Supplementary Material

E. F. Koslover, C. K. Chan, and J. A. Theriot

S1. HAAR WAVELET

The Haar wavelet is commonly used to approximate velocities over different time-scales from noisy trajectories. For instance, it has been employed in single particle tracking studies to distinguish actively transported vesicles from those diffusing passively in the cytoplasm[1]. This wavelet is defined by the coefficients,

$$w_j = \frac{1}{n(n+1)} \begin{cases} 1 & \text{for } 0 < j \leq n \\ -1 & \text{for } -n \leq j < 0 \\ 0 & \text{for } j = 0. \end{cases} \tag{S1}$$

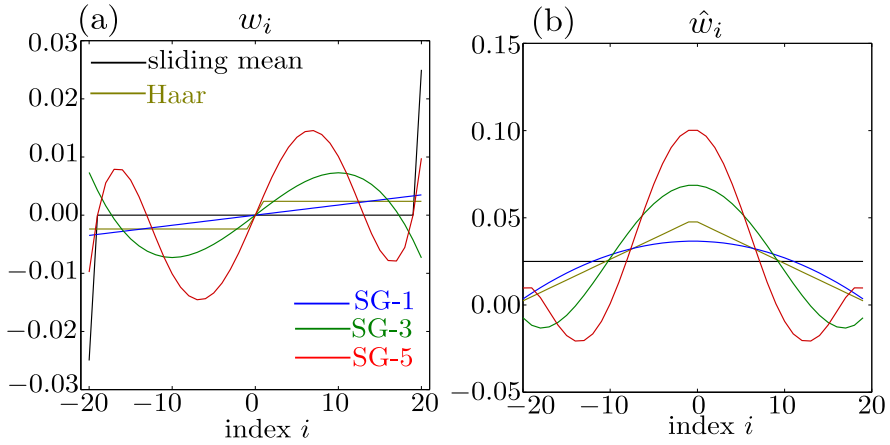

FIG. S1: Coefficients for different wavelet shapes, shown for span $n = 20$. (a) w_i coefficients for the trajectory positions. (b) $\hat{w}_i$ coefficients for the trajectory steps. Wavelet shapes shown are the sliding mean wavelet (black), Haar wavelet (yellow), and $1^{\text{st}}, 3^{\text{rd}}$, and 5^{th} order Savitzky-Golay wavelets (blue, green, and red, respectively).

The simple form of the Haar wavelet enables direct calculation of the summations in Eq. 10 to yield a closed form solution for the rescaling functions,

$$A^{(n)}_{k,\mathrm{Haar}} = \begin{cases} \frac{1}{60n^2(n+1)^2}\left[3k^5 - k^4(10n+5) + k^3(40n^2+40n-5) - k^2(80n^3+120n^2+30n-5)\right. & \\ \left. +k(60n^4+120n^3+80n^2+20n+2)\right] & \text{for } k \le n \\ \frac{1}{60n^2(n+1)^2}\left[-k^5 + k^4(10n+5) - k^3(40n^2+40n+5) + k^2(80n^3+120n^2+30n-5)\right. & \\ \left. -k(80n^4+160n^3+60n^2-20n-6) + (44n^5+110n^4+80n^3+10n^2-4n)\right] & \text{for } n < k < 2n \\ \frac{(2n+1)(3n^2+3n+4)}{30n(n+1)} & \text{for } k \ge 2n \end{cases} \tag{S2}$$

$$B^{(n)}_{k,\mathrm{Haar}} = \begin{cases} \frac{1}{n^2(n+1)^2}\left[-k^3 + k^2(2n+1) - 4kn(n+1) + 2n(n+1)(n^2+n+1)\right] & \text{for } k \le n \\ \frac{1}{3n^2(n+1)^2}\left[k^3 - 3k^2(2n+1) + 2k(6n^2+6n+1) + 2n^3(3n+4) - 12n^2(n+1) - 2n\right] & \text{for } n < k < 2n \\ \frac{6n^2-2n+2}{3n(n+1)} & \text{for } k \ge 2n \end{cases}$$

S2. SAVITZKY-GOLAY WAVELETS

The BNEW method relies on the drift velocity varying slowly in time, so that it can be effectively subtracted out by applying a wavelet that calculates a local estimate of the velocity across a window of span n. In the case where the drift velocity is not constant in time, one can estimate it more effectively by performing a less aggressive smoothing of the data. This can be done by either decreasing the span n, or by using a non-linear local approximation for the velocity. In either case, decreasing the smoothing also results in a more noisy estimate for the local velocity.

A common data smoothing technique uses the Savitzky-Golay filter[2], which involves fitting a local polynomial to each window of the data series. Higher-order polynomial fits provide less aggressive smoothing of the data. In practice, the Savitzky-Golay filter is applied by convolving each window of the data with a set of weights based on the least-square regression to the polynomial of a given degree[3]. Consequently, this smoothing technique can be formulated as a wavelet, with the particular set of weights w_i defined below.

To find the best-fit polynomial of degree d to data points p_i in a window $-n \le i \le n$, we can perform ordinary least-squares regression. The coefficients $a_0, \ldots, a_d$ of such a polynomial are given by

$$\vec{a} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{p} \tag{S3}$$

where $\mathbf{A}$ is a $(2n+1) \times (d+1)$ matrix with $A_{i,j} = i^j, -n \le i \le n, 0 \le j \le d$. The local first derivative (velocity) approximation at the center of the window is given by a_1/Δ, where Δ is the time step. The weights for each of the data points used to compute this approximation are simply the second ($j = 1$) row of the matrix $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$. That is, the Savitzky-Golay wavelet is defined by

$$\vec{w} = (0, 1, 0, \ldots, 0) \cdot (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T. \tag{S4}$$

We note that since the best fit polynomial to a trajectory with constant velocity $p_i = yi + z$ has $a_1 = y$, then the required constraints on the wavelet shape (Eq. 4) are satisfied. It can be shown[3]

that the Savitzky-Golay filter for calculating first derivatives is identical for degree d and $d+1$ where d is an odd integer. We thus limit our discussion to filters of odd degree. The wavelet associated with the degree d Savitzky-Golay filter will be abbreviated as SG-d. The shape of such wavelets in terms of both the weights w_i associated with the positions and the equivalent weights $\hat{w}_i$ associated with the trajectory steps are plotted in Fig. S1. Unless otherwise noted, SG-3 wavelets are used throughout this work, for reasons discussed in Supplemental Section S6.

S3. BNEW ANALYSIS FOR FRACTIONAL BROWNIAN MOTION

We now analyze the behavior of the BNEW method when applied to particle trajectories where the stochastic component of the motion corresponds to fractional Brownian motion[4]. Discrete increments from a fractional Brownian motion process $\vec{p}(t)$ at time steps of size Δ are referred to as discrete fractional Brownian noise, and defined by

$$\vec{v}_i = \frac{\vec{p}[(i+1)\Delta] - \vec{p}[i\Delta]}{\Delta}. \tag{S5}$$

The $\vec{v}_i$ are identically normally distributed, with mean 0, variance $4D\Delta$, and covariance function (for $k \geq 1$) [5, 6]

$$\langle \vec{v}_i \cdot \vec{v}_{i+k} \rangle = 2D\Delta^{\alpha-2} \left(|k+1|^\alpha - 2|k|^\alpha + |k-1|^\alpha \right), \tag{S6}$$

This model is appropriate for thermal motion in a viscoelastic material. Note that throughout this work we assume that the stochastic motion occurs independently in each of two dimensions, though this approach can be easily extended to thre-dimensional trajectories. Sampling of discrete fractional Brownian noise for purposes of simulation was carried out using the circulant embedding approach for generating stationary processes[5].

When BNEW analysis is applied to a trajectory consisting of constant drift, fractional Brownian motion, and localization error, the adjusted MSD is derived from Eq. 6, 10 as

$$\begin{aligned} MSD_k^{(n)} = \left\langle |\vec{p}_k^{(n)} - \vec{p}_0^{(n)}|^2 \right\rangle =& 4D\Delta^\alpha A_k^{(n)} + 4\epsilon^2 B_k^{(n)} \\ &+ 4D\Delta^\alpha \sum_{i=-n}^{k+n-2} \sum_{j=1}^{k+n-2-i} c_i^{(n,k)} c_{i+j}^{(n,k)} \left((j+1)^\alpha - 2(j)^\alpha + (j-1)^\alpha \right), \end{aligned} \tag{S7}$$

where $c_i^{(n,k)}$ refers to the coefficients in Eq. 8 associated with a particular combination of wavelet span n and time separation k.

Rescaling the adjusted MSD by $B_k^{(n)}$ as for diffusive motion gives

$$\widehat{\mathrm{MSD}}_k^{(n)} = 4D\Delta^\alpha \tilde{A}_k^{(n)} + 4\epsilon^2, \tag{S8}$$

where $\tilde{A}_k^{(n)}$ is plotted as a function of the rescaled time $\hat{t}_k = A_k^{(n)}/B_k^{(n)}$ in Fig. S2a,b, for both the Haar wavelet and the 3$^{\mathrm{rd}}$ order Savitzky-Golay wavelet. The function $\tilde{A}_k^{(n)}$ depends on α and on the shape of the wavelet used, but not on the parameters D, ϵ.

We note that the results for different wavelet spans n collapse approximately onto a power-law curve, with downward "hooks" at larger values of k. So long as the maximal k values are relatively small compared to the span n, then $\tilde{A}_k^{(n)}$ can be approximately fit by the functional form,

$$\tilde{A}_k^{(n)} = f(\alpha)\hat{t}_k^\alpha + g(\alpha)^2 \tag{S9}$$

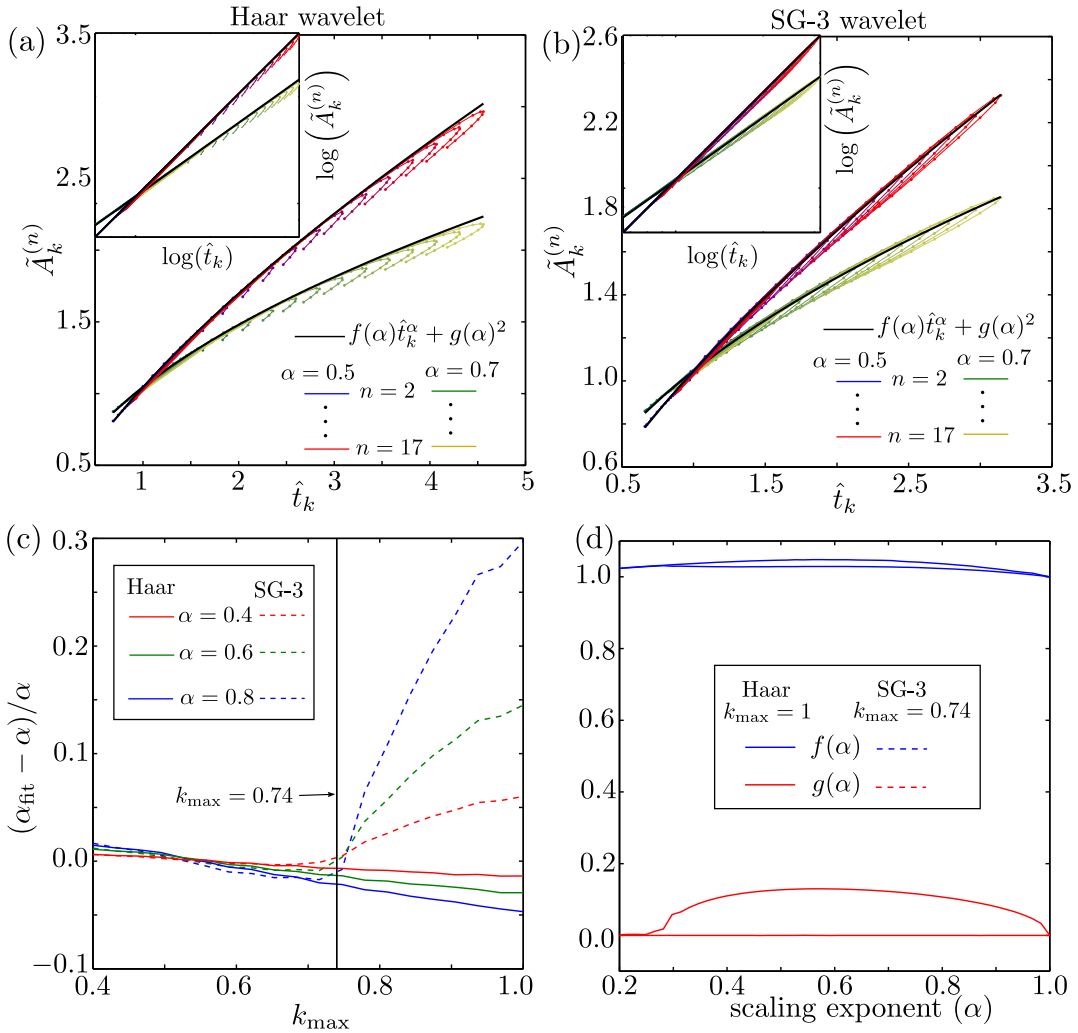

FIG. S2: BNEW analysis applied to trajectories consisting purely of fractional Brownian motion (no localization error or drift). (a) Function $\tilde{A}_k^{(n)}$ plotted against the rescaled time, for $\alpha = 0.5$ and $\alpha = 0.7$. Haar wavelet spans $2 \leq n \leq 17$ are included. Black lines correspond to the functional form in Eq. S9. Inset shows same data plotted on log-log axes. (b) Same results plotted using 3$^{\rm rd}$ order Savitzky-Golay (SG-3) wavelets for the wavelet analysis. (c) Fractional bias in fitted scaling exponent $\alpha_{\rm fit}$, as a function of maximal k value used in the fits, for Haar wavelets (solid lines) and SG-3 wavelets (dashed lines). (d) Function values $f(\alpha)$ (blue) and $g(\alpha)$ (red), for Haar wavelets (solid line) and SG-3 wavelets (dashed line), obtained by fitting $\tilde{A}_k^{(n)}$ to Eq. S9 for $2 \leq n \leq 17$, $1 \leq k \leq \lfloor k_{\rm max} n \rfloor$

.

We are interested primarily in extracting an accurate estimate of the viscoelastic scaling parameter α from such fits, in order to determine whether the particle motion is consistent with a viscous or a viscoelastic environment. Using wavelet spans $2 \leq n \leq 17$ for the fits, we plot the error in the parameter $\alpha_{\rm fit}$ depending on the maximal k values used in the fit. Specifically only values of k such that $1 \leq k \leq \lfloor k_{\rm max} n \rfloor$ are used for each span n. When fitting data from a limited number of trajectories, larger values of $k_{\rm max}$ would retain more information from the data, allowing for a less noisy fit. However, at higher values of $k_{\rm max}$, the downward "hooks" in the rescaled curves lead to incorrect estimations of the scaling parameter. We thus use the maximal value that allows for an accurate estimation of $\alpha_{\rm fit}$ (see Fig. S2c). Namely, $k_{\rm max} = 0.74$ is used for all data analysis using

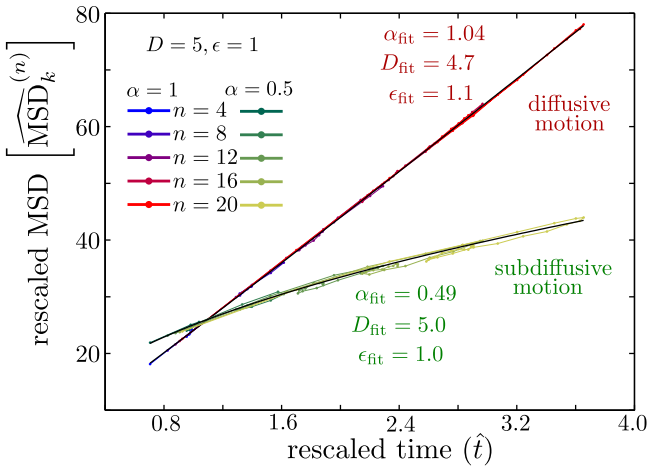

FIG. S3: Result of BNEW analysis with 3rd-order Savitzky-Golay wavelets on the same simulated trajectory dataset described in Fig. 1. Time separations $1 \leq k \leq \lfloor 0.74n \rfloor$ are shown for each wavelet span n. Fitted curves are shown in black, and the estimated values from the fits are listed for both the diffusive and the subdiffusive case.

3rd order Savitzky-Golay wavelets, while $k_{\max} = 1$ is used for analysis with Haar wavelets (*e.g.* Fig. 1).

The functions $f(\alpha)$ and $g(\alpha)$ are necessary for determining the effective diffusion coefficient (D) and localization error (ϵ), respectively. In particular the rescaled adjusted MSD is fit to,

$$\widehat{\rm MSD}_k^{(n)} = 4\hat{D}\Delta^\alpha \hat{t}_k^\alpha + 4\hat{\epsilon}^2, \tag{S10}$$

and the parameters D and ϵ are then extracted by

$$D_{\rm fit} = \frac{\hat{D}_{\rm fit}}{f(\alpha_{\rm fit})}, \tag{S11}$$

$$\epsilon_{\rm fit} = \sqrt{\hat{\epsilon}_{\rm fit}^2 - D_{\rm fit}\Delta^{\alpha_{\rm fit}} g(\alpha_{\rm fit})^2}. \tag{S12}$$

In the limit of diffusive motion ($\alpha = 1$), we have $f(1) = 1, g(1) = 0$, recovering Eq. 11. The functions f and g are determined numerically, by fitting $\tilde{A}_k^{(n)}$ as a function of $\hat{t}_k$ for $2 \leq n \leq 17$ for a range of α values (Fig. S2d).

The validity of this procedure for extracting estimates of α, D, ϵ to characterize the stochastic component of the motion is demonstrated for simulated trajectories in Fig. 1, using the Haar wavelet form. Analysis of the same trajectories with a 3rd order Savitzky-Golay wavelet is shown in Fig. S3.

S4. STATISTICS OF A PERSISTENT RANDOM WALK

To explore the effect of time-varying drift on the BNEW analysis method, we focus on a concrete model for the drift velocity as a continuous-time persistent random walk. This model has been used to describe the animal behavior[7], locomotion of motile cells[8], and paths of semiflexible polymers (where it is termed the worm-like chain)[9, 10]. It is appropriate for situations where there is a single characteristic timescale of persistence for the dynamical process. More complicated dynamics with multiple timescales can generally be expressed as sums of multiple persistent random walks.

One realization of a persistent random walk is motion with a constant speed (γ) along an orientation that varies diffusively in time with decorrelation time τ. In this section we derive some useful results regarding multi-point velocity correlations in this model. The propagator for the

persistent random walk in two dimensions satisfies

$$\frac{\partial G(\theta|\theta_0;t)}{\partial t} = \frac{1}{\tau}\frac{\partial^2 G(\theta|\theta_0;t)}{\partial \theta^2}, \tag{S13}$$

where $G(\theta|\theta_0;t)$ is the distribution of orientations for the velocity vector $\vec{u}$ at time t, given an orientation along θ_0 at time 0. This propagator can be expressed as

$$\begin{aligned} \partial G(\theta|\theta_0;t) &= \frac{1}{2\pi}\sum_{m=-\infty}^{\infty} g_m(t) e^{im(\theta-\theta_0)} \\ g_m(t) &= e^{-m^2 t/\tau}. \end{aligned} \tag{S14}$$

The velocity autocorrelation function is,

$$\begin{aligned} \langle \vec{u}(t)\vec{u}(0)\rangle =& \gamma^2 \int_0^{2\pi} d\theta \int_0^{2\pi} d\theta_0 \cos(\theta-\theta_0) G(\theta|\theta_0;t) \\ =& \frac{\gamma^2}{2}[g_1(t)+g_{-1}(t)] = \gamma^2 e^{-t/\tau} \end{aligned} \tag{S15}$$

At times shorter than the correlation time τ, the persistent random walk behaves like a ballistic motion with speed γ, while at longer times, the velocity orientation decorrelates and the motion appears diffusive with effective diffusion coefficient $\gamma^2\tau$. As a model for drift velocity, this formalism is convenient because it interpolates between constant drift ($\tau \to \infty$) and rapidly time varying drift $\tau \to 0$. In cases where the correlation time becomes comparable to the time step of the observed particle motion ($\tau \leq \Delta$), this drift velocity can no longer be distinguished from the diffusive component of the particle trajectories and BNEW analysis will not be accurate. In Fig. 2, we explore the bias and error introduced into the BNEW analysis in cases where the correlation time is finite but still much larger than the time step Δ.

In order to calculate errors in the fit parameters obtained with the BNEW method, we require also the fourth-order correlations in the particle velocities. For the persistent random walk, these correlations are given by,

$$\begin{aligned} \left\langle u_{i_0}^{(x)} u_{i_1}^{(x)} u_{i_2}^{(x)} u_{i_3}^{(x)} \right\rangle =& \left(\frac{\gamma}{2\pi}\right)^4 \int\int\int\int d\theta_0 d\theta_1 d\theta_2 d\theta_3 \cos\theta_0 \cos\theta_1 \cos\theta_2 \cos\theta_3 \\ & \times G(\theta_3|\theta_2,t_3) G(\theta_2|\theta_1,t_2) G(\theta_1|\theta_0,t_1) \\ =& \left(\frac{\gamma}{2}\right)^4 [2 g_1(t_1) g_2(t_2) g_1(t_3) + 4 g_1(t_1) g_0(t_2) g_1(t_3)] \\ =& \frac{\gamma^4}{4} e^{-(t_1+t_3)/\tau}\left[\frac{1}{2}e^{-4t_2/\tau}+1\right], \end{aligned} \tag{S16}$$

where $t_j = (i_j - i_{j-1})\Delta$ and the indices are ordered as $i_0 \leq i_1 \leq i_2 \leq i_3$. Similarly, we have

$$\begin{aligned} \left\langle u_{i_0}^{(x)} u_{i_1}^{(x)} u_{i_2}^{(y)} u_{i_3}^{(y)} \right\rangle &= \frac{\gamma^4}{4} e^{-(t_1+t_3)/\tau}\left[-\frac{1}{2}e^{-4t_2/\tau}+1\right] \\ \left\langle u_{i_0}^{(x)} u_{i_1}^{(y)} u_{i_2}^{(x)} u_{i_3}^{(y)} \right\rangle &= \left\langle u_{i_0}^{(x)} u_{i_1}^{(y)} u_{i_2}^{(y)} u_{i_3}^{(x)} \right\rangle = \frac{\gamma^4}{8} e^{-(t_1+4t_2+t_3)/\tau}, \end{aligned} \tag{S17}$$

and other combinations of the x, y dimensions can be obtained by symmetry.

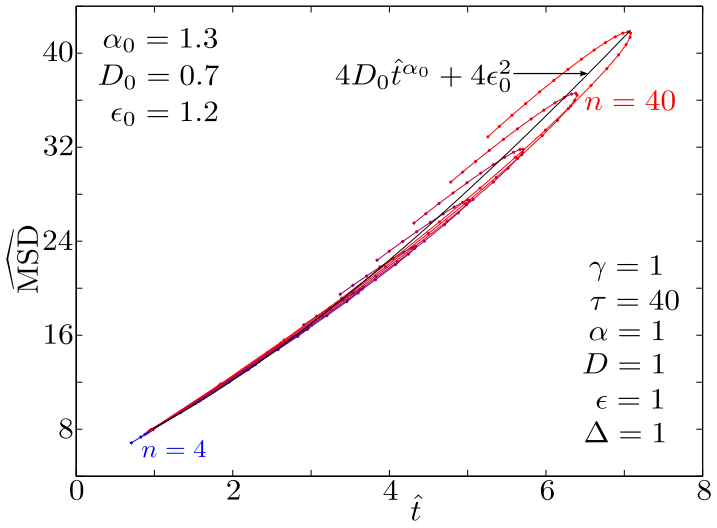

FIG. S4: Analytically calculated values of the rescaled adjusted MSD for trajectories consisting of persistent random walk drift with parameters $\tau = 40, \gamma = 1$, diffusive random motion $(\alpha = 1, D = 1)$, and localization error $(\epsilon = 1)$, sampled at time intervals of size $\Delta = 1$. SG-3 wavelets were used for the BNEW analysis, with spans between $n = 4$ and $n = 40$ (colored from blue to red). Black line corresponds to the power-law fit, showing the bias away from a linear function.

S5. ERROR ANALYSIS IN THE PRESENCE OF TIME-VARYING DRIFT

In this section, we calculate the bias and error in the fitted coefficients $\alpha_{\text{fit}}, D_{\text{fit}}, \epsilon_{\text{fit}}$, when the drift-velocity $\vec{u}$ varies in time as a persistent random walk (defined in Supplemental Section S4). Throughout this section, we restrict to the case with diffusive stochastic motion ($\alpha = 1$). The rescaled, adjusted MSD after BNEW analysis using wavelets of span n is

$$\widehat{\text{MSD}}_k^{(n)} = 4D\Delta t_k^{(n)} + 4\epsilon^2 + \frac{1}{B_k^{(n)}} \sum c_i^{(n,k)} c_j^{(n,k)} \langle \vec{u}_i \cdot \vec{u}_j \rangle \,, \tag{S18}$$

where $t_k^{(n)} = A_k^{(n)}/B_k^{(n)}$, as defined by Eq. 10, and the drift correlation function is given by Eq. S15, with $t = |i - j|\Delta$. The presence of the time-varying drift velocity causes the rescaled MSD curves for different wavelet spans to deviate from a universal line, increasing super-linearly and curling upwards at higher k values (Fig. S4). We find the bias in the fitted coefficients by fitting the power law Eq. 12 to the analytically calculated values of $\widehat{\text{MSD}}_k^{(n)}$ for $2 \le n \le n_{\max}$, $1 \le k \le \lfloor 0.74n \rfloor$. The parameter values from these fits are denoted $\alpha_0, \hat{D}_0, \hat{\epsilon}_0$.

In estimating the bias and the mean squared error of the fit parameters (described below), we linearize the fit function (Eq. 12) in terms of the parameters $\alpha, \hat{D}, \hat{\epsilon}$ in the neighborhood of $\alpha_0, \hat{D}_0, \hat{\epsilon}_0$. That is, we take

$$g(t_k^{(n)}, \vec{\chi}) \approx g(t_k^{(n)}, \vec{\chi}_0) + \mathbf{Z}(\vec{\chi} - \vec{\chi}_0), \quad Z_{i,j} = \frac{\partial g(t_i, \vec{\chi})}{\partial \chi_j}. \tag{S19}$$

Here, $\vec{\chi} = (\alpha, \hat{D}, \epsilon)$ is a vector of parameter estimates and $\mathbf{Z}$ is the matrix of partial derivatives of the fitted function g with respect to the parameters. This approximation is valid in the case where a large number of particle tracks are sampled, so that the measured MSD does not deviate far from the analytically calculated value (Eq. S18). In this case, the bias, defined as $\langle \vec{\chi}_{\text{fit}} - \vec{\chi} \rangle$ can be approximated as $\vec{\chi}_0 - \vec{\chi}$. The approximation breaks down in the case of limited sampling or when ϵ approaches 0, at which point additional bias inherent to the nonlinear regression can arise. To retain tractable error approximations, however, we focus on the linearizeable case. A comparison of bias and error estimates from simulated data as compared to our analytical approximations is shown in Fig. S5, for parameter values relevant to experimental data on lysosome motion in

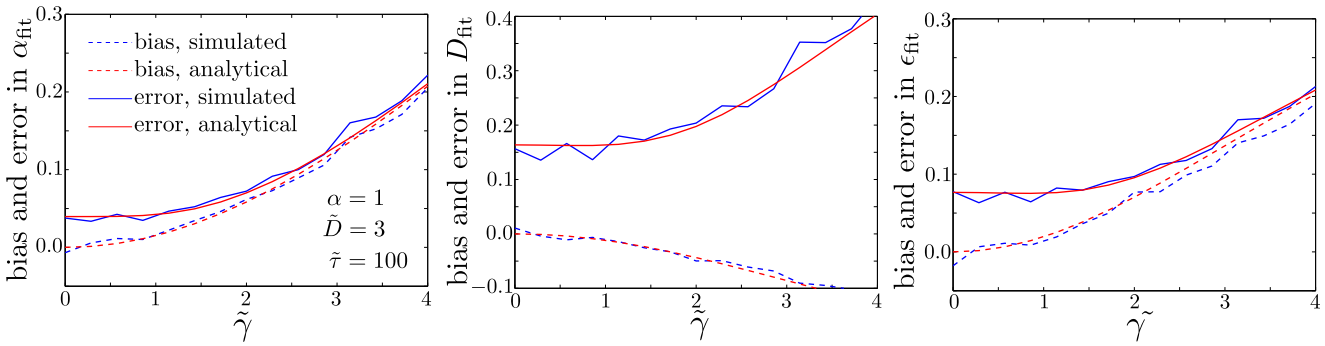

FIG. S5: Comparison of bias (dashed lines) and root mean squared error (solid lines) in fitted parameters $(\alpha_{\rm fit}, D_{\rm fit}, \epsilon_{\rm fit})$ as a function of dimensionless drift magnitude $\tilde{\gamma}$. Blue curves are from simulated data; red curves are from approximate analytical calculations. Errors are calculated for 200 trajectories of length 200 timepoints, with dimensionless parameters $\alpha = 1, \tilde{D} = 3, \tilde{\tau} = 100$, using SG-3 wavelets with spans $2 \leq n \leq 17$. Simulated values were averaged over 100 datasets.

.

HL60 cells (see Results section). We note that due to the presence of time-varying drift and the assumption of diffusive stochastic motion, we have $\alpha_0 \geq \alpha = 1$, so that we take $D_{\rm fit} = \hat{D}_{\rm fit}$, and $\epsilon_{\rm fit} = \hat{\epsilon}_{\rm fit}$ when calculating the bias and error.

We calculate an estimation of the total mean squared error (S^2) in the fitted parameters as,

$$S_\alpha^2 = (\alpha_0 - \alpha)^2 + \left\langle (\alpha_{\rm fit} - \alpha_0)^2 \right\rangle \tag{S20}$$

$$S_D^2 = (D_0 - D)^2 + \left\langle (D_{\rm fit} - D_0)^2 \right\rangle \tag{S21}$$

$$S_\epsilon^2 = (\epsilon_0 - \epsilon)^2 + \left\langle (\epsilon_{\rm fit} - \epsilon_0)^2 \right\rangle, \tag{S22}$$

where the first term arises from the bias due to deviation of the rescaled MSD from a universal line and the second term arises from noise in the sampled values. The brackets $\langle \ldots \rangle$ indicate averaging over a large set of trajectories, with fitted parameters extracted from the pooled set. These noise terms are taken from the diagonal of the covariance matrix of the fitted parameters, calculated as described below.

Under the linearization assumption, the covariance matrix of fitted parameters from performing a non-linear least-square fit to Eq. 12 is given by[11],

$$\mathbf{E} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{M}\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}, \tag{S23}$$

where $\mathbf{M}$ is the covariance matrix of the individual data points used in the fit and $\mathbf{Z}$ is the matrix of partial derivatives of the fitted function with respect to the fit parameters. Namely,

$$\begin{aligned} Z_{m,\alpha} &= 4D_0(\Delta\hat{t}_m)^{\alpha_0} \log\left(\Delta\hat{t}_m\right), \\ Z_{m,\hat{D}} &= 4(\Delta\hat{t}_m)^{\alpha_0} \\ Z_{m,\hat{\epsilon}} &= 4\epsilon_0 \end{aligned} \tag{S24}$$

where the index m runs over all individual data points used for fitting, including all combinations of wavelet span n and time interval k for the rescaled adjusted MSD.

Calculating the covariance of the individual data points is complicated by the fact that each

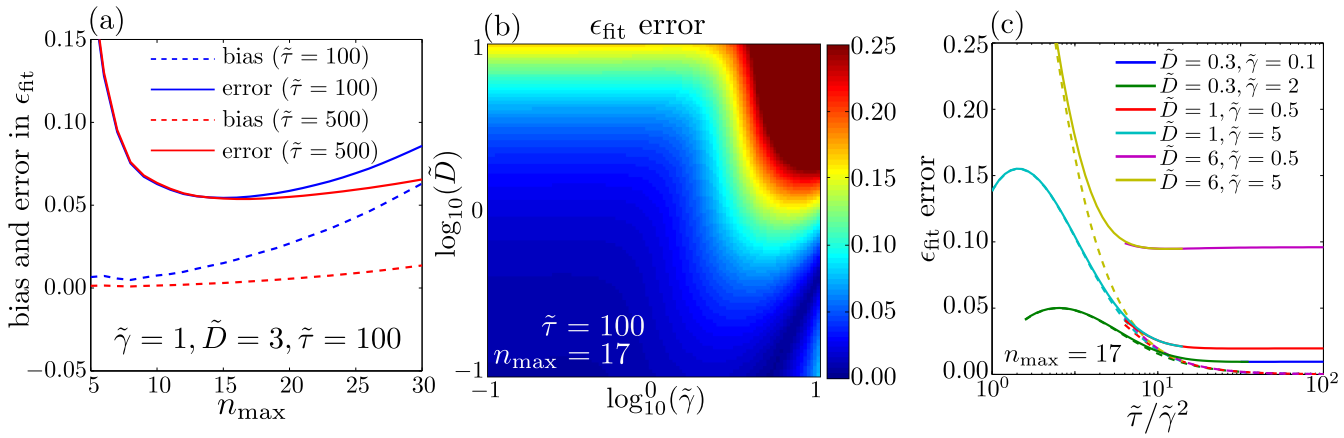

FIG. S6: Bias and root mean square error in the estimated value of the localization error $\epsilon_{\rm fit}$. (a) Bias (dashed curves) and error (solid curves) plotted as a function of maximal wavelet span $n_{\rm max}$, for $\tilde{D}=3, \tilde{\gamma}=1$, shown for drift correlation time $\tilde{\tau}=100$ (blue) and $\tilde{\tau}=500$ (red). (b) Root mean square error as a function of the dimensionless diffusion constant $\tilde{D}$ and drift magnitude, $\tilde{\gamma}$, assuming a persistece time of $\tilde{\tau}=100$. (c) Bias and error as a function of the compound parameter $\tilde{\tau}/\tilde{\gamma}^2$. SG-3 wavelets were used throughout and errors are calculated assuming 400 tracks of length 200 timesteps each.

adjusted MSD includes an average over many tracks and many time windows from each individual track. For simplicity, we assume in our error estimates that the individual particle tracks are completely independent from each other. Time averaging in the MSD calculation, however, can involve many non-independent intervals, and the co-dependence of these intervals must be taken into account when calculating data point covariance. We focus on the time average only, noting that the elements of the covariance matrix simply scale inversely with the number of tracks in the ensemble average, assuming individual tracks are independent of each other. For a track composed of $N+1$ sequential positions $\vec{p}_0, \ldots, \vec{p}_N$, the adjusted mean scaled displacement (for a span n and time separation k) is given by

$$\begin{aligned} \mathrm{MSD}_k^{(n)} &= \frac{1}{\mathcal{N}_k} \sum_{l=n_{\rm max}}^{N-n_{\rm max}-k} \eta_l^{(n,k)}, \\ \eta_l^{(n,k)} &= \left| \bar{p}_{l+k}^{(n)} - \bar{p}_l^{(n)} \right|^2 \\ &= \left| \Delta \sum_{i=-n}^{k+n-2} c_i^{(n,k)} \left(\vec{u}_{l+i} + \vec{v}_{l+i}\right) + \sum_{i=-n}^{k+n-1} \hat{c}_i^{(n,k)} \xi_{l+i} \right|^2 \\ \mathcal{N}_k &= N - k - 2n_{\rm max} + 1 \end{aligned} \tag{S25}$$

where the first and last $n_{\rm max}$ data points are dropped for all the different wavelet spans $n \le n_{\rm max}$ used in the BNEW analysis to avoid edge effects. The covariance matrix elements for the rescaled MSD are then,

$$\begin{aligned} M_{m_1,m_2} &= \frac{\mathrm{cov}\left(\mathrm{MSD}^{(m_1)}, \mathrm{MSD}^{(m_2)}\right)}{B^{(m_1)} B^{(m_2)}} \\ &= \frac{\sum_{l_1,l_2} \mathrm{cov}\left(\eta_{l_1}^{(m_1)}, \eta_{l_2}^{(m_2)}\right)}{B^{(m_1)} B^{(m_2)} \mathcal{N}_{k_1} \mathcal{N}_{k_2}} \end{aligned} \tag{S26}$$

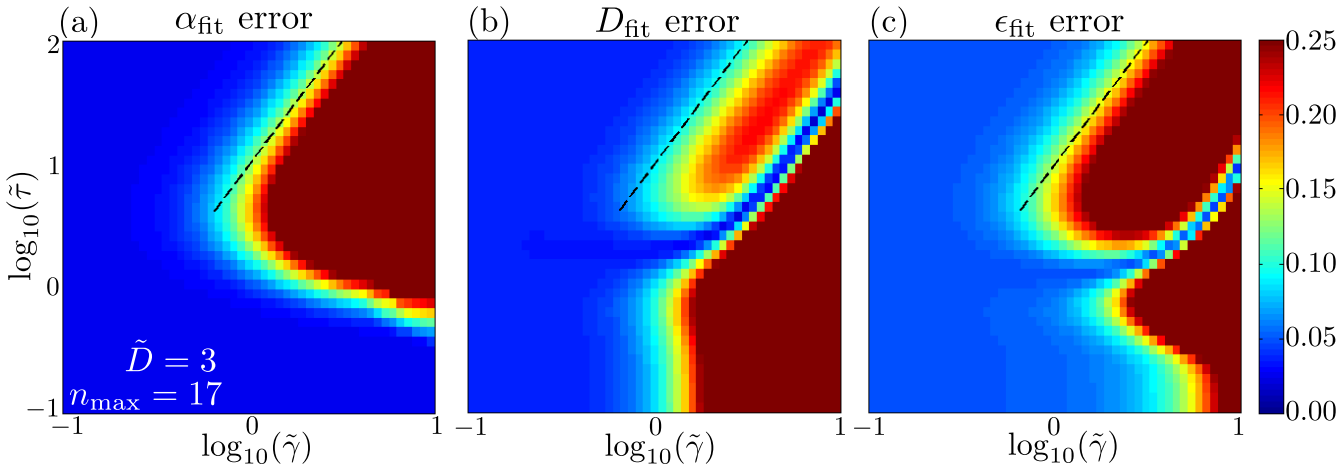

FIG. S7: Root mean square error in fit parameters (a) $\alpha_{\rm fit}$, (b) $D_{\rm fit}$, (c) $\epsilon_{\rm fit}$, as a functin of dimensionless magnitude ($\tilde{\gamma}$) and correlation time $\tilde{\tau}$ of the drift velocity. All calculations use SG-3 wavelets with spans $2 \leq n \leq 17$, assuming 400 tracks of length 200 timesteps each. Dashed black lines mark a constant value of $\tilde{\tau}/\tilde{\gamma}^2$.

where, again, the m indices refer to pairs of wavelet spans n and time separations k. Noting that the average $\langle \eta_{l_1} \eta_{l_2} \rangle$ depends only on the difference between the indices ($q = l_1 - l_2$), and assuming diffusive stochastic motion ($\langle \vec{v}_i \cdot \vec{v}_j \rangle = 0$ for $i \neq j$), it can be shown that

$$\begin{aligned}\mathrm{cov}\left(\eta_l^{(m_1)}, \eta_{l+q}^{(m_2)}\right) =& (\gamma\Delta)^4 H^{(u)}_{m_1,m_2}(q) + (2D\Delta)^2 H^{(v)}_{m_1,m_2}(q) + \epsilon^4 H^{(\xi)}_{m_1,m_2}(q) + 4\gamma^2 D\Delta^3 F^{(u)}_{m_1,m_2}(q) F^{(v)}_{m_1,m_2}(q) \\ &+ 2\gamma^2\epsilon^2\Delta^2 F^{(u)}_{m_1,m_2}(q) F^{(\xi)}_{m_1,m_2}(q) + 4D\epsilon^2\Delta F^{(v)}_{m_1,m_2}(q) F^{(\xi)}_{m_1,m_2}(q),\end{aligned} \tag{S27}$$

where the $\mathbf{H}$ and $\mathbf{F}$ matrices that couple drift velocity, diffusion, and localization error across time steps are defined below. The fourth-order covariance of the drift velocity is

$$H^{(u)}_{m_1,m_2}(q) = \frac{1}{\gamma^4} \sum_{i_1,j_1,i_2,j_2} c_{i_1}^{(m_1)} c_{j_1}^{(m_1)} c_{i_2}^{(m_2)} c_{j_2}^{(m_2)} \mathrm{cov}\left(\vec{u}_{i_1} \cdot \vec{u}_{j_1}, \vec{u}_{i_2+q} \cdot \vec{u}_{j_2+q}\right) \tag{S28}$$

whose matrix elements can be calculated using Eq. S15, S16, S17. The corresponding covariance matrices $\mathbf{H}^v$, $\mathbf{H}^\xi$ for the diffusive velocities and localization errors, respectively, can be simplified to

$$\begin{aligned} H^{(v)}_{m_1,m_2}(q) &= 4\sum_{i\in T}\left[c_i^{(m_1)} c_{i-q}^{(m_2)}\right]^2 + 8\sum_{i,j\in T; j>i} c_i^{(m_1)} c_j^{(m_1)} c_{i-q}^{(m_2)} c_{j-q}^{(m_2)} \\ H^{(\xi)}_{m_1,m_2}(q) &= 4\sum_{i\in T}\left[\hat{c}_i^{(m_1)} \hat{c}_{i-q}^{(m_2)}\right]^2 + 8\sum_{i,j\in T; j>i} \hat{c}_i^{(m_1)} \hat{c}_j^{(m_1)} \hat{c}_{i-q}^{(m_2)} \hat{c}_{j-q}^{(m_2)} \end{aligned}$$

where the range of indices in the summations is $T = [\max(-n_1, -n_2+q), \min(n_1+k_1-1, n_2+k_2-1+q)]$ and $c^{(n,k)}_{n+k-1} = 0$ by convention. In the above, we make use of the fact that both the diffusive velocities and the localization errors in each dimension are independent and normally distributed with mean 0, and that the fourth moment of a normal distribution is three times the variance.

Due to the independence between different components of the particle motion (drift, diffusion,

and localization error), the cross-covariance terms can be factored into individual second-order correlations. For instance,

$$\operatorname{cov}\left(\vec{u}_{i_1}\cdot\vec{v}_{j_1},\vec{u}_{i_2+q}\cdot\vec{v}_{j_2+q}\right)=\frac{1}{2}\left\langle\vec{u}_{i_1}\cdot\vec{u}_{i_2+q}\right\rangle\left\langle\vec{v}_{i_1}\cdot\vec{v}_{i_2+q}\right\rangle, \tag{S29}$$

where the factor of $1/2$ arises from the two independent dimensions of the diffusive velocity. We therefore define the following quantities to complete the covariance calculation in Eq. S27:

$$\begin{aligned}
F^{(u)}_{m_1,m_2}(q) &= \frac{1}{\gamma^2}\sum_{i_1,i_2} c_{i_1}^{(m_1)}c_{i_2}^{(m_2)}\left\langle\vec{u}_{i_1}\cdot\vec{u}_{i_2+q}\right\rangle \\
&= \sum_{i_1,i_2} c_{i_1}^{(m_1)}c_{i_2}^{(m_2)}e^{(-|i_2+q-i_1|\Delta/\tau)} && \text{(S30)}\\
F^{(v)}_{m_1,m_2}(q) &= 2\sum_{i\in T} c_i^{(m_1)}c_{i-q}^{(m_2)} && \text{(S31)}\\
F^{(\xi)}_{m_1,m_2}(q) &= 2\sum_{i\in T} \hat{c}_i^{(m_1)}\hat{c}_{i-q}^{(m_2)} && \text{(S32)}
\end{aligned}$$

The indices in Eq. S30 are taken from the range $-n_1\le i_1\le n_1+k_1-2, -n_2\le i_2\le n_2+k_2-2$. After computing the matrices $\mathbf{H}^{(u)},\mathbf{H}^{(v)},\mathbf{H}^{(\xi)},\mathbf{F}^{(u)},\mathbf{F}^{(v)},\mathbf{F}^{(\xi)}$, the covariance matrix $\mathbf{M}$ for the time-averaged, rescaled MSD can be calculated with the aid of Eq. S27 via

$$\begin{aligned}
M_{m_1,m_2} =& \frac{1}{B^{(m_1)}B^{(m_2)}\mathcal{N}_{k_1}\mathcal{N}_{k_2}}\Big[\min\left(\mathcal{N}_{k_1},\mathcal{N}_{k_2}\right)\operatorname{cov}\left(\eta_l^{(m_1)},\eta_l^{(m_2)}\right)+ \\
&+\sum_{q=1}^{\mathcal{N}_{k_2}-1}\min\left(\mathcal{N}_{k_1},\mathcal{N}_{k_2}-q\right)\operatorname{cov}\left(\eta_l^{(m_1)},\eta_{l+q}^{(m_2)}\right)+\sum_{q=1}^{\mathcal{N}_{k_1}-1}\min\left(\mathcal{N}_{k_2},\mathcal{N}_{k_1}-q\right)\operatorname{cov}\left(\eta_l^{(m_2)},\eta_{l+q}^{(m_1)}\right)\Big]
\end{aligned} \tag{S33}$$

We note that while the individual $\mathbf{H}^{(u)}(q)$ matrices are time-consuming to compute for any given value of the drift velocity correlation time τ, this computation need be done only once to find the covariance as a function of the remaining parameters (γ, D, ϵ).

The bias and error in the fitted scaling exponent α_{fit} and diffusion coefficient D_{fit} are shown in Fig. 2. The corresponding results for the fitted localization error ϵ_{fit} are plotted in Fig. S6. We note that the bias in the estimated localization error is determined primarily by the mean square displacement due to the time-varying component of the drift velocity that does not get corrected by the BNEW analysis. Thus, for cases with long correlation time ($\tilde{\tau}>10$) and relatively small drift, this bias is a function of the compound parameter $\tilde{\tau}/\tilde{\gamma}^2$ (Fig. S6c). When the effect of drift velocity is small, the overall error in ϵ_{fit} is determined by the sampling error, which increases with increasing $\tilde{D}$ (increasing magnitude of diffusion relative to localization error).

For less persistent drift ($1<\tau<10$), the bias in the scaling exponent drops back down as the persistent random walk itself approaches diffusive behavior over the timespan of the wavelet (Fig. S7a). As the persistence time becomes small relative to the time step ($\tau\lesssim 1$) the drift velocity as treated here (with displacement over each discrete step given by $\Delta\vec{u}_i$) approaches a chain of freely jointed steps. The average fitted diffusion coefficient then approaches the sum of the true diffusion and the effective diffusion arising from the drift ($\tilde{D}_0\approx\tilde{D}+\tilde{\gamma}^2/4$). The relative bias in D_{fit} goes to $\gamma^2/4D$ for small $\tilde{\tau}$ (Fig. S7b).

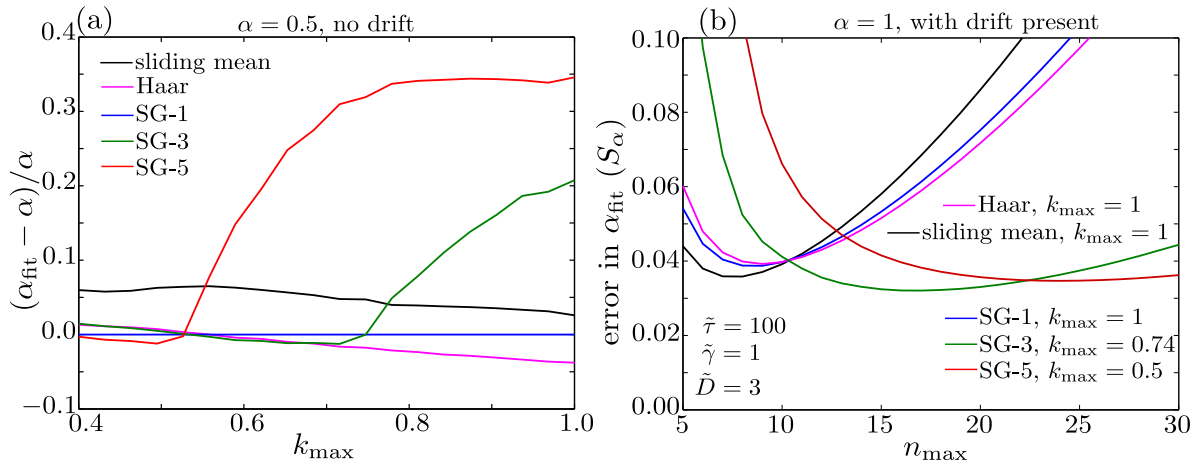

FIG. S8: Selecting optimal wavelet shape based on error estimates. (a) Fractional error in fitted exponent for the case of pure fractional Brownian motion, for different wavelet shapes, analogous to Fig. S2c. Fits are done for wavelet spans $2 \leq n \leq 17$ and $1 \leq k \leq \lfloor k_{\rm max} n \rfloor$. (b) Overall root mean squared error in the fitted scaling parameter for different wavelet shapes, as a function of maximal allowed wavelet span $n_{\rm max}$. The range of time separations k used for the fitting is selected based on panel (a). Errors are calculated assuming 400 trajectories of length 200 steps each.

S6. SELECTION OF WAVELET SHAPE

To select among a number of possible wavelet forms for use in BNEW analysis, we compare the estimated total error in the viscoelastic scaling $\alpha_{\rm fit}$ for a specific set of parameters relevant to the experimental data discussed in the Results section. The wavelet forms considered (plotted in Fig.S1) include the simplest sliding mean wavelet (characterized by constant step weights $\hat{w}_i = 1/(2n)$), the Haar wavelet (Supplemental Section S1), and different order Savitzky-Golay wavelets (SG-1, SG-3, SG-5, defined in Supplemental Section S2). The error in the fitted parameters depends additionally on the choice of $k_{\rm max}$, which sets how many time separations are considered for each wavelet span ($1 \leq k \leq \lfloor k_{\rm max} n \rfloor$). While increasing $k_{\rm max}$ up to $k_{\rm max} = 1$ will generally decrease the total error, it can also lead to over-estimating the scaling coefficient α in the case of subdiffusive stochastic motion, as discussed in Section S3. When fitting experimental data, we do not know *a priori* whether the particles are moving in a viscous or a viscoelastic medium, and thus want to avoid this source of bias in the α estimates. Consequently, while we use $k_{\rm max} = 1$ for the sliding mean, Haar, and SG-1 wavelets, we set $k_{\rm max} = 0.74$ for the SG-3 wavelet, and $k_{\rm max} = 0.5$ for the SG-5 wavelet, so that in the absence of drift, the scaling can be accurately characterized by $\alpha_{\rm fit}$ (see Fig.S8a and S2c). As plotted in Fig. S8b, the SG-3 wavelet gives the lowest error in the scaling estimate at an appropriately selected $n_{\rm max}$. These 3$^{\rm rd}$-order Savitzky-Golay wavelets are used throughout this work, unless otherwise specified. We note that increasing the degree of the polynomial fit in the SG wavelets also increases the optimal $n_{\rm max}$ as the less aggressive data smoothing more effectively subtracts out a time-varying drift. Selection of the appropriate $n_{\rm max}$ is discussed in Section S9.

S7. ERRORS FROM FITTING ORDINARY MSD CURVES

In order to compare the performance of the BNEW method to the traditional approach of fitting uncorrected mean squared displacement curves, we calculate the error in the fit parameters for the

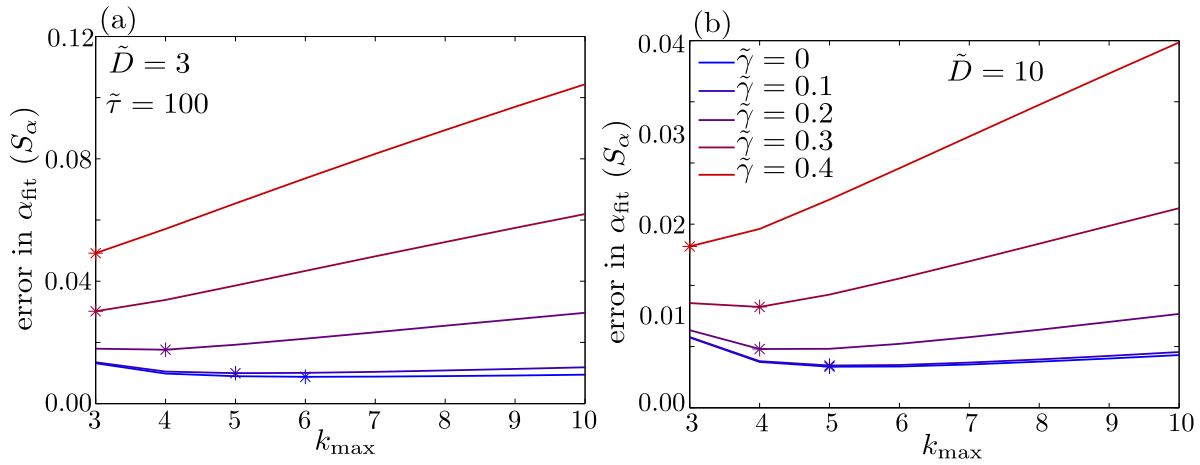

FIG. S9: Root mean squared error (S_α) in the fitted scaling parameter using ordinary MSD curves, without BNEW analysis. Time points $1 \leq k \leq k_{\rm max}$ are used for fitting. (a) For dimensionless diffusion constant $\tilde{D} = 3$, plotted for several values of drift velocity magnitude $\tilde{\gamma}$. Starred points indicate optimal value of $k_{\rm max}$ for each drift velocity. (b) For diffusion constant $\tilde{D} = 10$ with the same values of drift velocity. Errors are calculated assuming 400 trajectories with 200 timepoints each.

latter approach in the presence of drift velocities that behave as a persistent random walk. These errors can be obtained by applying the analysis in Section S5 with the "null" wavelet defined by $w_i = 0$. The corresponding scaling functions are $A_k = k, B_k = 1$, so that in the absence of drift, the rescaled MSD takes the form of the traditional mean-squared displacement,

$$\widehat{\rm MSD}_k = 4Dk + 4\epsilon^2. \tag{S34}$$

The drift velocity results in biased estimation of the parameters for stochastic motion, and the bias becomes more pronounced if a larger range of time separations k is used for the fit. For non-trivial magnitudes of the drift, the smallest mean squared error in the estimated parameters is achieved by taking the lowest possible $k_{\rm max} = 3$ (Fig. S9). This is the value used for the comparison to BNEW analysis in Fig. 3. It is interesting to note that even in the absence of persistent drift ($\gamma = 0$), the optimal range of k when fitting ordinary MSD curves is quite small: $k_{\rm max} = 6$ for tracks of length 200 timepoints, with $\tilde{D} = 3$, and even smaller values for higher dimensionless diffusion constant $\tilde{D}$. This effect has been pointed out previously based on similar calculations of the error associated with performing regressions on mean squared displacement curves.[12]

S8. BNEW ANALYSIS OF FRACTIONAL BROWNIAN MOTION WITH TIME-VARYING DRIFT

In the case where the stochastic component of the particle trajectories corresponds to fractional Brownian motion, the BNEW method is more sensitive to time varying drift than in the purely diffusive case. Intuitively, this is to be expected, since fractional Brownian motion results in smaller overall displacement of the particle relative to the drift, thus making this component of the motion more difficult to isolate for a given set of parameters $\tilde{\gamma}, \tilde{\tau}, \tilde{D}$. The bias in the estimated parameter values can be calculated by fitting the power-law curve to the analytical equation for rescaled

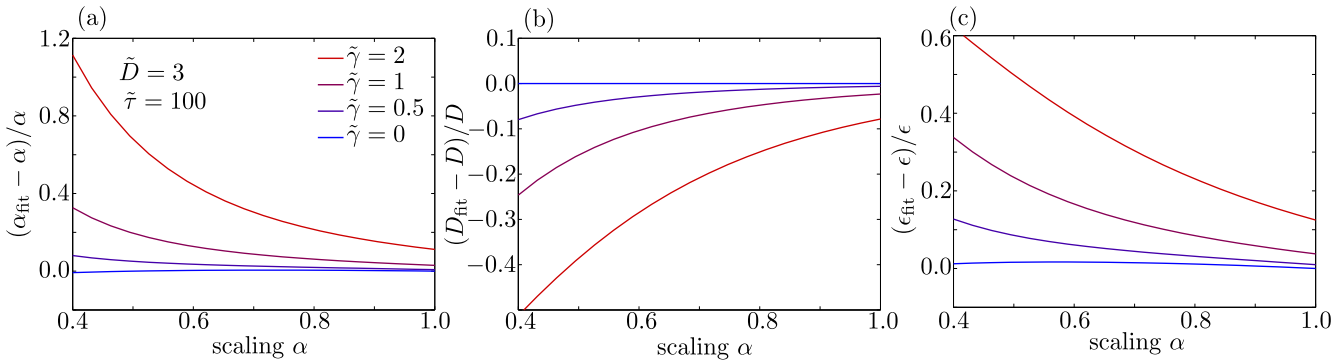

FIG. S10: Relative bias in fitted parameters from BNEW analysis of trajectories composed of fractional Brownian motion, localization error, and persistent random drift (with correlation time $\tilde{\tau} = 100$ and magnitude $\tilde{\gamma}$). (a) Bias in scaling exponent α; (b) bias in diffusion coefficient D; (c) bias in localization error ϵ. Dimensionless units are used throughout. SG-3 wavelets with spans $2 \leq n \leq 17$ and time separations $1 \leq k \leq \lfloor 0.74n \rfloor$ are used for the fits.

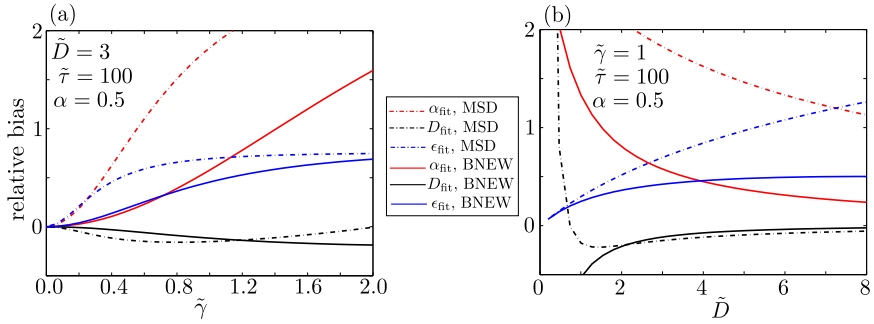

FIG. S11: Relative bias in fitted parameters from BNEW analysis (solid lines) as compared to fitting ordinary MSD curves (dashed lines), for trajectories composed of fractional Brownian motion (with scaling $\alpha = 0.5$), localization error, and persistent random drift (with correlation time $\tilde{\tau} = 100$ and magnitude $\tilde{\gamma}$). (a) Plotted versus drift magnitude; (b) plotted versus diffusion coefficient. Dimensionless units are used throughout. SG-3 wavelets with spans $2 \leq n \leq 17$ and time separations $1 \leq k \leq \lfloor 0.74n \rfloor$ are used for the fits.

MSD,

$$\widehat{\mathrm{MSD}}_k^{(n)} = 4D\Delta\tilde{A}_k^{(n)} + 4\epsilon^2 + \frac{1}{B_k^{(n)}} \sum c_i^{(n,k)} c_j^{(n,k)} \left\langle \vec{u}_i \cdot \vec{u}_j \right\rangle , \tag{S35}$$

where $\tilde{A}$ is defined by Eq. S7, S8. The power-law fits are done as a function of $t_k^{(n)}$.

The bias in the fitted coefficients is plotted as a function of the scaling exponent α for the fractional Brownian motion in Fig. S10. We note that for smaller values of α, the magnitude of drift velocity ($\tilde{\gamma}$) at which the BNEW analysis can give accurate estimates is more limited. Nonetheless, the accuracy of this method remains significantly better than fitting ordinary MSD curves for small magnitude drift (Fig. S11). We note that the advantage of the BNEW method in this case is particularly striking when determining the scaling coefficient $\alpha_{\rm fit}$, which can be estimated to within 50% for $\tilde{\gamma} < 1$ (when $\tilde{D} = 3, \tilde{\tau} = 100$). A fit of the ordinary MSD curve would require $\tilde{\gamma} < 0.3$ for the same level of accuracy.

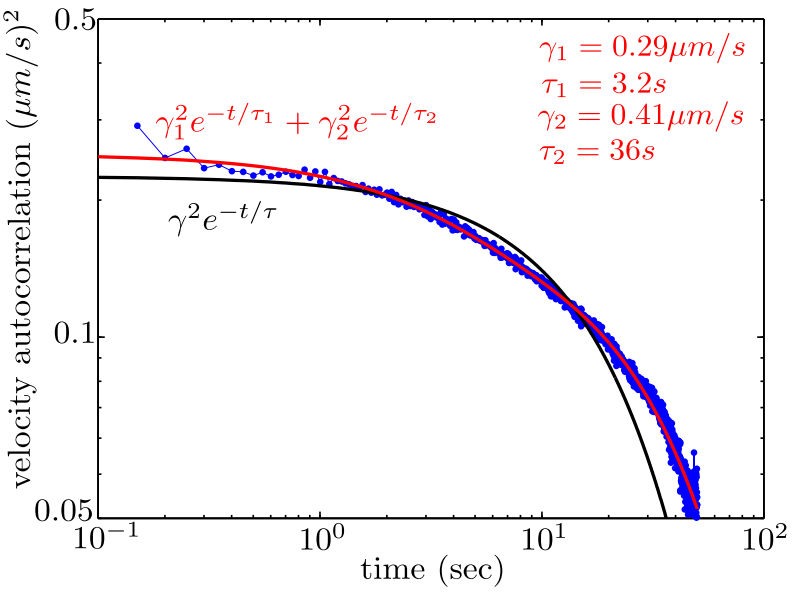

FIG. S12: Velocity autocorrelation function C_v (blue) for pooled tracks from 93 HL60 cells. Single exponential fit is shown in black and double exponential in red.

S9. SELECTION OF $n_{\max}$ FOR ANALYZING LYSOSOME MOTION

To determine the appropriate wavelet spans for the BNEW analysis of lysosome motion in neutrophils, it is necessary to first calculate an approximation of the drift velocity magnitude and correlation over time. We pool together lysosome tracks from all 93 cells and calculate the velocity autocorrelation function, defined as

$$C_v(t) = \frac{1}{\Delta^2} \langle [\vec{p}(t' + \Delta) - \vec{p}(t')] \cdot [\vec{p}(t' + t + \Delta) - \vec{p}(t' + t)] \rangle \tag{S36}$$

where the average is done over all tracks and all time points t'. The velocity autocorrelation contains components from the drift velocity, the stochastic motion and the localization error. However, if we assume that the stochastic particle motion is diffusive, then this should contribute only to the $t = 0$ correlation timepoint and if the localization error is uncorrelated in time, then it should contribute to the $t = 0$ and $t = \Delta$ timepoints only. We thus consider the autocorrelation function for $t \geq 2\Delta$ only to focus specifically on the drift velocity.

This autocorrelation is not well fit by a single exponential, but can be approximately fit to a double exponential decay (Fig. S12). We use this approximate functional form to select the optimal maximal wavelet span $n_{\max}$ to use in the BNEW analysis. Specifically, we approximate the drift as the composition of two independent persistent random walks with correlation times $\tau_1 = 1.1$ sec and $\tau_2 = 25$ sec and magnitudes $\gamma_1 = 0.49\mu m/s$ and $\gamma_2 = 0.58\mu m/s$, respectively. The error in the fitted parameters can then be calculated using a generalization of Eq. S27,

$$\begin{aligned} \mathrm{cov}\left(\eta_l^{(m_1)}, \eta_{l+q}^{(m_2)}\right) =& (\gamma_1\Delta)^4 H^{(u,1)}_{m_1,m_2} + (\gamma_2\Delta)^4 H^{(u,2)}_{m_1,m_2} + (2D\Delta)^2 H^{(v)}_{m_1,m_2} + \epsilon^2 H^{(\xi)}_{m_1,m_2} \\ &+2\Delta^2 \left[\gamma_1^2 F^{(u,1)}_{m_1,m_2} + \gamma_2^2 F^{(u,2)}_{m_1,m_2}\right] \left[2D\Delta F^{(v)}_{m_1,m_2} + \epsilon^2 F^{(\xi)}_{m_1,m_2}\right] \\ &+4D\epsilon^2\Delta F^{(v)}_{m_1,m_2} F^{(\xi)}_{m_1,m_2} + 2\gamma_1^2\gamma_2^2\Delta^4 F^{(u,1)}_{m_1,m_2} F^{(u,2)}_{m_1,m_2}, \end{aligned} \tag{S37}$$

where $\mathbf{H}^{(u,1)}, \mathbf{H}^{(u,2)}$ refer to the fourth order correlation matrices (Eq. S28) and $\mathbf{F}^{(u,1)}, \mathbf{F}^{(u,2)}$ to the

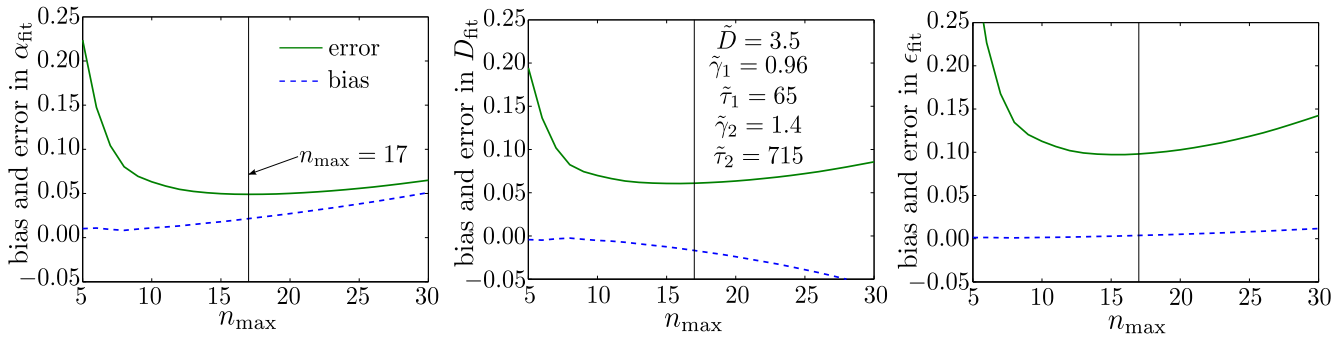

FIG. S13: Analytically approximated errors in estimated fit parameters with the BNEW method, assuming drift velocity composed of two independent persistent random walks with dimensionless correlation times $\tilde{\tau}_1, \tilde{\tau}_2$ and magnitudes $\tilde{\gamma}_1, \tilde{\gamma}_2$, respectively. Stochastic motion is assumed diffusive with dimensionless diffusion constant $\tilde{D}$. Time units are non-dimensionalized by the time step $\Delta = 0.05s$ and length units by the localization error $\epsilon = 0.015\mu m$. The drift velocity parameters are selected based on the pooled velocity autocorrelation function for all cells (Fig. S12), and the diffusion constant and localization error based on BNEW analysis results with SG-3 wavelets of span $2 \leq n \leq 17$ and time separations $1 \leq k \leq \lfloor 0.74n \rfloor$ (Fig. 5). Bias and error in $\tilde{D}_{\rm fit}$ are shown as fractions of $\tilde{D}$. Bias and error in $\alpha_{\rm fit}, \tilde{\epsilon}_{\rm fit}$ are shown in dimensionless units. Error calculations are done for a total of 391 tracks of length 118 time points each.

second order correlation matrices (Eq. S30), based on τ_1 and τ_2, respectively.

The analytically calculated errors in the fitted parameters for this approximate double-exponential drift velocity are plotted in Fig. S13 as a function of the maximal wavelet span $n_{\rm max}$. For this analysis, we assume 391 tracks of length 118 time points each, the mean values for the population of HL60 cells.

We find that the minimum error in estimating the viscoelastic scaling α occurs for $n_{\rm max} = 17$, which we use for the analysis of the experimental data. The dimensionless values of the diffusion coefficient ($\tilde{D} \approx 3.5$) and drift velocities ($\tilde{\gamma}_1 \approx 1, \tilde{\gamma}_2 \approx 1.4$) used to select the optimal $n_{\rm max}$ were taken from the BNEW analysis statistics of the population of HL60 cells (Fig. 5). The choice of wavelet span used for the analysis is thus self-consistent with the parameters obtained from the analysis itself.

S10. QUANTIFYING LOCALIZATION ERROR WITH A FIXED CELL

We obtain an independent measurement of the localization error inherent in our imaging protocol by tracking labeled lysosome particles in several HL60 cells fixed with formaldehyde/glutaraldehyde (see Experimental Methods). The mean squared displacement of the particles is approximately constant with time for short times (Fig. S14a), and the distribution of individual step sizes is shown in Fig. S14b. The localization error is estimated from the root mean squared displacement over a single time step,

$$\epsilon = \sqrt{\left\langle \left| \vec{p}(\Delta) - \vec{p}(0) \right|^2 \right\rangle / 4} \approx 0.014\mu m \tag{S38}$$

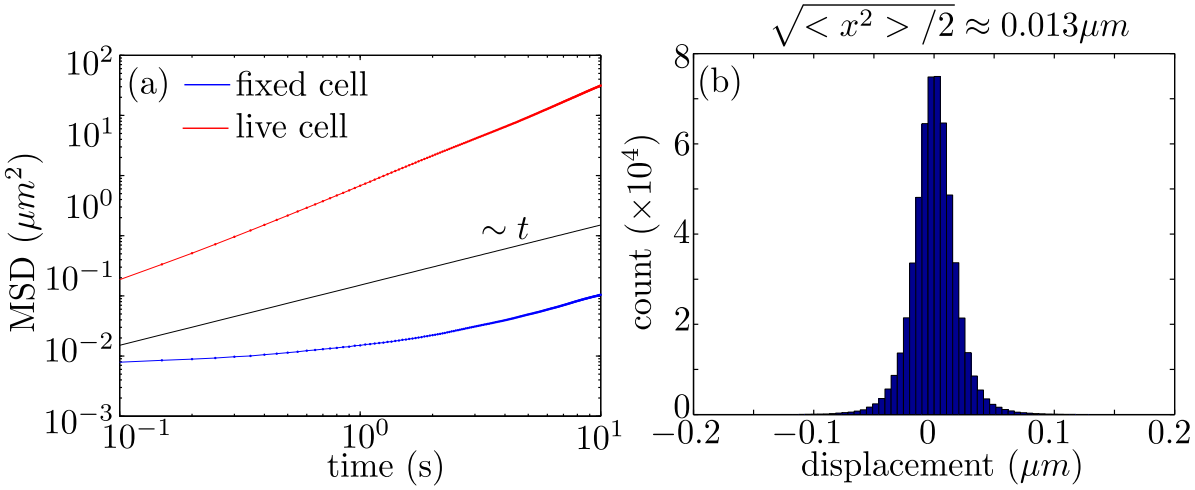

FIG. S14: Particle trajectory analysis for fixed cells (statistics from 11 cells total). (a) Mean squared displacement of lysosomes in fixed HL60 cells (blue) and a live cell (red, cell shown in Fig. 4). Linear scaling indicated with black line. (b) Histogram of step sizes for lysosome displacement in fixed cells, along a single dimension, over one time step ($\Delta = 0.05s$).

[1] K. Chen, B. Wang, J. Guan, and S. Granick, ACS Nano **7**, 8634 (2013).
[2] A. Savitzky and M. J. Golay, Anal Chem **36**, 1627 (1964).
[3] R. W. Schafer, IEEE Signal Proc Mag **28**, 111 (2011).
[4] W. Deng and E. Barkai, Phys Rev E **79**, 011112 (2009).
[5] D. P. Kroese and Z. I. Botev, arXiv preprint arXiv:1308.0399 (2013).
[6] S. C. Weber, J. A. Theriot, and A. J. Spakowitz, Phys Rev E **82**, 011913 (2010).
[7] P. Turchin, *Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants*, vol. 1 (Sinauer Associates Sunderland, 1998).
[8] D. Selmeczi, S. Mosler, P. H. Hagedorn, N. B. Larsen, and H. Flyvbjerg, Biophys J **89**, 912 (2005).
[9] C. Bouchiat, M. Wang, J.-F. Allemand, T. Strick, S. Block, and V. Croquette, Biophys J **76**, 409 (1999).
[10] H. Yamakawa, *Helical wormlike chains in polymer solutions* (Springer Science & Business Media, 2012).
[11] T. Kariya and H. Kurata, *Generalized least squares* (John Wiley & Sons, 2004).
[12] X. Michalet, Phys Rev E **82**, 041914 (2010).