Eq. 4

The crucial simplification that allows us to compute the tFCS curve explicitly for linear models relies on two general properties of the multivariate Gaussian distribution: that the difference between two multivariate Gaussian random variables is itself multivariate Gaussian, and that the Fourier transform of the multivariate Gaussian density is a multivariate Gaussian density. Here we verify that Eq. 4 holds whenever $\bar{\mathbf{x}}_{m\mu}$ and $\bar{\mathbf{y}}$ are jointly Gaussian by direct computation of the result.

We first write the definition of the average:

$$\left\langle e^{i\mathbf{k}^{\mathrm{T}}(\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}})}\right\rangle = \int \mathrm{d}^n\bar{\mathbf{x}}_{m\mu}\mathrm{d}^n\bar{\mathbf{y}}\, e^{i\mathbf{k}^{\mathrm{T}}(\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}})}p(\bar{\mathbf{x}}_{m\mu},\bar{\mathbf{y}}) \tag{S1}$$

where $n$ is the dimension of $\bar{\mathbf{x}}_{m\mu}$ and $\bar{\mathbf{y}}$ ($n = 6$ in the case of 3-dimensional tracking) and $p(\bar{\mathbf{x}}_{m\mu},\bar{\mathbf{y}})$ is the Gaussian joint probability density of these variables. We insert this joint density, in terms of its inverse Fourier transform for convenience, denoting the mean vector by $\mu$ and the covariance matrix by $\Sigma$:

$$\left\langle e^{i\mathbf{k}^{\mathrm{T}}(\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}})}\right\rangle = \int \mathrm{d}^n\bar{\mathbf{x}}_{m\mu}\mathrm{d}^n\bar{\mathbf{y}}\, e^{i\mathbf{k}^{\mathrm{T}}(\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}})}\int \frac{\mathrm{d}^{2n}\mathbf{q}}{(2\pi)^{2n}}\exp\left\{-i\mathbf{q}^{\mathrm{T}}\left(\begin{bmatrix}\bar{\mathbf{x}}_{m\mu}\\ \bar{\mathbf{y}}\end{bmatrix}-\mu\right)-\frac{1}{2}\mathbf{q}^{\mathrm{T}}\Sigma\mathbf{q}\right\} \tag{S2}$$

$$= \int \frac{\mathrm{d}^{2n}\mathbf{q}}{(2\pi)^{2n}}\mathrm{d}^n\bar{\mathbf{x}}_{m\mu}\mathrm{d}^n\bar{\mathbf{y}}\,\exp\left\{-i\left(\mathbf{q}-\begin{bmatrix}\mathbf{k}\\ -\mathbf{k}\end{bmatrix}\right)^{\mathrm{T}}\begin{bmatrix}\bar{\mathbf{x}}_{m\mu}\\ \bar{\mathbf{y}}\end{bmatrix}+i\mathbf{q}^{\mathrm{T}}\mu-\frac{1}{2}\mathbf{q}^{\mathrm{T}}\Sigma\mathbf{q}\right\}. \tag{S3}$$

In this form we can directly compute the integrals over $\bar{\mathbf{x}}_{m\mu}$ and $\bar{\mathbf{y}}$ by using the identity

$$\int \mathrm{d}x e^{ikx} = 2\pi\delta(k) \equiv \tilde{f}(k), \tag{S4}$$

which is easily demonstrated by using the fact that $\tilde{f}(k)$ is the Fourier transform of $f(x) = 1$, i.e.

$$\int \frac{\mathrm{d}k}{2\pi}e^{-ikx}\tilde{f}(k) = 1, \tag{S5}$$

which is clearly satisfied by Eq. S4. Application of this identiy to compute the integrals over $\bar{\mathbf{x}}_{m\mu}$ and $\bar{\mathbf{y}}$ yields a factor of $(2\pi)^{2n}\delta^{2n}\left(\mathbf{q}-\begin{bmatrix}\mathbf{k}\\ -\mathbf{k}\end{bmatrix}\right)$, which trivializes computation of the remaining integral over $\mathbf{q}$:

$$\int \mathrm{d}^{2n}\mathbf{q}\,\delta^{2n}\left(\mathbf{q}-\begin{bmatrix}\mathbf{k}\\ -\mathbf{k}\end{bmatrix}\right)e^{i\mathbf{q}^{\mathrm{T}}\mu-\frac{1}{2}\mathbf{q}^{\mathrm{T}}\Sigma\mathbf{q}} = \exp\left\{i\begin{bmatrix}\mathbf{k}^{\mathrm{T}} & -\mathbf{k}^{\mathrm{T}}\end{bmatrix}\mu-\frac{1}{2}\begin{bmatrix}\mathbf{k}^{\mathrm{T}} & -\mathbf{k}^{\mathrm{T}}\end{bmatrix}\Sigma\begin{bmatrix}\mathbf{k}\\ -\mathbf{k}\end{bmatrix}\right\}. \tag{S6}$$

Now we are left to insert the values for $\mu$ and $\Sigma$. From their definition, these are given by

$$\mu = \begin{bmatrix}\langle\bar{\mathbf{x}}_{m\mu}\rangle\\ \langle\bar{\mathbf{y}}\rangle\end{bmatrix}, \Sigma = \begin{bmatrix}\left\langle\bar{\mathbf{x}}_{m\mu}\bar{\mathbf{x}}_{m\mu}^{\mathrm{T}}\right\rangle & \left\langle\bar{\mathbf{x}}_{m\mu}\bar{\mathbf{y}}^{\mathrm{T}}\right\rangle\\ \left\langle\bar{\mathbf{y}}\bar{\mathbf{x}}_{m\mu}^{\mathrm{T}}\right\rangle & \left\langle\bar{\mathbf{y}}\bar{\mathbf{y}}^{\mathrm{T}}\right\rangle\end{bmatrix}. \tag{S7}$$

We multiply the matrices in Eq. S6 to find:

$$\left\langle e^{i\mathbf{k}^{\mathrm{T}}(\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}})}\right\rangle = \exp\left\{i\mathbf{k}^{\mathrm{T}}\left\langle\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}}\right\rangle - \frac{1}{2}\mathbf{k}^{\mathrm{T}}\left\langle\left(\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}}\right)\left(\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}}\right)^{\mathrm{T}}\right\rangle\mathbf{k}\right\} \tag{S8}$$

which, after assuming $\langle\bar{\mathbf{x}}_{m\mu}-\bar{\mathbf{y}}\rangle = 0$, yields the result in Eq. 4.

# 2 Full $g_2(\tau)$ for the examples in the text

In this section we present the complete forms for $g_2(\tau)$ for the various examples described in the accompanying manuscript. In all examples we assume that the dynamics of the tracking stage can be treated as first-order, with $\mathbf{A} = \gamma$, $\mathbf{B} = -\gamma$, $\mathbf{C} = 1$ and, as in the manuscript, we account for systematic differences between Cartesian axes by expressing each $g_2(\tau)$ as a product over independent terms for the $x$, $y$, and $z$ axes.

## 2.1 Harmonically-bound dyes

The pedagogical example of the harmonically-bound dyes provides a special case for the computation of $g_2(\tau)$ because the terms $\lambda^{\tau}_{m\mu}$ only depend on the relationship between $m$ and $\mu$ (i.e. whether they are different or not), not on the absolute value of either. In this case we may divide terms of the form $(\lambda^0_{mm})_\alpha + \frac{1}{4}w^2_\alpha + \sigma^0_\alpha$ out of both the numerator and denominator of Eq. 12 to yield the relatively simple formula

$$g_2(\tau) = \frac{1}{M^2}\sum_{m,\mu=1}^{M}\prod_{\alpha\in\{x,y,z\}}\left[1-\left(\frac{\lambda^{\tau}_{m\mu}+\sigma^{\tau}_{\alpha}}{(\lambda^0_{mm})_\alpha+\frac{1}{4}w^2_\alpha+\sigma^0_\alpha}\right)^2\right]^{-1/2} - 1, \tag{S9}$$

into which we insert the dynamic parameters in Eqs. 19 and 20. We find

$$\begin{aligned} g_2(\tau)|_{\text{harmonic}} = \frac{1}{M}\prod_{\alpha\in\{x,y,z\}}&\left\{1-\left[\frac{a^2\left(e^{-\beta\tau}\pm\frac{\gamma^2_\alpha e^{-\beta\tau}-\gamma_\alpha\beta e^{-\gamma_\alpha\tau}}{N(\gamma^2_\alpha-\beta^2)}\right)+\sigma^0_\alpha e^{-\gamma_\alpha\tau}}{a^2\left(1\pm\frac{\gamma_\alpha}{N(\gamma_\alpha+\beta)}\right)+\frac{1}{4}w^2_\alpha+\sigma^0_\alpha}\right]^2\right\}^{-1/2} \\ &+\frac{M-1}{M}\prod_{\alpha\in\{x,y,z\}}\left\{1-\left[\frac{a^2\left(\pm\frac{\gamma^2_\alpha e^{-\beta\tau}-\gamma_\alpha\beta e^{-\gamma_\alpha\tau}}{N(\gamma^2_\alpha-\beta^2)}\right)+\sigma^0_\alpha e^{-\gamma_\alpha\tau}}{a^2\left(1\pm\frac{\gamma_\alpha}{N(\gamma_\alpha+\beta)}\right)+\frac{1}{4}w^2_\alpha+\sigma^0_\alpha}\right]^2\right\}^{-1/2} - 1, \end{aligned} \tag{S10}$$

where $\pm$ distinguishes between experiments in which the tracking and tFCS dyes are different (+) and the same (-).

In the text we describe how the limit $M, N \to \infty$ corresponds to the motion of a solid particle; that limit follows easily from Eq. S10,

$$g_2(\tau)|_{\text{solid}} = \left\{\left[1-\left(\frac{\sigma^0_x e^{-\gamma_x\tau}}{a^2+\frac{1}{4}w^2_x+\sigma^0_x}\right)^2\right]\left[1-\left(\frac{\sigma^0_y e^{-\gamma_y\tau}}{a^2+\frac{1}{4}w^2_y+\sigma^0_y}\right)^2\right]\left[1-\left(\frac{\sigma^0_z e^{-\gamma_z\tau}}{a^2+\frac{1}{4}w^2_z+\sigma^0_z}\right)^2\right]\right\}^{-1/2} - 1 \tag{S11}$$

## 2.2 $\lambda_{m\mu}$ for a general dye-labeled polymer

For a set of probe dyes located at the contour positions $\{u_m\}$, and a set of tracking dyes located at $\{v_n\}$, we use the equations from the text to compute the dynamic parameter

$$\lambda^{\tau}_{m\mu}\big|_{\mathrm{RZ}/M/N} = \frac{2r_0^2}{\pi^2}\sum_{q=1}^{\infty}\frac{1}{q^2}\left\{e^{-\tau/\tau_q}\left[c_q^m c_q^\mu - \frac{\gamma\tau_q s_q}{\gamma^2\tau_q^2-1}\left[(\gamma\tau_q+1)c_q^m + (\gamma\tau_q-1)c_q^\mu - \gamma\tau_q s_q\right]\right] + e^{-\gamma\tau}\frac{\gamma\tau_q s_q}{\gamma^2\tau_q^2-1}\left[2c_q^m - s_q\right]\right\}, \tag{S12}$$

where we have used shorthand notation to indicate cosine terms and sums of cosine terms according to

$$c_q^m \equiv \cos\left[\frac{q\pi u_m}{L}\right], s_q \equiv \frac{1}{N}\sum_{n=1}^{N}\cos\left[\frac{q\pi v_n}{L}\right]. \tag{S13}$$

These equations make it possible to calculate $g_2(\tau)$ for a Rouse/Zimm polymer with any combination of tracking and tFCS dye positions; we do not provide that formula explicitly here because of its extreme complexity and because we do not use it in the accompanying manuscript.

## 2.3 Rouse/Zimm polymer with 1 terminal tFCS dye and dense tracking dyes

The typical assumption regarding intercalating dyes is that they distribute themselves essentially uniformly along the backbone of genomic DNA molecules. We can compute the statistics of $s_q$ under this uniform assumption, finding $\langle s_q\rangle = 0$ and $\mathrm{Var}[s_q] = (2N)^{-1}$. Therefore, when $N$ is large we can substitute $s_q \approx 0$, indicating that the intramolecular motion does not influence the dynamics of the tracking system, in order to simplify Eq. S12:

$$\lambda^{\tau}_{m\mu}\big|_{RZ/M/\infty} = \frac{2r_0^2}{\pi^2}\sum_{q=1}^{\infty}\frac{1}{q^2}e^{-\tau/\tau_q}\cos\left[\frac{q\pi u_m}{L}\right]\cos\left[\frac{q\pi u_\mu}{L}\right]. \tag{S14}$$

For the specific example with a single probe dye on one molecular terminus, the cosine terms are equal to 1. Assuming simple first-order dynamics for the stage response, the result is

$$g_2(\tau)\big|_{\mathrm{RZ/terminal/dense}} = \prod_{\alpha\in\{x,y,z\}}\left[1 - \left(\frac{\frac{2r_0^2}{\pi^2}\sum_{q=1}^{\infty}\frac{1}{q^2}e^{-\tau/\tau_q} + \sigma_\alpha^0 e^{-\gamma_\alpha\tau}}{\frac{1}{3}r_0^2 + \frac{1}{4}w_\alpha^2 + \sigma_\alpha^0}\right)^2\right]^{-1/2} - 1. \tag{S15}$$

## 2.4 Rouse/Zimm polymer with sparse probe dyes and dense tracking dyes

In the case where there are multiple dyes in different locations along a polymer chain, the dynamics of the dyes are no longer identical because different parts of the chain move differently; for example, the ends of the chain have a larger RMS displacement from the center of mass and a slower relaxation to equilibrium than the chain midpoint. In this case the full form of Eq. 12 must be used to calculate $g_2(\tau)$.

One useful fact for simplifying our final expression comes from the fact that $\lambda^0_{mm}$ can be calculated exactly for the Rouse and Zimm polymers. It can be shown[1] that

$$\lambda^0_{mm} = \frac{2r_0^2}{\pi^2}\sum_{q=1}^{\infty}\frac{1}{q^2}\cos^2\left(\frac{q\pi u_m}{L}\right) = r_0^2\left[\left(\frac{u_m}{L} - \frac{1}{2}\right)^2 + \frac{1}{12}\right]. \tag{S16}$$

We now find the tFCS curve, using this result where needed:

$$g_2(\tau)|_{\mathrm{RZ/M/dense}} = \mathcal{N} \sum_{m,\mu=1}^{M} \prod_{\alpha\in\{x,y,z\}} \left\{ \left\{ r_0^2 \left[ \left( \frac{u_m}{L} - \frac{1}{2} \right)^2 + \frac{1}{12} \right] + \frac{1}{4} w_\alpha^2 + \sigma_\alpha^0 \right\} \left\{ r_0^2 \left[ \left( \frac{u_\mu}{L} - \frac{1}{2} \right)^2 + \frac{1}{12} \right] + \frac{1}{4} w_\alpha^2 + \sigma_\alpha^0 \right\} \right.$$
$$\left. - \left\{ \frac{2r_0^2}{\pi^2} \sum_{q=1}^{\infty} \frac{1}{q^2} e^{-\tau/\tau_q} \cos\left[ \frac{q\pi u_m}{L} \right] \cos\left[ \frac{q\pi u_\mu}{L} \right] + \sigma_\alpha^0 e^{-\gamma_\alpha \tau} \right\}^2 \right\}^{-1/2} - 1 \quad \text{(S17)}$$

with the normalization constant given by

$$\mathcal{N} = \left\{ \sum_{m=1}^{M} \prod_{\alpha\in\{x,y,z\}} \left[ r_0^2 \left( \left( \frac{u_m}{L} - \frac{1}{2} \right)^2 + \frac{1}{12} \right) + \frac{1}{4} w_\alpha^2 + \sigma_\alpha^0 \right]^{-1/2} \right\}^{-2}. \quad \text{(S18)}$$

### 2.5 Rouse/Zimm polymer with one probe dye and one tracking dye, on opposite molecular termini

In this case we may again use Eq. S9, because there is only a single tFCS dye. Inserting $u_m = 0$ and $v_n = L$ into Eq. S12, we find

$$g_2(\tau)|_{\mathrm{RZ/terminal/terminal}} =$$
$$\prod_{\alpha\in\{x,y,z\}} \left\{ 1 - \left[ \frac{\frac{2r_0^2}{\pi^2} \sum_{q \text{ odd}} \frac{1}{q^2} \frac{(4\gamma_\alpha^2\tau_q^2 - 1)e^{-\tau/\tau_q} - 3\gamma_\alpha\tau_q e^{-\gamma_\alpha\tau}}{\gamma_\alpha^2\tau_q^2 - 1} + \frac{2r_0^2}{\pi^2} \sum_{q \text{ even}} \frac{1}{q^2} \frac{\gamma_\alpha\tau_q e^{-\tau/\tau_q} - e^{-\gamma_\alpha\tau}}{\gamma_\alpha^2\tau_q^2 - 1} + \sigma_\alpha^0 e^{-\gamma_\alpha\tau}}{\frac{2r_0^2}{\pi^2} \sum_{q \text{ odd}} \frac{1}{q^2} \frac{4\gamma_\alpha\tau_q + 1}{\gamma_\alpha\tau_q + 1} + \frac{2r_0^2}{\pi^2} \sum_{q \text{ even}} \frac{1}{q^2} \frac{1}{\gamma_\alpha\tau_q + 1} + \frac{1}{4} w_\alpha^2 + \sigma_\alpha^0} \right]^2 \right\}^{-1/2} - 1 \quad \text{(S19)}$$

## 3 The effect of photobleaching on intramolecular tFCS

One concern that we have not addressed is that the fluorescent dyes attached to the tracked molecules will have a tendency to bleach over time. This bleaching will change the observed intramolecular dynamics, because the configuration of intact dyes will change, and will also result in a dimmer molecule with larger measurement noise. Depending on the labeling scheme, these effects of bleaching can affect either or both of the tracking and probe fluorescence statistics.

The relatively dense POPO-3 labels that we use cause the tracking error statistics to be dominated by center-of-mass motion rather than intramolecular motion. In this case, a large fraction of the POPO-3 dyes must be bleached in order to induce a noticeable change in the tracking statistics. Fortunately, the POPO dyes tend to be quite stable when used with low excitation intensities and in the presence of 2-mercaptoethanol. Figure S1 shows a particularly striking example, where essentially no photobleaching is observed on either POPO-1 or POPO-3 for 100s. We display the POPO-1 fluorescence directly in the figure. The POPO-3 fluorescence in these experiments is actually constant at 20kHz, maintained by a feedback loop that gradually increases the laser intensity to offset any losses due to photobleaching. We can therefore infer the extent of POPO-3 photobleaching from the laser intensity required to maintain this count rate. It is clear from the figure that photobleaching is a minimal concern in our experiments.

Of course, photobleaching may be a challenge in future experiments. It is possible to build photobleaching into our results from the accompanying manuscript, at least in an approximate manner. For example, we could include probe dye bleaching by making the $b_m$ into random variables that equal 1 at time zero and jump to 0 (representing bleaching) at some later time, with the probability of a jump at time $t$ dependent on the excitation intensity of the dye at that time. Such an approach would be difficult to solve explicitly, and we believe that such considerations are beyond the scope of this paper, so we will not consider them in any detail here.

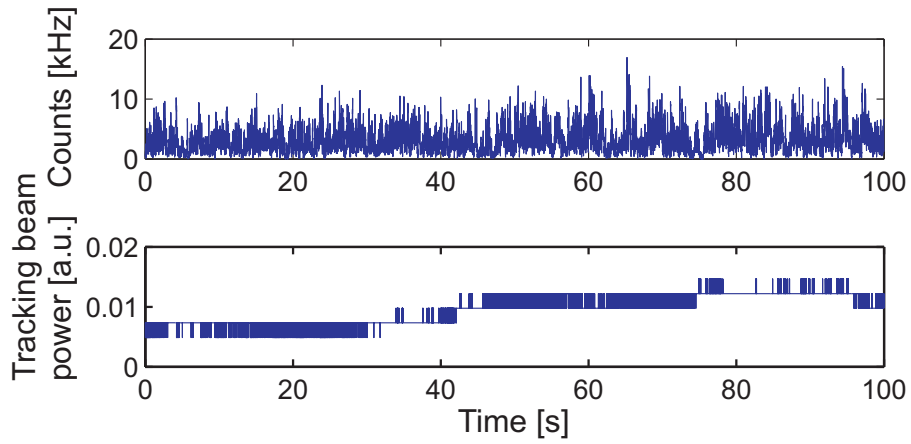


Figure S1: Fluorescence data from POPO-1/POPO-3 DNA conjugate with a prepared POPO-1 density of 1 dye per 24000 bp. Top: POPO-1 fluorescence. Bottom: tracking laser beam power required to achieve 20kHz count rate.

# 4 Combining fluorescence data from sparsely labeled molecules

In our experiments on DNA labeled sparsely with intercalating dyes, we make prolonged fluorescence measurements on individual molecules in order to infer the statistics of the molecules' motion. However, as we demonstrate, there are systematic differences between the fluorescence statistics of different molecules that are consistent with the expected differences due to random label configurations. There are two consequences of this fact:

1. It is impossible to precisely interpret any individual fluorescence trajectory, because we cannot know the molecule's dye configuration. In order to properly make inferences regarding the underlying dynamics of the DNA, we must combine data from many molecules in order to average away this configuration uncertainty.

2. Differences in the mean fluorescence rates between individual molecules result in differences in the measurement noise between those molecules, with brighter molecules producing less noisy tFCS curves. Fluorescence data can therefore only be combined by heeding proper care to these systematic complications.

In this section, we determine the appropriate means for combining fluorescence data from multiple molecules with random label configurations.

We initially process our data by computing $g_2(\tau)$ for each trajectory, yielding a set of curves that we denote $\{\hat{g}_i(\tau)\}$. Each molecule has its own true tFCS curve $g_i(\tau)$, which differs from $\hat{g}_i(\tau)$ due to the photon-counting noise inherent in optical measurements. Furthermore, there exists a single true underlying tFCS curve $g^*(\tau)$ for the DNA molecule that is equal to the ensemble average of $g_i(\tau)$ over all possible label configurations. The goal of our experiment is to determine $g^*(\tau)$ from the measured $\hat{g}_i(\tau)$. We define the measurement noise $\xi_i(\tau) = \hat{g}_i(\tau) - g_i(\tau)$, and the configuration noise $\eta_i(\tau) = g_i(\tau) - g^*(\tau)$. The measured tFCS curves can be expressed in terms of the underlying $g^*(\tau)$ and these noises:

$$\hat{g}_i(\tau) = g^*(\tau) + \eta_i(\tau) + \xi_i(\tau). \tag{S20}$$

In order to determine the appropriate rule for combining data, we require the distributions of the noise terms $\eta_i$ and $\xi_i$. These quantities are very complicated, so we will not attempt to determine their distributions

rigorously. We instead accept approximate statistics for $\eta_i(\tau)$ and $\xi_i(\tau)$ that, although inexact, will provide a great improvement over more naive data-combining approaches.

The measurement noise for a fluorescence process with a constant rate can be reliably estimated by a Gaussian, with statistics

$$\langle \xi_i(\tau) \rangle = 0 \tag{S21}$$

$$\mathrm{Var}[\xi_i(\tau)] \equiv f(\Gamma_i, T_i, \tau) = \Gamma_i^2 T_i \Delta_\tau (1 + 2\Gamma_i \Delta_\tau) \tag{S22}$$

where $\Gamma_i$ is the mean fluorescence rate of molecule $i$, $T_i$ is the duration of the $i^{\text{th}}$ molecule's fluorescence trajectory, and $\Delta_\tau$ is the width of the correlation bin at time $\tau$[2]. While the fluctuating rate in tFCS experiments leads to substantially more complicated statistics[3], this expression provides a first-order estimate of the noise in our measurements.

The statistics of $\eta_i$ are extremely difficult to calculate exactly because any number of dyes may be present at any location on the molecule, and they cannot be Gaussian because $g_2(\tau) \geq 0$ for all label configurations according to Eq. S18. We probe the distribution of $\eta_i(\tau)$ using Monte Carlo simulations, in which we assign random dye numbers and positions according to the mean dye density in the experiment and compute the resulting tFCS curves. We find that the distribution of $\eta_i(\tau)$ is unimodal with a local maximum and a long positive tail. While not a Gaussian, $\eta_i(\tau)$ is also not horribly unlike a Gaussian and so we are willing to accept a Gaussian approximation due to its analytic simplicity. We will use the statistics

$$\langle \eta_i(\tau) \rangle = 0 \tag{S23}$$

$$\mathrm{Var}[\eta_i(\tau)] \equiv \kappa_\tau^2, \tag{S24}$$

the first of which follows by definition and the second of which we may compute numerically.

We will choose our estimate $h(\tau)$ for $g^*(\tau)$ via maximum-likelihood estimation. Under the Gaussian assumptions, the log-likelihood associated with a function $h(\tau)$ given the measurements $\{\hat{g}_i(\tau)\}$ is given by

$$\ell\left(\{\hat{g}_i(\tau)\}\right) = -\sum_i \frac{\left(\hat{g}_i(\tau) - h(\tau)\right)^2}{2(\kappa_\tau^2 + f(\Gamma_i, T_i, \tau))}, \tag{S25}$$

which is maximized by choosing

$$h(\tau) = \left( \sum_i \frac{1}{\kappa_\tau^2 + f(\Gamma_i, T_i, \tau)} \right)^{-1} \sum_i \frac{1}{\kappa_\tau^2 + f(\Gamma_i, T_i, \tau)} \, \hat{g}_i(\tau). \tag{S26}$$

As we may have expected this result is simply a weighted average, and the weights that we calculated provide an approximate accounting for the relationship between the measurement noise and the systematic configuration uncertainty. The data in Fig. 4 of the accompanying manuscript for the POPO-1/POPO-3 conjugate is this $h(\tau)$.

# 5 Related Techniques

Intramolecular tFCS is the most recent of several techniques that have been used for similar purposes for decades. Here we discuss the relationships between these techniques.

## 5.1 Dynamic light scattering and FCS

In dynamic light scattering (DLS), the most traditional method for studying the dynamics of polymers, a beam of light illuminates a sample containing the polymer and the scattered fraction of the beam is measured by a detector at a fixed angle from the sample. As the path lengths between the scattering centers and the detector change due to motion of the sample molecules, the scattered amplitude changes due to constructive

or destructive interference. The autocorrelation function of the scattered amplitude is given by the *dynamic structure factor*

$$S(\mathbf{q},\tau) = \frac{1}{M}\sum_{m,\mu=1} b_m b_\mu \left\langle e^{i\mathbf{q}^T(\mathbf{x}_\mu^{t+\tau}-\mathbf{x}_m^t)}\right\rangle, \tag{S27}$$

where $M$ is the number of scattering centers, $b_m$ are scattering amplitudes, $\mathbf{q}$ is the wave vector of the scattered light and the $\mathbf{x}_m^t$ are the time-dependent scattering center positions.

For many years DLS was an important tool for determining the sizes and diffusion coefficients of polymers; however, it was always limited by the relatively small and homogeneous scattering cross-sections of the backbones of most polymers. Distinguishing characteristics of the intramolecular fluctuations were impossible to resolve because measurements contained superpositions of signals from too many nearly uncorrelated scattering centers. As a consequence, DLS was most sensitive only to the lowest-order modes of the molecules' motion.

Some success was had in improving specificity using isotopic labeling and neutron scattering in place of light scattering[4], but ultimately little was learned regarding the intramolecular dynamics of polymers until the introduction of intramolecular fluorescence methods. Fluorescence solves both the sensitivity and specificity problems of DLS, since dyes have very high absorption cross-sections and can often be located in very specific positions on the polymer chain. The statistics of FCS have been shown to be closely related to those of DLS, since $g_2(\tau) \propto \int \mathrm{d}^3\mathbf{q}\, S(\mathbf{q},\tau)\tilde{\phi}(\mathbf{q})\tilde{\phi}(-\mathbf{q})$[5], where $\tilde{\phi}$ is the Fourier-transformed beam intensity as defined earlier.

The interesting feature of stationary FCS evident from the dynamic structure factor is its dependence only on the position *difference*, $\mathbf{x}_\mu^{t+\tau}-\mathbf{x}_m^t$, and not on the absolute positions of the dyes. This occurs because the dye positions are not correlated with the laser position, so that when averages of the form $\left\langle e^{i\bar{\mathbf{k}}^{\mathrm{T}}\bar{\mathbf{x}}_{m\mu}}\right\rangle$ are computed, the uniform dye distribution contributes a term $\delta^3(\mathbf{k}+\mathbf{k}')$ as shown in [5]. By contrast, in tFCS the tracking system induces correlations between the laser and dye positions, so this delta function is generalized to a Gaussian with variance given by the inverse of the tracking error variance. This is quite a significant difference, accounting for the different functional forms of the tFCS and stationary FCS curves.

## 5.2 Camera-based correlation measurements

A new type of correlation measurement was proposed by Cohen and Moerner in their measurements on $\lambda$-phage DNA immobilized in an electro-osmotic trap[6]. They imaged the fluorescence from their DNA molecules onto a CCD camera, yielding a two-dimensional time-dependent intensity map $I(\mathbf{x},t)$ with a time resolution of 4.5ms, and so were able to measure spatially sensitive correlation functions such as $C(\mathbf{x}_1,\mathbf{x}_2,\tau) = \langle I(\mathbf{x}_1,t)I(\mathbf{x}_2,t+\tau)\rangle$ .We can now derive the predicted curves for $C(\mathbf{x}_1,\mathbf{x}_2,\tau)$ by following the general approach used to derive the tFCS statistics in the accompanying manuscript. For simplicity, we treat the image as continuous rather than as composed of pixels; furthermore we assume that the depth of field of the imaging system is much larger than the tracked molecule, so that each dye on the molecule produces an identically shaped spot in the image. We let $\psi(\mathbf{x})$ be the point-spread function of the microscope, which we will assume to be Gaussian with waist $w_\psi$. The detected image at time $t$ is

$$I(\mathbf{x},t) = \sum_{m=1}^{M} b_m \int \frac{\mathrm{d}^2\mathbf{k}}{(2\pi)^2} e^{-i\mathbf{k}^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_m)}\tilde{\psi}(\mathbf{k}), \tag{S28}$$

from which $C(\mathbf{x}_1,\mathbf{x}_2,\tau)$ and the lower-dimensional curves that are fit to the data in [6], $G(\tau) \equiv \int \mathrm{d}^2\mathbf{x}\, C(\mathbf{x},\mathbf{x},\tau)$ and $\tilde{S}(\mathbf{q},t) \equiv \int \mathrm{d}^2\mathbf{x}\, C(\mathbf{x},\mathbf{x}+\mathbf{q},\tau)$, are easily found. If, as the authors do, we assume that the feedback system cancels the translational Brownian motion without error, we have

$$G(\tau) = \sum_{m,\mu=1}^{M} \frac{\pi^{-1} b_m b_\mu}{w_\psi^2 + \frac{2}{3}\left\langle |\mathbf{r}_\mu^{t+\tau} - \mathbf{r}_m^t|^2\right\rangle}, \tag{S29}$$

where, as in earlier sections, we have assumed $x_m^t$ is Gaussian and that there is no coupling between orthogonal Cartesian axes. This result is identical (to within a normalization constant) to the stationary FCS curve with a diffusion coefficient of zero. A similar calculation yields $\tilde{S}(\mathbf{q},\tau) = S_0(\mathbf{q},\tau)\tilde{\phi}(\mathbf{q})^2$, where $S_0(\mathbf{q},\tau)$ is the dynamic structure factor from Eq. S28 but for a molecule with zero diffusion coefficient.

Our computed $G(\tau)$ and $\tilde{S}(\mathbf{q},\tau)$ are both different from those used for fitting in [6]. The differences are subtle, and to an extent can be attributed simply to our accounting for the nonzero width of the point spread function of the microscope. The ease with which we were able to make this correction is a testament to the strength of the approach to intramolecular tFCS calculations that we have developed in this paper. In fact, yet another simple generalization is all that is required to further account for trapping inaccuracy due to localization noise in spatially resolved correlation measurements.

# References


[1] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, San Diego, CA, 2000.

[2] Andrew J. Berglund. *Feedback Control of Brownian Motion for Single-Particle Fluorescence Spectroscopy*. PhD thesis, California Institute of Technology, 2006.

[3] Dennis E. Koppel. Statistical accuracy in fluorescence correlation spectroscopy. *Phys. Rev. A*, 10:1938–1945, 1974.

[4] M. Daoud, J. P. Cotton, B. Farnoux, G. Jannink, G. Sarma, H. Benoit, R. Duplessix, C. Picot, and P. G. de Gennes. Solutions of flexible polymers. neutron experiments and interpretation. *Macromolecules*, 8 (6):804–818, 1975.

[5] J. Rička and Th. Binkert. Direct measurement of a distinct correlation function by fluorescence cross correlation. *Phys. Rev. A*, 39:2646–2652, 1989.

[6] Adam E. Cohen and W. E. Moerner. Internal mechanical response of a polymer in solution. *Phys. Rev. Lett.*, 98:116001, 2007.