Electronic Supplementary Material (ESI) for Soft Matter.

Supplementary information

for

Complex nearly immotile behaviour of enzymatically driven cargos

O. Osunbayo†, C.E. Miles†, F. Doval, B.J.N. Reddy, J.P. Keener, M.D. Vershinin

Supplement Text 1

We seek a theoretical model which exhibits the non-Gaussian diffusivity distribution phenotype, yet requires only the most minimal features of found in both our *in vivo* and *in vitro* model systems: active enzymatically-driven displacement and passive processes such as diffusion. Certainly, many "passive" processes involve transitions between distinct energy states (e.g. attachment/detachment) and so their temperature behavior may be non-linear, often Arrhenius. Indeed, even bead diffusion in water (Fig. S1) is non-linear with temperature due to the thermal variation of water viscosity. However, on biological temperature scales, the linear approximation is often appropriate. By comparison, active biological processes show strong temperature dependence over a very small temperature range. The activity of many enzymes can double over just 10° C. The distinction between active and passive processes is further motivated by biology, in the sense that enzymatic processes usually have different modes of regulation in cells. It is therefore often desirable to learn from simple experimental observations whether changes are driven by enzymatic activity. For example, in the case of cytoskeletal transport it is often desirable to know whether particle immotility is due to its lack of attachment to the cytoskeletal filaments or due to collective action of engaged and enzymatically active molecular motors.

The minimal model we seek ideally incorporates just one passive and one active process with the understanding that full system behavior is governed by many such processes. For example, powerstroke displacement is enzymatically driven but so is the transition between high and low motor affinity states for the filament. We aim here to avoid considering such details in favor of model simplicity. The active process in our model is represented as simple jumps in either direction which are assumed to have fixed characteristic distance δ and occur at the same rate λ, so that the system is symmetric. These considerations lead to a bi-directional jump diffusion process, described in 1 dimension by the probability density of the cargo's temporal displacement $p(x,t)$, which is governed by the Fokker-Planck equation[SR1]

$$\partial_t p(x,t) = D\partial_{xx} p(x,t) - 2\lambda p(x,t) + \lambda p(x-\delta,t) + \lambda p(x+\delta,t). \quad (1)$$

When the rate of jumps λ is set to zero, this reduces to the classical passive diffusion process, hence D includes all passively driven fluctuations. This system is highly reminiscent of the more general continuous-time random walk (CTRW) processes (18) but our model lacks aging and long tailed correlations (Fig. S5). The advantage of the above simplifications is that all the moments μ_j of the distribution $p(x,t)$ can be computed analytically.

We can relate the theoretical quantities to experimental observables (cargo position along a filament) by considering the observations of a single trajectory denoted $\{x_0, x_1, x_2, \ldots, x_N\}$, where each observation of

cargo location x_i occurs at a fixed time interval Δt after x_{i-1} (typically, this is the frame timing in the video record of cargo motion). Arguably the simplest estimate of diffusivity in 1D is:

$$D_s^{(i)} := \frac{[x_{i+1} - x_i]^2}{2\Delta t} \tag{2}$$

That is, for the i[th] time window of length Δt, we associate a sample of the diffusivity. Although the actual procedure used in our data analysis is considerably more complex, we can use these observations as an estimator for *D* along an observed trajectory:

$$D_{est} = \frac{1}{N}\sum_{i=1}^{N} D_s^{(i)} \tag{3}$$

Here D_{est} is a random variable representing the empirical diffusivity distribution. We are thus simplifying this diffusivity estimate to a mean squared displacement with lag time of 1 frame. This choice neglects information from the MSD curve for higher time lags. The actual practice of estimating D for a path is considerably more complicated – it is a weighted linear regression(17) which also tends to "distrust" information for higher time lags (by weighing them lower) albeit not as drastically as our choice. Our estimator is more tractable for closed form calculations and provides intuition for the curious statistics observed in experiments. In the actual procedure used for experimental data analysis, samples are pair point averages and are thus *not* truly independent. Hence, the above are lower bounds for the variance and skewness computed in this supplement. However, the estimator considered here produces close agreement qualitatively and quantitatively with the more involved overlapping window procedure (Fig. S2,S3) while providing intuition due to its analytical tractability.

Let $p(x,t)$ be the probability density of a jump-diffusion process in one dimension governed by the Fokker-Planck equation (1). Ultimately, we are interested in the distribution of diffusion coefficients. That is, we associate an empirical diffusivity with each trajectory, reflecting how the squared displacement scales with time. Of course, the time in the model is a continuous rather than discrete variable, but we can use time increments of Δt to sample x(t) which then produces a random variable $D_*(\Delta t)$ analogous to the experimentally defined $D_s^{(i)}$. We emphasize the important observation that, due to ergodicity of the system, $D_\star(\Delta t)$ has the same statistics as $D_s^{(i)}$. Consequently, computing the statistics of the estimator D_{est} reduces to computing those of $D_\star(t)$.

Moments of $D_\star$ can be related to moments of x by $\langle D_\star{}^j \rangle = \mu_{2j} / (2t)^j$ where $\mu_j := \langle x^j \rangle$.

Multiplying (1) by x^j and integrating, we find the explicit evolution of these moments

$$\frac{d\mu_j}{dt} = Dj(j-1)\mu_{j-2} + 2\lambda \sum_{k=1}^{\lfloor j/2 \rfloor} \binom{j}{2k} \delta^{2k} \mu_{j-2k}, \tag{4}$$

where $\lfloor \cdot \rfloor$ is the floor operator.

This system of ordinary differential equations is "lower triangular" in the sense that μ_j only depends on μ_k for $k < j$. That is, we can solve these moments equations sequentially. For the first moment, we have $d\mu_1/dt = 0$, so that $\mu_1 \equiv 0$. This is intuitive, as diffusion and symmetric jumps produce no displacement on average.

The remaining equations of interest are

$$\begin{aligned} \frac{d\mu_2}{dt} &= 2D + 2\lambda\delta^2, \\ \frac{d\mu_4}{dt} &= 12\left[D + \lambda\delta^2\right]\mu_2 + 2\lambda\delta^4, \\ \frac{d\mu_6}{dt} &= 30\left[D + \lambda\delta^2\right]\mu_4 + 30\lambda\delta^4\mu_2 + 2\lambda\delta^6. \end{aligned} \tag{5}$$

The most notable feature of these statistics that we utilize is that the MSD for this process is linear and the slope can be estimated from experimental data which allows us to compute the mean of the empirical diffusivity distribution over several paths:

$$E\left[D_{est}\right] = E[D_\star(\Delta t)] = \frac{\mu_2}{2(\Delta t)} = D + \lambda\delta^2. \tag{6}$$

As expected, if λ=0, the mean of the empirical distribution indeed recovers the true underlying passive diffusion coefficient. We can also see that for enzymatically driven jumps, whose rate λ scales with temperature per Arrhenius law, the effective diffusion coefficient will have identical scaling with the same activation energy. This suggests that our intuition is correct: if passive diffusion is dominant, the empirical diffusivity should scale linearly, otherwise it will scale as a Boltzmann factor.

We can proceed on to the variance, using

$$\tilde{\sigma}^2 := var[D_\star(\Delta t)] = \frac{1}{(2(\Delta t))^2}\left[\mu_4 - \mu_2^2\right] = 2\delta^4\lambda^2 + 2D^2 + 4\delta^2 D\lambda + \frac{\delta^4\lambda}{2(\Delta t)}. \quad (7)$$

Notably, if $\lambda = 0$ (i.e. no jumps, classical diffusion), there is no dependence on Δt (the sampling time). However, with $\lambda \neq 0$, $\tilde{\sigma}^2 \to \infty$ as $\Delta t \to 0$.

The next moment, the skewness, is found to be

$$\tilde{g} := skew[D_\star(\Delta t)] = \frac{32D^3(\Delta t)^2 + 96\delta^2 D^2\lambda(\Delta t)^2 + 24\delta^4 D\lambda(\Delta t)(4\lambda(\Delta t)+1) + \delta^6\lambda(8\lambda(\Delta t)(4\lambda(\Delta t)+3)}{(\Delta t)\left(4D^2(\Delta t) + 8\delta^2 D\lambda(\Delta t) + \delta^4\lambda(4\lambda\Delta t + 1)\right)\sqrt{8D^2 + 16\delta^2 D\lambda + \frac{2\delta^4\lambda(4\lambda(\Delta t)+}{(\Delta t)}}} \quad (8)$$

This result leads to several important observations. First, in the absence of jumps (λ=0), the skewness of each sample is 2√2, as expected for the chi-squared distribution with 1 degree of freedom. Second, the skewness is a non-monotonic function of D and λ. Therefore, skewness of diffusivities (unlike mean diffusivity) is not expected to follow Arrhenius behavior as lambda changes with temperature. It may in fact grow or decline or go through a peak depending on the specific parameter values. Finally, we see that when the system approaches immotility (D and λ are both small while other variables are finite), the skewness reduces to approximately:

$$\tilde{g} := skew[D_\star(\Delta t)] \sim \frac{1}{\sqrt{2\lambda(\Delta t)}} \quad (8a)$$

An identical limit arises if Δt is taken to be small while other variables are held finite. Indeed, if $\lambda << 1/\Delta t$ then the situation is similar in spirit to classical CTRW models: the tails become relatively long and lead to strong deviations from Brownian behavior (Fig. S2A). What is particularly curious is that in this limit, step size is unimportant so that it is possible to get highly skewed distributions of diffusivities from an otherwise experimentally obscure active process. In practical terms, this implies that skewed distribution of diffusivities may serve as the first clue that the apparently diffusive process has complex constituents. The temperature dependent studies may then be used to try and separate active from passive contributions to the overall random process.

We note for completeness that the skewness does indeed decay with the number of samples as predicted by the Central Limit Theorem. Using the definition of D_{est} and that the samples $D_s^{(i)}$ are independent, the statistics of the collection of diffusion coefficients over many paths becomes

$$var[D_{est}] = \frac{\tilde{\sigma}^2}{N},$$
$$skew[D_{est}] = \tilde{g}/\sqrt{N}. \quad (9)$$

Here N is the number of samples per path. Note that the rate of convergence is dependent on the magnitude of the skewness. Hence, as skewness is enhanced, convergence (to a Gaussian) becomes quite slow (Fig. S2), even without violating the central limit theorem or ergodicity. In all cases, skewness estimation from data is robust when the number of trajectories or paths (M) is large (Fig. 2B).

This result is quite intuitive as a sampling issue. As experimental setups (and consequently, this random variable) are "snapshots" of the process and we have no memory effects built into our model, each sample is an independent identically distributed random variable representing diffusion and jumps occurring in a fixed amount of time (sampling time). The relationship between the sampling rate and the rate at which these jumps occur is therefore crucial in determining the statistics. In such cases, limited experimental data sets can still produce strongly skewed distributions for slopes of mean squared displacement curves without violating ergodicity. This observation has been made in other branches of scientific literature (for instance, finance[SR2]) but seems to be underappreciated in this context.

Supplement Text 2

The model in Supplement Text 1 represents mean-field behavior where individual degrees of freedom are not explicitly accounted for. This approximation is appropriate for non-processive motors in our experiments because our motor ensembles are relatively large: small motor ensembles would not stay associated with microtubules for extended times. However, the microscopic picture we have is bead-microtubule coupling via variable number of non-processive motors at a variety of relative binding positions and hence a variety of coupling strengths. In this picture, when a motor detaches or attaches to the microtubule, the system evolves to a new overall state. We would then expect that after some lag time, memory effects would become negligible and mean field behavior describes the ensemble dynamics, which is supported by experimental observations (Fig. S4): Beyond about 1 second lag time, MSD curves do show a linear trend. Therefore, on long lag time scales the system becomes a good model of apparently diffusive behavior. Time–averaged MSD (TA-MSD) analysis(5, 6, 41) of motility data (Fig. S5) showed robust convergence, suggesting that the sub-diffusive process is ergodic.

Our analysis also revealed a broad distribution of anomalous exponents from near zero to slightly above unity (Fig. S4B). Generally, the distribution shifted lower with declining temperature. This is naturally explained by the above microscopic picture: as temperature gets lower, individual motor dynamics and thus also ensemble dynamics slows down leading to a flattening and leveling off of the MSD curve.

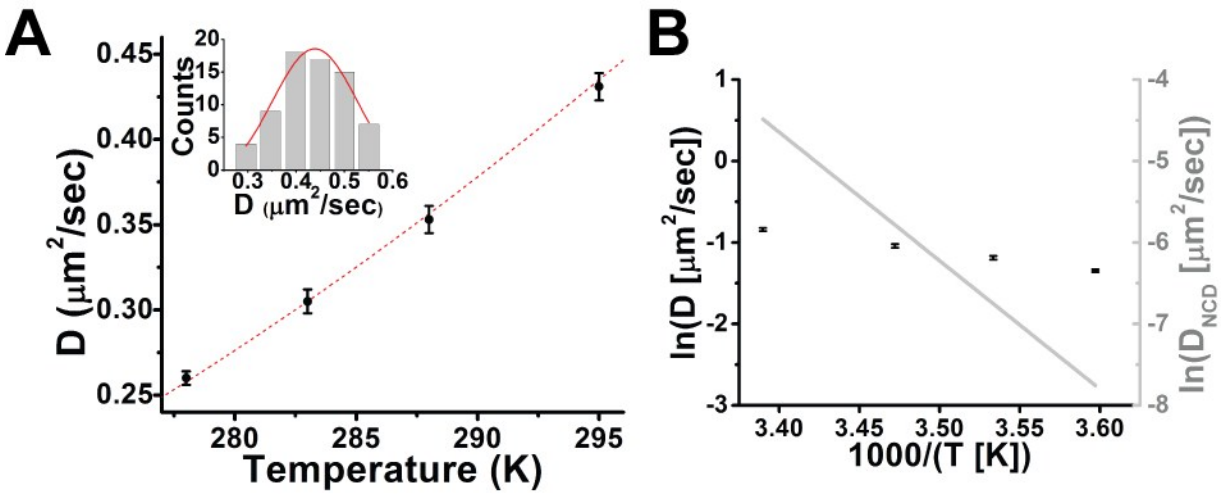

Fig. S1. Diffusion coefficients for motion of 1μm diameter beads in water. Weak temperature dependence for diffusion coefficient D is evident on both linear (A) and Arrhenius (B) plots. Fit to Stokes-Einstein equation for bead radius 505 nm is shown (A, red dashed line); nominal bead radius is 499±19 nm. (inset) Histogram of diffusion coefficients D at 22 °C and a Gaussian fit (red). (B) Bead diffusion in water (black) is superimposed on linear trend for N340K NCD diffusion.

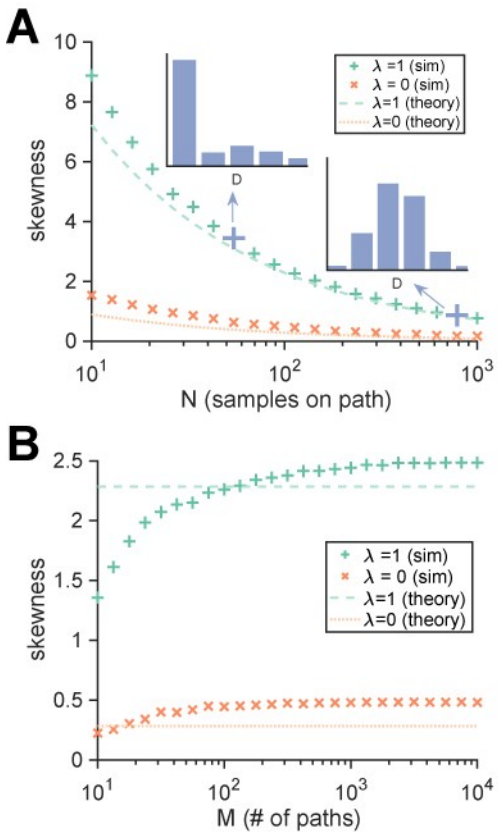

Fig. S2. Analytic and simulation predictions for the skewness of the distribution of empirical diffusion coefficients. Classical diffusion (λ=0) and jump diffusion (λ=1) cases are shown. (A) Skewness decays asymptotically to zero as a function of the number of points on each path (N). Simulation results are slightly but consistently higher than analytic predictions due to estimator bias. Inset histograms illustrate the skewness of the distribution of empirical diffusion coefficients for the indicated points in the main plot (blue crosses). (B) Simulated skewness as a function of the number of trajectories (M) shows asymptotic convergence to analytic prediction (up to estimator bias) for large M. Analytic curves: dashed lines. Simulated results: crosses. Parameters (unless noted): D=1, M=1000, N=100, Δt=1e-3, delta=0.8.

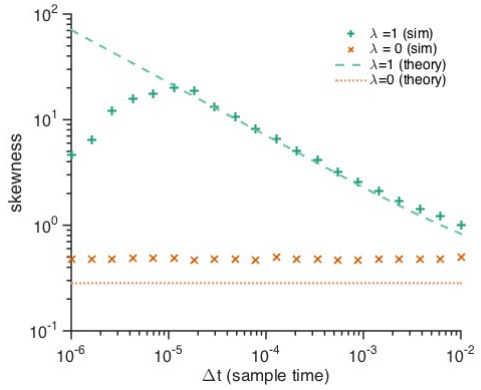

Fig. S3. Simulated skewness for small Δt. We examine the prediction from equation 4 in Supplement Text 1 that as Δt becomes much smaller than $1/\lambda$, the skewness should increase. The simulated skewness does indeed grow as expected within a wide range of dt. Of course, there is a technical limit on the number of samples (N) we can reasonably process, but we can fix N to be large and vary the sample time Δt (meaning that the total simulation time N*Δt also varies). Consequently, for any finite simulation it is always possible to examine low enough values of Δt, so that the length of simulation is not long enough to capture accurate statistics of the jump component. We extended our modeling up to this range of Δt.

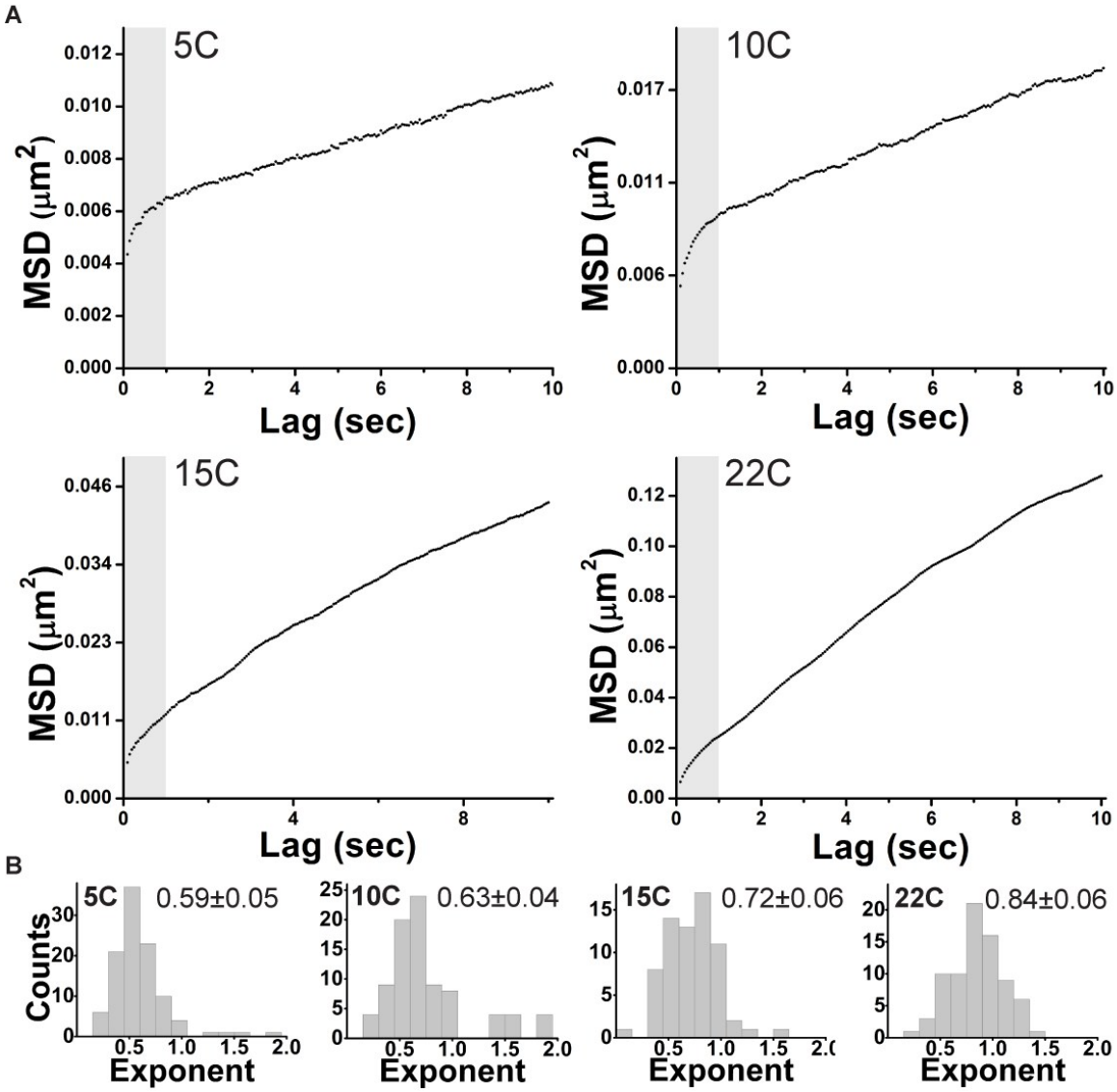

Fig. S4. Anomalous features of bead motility. (A) MSD curves were computed for 5, 10, 15, 22 °C data as labeled. (B) Sub-diffusive anomalous exponents for MSD records below 1 second (grey region in (A)) at 5, 10, 15, 22 °C as labeled. Peak locations and 95% C.I. are shown for each panel. Anomalous exponents were estimated via log-log linearization after adjustment for noise[SR3]. Each count corresponds to a full distinct bead trajectory.

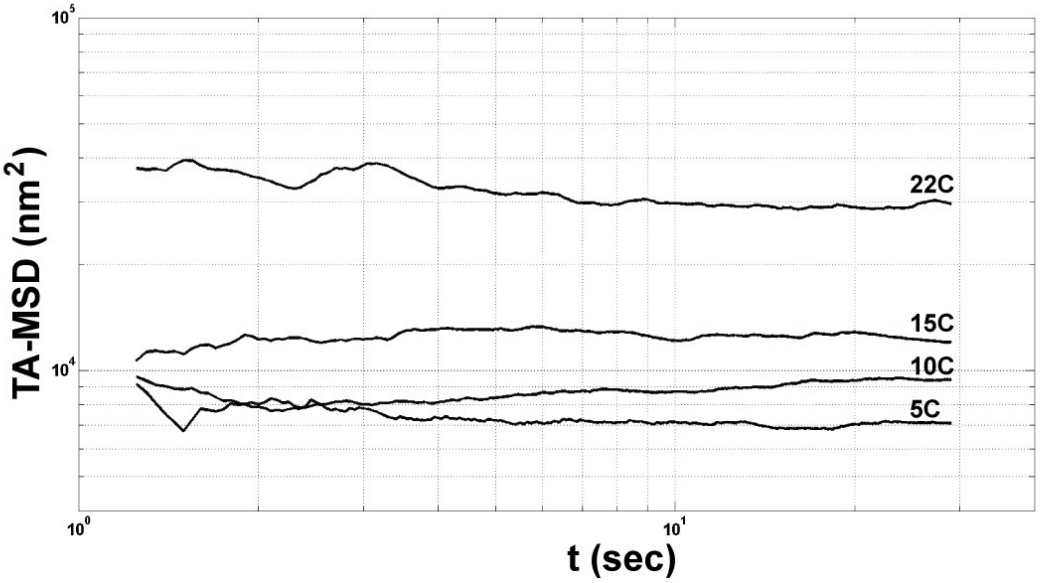

Fig. S5. TA-MSD curves show no definitive signature of aging. Given a discrete trajectory r(frame number), one can compute squared displacements within the trajectory $[r(\tau+s)-r(\tau)]^2$ for some lag s. Here, τ can vary from 1 to some chosen frame number t-s. The mean of the squared displacements computed as a function of t at a fixed s (here 1 sec) and then averaged across trajectories is a common measure of the dependence of MSD analysis on total measurement time. Frame numbers have been converted to frame timings for the purpose of plotting. The temperature is indicated just above each curve.

Supplement References

SR1. Gardiner, Crispin. 2009. Stochastic Methods - A Handbook for the Natural and Social Sciences. Springer-Verlag Berlin Heidelberg.

SR2. Aït-Sahalia, Y. 2004. Disentangling diffusion from jumps. J. Financ. Econ. 74: 487–528.

SR3. Kepten, E., A. Weron, G. Sikora, K. Burnecki, and Y. Garini. 2015. Guidelines for the Fitting of Anomalous Diffusion Mean Square Displacement Graphs from Single Particle Tracking Experiments. PLOS ONE. 10: e0117722.