Chaos
An Interdisciplinary Journal of Nonlinear Science

RESEARCH ARTICLE | OCTOBER 13 2023

Parameter estimation of the fractional Ornstein–Uhlenbeck process based on quadratic variation

Joanna Janczura ; Marcin Magdziarz ; Ralf Metzler

Chaos 33, 103125 (2023)
https://doi.org/10.1063/5.0158843

Parameter estimation of the fractional Ornstein–Uhlenbeck process based on quadratic variation

Cite as: Chaos 33, 103125 (2023); doi: 10.1063/5.0158843
Submitted: 18 May 2023 · Accepted: 26 September 2023 ·
Published Online: 13 October 2023

Joanna Janczura,[1] Marcin Magdziarz,[1,a)] and Ralf Metzler[2,3]

AFFILIATIONS

[1]Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland
[2]Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam-Golm, Germany
[3]Asia Pacific Centre for Theoretical Physics, Pohang 37673, Republic of Korea

[a)]**Author to whom correspondence should be addressed:** marcin.magdziarz@pwr.edu.pl

ABSTRACT

Modern experiments routinely produce extensive data of the diffusive dynamics of tracer particles in a large range of systems. Often, the measured diffusion turns out to deviate from the laws of Brownian motion, i.e., it is anomalous. Considerable effort has been put in conceiving methods to extract the exact parameters underlying the diffusive dynamics. Mostly, this has been done for unconfined motion of the tracer particle. Here, we consider the case when the particle is confined by an external harmonic potential, e.g., in an optical trap. The anomalous particle dynamics is described by the fractional Ornstein–Uhlenbeck process, for which we establish new estimators for the parameters. Specifically, by calculating the empirical quadratic variation of a single trajectory, we are able to recover the subordination process governing the particle motion and use it as a basis for the parameter estimation. The statistical properties of the estimators are evaluated from simulations.

Recent advances in single particle tracking experiments confirmed that various complex systems display anomalous, nonlinear in time mean squared displacement. Here, we analyze the dynamics of anomalous subdiffusive fractional Ornstein–Uhlenbeck process. This process is conveniently described by the fractional (in time) Fokker–Planck equation in the presence of harmonic potential. Equivalently, it can be described by subordinating the classical Ornstein–Uhlenbeck process. Using this fact, we establish dedicated estimators for the parameters of the anomalous diffusion of a test particle in a confining harmonic potential. To contract the estimators, we employ the empirical quadratic variation of a single trajectory. We evaluate the statistical properties of the estimators by extensive simulations.

I. INTRODUCTION

After Einstein and Smoluchowski published their theories for diffusion,[1,2] Perrin realized the potential to extract Avogadro's number from diffusive measurements,[3] an approach that was further improved by Nordlund[4] and a whole series of scientists. In turn, Smoluchowski derived the reaction rate of molecules following their diffusive encounter.[5] Today, measuring diffusive properties of tracer particles in complex environments such as soft and bio matter inform about materials properties and biological processes in living cells or tissues.[6–9]

In many examples, the measured diffusive dynamics is anomalous in the sense that the mean squared displacement (MSD) of the tracked particle follows the power-law form[11–13]

$$\langle x^2(t)\rangle \simeq K_\alpha t^\alpha, \tag{1}$$

where K_α with dimensions length/time$^\alpha$ is the generalized diffusion coefficient and α is the anomalous diffusion exponent. We distinguish subdiffusion ($0 < \alpha < 1$) and superdiffusion ($\alpha > 1$), including the limiting cases of normal Brownian diffusion with $\alpha = 1$ and wave-like, ballistic propagation for $\alpha = 2$. As shown by Einstein and Smoluchowski, Brownian motion emerges from a stochastic process when its increments are independent, identically distributed random variables with finite variance.[14,15] Violation of one or more of these conditions gives rise to anomalous diffusion.

In cases where the MSD is not well defined (for example, in the case of stable processes), one can replace MSD in (1) with the fractional moments

$$\langle |x(t)|^q \rangle \sim t^{cq}. \tag{2}$$

Then, the case $c > 1/2$ is typical for superdiffusion and $c < 1/2$ is characteristic for subdiffusion.

Following the detailed diffusive characteristics unveiled by experiments in many different systems, a large range of different anomalous diffusion processes have been established.[16–18] Given the variety of possible processes that may effect the MSD (1) with given values of α and K_α, from an application point of view, it is important to extract the underlying parameters but also the exact anomalous diffusion process for a given system. Different approaches have been conceived, such as complementary stochastic observables,[17,19–24] Bayesian maximum likelihood methods,[25–28] q-moments,[29] p-variation,[30,31] as well as machine learning approaches.[32–36]

Most of these approaches are designed for the analysis of unconfined motion. Here, we establish dedicated estimators for the parameters of the anomalous diffusion of a test particle in a confining harmonic potential. We describe the dynamics of the particle by the fractional Ornstein–Uhlenbeck process,[10,11] which corresponds to a subdiffusive continuous time random (CTRW) experiencing a restoring linear force. Subdiffusive CTRWs are characterized as jump processes whose jump lengths are drawn from a Gaussian density, while they experience random waiting times between successive jumps that are governed by long-tailed densities of the form $\psi(\tau) \simeq \tau^{-1-\alpha}$ with $0 < \alpha < 1$ such that the characteristic waiting time $\langle \tau \rangle$ diverges.[11,16,37]

Long-tailed waiting time densities were experimentally observed for the motion of potassium channels in cell membranes,[38] tracer motion in actin gels,[39,40] messenger ribonucleoprotein transport in neuron cells,[41] functionalized colloidal particles along a complementarily functionalized surface,[42] in weakly chaotic systems,[43] in porous media,[44] and even for the search mode of tracked kites.[45] In simulations, power-law waiting times were, i.a., observed for the motion of drug molecules in silica slabs.[46]

Harmonic oscillators are prototype systems in physics. In the context of single particle tracking, they naturally occur when the tracked particle is initially placed at the bottom of a harmonic trap potential in an optical tweezer setup and the stiffness of the trap is chosen to be low.[8,47–49] Alternatively, a tracer attached to a flexible polymer, whose other end is fixed, will experience a Hookean force.[50] We also mention internal protein dynamics monitored by the relative motion of two labeled amino acids, which are linked by a number of amino acids along the sequence.[51] Harmonic confinement can also be viewed as a simple model for home-ranging effects in the movement dynamics of animals.

In this paper, we employ the idea of quadratic variation and propose dedicated estimators for the associated model parameters. We verify their statistical properties using simulations of the underlying stochastic process. The paper is organized as follows. In Sec. II, we introduce the fractional Ornstein–Uhlenbeck process and recall its main properties. Next, in Sec. III, we propose new estimators for the parameters of this process. The accuracy of these estimators is then evaluated on the basis of extensive simulations in Sec. IV. In Sec. V, we draw our conclusions.

II. FRACTIONAL ORNSTEIN–UHLENBECK PROCESS

The fractional Ornstein–Uhlenbeck process $X_\alpha(t)$ is given by the subordination scheme,[52–55]

$$X_\alpha(t) = X(S_\alpha(t)), \quad t \in [0, T], \tag{3}$$

where $X(t)$ is the standard Ornstein–Uhlenbeck process, given as the solution of the Îto stochastic differential equation[14,56]

$$dX(\tau) = -\lambda_1 X(\tau) d\tau + (2K_1)^{1/2} dB(\tau), \quad X(0) = 0, \tag{4}$$

where $\lambda_1 > 0$ and $K_1 > 0$. Here, the stochastic driving force is the differential of the Wiener process $B(\tau)$. Moreover, λ_1 corresponds to the strength of the trapping harmonic potential, corresponding to the inverse correlation time of the Ornstein–Uhlenbeck process. The mapping of the Brownian Ornstein–Uhlenbeck process $X(\tau)$ as a function of the time τ (operational time corresponding to the number of steps[57,58]) to the process time t of the fractional Ornstein–Uhlenbeck process $X_\alpha(t)$ is given in terms of a subordinator.[59–61] In our notation, this mapping is based on $S_\alpha(t)$, the inverse α-stable subordinator, defined as[52–55]

$$S_\alpha(t) = \inf\{\tau > 0 : U_\alpha(\tau) > t\}, \tag{5}$$

where $U_\alpha(t)$ is the α-stable subordinator with Laplace transform $\langle \exp(-uU_\alpha(\tau)) \rangle = \exp(-\tau u^\alpha)$.

Sample trajectories of the process $X_\alpha(t)$ are plotted in Fig. 1. Note that the subordination in formula (3) changes the time scale of the Ornstein–Uhlenbeck process. For every jump of the process $U_\alpha(\tau)$, there is a flat period in its inverse, $S_\alpha(t)$, yielding equal values of the process $X_\alpha(t)$ over the corresponding waiting time. The constant periods are clearly visible in Fig. 1, and their length is longer for lower values of the parameter α, i.e., heavier tails of the corresponding α-stable distribution.

The probability density function (PDF) $w(x, t)$ of the process $X_\alpha(t)$ is the solution of the fractional Fokker-Planck equation[10,11,53]

$$\frac{\partial w(x,t)}{\partial t} = {}_0D_t^{1-\alpha} \left[-\lambda_\alpha x + K_\alpha \frac{\partial^2}{\partial x^2} \right] w(x,t), \tag{6}$$

where λ_α is the generalized inverse correlation time of the fractional Ornstein–Uhlenbeck process with dimension $(\sec^{-\alpha})$.[11] The operator ${}_0D_t^{1-\alpha}$ with $\alpha \in (0, 1)$ is the fractional Riemann–Liouville derivative.[11] Hence, the fractionality of the considered process is related to time.

III. ESTIMATION OF PARAMETERS

In this section, we propose new estimators for the parameters of the fractional Ornstein–Uhlenbeck process $X_\alpha(t)$, based on properties of the corresponding quadratic variation $V^2(t)$. Specifically, we use the fact that the inverse subordinator $S_\alpha(t)$ is linked with the quadratic variation of $X_\alpha(t)$ in the following way:[23]

$$V^{(2)}(t) = 2KS_\alpha(t). \tag{7}$$

Hence, the subordinator $S_\alpha(t)$ can be recovered from the process $V^2(t)$. Having a trajectory of the process $X_\alpha(t)$, its quadratic

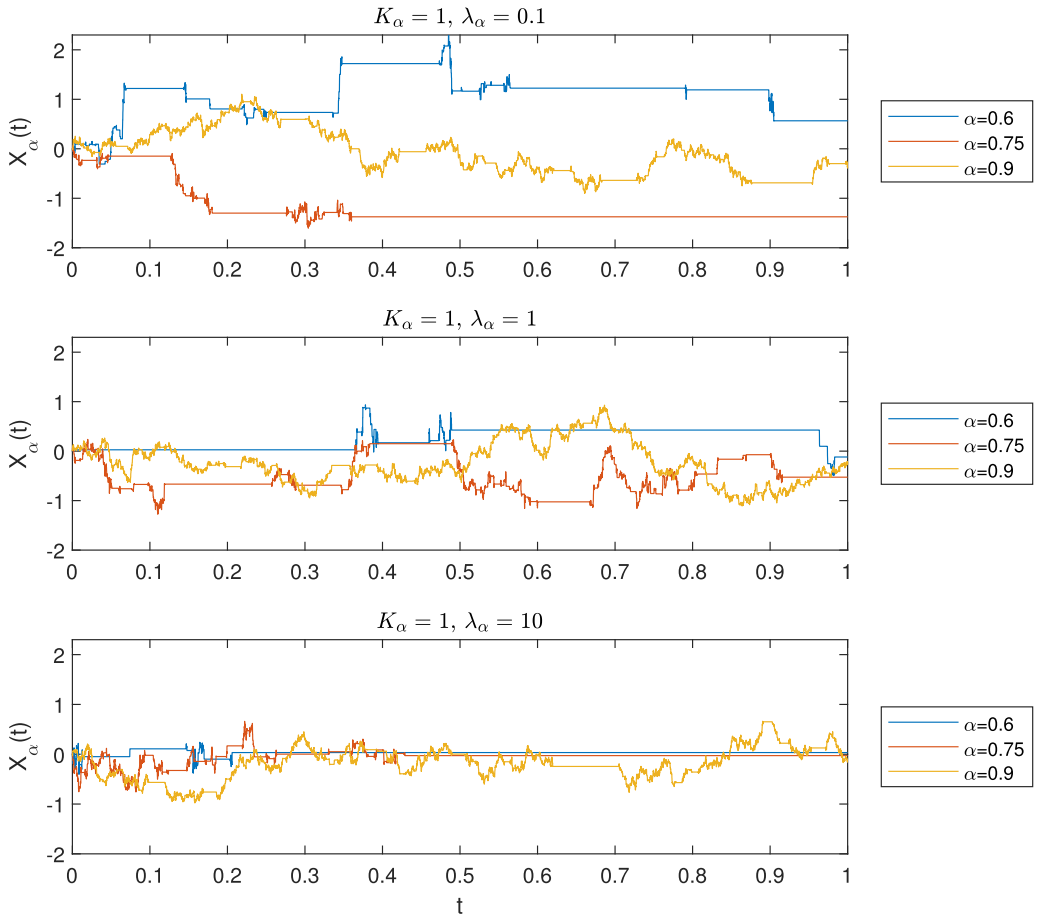

FIG. 1. Sample trajectories of the fractional Ornstein–Uhlenbeck process with $\lambda_\alpha = 0.1$ (upper panel), $\lambda_\alpha = 1$ (middle panel), or $\lambda_\alpha = 10$ (bottom panel) and $\alpha \in \{0.6, 0.75, 0.9\}$, $K_\alpha = 1$, for time step $\Delta t = 10^{-4}$.

variation can be estimated using its empirical form

$$\hat{V}^2(t) = \sum_{k=1}^{n} \left[X_\alpha(t_k) - X_\alpha(t_{k-1})\right]^2, \tag{8}$$

where $(t_0, t_1, \ldots, t_n)$ is a given partition of the interval $[0, t]$.

In the subordinator picture, flat periods of $S_\alpha(t)$ correspond to the jumps of the process $U_\alpha(t)$. Hence, their lengths constitute an independent, identically distributed sample of totally skewed α-stable random variables. This sample can be approximately recovered from a trajectory of $X_\alpha(t)$, using the formulas (7) and (8), as the sample of the lengths of constant periods in the estimated quadratic variation process $\hat{V}^2(t)$. Moreover, the value of the parameter α can be estimated from that sample using one of the known estimation methods for the parameters of the α-stable distribution. In the following, we will use the regression[62] as well as McCulloch[63] methods, for other methods see, e.g., Refs. 64 and 65.

Next, we introduce the estimator of the scale parameter K_α. To this end, we use the quadratic variation properties (7) and (8) along with the fact that[66]

$$\langle S_\alpha(t) \rangle = \frac{t^\alpha}{\Gamma(\alpha + 1)}, \tag{9}$$

where Γ is the complete Gamma function.

Taking the last value from the considered interval $[0, T]$ in (9) and rearranging Eq. (7), we obtain the estimator $\hat{K_\alpha}$,

$$\hat{K_\alpha} = \frac{\Gamma(\hat{\alpha} + 1)}{2T^{\hat{\alpha}}} \hat{V}^{(2)}(T), \tag{10}$$

where $\hat{\alpha}$ is the value of the previously estimated α parameter.

We now use the least squares technique to derive a new estimator for the parameter λ_α. We recall that $X_\alpha(t) = X(S_\alpha(t))$, where X satisfies Eq. (4). Therefore, X_α is the solution of

$$dX_\alpha(t) = -\lambda_\alpha X_\alpha(t) dS_\alpha(t) + (2K_\alpha)^{\frac{1}{2}} dB(S_\alpha(t)), \tag{11}$$

and rearranging we have

$$\frac{dX_\alpha(t)}{dS_\alpha(t)} + \lambda_\alpha X_\alpha(t) = (2K_\alpha)^{\frac{1}{2}} b_\alpha(t), \tag{12}$$

where $b_\alpha(t) = \frac{dB(S_\alpha(t))}{dS_\alpha(t)}$ is the subdiffusive white noise. Now, to obtain the least squares estimator of λ_α, we want to minimize the integral from the square of the left side of the above equation, i.e., to minimize the following expression with respect to λ_α,

$$\int_0^T \left| \frac{dX_\alpha(t)}{dS_\alpha(t)} + \lambda_\alpha X_\alpha(t) \right|^2 dS_\alpha(t) = \int_0^T \left| \frac{dX_\alpha(t)}{dS_\alpha(t)} \right|^2 dS_\alpha(t) + 2\lambda_\alpha \int_0^T X_\alpha(t) dX_\alpha(t) + \lambda_\alpha^2 \int_0^T X_\alpha^2(t) dS_\alpha(t). \quad (13)$$

Finally, the above expression is the standard quadratic polynomial with respect to λ_α. Its minimum is attained for

$$\hat{\lambda}_\alpha = -\frac{\int_0^T X_\alpha(t) dX_\alpha(t)}{\int_0^T X_\alpha^2(t) dS_\alpha(t)}, \quad (14)$$

Formula (14) describes our new estimator of λ_α. Note that for practical uses, the integrals in (14) are approximated using partial sums and $S_\alpha(t)$ is recovered from Eqs. (7), (8), and (10), i.e.,

$$\hat{S}_\alpha(t) = \frac{\hat{V}^2(t)}{2\hat{K}_\alpha}. \quad (15)$$

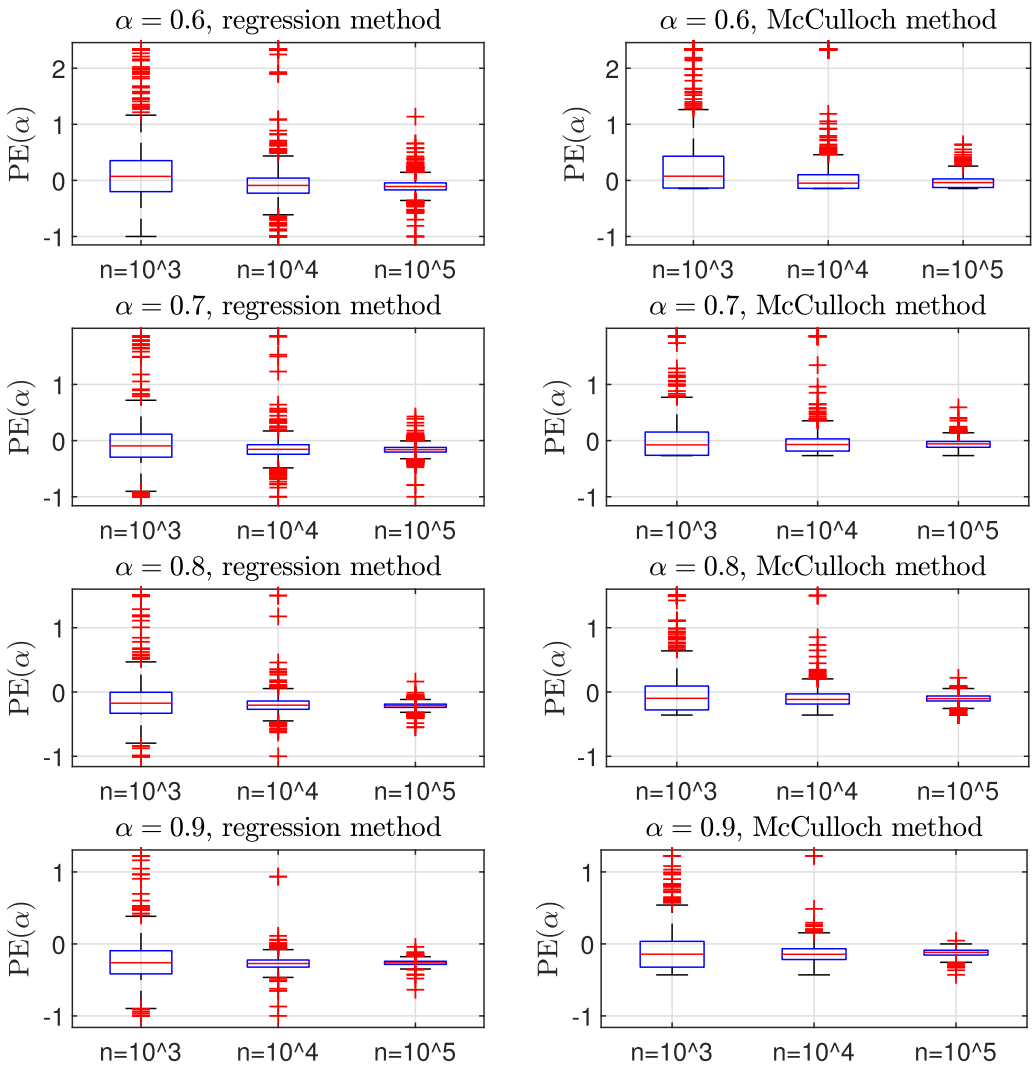

FIG. 2. Boxplots of the percentage errors (PEs) of the α-estimator obtained from 10^3 simulated trajectories of the fractional Ornstein–Uhlenbeck process using the regression (left panels) or McCulloch (right panels) methods. Results for different α parameters (0.6, 0.7, 0.8, or 0.9) are plotted in different rows. The other parameters were set to $K_\alpha = 1$ and $\lambda_\alpha = 1$, and we chose $T = 1$. The number of trajectory points was equal to $n = 10^3$, $n = 10^4$, or $n = 10^5$.

IV. SIMULATIONS

In order to verify the performance of the proposed estimators, we simulate 10^3 trajectories of the fractional Ornstein–Uhlenbeck process with $n = 10^3$, $n = 10^4$, or $n = 10^5$ points on the time interval [0,1], i.e., with time steps $\Delta t = \frac{T}{n}$ equal to 10^{-3}, 10^{-4}, and 10^{-5}, respectively. The parameters are set to $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$, $K_\alpha \in \{0.1, 1, 10\}$, $\lambda_\alpha \in \{0.1K, K, 10K\}$. For practical illustration, in the Appendix, we also show sample results obtained for a lower number of trajectories.

We start with the calculation of the percentage errors (PEs) of the estimated values. The PE of an estimator is obtained as

$$\mathrm{PE}(\theta) = \frac{\hat{\theta} - \theta}{\theta}, \tag{16}$$

where $\hat{\theta}$ is the estimator of the parameter θ.

In Fig. 2, we show the boxplots of the obtained PEs of the α estimator based on the regression as well as the McCulloch methods. The other parameters were set to $K_\alpha = 1$ and $\lambda_\alpha = 1$. Recall that a boxplot is a visualization of a distribution. Concretely, the box indicates the range between the 25th and 75th percentile, the central red line is the median and the whiskers extend to the most extreme data points not considered to be outliers, while the red crosses mark the outliers. In all cases, the widths of the boxplots decrease with the sample size. This convergence is slightly faster for the regression method. However, especially for larger values of α both estimators are biased, yielding lower estimates than the true value. This effect is more pronounced for the regression method. In this case, for $\alpha \geq 0.7$, the true value lies outside the range expected from the estimator distribution. Hence, for the estimation of K_α and λ_α parameters, we will use α estimates obtained from the McCulloch method. Bias of the estimators is caused by the fact that the data are observed in the discrete time point only; therefore, in the analyzed samples, observations shorter than Δt are not available. Also, it is a general rule that the estimators of stable index perform worse for α close to two.

The boxplots of the PEs of $\hat{\lambda_\alpha}$ and $\hat{K_\alpha}$ are plotted in Fig. 3. The other parameters were set to $\alpha = 0.7$ and $K_\alpha = 1$ for the λ_α parameter or $\alpha = 0.7$ and $\lambda_\alpha = 1$ for the K_α estimation, respectively. For both parameters, the convergence of the estimator distribution with a growing number of trajectory points is barely visible. In the case of the K_α parameter estimates, the PEs have a similar distribution for all choices of the true parameter with medians close to the true value. For the λ_α parameter, the situation is different. Here, the PEs are much larger for smaller values of λ_α than K_α. The most accurate estimates were obtained for $\lambda_\alpha = 10$ with the smallest percentage errors and median close to the true value. Note that for relatively large K_α and relatively small λ_α, the strength of the noise obscures the action of harmonic potential. In such cases, as expected, the estimation of λ_α is biased.

Next, we calculate the means and standard errors (se) of the estimated values for all considered combinations of the process parameters and the longest trajectories with $n = 10^5$ points. Recall that the standard error of the estimator is equal to the standard deviation of the difference between the estimator and the true value of the parameter,

$$\mathrm{se}(\hat{\theta}) = \sqrt{\langle(\hat{\theta} - \theta)^2\rangle}. \tag{17}$$

The obtained values are given in Table I. The corresponding boxplots of the obtained estimates are depicted in Figs. 4–6 for α, K_α, and λ_α, respectively. For the estimation of K_α and λ_α, we use the α

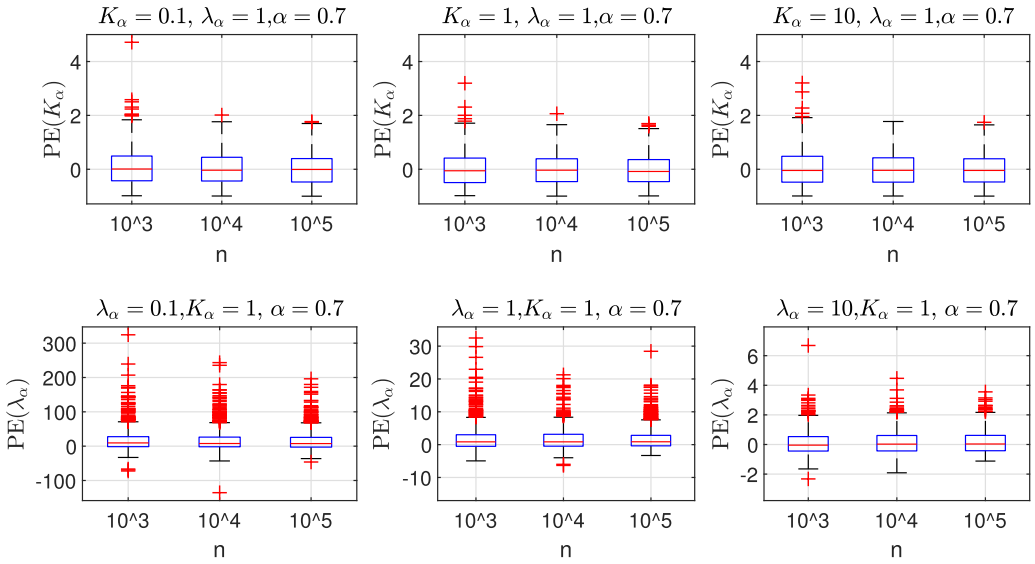

FIG. 3. Boxplots of the PEs of K_α (upper panels) and λ_α (bottom panels) estimator obtained from 10^3 simulated trajectories of the fractional Ornstein–Uhlenbeck process with different values of K_α and λ_α. The other parameters were set to $\lambda_\alpha = 1$, $\alpha = 0.7$, $T = 1$, or $K_\alpha = 1$, $\alpha = 0.7$, and $T = 1$ for the estimation of K_α or λ_α, respectively. The number of trajectory points was equal to $n = 10^3$, $n = 10^4$, and $n = 10^5$.

TABLE I. Means and standard errors (se) of the estimated parameters based on 10^3 trajectories of the fractional Ornstein–Uhlenbeck process simulated with $n = 10^5$ points on the interval [0,1] with different values of the parameters α, K_α, and λ_α. The parameter α is estimated using the regression (reg) or McCulloch (MC) methods.

Simulated parameters			Estimated parameters							
			α - reg		α - MC		K_α		λ_α	
α	K_α	λ_α	Mean	se	Mean	se	Mean	se	Mean	se
0.60	0.10	0.01	0.54	0.11	0.59	0.10	0.10	0.07	1.55	2.77
0.60	0.10	0.10	0.54	0.12	0.59	0.10	0.10	0.07	1.67	2.75
0.60	0.10	1.00	0.54	0.12	0.60	0.12	0.10	0.07	2.64	3.32
0.60	1.00	0.10	0.55	0.15	0.60	0.13	1.00	0.66	1.61	2.73
0.60	1.00	1.00	0.54	0.09	0.58	0.07	1.02	0.70	2.67	3.01
0.60	1.00	10.00	0.54	0.12	0.59	0.10	0.97	0.65	11.18	8.11
0.60	10.00	1.00	0.54	0.14	0.60	0.14	9.68	6.50	2.76	3.38
0.60	10.00	10.00	0.53	0.10	0.59	0.08	9.94	6.72	11.41	8.00
0.60	10.00	100.00	0.54	0.10	0.59	0.08	9.23	6.03	85.00	55.63
0.70	0.10	0.01	0.59	0.06	0.65	0.06	0.10	0.06	1.27	2.49
0.70	0.10	0.10	0.59	0.05	0.65	0.06	0.10	0.06	1.71	2.92
0.70	0.10	1.00	0.59	0.09	0.66	0.09	0.10	0.06	2.71	3.30
0.70	1.00	0.10	0.59	0.09	0.66	0.10	0.99	0.58	1.68	2.85
0.70	1.00	1.00	0.59	0.07	0.65	0.06	0.97	0.56	2.73	3.26
0.70	1.00	10.00	0.59	0.07	0.66	0.07	0.99	0.58	11.54	7.50
0.70	10.00	1.00	0.59	0.07	0.66	0.08	9.84	5.73	2.70	3.17
0.70	10.00	10.00	0.59	0.11	0.66	0.11	9.97	5.53	12.02	7.67
0.70	10.00	100.00	0.59	0.09	0.66	0.09	9.65	5.66	94.60	55.71
0.80	0.10	0.01	0.63	0.07	0.72	0.08	0.10	0.04	1.70	2.93
0.80	0.10	0.10	0.63	0.07	0.72	0.08	0.10	0.04	1.77	3.03
0.80	0.10	1.00	0.63	0.07	0.72	0.07	0.10	0.05	2.56	3.21
0.80	1.00	0.10	0.63	0.04	0.72	0.05	0.99	0.46	1.54	2.58
0.80	1.00	1.00	0.63	0.04	0.72	0.05	0.96	0.46	2.82	3.36
0.80	1.00	10.00	0.63	0.04	0.72	0.05	0.99	0.45	11.79	6.76
0.80	10.00	1.00	0.63	0.05	0.72	0.07	9.82	4.43	2.75	3.10
0.80	10.00	10.00	0.63	0.04	0.72	0.05	9.73	4.50	11.35	6.44
0.80	10.00	100.00	0.63	0.06	0.72	0.07	9.86	4.56	99.59	47.40
0.90	0.10	0.01	0.66	0.03	0.79	0.05	0.10	0.03	1.88	3.10
0.90	0.10	0.10	0.66	0.03	0.79	0.05	0.10	0.03	1.73	2.91
0.90	0.10	1.00	0.66	0.05	0.79	0.06	0.10	0.03	2.71	3.25
0.90	1.00	0.10	0.66	0.03	0.78	0.05	0.97	0.31	1.93	3.01
0.90	1.00	1.00	0.66	0.03	0.79	0.05	0.97	0.32	2.69	3.30
0.90	1.00	10.00	0.66	0.03	0.79	0.05	0.95	0.32	11.01	5.83
0.90	10.00	1.00	0.66	0.05	0.79	0.06	9.64	3.14	2.72	3.16
0.90	10.00	10.00	0.66	0.03	0.78	0.05	9.51	3.10	11.59	5.93
0.90	10.00	100.00	0.66	0.04	0.79	0.05	9.70	3.11	98.57	34.10

parameter estimates based on the McCulloch method, as it yielded more reliable results.

The results obtained for the α estimators confirm the conclusions drawn from the boxplots of the PEs. The regression estimator yields lower standard errors than the McCulloch one, but, concurrently, it is more biased. The values of the other parameters, i.e., λ_α and K_α, have no influence on the results of the α estimation. The length of the time interval here and in the following two cases is chosen as $T = 1$.

For the K_α estimation, the obtained means are close to the true values for all chosen parameters. The standard errors are proportional to the true parameter values and they clearly decrease with increasing α. Recall that for larger α the waiting periods of the process $X_\alpha(t)$ are typically shorter.

The λ_α estimation results yield a different picture. Now, the obtained values depend markedly on the choice of the K_α parameter. When λ_α is much smaller than K_α, the standard errors of the estimator are very high, resulting in mean values that are far from

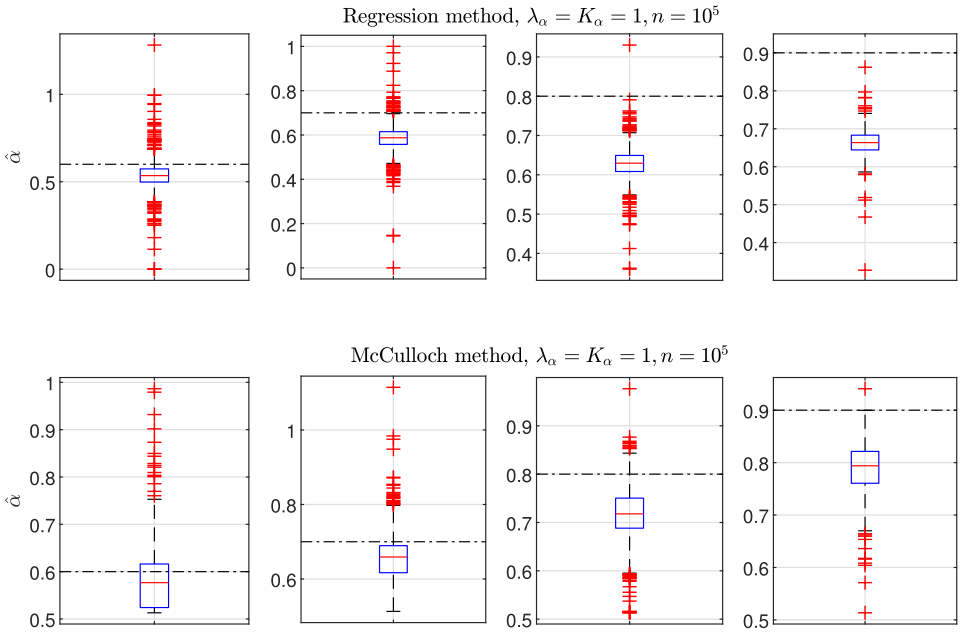

FIG. 4. Boxplots of the estimated values of α obtained from 10^3 simulated trajectories of the fractional Ornstein–Uhlenbeck process and the regression (upper panels) or McCulloch (bottom panels) methods. The other parameters were set to $\lambda_\alpha = 1$, $K_\alpha = 1$, $T = 1$, and $n = 10^5$. True values of α are marked with black dashed lines.

the simulated values. When λ_α is larger than K_α, the standard errors are proportionally smaller. Again, there is a visible dependence of the estimation results on the simulated α values. The standard errors of $\hat{\lambda}_\alpha$ are smaller for larger values of α.

A. Estimation in the presence of localization errors

We also analyze the performance of the estimation procedure in the presence of localization errors. Following Ref. 67, we consider static errors, which occur due to imperfect measurement,

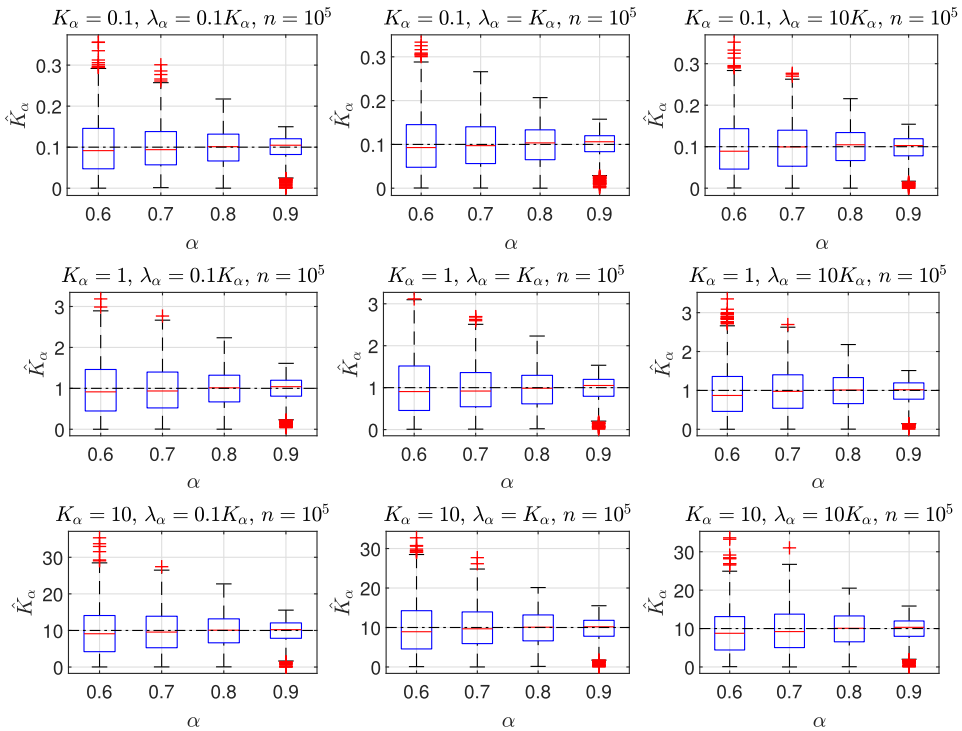

FIG. 5. Boxplots of the estimated values of K_α obtained from 10^3 simulated trajectories of the fractional Ornstein–Uhlenbeck process. True values of K are marked with black dashed lines.

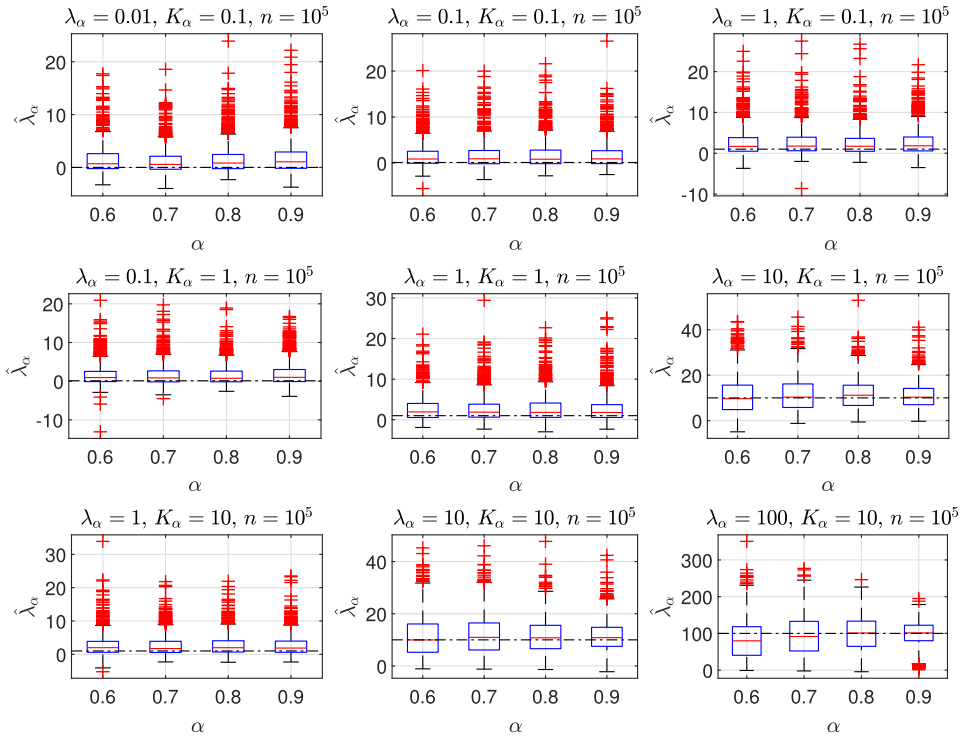

FIG. 6. Boxplots of the estimated values of λ_α obtained from 10^3 simulated trajectories of the fractional Ornstein–Uhlenbeck process. True values of λ_α are marked with black dashed lines.

as well as dynamic errors, which come from limited time resolution of the measurement. We assume, in a standard manner, that the static error e_t is an additive random noise following a Gaussian distribution of the form

$$p(e_t) = \sqrt{\frac{1}{4\pi\sigma_e^2\Delta t}}\exp\left(-\frac{e_t^2}{4\sigma_e^2\Delta t}\right) \tag{18}$$

independent of the previous measurements. The noisy simulated trajectories are then simply obtained as

$$\widetilde{X}_\alpha(t_k) = X_\alpha(t_k) + e_{t_k}. \tag{19}$$

On the other hand, trajectories measured with dynamic localization errors are the averages of the real process

$$\overline{X}_\alpha(t) = \frac{1}{\tau_e}\int_0^{\tau_e} X_\alpha(t-\xi)d\xi \tag{20}$$

depending on the exposure time τ_e. Here, we simulate trajectories of the fractional Ornstein–Uhlenbeck process on a dense grid of $n = 10^5$ points on the interval $[0, 1]$ and then average them over subintervals as

$$\overline{X}_\alpha(t_k) = \frac{1}{n_e}\sum_{i=0}^{n_e-1} X_\alpha(t_{k-i}), \tag{21}$$

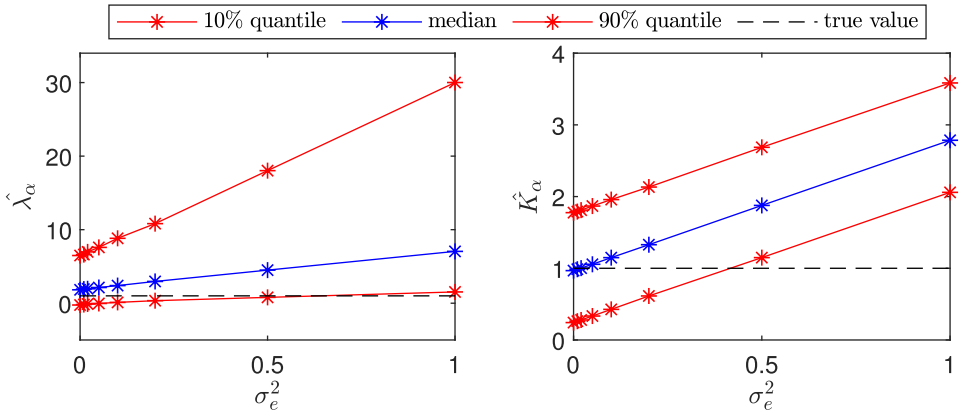

FIG. 7. Medians and 10% or 90% quantiles of the estimated values of λ_α (left panel) or K_α (right panel) for trajectories with static localization errors, $\widetilde{X}_\alpha$. The true simulated values of the parameters are marked with black dashed lines. The α parameter was set to 0.7, while $T = 1$ and $\Delta t = 10^{-5}$. The quantiles were calculated from 1000 repetitions of the estimation procedure.

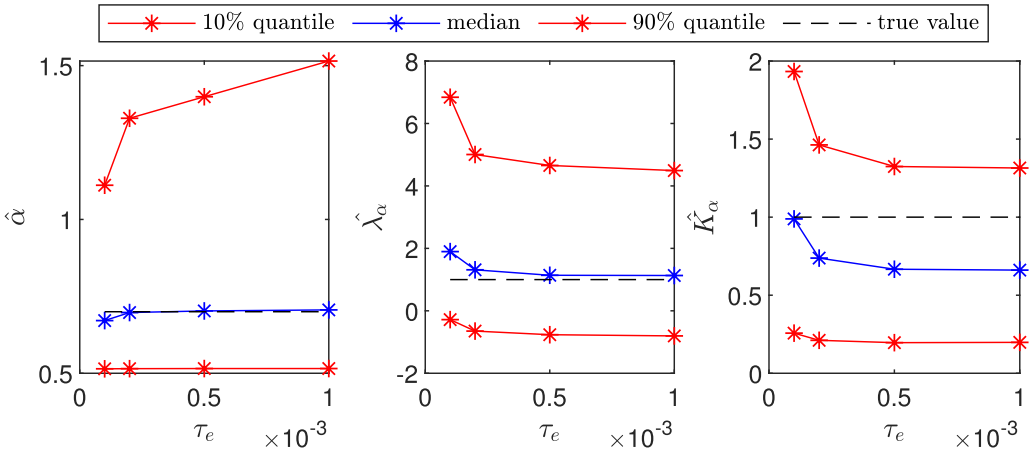

FIG. 8. Medians and 10% or 90% quantiles of the estimated values of α (left panel), λ_α (middle panel), or K_α (right panel) for trajectories with dynamic localization errors, $\overline{X}_\alpha$. The true simulated values are marked with black dashed lines. The quantiles were calculated from 1000 repetitions of the estimation procedure.

with $n_e = \tau_e/\Delta t$ being the number of the corresponding time steps.

The estimation results for trajectories with static localization errors for different values of the parameter σ_e^2 are plotted in Fig. 7. The parameters of the fractional Ornstein–Uhlenbeck process were set to $\alpha = 0.7, \lambda_\alpha = 1, K_\alpha = 1, T = 1$, and $\Delta t = 10^{-5}$. The estimation procedure was repeated for 1000 independent trajectories. Note that adding a random noise to the positions of the trajectories also adds variability to the constant time periods. Estimation of α would require setting an additional tolerance level, below which the variability is only due to the noise. Hence, for simplicity, we assume here that the α parameter is known. The static localization errors increase the variability of the λ_α parameter estimation. With higher variance of the errors, the values of $\hat{\lambda}_\alpha$ are also higher. This effect is likely caused by the errors of the estimation of K_α, whose values grow linearly with the variability of the static localization errors.

The estimation results for trajectories with dynamic localization errors for different values of τ_e are plotted in Fig. 8. The parameters of the fractional Ornstein–Uhlenbeck process were set to $\alpha = 0.7, \lambda_\alpha = 1$, and $K_\alpha = 1$. The estimation procedure was repeated for 1000 independent trajectories.

The variability of the α estimation is higher, with higher values of τ_e, due to shorter samples of constant time periods in $\overline{X}_\alpha$. However, the median value of $\hat{\alpha}$ is similar for all considered cases of dynamic errors. Estimation of the generalized diffusion coefficient from the averaged values on subintervals leads to lower than simulated values of K_α. The effect of decreasing values of the estimator with increasing size of dynamic errors is also visible in the case of λ_α. However, here it does not lead to a lower accuracy for the considered levels of τ_e.

V. CONCLUSIONS

We here introduced new estimators for the parameters of the fractional Ornstein–Uhlenbeck process. Using the empirical quadratic variation of a single trajectory, we were able to estimate the subordinator process. It was further used as a basis for the parameter estimation of the anomalous diffusion exponent α, the generalized diffusion coefficient K_α, and the generalized inverse time scale λ_α.

To verify the performance of the estimators, we performed a simulation study. In the case of the α parameter estimation, we found that the regression method yields lower standard errors than the McCulloch method, but at the same time it is more biased, especially for larger values of the α parameter. For both methods, increasing the sample size decreases the variance of the estimators. However, both methods lead to the underestimation of the α values. This effect is more pronounced for the regression method and larger values of α. In the case of the K parameter, the proposed estimator yields reliable results, as the obtained means are close to the true values. The standard errors are proportional to the true parameter values. Concurrently, the accuracy of the λ_α parameter estimation is highly dependent on the relation between λ_α and K_α. The most precise results were obtained when λ_α is higher than K_α. Finally, we found that the variance of the λ_α as well as K_α estimators is higher for lower values of α. There is no such dependence between the results obtained for the α parameter estimation and the chosen values of λ_α and K_α.

The results obtained here are expected to be useful for direct parameter estimates from experiments and simulations. They may also be useful as benchmarks for dedicated Bayesian or machine learning-based estimators.

ACKNOWLEDGMENTS

This work was supported by NCN Sonata Bis under Grant No. 2019/34/E/ST1/00360. R.M. acknowledges support from the German Science Foundation (DFG, Grant No. ME 1535/12-2).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Joanna Janczura: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Visualization (equal); Writing – original draft (equal). **Marcin Magdziarz:** Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Supervision (equal); Writing – original draft (equal). **Ralf Metzler:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Writing – original draft (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX: ADDITIONAL SIMULATIONS

In Figs. 9 and 10, we show the estimation results corresponding to a lower number of trajectories, representing a situation that is often encountered in experiments.

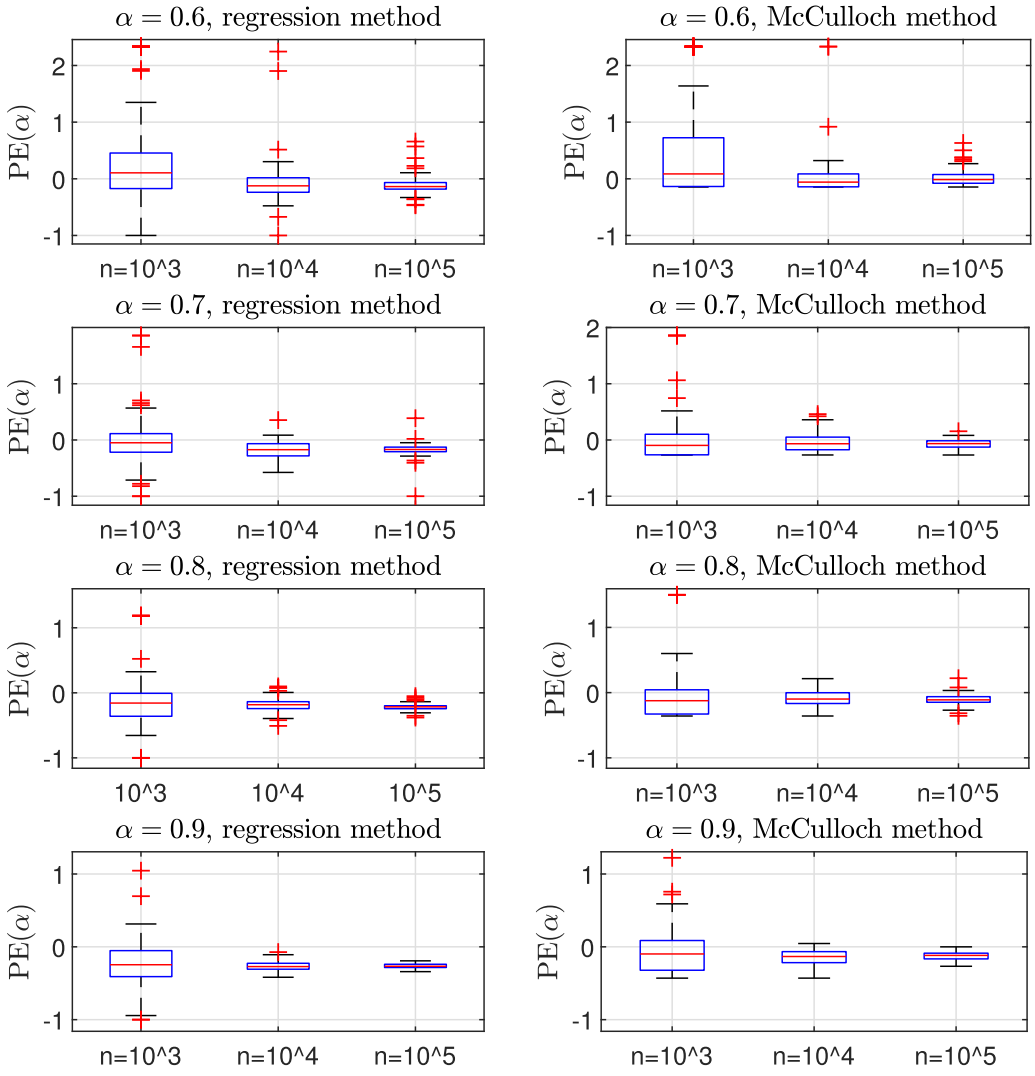

FIG. 9. Boxplots of the PEs of the α estimator obtained from 10^2 simulated trajectories of the fractional Ornstein–Uhlenbeck process using the regression (left panels) or McCulloch (right panels) methods. Results for different α parameters (0.6, 0.7, 0.8, or 0.9) are plotted in different rows. The other parameters were set to $K_\alpha = 1$, $\lambda_\alpha = 1$, and $T = 1$. The number of trajectory points was equal to $n = 10^3$, $n = 10^4$, and $n = 10^5$.

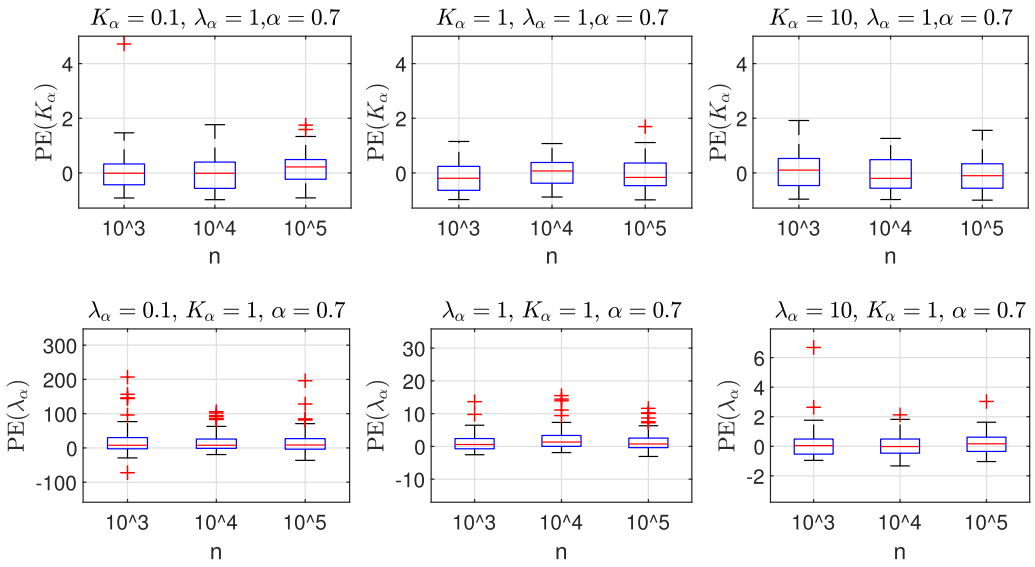

FIG. 10. Boxplots of the PEs of the K_α (upper panels) and λ_α (bottom panels) estimators obtained from 10^2 simulated trajectories of the fractional Ornstein–Uhlenbeck process with different values of K_α or λ_α. The other parameters were set to $\lambda_\alpha = 1$, $\alpha = 0.7$, and $T = 1$ or $K_\alpha = 1$, $\alpha = 0.7$, $T = 1$ for the K_α and λ_α estimation, respectively. The number of trajectory points was equal to $n = 10^3$, $n = 10^4$, or $n = 10^5$.

REFERENCES

[1] A. Einstein, Ann. Phys. (Leipzig) **322**, 549 (1905).
[2] M. von Smoluchowski, Ann. Phys. (Leipzig) **21**, 756 (1906).
[3] J. Perrin, Compt. Rend. (Paris) **146**, 967 (1908).
[4] I. Nordlund, Z. Phys. Chem. **87**, 40 (1914).
[5] M. von Smoluchowski, Physikal. Zeitschr. **17**, 557 (1916).
[6] D. Wirtz, Annu. Rev. Biophys. **38**, 301 (2009).
[7] K. Nørregaard, R. Metzler, C. Ritter, K. Berg-Sørensen, and L. Oddershede, Chem. Rev. **117**, 4342 (2017).
[8] S. Scott *et al.*, Phys. Chem. Chem. Phys. **25**, 1513 (2023).
[9] X. S. Xie, P. J. Choi, G.-W. Li, N. K. Lee, and G. Lia, Annu. Rev. Biophys. **37**, 417 (2008).
[10] R. Metzler, E. Barkai, and J. Klafter, "Anomalous diffusion and relaxation close to thermal equilibrium: A fractional Fokker-Planck equation approach, Phys. Rev. Lett. **82**, 3563 (1999).
[11] R. Metzler and J. Klafter, "The random walk's guide to anomalous diffusion: A fractional dynamics approach," Phys. Rep. **339**, 1 (2000).
[12] F. Höfling and T. Franosch, Rep. Progr. Phys. **76**, 046602 (2013).
[13] E. Barkai, Y. Garini, and R. Metzler, Phys. Today **65**, 29 (2012).
[14] N. van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1981).
[15] P. Lévy, *Processus Stochastiques et Mouvement Brownien* (Gauthier-Villars, Paris, 1948).
[16] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
[17] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, Phys. Chem. Chem. Phys. **16**, 24128 (2014).
[18] I. M. Sokolov, Soft Matter **8**, 9043 (2012).
[19] Y. Meroz and I. M. Sokolov, Phys. Rep. **573**, 1 (2015).
[20] V. Tejedor, O. Benichou, R. Voituriez, R. Jungmann, F. Simmel, C. Selhuber-Unkel, L. Oddershede, and R. Metzler, Biophys. J. **98**, 1364 (2010).
[21] E. Aghion, P. G. Meyer, V. Adlakha, H. Kantz, and K. E. Bassler, New J. Phys. **23**, 2 (2021).
[22] O. Vilk, E. Aghion, T. Avgar, C. Beta, O. Nagel, A. Sabri, R. Sarfati, D. K. Schwartz, M. Weiss, D. Krapf, R. Nathan, R. Metzler, and M. Assaf, Phys. Rev. Res. **4**, 033055 (2022).
[23] M. Magdziarz and J. Klafter, Phys. Rev. E **82**, 011129 (2010).
[24] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter, Phys. Rev. Lett. **103**, 180602 (2009).
[25] J. A. Sharp, A. P. Browning, K. Burrage, and M. J. Simpson, J. Roy. Soc. Interface **19**, 20210940 (2021).
[26] J. Krog, L. H. Jacobsen, F. W. Lund, D. Wüstner, and M. A. Lomholt, J. Stat. Mech. **2018**, 093501 (2018).
[27] S. Thapa, M. A. Lomholt, J. Krog, A. G. Cherstvy, and R. Metzler, Phys. Chem. Chem. Phys. **20**, 29018 (2018).
[28] S. Thapa, S. Park, Y. Kim, J. H. Jeon, R. Metzler, and M. A. Lomholt, J. Phys. A **55**, 19 (2022).
[29] G. A. Gottwald and I. Melbourne, J. Stat. Mech. **2016**, 123205 (2016).
[30] C. Hein, P. Imkeller, and I. Pavlyukevich, *Recent Development in Stochastic Dynamics and Stochastic Analysis* (World Scientific, 2010), pp. 161–175.
[31] M. Magdziarz, J. Ślęzak, and J. Wójcik, J. Phys. A: Math. Theor. **46**, 325003 (2013).
[32] G. Muñoz-Gil, M. A. Garcia-March, C. Manzo, J. D. Martín-Guerrero, and M. Lewenstein, New J. Phys. **22**, 013010 (2020).
[33] S. Bo, F. Schmidt, R. Eichhorn, and G. Volpe, Phys. Rev. E **100**, 1 (2019).
[34] G. Muñoz-Gil, G. Volpe, M. A. Garcia-March, E. Aghion, A. Argun, C. B. Hong, T. Bland, S. Bo, J. A. Conejero, N. Firbas, Ò. Garibo i Orts, A. Gentili, Z. Huang, J.-H. Jeon, H. Kabbech, Y. Kim, P. Kowalek, D. Krapf, H. Loch-Olszewska, M. A. Lomholt, J.-B. Masson, P. G. Meyer, S. Park, B. Requena, I. Smal, T. Song, J. Szwabiński, S. Thapa, H. Verdier, G. Volpe, A. Widera, M. Lewenstein, R. Metzler, and C. Manzo, Nat. Commun. **12**, 6253 (2021).
[35] H. Seckler and R. Metzler, Nat. Commun. **13**, 6717 (2022); H. Seckler, J. Szwabinski, and R. Metzler, J. Phys. Chem. Lett. **14**, 7910 (2023).
[36] P. Kowalek, H. Loch-Olszewska, Ł. Łaszczuk, J. Opała, and J. Szwabiński, J. Phys. A **55**, 24 (2022).
[37] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).
[38] A. V. Weigel, B. Simon, M. M. Tamkun, and D. Krapf, Proc. Natl. Acad. Sci. U.S.A. **108**, 6438 (2011).
[39] I. Y. Wong, M. L. Gardel, D. R. Reichman, E. R. Weeks, M. T. Valentine, A. R. Bausch, and D. A. Weitz, *ibid.* **92**, 178101 (2004).
[40] M. Levin, G. Bel, and Y. Roichman, J. Chem. Phys. **154**, 144901 (2021).
[41] M. S. Song, H. C. Moon, J.-H. Jeon, and H. Y. Park, Nat. Commun. **9**, 344 (2018).

[42]Q. Xu, L. Feng, R. Sha, N. C. Seeman, and P. M. Chaikin, Phys. Rev. Lett. **106**, 228102 (2011).
[43]T. H. Solomon, E. R. Weeks, and H. L. Swinney, Phys. Rev. Lett. **71**, 3975 (1993).
[44]Y. Edery, A. Guadagnini, H. Scher, and B. Berkowitz, Water Resour. Res. **50**, 1490 (2014).
[45]O. Vilk, Y. Orchan, M. Charter, N. Ganot, S. Toledo, R. Nathan, and M. Assaf, Phys. Rev. X **12**, 031005 (2022).
[46]A. Díez Fernandez, P. Charchar, A. G. Cherstvy, R. Metzler, and M. W. Finnis, Phys. Chem. Chem. Phys. **22**, 27955 (2020).
[47]J.-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sørensen, L. Oddershede, and R. Metzler, Phys. Rev. Lett. **106**, 048103 (2011).
[48]J.-H. Jeon, N. Leijnse, L. B. Oddershede, and R. Metzler, New J. Phys. **15**, 045011 (2013).
[49]M. A. Taylor, J. Knittel, and W. P. Bowen, New J. Phys. **15**, 023018 (2013).
[50]T. Roy, K. Szuttor, J. Smiatek, C. Holm, and S. Hardt, Polymers **11**, 488 (2019).
[51]H. Yang, G. Luo, P. Karnchanaphanurach, T.-M. Louie, I. Reich, S. Cova, L. Xun, and X. S. Xie, Science **302**, 262 (2003).
[52]M. Magdziarz, J. Stat. Phys. **135**, 763 (2009).
[53]M. Magdziarz, Stoch. Proc. Appl. **119**, 3238 (2009).
[54]M. Magdziarz, A. Weron, and J. Klafter, Phys. Rev. Lett. **101**, 210601 (2008).
[55]M. Magdziarz, A. Weron, and K. Weron, Phys. Rev. E **75**, 016708 (2007).
[56]W. T. Coffey and Y. P. Kalmykov, *The Langevin Equation* (World Scientific, Singapore, 2012).
[57]A. Chechkin and I. M. Sokolov, Phys. Rev. E **103**, 032133 (2021).
[58]J. Klafter and I. M. Sokolov, *First Steps in Random Walks. From Tools to Applications* (Oxford University Press, Oxford, 2011).
[59]S. Bochner, *Harmonic Analysis and the Theory of Probability* (Berkeley University Press, Berkeley, CA, 1960).
[60]W. Feller, *An Introduction to Probability Theory and its Application* (Wiley, New York, NY, 1970).
[61]H. C. Fogedby, Phys. Rev. E **50**, 1657 (1994).
[62]I. A. Koutrouvelis, J. Am. Stat. Assoc. **75**, 918 (1980).
[63]J. H. McCulloch, Commun. Stat. Simul. **15**, 1109 (1986).
[64]J. P. Nolan, *Levy Processes* (Birkhäuser, 2001), pp. 379–400.
[65]W. H. DuMouchel, Ann. Stat. **1**, 948–957 (1973).
[66]M. Magdziarz, Stoch. Models **26**, 256 (2010).
[67]V. Sposini, D. Krapf, E. Marinari *et al.*, Commun. Phys. **5**, 305 (2022).