Optimal parameters for anomalous-diffusion-exponent estimation from noisy data

Yann Lanoiselée,[1,*] Grzegorz Sikora,[2,†] Aleksandra Grzesiek,[2,‡] Denis S. Grebenkov,[1,§] and Agnieszka Wyłomańska[2,∥]

[1]*Laboratoire de Physique de la Matière Condensée (UMR 7643), CNRS – Ecole Polytechnique, 91128 Palaiseau, France*

[2]*Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wroclaw University of Science and Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland*

(Received 1 August 2018; published 26 December 2018)

The most common way of estimating the anomalous-diffusion exponent from single-particle trajectories consists in a linear fitting of the dependence of the time averaged mean square displacement on the lag time at the log-log scale. However, various measurement noises that are unavoidably present in experimental data can strongly deteriorate the quality of this estimation procedure and bias the estimated exponent. To investigate the impact of noises and to improve the estimation quality, we compare three approaches for estimating the anomalous-diffusion exponent and check their efficiency on fractional Brownian motion corrupted by Gaussian noise. We discuss how the parameters of both the anomalous-diffusion model and the estimation techniques influence the estimated exponent. We show that the conventional linear fitting is the least optimal method for the analysis of noisy data.

DOI: 10.1103/PhysRevE.98.062139

I. INTRODUCTION

Anomalous diffusion processes are widely discussed in the literature, in particular, in the context of single-particle trajectories analysis [1–12]. The anomalous-diffusive behavior is manifested by nonlinear time growth of the mean square displacement (MSD), $\langle X^2(n)\rangle \simeq 2D_\beta(n\delta)^\beta$, where n is the integer lag time, δ the frame duration, β is the anomalous-diffusion exponent, D_β is the generalized diffusion coefficient (in units $\mathrm{m}^2/\mathrm{s}^\beta$), and $\langle . \rangle$ denotes the ensemble average (EA) over the probability distribution of $X(n)$. Depending on the β parameter one can distinguish between subdiffusive ($\beta < 1$), diffusive ($\beta = 1$), and superdiffusive ($\beta > 1$) behavior [2,13–20]. However, due to a limited number of trajectories in many experiments, the EA MSD needs to be replaced by the time average (TA) MSD calculated from a single trajectory. For a vector of observations $X(1), X(2), \dots, X(N)$ of length N, the TAMSD at the lag time $n\delta$ is defined as

$$M_N(n) = \frac{1}{N-n}\sum_{k=1}^{N-n}[X(k+n) - X(k)]^2. \tag{1}$$

For an ergodic process with stationary increments, TAMSD converges to EAMSD in the limit $N \to \infty$, $M_{N\to\infty}(n) = \langle X^2(n)\rangle$, i.e., the distribution of TAMSD converges to a Dirac delta function centered on the value of EAMSD. Consequently, for $1 \leqslant n \ll N$ the mean TAMSD scales as

$$\langle M_N(n)\rangle \simeq 2D_\beta(n\delta)^\beta. \tag{2}$$

The TAMSD is one of the classical tools used for estimation of the anomalous-diffusion exponent β. The procedure of estimation is simple: the TAMSD is plotted versus the lag time n at the log-log scale and the estimated β parameter is the slope of the expected straight line, fitted by using the least squares method [8,11].

The classical pure anomalous-diffusion models include fractional Brownian motion (fBm) [21,22], fractional Lévy stable motion [23], and continuous-time random walk [24,25]. In this paper, we focus on the fBm that is a non-Markovian generalization of Brownian motion and one of the most fundamental models of stochastic motion. Specifically, it is the only self-similar Gaussian process with stationary increments. The fBm can also be related to generalized Langevin processes with power law decaying friction kernels, an attractive framework for many physical systems [26–29].

One of the main statistical challenges in the experimental data analysis is the proper model recognition and the precise estimation of the best model parameters. In this paper, we focus on the estimation of the parameters of noisy anomalous diffusion in which "pure" (i.e., noiseless) fBm is progressively corrupted by Gaussian white noise. We propose two alternative approaches for anomalous-diffusion exponent estimation and compare them to the common linear fitting on simulated data. Moreover, we discuss how the parameters of the considered model influence the estimation results. A similar problem was discussed in [30–32] in the case of ordinary Brownian motion.

The rest of the paper is organized as follows: in the next section we formulate the problem. In Sec. III we propose and compare three approaches for anomalous parameters estimation. In Sec. IV we check the efficiency of the proposed

*yann.lanoiselee@polytechnique.edu
†grzegorz.sikora@pwr.edu.pl
‡aleksandra.grzesiek@pwr.edu.pl
§denis.grebenkov@polytechnique.edu
∥agnieszka.wylomanska@pwr.edu.pl

estimation techniques on simulated data. The last section concludes the paper.

II. PROBLEM FORMULATION

The classical approach for estimating the parameters D_β and β from Eq. (2) for pure anomalous diffusion consists in a linear fitting. More precisely, the TAMSD is first calculated from a vector of positions according to Eq. (1). Then, taking the logarithm of both sides of the formula (2) one can estimate the parameters using the classical least squares method in linear regression. The details of this approach are presented for instance in [33]. Usually, the parameters are estimated by using integer lag times $n \in [1, n_{\max}]$. The accuracy of the estimation decreases as $n_{\max}$ gets larger. In spite of its numerous applications in practice, the approach has some drawbacks.

Even if the experimental data exhibit a behavior adequate to some theoretical model of anomalous diffusion, it is always disturbed by measurement noise [11]

$$\tilde{X}(n) = X(n) + \xi(n), \tag{3}$$

where $X(n)$ is a pure anomalous-diffusion process with D_β and β parameters, and $\xi(n)$ denotes noise, which is assumed to be independent of $X(n)$ and normally distributed with mean zero and variance σ^2. The EAMSD reads then

$$\langle \tilde{X}^2(n)\rangle = 2D_\beta(n\delta)^\beta + \sigma^2. \tag{4}$$

Figure 1 shows that the noise term σ^2 makes the TAMSD (as well as the EAMSD) flat, until the contribution from anomalous diffusion becomes dominant: $2D_\beta(n\delta)^\beta \gg \sigma^2$. To avoid such a noise dominated region, it is natural to perform the fitting from $n_{\min}$ to $n_{\max}$, with some $n_{\min} > 1$. This is the first problem discussed in this paper. We check by simulations how the noise term σ^2 influences the estimation results and how the selection of the $n_{\min}$ and $n_{\max}$ in the classical estimation algorithm can change the estimation efficiency.

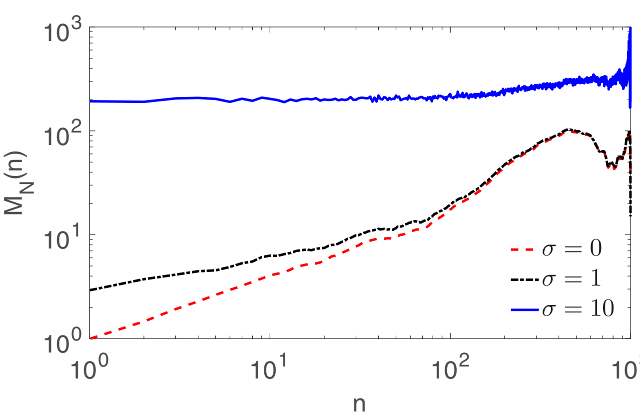

FIG. 1. The TAMSD of the fractional Brownian motion with $\beta = 0.6$ and $D_\beta = 1/2$, corrupted by white noise with three different values of the standard deviation $\sigma = \{0, 1, 10\}$. Arbitrary units are used.

The second considered problem can be formulated as follows: even if the fitting is performed over the window from $n_{\min}$ to $n_{\max}$, it is not enough to get efficient estimators of D_β and β from the *linear* fit because departures from the linear shape of MSD is increasing with σ (see Fig. 1). In this paper, we propose two alternative approaches for estimating β and D_β via a nonlinear fitting. Both approaches assume the toy model defined in Eq. (3), i.e., the noise term is taken into consideration. Then, we compare the estimation results for the proposed methods with the classical method where the model is just anomalous-diffusive process $X(n)$. Moreover, we check also the influence of σ, $n_{\min}$, and $n_{\max}$ on the estimation results for the different approaches. While the simulations will be presented for the selected anomalous-diffusion model $\tilde{X}(n)$ in (3), namely fBm, the problem is relevant for any ergodic process showing anomalous diffusion.

III. ANOMALOUS-DIFFUSION-EXPONENT ESTIMATION

In this section, we describe three approaches used to estimate the anomalous-diffusion exponent β. Although we focus on the anomalous-diffusion-exponent estimation, the presented approaches are also useful for estimating the diffusion coefficient D_β.

A. Approach I

The classical approach I consists in taking the logarithm of both sides of Eq. (4) and expanding the right-hand side to the first order of with respect to the small parameter $\frac{\sigma^2}{2D_\beta(n\delta)^\beta} \ll 1$. The relation becomes

$$\ln(\langle \tilde{X}^2(n)\rangle) = \ln(2D_\beta) + \beta\ln(n\delta) + \frac{\sigma^2}{2D_\beta(n\delta)^\beta} + O[(n\delta)^{-2\beta}], \tag{5}$$

then with the variable $u_\delta = \ln(n\delta)$ we get

$$\ln(\langle \tilde{X}^2(u_\delta)\rangle) = \ln(2D_\beta) + \beta u_\delta + \frac{\sigma^2}{2D_\beta}e^{-\beta u_\delta} + O(e^{-2\beta u_\delta}). \tag{6}$$

There is a linear dependence of $\ln(\langle \tilde{X}^2(u_\delta)\rangle)$ on u_δ with an exponentially decaying (in log-log coordinates) correction related to the noise term σ^2. In the limit of small σ^2 (respectively large u_δ), the noise effect disappears and the estimation is reduced to a linear regression. In this approach, we estimate the β exponent in a similar way as for pure fBm. The details of this approach one can find for instance in [33] therefore we only sketch the idea. In order to estimate the anomalous-diffusion exponent β for pure fBm, one needs to calculate TAMSD from the trajectory $X(1), X(2), \ldots, X(N)$ of length N at the points $n_{\min}, \ldots, n_{\max}$ and then fit $\ln[M_N(n)]$ by a linear function of a form $\alpha + \beta u_\delta$ for $u_\delta = \ln(n_{\min}\delta), \ldots, \ln(n_{\max}\delta)$. The exact form of the estimator $\hat{\beta}$ from the least squares method to linear fitting reads

$$\hat{\beta} = \frac{(\Delta n + 1)\sum_{n=n_{\min}}^{n_{\max}} \ln(n)\ln[M_N(n)] - \sum_{n=n_{\min}}^{n_{\max}} \ln(n)\left(\sum_{n=n_{\min}}^{n_{\max}} \ln[M_N(n)]\right)}{(\Delta n + 1)\sum_{n=n_{\min}}^{n_{\max}} \ln^2(n) - \left(\sum_{n=n_{\min}}^{n_{\max}} \ln(n)\right)^2}, \quad (7)$$

where $\Delta n = n_{\max} - n_{\min}$. In approach I, for small values of σ, we can neglect the noise term and use the estimator for pure fBm. A similar calculation to the one in [33] shows that the estimator of Eq. (7) is consistent and asymptotically unbiased. Another choice of the estimator is possible in the case $n_{\min} = 1$ [33].

Due to the finite trajectory length $N < \infty$, the selection of $n_{\max}$ too close to N would result in large fluctuations because of the small number of data points contributing to the average. Simulations allow checking the effects of both $n_{\min}$ and $n_{\max}$ on the estimation quality at a given level of noise (see Sec. IV).

B. Approach II

Approach II is one of the alternatives to the classical method presented above. In contrast to approach I, we estimate the β parameter taking into account the noise term $\xi(n)$ in the model (3).

The idea is to perform a fitting using the exact formula of the EAMSD in Eq. (4),

$$M_N(n) \sim 2D_\beta(n\delta)^\beta + \sigma^2, \quad (8)$$

where $\sim$ means the equality in the expected value. Similarly to approach I, we calculate $M_N(n)$ for the lag times $n_{\min}, \dots, n_{\max}$. However, in this case, the fitting function $\hat{M}(n)$ is defined as follows:

$$\hat{M}(n) = 2\hat{D}_\beta(n\delta)^{\hat{\beta}} + \hat{\sigma}^2, \quad (9)$$

where $\hat{\beta}$, $\hat{D}_\beta$, and $\hat{\sigma}$ are three fitting parameters. In order to estimate these parameters, one has to the minimize the error function $\Upsilon = \sum_{n=n_{\min}}^{n_{\max}}[M_N(n) - \hat{M}(n)]^2$. The minimum is found when the gradient of Υ with respect to the fitting parameters is equal to zero and the error is the lowest. The function $\hat{M}(n)$ has a nonlinear dependence on $n\delta$ making it dependent on the fitting parameters themselves. In this case, the explicit expression of an estimator $\hat{\beta}$ is not accessible and has to be calculated numerically. The nonlinear fitting methodology is described in the Appendix.

C. Approach III

In the last approach, we make a simple transformation of the TAMSD to reduce the number of fitting parameters. We take into account the fact that the fitting starts at the point $n_{\min}$ by subtracting from $M_N(n)$, $n \in [n_{\min}, \dots, \tau_{\max}]$ the term $M_N(n_{\min})$ so that the formula (8) becomes

$$M_N(n) - M_N(n_{\min}) \sim 2D_\beta\delta^\beta\left(n^\beta - n_{\min}^\beta\right). \quad (10)$$

This transformation removes the noise term, at the cost of a more complex dependence on β. Thus, in approach III, we calculate $M_N(n) - M_N(n_{\min})$ at lag times $n_{\min}, \dots, n_{\max}$ and then fit them by the function

$$\hat{M}(n) = 2\hat{D}_\beta\delta^{\hat{\beta}}\left(n^{\hat{\beta}} - n_{\min}^{\hat{\beta}}\right). \quad (11)$$

Similar to approach II, the function $\hat{M}(n)$ is nonlinear, so an iterative numerical procedure is necessary for error minimization.

It is worthwhile to highlight that only approach I provides unique and explicit estimators of the parameters whereas estimators obtained by nonlinear methods (approaches II and III) depend on different assumptions (such as the choice of the optimization method, the starting point of iterations, the stopping criteria, etc.) and the estimators are not given in the explicit form. Therefore the theoretical properties of the estimators obtained using nonlinear fitting were not well studied. This may be the reason why the linear fitting is the most popular.

We should mention that in approaches I and III we estimate two parameters while in approach II we estimate three. With one extra fitting parameter, one may expect approach II to be the best one. However, as we indicate in the simulation study, this is not true in general. We will show that approach III is the best one for some parameters.

IV. OPTIMAL PARAMETERS FOR β ESTIMATION

In this section, we investigate the β parameter estimation by Monte Carlo simulations. We simulate single-particle trajectories by the model in Eq. (3), where $X(n)$ is a fBm with a given β and $D_\beta = 1/2$ and $\delta = 1$. We compare three approaches and check which one is the most efficient for the β parameter estimation. Although similar techniques can be used for estimating the diffusion parameter D_β, we do not consider this option.

We test the three fitting approaches on three representative cases of the fBm: (i) sub-diffusive antipersistent motion with $\beta = 0.6$, (ii) diffusive Markovian motion with $\beta = 1$, and (iii) superdiffusive persistent motion with $\beta = 1.4$, all with the the same generalized diffusion coefficient $D_\beta = 1/2$. In order to be closer to experimental conditions, three levels of noise are tested, classified from none to high noise with the standard deviation taking values $\sigma = \{0, 1, 10\}$. In every case the fitting is performed using the TAMSD calculated from a single trajectory. The efficiency of each method is measured in terms of the accuracy $A(\%)$ of the estimation, which is the percentage of the estimated exponent which falls in the range $\beta - 0.2 < \hat{\beta} < \beta + 0.2$. The quantity A is calculated from the estimated distribution of $\hat{\beta}$ obtained from $M = 1000$ realizations.

First, we determine which approach from Sec. III is the most accurate for each couple $n_{\min}, n_{\max}$ and level of noise σ. For a better comparison, it is convenient to replace $n_{\max}$ by the time window width

$$\Delta n = n_{\max} - n_{\min}.$$

Figure 2 shows the result for long trajectories ($N = 1000$). Even in the case without noise (first row), the classical

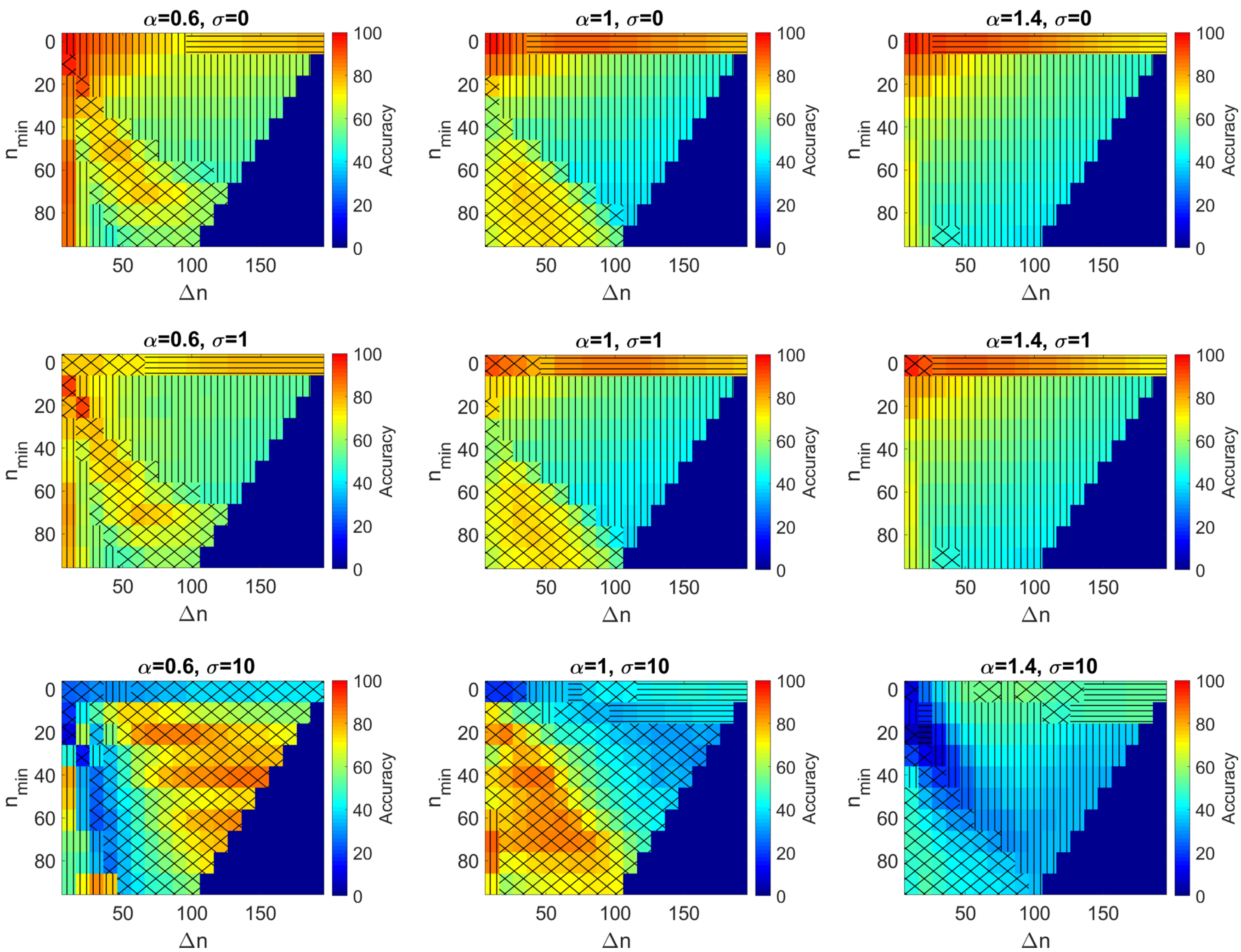

FIG. 2. Percentage of estimations in the range $\beta - 0.2 < \hat{\beta} < \beta + 0.2$ for the best approach among the three tested. From left to right $\beta = \{0.6, 1, 1.4\}$, from top to bottom the standard deviation of white noise in the range $\sigma = \{0, 1, 10\}$ with $D_\beta = 1/2$ and $N = 1000$. Each pair $[n_{\min}, \Delta n]$, where $\Delta n = n_{\max} - n_{\min}$ with $n_{\min} \in [1, 11, \ldots, 191]$ and $n_{\max} \in [11, \ldots, 201]$, is considered. Dashed patterns highlight the most accurate approach with horizontal, vertical, and crossed lines respectively associated to approaches I, II, and III; colors highlight the corresponding best accuracy score from dark blue (accuracy = 0) to light red (accuracy = 100%); dark blue triangular regions ($n_{\max} > N/5$) were not tested.

approach I performs rather poorly as it can outperform other approaches only in the case where $n_{\min} = 1$ and $\Delta n > 110, 40, 30$ for $\beta = 0.6, 1, 1.4$, respectively. Approach III gives satisfactory results in the lower triangle where roughly $n_{\min} > \Delta n$, from anti-persistent to diffusive motion ($\beta \leqslant 1$) while it is approach II that is better for $\beta > 1$ in this region. In all other situations, approach II is the best.

Conclusions drawn from the first row of Fig. 2 are also applicable to small noise (second row, $\sigma = 1$). This is understandable as the TAMSD is affected by this level of noise only around $n \approx 1$ (see Fig. 1). In contrast, for large noise ($\sigma = 10$, third row), the TAMSD is affected by the noise on a longer time range, and the results are different: (i) for $\beta = 0.6$ the best estimation is achieved by increasing both $n_{\min}$ and Δn; (ii) for $\beta = 1$ the best score is achieved by increasing $n_{\min}$ but keeping Δn not too high; (iii) for $\beta = 1.4$, the quality of the estimations is poor in every case. The last result is counterintuitive as the impact of noise is reduced as β increases [see Eq. (5)] but this reduction is not enough at small n, for instance, noise still presents one-third of the MSD $[2D_\beta(n_c\delta)^{1.4}/\sigma^2 = 2]$ at $n_c \approx 44$. At longer lag time n (but still with $n < N/20$), the positive autocorrelations slow down the self-averaging so the distribution of the TAMSD is wider [33]. The combination of noise and the wider distribution of TAMSD prevents obtaining a correct estimation of the superdiffusive exponent with trajectories of length $N = 1000$.

Figure 3 shows the results for trajectories of length $N = 100$. Such short trajectories are often encountered in biological applications. In this case, the distribution of the TAMSD is wider, thus the estimation is more difficult. When there is no noise, it is still possible to achieve a good estimation for small $n_{\min}$ and Δn while the presence of even mild noise makes the estimation unreliable. In the regime of strong noise, the estimation is so bad that it would make no difference to uniformly pick an exponent in the range $\beta \in [0, 2]$. Thus, for short trajectories the TAMSD is not appropriate.

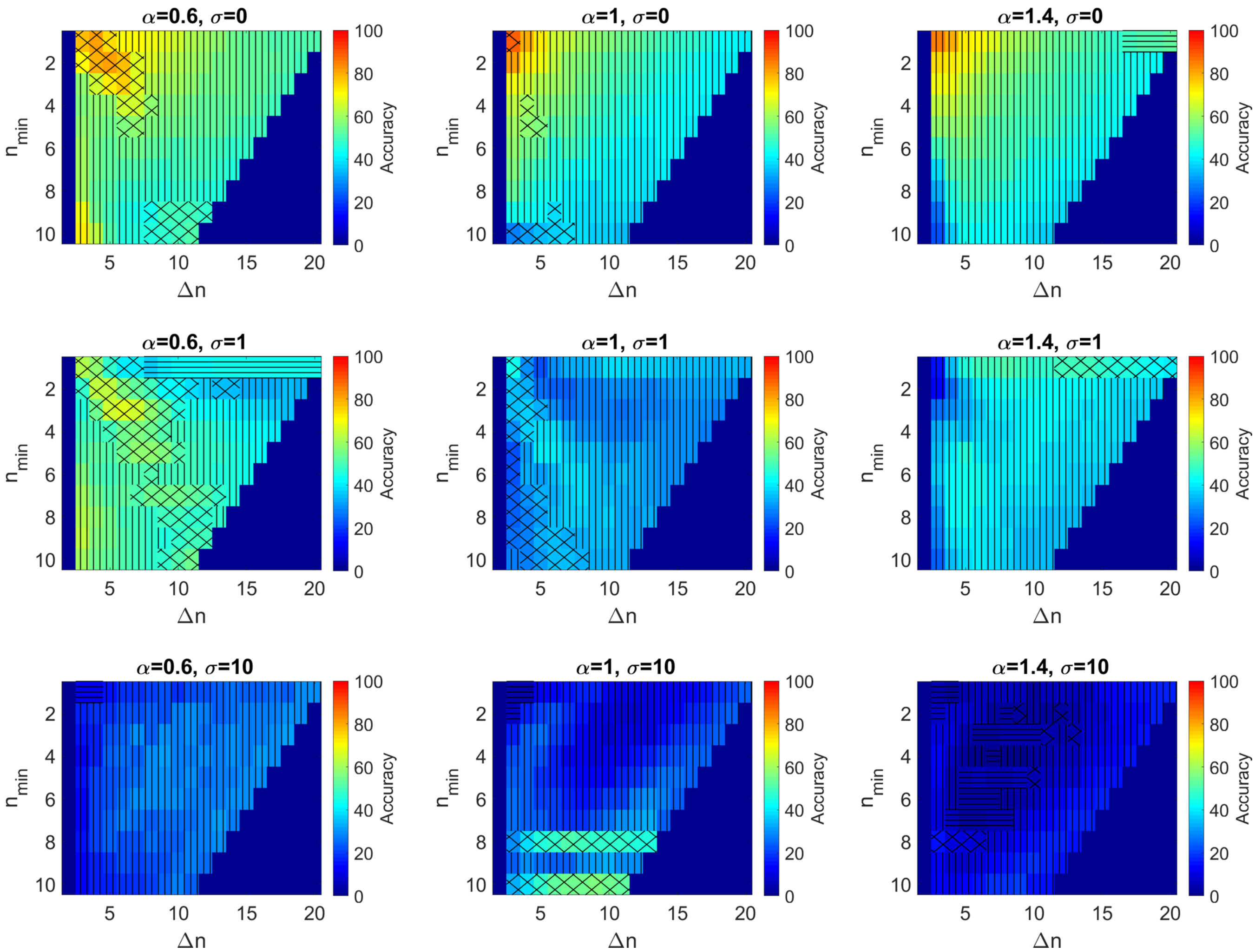

FIG. 3. Percentage of estimations in the range $\beta - 0.2 < \hat{\beta} < \beta + 0.2$ (accuracy) for the best approach among the three tested. From left to right $\beta = \{0.6, 1, 1.4\}$, from top to bottom the standard deviation of white noise in the range $\sigma = \{0, 1, 10\}$ with $D_\beta = 1/2$ and $N = 100$. Each pair $[n_{\min}, \Delta n]$, where $\Delta n = n_{\max} - n_{\min}$ with $n_{\min} \in [1, 2, \dots, 19]$ and $n_{\max} \in [3, \dots, 20]$, is considered. Dashed patterns highlight the most accurate approach with horizontal, vertical, and crossed lines respectively associated to approaches I, II, and III; colors highlight the corresponding best accuracy score from dark blue (accuracy = 0) to light red (accuracy = 100%); dark blue triangular regions ($n_{\max} > N/5$) were not tested.

What are the best approaches and the optimal estimation parameters that maximize the accuracy? Looking at Figs. 2 and 3, one can see that there is neither "the best approach" nor the unique optimal values for $n_{\min}$ and $n_{\max}$. In turn, we can determine the best approach and the optimal parameters for each combination of the trajectory length N, exponent β, and level of noise σ. The results are gathered in Table I. Strikingly, the commonly used approach I is nowhere the

TABLE I. Summary of the best approaches for two lengths of trajectory, $N = 100$ and $N = 1000$, with three level of noise $\sigma = 0, 1, 10$, for three cases of the exponent $\beta = 0.6, 1, 1.4$. The table shows the best approach with the corresponding parameters $n_{\min}$ and $\Delta n = n_{\max} - n_{\min}$ with $n_{\min} \in [1, N/100 + 1, \dots, 19N/100 + 1]$ and $n_{\max} \in [N/100 + 1, \dots, 20N/100 + 1]$, and the corresponding accuracy $A(\%)$.

		$\beta = 0.6$				$\beta = 1$				$\beta = 1.4$			
N	σ	Approach	$n_{\min}$	Δn	$A(\%)$	Approach	$n_{\min}$	Δn	$A(\%)$	App.	$n_{\min}$	Δn	$A(\%)$
	0	II	1	10	98	II	1	10	97	II	1	10	98
1000	1	III	11	10	92	III	1	10	89	III	1	10	86
	10	III	41	150	86	III	71	10	88	III	1	90	55
	0	III	2	4	82	III	1	2	88	II	1	2	86
100	1	III	3	5	66	III	1	2	44	III	1	7	50
	10	II	7	12	32	III	10	8	56	II	10	8	21

best. For $N = 1000$, when there is no noise, the best choice is approach II in every case, with $n_{\min} = 1$ and $\Delta n = 10$ (note that the optimal value for Δn can be even smaller, as the value 10 came from the discrete exploration of the parameters' space). In the presence of noise, approach III is the best, with progressively increasing $n_{\min}$ and Δn as the noise level increases. For $N = 100$, the noise impacts significantly the accuracy. Even for $\sigma = 1$, accuracy drops to $\approx$50% emphasizing that precise estimation based on such a short trajectory requires a very good experimental signal to noise ratio.

From the statistical point of view, the linear fitting (approach I) is correct if we consider the pure fBm. This point is discussed for instance in [33] where the theoretical properties of the anomalous-diffusive parameter estimator are discussed for the fBm. However, as we indicated in this paper, approach I is not appropriate when the model under consideration is fBm disturbed by the additional noise. Therefore, it is reasonable to propose the alternative approaches which are more effective in the considered cases.

V. CONCLUSIONS

We studied the problem of estimation of the anomalous-diffusion exponent for processes in which a pure anomalous-diffusion model is corrupted by independent noise. We propose two alternative approaches that can be used for estimating the anomalous-diffusion exponent. We indicate their advantages and limitations and check their efficiency by Monte Carlo simulations. We show that the classical estimation fails in every case. Our study shows that none of the approaches is the best for all cases and we cannot identify the best approach for the considered problem. We indicate how the model parameters, as well as parameters of the estimation techniques, may influence the results. The presented discussion and results can be useful for a more reliable statistical analysis of single-particle trajectories in cell biology and other fields. The presented study can be extended to other types of perturbations, e.g., other types of noise ($1/f$ noise, colored noise, etc.) or even another process which may contaminate the pure anomalous-diffusive motion. This study may be interesting for experimentalists from different areas and will be considered in our future research.

ACKNOWLEDGMENTS

D.S.G. acknowledges the support under Grant No. ANR-13-JSV5-0006-01 of the French National Research Agency. A.G. and A.W. would like to acknowledge the support of NCN OPUS Grant No. UMO-2016/21/B/ST1/00929, "Anomalous diffusion processes and their applications in real data modeling."

APPENDIX: NONLINEAR FITTING

Nonlinear fitting in approaches II and III consists in finding the parameters $\hat{D}_\beta$, $\hat{\beta}$, and possibly $\hat{\sigma}$ [see (9) and (11)] that minimize the sum of squared errors. For a nonlinear problem, there is no explicit expression for the estimator, and one has to perform the minimization procedure by numerical methods. In this article, we use a trust region method [34,35] to perform nonlinear least square fitting with MATLAB. In order to reduce the calculation time and avoid nonphysical values of parameters, some constraints are imposed on the parameters. All parameters are positive, the exponent $\hat{\beta}$ cannot exceed the ballistic regime, $\hat{\beta} \leqslant 2$, and the noise is necessarily smaller than the TAMSD at $n = 1$ so $\hat{\sigma}^2 \in [0, M_N(1)]$. There is no evident upper bound for the generalized diffusion coefficient so we assume $\hat{D}_\beta \in [0, \infty)$. For the minimization procedure, a crucial point is the choice of the stopping criterion ϵ. The iteration is interrupted when the relative change in the error function $\frac{|\Upsilon_{i+1}-\Upsilon_i|}{1+|\Upsilon_i|} < \epsilon$. Choosing ϵ too large forces the algorithm to stop before convergence, resulting in poor estimation. Conversely, taking ϵ too small makes the minimization longer because the random nature of the TAMSD imposes a lower limit on the possible precision obtained. In our case, M_N does not follow exactly the theoretical MSD as the TAMSD, evaluated over a single realization of a stochastic process of finite length N, is itself random. Thus one cannot expect a perfect match between $\hat{M}$ and M_N, in other words, there is a distribution of the minima for the function Υ which is determined by the fluctuations of the TAMSD, depending on $n_{\min}$, $n_{\max}$, N, and β, moreover, the presence of a white noise increases uncertainty and thus increases the optimal ϵ. The best ϵ is the largest possible value for which the estimation remains unchanged. In this article we chose $\epsilon = 0.01$ as a good compromise between speed and precision. Finally we initiated the fitting parameters with the parameters of Brownian motion with $\hat{\beta}_0 = 1$, $\hat{D}_{\beta_0} = M_N(1)$, and the noise $\hat{\sigma}_0^2 = M_N(1)/2$. The noise level can also be initiated according to the expected experimental measurement error.

[1] D. Arcizet, B. Meier, E. Sackmann, J. O. Rädler, and D. Heinrich, Phys. Rev. Lett. **101**, 248103 (2008).
[2] R. Metzler, V. Tejedor, J.-H. Jeon, Y. He, W. H. Deng, S. Burov, and E. Barkai, Acta Phys. Pol. B **40**, 1315 (2009).
[3] V. Tejedor, O. Bénichou, R. Voituriez, R. Jungmann, F. Simmel, C. Selhuber-Unkel, L. B. Oddershede, and R. Metzler, Biophys. J. **98**, 1364 (2010).
[4] M. Magdziarz and A. Weron, Phys. Rev. E **84**, 051138 (2011).
[5] D. S. Grebenkov, Phys. Rev. E **84**, 031124 (2011).
[6] A. A. Andreanov and D. S. Grebenkov, J. Stat. Mech. (2012) P07001.
[7] D. S. Grebenkov, Phys. Rev. E **88**, 032140 (2013).
[8] N. Gal, D. Lechtman-Goldstein, and D. Weihs, Rheol. Acta **52**, 425 (2013).
[9] Y. Meroz, I. M. Sokolov, and J. Klafter, Phys. Rev. Lett. **110**, 090601 (2013).
[10] S. Türkcan and J.-B. Masson, PloS One **8**, 0082799 (2013).

[11] E. Kepten, A. Weron, G. Sikora, K. Burnecki, and Y. Garini, PLOS One **10**, e0117722 (2015).
[12] Y. Lanoiselée and D. S. Grebenkov, Phys. Rev. E **93**, 052146 (2016).
[13] I. M. Tolić-Nørrelykke, E.-L. Munteanu, G. Thon, L. Oddershede, and K. Berg-Sørensen, Phys. Rev. Lett. **93**, 078102 (2004).
[14] I. Golding and E. C. Cox, Phys. Rev. Lett. **96**, 098102 (2006).
[15] C. Wilhelm, Phys. Rev. Lett. **101**, 028101 (2008).
[16] J. Szymanski and M. Weiss, Phys. Rev. Lett. **103**, 038102 (2009).
[17] E. Sackmann, F. Keber, and D. Heinrich, Annu. Rev. Condens. Matter Phys. **1**, 257 (2010).
[18] 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).
[19] E. Bertseva, D. S. Grebenkov, P. Schmidhauser, S. Gribkova, S. Jeney, and L. Forró, Eur. Phys. J. E **35**, 63 (2012).
[20] P. C. Bressloff and J. M. Newby, Rev. Mod. Phys. **85**, 135 (2013).
[21] J. Beran, *Statistics for Long-Memory Processes* (Chapman & Hall, New York, 1994).
[22] B. B. Mandelbrot and J. W. Van Ness, SIAM Rev. **10**, 422 (1968).
[23] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes* (Chapman & Hall, London, 1994).
[24] R. Metzler and J. Klafter, Phys. Rep. **339**, 1 (2000).
[25] R. Metzler and J. Klafter, J. Phys. A **37**, R161 (2004).
[26] J. Klafter, S. C. Lim, and R. Metzler, *Fractional Dynamics. Recent Advances* (World Scientific, Singapore, 2011).
[27] K. Burnecki and A. Weron, Phys. Rev. E **82**, 021130 (2010).
[28] D. S. Grebenkov, Phys. Rev. E **83**, 061117 (2011).
[29] G. Sikora, K. Burnecki, and A. Wyłomańska, Phys. Rev. E **95**, 032110 (2017).
[30] X. Michalet, Phys. Rev. E **82**, 041914 (2010).
[31] X. Michalet and A. J. Berglund, Phys. Rev. E **85**, 061916 (2012).
[32] D. Ernst and J. Köhler, Phys. Chem. Chem. Phys. **15**, 3429 (2013).
[33] G. Sikora, M. Teuerle, A. Wyłomańska, and D. Grebenkov, Phys. Rev. E **96**, 022132 (2017).
[34] T. F. Coleman and Y. Li, Math. Prog. **67**, 189 (1994).
[35] T. F. Coleman and Y. Li, SIAM J. Optim. **6**, 418 (1996).