Supporting Information:
Intrinsic cell-to-cell variance from experimental single-cell motility data

Anton Klimek,[1] Johannes C. J. Heyn,[2] Debasmita Mondal,[3,4] Sophia Schwartz,[5]
Joachim O. Rädler,[2] Prerna Sharma,[3,6] Stephan Block,[5] and Roland R. Netz[1]

[1]*Fachbereich Physik, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Faculty of Physics and Center for NanoScience,*
Ludwig Maximilians Universität, 80539 München, Germany
[3]*Department of Physics, Indian Institute of Science, 560012 Bangalore, India*
[4]*James Franck Institute, University of Chicago, 60637 Chicago, USA*
[5]*Fachbereich Chemie und Biochemie,*
Freie Universität Berlin, 14195 Berlin, Germany
[6]*Department of Bioengineering, Indian Institute of Science, 560012 Bangalore, India*

CONTENTS

I. AVERAGING VELOCITIES OVER POPULATIONS CAN RESULT IN NON-GAUSSIAN DISTRIBUTIONS

We show in the main text that the velocity distributions of the considered moving objects are Gaussian on the single-individual level in Fig. 2. Even though all moving objects are well described by Gaussian distributions on the individual level, averaging over the entire population can lead to non-Gaussian effects. In Fig. S1 we show that averaging with respect to the ensemble-averaged mean velocity $\overline{v}$ and the ensemble-averaged standard deviation $\overline{\sigma}$ according to

$$v = (v_{\text{ind}} - \overline{v})/\overline{\sigma} \tag{S1}$$

can lead to deviations from Gaussian behavior. These deviations are negligible for the beads shown in Fig. S1a. For the cancer cells deviations from Gaussian behavior are slightly increased when using Eq. (S1) compared to the rescaling on the individual level, as shown in Fig. S1b. For comparison, we show a fit of the Laplace distribution defined as $p(v) = e^{-|v/b|}/2b$ to the average data (black symbols) as an orange line in Fig. S1b. The Laplace distribution describes the center of the distribution slightly better than a Gaussian, however, the wings of the distribution are better described by a Gaussian. The deviations from purely Gaussian behavior for the black symbols in Fig. S1b could arise from non-local interaction with the environment, i.e. the substrate protein density in case of the cancer cells [31]. The possible influence of non-Gaussian effects on the GLE description is not considered in this work. For the algae cells the deviations from Gaussian behavior are quite drastic, when rescaling according to Eq. (S1). The results of Fig. S1 showcase that it is important to analyze motility data on the single-individual level and that averaging of observables over individual moving objects can lead to non-Gaussian effects.

II. MEMORY KERNEL EXTRACTION

Multiplying the GLE Eq. (1) by the initial velocity $\dot{x}(t_0)$, averaging over the random force and integrating from t_0 to t leads to

$$(C_{vv}(t) - C_{vv}(0)) = -\int_0^t C_{vv}(s)G(t-s)ds\,, \tag{S2}$$

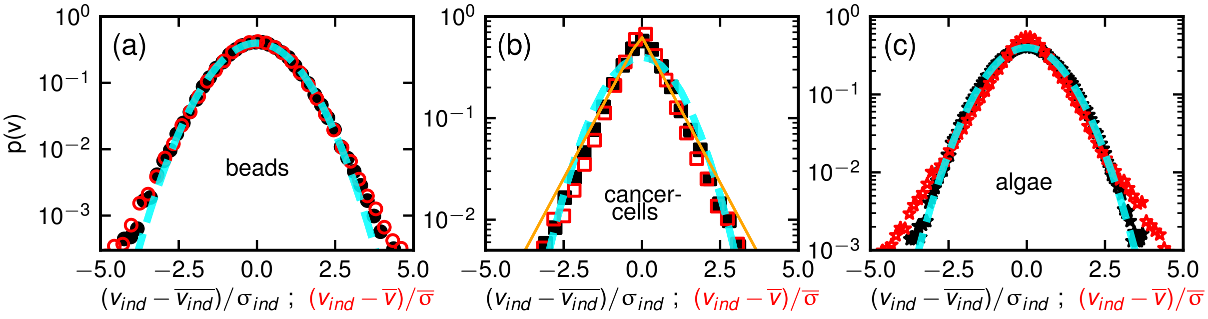

FIG. S1. Average velocity distribution for beads (a), cancer cells (b), algae (c). The black symbols correspond to the averages shown in Fig. 2b,e,h and result from rescaling with respect to individual trajectory means $\overline{v_{\text{ind}}}$ and standard deviations σ_{ind} according to the black x-axis label. Red symbols result from rescaling by the population mean $\overline{v}$ and standard deviation $\overline{\sigma}$ according to Eq. (S1). The dashed cyan line represents the standard normal distribution. The orange line in (b) corresponds to a fit of the Laplace distribution to the black data points.

where we used that $\langle \dot{x}(t_0) F_R(t) \rangle = 0$ [33–35], set $t_0 = 0$ and introduced the integral kernel

$$G(t) = \int_0^t \Gamma(s) ds \,. \tag{S3}$$

In order to invert Eq. (S2), we discretize it. Since $C_{vv}(t)$ is even while $G(t)$ is odd, we discretize $G(t)$ on half steps and $C_{vv}(t)$ on full steps and obtain [16]

$$G_{i+1/2} = \frac{2(C_{vv}^0 - C_{vv}^{i+1})}{\Delta(C_{vv}^1 + C_{vv}^0)} - \sum_{j=1}^{i} G_{i-j+1/2} \frac{C_{vv}^{j+1} + C_{vv}^j}{C_{vv}^1 + C_{vv}^0} \,. \tag{S4}$$

The kernel Γ_i is obtained by the discrete derivative $\Gamma_i = \frac{G_{i+1/2} - G_{i-1/2}}{\Delta}$ with the initial value $\Gamma_0 = 2G_{1/2}/\Delta$.

III. GOODNESS OF FIT TO INDIVIDUAL TRAJECTORY DATA

In Fig. 4 of the main text we show single-trajectory fit results of the models Eqs. (6), (7), (8) to randomly chosen single trajectory data. Since we cannot show the fit results of all individuals, we instead show the distribution of R^2 values for all fits in Fig. S2. Here, R^2 is defined using the cost function E_{cost} of Eq. (16) as

$$R^2 = 1 - \frac{E_{\text{cost}}}{\sum_{i=0}^{n-1} (C_{vv}^{\text{exp}}(i\Delta) - \overline{C_{vv}^{\text{exp}}})^2} \,, \tag{S5}$$

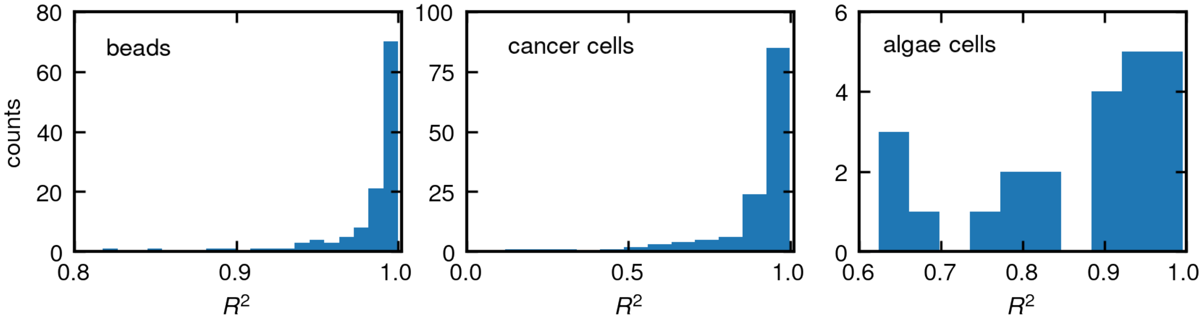

FIG. S2. Histogram of the R^2 values defined in Eq. (S5) for the fits of the respective model Eqs. (8),(7),(6) to individual trajectories for polystyrene beads (a), breast cancer cells (b), algae cells (c).

where $\overline{C_{vv}^{\mathrm{exp}}}$ denotes the mean value of the VACF over the n fitted time steps. Therefore, R^2 gives an estimate for the goodness of the fit, where a value of zero corresponds to a bad fit that is equivalent to assuming the function to be the mean everywhere and a value of one corresponds to perfect agreement between the fit and the data.

Naturally, for noisy data obtained from experiments, there is some spread in the R^2 values, as shown in Fig. S2. Nevertheless, it can be seen that most of the R^2 values of our extraction are close to one, thus indicating that our different fit models describe the respective data accurately on the individual trajectory level. To further undermine this point, we show the fit results with the highest R^2 value for the respective individuals in Fig. S3. Here, the data agrees perfectly with the fit results. The R^2 values of the respective fits for the randomly chosen individuals in Figs. 4a,e,h are 1.000, 0.964, 0.953.

IV. EFFECTIVE KERNEL FOLLOWS UNIQUELY FROM CORRELATION FUNCTION

Fourier transformation of Eq. (1) and Eq. (2) leads to

$$\tilde{v}(\omega) = \frac{\tilde{F}_R(\omega)}{\tilde{\Gamma}_v^+(\omega) + i\omega} \tag{S6}$$

with the single-sided Fourier transform defined as $\tilde{\Gamma}_v^+(\omega) = \int_0^\infty e^{-i\omega t}\Gamma_v(t)dt$ and

$$\langle \tilde{F}_R(\omega)\tilde{F}_R(\omega')\rangle = 2\pi\delta(\omega+\omega')\tilde{\Gamma}_R(\omega')\,. \tag{S7}$$

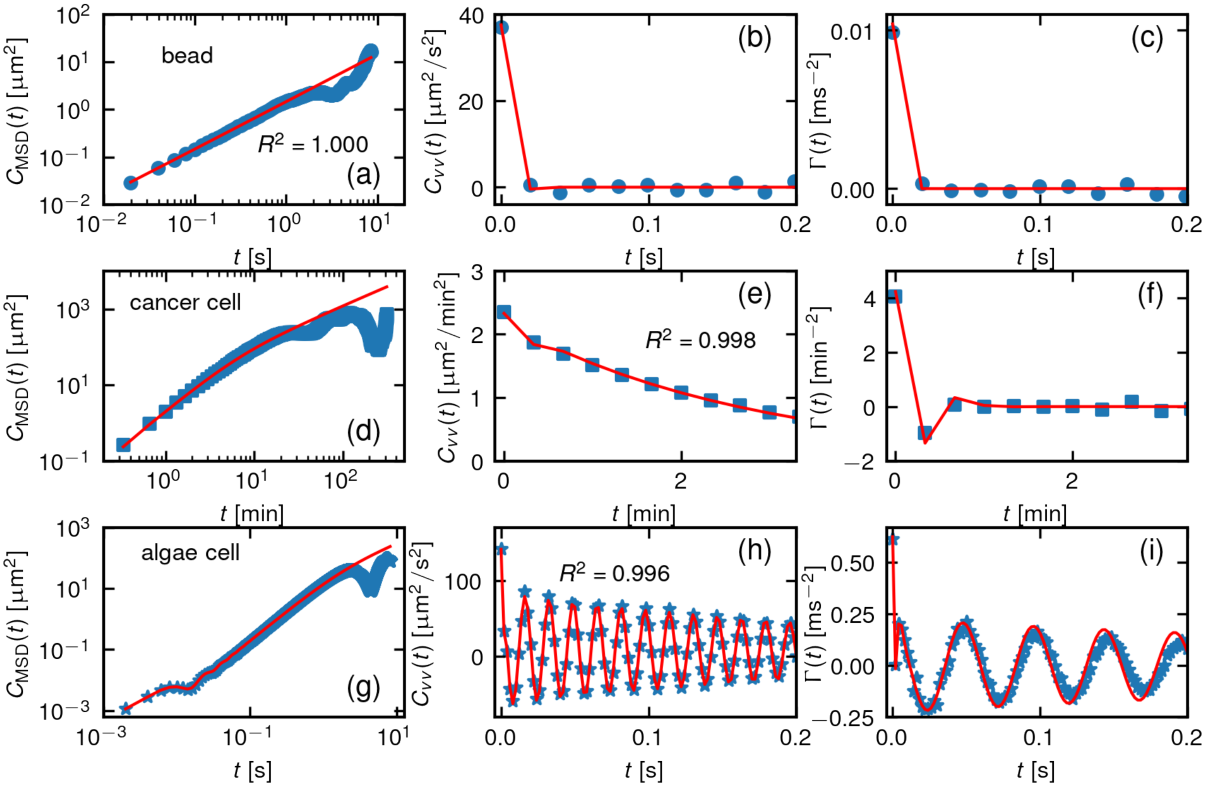

FIG. S3. Experimental results for the MSD, $C_{\mathrm{MSD}}(t)$, VACF, $C_{vv}(t)$, and friction kernel $\Gamma(t)$, of a single bead (a)-(c), a single cancer cell (d)-(f) and a single algae cell (g)-(i) (blue symbols), chosen as the objects with the highest R^2 value, which is given in the respective plot. The red lines denote the fit result.

From the Fourier transform of the VACF

$$\begin{aligned}\tilde{C}_{vv}(\omega) &= \int_{-\infty}^{\infty} dt e^{-i\omega t}\langle v(0)v(t)\rangle \\ &= \int_{-\infty}^{\infty} e^{-i\omega t}dt \int_{\infty}^{\infty} e^{i\omega t}\frac{d\omega}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\langle\tilde{v}(\omega)\tilde{v}(\omega')\rangle\end{aligned} \tag{S8}$$

we obtain by inserting Eqs. (S6) and (S7)

$$\tilde{C}_{vv}(\omega) = \frac{\tilde{\Gamma}_R(\omega)}{(\tilde{\Gamma}_v^+(\omega)+i\omega)(\tilde{\Gamma}_v^+(-\omega)-i\omega)}\,. \tag{S9}$$

Equating the non-equilibrium and the surrogate VACF with $\Gamma_R(t)/B = \Gamma_v(\mid t\mid) = \Gamma(\mid t\mid)$ leads to

$$\frac{\tilde{\Gamma}_R(\omega)}{\mid\tilde{\Gamma}_v^+(\omega)+i\omega\mid^2} = \frac{B\tilde{\Gamma}(\omega)}{\mid\tilde{\Gamma}^+(\omega)+i\omega\mid^2}\,. \tag{S10}$$

We now demonstrate that Eq. (S10), which follows from the GLE Eq. (1), can be solved for $\tilde{\Gamma}(\omega)$ and $\tilde{\Gamma}^{+}(\omega)$.

Multiplying Eq. (1) by $v(t_0)$, inserting $\Gamma(|t|) = \Gamma_R(t)/B = \Gamma_v(|t|)$ and averaging the entire equation leads to the Volterra equation

$$\frac{d}{dt}C_{vv}(t) = -\int_0^t \Gamma(t-s)C_{vv}(s)ds\,, \tag{S11}$$

where we used that the random force is not correlated with the initial velocity at the projection time, i.e. $\langle v(t_0)F_R(t)\rangle = 0$, and set $t_0 = 0$. For a given $\Gamma(t)$ one can find a solution of Eq. (S11) in terms of $C_{vv}(t)$. Inversely, for given correlation function $C_{vv}(t)$ one can solve Eq. (S11) in terms of the friction kernel $\Gamma(t)$ by Laplace transformation. With the definition of the Laplace transform of a function $f(t)$

$$\hat{f}(q) = \int_0^\infty f(t)e^{-qt}dt, \tag{S12}$$

the unique solution for $\Gamma(t)$ in Laplace space is given by

$$\hat{\Gamma}(q) = \left(\frac{C_{vv}(0)}{\hat{C}_{vv}(q)} - q\right). \tag{S13}$$

The existence of a unique friction kernel for any given input correlation function $C_{vv}(t)$ assures that one can always find an effective friction kernel $\Gamma(t)$ that, when employed in the GLE and using $\Gamma(|t|) = \Gamma_R(t)/B = \Gamma_v(|t|)$, reproduces the two-point correlation functions. Thus, every non-equilibrium model described by Eq. (1) with $\Gamma_R(t)/B \neq \Gamma_v(|t|)$ can be mapped on an effective model with $\Gamma(|t|) = \Gamma_R(t)/B = \Gamma_v(|t|)$ determined by Eq. (S13), because the Green's function, which completely describes a Gaussian process, is solely given in terms of the positional two-point correlation function [12, 40].

V. BEAD RADIUS VARIANCE DETERMINED BY ATOMIC FORCE MICROSCOPY

In order to estimate the bead radius variance independently of the MSD analysis, we determine the radius variance from 220 beads via atomic force microscopy (AFM). The radius distribution is shown in Fig. S4, where the mean radius is $\langle r\rangle = 0.51\,\mu$m and the standard deviation is $\Delta r = 0.02\,\mu$m. This is in line with the values of $\langle r\rangle = 0.50\,\mu$m and standard deviation $\Delta r = 25\,$nm provided by the manufacturer. These values also agree with the estimated radius variance from the motility data in the main text of $\Delta r = 86 \pm 63\,$nm.

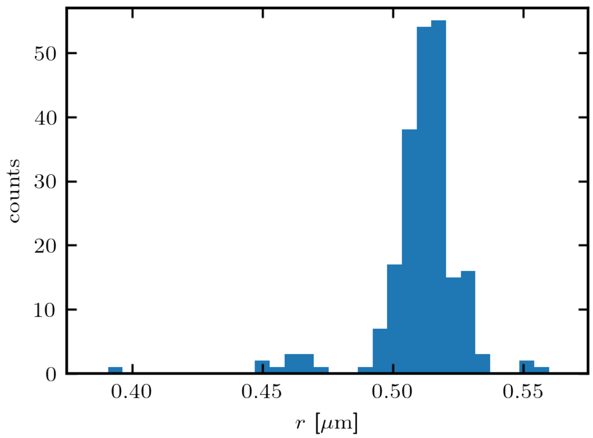

FIG. S4. Distribution of measured bead radii via AFM, where the exact sample preparation and AFM settings are explained in the text.

The sample of beads is prepared on freshly cleaved mica. Before the AFM measurement the beads (FluoSpheres NeutrAvidin-Labeled Microspheres, 1.0 μm ThermoFisher Scientific) are diluted 1:100 in deionized water and sonicated for 10 min to reduce bead aggregation. The sample is then left to dry overnight. AFM images are obtained using a MultiMode 8 atomic force microscope (Veeco, Santa Barbara, CA) with a Nanoscope V controller and NanoScope software. Imaging is performed with TappingMode in air using AC160TS cantilevers (Asylum Research by Oxford Instruments, LOT#753014). The AFM images with a frame size of 10 μm × 10 μm are recorded at a scan rate of 1 Hz over 256 lines, with an integral gain of 5 and proportional gain of 10, without z-limit.

VI. LEAST-SQUARE FIT PARAMETER DISTRIBUTIONS OF NOISY EXPONENTIALS: THE MEDIAN IS MORE ACCURATE THAN THE MEAN

In Fig. S5 we show the fit results of 10^5 least-square fits to the noisy function

$$f(t) = \alpha_0 e^{-\beta_0 t} + \gamma_0 + \zeta(t) \tag{S14}$$

with the fit function

$$f^{\text{fit}}(t) = \alpha e^{-\beta t} + \gamma\,, \tag{S15}$$

where $\zeta(t)$ describes Gaussian noise with zero mean $\langle \zeta(t) \rangle = 0$ and variance $\langle \zeta(0)\zeta(t) = 5\delta(t)$. The form of Eq. (S14) is motivated by the fact that VACFs often contain exponen-

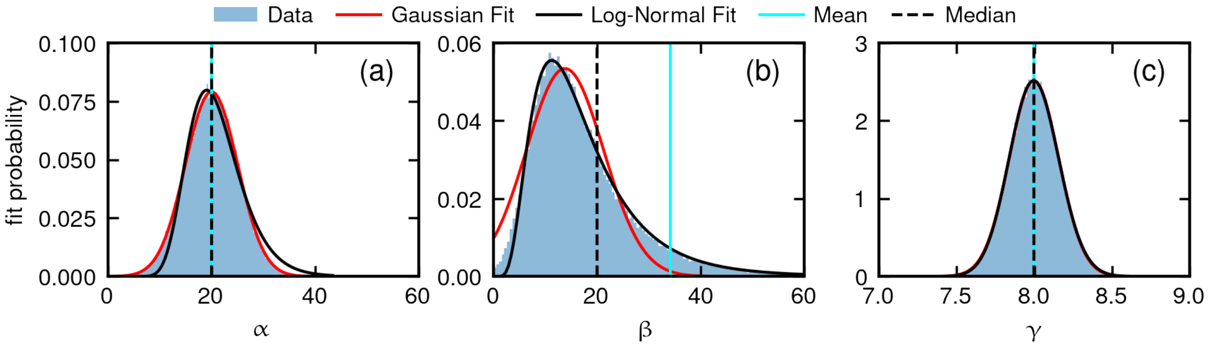

FIG. S5. Distributions of fitted parameters (a) α, (b) β, (c) γ of 10^5 fits of Eq. (S15) to data sets of Eq. (S14) with input parameters $\alpha_0 = 20$, $\beta_0 = 20$, $\gamma_0 = 8$, $\langle\zeta(0)\zeta(t) = 5\delta(t)$ for t between $0 - 100$ in steps of 0.1. The mean and median of the data are presented as vertical lines and are given by $\langle\alpha\rangle = 19.98$, $\alpha^{\text{med}} = 20.00$, $\langle\beta\rangle = 34.18$, $\beta^{\text{med}} = 19.98$, $\langle\gamma\rangle = 8.00$, $\gamma^{\text{med}} = 8.00$.

tially decaying components. Since our model invokes fitting to VACFs with exponential components with a least-square fit according to Eq. (16) in the main text, here, we discuss the resulting parameter distributions of such fits. The distribution of fitted values for the exponential prefactor α and the constant γ are well described by Gaussian distributions, see Figs. S5a,c, whereas the distribution of the exponential factor β is better described by a log-normal distribution, as shown in Fig. S5b, which has the form

$$p(\beta) = \frac{\exp(-\frac{(\ln(\beta)-\ln(\beta^{\text{med}}))^2}{2\sigma^2})}{\beta\sigma\sqrt{2\pi}}\,, \tag{S16}$$

where β^{med} denotes the median of the distribution and σ determines the width of the distribution.

It turns out that the actual input parameter of $\beta_0 = 20$ is better described by the median of the distribution of fitted parameters compared to the mean, which we show in Fig. S5b. For the parameters α and γ the median and mean describe the actual input parameters equally well, as seen in Fig. S5a,c. As the median accurately describes the true value of Gaussian distributed parameters as well as Log-normal distributed parameters from least-squares fits, we choose it as a representation of the center of distributions rather than the mean.

VII. ANALYTICAL ESTIMATION OF PARAMETER VARIANCE

We compare our parameter variance estimate from simulations with analytical estimates. As an example, we estimate the mean squared velocity $B = \langle v^2 \rangle = C_{vv}(0)$ using the persistent random walk (PRW) model Eq. (7). Due to our extraction of the model parameters from the VACF Eq. (15), the statistical error of the parameters relates closely to the statistical error of the VACF. We consider a trajectory composed of $n + 1 = L/\Delta$ positions which produces a VACF with n values, where the mean squared velocity of the continuous trajectory is estimated by the finite difference velocities Eq. (11) as

$$B_{\text{est}} = \frac{1}{n} \sum_{i=1}^{n} v_i^2 \,. \tag{S17}$$

Here, it should be noted that this estimate deviates from the estimation of our fit model including localization noise effects Eq. (15). Especially in cases with localization noise, the estimate B_{est} in Eq. (S17) overestimates the true value of B, as displacements are on average increased by the noise. On the other hand, a finite time step discretization leads to decreased values B_{est}, as a finite time step does not capture instantaneous velocities but rather corresponds to an average over the interval of the time step. Our extraction of parameters by a least squares fit of Eq. (15) accounts for discretization and localization noise effects, but is not analytically tractable.

The moments of v follow from the velocity distribution $\rho(v)$ as

$$\langle v^k \rangle = \int_{-\infty}^{\infty} v^k \rho(v) dv \,. \tag{S18}$$

We want to determine the dependence of the variance of our estimation B_{est} on the length of the trajectories

$$\sigma^2(B_{\text{est}}) = \langle B_{\text{est}}^2 \rangle - \langle B_{\text{est}} \rangle^2 \,. \tag{S19}$$

Inserting Eq. S17 into Eq. S19 leads to

$$\begin{gathered} \sigma^2(B_{\text{est}}) = \\ \tfrac{1}{n^2} \langle \textstyle\sum_{i=1}^{n} v_i^2 \sum_{j}^{n} v_j^2 \rangle - (\tfrac{1}{n} \textstyle\sum_{i=1}^{n} \langle v_i^2 \rangle)^2 \\ = \tfrac{1}{n^2} (n \langle v^4 \rangle + \textstyle\sum_{i \neq j}^{n} \langle v_i^2 v_j^2 \rangle - n^2 \langle v^2 \rangle^2) \,. \end{gathered} \tag{S20}$$

If the velocities at times $i \neq j$ are independent of each other, the correlation of the squared velocities can be written as

$$\langle v_i^2 v_j^2 \rangle = \langle v_i^2 \rangle \langle v_j^2 \rangle \; = \langle v^2 \rangle^2, \tag{S21}$$

thus, the variance of B_{est} follows as

$$\sigma^2(B_{\mathrm{est}}) = \frac{1}{n^2}\left(n\langle v^4\rangle + n(n-1)\langle v^2\rangle^2 - n^2\langle v^2\rangle^2\ \right) = \frac{1}{n}\left(\langle v^4\rangle\ - \langle v^2\rangle^2\right)\,. \tag{S22}$$

In fact, the assumption of uncorrelated squared velocities is not true for most experimental trajectories, indicated by the non-zero values of the VACF for $t > 0$, meaning one needs to consider the correlation between the squared velocities over time. This results in a different estimate for the variance of B_{est} [42], which can be written as

$$\begin{aligned}\sigma^2(B_{\mathrm{est}}) = {} & \frac{1}{n}(\langle v^4\rangle - \langle v^2\rangle^2 \\ & + 2\sum_{k=1}^{n-1}(1-k/n)\langle v_0^2 v_k^2\rangle)\,,\end{aligned} \tag{S23}$$

where we used that the correlation function only depends on time differences $k = |i-j|$ and that the kth term $\langle v_0^2 v_k^2\rangle$ occurs $n-k$ times in the sum. If one does not know the analytic form of $C_{v^2v^2}(t) = \langle v^2(0)v^2(t)\rangle$, one can obtain the function from the data similarly to Eq. 12 for $C_{vv}(t)$, where one usually introduces a cutoff length n_{cut} to estimate the variance as

$$\sigma^2(B_{\mathrm{est}}) = \frac{1}{n}\left(C^0_{v^2v^2} - (C^0_{vv})^2 + 2\sum_{k=1}^{n_{\mathrm{cut}}} C^k_{v^2v^2}\right) \tag{S24}$$

in order to avoid using the estimated correlation function for large times $t > n_{\mathrm{cut}}\Delta$, where the noise is stronger due to less averaging, see Eq. 12. Similarly to the VACF Eq. (19) one can calculate the correlation of the squared velocities for the PRW (which is described by the GLE Eq. (1) with the instantaneous friction kernel Eq. (7)) starting from the formal solution for the velocity

$$v(t) = v(0)e^{-t/\tau_m} + \int_0^t F_R(s)e^{-(t-s)/\tau_m}ds\,. \tag{S25}$$

The squared velocity correlation function then takes the form

$$\langle v^2(0)v^2(t)\rangle = \left\langle v^2(0)\left[v^2(0)e^{-2t/\tau_m} + 2v(0)e^{-t/\tau_m}\int_0^t F_R(s)e^{-(t-s)/\tau_m}\,ds + \left(\int_0^t F_R(s)e^{-(t-s)/\tau_m}\,ds\right)^2\right]\right\rangle\,. \tag{S26}$$

Inserting the first and second moment of the random force $\langle F_R(s)\rangle = 0$ and $\langle F_R(s)F_R(s')\rangle = 2B\delta(s-s')/\tau_m$ respectively and evaluating the integrals, one finds

$$\langle v^2(0)v^2(t)\rangle = 3\langle v^2(0)\rangle^2 e^{-2t/\tau_m} + B\langle v^2(0)\rangle\left(1-e^{-2t/\tau_m}\right)\,, \tag{S27}$$

where we used that the fourth moment of a Gaussian is $\langle v(0)^4\rangle = 3\langle v^2(0)\rangle^2$. Now inserting $\langle v^2(0)\rangle = B$ yields

$$C_{v^2v^2}(t) = B^2(1 + 2e^{-2t/\tau_m})\,. \tag{S28}$$

In Fig. S6a we show extracted PRW parameter distributions from simulations with identical input parameters for different trajectory lengths L, where one can clearly see the variance decreasing for increasing trajectory length. In Figs. S6b-d this decrease is presented for different localization noise widths σ_{loc}. The analytic prediction for uncorrelated velocities Eq. (S22), shown as black dashed line in Figs. S6b-d, is clearly below the observed variance, since the velocities in the PRW are intrinsically correlated, but correctly describes the powerlaw dependence on L. For increasing localization noise, the relative variance $\sigma^2(B)/B^2$ decreases for B_{est} (crosses in Figs. S6b-d), because the mean increases by

$$\langle B_{\text{est}}\rangle = \frac{2C^{\text{theo}}_{\text{MSD}}(\Delta) + 4\sigma_{\text{loc}}}{2\Delta^2} \approx B + \frac{2\sigma_{\text{loc}}}{\Delta^2}\,, \tag{S29}$$

which follows directly from Eq. (15) of the main text. The approximation in the latter part of Eq. (S29) is accurate when the time step is smaller than the length of the ballistic regime in the MSD; in the case of the PRW, this means it is exact for $\Delta < \tau_m$. The localization noise increases the estimation of the mean squared velocity, as seen in Eq. (S29), and it leads to a decorrelation of consecutive velocities, as seen for instance in Figs. 4e,f in the main text by the dip in the VACF at $t = \Delta$. Therefore, the estimator B_{est} Eq. (S17) comes closer to the prediction for uncorrelated velocities for increasing localization noise, as seen by comparing the crosses in Fig. S6a to the crosses in Fig. S6b and c.

The mean of B_{est} according to Eq. (S17) is shown in the upper right corner of Figs. S6b-d, which increases with increasing localization noise, as expected from Eq. (S29). Here, we also show that B determined by our fit of Eq. (16), denoted by B_{fit}, gives accurate estimates close to the input value of B even for high localization noise width σ_{loc}. Moreover, the variance B_{fit}, shown as circles in Figs. S6b-d, is close to the analytical estimate for $\sigma(B_{\text{est}})$ according to Eq. (S24), shown as the cyan line in Figs. S6b-d. However, this close match is not guarantied for arbitrary models, since the derivation cannot be done analytically for nonlinear fit models, such as our fit by Eq. (16). Additionally, estimating the parameter variance of a trajectory ensemble of different trajectory lengths L further complicates the matter. All these complications are avoided by estimating the parameter variance by simulating the underlying process based on the GLE Eq. (1), where all experimental settings, such as time

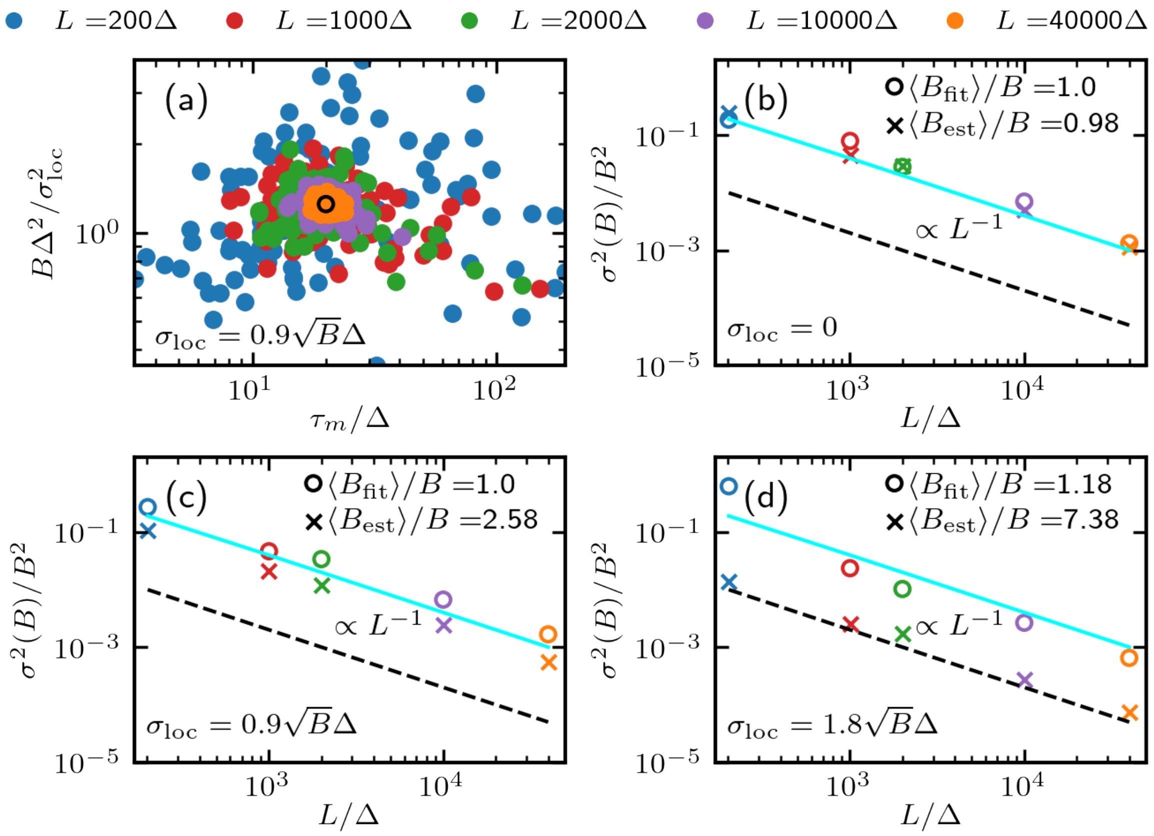

FIG. S6. (a) Distributions of PRW parameters B, τ_m extracted from simulated trajectories of different lengths in different colors, where each distribution contains $N = 100$ points and the input is shown as a black empty circle, $B\Delta^2/\sigma_{\rm loc} = 1.25$, $\tau_m/\Delta = 20$. (b)-(d) The variance of the estimator for B in Eq. (S17) is shown as crosses, the estimation of B by our fit Eq. (16) in the main text as circles for different trajectory lengths estimated from $N = 100$ simulated trajectories with the input as in (a), but with (b) $\sigma_{\rm loc} = 0$, (c) $\sigma_{\rm loc} = 0.9\sqrt{B}\Delta$ and (d) $\sigma_{\rm loc} = 1.8\sqrt{B}\Delta$. The respective mean values of the two different estimates of B for trajectory length $L/\Delta = 4 \times 10^4$ are given in the upper right corner. The black dashed line represents the expected variance for uncorrelated data given in Eq. (S22) and the cyan line the expected variance given by Eq. (S24) using the analytical result for the correlation function given in Eq. (S28) with $n_{\rm cut} = n = L/\Delta - 1$. As dictated by the law of large numbers, the variance is inversely proportional to the number of data points, i.e. inversely proportional to the trajectory length L.

step, different trajectory lengths and localization noise, can easily be included, as explained in the main text.

VIII. TEST METHOD WITH KNOWN INTRINSIC PARAMETER DISTRIBUTION

In order to show, that the statistical test explained in the Methods actually results in correct estimates for the cell-to-cell variance, we test our method for synthetic data from a known intrinsic parameter distribution. For the PRW model Eq. (7) the parameter vector has three entries $\vec{z} = (\tau_m,\, B,\, \sigma_{\mathrm{loc}})$, so we start by drawing N parameter vectors from a Gaussian distribution with covariance matrix

$$\mathrm{Cov}^{\mathrm{inp}*} = \begin{pmatrix} 22.5 & 68.2 & 0.17 \\ 68.2 & 440.5 & -0.017 \\ 0.17 & -0.017 & -0.012 \end{pmatrix} \tag{S30}$$

and mean

$$\langle \vec{z} \rangle = (10\,\mathrm{min}, 1.25\,\mu\mathrm{m}^2/\mathrm{min}^2, 0.1\,\mu\mathrm{m})\,, \tag{S31}$$

where the mean is in the order of the parameters exhibited by the cancer cells as seen in Fig. 5e. For clarity we omit the units of the covariance in Eq. (S30). A Gaussian distribution with covariance given by Eq. (S30) and mean given by Eq. (S31) is shown as violet empty symbols in Fig. S7a. We show the axes in Fig. S7 in relative units because this way the results are better comparable across different systems with different units and the simulation results only vary if the ratios of parameters change. Now, we simulate a trajectory for each of the N parameter vectors $\vec{z}$ of this known distribution of length $L = 100\,\Delta$ with $\Delta = 0.5\,\mathrm{min}$, which is a typical length for single-cell tracking data.

In order to obtain the correct continuous time behavior, we simulate trajectories with a much smaller time step than the discretization of the trajectories with $h = 0.05\Delta$ and then obtain trajectories by using only every $\Delta/h = 20$th step of the simulation and adding localization noise drawn from the Gaussian distribution of width σ_{loc} (for simulations of the algae cells we use $h = 0.005\Delta$). The extracted parameters from these simulated trajectories are denoted by the blue symbols shown in Fig. S7a. Since we know the exact distribution that led to this extracted parameter distribution, we can now test whether our method

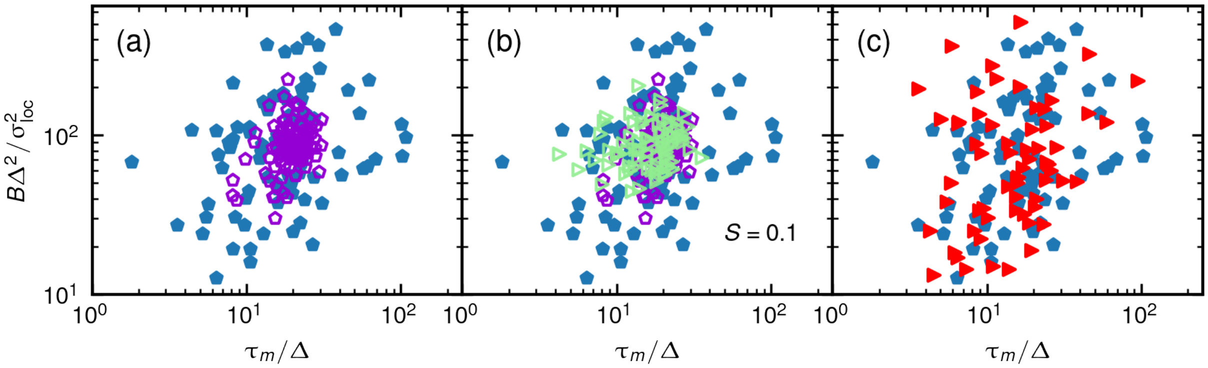

FIG. S7. (a) $N = 100$ randomly drawn PRW parameter vectors from a Gaussian distribution with covariance and mean given in Eqs. (S30), (S31) respectively, are shown as empty violet pentagons. Filled blue pentagons denote the parameters extracted from simulations using the empty violet pentagon input parameters for trajectories of length $L = 100\,\Delta$. (b) A Gaussian distribution of empty green triangles with $S = 0.1$ compares well to the original violet pentagon distribution that was used to create trajectories and extract the synthetic distribution of blue pentagons. (c) The red triangles represent the parameters extracted from the simulation using the green triangles in (b) as input parameters.

reproduces the spread of the original distribution. The ratio S defined in Eq. (9) of the input covariance $\mathrm{Cov}^{\mathrm{inp}*}$ to the covariance of the distribution extracted from the synthetic data, here denoted as $\mathrm{Cov}^{\mathrm{exp}}$, is $S = 0.11$. Indeed we find that the estimated covariance ratio using the statistical test introduced in the Methods is $S^* = 0.17 \pm 0.08$. An exemplary distribution with $S = 0.1$ is shown in Fig. S7b as empty green triangles and agrees well with the true original distribution shown as empty violet pentagons. Hence, the distribution extracted from simulations using the empty green triangles in Fig. S7b shown in Fig. S7c (red triangles) agrees with the distribution extracted from the synthetic data (blue symbols).

We note that if the type of the underlying distribution is known, one can easily adapt the method to use an input distribution different from a Gaussian distribution. Nevertheless, in the most common case where the distribution type is not known, the Gaussian approach is most likely to describe the distribution.

Additionally, we create synthetic distributions of $N = 100$ PRW trajectories originating all from the same parameter vector $B\Delta^2/\sigma^2_{\mathrm{loc}} = 80$, $\tau_m/\Delta = 25$ close to the values of the breast cancer cell data. All parameters being the same means $\mathrm{Cov}^{\mathrm{inp}*} = 0$ and hence

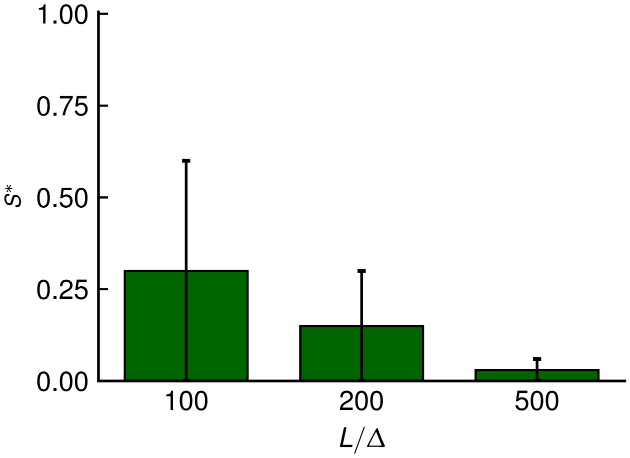

FIG. S8. (a) Estimated variance ratio S^* shown in green for synthetic distributions with all identical input parameters $B\Delta^2/\sigma_{\text{loc}}^2 = 80$, $\tau_m/\Delta = 25$ ($S = 0$) and different trajectory lengths L for $N = 100$ synthetic trajectories.

the expected result is $S^* = 0$. We create synthetic distributions by simulations of length $L = 500\,\Delta$, $L = 200\,\Delta$ and $L = 100\,\Delta$ respectively and show the results for the estimated variance ratio S^* in Fig. S8. All of the results agree with the expected value of $S^* = 0$, but the uncertainty is higher for shorter trajectories, since the spread and the fluctuations of the fitted correlation functions are higher for shorter trajectories.