Power spectrum and Allan variance methods for calibrating single-molecule video-tracking instruments

Bob M. Lansdorp and Omar A. Saleh

Citation: Rev. Sci. Instrum. 83, 025115 (2012); doi: 10.1063/1.3687431
View online: http://dx.doi.org/10.1063/1.3687431
View Table of Contents: http://aip.scitation.org/toc/rsi/83/2
Published by the American Institute of Physics
Power spectrum and Allan variance methods for calibrating single-molecule video-tracking instruments

Bob M. Landoor¹ and Omar A. Saleh², a)

¹Materials Department, University of California Santa Barbara, Santa Barbara, California 93106, USA
²Materials Department and Biomolecular Science and Engineering Program, University of California Santa Barbara, Santa Barbara, California 93106, USA

(Received 14 October 2011; accepted 2 February 2012; published online 29 February 2012)

Single-molecule manipulation instruments, such as optical traps and magnetic tweezers, frequently use video tracking to measure the position of a force-generating probe. The instruments are calibrated by comparing the measured probe motion to a model of Brownian motion in a harmonic potential well; the results of calibration are estimates of the probe drag, $\alpha$, and spring constant, $\kappa$. Here, we present both time- and frequency-domain methods to accurately and precisely extract $\alpha$ and $\kappa$ from the probe trajectory. In the frequency domain, we discuss methods to estimate the power spectral density (PSD) from data (including windowing and blocking), and we derive an analytical formula for the PSD which accounts both for aliasing and the filtering intrinsic to video tracking. In the time domain, we focus on the Allan variance (AV): we present a theoretical equation for the AV relevant to typical single-molecule setups and discuss the optimal manner for computing the AV from experimental data using octave-sampled overlapping bins. We show that, when using maximum-likelihood methods to fit to the data, both the PSD and AV approaches can extract $\alpha$ and $\kappa$ in an unbiased and low-error manner, though the AV approach is simpler and more robust. © 2012 American Institute of Physics. [http://dx.doi.org/10.1063/1.3687431]

I. INTRODUCTION

Single-molecule manipulation (SMM) instruments, such as the magnetic tweezer (MT), optical trap (OT), or atomic force microscope (AFM), measure properties of single molecules by tethering the molecule to a probe (i.e., a magnetic bead in the MT, a dielectric sphere in the OT, or a cantilever in the AFM), applying force to the probe, and evaluating the molecule’s response by measuring the probe’s position (Fig. 1). Performing a SMM experiment with accuracy and precision requires calibration of the applied force.

While approximations to this expression have been made before,⁵ our expression is exact and, to the best of our knowledge, novel. We show that finite-duration (spectral leakage) effects can be removed by calculating the windowed, blocked PSD, $\tilde{P}_{b,w}(f)$, using established discrete Fourier transform (DFT) algorithms.⁴, 6 Finally, following Nørrelykke and Flyvbjerg,⁴ we show that bias-free fitting of $P_{A,B}$ to $\tilde{P}_{b,w}(f)$ can be achieved using a maximum-likelihood estimate (MLE) algorithm.

In this article, we present and compare two approaches to low-error, bias-free $\alpha$ and $\kappa$ estimation that remove or sidestep each of the three difficulties. Both strategies are formulated for a particular instrumental response function corresponding to zero-dead-time video-based tracking of probe position; that is, tracking by imaging the probe at a frame rate $f_s$ with the shutter open for a time $\tau = 1/f_s$. Our approaches remove all three sources of bias, and so using either time-domain or frequency-domain analysis procedures; this contrasts with prior work that focused on only some of the sources of bias, and used only frequency-domain methods.¹–⁴

In the frequency domain, we derive a closed-form expression for the measured probe PSD, $P_{A,B}(f)$, that accounts for both aliasing and the video-tracking instrument response. While approximations to this expression have been made before,⁵ our expression is exact and, to the best of our knowledge, novel. We show that finite-duration (spectral leakage) effects can be removed by calculating the windowed, blocked PSD, $\tilde{P}_{b,w}(f)$, using established discrete Fourier transform (DFT) algorithms.⁴, 6 Finally, following Nørrelykke and Flyvbjerg,⁴ we show that bias-free fitting of $P_{A,B}$ to $\tilde{P}_{b,w}(f)$ can be achieved using a maximum-likelihood estimate (MLE) algorithm.

¹Electronic mail: saleh@engineering.ucsb.edu.
Our second strategy is based on a time-domain analysis, the Allan variance (AV). The AV is a two-sample difference measurement that is well-known in the frequency-stability literature and has recently been introduced to the SMM community. Since the AV is a time-domain analysis, neither aliasing nor spectral leakage occur. Further, the calculation of Allan variance intrinsically assumes a finite averaging time that exactly matches the video-tracking instrument function, so no additional correction for the instrument needs to be applied. Thus, AV is well-suited to the problem of determining \( \alpha \) and \( \kappa \), since no corrections must be made for the instrument response, aliasing or spectral leakage. However, using the AV is not completely free of subtleties: we show that obtaining low-error, bias-free parameter estimates requires that the AV is estimated from data using bins that fully overlap, and that range in length by factors of 2. Finally, we point out that biased fitting is also an issue with the AV due to the non-normally distributed noise in AV estimates, and estimating the noise distribution is more involved for the AV than the PSD. To overcome this, we extend the Nørrelykke–Flybjerg strategies to the Allan variance, and show that unbiased parameter estimates can be found from an MLE approach that accounts for the actual distribution of Allan variance estimates.

Upon comparing the two strategies, we find that the AV and the PSD, when applied in an optimal manner, can both estimate \( \alpha \) and \( \kappa \) with zero-bias and with nearly-identical error. However, use of the AV is conceptually simpler, since fewer corrections need to be applied, and more practical, since the AV compresses the data more efficiently than the PSD. Furthermore, AV does not require the choice of a data compression (blocking) parameter; thus, the AV calculation has less indeterminacy.

II. POWER SPECTRAL DENSITY

A. Langevin dynamics and the physical PSD

For micron-scale probe beads trapped by optical or magnetic forces in water, inertial effects can be ignored up to roughly megahertz frequencies, and thus are insignificant at typical video-tracking frequencies of approximately 100 Hz. Thus, the bead position \( x(t) \) vs. time \( t \) obeys the overdamped Langevin equation of motion:

\[
\kappa x + \alpha \dot{x} = F_L,
\]

where \( \kappa \) is the spring constant of the system and \( \alpha = 6\pi \eta r \) is the dissipation of the spherical bead of radius \( r \) in a solution of viscosity \( \eta \), and \( F_L \) is the Langevin force, a time-varying stochastic force that, in thermal equilibrium, obeys the fluctuation-dissipation relation \( \langle F_L(t + \tau)F_L(t) \rangle = 2\kappa k_B T \delta(t) \) given thermal energy \( k_B T \). The power spectral density of bead motion, \( P(f) \), where \( f \) is the frequency in hertz, can be calculated by taking the magnitude of the Fourier transform of Eq. (1) to obtain

\[
P(f) = \frac{k_B T}{2\pi^2 \alpha(f_c^2 + f^2)},
\]

where \( f_c = \kappa/2\pi\alpha \). We use two-sided power spectra throughout so that integrating \( P(f) \) on the range \((-\infty, \infty)\) results in the equipartition result \( \langle x^2 \rangle = k_B T/\kappa \).

B. An analytical expression for the PSD function, accounting for aliasing and the instrument function

The PSD of Eq. (2) corresponds to the true physical motion of the bead; any actual measurement \( \hat{P}(f) \) will deviate from \( P(f) \) due to the distorting effects of aliasing and the instrumentation. In video-tracking experiments in which the bead position is measured by a camera with a shutter time \( \tau_s \), any motion of the bead that occurs on timescales faster than \( \tau_s \) is diminished by the measurement; i.e. video tracking introduces a low-pass filter. The low-pass filtering effect of the camera can be quantitatively accounted for: when the illumination is constant over \( \tau_s \), the measured position \( x_i \) from frame \( i \) at time \( t_i \) corresponds to the average over the interval \( (t_i - \tau_s/2, t_i + \tau_s/2) \):

\[
x_i = \frac{1}{\tau_s} \int_{t_i - \tau_s/2}^{t_i + \tau_s/2} x(t) \, dt = [x * \Pi_{\tau_s}](t_i),
\]

which we have written in the second line as the convolution of \( x(t) \) with the boxcar function \( \Pi_{\tau_s}(t) \). We define \( \Pi_{\tau_s}(t') \) to have value 1/\( \tau_s \) for \( |t'| < \tau_s/2 \), and to be zero otherwise. The frequency-space equivalent of convolution is simply multiplication of the Fourier transforms of the two convolved functions. Thus, since the Fourier transform of \( \Pi_{\tau_s}(t) \) is \( \sin(\pi f \tau_s)/(\pi f \tau_s) \), the PSD of bead motion including the boxcar response, \( P_B(f) \), is

\[
P_B(f) = P(f)H(f),
\]
where the instrument response $H(f)$ is

$$H(f) = \frac{\sin^2(\pi f \tau)}{(\pi f \tau)^2}.$$  

(5)

$P_B(f)$ only accounts for the effect of instrumental low-pass filtering in distorting the measured PSD; it does not account for the further distortions created by aliasing. The aliasing effect causes the PSD at a particular positive frequency $f'$, with $0 < f < f_s/2$, to contain the summed power at all frequencies in the series $|f + n f_s|$, where $n \in \{-1, 0, 1, \ldots\}$. Thus, the measured PSD accounting for both aliasing and boxcar filtering is given by

$$P_{A,B}(f) = \sum_{n=-\infty}^{\infty} P_B(|f + n f_s|)$$

(6)

with $P_B$ given by Eqs. (4) and (5).

In the special case that the sampling frequency is the inverse of the shutter time, $f_s = 1/\tau_s$, the sum in Eq. (6) can be performed analytically:

$$P_{A,B}(f) = \frac{2 k_B T \alpha}{\kappa^3} \left( \frac{2 \alpha f_s}{\kappa} \right) \left( \frac{\frac{\pi}{f_s}}{\kappa} \sinh \left( \frac{\kappa}{ \alpha f_s} \right) \sinh \left( \frac{\alpha f_s}{\kappa} \right) - \cosh \left( \frac{\kappa}{ \alpha f_s} \right) \cosh \left( \frac{\alpha f_s}{\kappa} \right) \right).$$

(7)

Equation (7) can be verified in multiple ways. First, we can use $P_{A,B}$ and the Wiener–Kinchin relation to find the variance of the measured signal in the time domain; the result matches previously predicted results for boxcar-averaged probe positions. 

Second, we can compare $P_{A,B}$ with $P_A(f)$, the PSD expression that accounts only for aliasing (and not instrumental low-pass filtering):

$$P_A(f) = \frac{k_B T}{\kappa f_s} \left( \frac{\sinh \left( \frac{\kappa}{\alpha f_s} \right)}{\cosh \left( \frac{\kappa}{\alpha f_s} \right) - \cos \left( \frac{2\pi f}{f_s} \right)} \right).$$

(8)

$P_A$ corresponds to the PSD measured by an instrument with sampling frequency $f_s$, but with an infinitely fast measurement interval (i.e., the boxcar measurement function, $\Pi_\tau$, is replaced by a delta function). We compare $P_A(f)$ and $P_{A,B}(f)$ to experimental data in Fig. 2. The data are clearly best-described by $P_{A,B}$, lending confidence to the expression in Eq. (7).

We further verified the prediction of Eq. (7) through simulation of probe trajectories in the time domain, computation of the PSDs of the simulated trajectories, and comparison to $P_{A,B}$. We simulated discrete particle trajectories using an algorithm based on the solution to the Smoluchowski equation for a Brownian particle in a harmonic well. To account for the averaging effect of the boxcar filter, we simulated discrete points with a short time step of $\delta = \tau$, then averaged together consecutive clusters of $\tau/\delta$ points, thus arriving at a set of discrete positions $x_i$ spaced by time $\tau$. We did this for multiple time steps $\delta = \tau m$, with $m = 1, 2, 4, 8, 16$. Figure 3 shows comparisons of the simulation with both $P_{A,B}$ and $P_A$ for both $f_\delta < f_s$ and $f_\delta \approx f_s$. For $m = 1$, there is no averaging of the time-domain signal, so the simulated PSDs match $P_A$, as expected. As $m$ increases, the simulated PSDs converge on $P_{A,B}$.
This directly indicates that Eq. (7) correctly accounts for the camera’s averaging effect on the measured probe position.

C. Estimating the PSD from a time series of data

Spectral leakage and windowing: Given a set of \( N \) discrete measurements of probe position vs. time, an estimate, \( \tilde{P}(f) \), of the PSD can be made from the discrete Fourier transform. Directly comparing \( \tilde{P}(f) \) to \( P_{A,B}(f) \), (Eq. (7)) leads to difficulties due to the finite duration of the measurement: the frequency-domain analysis of the Langevin equation implicitly assumes that the trajectory is infinitely long, while any measurement will naturally be of finite duration. This truncation leads to the phenomenon of spectral leakage, in which some power present in the signal at one frequency shows up at other frequencies. For white noise, spectral leakage has no effect, as neighboring frequencies leak equal and opposite power to each other, resulting in zero net change. However, for limited bandwidth signals (such as the Lorentzian, where the power is focused in the range \( f < f_c \)), spectral leakage leads to PSD distortions.

Spectral leakage can be minimized by “windowing” the input data.\(^4\) In particular, given \( N \) measurements of probe position, \( x_j, j \in \{0, 1, 2, \ldots, N - 1\} \), at sampling rate \( f_s \), the PSD can be estimated from:

\[
\tilde{P}_{w,N} = \frac{1}{f_s N} \sum_{j=0}^{N-1} w_j x_j e^{-2\pi i jk/N},
\]

where \( w_j \) is the windowing function, which can be chosen from a range of standard functions, all of which taper to zero at the edges of the data trace.\(^18\) Here, we use a Hann window, defined by

\[
w_j = \frac{\sqrt{8}}{3} \sin^2 \left( \frac{\pi j}{N} \right).
\]

Application of the window reduces the total power in the PSD in a frequency-independent manner; this is corrected by the leading factor of \( \sqrt{8}/3 \).

Blocking: In the absence of stochastic noise, the windowed PSD \( \tilde{P}_{w,N} \) would converge on the analytical aliased filtered Lorentzian \( P_{A,B} \). However, a feature of the discrete Fourier transform is that increasing \( N \) does not decrease the stochastic noise; instead, it increases the number and density of estimates in frequency space. Moreover, while the noise in \( x_j \) is normally distributed, the stochastic noise in a single power spectrum is distributed exponentially, complicating least-squares fitting routines that assume normally distributed noise.

The stochastic noise can be decreased and the noise distribution made more normal by compressing the data. Typically, this is done by blocking:\(^1\) the data trace is separated into \( b \) blocks of length \( m \), the PSD of each block, \( \tilde{P}_b \), is calculated, and averaged with all others to give the blocked PSD \( \tilde{P}_b = \sum \tilde{P}_m/b \). Blocking should be adjusted when also windowing, as application of the window minimizes the contribution of data near the termini of each block. All of the data can be efficiently utilized by following Welch’s method;\(^6\) instead of blocking into non-overlapping bins, \( b = Nm \), the data trace is blocked into half-overlapping bins (Fig. 4), giving \( b = 2Nm - 1 \). The windowed PSD of each block, \( \tilde{P}_{w,m} \), is calculated, and averaged with all others to give the final blocked, windowed PSD estimate \( \tilde{P}_{w,b} = \sum \tilde{P}_{w,m}/b \). This insures efficient exploitation of all of the available data while keeping each bin statistically independent of its neighbors.

The noise distribution in a PSD: Recently, Nørrelykke and Flyvbjerg\(^1\) have pointed out that, even after blocking, the distribution of noise in the estimated PSD is typically still far from normally distributed, but that fitting can proceed in a precise manner by accounting for the actual distribution: averaging \( b \) spectra together causes the noise distribution to be the convolution of \( b \) exponential distributions, which is a Gamma distribution. In a Gamma distribution, the probability \( f \) of observing a certain value \( x \) depends on parameters \( \eta, \theta \) as

\[
f(x; \eta, \theta) = x^{\eta-1} e^{-x/\theta}/\theta^{\eta} \Gamma(\eta),
\]

where \( \Gamma(\eta) \) is the gamma function. The two free parameters are the shape parameter \( \eta \) and the scale parameter \( \theta \). For a PSD calculated from \( b \) blocks, and whose true value is given by \( P_{A,B}(f) \), the probability of observing \( \tilde{P}_{b,w} \) is proportional to \( f(\tilde{P}_{b,w}; b, P_{A,B}(f))/b \); that is, the shape is \( \eta = b \) and the scale is \( \theta = P_{A,B}(f)/b \). Explicitly using the Gamma distribution allows for the application of bias-free maximum-likelihood estimation.\(^1\) We discuss this below (Sec. IV), in conjunction with a similar discussion of fitting to the Allan variance.
III. ALLAN VARIANCE

A. Definition of the Allan variance

An alternative to frequency-domain analysis of probe fluctuations is the AV, denoted by $\sigma^2(\tau)$. The AV is a time-domain measure: it is half of the ensemble-averaged variance of the difference between two consecutive samples of position,21,22 where each sample is itself a local average of probe position (Eq. (12)). The timescale $\tau$ denotes both the time between consecutive samples, and the time over which each sample is averaged. Thus, for a probe position $x(t)$ and integer $j$, the AV is defined by

$$\sigma^2(\tau) = \frac{1}{2}(\bar{x}_{\tau,j+1} - \bar{x}_{\tau,j})^2, \tag{12}$$

where $\bar{x}_{\tau,j}$ is given by

$$\bar{x}_{\tau,j} = \frac{1}{\tau} \int_{(j-\frac{1}{2})\tau}^{(j+\frac{1}{2})\tau} x(t)\,dt \tag{13}$$

$$= [x \ast \Pi_{\tau}](j \tau). \tag{14}$$

The identical forms of Eqs. (3) and (14) emphasize that the camera measures an average probe position, exactly as the instrument function is well-approximated by the boxcar function introduced in Eq. (3).

B. An analytical expression for the AV of a damped harmonic oscillator

In analogy to Sec. II B, we derive an analytical expression for the AV for a Brownian probe in a harmonic well. By expanding the product in Eq. (12), we find that the AV is related to the variance and autocorrelation of $\bar{x}_{\tau}$:

$$\sigma^2(\tau) = \langle \bar{x}_{\tau}^2 \rangle - \langle \bar{x}_{\tau,j+1} \bar{x}_{\tau,j} \rangle, \tag{15}$$

where we have made use of the stationary nature of the process, so $\langle \bar{x}_{\tau,j+1}^2 \rangle = \langle \bar{x}_{\tau,j}^2 \rangle = \langle \bar{x}_{\tau}^2 \rangle$. The Wiener–Khinchin theorem can be used to relate the variance and autocorrelation of $\bar{x}_t$ to its PSD; in turn, given the properties of convolution, the PSD of $\bar{x}_t$ depends on the PSD of $x(t)$ and of $\Pi_{\tau}$. This allows the AV to be related to $P(f)$, giving:

$$\sigma^2(\tau) = \int_{-\infty}^{\infty} \frac{4\sin^4(\pi f \tau)P(f)}{\pi f^2} \,df. \tag{16}$$

Using Eq. (2), this integral can be performed to obtain an analytical expression for the AV of an overdamped bead in a harmonic well:

$$\sigma^2_{SMM}(\tau) = \frac{2k_BT\alpha}{\kappa^2\tau} \left(1 + \frac{2\alpha}{\kappa \tau} e^{-\frac{\tau}{2\tau}} - \frac{\alpha}{2\kappa \tau} e^{-\frac{\tau}{2\tau}} - \frac{3\alpha}{2\kappa \tau} \right). \tag{17}$$

This expression has long been known in the frequency-stability literature,23 but we have rewritten it here in terms of $\alpha$ and $\kappa$ with special application to SMM experiments. For long times, $\tau \gg \alpha/\kappa$, the AV reduces to $\sigma^2_{SMM}(\tau) \approx 2k_BT\alpha/\kappa^2\tau$, consistent with previously published approximations.10–12 The large time limit can be understood by noting that neighboring $\bar{x}_t$ values become uncorrelated for large $\tau$, so the variance arises from the standard error in determining the mean of a normally distributed value, i.e., variance $\sim 1/(\# \text{ of points}) \sim 1/\tau$. In the opposite limit, $\tau \ll \alpha/\kappa$, we find $\sigma^2(\tau) \approx 2k_B T \alpha/3\kappa$; that the variance increases linearly with $\tau$ arises from the purely diffusive nature of the probe motion at short times. Thus, a feature of the AV of a damped, Brownian harmonic oscillator is that it increases at short times and decreases at long times (Fig. 5); the location of the intervening peak can be found numerically to be $\tau_{\text{max}} \approx 1.89/\alpha/\kappa$, i.e., the peak scales with the system’s correlation time, $\alpha/\kappa$.

C. Estimating the AV from a time series of data

**Overlapping AV:** For a discrete set of measurements of probe positions $x_j$ at times $t_j = j \tau$, for $j = 1, \ldots, N$, the “standard” AV can be calculated for timescale $\tau = m\tau$, by computing the mean of successive bins of $m$ points, then taking the mean-squared difference of the resulting $N/m - 1$ neighboring pairs of bins. So, in particular, the first pair of bins corresponds to the sets of points $(1, \ldots, m)$ and $(m + 1, \ldots, 2m)$, the second pair are points $(m + 1, \ldots, 2m)$ and $(2m + 1, \ldots, 3m)$, and so on, through the $N$th point. However, the standard AV is known to not fully utilize the data.24 A superior option is the “overlapping” AV (OAV), which uses every possible $m$-bin (Fig. 4). In particular, the first OAV pair of bins is $(1, \ldots, m)$ and $(m + 1, \ldots, 2m)$, the second pair is $(2, \ldots, m + 1)$ and $(m + 2, \ldots, 2m + 1)$, and so on, through the $N$th point. The
standard and overlapping AVs are identical for \( m = 1 \) and \( m = N/2 \), but elsewhere the OAV more efficiently exploits the available data, and so provides a better estimate of the true AV. The experimentally derived OAV can be computed from \(^{29}\)

\[
\bar{\sigma}_m^2 = \frac{1}{2(N - 2m)(m \tau_c)^2} \sum_{k=1}^{N-2m} (x_{k+2m} - 2x_{k+m} + x_k)^2.
\]

We use only the OAV in calculations throughout this article.

*Octave sampling:* The OAV can be determined for any integer value of the bin size, \( m = 1, 2, \ldots, N/2 \) (see Fig. 5). However, upon estimating and plotting the OAV, it is clear that neighboring values of \( m \) have a correlated noise, which stems from the nearly-identical sets of points used to calculate the OAV for similar values of \( m \). While calculating all possible \( m \) values might be useful for visualizing the data, the noise correlations preclude most fitting algorithms, which assume independent error in the fitted points. Allan pointed out that this issue can be solved by octave sampling, \(^{36}\) i.e., only using bin sizes in powers of 2: \( m = 1, 2, 4, 8, \ldots \) up to \( m \leq N/2 \). Using bins of \( 2^m \) ensures a minimal correlation between successive estimates, and improves the results of fitting, as discussed in Sec. V and shown in Fig. 7.

*Error in the AV estimate:* At the root of the AV is the difference \( \bar{x}_{j+1} - \bar{x}_j \), which is normally distributed with mean zero. Thus, an estimate for the AV (which is a sum of squares of such differences) is Gamma distributed. The estimated blocked and windowed PSD, \( \tilde{\mathbf{P}}_{b,w} \) (Sec. II C), is also Gamma distributed, and \( \tilde{\mathbf{P}}_{b,w} \) has the same shape factor, \( \eta = b \), for all frequencies \( f_k \). However, the estimated AV does not have a constant shape factor because the number of available differences depends on the bin length \( m \); this is the major complicating factor in working with the AV relative to the PSD.

Generally, the shape factor is given by the number of exponential distributions that are convolved to form the metric (this is why \( \eta = b \) for the PSD). The standard AV for a given \( m \) is calculated from \( N/m - 1 \) differences; since each difference has a Chi-Square distribution, two of which must be convolved to get an exponential distribution, the shape factor is

\[
\eta_{AV}(m) = \frac{1}{2} \left( \frac{N}{m} - 1 \right). \quad (19)
\]

As discussed, the OAV more efficiently exploits the available data, meaning that there are more available degrees of freedom per \( m \). Thus, for the OAV, the shape factor of Eq. (19), while nearly correct, is an underestimate. The exact calculation of the available degrees of freedom for the OAV is quite involved, and beyond the scope of this paper; it has been the subject of much discussion in the literature. \(^{20,27}\) However, we have found that using \( \eta \) from the relatively simple Eq. (19) gives near-identical results to more exact, involved estimates, so we use Eq. (19) in all computations discussed below.
where \( \eta_k \) is the shape parameter for point \( k \). Eq. (20) applies to both the PSD and AV. For the PSD, the shape factor is constant for all points, and equal to the number of blocks, \( \eta_k = b \), while the true value is found from Eq. (7) as \( P_{\alpha,\kappa}(f_\tau) \). For the AV, the shape factor is point-dependent, \( \eta_k = \eta_{AV} \), where \( \eta_{AV} \) is defined in Eq. (19). The true value is found from Eq. (17) as \( \sigma^2_{\text{MM}}(\tau = k\tau_\epsilon, \alpha, \kappa) \). In either case, minimization of Eq. (20) with respect to \( (\alpha, \kappa) \) can be carried out relatively efficiently using most numerical analysis software; we use the MLE approach to generate the results described below.

V. BIAS AND ERRORS IN AV AND PSD FITS

To test and compare the AV and PSD algorithms, we performed a set of simulations of probe trajectories using a variety of input parameters and tested the ability of the algorithms to correctly extract the known parameters. In particular, we simulated video-tracking trajectories using a frequency-space algorithm, as discussed in the Appendix. All trajectories contained \( N = 4096 \) points sampled at \( f_c = 100 \) Hz. The drag was \( \alpha = 10^{-5} \) pN s/nm, as appropriate for a spherical bead of diameter 1 \( \mu \)m in water. We varied the spring constant, \( \kappa \), between \( \approx 1.4 \times 10^{-5} \) and \( \approx 6.8 \times 10^{-3} \) pN/nm; this gave a set of corner frequencies \( f_c \) that ranged from \( \approx 0.2 \) to \( \approx 100 \) Hz.

Each trajectory was processed according to Secs. II C and III C to give estimates of the PSD and AV, and fit using the MLE approach to extract the optimal parameters \( (\alpha, \kappa) \).

All numerical work was carried out in Mathematica, with the MLE minimization carried out using the FindMinimum function.

To emphasize the need for windowing in the PSD calculation, we compared the best-fit parameters from PSDs that did or did not use a Hann Window (Fig. 6) for a variety of block lengths \( m \). When windowing, we used half-overlapping samples (blocks), while the unwindowed analysis used non-overlapping samples; so, for example, for blocks of length \( m = N/4 = 1024 \), the windowed PSD was calculated from \( b = 7 \) blocks, while the unwindowed PSD used \( b = 4 \) blocks. Both methods performed similarly in estimating the spring constant, giving unbiased estimates with relatively low noise over a broad range of \( f_c \). In contrast, the use of windowing was clearly superior in the estimate of drag: the unwindowed method gave significant bias and increased stochastic error, in the estimate of \( \alpha \), particularly at lower \( f_c \). For both parameters, the stochastic error in the windowed estimates decreased as \( m \) decreased, with \( m = N/8 = 512 \) points or fewer needed to obtain the minimal error. We attribute this to end-effects: using a single block \( (m = N) \) and applying a window cause a loss of information near each terminus of the trace, which increases the error. Instead, using many overlapped blocks decreases the information lost at the termini (because the blocks are shorter), while overlapping avoids information loss in the interior of the trace.

The results of the AV analysis (Fig. 7) demonstrate the importance of octave-sampling: we compared the best-fit parameters from AVs estimated using bins of all possible lengths (all-\( \tau \)), \( m = 1, 2, 3, \ldots, N/2 \), and those estimated from octave-bins, \( m = 1, 2, 4, \ldots, N/2 \). The octave-sampled method
VI. EXPERIMENTAL RESULTS

To demonstrate the applicability of our AV fitting method to experimental data, we use a magnetic tweezer instrument to stretch double-stranded DNA molecules over a range of forces.\(^{28}\) From each particle trajectory taken at a constant force, we calculate the octave-sampled, overlapping AV, and use MLE fitting of Eq. (17) to estimate \(\alpha_x, \alpha_y, \alpha_z, \kappa_x, \kappa_y,\) and \(\kappa_z\). We then compare these parameters to those independently estimated by theories of bead drag or polymer elasticity.

In Figure 8 we plot the drag coefficients \(\alpha_s, \alpha_y, \alpha_z\) as a function of the average bead height. As the magnetic force is decreased, the bead height decreases and the drag coefficient of the bead increases due to surface effects, which can be theoretically predicted using the Faxén correction.\(^{29}\) Both \(\alpha_x\) and \(\alpha_z\) closely match the expected behavior. We consistently found that \(\alpha_y > \alpha_s\), which we attribute to the rotational constraints imposed on the magnetic beads by the applied magnetic field. At high force, the fitting routine was unable to accurately determine \(\alpha_z\), since the corner frequency of the system approached and exceeded the Nyquist frequency of our camera.

In Figure 8 we also present the experimentally derived \(\kappa_x, \kappa_y, \kappa_z\) values as a function of extension. At each measured tether extension \(L\), we can generate three independent predictions for \(\kappa_x, \kappa_y, \kappa_z\) by, first, using the worm-like chain model for double-stranded DNA (Ref. 30) to convert \(L\) into an estimate of force, \(F\). Then, we use the inverted-pendulum model for magnetic tweezers\(^{15}\) to convert \(F\) and \(L\) into expected stiffness through the relations: \(\kappa_z = \partial F/\partial z\), \(\kappa_x = F/L,\) and \(\kappa_y = F(L + R)\). For all predictions, we use a persistence length of 53 nm, consistent with previous measurements,\(^{30}\) and a contour length of 1720 nm. Although we do not correct for bead misstereing,\(^{31}\) we nonetheless find a remarkably good agreement between theory and experiment over a wide range of forces in all six measured parameters.

VII. CONCLUSION

In this article, we have presented two comprehensive and novel strategies for calibrating the spring constant \(\kappa\) and measuring the drag \(\alpha\) in a SMM instrument that uses full-frame video tracking to measure probe position. These strategies are based on the AV and PSD. In particular, given \(N\) discrete measurements of position \(x_j(t)\), with \(j \in \{0, 1, 2, \ldots, N - 1\}\), we suggest the following fitting procedure for the PSD:

1. Divide the trace into half-overlapping bins of length \(m \leq N/8\), resulting in \(b = 2N/m - 1\) bins.
2. For each bin, multiply \(x_j\) by a Hann window and calculate a windowed PSD \(\tilde{P}_{w,m}\) for each bin (Eq. (9)).
3. Average the Power spectra over all bins to obtain a single blocked and windowed \(\tilde{P}_{b,w}\).
4. Minimize the cost function in Eq. (20) using \(y_k = P_{b,w}(\alpha, \kappa, f)\) and \(\bar{y}_k = \tilde{P}_{b,w}(f_k)\) to obtain \(\alpha\) and \(\kappa\).

The corresponding recipe for calculating \(\alpha\) and \(\kappa\) using an AV fit to the same data is

1. Calculate the octave-sampled OAV \(\sigma^2_m(\tau)\) using Eq. (18).
2. Calculate the shape factor \(\eta_{AV}(m)\) from Eq. (19).
3. Minimize the cost function in Eq. (20) using \(\eta_{AV}(m), y_k = \sigma^2_{MM}(\tau, \alpha, \kappa),\) and \(\bar{y}_k = \sigma^2_{m}(\tau)\) to obtain \(\alpha\) and \(\kappa\).

We conclude by comparing the AV and PSD approaches: as can be seen from the results (Figs. 6 and 7), both the AV and
PSD algorithms are capable of unbiased, low-error parameter estimates of equal quality. Thus to differentiate between the two we focus on ease-of-use, both in terms of general complexity and, in particular, for the numerical minimization required by the MLE algorithm. For both criteria, we find that the AV algorithm outperforms the PSD: In terms of complexity, while both the AV and PSD require fitting with a MLE algorithm rather than least-squares, the PSD approach requires applying multiple corrections, particularly for instrument response, discrete sampling (aliasing), and finite duration of the data (spectral leakage). The AV is not subject to any of these three issues, as the instrument response intrinsic to video tracking exactly matches the assumptions made in calculating the AV, and the PSD finite-sampling effects that manifest in the DFT sum are simply not present in the time-domain AV calculation. Further algorithmic simplicity is achieved by the AV in that no choice of blocking number, \( b \), needs to be made; this is particularly important since we show that the choice of \( b \) has a minor, but significant, effect on the fitting results of the PSD.

Numerical minimization of the cost function (Eq. (20)) is also more straightforward with the AV because it achieves superior data compression without sacrificing accuracy or precision. The octave-sampled OAV for a data trace of \( N \) points contains \( \log_2 N \) \( N \) points, while the estimated PSD (using the maximum bin length that still reduces stochastic error) contains \( N/8 \) points. Therefore, the total number of points that must be fit grows linearly with \( N \) for the PSD, but only logarithmically with \( N \) for the AV. For example, for the \( N = 4096 \) traces simulated here, the AV contains 12 points, while the PSD contains 513 points. This means that the MLE cost function (Eq. (20)) contains only 12 terms when working with the AV, but contains 513 terms for the PSD. Because of the advantages in simplicity, speed, and robustness of fitting with the AV as compared to the PSD, we suggest that the Allan variance is the ideal calibration method for many SMM instruments.

ACKNOWLEDGMENTS

This work was supported by the NIH under Grant No. 1R21GM079581-01A1, and the NSF under Grant No. PHY-0748564. B.L. acknowledges support from an NSERC PGS-D fellowship. We thank T. M. Squires, D. B. McIntosh, and A. A. M. Saleh for helpful discussions.

APPENDIX: EFFICIENT SIMULATION OF VIDEO-TRACKING TRAJECTORIES IN FREQUENCY SPACE

The efficient simulation of probe trajectories is important as a tool to test calibration algorithms. A typical strategy to simulate particle trajectories is to recast the Langevin equation, Eq. (1), as an equation of probability flux (a Smoluchowski equation), then apply an iterative scheme to generate successive discrete time points in the trajectory. The discrete points generated in this way are not a realistic representation of data measured by a video-tracking instrument, as the method implies that the instrument has an infinitely-fast time response. A more correct approach is to generate points with a small time step, then to average groups of points together, thus approximating the averaging effect of the camera. However, this can be quite onerous as it requires the simulation of many more points than are finally desired, particularly when \( f_s \) approaches \( f_t \) (see Fig. 3).

There is an alternative approach: since the Fourier transform of a trajectory contains all the information needed to recapitulate the trajectory, it is possible to use \( P_{A,B} \) (Eq. (7)) as an efficient simulation tool. In particular, we can get an \( N \)-point trajectory by simulating \( N \) points in the frequency domain that are distributed according to \( P_{A,B} \), then applying an inverse Fourier transform. Here, we briefly describe this approach.

Consider a set of points \( x_j \) that represent a probe trajectory at \( N \) time points \( t_j = j \tau_c = j f_t \); we assume \( N \) is an even number. We define the Fourier transform of \( x_j \) to be \( \hat{x}_k \), which is a complex number and is defined for a set of \( N \) frequencies \( f_k = k f_t / N \), with \( k = -(N/2) + 1, ..., 0, 1, ..., N/2 \). All \( x_j \) are real numbers, which constrains the allowed values of \( \hat{x}_k \):

\[
\begin{align*}
\text{Re}(\hat{x}_k) &= \text{Re}(\hat{x}_{-k}), \\
\text{Im}(\hat{x}_k) &= -\text{Im}(\hat{x}_{-k}), \quad (A1) \\
\text{Im}(\hat{x}_0) &= \text{Im}(\hat{x}_{N/2}) = 0.
\end{align*}
\]

Thus, it suffices to choose \( N \) random values, one each for \( k = 0 \) and \( k = N/2 \), and two each (one imaginary and one real) for \( k = 1, ..., (N/2) - 1 \), suggesting the following algorithm:

1. Use a random number generator to generate a set of \( N \) points, \( q_j \) for \( j = 0, ..., N - 1 \), that are normally distributed with mean zero and unit variance.
2. Set the real-only values at zero frequency and the Nyquist frequency by multiplying a random number by the expected value at those frequencies:

\[
\begin{align*}
\hat{x}_0 &= P_{A,B}(0) q_0, \\
\hat{x}_{N/2} &= P_{A,B}(f_s/2) q_{N/2}. \quad (A2)
\end{align*}
\]

3. Set the remaining complex values, \( \hat{x}_k \) for \( k = 1, ..., (N/2) - 1 \), by using different random numbers for the real and imaginary components. Each is scaled by \( P_{A,B} \), but must be divided by \( \sqrt{2} \) to account for ratio between the expected value of the magnitude and each complex component:

\[
\hat{x}_k = \frac{1}{\sqrt{2}}(P_{A,B}(f_k) q_k + i P_{A,B}(f_k) q_{k+N/2}). \quad (A3)
\]

4. Apply the constraints, Eq. (A1), to find \( \hat{x}_k \) for all \( k \).
5. Calculate \( x_j \) by applying a discrete inverse Fourier transform to \( \hat{x}_k \):

\[
x_j = \sqrt{\frac{f_s}{N}} \sum_k \hat{x}_k e^{2\pi i j k / N}. \quad (A4)
\]