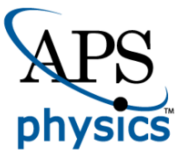

Decorrelation correction for nanoparticle tracking analysis of dilute polydisperse suspensions in bulk flow

John Hartman and Brian Kirby

Phys. Rev. E **95**, 033305 — Published 14 March 2017

DOI: 10.1103/PhysRevE.95.033305

Accurate correction of bulk flow in nanoparticle tracking analysis for dilute, polydisperse suspensions

John Hartman and Brian Kirby*
Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY 14853, USA
(Dated: February 16, 2017)

Nanoparticle tracking analysis, a multi-probe single particle tracking technique, is a widely used method to quickly determine the concentration and size distribution of colloidal particle suspensions. Many popular tools remove non-Brownian components of particle motion by subtracting the ensemble-averaged displacement at each time step, which is termed 'de-drifting'. Though critical for accurate size measurements, de-drifting is shown here to introduce significant biasing error and can fundamentally limit the dynamic range of particle size that can be measured for dilute, heterogeneous suspensions, such as biological extracellular vesicles. We report a more accurate estimate of particle mean-squared displacement, which we call Decorrelation Analysis, that accounts for correlations between individual and ensemble particle motion, which are spuriously introduced by de-drifting. Particle tracking simulation and experimental results show that this approach more accurately determines particle diameters for low concentration, polydisperse suspensions when compared with standard de-drifting techniques.

PACS numbers: 87.16.dp, 87.80Nj, 87.16.A-, 05.40.Jc
Keywords: Nanoparticle tracking analysis, Extracellular vesicles, Brownian motion

I. INTRODUCTION

The analysis of multi-probe spatial positions tracked over time provides important statistical information used in rendering super resolution microscopy images [1], investigating cellular transport mechanisms in biological systems [2, 3], characterizing local rheological properties of heterogeneous fluids, soft matter or cross-linked polymer gels [4–6], and measuring the particle size distribution (PSD) of nanoparticles in suspension [7, 8]. Empirical single particle tracking (SPT) trajectories are necessarily a superposition of thermal Brownian motion and particle localization errors. Additional sources of non-Brownian particle behavior – such as platform vibration, mechanical drift, intentional or artifactual ballistic motion, coupled motion and potential wells – confound statistical analyses that consider Brownian motion only [9–13]. Therefore, approaches that extract Brownian motion from a complicated mobility landscape are necessary for reliable analyses in a variety of contexts.

When deviations from Brownian behavior are well understood, solving in parallel for both Brownian and non-Brownian motion statistics often produces optimally accurate estimates [14–16]. These optimally accurate estimators are generally derived from parametric models of non-Brownian motion or other errors [17–19]. For example, random motion overlayed by uniform, constant directed motion leads to first- and second-order terms for mean-squared displacement (MSD), which can be determined simultaneously using maximum likelihood estimation [20]. Orthogonal experiments that reduce the dimension of the parametric model are also effective, such as using immobile beads to estimate localization error [21, 22]. Other models rely on particle interactions and thereby forego any requirement to consider the underlying motion present in the bulk fluid [23]. However, accuracy in parameter estimation does not necessarily mean that the model is appropriate for the given data set. For example, unbiased covariance-based estimators can outperform 'optimal' maximum likelihood estimators when particle trajectories span relatively few time steps [24].

The most widely used approaches serially extract Brownian behavior by first removing non-Brownian components of particle motion [25]. The effect is to revert a non-stationary process, such as random motion superposed with heterogeneous ballistic motion, to a stationary process. The standard strategy in SPT has been to subtract an estimated non-Brownian motion defined by the ensemble-averaged displacement in each time step, a process called 'de-drifting' or 'de-trending' in the literature [25, 26]. Additional processing of this estimate, such as convolution with a uniform weighting function, assumes Brownian and non-Brownian motion have separable characteristic timescales. Choosing the span of the weighting function is commonly subjective, such as in the popular IDL code [12], although first performing an ellipticity analysis on SPT trajectories may offer reasonable values for window span [27]. Less subjective approaches have been recently developed using multiresolution wavelet analysis with universal thresholding, which have been employed to correct spatiotemporally heterogeneous advection of intracellular probes in migrating cells [28] and to identify particular time steps in which displacements are influenced by hydrodynamic interactions in the diffusion of confined probe particles [29]. To process the estimated non-Brownian motion relies on large SPT data sets obtained by tracking large ensembles over numerous time steps.

* Also at Department of Medicine, Hematology/Medical Oncology, Weill Cornell Medicine; kirby@cornell.edu

Current de-drifting methods are important tools for SPT but require numerous, monodisperse particles. De-drifting using small ensembles (those with few particles) can erroneously reduce measured Brownian motion because ensemble-averaged displacement and component particle displacements remain significantly correlated. To prevent this effect, one can analyze only the components of particle displacements orthogonal to the estimated drift vector, but displacement data sets are inherently halved [30]. For identical particles, Bessel's correction can provide accurate determination of MSD after de-drifting with small ensembles [26]. This correction factor extended the application of de-drifting to heterogeneous non-Brownian motion by utilizing small ensembles of homogeneous particles from local domains [31]. Yet, no correction factor has been developed for small ensembles composed of particles with heterogeneous diffusive mobility. Savin et al. considered monodisperse probe particles in a complex fluid with heterogeneous viscosity but used overlapping displacements, requiring a block-averaging technique in addition to de-drifting [32].

Residual correlations induce new and more complicated effects when they exist in suspensions with heterogeneous diffusive mobility as compared to homogeneous particle suspensions. These effects prevent the proper application of de-drifting in several important SPT applications. De-drifting in nanoparticle tracking analysis (NTA) of highly disperse suspensions necessarily uses ensembles of particles with different diffusive mobility. Improved measurement accuracy could be influential for characterizing the PSD of biological extracellular vesicles, where dilute, polydisperse particle suspensions are commonly measured in applied fluid flow [33, 34]. In addition, de-drifting in passive particle tracking microrheology of complex media with heterogeneous viscoelastic properties may require ensembles of monodisperse probes that are localized to regions with uniform collective motion but different rheological properties. For these cases, standard (uncorrected) de-drifting or de-drifting with a single correction factor per ensemble do not isolate Brownian motion, and instead introduce artifacts into particle spatial positions, as demonstrated in this work. For clarity, our work is presented in a single practical context – NTA, wherein the desired output from analyzing numerous single particle trajectories is the PSD. Heterogeneous diffusive mobility is clearly illustrated by heterogeneous particle diameters, where we implement the Stokes-Einstein relation to map between the two parameters.

This work is structured in four main sections. First, the methodology of SPT simulations emulating NTA of suspensions composed of heterogeneous particles embedded in a Newtonian fluid and NTA experiments of well-defined polystyrene bead mixtures suspended in water are described. In Section III A, the removal of bulk fluid motion, i.e. drift, by standard de-drifting is presented in general form, replicating previous work on monodisperse suspensions (Section III B) and extending de-drifting to particle suspensions with arbitrary PSD (Section III C). The general form clarifies the origin of spurious biasing of particle diameter introduced by de-drifting for dilute, polydisperse particle diameter populations, and leads to predicted error for previous SPT experiments (Section III D) and of predicted dynamic range limitations for the practical scenario of a single large particle in a bath of small, identical particles (Section III E). Spurious biasing of particle diameter has two competing influences. First, the residual correlation ($\mathcal{O}(2/N)$) between the estimated drift used in de-drifting and individual particle displacements results in subtracting away true Brownian motion. Second, composition-dependent noise in the estimated non-Brownian motion adds spurious motion and may even dominate the de-drifted displacement of particles with lower mobility, i.e. larger particles in a high dispersity suspension. In Section IV A, we report a de-drifting method (Decorrelation Analysis) for SPT trajectories in multi-probe experiments that contain temporally heterogeneous non-Brownian motion and where ensembles may have heterogeneous diffusive mobility. Decorrelation Analysis (DA) effectively isolates isotropic Brownian motion, as validated by simulations and experimental data (Section IV B). Throughout this work, we demonstrate the effect of convolving the estimated drift vector with a linear weighting function by comparing 'raw' and 'smooth' analyses (Table I).

II. METHODS

A. Simulations

To investigate the effect of de-drifting on ensembles composed of few particles with heterogeneous diffusive mobility, we simulated 2D SPT trajectories in a Newtonian fluid using MATLAB® version R2013a (The MathWorks, Natick, Massachusetts, USA). In simulations characterizing artifacts introduced by standard de-drifting with no corrections (Fig. 1,2,3) all particle trajectories spanned the equivalent number of time steps, whereas in simulations benchmarking DA against other correction methods (Fig. 4, 5, 6) each particle trajectory spanned a random number of time steps between 100 and 1000 in order to better represent the experimental conditions of NTA. Particle diameters were sampled from a probability distribution with one, $f_{\mathrm{LogN}}(x;\mu_1,\sigma_1)$, or two lognormally distributed modes, $f_a(x) = \lambda_1 f_{\mathrm{LogN}}(x;\mu_1,\sigma_1) + \lambda_2 f_{\mathrm{LogN}}(x;\mu_2,\sigma_2)$, where λ is the proportion of sampling ($\sum_i \lambda_i = 1$), μ is the mean and σ is the standard deviation of the given mode. Diameters were translated into diffusion constants by use of the Stokes-Einstein relation, $D = k_b T/(3\pi\eta a)$, and used to generate particle trajectories consisting of displacements sampled from a zero-mean normal distribution $\sim \mathcal{N}(0, 2D\Delta t)$ for each spatial dimension, where k_b is the Boltzmann constant, T and η are the absolute temperature and dynamic viscosity, respectively, of the simulated Newtonian fluid, a is particle diameter, and D is

particle isotropic diffusion constant. Time-varying non-Brownian motion, typically termed 'drift' as its primary cause is often bulk fluid motion, was simulated for prescribed frequencies given the constant time step, and applied uniformly to simulated particles at each time step. In each dimension the superposition of Brownian (Δx^{Br}) and non-Brownian motion ($\boldsymbol{u}_{\mathrm{true,x}}\Delta t$) formed individual particle observed displacements,

$$\Delta x^{\mathrm{obs}}(t_i) = \Delta x^{\mathrm{Br}}(t_i) + \boldsymbol{u}_{\mathrm{true,x}}(t_i)\Delta t \ , \tag{1}$$

and subsequent subtraction of estimated drift formed de-drifted displacements (i.e. the inferred displacement of a particle after de-drifting),

$$\Delta x^{\mathrm{dd}}(t_i) = \Delta x^{\mathrm{Br}}(t_i) + \left(\boldsymbol{u}_{\mathrm{true,x}}(t_i) - \boldsymbol{u}_{\mathrm{est,x}}(t_i)\right)\Delta t \ . \tag{2}$$

To demonstrate the effect of convolving a linear weighting function with the estimated drift in the time domain, 'raw' (not time-averaged) values were compared to 'smooth' (time-averaged using top-hat weighting function) values (Table I).

Each particle track generated multiple sets of displacement data associated with an analysis, and these data sets were analyzed in parallel during a particular iteration of a simulation (see Table I for descriptions of analyses). Particle diameters were computed individually using the Stokes-Einstein relation and lognormal distribution parameters were determined using the native Expectation-Maximization algorithm GMDISTRIBUTION initialized with mean values from the native k-Means Clustering function KMEANS assuming two distribution modes. Probability density functions created from the determined lognormal distribution parameters were used to create distinct PSDs for each analysis approach.

B. Experiments

Experiments were performed using the NanoSight® NS300 (Malvern Instruments, Malvern, UK) equipped with a 532nm laser source, sCMOS camera and software version NTA 3.0 build 0068. 100- and 400nm polystyrene bead standards (Thermo Fisher Scientific, Waltham, MA, USA) were diluted in 1% phosphate-buffered saline (PBS). A Fusion 400 syringe pump (Chemyx Inc, Stafford, TX, USA) generated uniform, constant flow in the suspending fluid and was physically coupled to a high frequency oscillator to generate uniform, dynamic flows. The NanoSight® particle position data was exported and analyzed in MATLAB®. Some tracks were identified as 'false' and removed for all analyses. False tracks were identified using the exported 'true' or 'false' track designation from the NTA software with the additional criteria of a trajectory length threshold selected to prevent a biased over-sampling of smaller particles that are more likely to jump in and out of the camera field of view [18]. For quiescent bulk conditions, frame-by-frame suspension composition was controlled by sampling particles from specific size modes as specified by the size exported from NanoSight® and then dictating the starting frame of the selected particle tracks. This process of recreating an experiment with a controlled number of particles in each frame is possible only for quiescent bulk conditions. Plots of the % error of PSD parameters from experimental data use 'true' parameters that were measured in quiescent bulk conditions for the same sample.

III. DE-DRIFTING OF SPT TRAJECTORIES

A. Theory

SPT records individual particle's spatial positions over time and weaves these discrete positions together to create SPT trajectories that span any number of time steps. Observed particle position in one spatial dimension at time step t_i is given by $x(t_i) = x(t_0) + \sum_{l=0}^{i-1}\Delta x^{\mathrm{obs}}(t_l)$, where $\Delta x^{\mathrm{obs}}(t_l) = x(t_{l+1}) - x(t_l)$. One-dimensional observed particle displacement (Δx^{obs}) is determined by a combination of random thermal fluctuation and motion from a variety of possible non-Brownian 'drift' sources,

$$\Delta x^{\mathrm{obs}}(t_i) = \Delta x^{\mathrm{Br}}(t_i) + \boldsymbol{u}_{\mathrm{drift},x}(\boldsymbol{r}_i, t_i)\Delta t \ , \tag{3}$$

where $\boldsymbol{r}_i$ is two-dimensional particle position at time t_i and Δt is the constant time lag between observations. Although the source-agnostic drift term, $\boldsymbol{u}_{\mathrm{drift},x}$, can vary discretely in time and space, it is assumed constant over Δt.

De-drifting is elegant in concept – simply remove the drift analytically by subtracting an estimate, thereby including only Brownian motion in further analyses. The presence of drift, i.e. bulk fluid flow, generates a nonzero mean displacement, but de-drifting theoretically reverts particle displacement to a zero-mean process. The estimate is taken as the ensemble-averaged displacement between time steps, $\langle\Delta x^{\mathrm{obs}}\rangle_N$, where $\langle\cdot\rangle_N$ indicates the average of an N-particle ensemble. The respective dynamics of drift and Brownian motion are generally considered separable such that high-frequency components of the ensemble average are often filtered out to obtain lower frequency drift by itself. For example, the popular MATLAB® implementation of Crocker et al.'s IDL code uses a top-hat temporal filter [35]. The choice of weighting function determines the power of spectral components remaining in the estimated drift. Here, a general linear weighting function is applied as a central moving average and convolved with the ensemble-averaged displacement to show the de-drifted displacement,

$$\Delta x^{\mathrm{dd}}(t_i) = \Delta x^{\mathrm{obs}}(t_i) - \sum_{k=-\frac{w-1}{2}}^{\frac{w-1}{2}} \left\{W_k \cdot \langle\Delta x^{\mathrm{obs}}(t_{i+k})\rangle_N\right\} \ , \tag{4}$$

where the weighting function has odd-valued span w and coefficients that sum to unity ($\sum W_k = 1$). Because particles can leave and enter the camera focal plane, N is dependent on time step. The true spatial variation of drift determines the spatil length scale over which particles may be ensemble-averaged for this purpose – uniform bulk motion can be removed with a single ensemble whereas heterogeneous bulk motion may require spatially localized ensembles [31]. In this manuscript, we consider only uniform bulk motion because it is the expected form of directed transport in NTA experiments. Applications with spatially heterogeneous drift would require a method to define local ensembles, such as by machine learning methods [36]. Substitution of Eq. (3) into Eq. (4) leads to

$$\begin{aligned}\Delta x^{\mathrm{dd}}(t_i) = \quad & \Delta x^{\mathrm{Br}}(t_i) - \sum_{k=-\frac{w-1}{2}}^{\frac{w-1}{2}} \left\{ W_k \cdot \langle \Delta x^{\mathrm{Br}}(t_{i+k}) \rangle_N \right\} \\ & + \boldsymbol{u}_{\mathrm{drift},x}(\boldsymbol{r}_i, t_i)\Delta t \\ & - \sum_{k=-\frac{w-1}{2}}^{\frac{w-1}{2}} \left\{ W_k \cdot \langle \boldsymbol{u}_{\mathrm{drift},x}(\boldsymbol{r}_i, t_{i+k})\Delta t \rangle_N \right\} \ , \end{aligned} \tag{5}$$

where $\langle \cdot \rangle_N$ has been considered a linear operator such that $\langle A + B \rangle_N = \langle A \rangle_N + \langle B \rangle_N$ for arbitrary random variables A and B. For negligible spatiotemporal heterogeneity in the true drift given the particular ensemble and averaging window span, drift terms cancel and de-drifted displacement is dependent only on particle Brownian motion,

$$\Delta x^{\mathrm{dd}}(t_i) = \Delta x^{\mathrm{Br}}(t_i) - \sum_{k=-\frac{w-1}{2}}^{\frac{w-1}{2}} \left\{ W_k \cdot \langle \Delta x^{\mathrm{Br}}(t_{i+k}) \rangle_N \right\} \ . \tag{6}$$

As $N \longrightarrow 0$, the resulting de-drifted displacement for an individual particle can no longer be assumed fully independent of the other particles in the suspension.

It is well known that the variance of a sum of random variables $A_1 + A_2 + \cdots + A_N$ is the sum of all covariances,

$$\mathrm{var}\left(\sum_{m=1}^{N} b_m A_m \right) = \sum_{m,j}^{N} b_m b_j \mathrm{cov}\left(A_m, A_j \right) \ , \tag{7}$$

where b's are constants. Because de-drifted displacement is a sum of random variables all with zero mean, the variance of Δx^{dd} can be determined using Eq. (7) and is equivalent to the MSD. The Brownian motion of any two particles is not cross-correlated because hydrodynamic interactions are assumed negligible, and single-particle Brownian motion is not autocorrelated in time. Importantly, uncorrelated random variables have covariance equal to zero. Thus, the de-drifted MSD simplifies to a sum of variances,

$$\alpha_1^{\mathrm{dd}} = \left(1 - \frac{2W_0}{N} \right) \alpha_1^{\mathrm{Br}} + \sum_{k=-\frac{w-1}{2}}^{\frac{w-1}{2}} \left\{ W_k^2 \cdot \frac{1}{N} \langle \alpha^{\mathrm{Br}} \rangle_N \right\} \ , \tag{8}$$

where the subscript in α_m identifies a specific probe particle m. Eq. (8) is an important result of this work. The first term contains the independent Brownian motion of an arbitrary particle 1 that is to be measured, and the second term represents the influence of the other tracked particles. This equation demonstrates that de-drifting, though removing correlated bulk fluid motion, introduces spurious correlations in the motion of particles composing the ensemble used to estimate drift. Equivalently stated, de-drifting converts the original displacement data series to a stationary process, but does not recreate a process describing the independent motion of a single Brownian particle.

B. Monodisperse SPT

Monodisperse particle suspensions are composed of particles with essentially equivalent diameter and therefore equivalent diffusive mobility. In this section, we validate simulation results with the literature on monodisperse suspensions and describe the effect of time-averaging, i.e. smoothing, estimated drift on measured PSDs. An unsmoothed approach ('Raw de-drifting') is compared to a smoothed approach ('Smooth de-drifting') to characterize the effect of time-averaging (see Table I for description of analyses).

In dilute, monodisperse suspensions, Monte Carlo (MC)-averaged distribution means and variances show that de-drifting introduced spurious peak shifts in the PSD. Bulk fluid displacement approximated by the unprocessed ensemble-averaged measured particle displacement consistently over-estimated particle sizes, and the error magnitude was inversely related to the total number of tracked particles per time step ('Raw de-drifted' in Fig. 1a). Previous reports have made the same observations [26, 31]. By assuming that averaged particle fluctuations originate only from bulk fluid motion, this approach spuriously subtracted away a portion of particle fluctuation, reducing MSD and increasing inferred diameter. For raw de-drifting of a monodisperse particle suspension ($\alpha_1^{\mathrm{Br}} = \alpha_2^{\mathrm{Br}} = \cdots = \alpha_N^{\mathrm{Br}} = \alpha^{\mathrm{Br}}$), Eq. (8) simplifies to $\alpha_1^{\mathrm{dd}} = (1 - 1/N)\,\alpha^{\mathrm{Br}}$, where the coefficient $(1 - 1/N)$ has been described as an MSD correction factor [26, 31]. Thus, for monodisperse suspensions, ensemble-averages remain correlated to the individual parts with scaling $\mathcal{O}(\frac{1}{N})$. As $N \longrightarrow \infty$, ensemble-averaged estimates become correct.

Smoothing the estimated drift signal can reduce high-frequency components that remain correlated to individual particle motion. Smoothing effectively augments the number of particles in the ensemble by considering ad-

Analysis name	**Displacement data set**	**MSD calculation** [a]
Specified PSD (continuous)	-	-
Sampled PSD (discrete)	-	$\frac{4k_bT}{3\pi\eta a}$
No correction	Δx^{obs}	$\langle\|\Delta \boldsymbol{r}\|^2\rangle$
Perfect correction	Δx^{Br}	$\langle\|\Delta \boldsymbol{r}\|^2\rangle$
Raw de-drifting	Eq. (6) with $w = 1, W_0 = 1$	$\langle\|\Delta \boldsymbol{r}\|^2\rangle$
Smooth de-drifting	Eq. (6) with $W_k = 1/w$	$\langle\|\Delta \boldsymbol{r}\|^2\rangle$
Monodisperse correction [26]	Eq. (6) with $w = 1, W_0 = 1$	$\langle\|\Delta \boldsymbol{r}\|^2\rangle \cdot (1 - 1/N)$
Raw Decorrelation Analysis	-	Eq. (12) with $w = 1$
Smooth Decorrelation Analysis	-	Eq. (12) with $W_k = 1/w$

TABLE I. Methods to generate PSDs used in this work. 'Specified PSD' is a continuous probability density function generated from the summation of two lognormally-distributed random variables with specified parameters, and therefore is not associated with a set of displacements nor any calculation of MSD. 'Sampled PSD' contains the particle diameter values that were sampled from 'Specified PSD' in SPT simulations, and therefore is associated with the discrete MSD values used to generate particle trajectories. 'No correction' implies that observed particle trajectories were not modified in any manner. 'Perfect correction' only considers particles' Brownian motion, and therefore represents ideal values given instrument broadening. 'Monodisperse correction' utilizes the MSD correction factor from [26]. Both 'Raw de-drifting' and 'Smooth de-drifting' are presented in order to illustrate the effect of applying linear weighting functions to estimated non-Brownian motion. Accordingly, both 'Raw DA' and 'Smooth DA' forms are presented for clarity. Although only the x component of displacement is shown, the y component is equivalently determined, and $\Delta \boldsymbol{r}^2 = \Delta x^2 + \Delta y^2$ is used in calculations of MSD. [a] Individual particle MSDs were converted into suspension PSDs by using k-means Clustering to initialize Gaussian mixture model fitting of multivariate distribution parameters, as described in the Methods section.

jacent values in the estimated drift. Distributions determined from 20 particles per time step, $N_t = 20$, for example, were identical to distributions determined from 4 particles per time step but 20 particles per averaging window, $N_w = 20$ (Fig. 1a,b,c). The equivalent distribution shifts for these two conditions show that time averaging simply expands the number of particles used to estimate drift, systematically reducing spurious peak shifts at the cost of temporal resolution in the estimated drift.

C. Polydisperse SPT

Polydisperse particle suspensions are composed of particles with heterogeneous diameters. SPT simulations demonstrate that the magnitude and direction of particle diameter biasing introduced by standard de-drifting is not equivalent for both subpopulations of a bi-modal PSD. Results from this section are presented in the context of NTA but are also applicable for SPT-based microrheology measurements of complex fluids with heterogeneous rheological properties.

Similar to the monodisperse case, de-drifting the motion of particles sampled from bi-modal PSDs demonstrated an inverse relationship between the magnitude of spurious peak shifts and the total number of tracked particles. As the number of tracked particles increased, size measurements for both populations became more accurate (Fig. 1b,c).

Simulations of tracked particles sampled from two different bi-modal PSDs show that biases in the measured diameter are a function of suspension composition. Two subpopulations close in mean diameter both exhibited over-estimated particle diameters (Fig. 1b); for example, 200- and 400-nm particles appeared to be 291- and 457-nm, respectively. On the other hand, two subpopulations far apart in mean diameter exhibited a varied biasing of particle diameter (Fig. 1c), where 100nm particles appeared larger (152nm) but 400nm particles appeared smaller (355nm). The role of the suspension composition in de-drifting has been inherently neglected by studies of monodisperse suspensions.

The varied nature of particle diameter biasing in SPT simulations of bi-modal particle suspensions occurs because the net shift in size depends on two counteracting influences represented by the two main terms in Eq. (8). First, as described for monodisperse suspensions, residual correlation between the estimated drift and individual particle motion causes vector subtraction to erroneously remove a portion of Brownian motion. The first term of Eq. (8) shows that this error is $\mathcal{O}(\frac{2W_0}{N})$. To better understand the second influence, consider the limit case of a motionless particle. Because this particle has no inherent diffusive MSD, de-drifting adds, rather than subtracts, particle motion equivalent to the opposite of the estimated drift vector. Thus, the inferred particle diameter, which in theory is infinite, is instead a finite value fully determined by the motion of other particles in the ensemble. This limit case represents high dispersity suspensions composed of particles with widely varying diffusive mobility. Disparities in diffusive mobility create disparities between particles in the relative magnitude of noise introduced by subtracting the ensemble-averaged

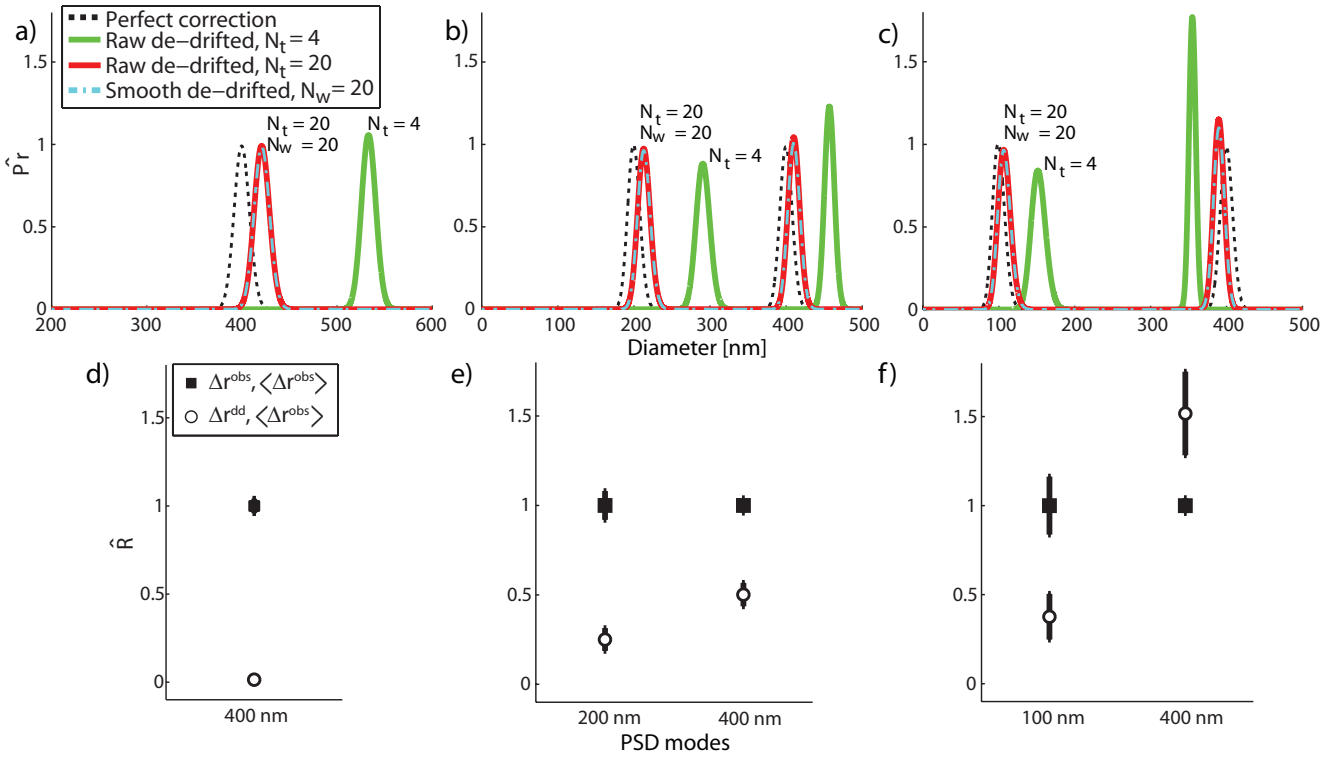

FIG. 1. **De-drifting causes particle number- and composition-dependent errors related to the correlation between estimated drift and particle displacement**. (a–c) $\hat{Pr}$ is the measured probability density distribution normalized by the 'Perfect correction'. PSDs represent MC-averaged (1000 iterations) distribution parameters measured from different total number particles per time step, N_t, or per window, N_w. The magnitude of peak shift was inversely related to the number of particles defined by the ensemble and the direction of shift depended on the suspension composition. Particles were sampled equally from lognormal distributions with standard deviation $\sigma = 8$nm and applied constant drift (1μm/s). (d–f) $\hat{R}_{xy}$ is the unbiased zero-lag cross-correlation function, Eq. (9), normalized by the 'Observed' value for each peak. Data points represent the MC average $\pm 2\sigma$ in the case of $N_t = 4$, 2e5 displacements per particle and no applied drift. Data points (circles) indicate whether the spurious peak shift was dominated by residual correlation (< 1) or introduced correlation (> 1). Legend indicates vectors being assessed for correlation.

displacement. Depending on the exact composition, this noise can dominate the inferred displacement of less mobile, larger particles. The second term of Eq. (8) shows that this error scales with $\mathcal{O}(\frac{wW_k^2}{N}\phi)$, where $\phi = \langle \frac{\alpha^{\mathrm{Br}}}{\alpha_1^{\mathrm{Br}}} \rangle_N$ is a measure of the dispersity relative to the particle being measured. The greater the dispersity or the smaller the relative mobility of the particle being measured, the greater the influence of this type of error.

In accordance with these scaling arguments, particle diameters sampled from a bi-modal PSD with peak means close in value were over-estimated, i.e. the dominating influence of de-drifting was the removal of a portion of Brownian motion (Fig. 1b). Individual de-drifted displacements and the ensemble average remained correlated because of small N_t, and noise associated with the ensemble average did not dominate the inferred displacement of larger particles.

Particle diameters sampled from a bi-modal PSD with peak means further apart exhibited varied biasing of inferred particle diameter because the dominant influence on biasing error was dependent on diameter (Fig. 1c). Correlated subtraction again introduced spurious peak shifts toward larger diameters, which dominated the net biasing of 100nm particles. For these smaller particles, the noise in the estimated drift did not drown out Brownian motion. Larger particles, however, have lower diffusive mobility and the relative magnitude of the noise was greater compared to particle Brownian motion. Uncorrelated vector subtraction of the ensemble-averaged displacement added noise whose magnitude dominated the inferred particle displacement, decreasing the inferred diameter of less-mobile 400nm particles.

To demonstrate the dominant source of particle size biasing, we traced the average value of the cross-correlation function between the estimated drift and individual particle displacements through de-drifting. Investigating only the correlations associated with the de-drifting step was achieved in simulations with no applied drift and by evaluating the cross-correlation function at zero lag for

each individual particle,

$$R_m = \frac{1}{T}\sum_{t=1}^{T} \Delta \boldsymbol{r}_m \langle \widetilde{\Delta \boldsymbol{r}}^{\,\mathrm{obs}} \rangle_N \ , \qquad (9)$$

where m indicates the arbitrary particle identification number, T is the total number of displacements in the trajectory and $\langle \widetilde{\Delta \boldsymbol{r}}^{\,\mathrm{obs}} \rangle_N$ is the complex conjugate of the estimated drift. The imaginary vector component was equal to the y-direction component of the estimated drift vector. Then, the cross-correlation function value for a particular size subpopulation was determined by averaging R_m over the particles composing that population. Simulations in which peaks spuriously shift toward larger diameters show that the observed displacement correlated most strongly with the estimated drift and that de-drifting reduced this correlation (Fig. 1d,e). Because the simulations here do not include an applied drift, reducing correlation corresponded to spurious removal of a portion of the true Brownian motion. Underestimated particle diameter, however, corresponded with increased correlation between particle de-drifted displacement and estimated drift, indicating that the estimated drift vector erroneously dominated the particle de-drifted displacement (Fig. 1f). Therefore, the fluctuating motion of the other particles in the suspension mostly determined the estimated diameter of less-mobile, larger particles.

D. Quantifying systematic bias of de-drifted particle MSD

There are circumstances when predicting the error caused by de-drifting may be important, such as to re-evaluate previous SPT results. In this section, we demonstrate a straightforward method to make analytical predictions of bias, and then show that these predictions agree perfectly with true bias calculated from simulated SPT trajectories (Fig. 2).

For a top-hat temporal filter, the linear weighting function coefficients are defined by $W_k = \frac{1}{w}\ \forall\ k$. By applying the Stokes-Einstein relation, Eq. (8) simplifies to,

$$\frac{a_1^{\mathrm{Br}}}{a_1^{\mathrm{dd}}} = \left(1 - \frac{2}{wN}\right) + \frac{1}{wN^2}\sum_{m=1}^{N} \frac{a_1^{\mathrm{Br}}}{a_m^{\mathrm{Br}}} \ , \qquad (10)$$

where all N particles were tracked over the same time steps and under the same conditions, a^{dd} is the 'de-drifted diameter' (inferred diameter after de-drifting) and a^{Br} is the true hydrodynamic diameter. Eq. (10) reiterates the scaling argument made above – that shifts in particle diameter depend on relative, rather than absolute, diameters in the particle suspension.

De-drifting introduces equivalent relative bias in the inferred particle diameter for monodisperse suspensions regardless of diameter. The magnitude of the peak shift reproduces previous results in the literature [26] and depends on the total number of particles. Furthermore, the

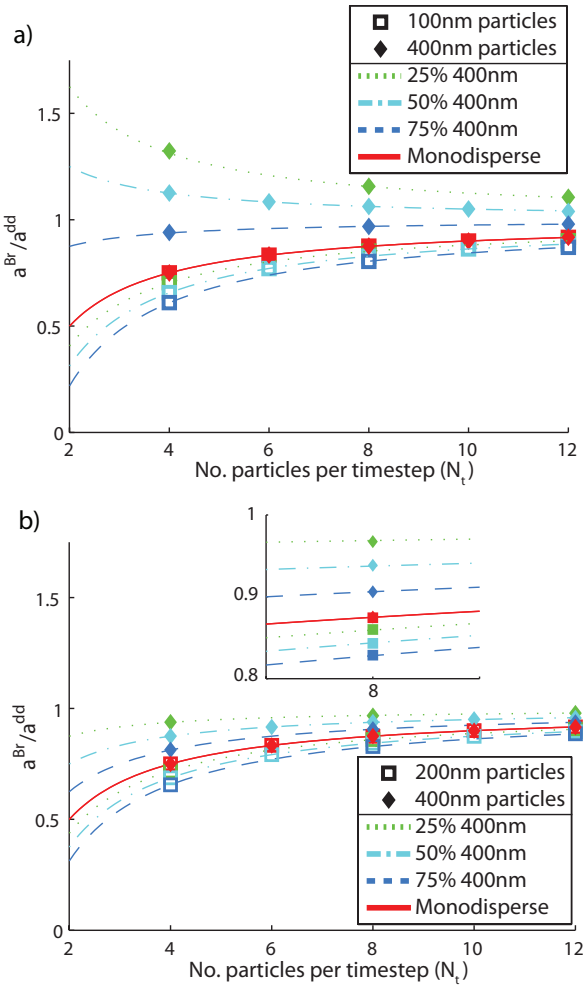

FIG. 2. **Spurious shifts from simulated particle motion are exactly predicted analytically**. Results from SPT simulations of bi-modal PSDs with 100nm- and 400nm-diameter peak means (a) or 200nm- and 400nm-diameter peak means (b) are shown. The y axes are a measure of the spurious peak shifting caused by de-drifting where values > 1 correspond to under-estimated diameter ($a^{\mathrm{dd}} < a^{\mathrm{Br}}$), values $= 1$ have no bias ($a^{\mathrm{dd}} = a^{\mathrm{Br}}$) and values < 1 correspond to over-estimated diameter ($a^{\mathrm{dd}} > a^{\mathrm{Br}}$). Lines represent the result of Eq. (10) given the relative number of each subpopulation that was simulated. Data points represent the average diameter ratio of the indicated size subpopulation (empty square or filled diamond) calculated from 10 simulations in which particles were tracked for 1e4 displacements with no bulk flow. Line color/style represent five cases corresponding to the number of small:large particles in ∞:0 (red/line), 3:1 (green/dotted), 1:1 (cyan/dot dash), 1:3 (blue/dash) and 0:∞ (red/line) ratios. The inset in (b) zooms in on the case $N = 8$ in order to better visualize the agreement between analytical and simulation results. Analytically calculated bias and true bias calculated from simulated trajectories are in perfect agreement, demonstrating that Eq. (10) predicts composition-dependent error introduced by de-drifting.

value of the y value in Fig. 2 indicates the direction of the peak shift, corroborating that de-drifting monodisperse suspensions always causes over-estimation of inferred particle diameter ($a^{\mathrm{Br}}/a^{\mathrm{dd}} < 1$).

De-drifting particle motion in SPT simulations of polydisperse suspensions, represented by particle sizes sampled from two different bi-modal PSDs, causes spurious shifts in inferred particle diameter that perfectly match the shifts predicted by analytical curves. Both the relative diameters of the size modes and the relative number of particles from each mode contribute to spurious peak shifts. MC simulations with 100- and 400-nm (Fig. 2a) and 200- and 400-nm (Fig. 2b) size modes shows both over- ($a^{\mathrm{dd}} > a^{\mathrm{Br}}$) and under-estimated ($a^{\mathrm{dd}} < a^{\mathrm{Br}}$) inferred particle diameter. There is perfect agreement between simulation results and analytical curves across mono- and poly-disperse suspensions with varying levels of dispersity, indicating that Eq. (10) is a straightforward analytical method to assess de-drifting-related biasing error for particle suspensions with arbitrary dispersity.

E. Limitations for single large particles

Accurate measurement of very low numbers of large particles in a bath of smaller particles is an important practical concern for NTA, especially for the characterization of tumor-derived extracellular vesicles, which are dilute, polydisperse and measured in applied bulk flow [33, 34]. Importantly, the large-diameter portion of the PSD has typically been reported as a small tail of the distribution or a small secondary peak [37], and thus these particles are likely strongly affected by the biasing effect described herein. To assess the limitations imposed by de-drifting in this relevant scenario, SPT simulations were performed using a single large particle of diameter a_{big} in a bath of identical small particles of diameter a_{bath}. Fig. 3 shows the inferred diameter of a particle ∞- (i.e. motionless), 100-, 10-, 5-, 3- and 1-times the diameter defined by the homogeneous bath particles and results are plotted as a function of *true* diameter ratio, $\zeta = a_{\mathrm{big}}/a_{\mathrm{bath}}$, *de-drifted* diameter ratio, $\zeta^{\mathrm{dd}} = a^{\mathrm{dd}}_{\mathrm{big}}/a_{\mathrm{bath}}$ and ensemble size per time step, N. Differences between ζ and ζ^{dd} indicate biasing in particle diameter caused by de-drifting.

In general, the accuracy of inferred diameter increases with N and decreases with diameter ratio, ζ^{dd}. More numerous particles ($N \longrightarrow \infty$) better estimate the true bulk fluid flow and reduce remaining correlation between particle de-drifted displacements and the estimated drift. Larger particles move less and their inferred diameters are more susceptible to noise in the estimated drift. Even with $N = 50$, a 1μm particle in a bath of 100nm particles would be measured as having a diameter of 0.865μm. Special mention should be made of the case $\zeta = 3$, in which the two counteracting biases essentially cancel out. It is easily shown from Eq. (10) that $\zeta = 3$ is the transition diameter ratio, at which net biasing transitions from over- to under-estimation.

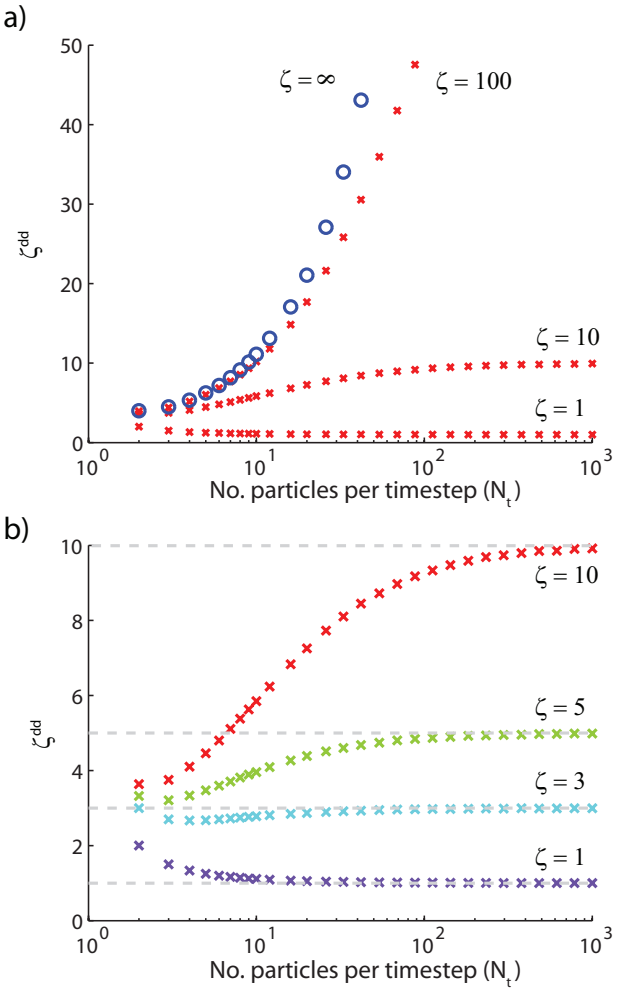

FIG. 3. **Accuracy of a single large particle in a bath of small particles**. Results from SPT simulations of $N-1$ bath particles and a single large particle over a range of diameters, $\zeta = a_{\mathrm{big}}/a_{\mathrm{bath}}$. Plots show (a) a broad range and (b) a practical range of ζ for vesicles. The y axes are the *de-drifted* diameter ratio, $\zeta^{\mathrm{dd}} = a^{\mathrm{dd}}_{\mathrm{big}}/a_{\mathrm{bath}}$, while ζ indicates the *true* diameter ratio, thus the de-drifted ratio approaches the known true ratio as $N \longrightarrow \infty$. Whether values approach ζ from ∞^+ or ∞^- indicates the dominant source of error.

IV. DE-CORRELATION ANALYSIS

Given that de-drifting removes motion associated with the bulk fluid yet introduces error by correlating particle motion, in this section we present an additional processing step that corrects for correlated particle motion. Low particle density causes de-drifting to systematically bias inferred particle diameter because the ensemble-averaged displacement either remains correlated to particle motion or contains sufficient noise to drown out less mobile, larger particle Brownian motion. We use the term

'Decorrelation Analysis' to describe the combination of removing bulk motion by de-drifting and removing introduced correlations by matrix inversion, as shown below. The solution provided here applies to suspensions of particles with arbitrary diameters and is straightforward to implement. The improved accuracy in measuring particle diameter permits reduced smoothing of the estimated drift and better maintains the temporal resolution of the analysis.

A. Theory

The MSD of particle Brownian motion, α^{Br}, is the desired quantity for size determination using NTA. In most studies to date, the de-drifted MSD (Eq. (8)) is used rather than the MSD of Brownian motion, which is valid only for large ensembles of particles. When ensembles are small, however, Eq. (8) must be simultaneously solved for α^{Br} for all tracked particles. Translating Eq. (8) into matrix form requires the approximation that all N particles are tracked for the duration of the averaging window centered at time step t_i, which is reasonable when averaging windows span much fewer time steps than particle trajectories. Thus,

$$\boldsymbol{\alpha}^{\mathrm{dd}} = \left[\left(1 - \frac{2W_0}{N} \right) \boldsymbol{\delta} + \left(\frac{1}{N^2} \sum_{k=-\frac{w-1}{2}}^{\frac{w-1}{2}} W_k^2 \right) \mathbb{1} \right] \cdot \boldsymbol{\alpha}^{\mathrm{Br}} \ , \tag{11}$$

where $\boldsymbol{\delta}$ is the identity matrix, $\mathbb{1}$ is an $N \times N$ matrix of ones, and both $\boldsymbol{\alpha}^{\mathrm{dd}}$ and $\boldsymbol{\alpha}^{\mathrm{Br}}$ are vectors containing de-drifted and Brownian motion MSDs, respectively, for all N particles. The dimensions of MSD vectors may vary in each time step as N changes. If the linear weighting function is an unweighted moving average ($W_k = \frac{1}{w} \ \forall \ k$), the bracketed matrix can be inverted analytically using a method for sums of matrices [38],

$$\boldsymbol{\alpha}^{\mathrm{Br}} = \frac{w}{wN-2} \left(N\boldsymbol{\delta} - \frac{1}{wN-1}\mathbb{1} \right) \cdot \boldsymbol{\alpha}^{\mathrm{dd}} \ , \tag{12}$$

providing a straightforward calculation of de-drifted and de-correlated MSD related only to the Brownian motion of each particle. The identity matrix term contains the correction for correlated subtraction whereas the ones matrix term contains the correction for uncorrelated subtraction. This solution reflects the concept of ergodicity: time-averaging for $w \longrightarrow \infty$ is equivalent to ensemble-averaging for $N \longrightarrow \infty$. Consistent with expectations, as N or $w \longrightarrow \infty$, de-drifting effectively removes bulk fluid flow without causing bias in the inferred MSD. Time-averaging, however, will reduce the temporal resolution of the estimated drift, which is important for dynamic bulk flows. Eq. (12) represents particle DA MSDs over time, and therefore a single particle has as many definitions as time steps in its trajectory. An unweighted average of a particle's DA MSD signal provides a good estimate for mapping to a single value for calculating particle diameter.

B. Benchmarking

In this section, simulations and experimental data validate the gains in measurement accuracy attributable to a follow-on de-correlation step (Eq. (12)) relative to de-drifting alone and to a monodisperse correction [26]. As opposed to current processing approaches to increase accuracy such as time averaging the estimated drift signal, de-correlation does not require reduced temporal resolution.

Uniform, constant drift

DA (Eq. (12)) was benchmarked against standard de-drifting approaches using SPT simulations and experiments. SPT simulations were implemented with real-world conditions, in which particle tracks do not contain equivalent statistical information. Particle trajectory spans of varying length (100–1000 time steps) begin at random time points within the simulation and the number of particles for each time step varied around a specified average composition. To reduce instrument broadening [39], particle trajectories spanned many time steps, and to reduce distribution fitting error, numerous particle tracks were simulated. Simulations applied a constant, uniform bulk fluid flow; however, experiments were performed with no applied bulk flow because syringe pump jitter can introduce dynamic drift [40]. A 'true' measurement for % error of PSD parameters was obtained without any drift correction. For the purposes of evaluating error introduced by different drift correction methods, zero bulk flow and constant, uniform bulk flow are equivalent.

Data from particle tracking experiments (Fig. 4c,d) validated simulation results (Fig. 4a,b). The % errors of peak means determined from experiment and simulation data were in good agreement for a given diameter subpopulation (square or diamond markers) and particle ensemble number (marker size). For example, % error of peak mean for 'raw' de-drifted 100nm particles was 26% vs. 30% ($N_t = 6$), 12% vs. 14% ($N_t = 12$), 8% vs. 10% ($N_t = 18$) in simulation vs. experiment, respectively. Calculated % errors of peak standard deviations generally shifted toward ∞^+ in experimental analyses relative to simulations (red markers), i.e. distribution peaks were uniformly broader. Shorter particle trajectories in experiments compared to simulations may have caused this broadening because it was observed that NanoSight® artificially cuts particle trajectories short after ∼100 time steps.

Both simulated (Fig. 4a,b) and experimental results (Fig. 4c,d) reproduced the varied magnitude and direction of spurious PSD peak shifts introduced by standard

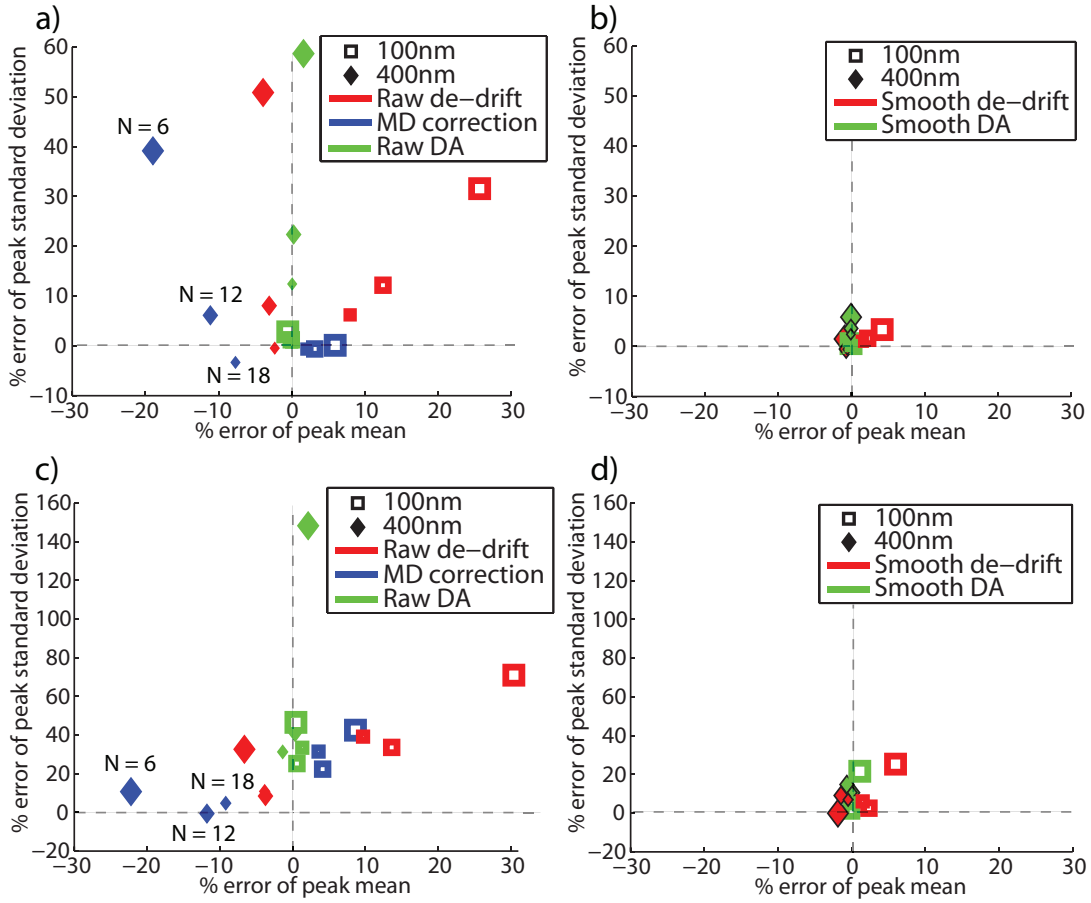

FIG. 4. **Simulations and experiments demonstrate that DA is more accurate than de-drifting for constant, uniform bulk flow measurement conditions**. Plots show % error of distribution parameters obtained from analysis of SPT simulations using the 'raw' estimated drift (a) and 'smooth' estimated drift (b), as well as from a particle tracking experiment using the 'raw' estimated drift (c) and 'smooth' estimated drift (d). Marker shape represents size mode, marker color represents the approach used for determining particle diameters, and marker size indicates the target number of particles in each time step – either 6 (large), 12, or 18 (small). (a)(b) For simulations, particle diameters were sampled from bi-modal 100- and 400-nm lognormal distributions (σ = 10nm) and prescribed trajectories with a random number of time steps (between 100–1000) in the presence of 65μm/s bulk fluid flow (∼10 seconds to traverse NanoSight® field of view). Plotted values represent averages from 100 simulations. (c)(d) For experimental data, 100nm- and 400nm-diameter polystyrene bead standards were tracked in quiescent fluid, and the number of particles in each time step was controlled in post-processing. In both simulations and experiments, de-correlating after de-drifting obtained the lowest % error of peak mean of all approaches for a given number of particles in each time step.

de-drifting approaches when used for dilute, polydisperse suspensions. For both 'raw' and 'smooth' de-drifting (red markers - Fig. 4), results corroborated previous figures – varying the total number of particles tracked in each time step modulated the magnitude of peak shifts, as illustrated by red markers moving away from the origin along the x axis for decreasing ensemble number. Plots also reproduced the different direction of peak shifts for 100- and 400nm subpopulations. The positive % error of peak mean values for 100nm subpopulations indicate spuriously over-estimated particle diameter, whereas the negative error values for 400nm subpopulations indicate spuriously under-estimated particle diameters. Two competing influences caused by de-drifting using the ensemble-averaged displacement are responsible: one, correlated subtraction removes Brownian motion, increasing inferred particle diameter; and two, improperly interpreted noise adds Brownian motion, decreasing inferred particle diameter.

DA (Eq. (12)) more accurately determined PSD peak means than de-drifting without a de-correlation step (Eq. (2)). Results in Fig. 4 show reduced % error magnitude for DA peak means (green markers) relative to de-drifted peak means (red markers). De-drifting introduced spurious peak shifts dependent on particle number, but the addition of a de-correlation step removed this dependence on particle density, thus all green markers regardless of time step ensemble number align at $x = 0$. A monodisperse correction method [26](blue markers) does not improve the peak mean estimate (Fig. 4a,c).

Smoothing the estimated drift signal dramatically improved inferred PSD peak standard deviation accuracy for both standard de-drifting and de-drifting plus de-correlation (Fig. 4b,d). Smoothing the estimated drift signal effectively augmented the ensemble size by considering the motion of particles in neighboring time steps, and larger ensembles will reduce error introduced by de-drifting. For experimental data analyzed by standard de-drifting with 6 particles per time step and a moving window average spanning 5 time steps ($N_w = 30$), the % error in the distribution standard deviation decreased from 32.6% to -0.1% (Fig. 4c,d). Similarly, corresponding simulations of $N_w = 30$ show a reduction from 148% to 11% in the % error of standard deviation.

Uniform, time-varying drift

When dynamic bulk flows contribute to particle motion, smoothing the estimated drift time series may not just reduce noise in that signal but also eliminate high-frequency components of true fluid motion. To demonstrate the effects of smoothing estimated drift under measurement conditions of high frequency oscillatory bulk flows (as expected with vibration in experimental setups), SPT simulations (Fig. 5a) and experiments (Fig. 5b) were performed with applied uniform, constant and uniform, oscillatory superimposed bulk flows. Simulated bulk flows matched experimental drift as estimated by Fourier-transform analysis of the experimental ensemble-averaged displacement time series. For calculating % error with experimental data, bi-modal distribution means and standard deviations were first determined under quiescent conditions. Experimental particle position data in quiescent fluid could be post-processed to control the number of particles per frame but the same approach is not possible with dynamic flows because particles at different time points in the experiment also experience different fluid motion.

Simulations show that smoothing the ensemble-averaged displacement time series, i.e. estimated drift, removed true bulk fluid motion and decreased measurement accuracy for dilute, bi-modal particle suspensions tracked in oscillatory bulk flows (Fig. 5a). Removing the true oscillatory fluid motion from the estimated drift vector meant that high-frequency components of bulk flow were misinterpreted as particle Brownian motion, and 'smooth' DA did not properly isolate particle Brownian motion. As a result, 'smooth' DA underestimated particle diameters and % error of peak mean spuriously shifted away from the origin (blue markers). Because true fluid motion was spuriously removed, the % error of peak mean inferred using a 'smooth' analysis (blue markers) approached diameters inferred from Eq. (4) with no drift subtraction step at all (data not shown). The conclusion from Fig. 5 is that, while smoothing can be advantageous under the proper conditions, smoothing can disadvantageous if bulk fluid motion contains high-frequency components.

Particle tracking experiments validated simulation results, demonstrating the importance of maintaining temporal resolution of the estimated drift signal when oscillatory bulk flows were present (Fig. 5b). Similar to simulation, a 'smooth' approach removed high frequency components of the true bulk fluid motion and the corresponding % error of peak mean (blue markers) approached values determined without a drift subtraction step (data not shown). On the other hand, a 'raw' approach successfully compensated for dynamic drift, as demonstrated by % error magnitudes close to the origin (green markers). Experimental validation shows that 'raw' and 'smooth' approaches obtained similar magnitude but oppositely signed % errors of peak standard deviations.

Mono-modal PSD with high dispersity

In Section III D, SPT simulations of suspensions with bi-modal PSDs demonstrated that, in a suspension with high dispersity, standard de-drifting could bias smaller particles toward larger diameters and larger particles toward smaller diameters, with a net biasing effect that peaks shifted toward each other. More generally, this effect leads to an overall compression of the PSD, which is important for interpreting broadly distributed populations. For example, NTA of biological extracellular vesicles often measures a PSD with a single, broad peak [41]. In this section, SPT simulations reflecting the important experimental scenario of a mono-modal PSD with high dispersity were used to compare DA with current de-drifting approaches, which do not account for heterogeneous particle diameters.

SPT simulations showed that de-drifting of suspensions with mono-modal PSDs introduced a spurious peak shift in the distribution mean, which was most accurately corrected by DA. Particle trajectory spans of varying length began at random time steps within the simulation such that the number of particles for each time step varied around 4 particles. Although these simulated suspensions were mono-modal, particles were not mono-disperse given the relative wide peak centered at 165nm. De-drifting spuriously shifted the MC-averaged distribution mean from 165nm to 201nm (open circles *vs.* red marks, Fig. 6a). A monodisperse correction factor partially reduced the spurious peak shift, obtaining an MC-averaged distribution mean of 152nm. However, neither de-drifting nor de-drifting with a monodisperse correction factor measured the MC-averaged distribution mean as accurately as DA, which found a peak mean at 167nm. Thus, the error in peak mean varied from 22%, -8% and 1% for raw de-drifting, a monodisperse correction and raw DA, respectively.

De-drifting spuriously narrowed the PSD relative to the PSD made up of sampled particle diameters ($\sigma =$ 75nm), and this effect was again best corrected by DA. De-drifting spuriously narrowed the PSD, determining an

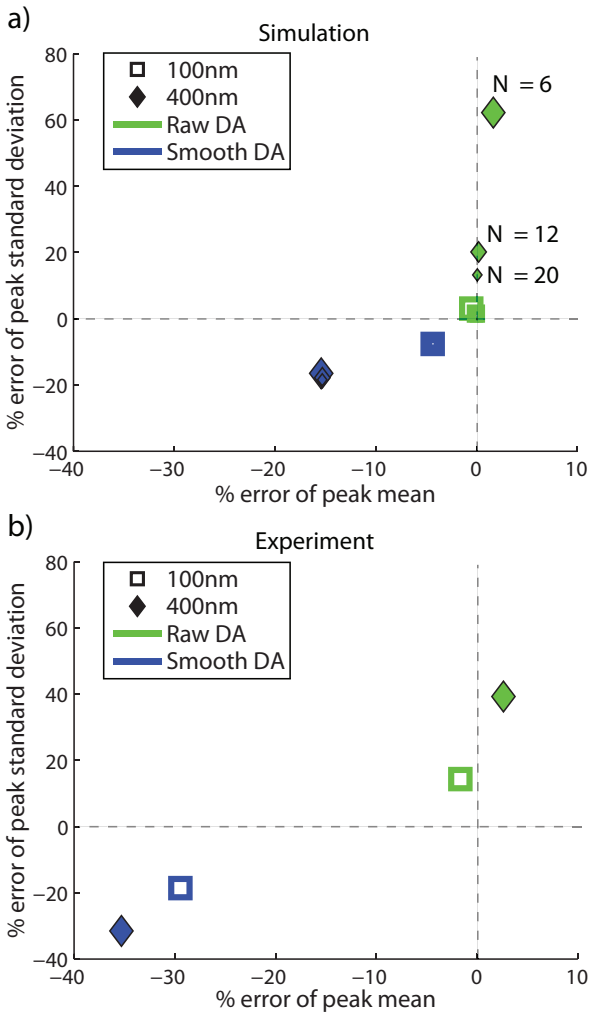

FIG. 5. **A 'raw' approach determined more accurate PSDs than a 'smooth' approach in oscillatory bulk flow conditions**. Plots show % error of distribution parameters obtained from DA of SPT simulations (a) and data from a particle tracking experiment (b) in which uniform bulk flows with constant and oscillatory components were present. The frequency spectrum of applied simulated bulk flow was matched to the frequency spectrum of the experimental ensemble-averaged displacement time series using Fourier-transform analysis. Marker shape represents size mode, marker color represents the approach used for determining particle diameters and, for simulations only, marker size indicates the target number of particles in each time step – either 6 (large), 12, or 18 (small). In both simulations and experiments, 'raw' DA bulk flow removal obtained the lowest % error of peak mean of all approaches for a given number of particles in each time step. Smoothing the estimated drift time series removed true oscillatory drift such that subtracting the 'smooth' estimated drift did not remove high frequency bulk flow components, spuriously underestimating inferred particle diameters.

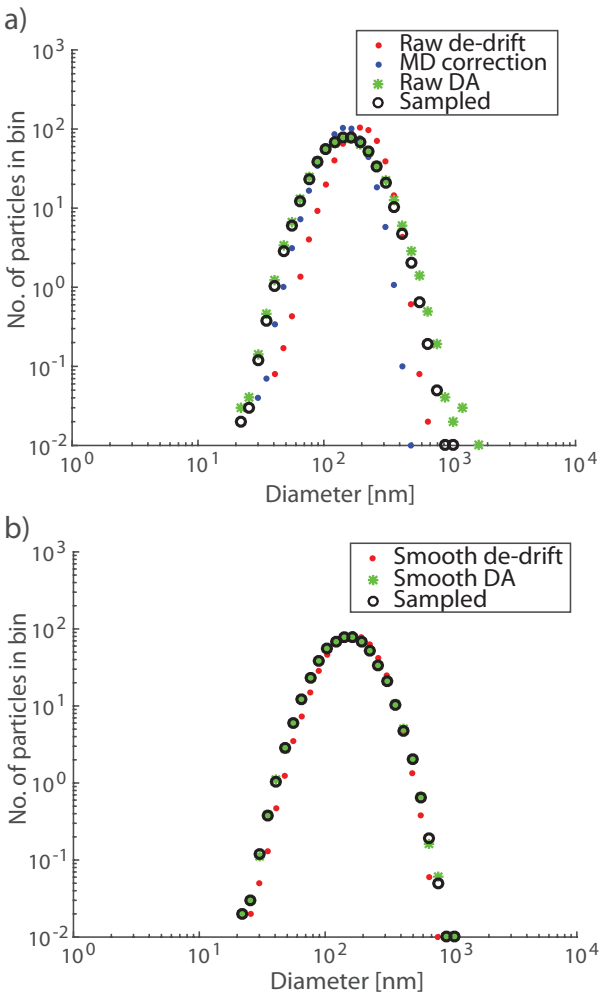

FIG. 6. **Simulations demonstrate that DA is more accurate than de-drifting for a broad, mono-modal PSD**. Plots show MC-averaged binned histograms of particle diameter obtained from analysis of SPT simulations using 'raw' estimated drift (a) and 'smooth' estimated drift ($w = 3$) (b). Particle diameters were sampled from a lognormal PSD (open circles) defined by $f_{\mathrm{LogN}}(x; \mu = 165\mathrm{nm}, \sigma = 75\mathrm{nm})$, and prescribed trajectories with a random number of time steps between 1000–2000 that began at random time steps such that $N_t \approx 4$ particles. DA (Eq. (12), green asterisks), is compared to standard de-drifting (Eq. (8), red marks) and to standard de-drifting with a monodisperse correction factor (blue marks) [26, 31]. Plotted values represent averages from 100 simulations. DA histogram values were nearly identical to binned sampled diameters, demonstrating high measurement accuracy regardless of particle diameter. On the other hand, current de-drifting methods both spuriously shift and spuriously narrow the PSD.

MC-averaged standard deviation of 65nm. A monodisperse correction factor further decreased the standard deviation to 49nm, whereas DA better estimated peak width, finding a standard deviation of 82nm. Thus, the error in peak standard deviation varied from −13%, −35% and 9% for raw de-drifting, a monodisperse correc-

tion and raw DA, respectively. Narrowed distributions follow directly from the varied biasing of particle diameters as a function of relative diffusive mobility – larger particles are spuriously measured smaller because noise in the estimated drift adds apparent Brownian motion whereas smaller particles are spuriously measured larger because Brownian motion is spuriously subtracted. Section III D isolated this pinching effect using two separate peaks in the PSD, which effectively showed individual particle diameter biasing because peaks were very narrow. In this section, Fig. 6 shows that the integrated individual particle diameter bias on a broad distribution spuriously narrows the distribution peak. The monodisperse correction factor intends to correct for these spurious shifts introduced by de-drifting, but cannot account for polydispersity. Thus, the ratio of σ/μ remains constant as all particle diameters are shifted equivalently. DA more accurately corrects spurious shifts caused by de-drifting because DA properly accounts for different particle diameters; however, peak width generally increases slightly. Peak broadening occurs because the corrections are functions of other particles in the frame, which are themselves of uncertain diameter. This leads to systematic deviation from a normal distribution to a t-distribution, which has wider tails.

Time averaging the estimated drift before vector subtraction greatly improved the measured distribution for both standard de-drifting and DA. The MC-averaged distribution parameters for raw and smooth de-drifting methods were $f_{\mathrm{LogN}}(x; \mu = 201\mathrm{nm}, \sigma = 65\mathrm{nm})$ and $f_{\mathrm{LogN}}(x; \mu = 175\mathrm{nm}, \sigma = 71\mathrm{nm})$, respectively, for an averaging window span $w = 3$ time steps. Similarly, time averaging improved the measured distribution parameters for raw and smooth DA, resulting in $f_{\mathrm{LogN}}(x; \mu = 167\mathrm{nm}, \sigma = 82\mathrm{nm})$ and $f_{\mathrm{LogN}}(x; \mu = 165\mathrm{nm}, \sigma = 75\mathrm{nm})$. The disadvantage of time averaging is reduced temporal resolution of high frequency bulk motion components, as demonstrated in Fig. 5. Importantly, Raw DA measured more accurate distribution means, showing that DA is preferred over standard de-drifting when the spectral components of bulk fluid motion are unknown.

V. DISCUSSION

NTA is an increasingly popular, SPT-based characterization method that accurately and quickly determines colloidal particle size. Though NTA has over a 100-year history [42], only in the last couple of decades have camera and image-processing technology progressed sufficiently to enable wide-spread adoption. Recent commercial products such as the NanoSight® (Malvern Instruments, Malvern, UK) [8] and the ZetaView (Particle Metrix, Meerbusch, Germany) as well as publicly available software have popularized particle tracking techniques [43]. NTA is often perceived as a characterization method well-suited for polydisperse suspensions because it analyzes each particle independently; however, we show that de-drifting correlates individual particle motion for dilute suspensions. As a result, de-drifting can introduce significant error.

Some of the limitations of de-drifting dilute suspensions have been known for at least a decade [25], but the solution presented to date was suitable only for monodisperse suspensions [26, 31]. Our work resolves known limitations of de-drifting associated with low particle concentration as well as identifies and resolves composition-dependent limitations. Thus, this work extends the measurement accuracy of NTA to dilute suspensions with arbitrary size distributions that are measured in uniform bulk flows with unknown dynamics. Future work could use local ensembles to further extend this work to nonuniform bulk flows [31]. Gains in the accuracy of NTA for dilute, polydisperse suspensions are important because the reliability with which NTA measures polydisperse PSDs remains an open question [41, 44]. For example, recent data characterizing patient placental extracellular vesicles using NanoSight® demonstrated that applied flow significantly affected results, but the study did not determine whether the difference was physical (issue with chamber design) or analytical (issue with algorithm) [45].

Since its introduction a decade ago, the NanoSight® has demonstrated a remarkable versatility of application [46]. In addition to standard diameter and concentration measurements, NanoSight® scattering intensity data has been used to identify several extracellular vesicle subpopulations with different refractive index [47]. To the authors' knowledge, our method is the first to use exported position data from NanoSight®, which allowed investigation of different particle tracking analysis methods without constructing a custom video microscopy suite or developing particle tracking algorithms. It was observed that NanoSight® artificially cuts particle trajectories short after ∼100 time steps. Interestingly, particles that remain visible are then re-tracked, and in this manner a single particle may produce multiple particle trajectories without ever exiting the field of view. Whether counting single particles multiple times is part of the proprietary analysis or simply an artifact of data exportation is unknown.

Decorrelation Analysis facilitates several fields of study. Investigations of extracellular vesicles have demonstrated several subpopulations normally shed from a variety of cell types, but the effect of various stimuli on vesicle production is not well characterized [48]. For example, the study of small molecule inhibitors that suppress microvesicle production are greatly benefited by accurate size measurements of dilute, polydisperse suspensions [33]. Similarly, microdevices using size-based separation methods to nanoparticle populations for downstream assaying must use accurate cutoff diameters in the design process [49–51]. Malvern Inc. recently developed the ability to perform simultaneous nanoparticle size, zeta potential and scattering intensity characterization [52]. The zeta potential measurement, however, relies on accurate analysis of non-Brownian motion as-

sociated with particle electrophoresis and bulk fluid electroosmosis and thermal effects [53]. Our work shows that care should be taken with zeta potential measures of heterogeneous particle populations. Because our approach has the benefit of increased accuracy with minimal loss of temporal resolution, DA could facilitate dynamic studies, such as the dynamics of protein aggregation [54]. DA can also be used in passive particle tracking microrheology to measure the local viscosity of a complex fluid, where arguments of heterogeneous diameters in homogeneous fluid are substituted with homogeneous diameters in heterogeneous fluid.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144153. Experiments were performed in the Nanobiotechnology Center shared research facilities at Cornell University.

[1] M. J. Rust, M. Bates, and X. Zhuang, Nat. Methods **3**, 793 (2006).
[2] D. Wirtz, Annu. Rev. Biophysics **38**, 301 (2009).
[3] V. Levi, Q. Ruan, and E. Gratton, Biophys. J. **88**, 2919 (2005).
[4] T. G. Mason, K. Ganesan, J. H. van Zanten, D. Wirtz, and S. C. Kuo, Phys. Rev. Lett. **79**, 3282 (1997).
[5] B. Wang, J. Kuo, S. C. Bae, and S. Granick, Nat. Mater. **11**, 481 (2012).
[6] N. Yang, J. L. Hutter, and J. R. de Bruyn, J. Rheol. **56**, 797 (2012).
[7] C. Finder, M. Wohlgemuth, and C. Mayer, Part. Part. Syst. Char. **21**, 372 (2004).
[8] A. Malloy and R. Carr, Part. Part. Syst. Char. **23**, 197 (2006).
[9] M. J. Saxton and K. Jacobson, Annu. Rev. Biophys. Biomol. Struct. **26**, 373 (1997).
[10] R. N. Ghosh and W. W. Webb, Biophys. J. **66**, 1301 (1994).
[11] T. G. Mason, A. Dhople, and D. Wirtz, in *MRS Proc. Stat. Mech. Phys. Bio.*, Vol. 463 (1997) pp. 153–158.
[12] J. C. Crocker and D. G. Grier, J. Colloid Interf. Sci. **179**, 298 (1996).
[13] J. Bewerunge, I. Ladadwa, F. Platten, C. Zunke, A. Heuer, and S. U. Egelhaaf, Phys. Chem. Chem. Phys. **18**, 18887 (2016).
[14] X. Michalet and A. J. Berglund, Phys. Rev. E **85**, 061916 (2012).
[15] P. K. Relich, M. J. Olah, P. J. Cutler, and K. A. Lidke, Phys. Rev. E **93**, 042401 (2016).
[16] N. Hoze and D. Holcman, Phys. Rev. E **92**, 052109 (2015).
[17] J. G. Walker, Meas. Sci. Technol. **23**, 065605 (2012).
[18] H. Saveyn, B. De Baets, O. Thas, P. Hole, J. Smith, and P. Van der Meeren, J. Colloid Interf. Sci. **352**, 593 (2010).
[19] M. Röding, H. Deschout, K. Braeckmans, A. Särkkä, and M. Rudemo, J. Microscopy **252**, 79 (2013).
[20] J. Mellnik, M. Lysy, P. A. Vasquez, N. S. Pillai, D. B. Hill, J. Cribb, S. A. McKinley, and M. G. Forest, J. Rheol. **60**, 379 (2016).
[21] T. Savin and P. S. Doyle, Biophys. J. **88**, 623 (2005).
[22] A. Kowalczyk, C. Oelschlaeger, and N. Willenbacher, Meas. Sci. Technol. **26** (2015).
[23] J. C. Crocker, M. T. Valentine, E. R. Weeks, T. Gisler, P. D. Kaplan, A. G. Yodh, and D. A. Weitz, Phys. Rev. Lett. **85**, 888 (2000).
[24] C. L. Vestergaard, P. C. Blainey, and H. Flyvbjerg, Phys. Rev. E **89**, 022726 (2014).
[25] J. C. Crocker and B. D. Hoffman, in *Methods Cell Biol.*, Vol. 83 (2007) pp. 141–178.
[26] C. J. Rowlands and P. T. C. So, Appl. Phys. Lett. **102**, 21913 (2013).
[27] P. C. Nelson, C. Zurla, D. Brogioli, J. F. Beausang, L. Finzi, and D. Dunlap, J. Phys. Chem. B **110**, 17260 (2006).
[28] E. Koslover, C. Chan, and J. Theriot, Biophys. J. **110**, 700 (2016).
[29] K. Chen, B. Wang, J. Guan, and S. Granick, ACS Nano **7**, 8634 (2013).
[30] E. J. Fong, Y. Sharma, B. Fallica, D. B. Tierney, S. M. Fortune, and M. H. Zaman, Ann. Biomed. Eng. **41**, 837 (2013).
[31] M. Khan and T. G. Mason, J. Phys.-Condens. Mat. **28** (2016).
[32] T. Savin and P. S. Doyle, Phys. Rev. E **76**, 021501 (2007).
[33] S. M. Santana, M. A. Antonyak, R. A. Cerione, and B. J. Kirby, Phys. Biol. **11**, 065001 (2014).
[34] C. Gardiner, Y. J. Ferreira, R. A. Dragovic, C. W. Redman, and I. L. Sargent, J. Extracell. Vesicles **2**, 19671 (2013).
[35] V. Pelletier, N. Gal, P. Fournier, and M. L. Kilfoil, Phys. Rev. Lett. **102**, 188303 (2009).
[36] J. Mellnik, P. A. Vasquez, S. A. McKinley, J. Witten, D. B. Hill, and M. G. Forest, Soft Matter **10**, 7781 (2014).
[37] A. Nicolet, F. Meli, E. van der Pol, Y. Yuana, C. Gollwitzer, M. Krumrey, P. Cizmar, E. Buhr, J. Pétry, N. Sebaihi, B. de Boeck, V. Fokkema, R. Bergmans, and R. Nieuwland, Meas. Sci. Technol. **27**, 035701 (2016).
[38] K. Miller, Math. Mag. **54**, 67 (1981).
[39] H. Qian, M. Sheetz, and E. Elson, Biophys. J. **60**, 910 (1991).
[40] Z. Li, S. Y. Mak, A. Sauret, and H. C. Shum, Lab Chip **14**, 744 (2014).
[41] E. van der Pol, F. Coumans, A. Grootemaat, C. Gardiner, I. L. Sargent, P. Harrison, A. Sturk, T. van Leeuwen, and R. Nieuwland, J. Thromb. Haemost. **12**, 1182 (2014).
[42] C. Bigg, Stud. Hist. Philos. Sci. **39**, 312 (2008).
[43] E. Meijering, O. Dzyubachyk, and I. Smal, in *Methods Enzymol.*, Vol. 504 (2012) pp. 183–200.
[44] W. Anderson, D. Kozak, V. a. Coleman, Å. K. Jämting, and M. Trau, J. Colloid Interf. Sci. **405**, 322 (2013).
[45] M. Tong, O. S. Brown, P. R. Stone, L. M. Cree, and

L. W. Chamley, Placenta **38**, 29 (2016).
[46] R. Carr and M. Wright, *Nanoparticle Tracking Analysis (NTA): the first 1000 reports*, Tech. Rep. (Malvern Instruments, 2015).
[47] C. Gardiner, M. Shaw, P. Hole, J. Smith, D. S. Tannetta, C. W. Redman, and I. L. Sargent, J. Extracell. Vesicles **3**, 25361 (2014).
[48] K. W. Witwer, E. I. Buzás, L. T. Bemis, A. Bora, C. Lasser, J. Lotvall, E. N. Nolte-'t Hoen, M. G. Piper, S. Sivaraman, J. Skog, C. Théry, M. H. M. Wauben, and F. Hochberg, J. Extracell. Vesicles **2**, 20360 (2013).
[49] S. M. Santana, M. A. Antonyak, R. A. Cerione, and B. J. Kirby, Biomed. Microdevices **16**, 869 (2014).
[50] K. Lee, H. Shao, R. Weissleder, and H. Lee, ACS Nano **9**, 2321 (2015).
[51] K. K. Zeming, N. V. Thakor, Y. Zhang, and C.-H. Chen, Lab Chip **16**, 75 (2016).
[52] A. Sikora, D. Bartczak, D. Geißler, V. Kestens, G. Roebben, Y. Ramaye, Z. Varga, M. Palmai, A. G. Shard, H. Goenaga-Infante, and C. Minelli, Anal.A Methods **7**, 9835 (2015).
[53] D. Griffiths, W. Bernt, P. Hole, J. Smith, A. Malloy, and R. Carr, in *NSTI-Nanotech 2011*, Vol. 1 (2011) pp. 4–7.
[54] D. T. Yang, X. Lu, Y. Fan, and R. M. Murphy, AIChE J. **60**, 1236 (2014).