



# CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Identifying diffusive motions in single-particle trajectories on the plasma membrane via fractional time-series models

Krzysztof Burnecki, Grzegorz Sikora, Aleksander Weron, Michael M. Tamkun, and Diego Krapf

Phys. Rev. E **99**, 012101 — Published 2 January 2019

DOI: [10.1103/PhysRevE.99.012101](https://doi.org/10.1103/PhysRevE.99.012101)

# Identifying diffusive motions in single-particle trajectories on the plasma membrane via fractional time series models

Krzysztof Burnecki,<sup>1,\*</sup> Grzegorz Sikora,<sup>1</sup> Aleksander Weron,<sup>1</sup> Michael M. Tamkun,<sup>2</sup> and Diego Krapf<sup>3,4</sup>

<sup>1</sup>*Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center,  
Wrocław University of Science and Technology, 50-370 Wrocław, Poland*

<sup>2</sup>*Department of Biomedical Sciences, Colorado State University, Fort Collins, CO 80523, USA*

<sup>3</sup>*Department of Electrical and Computer Engineering,  
Colorado State University, Fort Collins, CO 80523, USA*

<sup>4</sup>*School of Biomedical Engineering, Colorado State University, Fort Collins, CO 80523, USA*

(Dated: November 30, 2018)

In this paper we show that an autoregressive fractionally integrated moving average (ARFIMA) time series model can identify two types of motion of membrane proteins on the surface of mammalian cells. Specifically we analyze the motion of the voltage-gated sodium channel Nav1.6 and beta 2 adrenergic receptors. We find that the autoregressive (AR) part models well the confined dynamics whereas the fractionally integrated moving average (FIMA) model describes the non-confined periods of the trajectories. Since the Ornstein-Uhlenbeck process is a continuous counterpart of the AR model, we are also able to calculate its physical parameters and show their biological relevance. The fitted FIMA and AR parameters show marked differences in the dynamics of the two studied molecules.

PACS numbers: 02.50.Ey, 87.15.K-, 87.15.Vv, 05.40.Jc

## I. INTRODUCTION

The advent of single-molecule techniques over the past two decades has revolutionized molecular biophysics. Amongst these techniques, single-particle tracking (SPT) has emerged as a powerful approach to study a variety of dynamic processes [1–4]. Individual trajectories have been obtained for diverse biological systems, including measurements in cell membranes [5–12], the cytoplasm [13–17], and the cell nucleus [18–22]. The dynamics of molecules in living cells typically exhibit complex behavior with a high degree of temporal and spatial heterogeneities, due to different factors such as spatial constraints and complex biomolecular interactions [23]. Each mechanism governing the diffusion process has different characteristics, which can give important information regarding the biological system [4, 24–27].

A phenomenon often observed in single-molecule experiments is subdiffusion, with a characteristic sublinear mean-squared displacement (MSD), which largely departs from the classical Brownian motion (BM). Determining the mechanisms underlying anomalous diffusion in complex fluids, e.g., in the cytoplasm of living cells or in controlled *in vitro* experiments [28–30], is a challenging problem. Subdiffusion can be rooted in different physical origins including immobile obstacles, binding, crowding, and heterogeneities [31]. Some of the theoretical models employed to describe subdiffusion are the continuous-time random walk [8], obstructed diffusion [32, 33], fractional Brownian motion (FBM) [29, 34], diffusion in a fractal environment [35, 36], fractional

Lévy stable motion (FLSM) [37], and fractional Langevin equation (FLE) [16, 34, 38].

A discrete-time model that generalizes the above fractional models is the autoregressive fractionally integrated moving average (ARFIMA) process [39, 40]. From the physical point of view, it is a discrete-time analogue of FLE that incorporates the memory parameter  $d$  [41]. Other popular models of subdiffusive dynamics like FBM and FLSM are the limiting cases of the ARFIMA with different noises [42]. ARFIMA exhibits power-law long-time dependencies, similar to FBM and FLSM. Long-time dependencies result in anomalous diffusion, evident in a non-linear MSD [37]. In contrast to FBM and FLSM, ARFIMA can also describe various light- and heavy-tailed distributions and an arbitrary short-time dependence.

ARFIMA was previously suggested as an appropriate model for SPT dynamics for various biological experiments, namely the motion of individual fluorescently labeled mRNA molecules in bacteria [43, 44] and transient anomalous diffusion of telomeres in the nucleus of eukaryotic cells [45]. A special case of the ARFIMA process, namely FIMA was proposed as a useful tool for estimating the anomalous diffusion exponent for particle tracking data with measurement errors [46]. FIMA was also useful in introducing so-called calibration surfaces, which are an effective tool for extracting both the magnitude of the measurement error and the anomalous exponent for autocorrelated processes of various origins [47]. Since ARFIMA models were successful in analyzing data in other fields (econometrics, see 2003 Nobel Prize in Economic Sciences for C.W.J. Granger and R. Engel; finance and engineering [48–50]), many statistical tools and computer packages are readily available, e.g., ITSM [51].

---

\* Corresponding author: krzysztof.burnecki@pwr.edu.pl

In this article, we focus on two types of motion in the plasma membrane, namely free and confined. Transient confinement within membrane domains is a very common feature in the plasma membrane [23]. Here, we focus on confinement within stable clusters. We propose the ARFIMA as a suitable model to characterize both types of dynamics. In Sec. II we discuss basic building blocks of the ARFIMA process: the autoregressive (AR), fractionally integrated (FI) and moving average (MA) parts. We show that the AR process is a discrete counterpart of the continuous-time Ornstein-Uhlenbeck (O-U) process. We also study the relationship between the ARFIMA and mean-squared displacement (MSD). In Sec. III our fitting procedure (presented in detail in Appendix B) is applied to the motion of individual voltage-gated Na channels and beta 2 adrenergic receptors on the surface of live mammalian cells. In the described examples, we show that the increments of the non-confined parts of the trajectories are well-described by the FIMA process whereas the confined parts can be modeled by the AR. The analysis also suggests that the beta 2 receptors appear to be more subdiffusive in the free state and have a lower autoregressive parameter in the confined state. Finally, the distribution of the estimated ARFIMA noise sequence appears to change from Gaussian in the free state to non-Gaussian in the confined state.

## II. ARFIMA MODEL

The autoregressive fractionally integrated moving average (ARFIMA) process is a generalization of the classical autoregressive moving average (ARMA) process that introduces the fractionally integrated (FI) part with the long memory parameter  $d$  [39, 40], see Appendix A for a presentation of general ARFIMA processes.

In this paper we concentrate on a special case of the ARFIMA process, namely on the ARFIMA with AR and MA parts of order 1, which is denoted by ARFIMA(1,  $d$ , 1). The ARFIMA(1,  $d$ , 1) process  $X(t)$  for  $t = 0, \pm 1, \dots$  is defined as a stationary solution of the fractional difference equation [51]

$$(1 - B)^d (X(t) - \phi X(t - 1)) = Z(t) - \psi Z(t - 1), \quad (1)$$

where  $Z(t)$  is the noise (i.i.d. sequence usually Gaussian or in general belonging to the domain of attraction of Lévy stable law),  $|\phi| < 1$  and  $|\psi| < 1$  are autoregressive and moving average parameters, respectively, and  $B$  is the backshift operator:  $BX(t) = X(t - 1)$ . The fractional difference operator  $(1 - B)^d$  is defined by means of the binomial expansion, namely  $(1 - B)^d = \sum_{j=0}^{\infty} b_j(d) B^j$ , where  $b_j(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$  and  $\Gamma$  is the Gamma function. In the finite variance case we assume that the memory parameter  $|d| < 1/2$  and for the general Lévy  $\alpha$ -stable case we assume that  $\alpha > 1$  and  $|d| < 1 - 1/\alpha$  [53].

Let us emphasize here a very convenient ‘building block structure’ of the ARFIMA (1,  $d$ , 1) model:

TABLE I. Physical ( with and without measurement noise) and corresponding ARFIMA(1,  $d$ , 1) models.

Physical model	ARFIMA(1, $d$ , 1)
Confinement by a potential well (O-U)	AR(1)
O-U + noise	ARMA(1,1)
Brownian motion (BM)	MA(0)
BM + noise	MA(1)
FBM	FI( $d$ )
FBM + noise	FIMA( $d$ , 1)

ARFIMA(1,  $d$ , 1) = AR(1) + FI( $d$ ) + MA(1), where the AR(1), FI( $d$ ) and MA(1) are defined by the following equations:

$$X(t) - \phi X(t - 1) = Z(t), \quad (2)$$

$$(1 - B)^d X(t) = Z(t), \quad (3)$$

$$X(t) = Z(t) - \psi Z(t - 1), \quad (4)$$

respectively.

The FI( $d$ ) part, which is related to the fractional difference operator, leads to the regularly varying (power-law) correlations which are related to the classical definition of long-range dependence (long memory): lack of summability of correlations [54]. In this paper we will call all processes with power-law correlations long-range dependent even if the correlations are summable in contrast to the exponentially (so much faster) decaying correlations. Long-memory processes are often used to describe the dynamical behavior of complex systems. Consequently, ARFIMA models have already emerged in the physical literature, e.g., in the modeling of soft X-ray solar emissions [55, 56], heartbeat interval changes, air temperature changes [57], and the motion of molecules in live cells [43–45, 47].

The ARFIMA(1,  $d$ , 1) process offers a lot of flexibility in modeling long (power-law) and short (exponential or finite-time) dependencies by choosing the memory parameter  $d$  in the FI part, and appropriate AR(1) and MA(1) process parameters ( $\phi$  and  $\psi$ ). Hence, they can be tailored to different empirical data. To illustrate it, let us recall three standard models used in the biophysics literature: confinement by a potential well, FBM and Brownian motion. These models correspond to different components of the ARFIMA(1,  $d$ , 1) process, namely the AR(1) part, FI( $d$ ) part and partial sum process of MA(0) (which is MA(1) with  $\psi = 0$ , so a pure white sequence), respectively. Furthermore, the MA(1) which introduces a one-lag dependence can be associated with the measurement noise [46]. When added to the FI( $d$ ) process, we obtain the FIMA( $d$ , 1) model:  $(1 - B)^d X(t) = Z(t) - \psi Z(t - 1)$ , which corresponds to FBM with noise [46]. This leads to an efficient algorithm for extracting the magnitude of the measurement error for fractional dynamics based on the FIMA processes [47]. The relations between the different models are summarized in Table I.

If the sample comes from an ARFIMA(1,  $d$ , 1) pro-

cess with noise belonging to the domain of attraction of Lévy  $\alpha$ -stable law, then for large sample lengths the time-averaged mean-squared displacement (MSD)  $\overline{\delta^2(\tau)}$  behaves like

$$\overline{\delta^2(\tau)} \sim \tau^a, \quad (5)$$

where the bar stands for the time average,  $\sim$  denotes asymptotic behavior and  $a = 2d + 1$  [37]. Therefore, the memory parameter  $d$  controls the extent of the diffusion anomaly regardless of the underlying distribution. If  $d < 0$ , the process is subdiffusive, and if  $d > 0$ , the character of the process changes to superdiffusive.

Single-particle tracking data are commonly analyzed by the MSD of the trajectories. However, past work has shown that MSDs can be susceptible to errors and biases. To improve accuracy in single particle studies guidelines with respect to measurement length and maximum time lags have been proposed [58].

In this article, for the purpose of identifying diffusive motions in single-particle trajectories we apply two special cases of the ARFIMA(1,  $d$ , 1) model, namely AR(1) and FIMA( $d$ , 1). FIMA( $d$ , 1) is characterized by three parameters: fractional  $d$  giving rise to the long memory, the parameter  $\psi$  of the MA(1) part corresponding to the one-lag dependence, and the distribution of the noise sequence, which in the case of the Gaussian white noise is fully characterized by its variance  $\sigma^2$ . The second model is AR(1), which is characterized by the AR parameter  $\phi$  and the distribution of the noise sequence. This model is a discrete analogue of the continuous-time Ornstein-Uhlenbeck (O-U) process  $X(t)$  defined by the overdamped Langevin equation

$$\frac{dX(t)}{dt} = -\frac{kX(t)}{\gamma} + \xi(t), \quad (6)$$

where  $\xi(t)$  is white Gaussian noise with  $\langle \xi(t)\xi(t') \rangle = 2D\delta(t-t')$ . A discretization of equation (6) gives

$$X(t) = -\underbrace{\left(\frac{k}{\gamma}\Delta t - 1\right)}_{\phi} X(t-1) + \underbrace{\Delta t \xi(t)}_{Z(t)}, \quad (7)$$

which is the AR(1) model equation with  $\phi = -(k\Delta t/\gamma - 1)$  and Gaussian white noise  $Z(t) = \Delta t \xi(t)$  with variance  $\sigma^2 = 2D\Delta t$ .

### III. ANALYSIS OF INDIVIDUAL MEMBRANE PROTEIN TRAJECTORIES

Single-particle tracking on the plasma membrane of mammalian cells indicates that often molecules are subjected to transient confinement. Here we employ ARFIMA models to characterize the motion of Nav1.6 channels in the soma of transfected cultured rat hippocampal neurons and beta 2 adrenergic receptors (B2AR) in transfected human embryonic kidney

(HEK-293) cells. Nav1.6 channels were biotinylated at an extracellular site and labeled with streptavidin-conjugated CF640R. The B2AR were tagged with HA and labeled with anti-HA antibody conjugated to CF640. Transfected cells were imaged by total internal reflection microscopy at 20 frames/s and individual fluorescent molecules were tracked with the u-track algorithm [59]. Experimental details about cell culture, transfection, and imaging were published previously [36, 60].

Following Ref. [61], we employed an automated algorithm to detect changes in molecule dynamics. This algorithm was based on a sliding-window MSD. As a result, we obtained trajectories belonging to two states: free and confined. Figure 1 depicts two representative free and two representative confined trajectories. Trajectories from the free state resemble Brownian diffusion whereas confined-state trajectories appear as realizations of a stationary process.

We selected five long representative trajectories corresponding to each motion (free and confined). The shortest were 174 (free state) and 153 (confined state) points for the the Nav1.6, and 84 (free state) and 300 (confined state) points for the beta 2 receptors. We focused on the X-coordinate of the motions. We first fitted the ARFIMA( $d$ , 1) model to the increments of free-state trajectories. The simplest well-fitted model common to all trajectories is FIMA( $d$ , 1). Figures 2 and 3 show a scatter plot of the estimated  $d$  and  $\psi$  parameters with the corresponding 95% confidence intervals obtained by Monte Carlo simulations. For the beta 2 receptor trajectories, the memory parameter  $d$  and the moving average parameter  $\psi$  are usually lower than for Nav1.6 trajectories. The MSD anomalous exponent  $a = 2d + 1$  is also shown in the right axis of Figure 2. The detailed results of the FIMA identification and validation procedure are presented in the Supplemental Material [52].

To check the goodness of fit of the FIMA model, we calculated the mean-squared displacement (MSD) for 1000 simulated trajectories of the model with parameters given in the Supplemental material [52] and compared the results with the MSD values of the analyzed representative trajectories. We can see in Fig. 4 that the fitted FIMA processes reproduce the sample MSD well. Some of the empirical MSD values fall outside the interquartile range but they are always contained in the 95% confidence interval [62], see Figure S1 in the Supplemental material [52]. By examining the width of the boxplots, we can also observe that the variability of the MSD exponent in the beta 2 receptor case seems bigger than in the Nav1.6 case, which suggests another difference between Nav 1.6 and beta 2 receptor dynamics.

We also checked the distribution of the model residuals. First, following the procedure in Appendix D we found that the residuals belong to the domain of attraction of the normal law. Next, it appeared that the residuals can be treated as Gaussian since the test was rejected only for one trajectory from the confined state. Hence we conclude that the increments of the trajectories in the

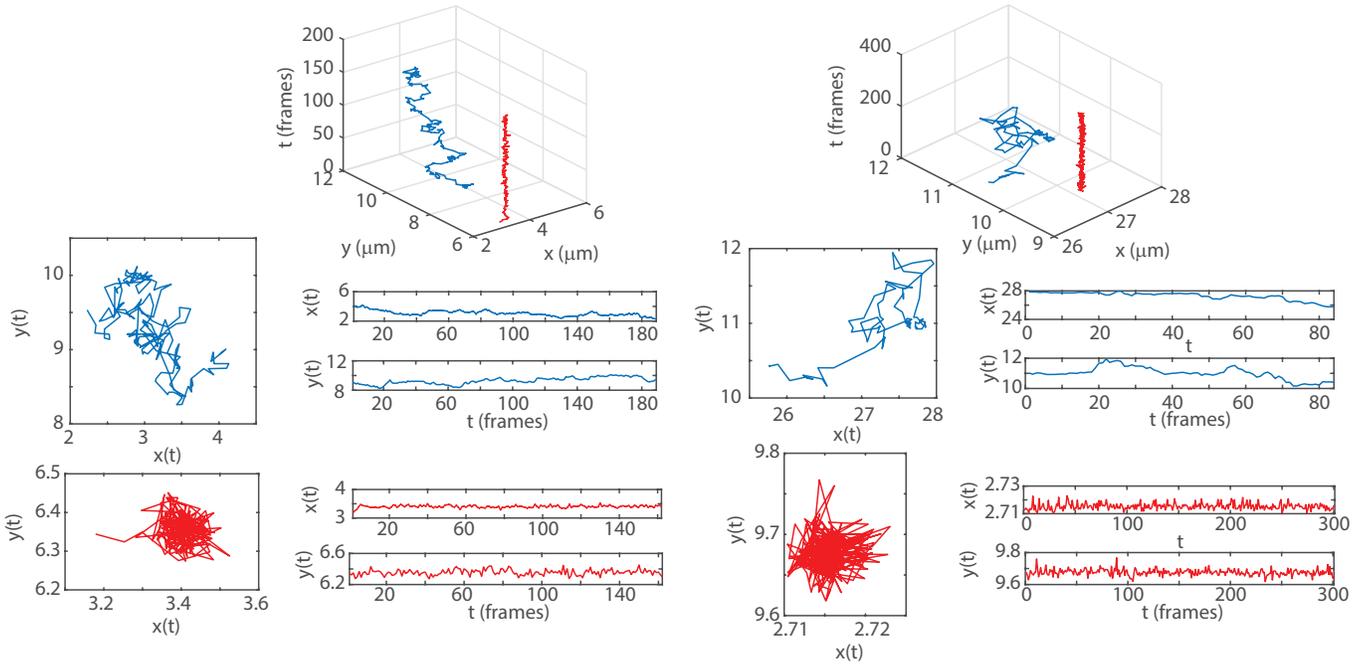


FIG. 1. Representative Nav1.6 (left panel) and beta 2 receptor (right panel) trajectories. The top panel illustrates the receptor dynamics in 3D, where the  $z$ -axis corresponds to time. The (upper) blue trajectories show molecules in the free state and the (lower) red trajectories show molecules in the confined state.

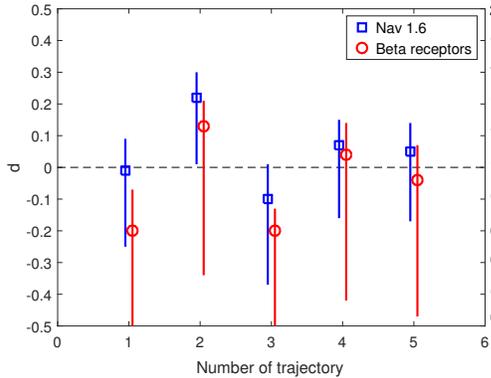


FIG. 2. Fitted  $d$  parameter (with the 95% confidence interval) of the FIMA( $d, 1$ ) model to the increments of the free state Nav1.6 and beta 2 receptor five representative trajectories.

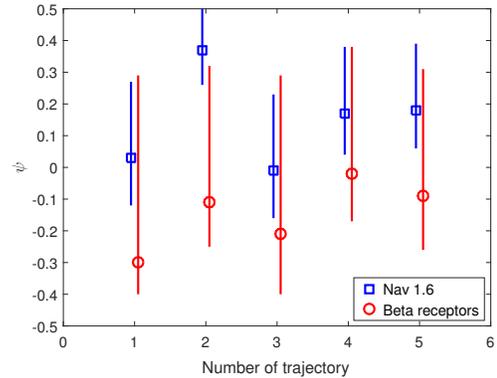


FIG. 3. Fitted  $\psi$  parameter (with the 95% confidence interval) of the FIMA( $d, 1$ ) model to the increments of the free state Nav1.6 and beta 2 receptor five representative trajectories.

free state can be modeled by the Gaussian FIMA( $d, 1$ ). Recall that this model represents a Gaussian fractional process with power-law memory, which corresponds to fractional Brownian motion with Hurst index  $H = d + 1/2$  [42].

We performed the same analysis for the confined case and we found that a simpler ARFIMA model describes well the data, namely AR(1). Its estimated parameters are depicted in Figure 5 with the corresponding 95% confidence intervals obtained by Monte Carlo simulations. Hence, with the use of the ARFIMA model we are able

to distinguish between these two different states. Moreover, we can see that for beta 2 receptor trajectories the autoregressive parameter is usually lower than for the Nav1.6 trajectories. The detail results of the AR identification and validation procedure are presented in the Supplemental Material [52].

As we showed in Section II, AR(1) corresponds to the O-U process. From a physical point of view, this model describes the motion of a Brownian particle in a harmonic potential with restoring force  $F = -kx$  and damping coefficient  $\gamma$  (Eq. 7). Therefore, we calculated the

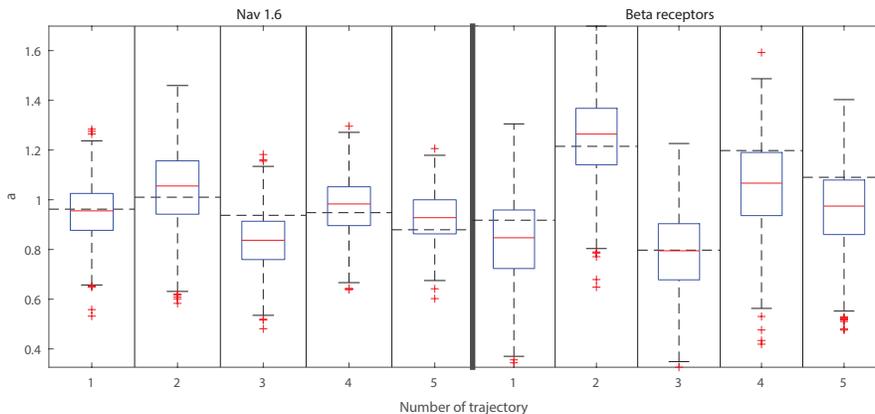


FIG. 4. Boxplots of estimated MSD exponents for 1000 simulated trajectories of the fitted FIMA processes for free state Nav1.6 and beta 2 receptor five representative trajectories. The dashed horizontal line stands for the MSD exponent obtained for the analyzed empirical trajectory.

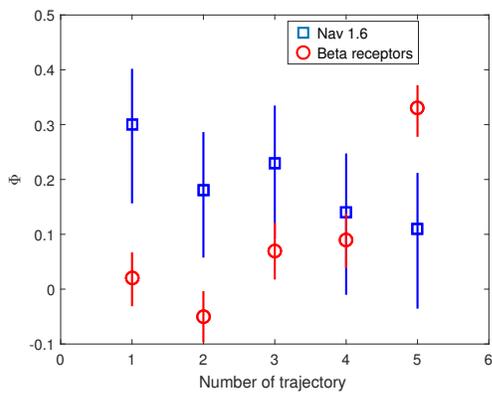


FIG. 5. Fitted  $\phi$  parameter (with the 95% confidence interval) of the AR(1) model to the confined state Nav1.6 and beta 2 receptor five representative trajectories.

corresponding O-U parameters and, in turn, the corresponding physical parameters that characterize the protein motion (Tables II-III). Namely, the diffusion is  $D = \sigma^2/(2\Delta t)$ , where  $\sigma^2$  is the variance of the noise term  $Z(t)$  and, in our time series,  $\Delta t = 0.05$  s. By considering the Einstein-Smoluchowski relation,  $\gamma = k_B T/D$  with  $k_B T$  being thermal energy at 37°C. Next the parameter  $\phi$  yields the constant  $k = \gamma(1 - \phi)/\Delta t$ . At last, by equipartition, the variance of the particle position, i.e., the square of the characteristic radius of the confining domain, is found,  $\langle X^2 \rangle = k_B T/k = \sigma^2/[2(1 - \phi)]$ . We note that in the analyzed trajectories, the confining radius of Nav1.6 channels is larger than the confining radius of B2AR, where the former yields a mean of 53 nm, while the latter has a mean of 23 nm.

To check the goodness of fit of the AR model, we calculated the mean-squared displacement (MSD) for 1000 simulated trajectories of the model with parameters given in the Supplemental material [52] and compared the results with the MSD values for the analyzed representa-

tive trajectories. We can see in Fig. 6 that the fitted AR processes reproduce the sample MSD well. Some of the empirical MSD values fall outside the interquartile range but they are always (even beta 2 receptor trajectory no. 3) contained in the 95% confidence interval [62], see Figure S2 in the Supplemental material [52]. By examining the width of the boxplots, we can also observe that the variability of the MSD exponent in the Nav 1.6 case seems bigger than in the beta 2 receptor case, which suggests another difference between Nav 1.6 and beta 2 receptor dynamics.

We also checked the distribution of the model residuals. Following the procedure in Appendix D, first we found that they belong to the domain of attraction of the normal law. However, in contrast to the free state, Gaussianity was rigorously rejected for almost all trajectories. NIG and  $t$  location-scale distributions were not rejected for most of the trajectories. In order to make a rigorous conclusion about the model residuals one should analyze more trajectories. Hence we conclude that the confined state trajectories can be modeled by the non-Gaussian AR(1).

For the detailed goodness-of-fit analysis of identified AR(1) and FIMA ( $d, 1$ ) models we refer the reader to the Supplemental Material [52].

#### IV. DISCUSSION AND CONCLUSIONS

In this paper we demonstrated that the ARFIMA(1,  $d$ , 1) model can identify two types of motions of membrane proteins. We analyzed five representative trajectories chosen from each of two categories: free and confined states of Nav1.6 and beta 2 adrenergic receptors. We found that the two special cases of the model, FIMA( $d, 1$ ) and AR(1), fully identify the free and confined state dynamics, respectively. These results allowed us to propose a new unified methodology to detect certain types of motion in complex systems.

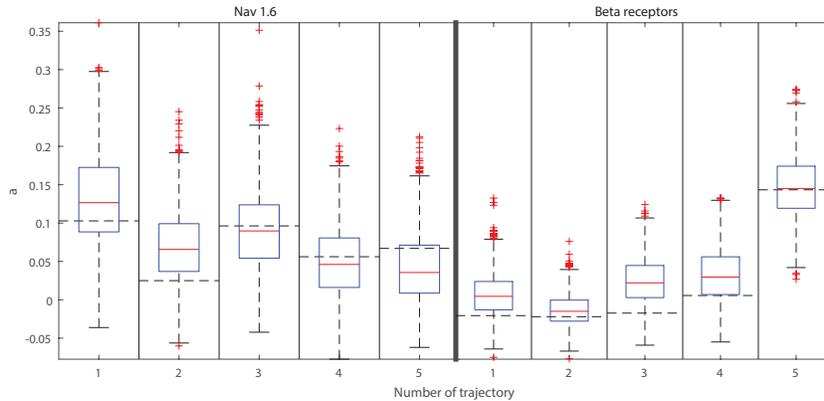


FIG. 6. Boxplots of estimated MSD exponents for 1000 simulated trajectories of the fitted AR processes for the confined state Nav1.6 and beta 2 receptor five representative trajectories. The dashed horizontal line stands for the MSD exponent obtained for the analyzed empirical trajectory.

TABLE II. Ornstein-Uhlenbeck process parameters for the representative confined state Nav1.6 trajectories.

Traj.	$\phi$	$\sigma^2 (nm^2)$	$D (\mu m^2/s)$	$\gamma (10^{-8} \text{ kg/s})$	$k (\text{pN}/\mu\text{m})$	$\langle X^2 \rangle (nm^2)$
1	0.30	1948	0.019	22.0	3.08	1391
2	0.18	7966	0.080	5.4	0.88	4857
3	0.23	4319	0.043	9.9	1.53	2804
4	0.14	4053	0.041	10.6	1.82	2356
5	0.11	6105	0.061	7.0	1.24	3430

In the free state, the beta 2 receptors appear to be more subdiffusive and, moreover, the moving average parameter is lower. In the confined state the autoregressive parameter for the beta 2 receptor trajectories seems to be lower than for Nav1.6 trajectories. Furthermore, the distribution changes from being Gaussian in the free to non-Gaussian in the confined state. Since the AR(1) is a discrete-time counterpart of the Ornstein-Uhlenbeck (O-U) process, we calculated the parameters of the corresponding O-U processes which, we found, are biologically meaningful. We note that the estimated FIMA memory parameters provide more accurate information on the subdiffusion type than the MSD exponents since they are more robust with respect to the measurement noise.

We would like to point out that a very popular model for subdiffusion continuous-time random walk (CTRW) [63] is not considered here, but it can be represented in the form of subordinated O-U process, i.e., AR with randomized time described by the inverse stable process [64, 65].

Accurate motion analysis often requires a transient motion classification [66, 67]. Many transient motion analysis algorithms employ either rolling windows analysis [68, 69] or hidden Markov modeling [70, 71]. Our studies show that one can also consider the ARFIMA model as a possible tool for such classification.

We believe that our methodology provides a simple unified way to gain deeper information into processes

leading to anomalous diffusion in single-particle tracking experiments. Finally, we note that in order to model the whole trajectories (free and confined parts of the trajectories together) the ARFIMA process is not enough. One possible extension is ARFIMA with noise described by the generalized autoregressive conditional heteroskedasticity (GARCH) model [72, 73]. Such models can be useful in description of changing diffusivity which results in so-called transient anomalous diffusion [74, 75]. ARFIMA combined with GARCH can describe both power-law decay of the autocorrelation function with arbitrary finite-lag effects (ARFIMA part) and changing diffusion exponent (GARCH part). This is a subject of ongoing work.

## ACKNOWLEDGMENTS

We thank Patrick Mannion, Sanaz Sadegh, and Laura Solé for useful discussions. K.B. and G.S. acknowledge the support by NCN Maestro Grant No. 2012/06/A/ST1/00258. Research of A.W. was supported by NCN-DFG Beethoven Grant No. 2016/23/G/ST1/04083. D.K. was supported by the National Science Foundation under grant No. 1401432. M.M.T acknowledges the support of the National Institutes of Health under grants R01GM109888 and RO1NS085142.

TABLE III. Ornstein-Uhlenbeck process parameters for the representative confined state beta 2 receptor trajectories.

Traj.	$\phi$	$\sigma^2(nm^2)$	$D(\mu m^2/s)$	$\gamma$ ( $10^{-8}$ kg/s)	$k$ (pN/ $\mu m$ )	$\langle X^2 \rangle (nm^2)$
1	0.02	553	0.006	77.4	15.18	282
2	-0.05	1327	0.0013	32.3	6.77	632
3	0.07	1485	0.015	28.8	5.36	799
4	0.09	1696	0.017	25.2	4.59	932
5	0.33	252	0.003	169.9	22.76	188

### Appendix A: General ARFIMA model

The ARFIMA( $p, d, q$ ) process  $X(t)$  for  $t = 0, \pm 1, \dots$  is defined as a solution of the equation

$$(1 - B)^d \Phi_p(B)X(t) = \Psi_q(B)Z(t),$$

where  $Z(t)$  is the noise sequence (Gaussian or, in general, in the domain of attraction of Lévy stable law) and  $B$  is the backshift operator, i.e.  $BX(t) = X(t - 1)$  and  $B^j X(t) = X(t - j)$  [40, 51]. Moreover the  $\Phi_p$  and  $\Psi_q$  are AR and MA polynomials respectively, known in classical time series theory:

$$\begin{aligned} \Phi_p(B)X(t) &= X(t) - \Phi_1 X(t - 1) - \dots - \Phi_p X(t - p), \\ \Psi_q(B)Z(t) &= Z(t) - \psi_1 Z(t - 1) - \dots - \psi_q Z(t - q). \end{aligned}$$

The crucial part of the above definition of ARFIMA series is the fractional difference operator  $(1 - B)^d$  defined as a power series

$$(1 - B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-B)^j = \sum_{j=0}^{\infty} b_j(d) B^j,$$

where  $b_j(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$  with asymptotic behavior  $b_j(d) \sim \Gamma(-d)^{-1} j^{-d-1}$  and  $\Gamma$  is the Gamma function.

In the finite variance case we assume that  $|d| < 1/2$  and for the general Lévy  $\alpha$ -stable case we assume that  $\alpha > 1$  and  $|d| < 1 - 1/\alpha$  [53]. The assumption about the exponent  $d$  ensures a proper definition of the operator for the Gaussian ARFIMA processes. The parameter  $d$  is called the memory parameter. From the physical point of view, it is known that ARFIMA is a discrete time analogue of the fractional Langevin equation that takes into account the memory parameter [41, 56].

In the Gaussian case, the autocovariance function  $r(k)$  of the ARFIMA process decays as  $k^{2d-1}$ . Moreover, for  $d > 0$  we have  $\sum_{k=0}^{\infty} |r(k)| = \infty$ , which serves as a classical definition of long memory [54]. For the Lévy  $\alpha$ -stable case with  $\alpha < 2$  the covariance does not exist and one has to replace it, e.g., with the codifference (see [76]). The codifference of the ARFIMA process was studied in [53], where it was proved that for  $d > 1 - 2/\alpha$  the ARFIMA possesses long-term dependence in the classical sense.

A partial sum of the ARFIMA process is asymptotically self-similar with the Hurst index equal to  $d + 1/\alpha$ , where  $\alpha$  is the index of stability [77]. As a consequence Gaussian ARFIMA(0, 0, 0), which is just the pure white

noise sequence, corresponds, in the limit sense, to BM. Similarly, ARFIMA(0,  $d$ , 0) corresponds to FBM with  $H = d + 1/2$ . For more information about the ARFIMA processes with their applications to biophysics see, e.g., Refs. [43, 47].

The ARFIMA process (in the literature also called FARIMA) is a generalization of the classical stationary discrete-time ARMA process to account for the long-range dependence (power-like decaying autocorrelation function) [51, 54].

The ARMA models provide a general framework for studying stationary short memory phenomena, i.e., processes with exponentially decaying autocovariance. These models consist of two broad classes of time series processes, namely the autoregressive (AR) and the moving average (MA). The ARMA is usually referred to as the ARMA( $p, q$ ) model where  $p$  is the order of the autoregressive part and  $q$  the order of the moving average part.

Let us now concentrate on the ARMA(1, 1) case which is sufficient for many studies. The process  $X(t)$  is ARMA(1, 1) if it is stationary and satisfies (for every  $t$ ) a linear difference equation with constant coefficients:

$$X(t) - \phi X(t - 1) = Z(t) - \psi Z(t - 1), \quad (\text{A1})$$

where  $t = 0, \pm 1, \dots$

The basic building blocks of the model are the AR(1):  $X(t) = \phi X(t - 1) + Z(t)$  and MA(1):  $X(t) = Z(t) - \psi Z(t - 1)$  processes, where  $\phi$  and  $\psi$  are real parameters and  $Z(t)$  is the noise term [51]. AR(1) stands for the autoregression and the explanatory variable is the observation immediately prior to our current observation. Its autocorrelation function  $r(k)$  decays as  $\phi^k \langle X^2(t) \rangle$ . The MA(1) part introduces one-lag dependence in the time series, namely  $X(t)$  is a stationary 1-correlated time series:  $X(s)$  and  $X(t)$  are independent whenever  $|t - s| > 1$ . The dependence is fully controlled by the parameter  $\psi$ . A stationary solution of ARMA(1, 1) equation exists if and only if  $\phi \neq \pm 1$ . If  $|\phi| < 1$ , then a unique stationary solution exists and is causal, since  $X(t)$  can be expressed in terms of the current and past values  $Z(s)$ ,  $s \leq t$ . Otherwise, if  $|\phi| > 1$ , then the solution is not causal since  $X(t)$  is then a function of  $Z(s)$ ,  $s \geq t$ . Moreover, if  $|\psi| < 1$  then  $X(t)$  is invertible, so the noise process  $Z(t)$  can be expressed in terms of past values  $X(s)$ ,  $s \leq t$  [51]. For the noise process  $Z(t)$  we only assume that it belongs to

the domain of attraction of Lévy  $\alpha$ -stable law for  $\alpha \leq 2$  [76, 78]. It can be a finite variance white noise (uncorrelated random variables with mean zero and variance  $\sigma^2$ , e.g., Gaussian or t-Student) or infinite variance i.i.d. sequence (e.g., Lévy  $\alpha$ -stable with  $\alpha < 2$  or Pareto).

The ARFIMA process is a  $d$ -differenced ARMA process, where  $d$  is a fractional memory parameter. As a consequence, ARFIMA(1,  $d$ , 1) process  $X(t)$  is defined as a stationary solution of the fractional difference equation [51]

$$(1 - B)^d (X(t) - \phi X(t-1)) = Z(t) - \psi Z(t-1). \quad (\text{A2})$$

### Appendix B: ARFIMA parameter estimation

In order to estimate the ARFIMA parameters one can apply the maximum likelihood estimation method, which is implemented, e.g., in ITSM [51], or its approximation given by the Whittle estimator [79, 80]. The Whittle estimator is particularly simple to implement in any computer software and it benefits from the elementary form of the ARFIMA spectral density. Briefly, let  $\{x_1, x_2, \dots, x_N\}$  be a trajectory of length  $N$ . For the model FIMA( $d$ , 1), we estimate the vector  $\beta = (\phi, d)$ . We denote the normalized periodogram by

$$\tilde{I}_N(\lambda) = \frac{\left| \sum_{t=1}^N x_t e^{-i\lambda t} \right|^2}{\sum_{t=1}^N x_t^2}, \quad -\pi \leq \lambda \leq \pi. \quad (\text{B1})$$

The Whittle estimator is defined as the vector argument  $(\phi, d)$  for which the following function attains its minimum value:

$$\int_0^\pi \tilde{I}_N(\lambda) W(\lambda, \phi) (2 - 2 \cos \lambda)^d d\lambda, \quad (\text{B2})$$

where

$$W(\lambda, \phi) = (1 - 2\phi \cos \lambda + \phi^2). \quad (\text{B3})$$

For the AR(1) the estimator minimizes

$$\int_0^\pi \tilde{I}_N(\lambda) (1 - 2\phi \cos \lambda + \phi^2) d\lambda. \quad (\text{B4})$$

### Appendix C: ARFIMA model validation

Having fitted the ARFIMA process, the next step is to investigate the residuals obtained either by fractional differencing the data [81] or as prediction errors [51].

- If there is no dependence among the residuals, we can regard them as observations of independent random variables and there is no further modeling to be done except to estimate their mean and variance [51].

- If there is a significant dependence among the residuals, we need to evaluate a more complex stationary time series model for the noise that accounts for the dependence, e.g. the generalized autoregressive conditional heteroskedasticity (GARCH) process [72, 73].

We now recall some simple tests for checking the hypothesis that the residuals are observed values of independent and identically distributed (i.i.d.) random variables. If they are, we conclude the model describes the data well.

First, we plot the sample autocorrelation function with its 95% confidence interval. In the Gaussian case about 95% of the sample autocorrelations should fall between the bounds  $\pm 1.96/\sqrt{n}$ . The same can be done for squares of the observations to check for a dependence in variance. Next, we apply the Ljung-Box, and McLeod-Li tests, which are portmanteau tests [51]. The Ljung-Box test relies on the sample autocorrelation function, which, at lag  $h$ , has a chi-squared distribution with  $h$  degrees of freedom. The McLeod-Li test is similar but on the squared data.

### Appendix D: ARFIMA residual distribution

Having found an appropriate ARFIMA model describing the data, we can identify the distribution underlying the noise sequence. Information about the distribution is helpful in determining confidence intervals for the estimated parameters and also in construction of prediction intervals. We now investigate the distribution of the residuals.

Following [82], we first check if the underlying distribution belongs to the domain of attraction of the Gaussian or non-Gaussian Lévy-stable distributions by examining its rate of convergence.

- If the results suggest the Gaussian domain of attraction, we consider three typical light-tailed probability distributions for the residuals of the ARFIMA model, namely Gaussian,  $t$  location-scale and normal inverse Gaussian (NIG).
- If the non-Gaussian Lévy stable domain of attraction is suggested, we test for Lévy  $\alpha$ -stable distributions.

The  $t$  location-scale distribution generalizes ordinary student's  $t$  distribution. The probability density function (PDF) of this distribution is given as

$$f_t(x) = \frac{n^{-\frac{1}{2}}}{\sigma B\left(\frac{n}{2}, \frac{1}{2}\right)} \left\{ 1 + \frac{[(x - \mu)/\sigma]^2}{n} \right\}^{-\frac{n+1}{2}}, \quad (\text{D1})$$

where  $B$  is the beta function  $B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$  and  $n$  denotes the degrees of freedom parameter. This distribution is useful for modeling data with heavier

tails than the normal. It approaches the Gaussian distribution as  $n$  approaches infinity and smaller values of  $n$  yield heavier tails.

A random variable  $X$  is said to have a NIG distribution if it has a PDF

$$f_{\text{NIG}}(x) = \frac{\alpha\delta}{\pi} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)} \frac{K_1 \left[ \alpha\sqrt{\delta^2 + (x-\mu)^2} \right]}{\sqrt{\delta^2 + (x-\mu)^2}}. \quad (\text{D2})$$

The NIG distribution, introduced in Ref. [83], is described by four parameters  $(\alpha, \beta, \delta, \mu)$ , where  $\alpha$  stands for tail heaviness,  $\beta$  for asymmetry,  $\delta$  is the scale parameter, and  $\mu$  is the location. The normalizing constant  $K_\lambda(t)$  in (D2) is the modified Bessel function of the third kind

with index  $\lambda$ , also known as the MacDonald function. The NIG distribution is more flexible than the Gaussian distribution since it allows for fat-tails and skewness. The Gaussian distribution arises as a special case by setting  $\beta = 0$ ,  $\delta = \sigma^2\alpha$ , and letting  $\alpha \rightarrow \infty$ .

In order to check the goodness of fit of the distributions considered here, we apply the Anderson-Darling test [84], which is based on the  $A^2$  statistic:

$$A^2 = N \int_{-\infty}^{\infty} \frac{[F_N(x) - F(x)]^2}{F(x)[1 - F(x)]} dF(x), \quad (\text{D3})$$

where  $F_N(x)$  and  $F(x)$  denote the empirical and theoretical distribution functions, respectively. The test is one of the most powerful statistical tools for detecting most departures from normality [84].

- 
- [1] F. Höfling and T. Franosch, *Rep. Prog. Phys.* **76**, 046602 (2013).
- [2] R. Metzler, J.-H. Jeon, A. G. Cherstvy and E. Barkai, *Phys. Chem. Chem. Phys.* **16**, 24128-24164 (2014).
- [3] C. Manzo and M. F. Garcia-Parajo, *Rep. Progr. Phys.* **78**, 124601 (2015).
- [4] D. Krapf, *Curr. Top. Membr.* **75**, 167-207 (2015).
- [5] M. Dahan, S. Lévi, C. Luccardini, P. Rostaing, B. Riveau and A. Triller, *Science* **302**, 442-445 (2003).
- [6] A. Sergé, N. Bertaux, H. Rigneault and D. Marguet, *Nat. Methods* **5**, 687-694 (2008).
- [7] S. Manley, J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz, *Nat. Methods* **5**, 155-157 (2008).
- [8] A. V. Weigel, B. Simon, M. M. Tamkun and D. Krapf, *Proc. Natl. Acad. Sci. USA* **108**, 6438 (2011).
- [9] A. Kusumi, K. G. Suzuki, R. S. Kasai, K. Ritchie and T. K. Fujiwara, *Trends Biochem. Sci.* **36**, 604-615 (2011).
- [10] D. Calebiro, F. Rieken, J. Wagner, T. Sungkaworn, U. Zabel, A. Borzi, E. Cocucci, A. Zürn and M. J. Lohse, *Proc. Natl. Acad. Sci. USA* **110**, 743-748 (2013).
- [11] A. V. Weigel, M. M. Tamkun and D. Krapf, *Proc. Natl. Acad. Sci. USA* **110**, E4591 (2013).
- [12] J. B. Masson, P. Dionne, C. Salvatico, M. Renner, C. G. Specht, A. Triller and M. Dahan, *Biophys. J.* **106**, 74-83 (2014).
- [13] M. Weiss, M. Elsner, F. Kartberg, and T. Nilsson, *Biophys. J.* **87**, 3518 (2004).
- [14] I.M. Tolić-Norrelykke, E.-L. Munteanu, G. Thon, L. Oddershede, and K. Berg-Sorensen, *Phys. Rev. Lett.* **93**, 078102 (2004).
- [15] I. Golding and E.C. Cox, *Phys. Rev. Lett.* **96**, 098102 (2006).
- [16] J.-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sorensen, L. Oddershede and R. Metzler, *Phys. Rev. Lett.* **106**, 048103 (2011).
- [17] Š. Bálint, I. V. Vilanova, A. Sandoval Álvarez and M. Lakadamyali, *Proc. Natl. Acad. Sci. USA* **110**, 3375-3380 (2013).
- [18] T. Kues, R. Peters and U. Kubitscheck, *Biophys. J.* **80**, 2954-2967 (2001).
- [19] V. Levi, Q. Ruan, M. Plutz, A. S. Belmont and E. Gratton, *Biophys. J.* **89**, 4275-4285 (2005).
- [20] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai and Y. Garini, *Phys. Rev. Lett.* **103**, 018102 (2009).
- [21] J. C. M. Gebhardt, D. M. Suter, R. Roy, Z. W. Zhao, A. R. Chapman, S. Basu, T. Maniatis and X S. Xie, *Nat. Methods* **10**, 421-426 (2013).
- [22] I. Izeddin, V. Récamier, L. Bosanac, I. I. Cissé, L. Boudarene, C. Dugast-Darzacq, F. Proux, O. Bénichou, R. Voituriez and O. Bensaude, *eLife* **3**, e02230 (2014).
- [23] D. Krapf, *Curr. Opin. Cell Biol.* **53**, 15-21 (2018).
- [24] E. Barkai, Y. Garini and R. Metzler, *Phys. Today* **65**, 29-35 (2012).
- [25] E. Kepten, I. Bronshtein and Y. Garini, *Phys. Rev. E* **87**, 052713 (2013).
- [26] I. Bronshtein, E. Kepten, I. Kanter, S. Berezin, M. Lindner, Abena B. Redwood, S Mai, S. Gonzalo, R. Foissner, Y. Shav-Tal and Y. Garini, *Nat. Commun.* **6**, 8044 (2015).
- [27] Y. Zhang and O. K. Dudko, *Annu. Rev. Biophys.* **45**, 117-134 (2016).
- [28] W. Pan, L. Filobelo, N. D. Q. Pham, O. Galkin, V. V. Uzunova and Peter G. Vekilov, *Phys. Rev. Lett.* **102**, 058101 (2009).
- [29] J. Szymański and M. Weiss, *Phys. Rev. Lett.* **103**, 038102 (2009).
- [30] G. Campagnola, K. Nepal, B. W. Schroder, O. B. Peersen, and D. Krapf, *Sci. Rep.* **5**, 17721 (2015).
- [31] J.-H. Jeon and R. Metzler, *Phys. Rev. E* **81**, 021103 (2010).
- [32] A. V. Weigel, S. Ragi, M. L. Reid, E. K. P. Chong, M. M. Tamkun, and D. Krapf, *Phys. Rev. E* **85**, 041924 (2012).
- [33] M. Hellmann, J. Klafter, D. W. Heermann and M. Weiss, *J. Phys. Condens. Matter* **23**, 234113 (2011).
- [34] E. Kepten, I. Bronshtein and Y. Garini, *Phys. Rev. E* **83**, 041919 (2011).
- [35] D. ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, 2000).
- [36] S. Sadegh, J. L. Higgins, P. C. Mannion, M. M. Tamkun and D. Krapf, *Phys. Rev. X* **7**, 011031 (2017).
- [37] K. Burneckı and A. Weron, *Phys. Rev. E* **82**, 021130

- (2010).
- [38] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini, *Phys Rev. Lett.* **103**, 018102 (2009).
- [39] J. R. M. Hosking, *Biometrika* **68**, 165 (1981).
- [40] C. W. J. Granger and R. Joyeux, *J. Time Series Anal.* **1**, 15 (1980).
- [41] M. Magdziarz and A. Weron, *Studia Math.* **181**, 47 (2007).
- [42] K. Burnecki, A. Weron, *J. Stat. Mech.* P10036 (2014).
- [43] K. Burnecki, M. Muszkieta, G. Sikora, A. Weron, *EPL* **98**, 10004 (2012).
- [44] K. Burnecki, *J. Stat. Mech.* P05015 (2012).
- [45] K. Burnecki, G. Sikora and A. Weron, *Phys. Rev. E* **86**, 041912 (2012).
- [46] K. Burnecki, E. Kepten, Y. Garini, G. Sikora and A. Weron, *Sci. Rep.* **5**, 11306 (2015).
- [47] G. Sikora, E. Kepten, A. Weron, M. Balcerek, K. Burnecki, *Phys. Chem. Chem. Phys.* **19**, 26566 (2017).
- [48] N. Crato and P. Rothman, *Econom. Lett.* **45**, 287 (1994).
- [49] G. Fouskitakis and S. Fassois, *IEEE Trans. Signal Process.* **47**, 3365 (1999).
- [50] L. Gill-Alana, *Economics Bulletin* **3**, 1 (2004).
- [51] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting* (Springer-Verlag, New York, 2002); *ITSM for Windows: A User's Guide to Time Series Modelling and Forecasting* (Springer-Verlag, New York, 1994).
- [52] See Supplemental Material
- [53] P. Kokoszka and M. S. Taqqu, *Stoch. Proc. Appl.* **60**, 19 (1995).
- [54] J. Beran, *Statistics for Long-Memory Processes* (Chapman & Hall, New York, 1994).
- [55] A. Stanislavsky, K. Burnecki, M. Magdziarz, A. Weron, and K. Weron, *Astrophys. J.* **693**, 1877 (2009).
- [56] K. Burnecki, J. Klafter, M. Magdziarz, and A. Weron, *Phys. A* **387**, 1077 (2008).
- [57] B. Podobnik, P. Ch. Ivanov, K. Biljakovic, D. Horvatic, H. E. Stanley, and I. Grosse, *Phys. Rev. E* **72**, 026121 (2005).
- [58] E. Kepten, A. Weron, G. Sikora, K. Burnecki and Y. Garini, *PLoS ONE* **10**(2), e0117722 (2015).
- [59] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid and G. Danuser, *Nat. Methods* **5**, 695 (2008).
- [60] E. J. Akin, L. Solé, B. Johnson B, M. E. Beheiry, J. B. Masson, D. Krapf and M. M. Tamkun, *Biophys. J* **111**, 1235 (2016).
- [61] A. Weron, K. Burnecki, E. J. Akin, L. Sole, M. Balcerek, M. M. Tamkun and D. Krapf, *Sci. Rep.* **7**, 5404 (2017).
- [62] J. Neyman, *Philos. Trans. A Math. Phys. Sci.* **236**, 333 (1937).
- [63] R. Metzler and J. Klafter, *Phys. Rep.* **339**, 77 (2000).
- [64] A. Weron, M. Magdziarz, *Phys. Rev. Lett.* **105**, 260603 (2010).
- [65] M. Magdziarz and A. Weron, *Ann. Phys.* **326**, 2431 (2011).
- [66] A. R. Vega, S.A. Freeman, S. Grinstein and K. Jaqaman, *Biophys. J.* **114**, 1018 (2018).
- [67] V. Briane, Ch. Kervrann and M. Vimond, *Phys. Rev. E* **97**, 062121 (2018).
- [68] Y.-L. Liu et al., *Biophys. J.* **111**, 2214 (2016).
- [69] G. Sikora, A. Wylomańska, J. Gajda, L. Solé, E.J. Akin, M. M. Tamkun and D. Krapf, *Phys. Rev. E* **96**, 062404 (2017)
- [70] N. Monnier et al., *Nat. Methods* **12**, 838 (2015).
- [71] T. Sungkaworn, M.-L. Jobin, K. Burnecki, A. Weron, M. J. Lohse and D. Calebiro, *Nature* **550**, 543 (2017).
- [72] R. Baillie, C-F. Chung, M. A. Tieslau, *J. Appl. Econometrics* **11**, 23, (1996).
- [73] S. Ling, W. K. Li, *J. Amer. Statist. Assoc.* **92**, 1184 (1997).
- [74] C. Manzo, J. A. Torreno-Pina, P. Massignan, G. J. Lapeyre, Jr., M. Lewenstein and M. F. Garcia Parajo, *Phys. Rev. X* **5**, 011021 (2015).
- [75] A. V. Chechkin, F. Seno, R. Metzler and I. M. Sokolov, *Phys. Rev. X* **7**, 021002 (2017).
- [76] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes* (Chapman & Hall, New York, 1994).
- [77] S. Stoev and M. S. Taqqu, *Fractals* **12**, 95 (2004).
- [78] A. Janicki and A. Weron, *A Simulation and Chaotic Behaviour of  $\alpha$ -Stable Stochastic Processes* (Marcel Dekker Inc., New York, 1994).
- [79] V. Pipiras and M. S. Taqqu, *Long-Range Dependence and Self-Similarity* (Cambridge University Press, Cambridge, 2017).
- [80] K. Burnecki and G. Sikora, *IEEE Trans. Signal Process.* **61**, 2825 (2013).
- [81] K. Burnecki and G. Sikora, *Chaos Soliton. Frac.* **102**, 456 (2017).
- [82] K. Burnecki, A. Wylomańska and A. Chechkin, *PLoS One* **10**, e0145604 (2015).
- [83] O. E. Barndorff-Nielsen, *Normal Inverse Gaussian Processes and the Modelling of Stock Returns* (Research Report 300. Department of Theoretical Statistics, University of Aarhus, 1995).
- [84] R. B. D'Agostino, M. A. Stephens, *Goodness-of-Fit Techniques* (Marcel Dekker, New York, 1986).