

# On the determination of the optimal parameters in the CAM model

Cite as: Chaos 31, 033113 (2021); <https://doi.org/10.1063/5.0032267>

Submitted: 08 October 2020 • Accepted: 01 February 2021 • Published Online: 02 March 2021

 Marco Bianucci and Riccardo Mannella



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Logistic map trajectory distributions: Renormalization-group, entropy, and criticality at the transition to chaos](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 033112 (2021); <https://doi.org/10.1063/5.0040544>

[Multifractal analysis of air and soil temperatures](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 033110 (2021); <https://doi.org/10.1063/5.0029658>

[Transport of coupled particles in rough ratchet driven by Lévy noise](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 033104 (2021); <https://doi.org/10.1063/5.0027116>

APL Machine Learning

Open, quality research for the networking communities

COMING SOON

LEARN MORE



# On the determination of the optimal parameters in the CAM model

Cite as: Chaos 31, 033113 (2021); doi: 10.1063/5.0032267

Submitted: 8 October 2020 · Accepted: 1 February 2021 ·

Published Online: 2 March 2021



View Online



Export Citation



CrossMark

Marco Bianucci<sup>1,a)</sup>  and Riccardo Mannella<sup>2,a)</sup> 

## AFFILIATIONS

<sup>1</sup>Istituto di Scienze Marine, Consiglio Nazionale delle Ricerche (ISMAR - CNR), 19032 Lerici (SP), Italy

<sup>2</sup>Dipartimento di Fisica, Università di Pisa, 56126 Pisa (PI), Italy

<sup>a)</sup>Authors to whom correspondence should be addressed: [marco.bianucci@cnr.it](mailto:marco.bianucci@cnr.it) and [riccardo.mannella@unipi.it](mailto:riccardo.mannella@unipi.it)

## ABSTRACT

In the field of complex systems, it is often possible to arrive at some simple stochastic or chaotic Low Order Models (LOMs) exploiting the time scale separation between leading modes of interest and fast fluctuations. These LOMs, although approximate, might provide interesting qualitative insights regarding some important aspects like the average time between two extreme events. Recently, the simplest example of a LOM with multiplicative noise, namely, a linear system with a linearly state dependent noise [also called correlated additive and multiplicative (CAM) model], has been considered as archetypal for numerous phenomena that present markedly non-Gaussian statistics. We show in this paper that the determination of the parameters of a CAM model from the (few) available data is far from trivial and that the actual most likely parameters might differ substantially from the ones determined directly from a (necessarily limited) short sequence of observations. We illustrate how this problem can be tackled, at least to the extent possible, using an approach that is based on Bayes' theorem. We shall focus on a CAM modeling the El Niño Southern Oscillation but the methodology can be extended to any phenomenon that can be described by a simplified LOM similar to the one examined here and where the available sequence of data is relatively short. We conclude that indeed a Bayesian approach can fix the problem.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0032267>

Since the work of Langevin,<sup>1</sup> simple stochastic differential equations have been used to model the dynamics of some “large scale” variable of interest “ $T$ ” of complex systems. Brownian motion and Ornstein–Uhlenbeck processes are notable examples. If the stochastic process, in the form of a very fast fluctuating force  $f(t)$  (a white noise), enters the model additively, then the central limit theorem applies, ensuring a very low probability for “large” (with respect to the standard deviation of the noise) events. In this case, even a quite short record of observational data could be sufficient to infer, with some accuracy, the values of the model parameters through fitting. Unfortunately, often, like in many large scale geophysical phenomena, the intensity of the fluctuating force is state dependent, namely, it is of the form  $g(T)f(t)$ , where  $f(t)$  is again a white noise and  $g(T)$  is some function of  $T$ . In this case, the noise is said to be “multiplicative,” and the model may exhibit fat tails. In the case of a weak deviation from the classical additive case, the following linear approximation applies  $g(T) \approx 1 + \beta T$ , where  $\beta$  is a parameter. Even for small  $\beta \sim 0.1$ – $0.3$ , large fluctuations of  $T$  are much more likely than in the additive case ( $\beta = 0$ ), but they might fail to show up if a short sequence of observational data

is available and fitted, leading to the inferred parameter's values that differ substantially from the actual most likely ones. Here, we consider the ENSO observation data as a benchmark, and we show that a Bayesian method can be really effective to fix the problem.

## I. INTRODUCTION

Large scale geophysical phenomena are the result of complex ocean–atmosphere–human activities–astronomical interactions, leading to a large diversity in amplitude, spatial pattern, and temporal evolution of events. Some important examples are El Niño Southern Oscillation (ENSO), the North Atlantic Oscillation (NAO), or the Sea Surface Temperature (SST) variability in the Gulf Stream and other strong currents. These phenomena have large global impacts<sup>2</sup> on society; thus, understanding the main features of their dynamics and the character of the recurrence of strong events is of crucial importance. Changing time and space scales, motile bacteria, migrating cells, and Brownian swimmers are crucial

to human life<sup>3–9</sup> as well. They can be modeled as mesoscopic *active Brownian particles* (ABPs), the dynamics of which is the result of the interaction with a very complex and fast environment.<sup>10</sup>

In all these cases, which have a big influence on human activities, and in many others less critical (e.g., laser fluctuations<sup>11,12</sup>), often the complexity and the large time scale separation between the dynamics of the slow variables of interest and the dynamics of the fast fluctuations allow for a description of the mean features of the dynamics of these phenomena by simple few degrees of freedom models or Low Order Models (LOMs, see, e.g., Refs. 13–21 for ENSO and Ref. 10 for ABP). This is the case, of course, for the original Brownian motion, chemical reactions or bistable systems,<sup>22–27</sup> and, in general, for statistical mechanics of matter.

The different parameters that appear in a LOM aim at encapsulating the intricate dynamics of systems, with typically an enormous number of degrees of freedom, in a few quantities. Observation data over different time spans can be used to derive the LOM parameters as a function of time, and predictions can be made regarding events in the future.

Since the 1990s, it has been emphasized that in systems driven by correlated additive and multiplicative (CAM) noises, the correlation between these noises is able to change the steady properties of the systems greatly (see, e.g., Refs. 11 and 28–34). In more recent years, the simplest one of such a class of LOMs, namely, a linear system with a linearly state dependent noise [see Eq. (1)], received some interest because it has been shown that it represents a physically sound simplification of many large scale geophysical phenomena<sup>18,35–39</sup> and mesoscopic ABP.<sup>10</sup> In particular, this LOM is able to catch the main non-Gaussian statistical features observed in these contexts. It has been used (to name just a few studies) to evaluate the impact of correlated noise in an energy depot model<sup>10</sup> or to model mid-latitude ocean–atmosphere local coupling.<sup>35</sup> The same LOM has been exploited in Ref. 37 to discuss the statistics of extreme weather events. In Ref. 18, Eq. (38) is the equivalent Fokker–Planck Equation (FPE), proposed to reproduce some ENSO statistical relevant quantities, like the equilibrium probability density function (PDF) and the timing of the recurrence of strong events. Moreover, it has been also shown, in Ref. 40, that strong ENSO events may not be more predictable than what can be accounted for by a bivariate linear system driven by CAM noise. It is then clear that this model is becoming more and more popular in oceanography (see, for example, Refs. 41, 42, and 26, 27, 43–46).

It is beyond the scope of the present work to discuss the details of how a linear model with CAM noise might emerge in different contexts. What is relevant here is that, for appropriate parameters, this simple LOM is a continuous process realization of Lévy flights;<sup>47</sup> thus, non-Gaussian fat tails in the PDF are in principle present. This clearly might lead to difficulties in the analysis and interpretation of experimental and observational data.

These difficulties have been recognized, for example, in Ref. 35, where the analytical PDF is not known; thus, parameter estimates, with the corresponding errors, are found by comparing the first four moments obtained from data observations and from the model, leading to substantial errors. The same method, involving the first four moments and a deep analysis of the errors, is found in Ref. 37. Although this approach is different from fitting data with the equilibrium analytical PDF or from evaluating directly the FPE

coefficients from their definition (two methods here exploited), any kind of direct fit of the data does not take into account the fact that the available observations span a relatively short time range and, given that the models are characterized by a non-Gaussian tail, in principle, the available observations might have failed to capture the non-Gaussian tail details.<sup>37,48</sup> This would lead to a wrong determination of the model parameters, and, in turn, a wrong determination of the non-Gaussian tail features with the consequences that the frequency of extreme events might be badly estimated. Thus, the following is a very relevant question: how good are the available observational data in order to derive the parameters that appear in such a LOM?

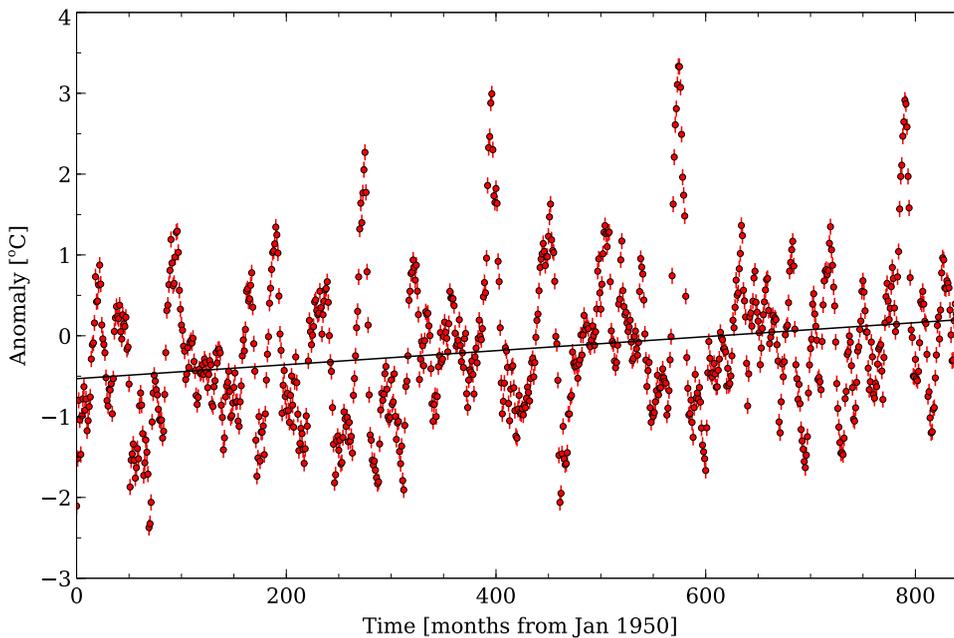
Here, we try to estimate the reliability of the inference procedure for such a model by using a Bayesian approach. Of course, the result is case dependent, and here we shall focus on ENSO events and data.

Because of the regularization effect<sup>49</sup> and the often (but not always) better result with respect to other inference approaches<sup>50</sup> (e.g., aka, frequentist), Bayesian techniques (see Ref. 51) have been widely exploited in high-dimensional estimation problems.<sup>52,53</sup> An application in the field of oceanography, with regard to the parameter estimations of a high-dimensional linear stochastic differential equation (SDE), a model for high-latitude sea surface temperature variability, can be found in Ref. 54. There, Bayes' approach is used to provide optimal probabilistic estimates for a very high-dimensional system, where the conditional PDF, typically called the conditional marginal likelihood, is assumed to be Gaussian. Other notable examples are in the field of acoustic in underwater ocean environment, where Bayesian treatments have been used to find bounds for field parameter estimations (e.g., in order to properly incorporate terms for the fluid flow's transverse speed and normal speed caused by oceanic streams and swells,<sup>53</sup> or to estimate ocean environmental parameters by exploiting wave modeling of acoustic waveguide propagation<sup>55</sup> or to predict the threshold signal-to-noise ratio<sup>56</sup>). In these cases, the focus is on deriving bounds on the mean-square error performance of the estimator for high-dimensional systems where the process dynamics and/or the observation equations are nonlinear. Bayesian techniques are also widely used in the multivariate nonlinear case. A good recent introductory paper is Ref. 57, which also provides algorithms and a few examples.

Our task is less ambitious. We are interested on a simple univariate CAM SDE and the Bayes relation *and surrogate sequences* are invoked to cure, at least as it is possible, the low reliability of the parameter estimates due to both the small number of data available (poor statistics) and the strongly non-Gaussian nature of large fluctuations (rare events matter).

## II. THE DYNAMICAL MODEL AND SOME PRELIMINARY CONSIDERATIONS

The typical time series we will focus on is the directly measured anomaly (deviation from the average) of the monthly Niño3 index, available in the NOAA database, starting from 1950, as shown in Fig. 1. This time series is interesting per se but here it is used mainly as a case study.



**FIG. 1.** The observed monthly Niño3 anomalies starting from 1950 (data from the NOAA<sup>63</sup>). The straight line is the detrending line (see the text).

### A. Detrending the data

Before we can carry out any analysis, we need to address the problem of which average is considered to derive the anomaly series. According to NOAA, the average is evaluated considering the last 30 years' data. We are dealing with a time series, however, which is approximately 70 years long, and it turns out that the average of the whole series is different from zero. Given that we are planning to interpret the anomaly data with a LOM in the form of an oscillator (for which the average anomaly can be expected to be zero, in the limit of a very long sequence) and although we are dealing with a finite time series (for which, on the contrary, the average anomaly could be different from zero), we decided to investigate the possibility of detrending the data, before we proceeded to analyze them (subtracting the average is equivalent to detrending with a constant), in order to bring the dynamics closer to what was expected for an oscillator. We tried low order polynomials, and we eventually opted to detrend the time series with a line,  $q(t) = a + b(t - 1950)$ , with the time  $t$  measured in years,  $a = -0.531 \pm 0.060$  °C and  $b = 0.0104 \pm 0.0015$  °C/year (shown as a straight line in Fig. 1). Detrending with higher order polynomials yielded coefficients with relative errors on the order of or larger than 100%, and we decided to rule them out.  $b \neq 0$  in our detrending line may bear climatic meanings but it is beyond the scope of this paper to investigate its role besides the mere use of a tool to detrend. In the following, we will use  $T$  to refer to the Niño3 anomaly detrended via  $q(t)$  for brevity. Also, the fits mentioned in this paper, including the one used to detrend the data, were done either using MINUIT<sup>58</sup> or NL2SOL,<sup>59,60</sup> in particular, the latter was used in fitting the distributions obtained from the surrogate trajectories (see the following text). When quoted, the errors on the fitted parameters are standard errors, related to the Hessian matrix at the minimum of the  $\chi^2$  (routine MINOS from MINUIT).

### B. The model

As discussed in Sec. I, a relatively simple LOM can be used as a model of ENSO: the purpose of this LOM is to generate time sequences that have statistical properties similar to the ones measured for the observed time series shown in Fig. 1.

The simplified model is described by a simple, one-dimensional stochastic differential equation (SDE),

$$dT = -\gamma T dt + \sqrt{2D}(1 + \beta T) dW, \quad (1)$$

where  $W$  is a standard Wiener process and  $\gamma$  is a frictional coefficient.  $D$  controls the intensity of the stochastic forcing and  $\beta$  is the relative intensity of the multiplicative part of the stochastic forcing.

This SDE has the associated Fokker-Planck equation given by Eq. (38) of Ref. 18, that for self-consistency we rewrite here as

$$\partial_t p(T, t) = [\gamma \partial_T T + \partial_T^2 D(1 + \beta T)^2] p(T, t), \quad (2)$$

which can be solved to find the equilibrium distribution, which reads<sup>20</sup>

$$p_s(T) = \beta f_\mu \left( \frac{\mu - 2}{1 + \beta T} \right) \quad \text{for } T > -1/\beta, \quad (3)$$

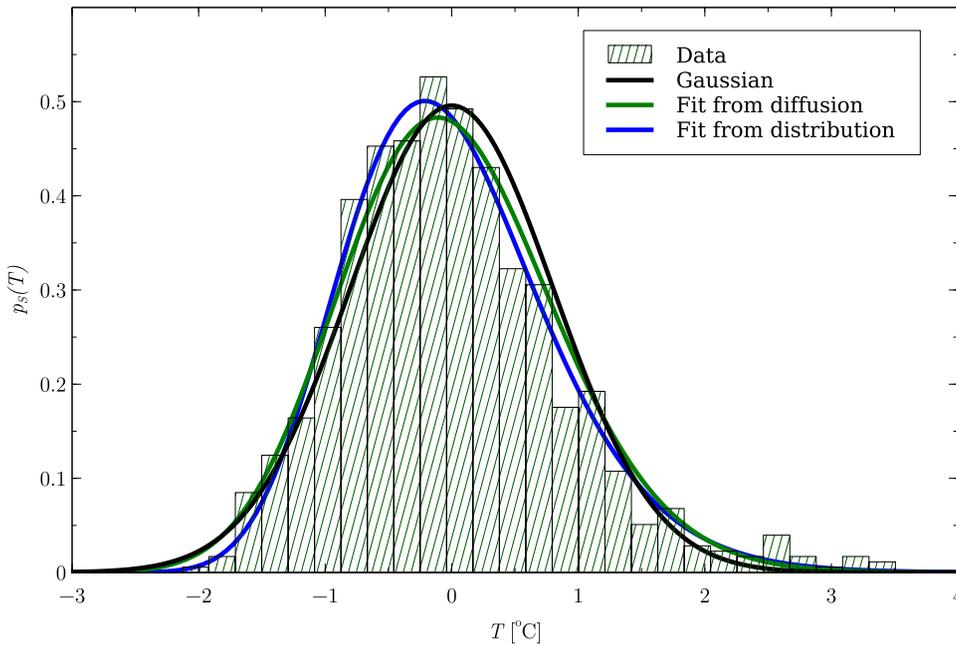
$$p_s(T) = 0 \quad \text{for } T \leq -1/\beta,$$

with

$$\mu \equiv 2 + \frac{\gamma}{D\beta^2}, \quad (4)$$

and where the Gamma-like density function  $f_\mu(x)$  is defined as

$$f_\mu(x) \equiv \frac{1}{(\mu - 2)\Gamma(\mu - 1)} e^{-x} x^{\mu-2}, \quad (5)$$



**FIG. 2.** The distribution of  $T$  obtained from Niño3 anomalies time series of Fig. 1 (histogram) and the best fitting Gaussian function of the data to show that the observed histogram is not Gaussian. By definition, the average of this distribution is zero; hence, the only free parameter is the width of the Gaussian. The curve “Fit from distribution” is a fit to the data with Eq. (3). The curve “Fit from diffusion” is obtained using in Eq. (3) the parameters obtained fitting the experimental drift and diffusion (see the text and Figs. 3 and 4). The reduced  $\chi^2$  of the Gaussian fit is twice the reduced  $\chi^2$  of the fit based on diffusion and four times the reduced  $\chi^2$  of the fit based on the distribution.

in which  $\Gamma(a)$  is the standard complete Gamma function. The dimensionless parameter  $\mu$  must be greater than 3 to have both a finite normalization of the reduced PDF and the existence of (at least) the first two moments. This stationary PDF depends only on the value of the  $\beta$  parameter and on the ratio between  $D$  and  $\gamma$ .

It is easy to check analytically that in the limit  $\beta \rightarrow 0$  and  $D/\gamma$  fixed, the stationary PDF in Eqs. (3)–(5) becomes a standard Gaussian with the same variance  $D/\gamma$ . However, for  $\beta \neq 0$ , the stationary PDF is non-Gaussian. In fact, for large positive  $T$ , it has a “power law” tail that makes possible high fluctuations of positive values of  $T$  (strong El Niño events). From Eq. (4), we have that the  $\mu$  parameter, which controls the decay of this “heavy” tail of the stationary PDF, strongly depends on the value of the  $\beta$  parameter. Although the average value of  $T$  is zero, the maximum of the stationary PDF is found at  $T_{max} = -2/(\beta\mu) \approx -2\beta D/\gamma$ , i.e., for fixed  $D/\gamma$  proportional to  $\beta$ . Notice also that for negative values of  $T$  (La Niña events), the probability of strong events is largely reduced and it has a threshold at  $T_{min} = -1/\beta$ . It is apparent that these non-Gaussian features are shared by the Niño3 data, as seen in Fig. 2. The three parameters  $\gamma$ ,  $D$ , and  $\beta$  should now be estimated from the observations.

A typical approach to infer these parameters is to find them from the observed stochastic trajectory of Fig. 1. The parameter  $\gamma$  can be evaluated from the drift. Looking at Eq. (1), taking the average over  $W$  on both sides after discretizing in time, we have

$$\langle \Delta T \rangle = -\gamma \Delta t \langle T \rangle, \tag{6}$$

where  $\Delta t$  is the time separation between the  $T$  observation and  $\langle \rangle$  stands for  $W$  averages.

The parameters  $D$  and  $\beta$  can be found from the fluctuations. From Eq. (2), we can derive

$$\frac{\partial \langle T^2 \rangle}{\partial t} = -2\gamma \langle T^2 \rangle + 2D(1 + \beta^2 \langle T^2 \rangle), \tag{7}$$

which yields, assuming stationarity and  $D\beta^2 \ll \gamma$ ,

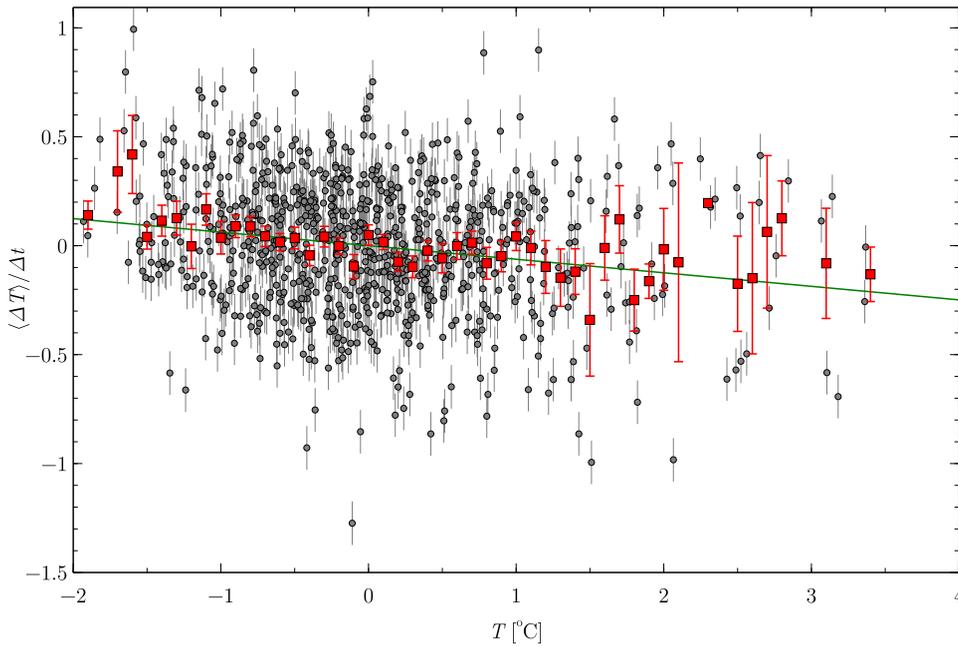
$$\langle T^2 \rangle = \frac{D}{\gamma - D\beta^2} \approx \frac{D}{\gamma}. \tag{8}$$

Equation (1), discretizing in time, recalling that  $(dW)^2 \sim dt$ , squaring the increment and taking the average over  $W$  on both sides, gives

$$\langle (\Delta T)^2 \rangle = 2D(1 + \beta T)^2 \Delta t + O((\Delta t)^2). \tag{9}$$

From the time series of the  $T$  observations, the difference between two subsequent measured temperatures is taken, and fits using Eqs. (6) and (9) are carried out. This approach makes sense as long as  $\Delta t$  (the frequency in time at which  $T$ 's are measured) is much smaller than the other typical time scales in the system ( $1/\gamma$ ,  $\langle T^2 \rangle/D$ ).

The results are shown in Figs. 3 and 4. The fits yield the parameters  $\gamma = 0.0622 \pm 0.0039 \text{ month}^{-1}$ ,  $D = 0.0445 \pm 0.0029 \text{ }^\circ\text{C}^2/\text{month}$ , and  $\beta = 0.078 \pm 0.029 \text{ }^\circ\text{C}^{-1}$ . We can also check whether  $\Delta t$  is sufficiently small:  $\Delta t$  is 1 month, and it should be (much) smaller than  $1/\gamma \approx 16.1$  months which is indeed the case. We used a second approach to evaluate  $\gamma$  from the drift data: we binned the temperature range in bins of  $0.1 \text{ }^\circ\text{C}$  and computed the quantity  $\langle \Delta T / \Delta t \rangle_{T_{bin}}$  in each bin. This is shown as red squares in Fig. 3. The fit to the red square with a function of  $f(T) = a + \beta T$  yields  $a = -0.002 \pm 0.011 \text{ }^\circ\text{C month}^{-1}$  and  $\gamma = -0.051 \pm 0.020 \text{ month}^{-1}$ , consistent with the former approach although with a much larger error.

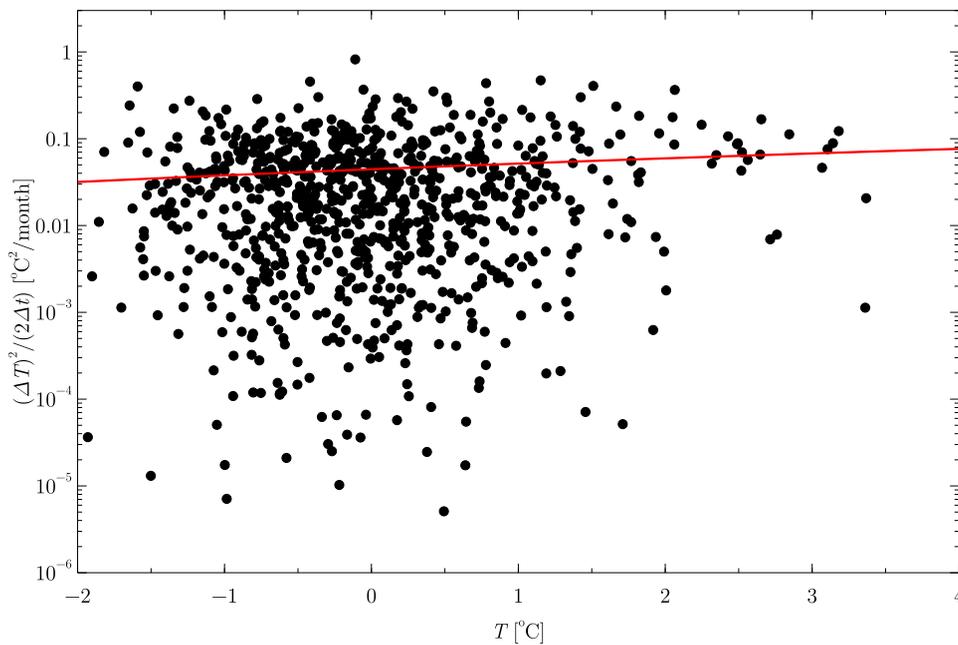


**FIG. 3.** The experimental drift [see Eq. (6)] obtained from the stochastic series of Fig. 1. Gray dots are the quantity  $\Delta T / \Delta t$ , computed for each  $T$ . Red squares are the quantity  $\langle \Delta T / \Delta t \rangle_{Tbin}$  with temperature bins of width  $0.1^\circ\text{C}$  to compute the averages. The straight line is the result of the best fit via the function  $f(T) = -\gamma T$  to the dots. The fit yields  $\gamma = 0.0622 \pm 0.0039 \text{ month}^{-1}$ .

Another possible approach to determine the unknown parameters is to build the experimental distribution of  $T$ ,  $P_{exp}(T)$ , and to fit this distribution using Eq. (3). The advantage of the latter approach is that it can be used even when  $\Delta t$  is not small compared to the other time scales. In this case, we can only determine  $\beta = 0.164 \pm 0.021 \text{ }^\circ\text{C}^{-1}$  and the ratio  $D/\gamma = 0.670 \pm 0.024 \text{ }^\circ\text{C}^2$ ,

obtaining  $D = 0.042 \pm 0.004 \text{ }^\circ\text{C}^2/\text{month}$  using the already determined value of  $\gamma$ .

On consideration of the errors on the parameters given by the different approaches, in the following, we are going to determine  $\gamma$  fitting the “gray” drift points of Fig. 3, and  $\beta, D/\gamma$  fitting the distribution functions.



**FIG. 4.** The experimental diffusion [see Eq. (9)] obtained from the stochastic series of Fig. 1. Dots are the quantity  $(\Delta T)^2 / (2\Delta t)$ , computed for each  $T$ . The solid line is the result of the best fit via the function  $f(T) = D(1 + \beta T)^2$ . The fit yields  $D = 0.0474 \pm 0.0026 \text{ }^\circ\text{C}^2/\text{month}$  and  $\beta = 0.078 \pm 0.029 \text{ }^\circ\text{C}^{-1}$ .

The value of  $\gamma$  found is consistent with what is expected on the grounds of phenomenological considerations, and similar values were obtained in the past ( $\gamma \approx 1/16 \text{ month}^{-1} \approx 0.0625^{17-20}$ ). The value of the quantity  $D/\gamma$  [see Eq. (8)] can be determined reasonably reliably as the variance of the data, obtaining  $D/\gamma \approx 0.8 \text{ }^\circ\text{C}^2$ . From this ratio and  $\gamma$ , we would get  $D \approx 0.046\text{--}0.050 \text{ }^\circ\text{C}^2/\text{month}$  consistent with what we found in the fits. The fit to the distribution gives a relatively small error on the quantity  $D/\gamma$ , incidentally. On the contrary, the value of the  $\beta$  parameter depends on “finer” features of the statistics of the data, such as the tails of the PDF, where the number of available data is small: this parameter can have a sizeable error, which is confirmed by the larger error found when a fit is done. In general, the fit done using the diffusion yields  $D$  with a minimal error but  $\beta$  with a large error  $\approx 30\%$ ; the fit using the distribution yields a much more accurate value of  $\beta$  than the diffusion fit and a very accurate value for  $D/\gamma$ .

The two procedures used to carry out the fit yield similar  $D$  values but different  $\beta$ s. Looking at Fig. 2, this results in two different curves: arguably, the curve with the parameters obtained fitting the distribution is closer to the data than the curve with the parameters obtained fitting via Eq. (9). Also, for large  $T > 0$ , the former curve has a fatter tail than the latter one. In the following, we will drop

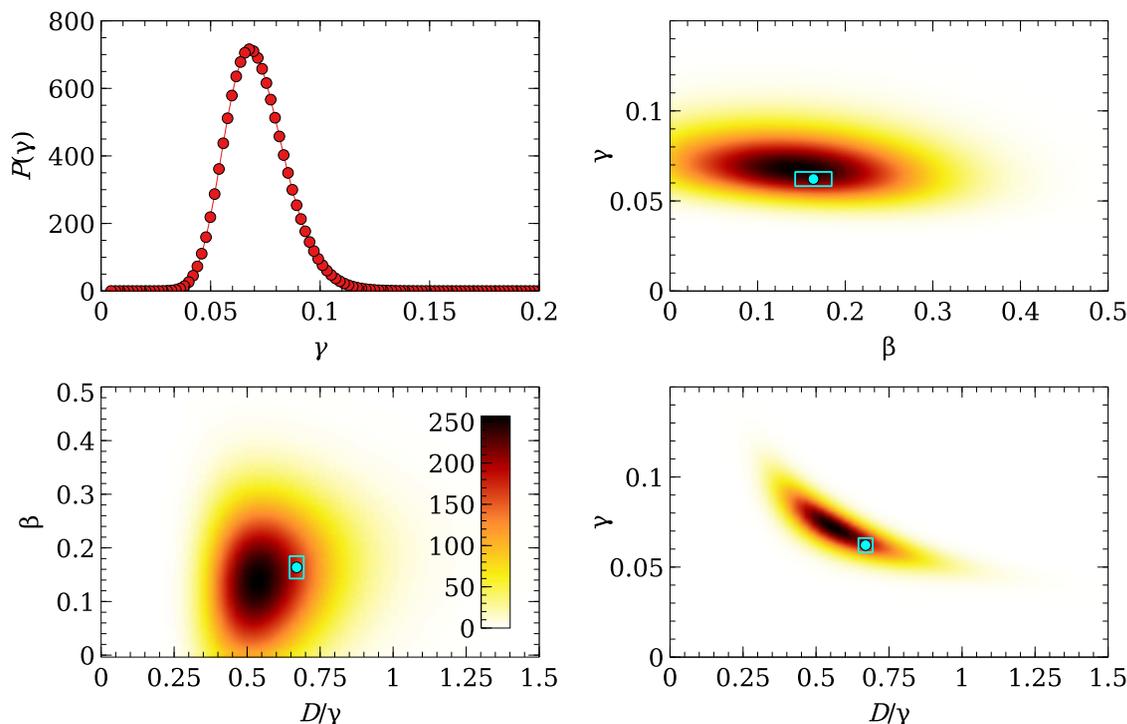
the units, assuming that  $D$ ,  $\gamma$ , and  $\beta$  are measured in  $^\circ\text{C}^2/\text{month}$ ,  $\text{month}^{-1}$ , and  $^\circ\text{C}^{-1}$ , respectively.

### III. THE PROBLEM OF SHORT STOCHASTIC SERIES

Let us assume that the model of Eq. (1) is indeed the “right” model with some known parameters. If we used this model to generate a single stochastic trajectory of the same length and sampling rate of the one shown in Fig. 1, which is then analyzed to derive the model parameters, do we get back the parameters we used for the simulation?

The reason behind this question is that it should be appreciated that the Niño3 time sequence is a little longer than 800 months. This means that we have a time series that is approximately 40 times longer than the typical time scales of the system under study and we may have a problem with the statistical significance of the data set: put it in other words, we are using only approximately 40 statistically independent data points to carry out the fits. In the presence of non-Gaussian tails, the fitted parameters might be very far from the “true” ones, much further than what the errors associated with the fitted parameters might suggest.

In order to address this problem, we carried out digital simulations of Eq. (1) with parameters taken from the fit of the available



**FIG. 5.** The probability distribution densities of parameters found fitting the surrogate trajectories. The “true” parameters used in the simulations are  $\gamma = 0.06217$ ,  $D = 0.04163$ , and  $\beta = 0.1637$  (cyan dot). The small rectangle around the cyan dot shows the error in the parameter determination when we fitted the original El Niño3 data. The top left plate shows the probability distribution of the  $\gamma$  values found fitting the surrogate drifts. The other three plates show the probability distributions as a function of two of the three parameters (shown on the axes) found fitting the surrogates trajectories, integrated over the third parameter (the parameter not shown on the axes). The maxima of the probability distributions are markedly different by the “true” parameters.

Niño3 data and generated stochastic trajectories of the same length of the Niño3 sequence. Then, we considered these trajectories as “surrogate” Niño3 data and we applied to them the same fitting procedure we used for the original Niño3 time series, starting with the detrending and obtaining some fitted parameters. The algorithm used for the numerical integration is the Heun scheme (see Ref. 61 for details) with a time step of  $10^{-2}$  months. Of course, each surrogate Niño3 time series will be different, but hopefully the fitted parameters we find will not differ much from the “true” ones used in the SDE. Unfortunately, it turns out that this is not the case.

Figure 5 summarizes the results obtained assuming as parameters in the SDE the values  $\gamma = 0.06217$ ,  $D = 0.04163$ , and  $\beta = 0.1637$ , simulating  $3 \times 10^6$  surrogate stochastic trajectories, building the corresponding equilibrium distribution, and then fitting them to infer  $\gamma$ ,  $D/\gamma$ , and  $\beta$ . The “true” parameters used in the simulations are marked by a cyan dot, and the small rectangle around the dot shows the experimental error we found fitting the El Niño3 data. We notice that the “true” parameters are fairly distant from the maximum of the distributions of the fitted parameters (darkest areas), and the distributions of the fitted parameters have a rather flat maximum. Also, looking at the probability distribution of the  $\beta$  and  $D/\gamma$  fitted parameters, the region where the distribution of the fitted values has the higher values lies below and to the left of the “true” values. This might have interesting consequences. The observed Niño3 anomalies’ time series of Fig. 1 is a single realization of an ideal model for which the corresponding parameters are unknown. When we fit it and infer the parameters, we obtain one particular set of  $\beta$  and  $D/\gamma$ . But Fig. 5 shows that if we used these fitted parameters to generate surrogate trajectories, which are in turn fitted to infer the parameters, we get a distribution for the parameters, which appears to yield smaller  $D$  and  $\beta$  than the ones used in the model. This implies that the parameters we obtain from the observed Niño3 time series are very likely underestimating the

“correct”  $D/\gamma$  and  $\beta$ . Figure 5 shows also the distribution of the  $\gamma$  values we obtained from the drift fit (top left plate). We notice that this distribution has a half width half height of approximately  $2.5 \times 10^{-2}$ , which is one order of magnitude larger than the error we obtained fitting the El Niño3 data, equal to  $3.9 \times 10^{-3}$ . In other words, the variability in the parameter we intrinsically expect due to the model is much larger than the error in the model parameter estimation.

The following question then arises: what are the “true” parameters we should use in the model, which are able to better characterize the expected dynamics, and how can we determine them?

#### IV. THE DETERMINATION OF THE MOST LIKELY MODEL PARAMETERS

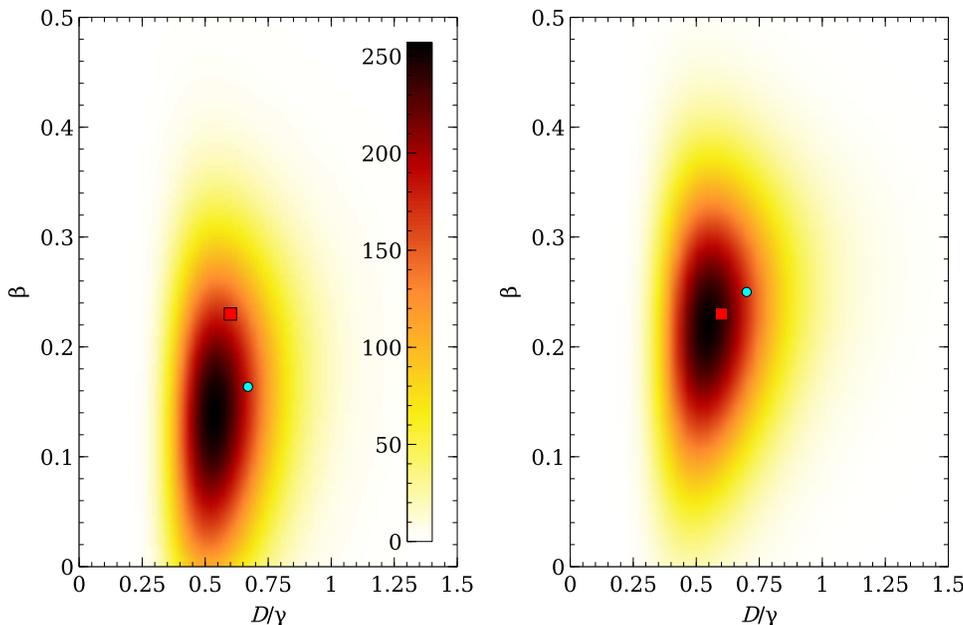
Figure 6 shows the probability distribution density of the fitted parameters obtained assuming two different sets of values (marked by cyan dots) for the parameters  $D$ ,  $\beta$ , and  $\gamma$  in the SDE when the surrogated trajectories were generated. The distributions are clearly different. In particular, the probability of obtaining  $D/\gamma$  and  $\beta$  at the point marked by a red square is very different for the two cases shown.

In the following, for compactness, we are going to use  $\theta \equiv \{D/\gamma, \beta, \gamma\}$ . We denote the parameters used in the simulations by  $\hat{\theta}$  and by  $\theta^*$  a specific set of  $D/\gamma$ ,  $\beta$ , and  $\gamma$  found in the fit. The distributions shown in Fig. 6 can be written as

$$P(D/\gamma, \beta, \gamma) \equiv P(\theta|\hat{\theta}). \quad (10)$$

This means that, assuming some parameter values (i.e.,  $\hat{\theta}$ ) in the SDE simulations, we obtained a distribution as a function of the fitted  $\theta$  in the surrogate data.

This distribution can be used to find the probability  $P(\theta^*)$  that the fitted parameters are found inside a small area centered



**FIG. 6.** The probability distribution density of parameters  $D/\gamma$ ,  $\beta$ , and  $\gamma$ , integrated over  $\gamma$ , found fitting the equilibrium distributions obtained from each surrogate trajectory, for two different sets of parameters  $D/\gamma$ ,  $\beta$ , and  $\gamma$  used in the SDE simulations, shown as cyan dots. On the left, parameters as in Fig. 5, on the right  $\gamma = 0.07$ ,  $D = 0.049$ , and  $\beta = 0.25$ .

on a specific value  $\theta^*$ : with reference to Fig. 6, for example, we may be interested in knowing the probability of finding the values  $D/\gamma$ ,  $\beta$ , and  $\gamma$  marked by a red square: the value of this probability is clearly different in the two cases shown, given that the distributions depend on  $\hat{\theta}$  (cyan dot). Because this probability depends on  $\hat{\theta}$ , really we have that  $P(\theta^*) = P(\theta^*|\hat{\theta})$ , i.e., the probability that we found  $\theta^*$  from the surrogates using  $\hat{\theta}$  in the SDE: as we change  $\hat{\theta}$ , the value of  $P(\theta^*)$  will change: looking at Fig. 6, the probability of  $P(\theta^*)$  of the red square is larger for the plate on the right than for the plate on the left.

Now, if we fixed  $\theta^*$  and changed  $\hat{\theta}$ , we could build a new probability distribution where we plot, as function of  $\hat{\theta}$ , the probability of observing  $\theta^*$ : this would be a distribution of the form  $P(\hat{\theta}) = P(\hat{\theta}|\theta^*)$ , meaning that this is the probability that, having found  $\theta^*$  in the surrogate, we had  $\hat{\theta}$  in the model.

We can find  $P(\hat{\theta}|\theta^*)$  using the Bayes theorem (that is a simple consequence of the definition of conditional probability from the

joint one),

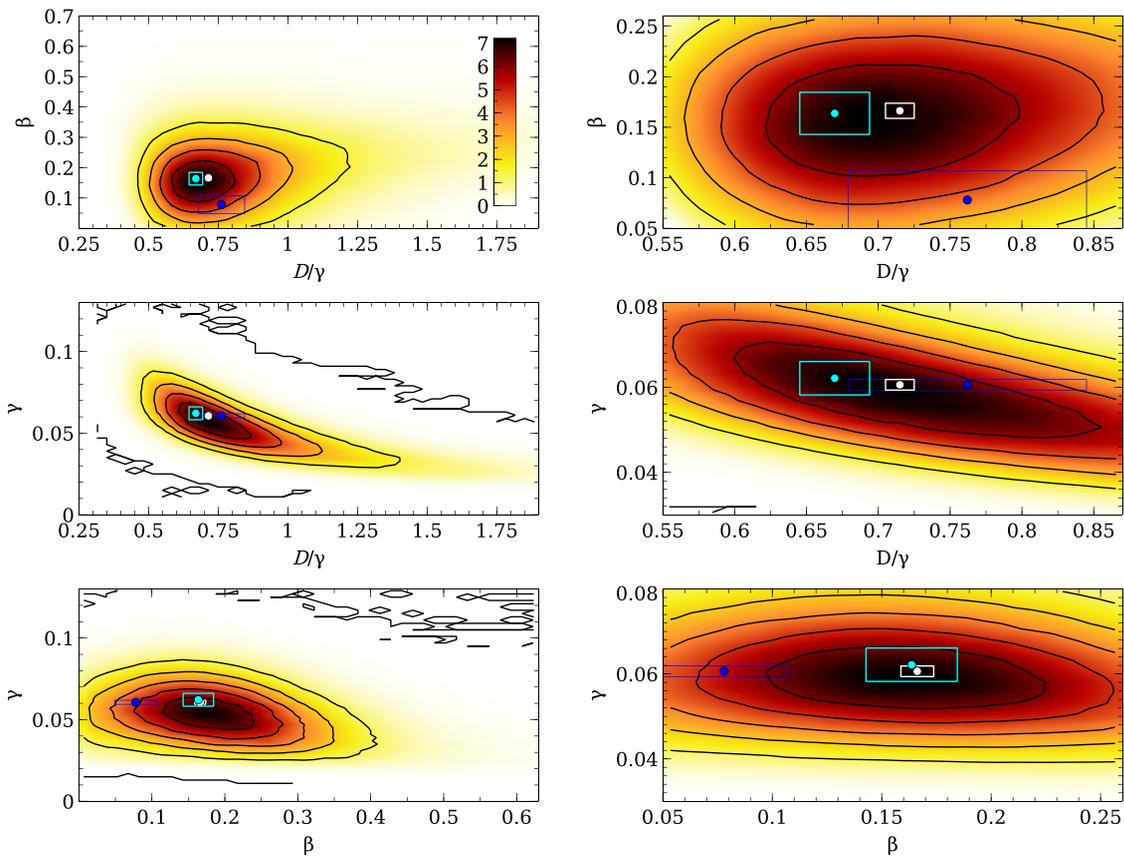
$$P(\hat{\theta}|\theta^*) = \frac{P(\theta^*|\hat{\theta})P(\hat{\theta})}{P(\theta^*)}, \tag{11}$$

where

$$P(\theta^*|\hat{\theta}) = P(\theta|\hat{\theta})\Big|_{\theta=\theta^*}. \tag{12}$$

The distribution  $P(\hat{\theta}|\theta^*)$  is the distribution we need to find the most likely value of the model parameters: it quantifies the probability that we had  $\hat{\theta}$  in the model if we observed  $\theta^*$  in a single trajectory.  $P(\hat{\theta}|\theta^*)$  follows from Eq. (11), if we knew the quantities on the right hand side. It will be clear in the following discussion that we need only to compute  $P(\theta^*|\hat{\theta})$  for different  $\hat{\theta}$ .

The computation of  $P(\theta^*|\hat{\theta})$  is carried out choosing randomly the model parameters  $\hat{\theta}$ , generating a trajectory via numerical integration of the SDE of Eq. (1), fitting it, and checking if the fitted



**FIG. 7.** Left: the probability distribution density of  $P(\hat{\theta}|\theta^*)$  with  $\theta^* = \{D/\gamma = 0.6696, \beta = 0.1637, \gamma = 0.0622\}$  and  $\hat{\theta} = \{D/\gamma, \beta, \gamma\}$  (the axis variables), integrated over  $\gamma$  (top),  $\beta$  (middle), and  $D/\gamma$  (bottom). This is the probability that, for a given set of parameters  $\hat{\theta}$  in the model, we found a trajectory that, after fitting, yielded parameter values close to  $\theta^*$ , shown as a cyan dot, the cyan rectangle marking the acceptance region around  $\theta^*$ . The white dots ( $\theta = \{D/\gamma = 0.715 \pm 0.01, \beta = 0.166 \pm 0.007, \gamma = 0.061 \pm 0.001\}$ ) are the maximum of  $P(\hat{\theta}|\theta^*)$ , with the associate error in the maximum determination; the blue dot is the  $\{D/\gamma, \beta, \gamma\}$  values obtained fitting the experimental drift and diffusion (the blue curve in Fig. 2). Right: zoom around  $\theta^*$ .

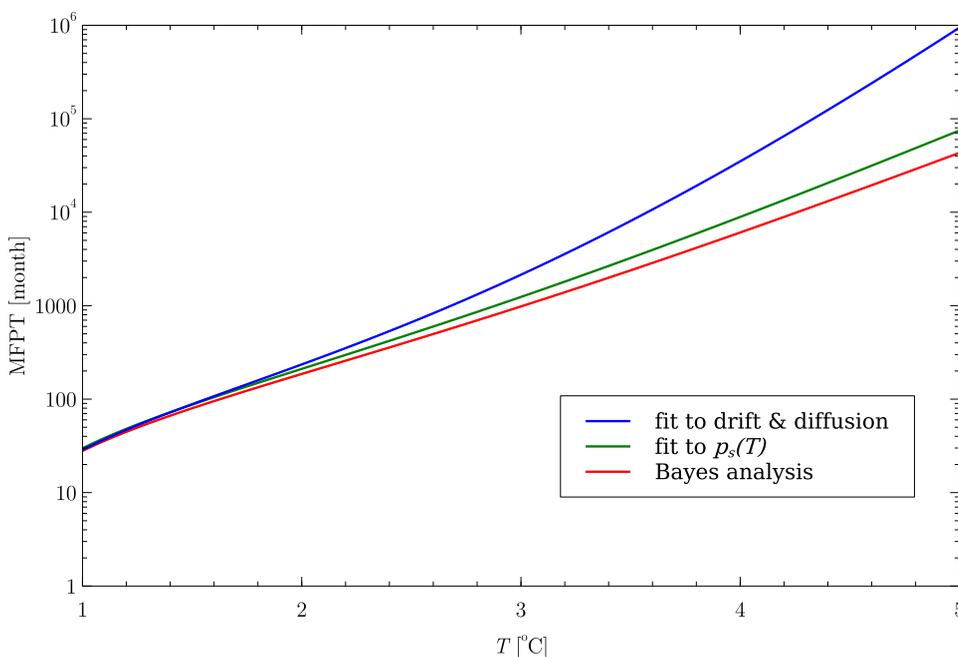
parameters are close to  $\theta^*$ : this loop is repeated until convergence of  $P(\theta^*|\hat{\theta})$  is achieved. Details are given in the Appendix.

The distribution  $P(\hat{\theta}|\theta^*)$  is shown in Fig. 7. The cyan dot marks  $\theta^*$ , the  $D/\gamma$ ,  $\beta$ , and  $\gamma$  parameters obtained from the fit of the Niño3 anomalies time data. The white dot marks the maximum of the distribution, i.e., the  $D/\gamma$ ,  $\beta$ , and  $\gamma$  values ( $\theta = \{D/\gamma = 0.715, \beta = 0.166, \gamma = 0.061\}$ ), which we should insert in the model to maximize the probability of observing  $\theta^* = \{D/\gamma = 0.6696, \beta = 0.1637, \gamma = 0.0622\}$  in a single realization. Note also that the maximum of this distribution is fairly wide. The determination of the white dot parameters is already a main result of this paper.

One of the reasons to build models of the El Niño dynamics is to be able to make predictions on how often we might expect to experience a given large ENSO event. In analytical terms, the answer to this question comes from evaluating the Mean First Passage Time for the SST to reach a given  $T$  threshold from the typical rest state (for details, see Ref. 19).

It is possible, now, to evaluate the MFPT for the model Eq. (2) using different  $\gamma$ ,  $D$ , and  $\beta$ . The results are summarized in Fig. 8.

The MFPT expected on the grounds of the parameters obtained fitting the drift and the diffusion of the known Niño3 anomalies time series is the curve shown in blue. The MFPT expected on the grounds of the parameters obtained fitting  $p_s(T)$  is the curve shown in green. The MFPT expected on the grounds of the model parameters that maximize the probability of finding the observed parameters in the fit of El Niño3 data is the curve shown in red. It should be appreciated that the blue curve yields unrealistic MFPT: according to the blue curve, ENSO events up to  $T = 3$  are expected to happen every 2200 months on average (once in every 180 years). On the other hand, the green and red curves yield more reasonable ENSO events: we get  $T = 3$  events every 1290 and 1000 months (107 and 80 years), respectively.

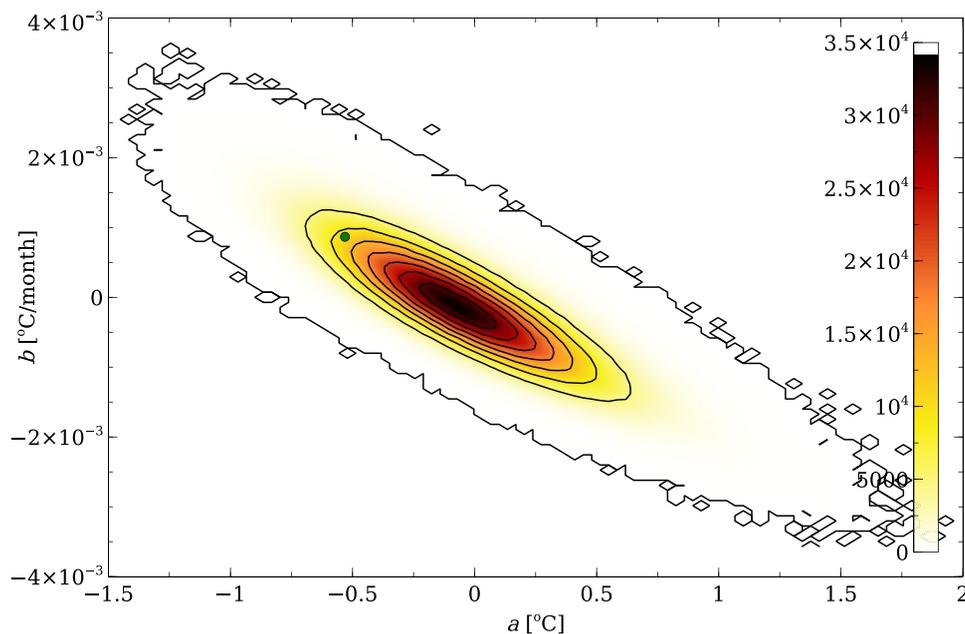


**FIG. 8.** The Mean First Passage Time to reach a given  $T$  threshold. The different curves shown refer to the parameters of the points Fig. 7 with the same color code. The blue curve corresponds to the parameters found fitting drift and diffusion. The green curve corresponds to the parameters found fitting  $p_s(T)$ . The red curve corresponds to the parameter values in the model, which maximize the probability of finding the observed parameters in the fit of the El Niño3 data.

A final word about the detrending we applied to the El Niño3 data. Given the model of Eq. (1), is the observed trend in the El Niño3 data due to poor statistics (generating a short stochastic sequence is bound to lead to a trend) or perhaps it might have a climatic meaning? In the simulations presented in Sec. III, as we mentioned, we needed to detrend the data. To evaluate if the observed trend in Fig. 1 is simply due to poor statistics or to some cause not represented by the stationary LOM of (1), we built, from the surrogate trajectories, the distribution of the observed detrending parameters  $a$  and  $b$ , which is shown in Fig. 9. We note that this distribution function has a peak at the origin, as expected, and that the probability of observing the detrending parameters used for the El Niño3 data, the El Niño3 parameters in a surrogate trajectory is only 1/3 of the probability of observing detrending parameters close to the origin. This means that the trend given by the straight line in Fig. 1 is likely due to external factors and not just to statistics. This is also consistent with the conclusion that in the time range of the used data (1950–2020), the ENSO intensity is increasing in average. A behavior that the running 30 years detrending procedure incorporated in the Niño3 anomalies data from NOAA (<https://psl.noaa.gov/data/climateindices/>) can smooth but not hide; a well known fact (see, e.g., Ref. 62 and references therein) that can be related to global warming.

## V. CONCLUSIONS

In this paper, we showed that attention must be given to the determination of model parameters in the presence of a non-Gaussian process with limited number of data. The reliability of the result is case dependent. The case of ENSO we have considered here is directly related to some recent works<sup>13–21</sup> where the model



**FIG. 9.** Detrending parameters found fitting the surrogate trajectories (see Sec. III) with a function of the form  $q(t) = a + bt$ . The green dot marks the detrending parameters used for the El Niño3 data.

of Eq. (1) has been used to infer, from data, the analytical probability distribution and the recurring time of ENSO events. The main conclusion of these works is that, due to the non-Gaussian nature of the model, the probability to have strong and very strong ENSO is far larger (orders of magnitude) than that obtainable from a Gaussian statistics. In particular,<sup>19</sup> “an important result that emerges [...] is that the average recurring time of very strong ENSO events is of the order of just some tens of years (about the human average lifetime), with important implications for society. This result is strictly related to the multiplicative nature of the perturbation to the ROM, as expressed by the parameter  $\beta$ .” Clearly, because of the simplicity of these LOMs, they provide only a first order estimate that needs to be corroborated in more complex models. However, catching the main large scale features of the physical phenomenon, these simple models should give important information about its statistics. It is then crucial to assess the reliability of such information.

At first sight, it seems that such a reliability is poor. In fact, if we use, for the model, the parameter obtained from the fit to observations, the results are most likely incorrect: when we generate surrogate stochastic ENSO data and fit them, the parameters we infer are different from the ones used in the model integrated to generate the trajectories (Fig. 5).

However, this is not a final conclusion. In fact, we then successfully address the problem of what are the most likely parameters that should be used in the model to maximize the probability of observing the parameters inferred from the real Niño3 time sequence. From these most likely parameters, we get more “optimized” models that give the results of Fig. 8. The conclusion is surprising because it essentially confirms that the non-Gaussian feature of the model leads to the aforementioned finding on recurrence time for intermediate to very strong ENSO events. We also concluded that additional care is needed in fitting: the parameters obtained by a fit to the data (drift and diffusion) yield unrealistic MFPT, whereas if

the probability distribution is built first and then fitted taking into account the specific model, the MFPT obtained appear to be more realistic.

Finally, we would like to add that the additional step of finding the optimal parameters via a Bayesian analysis should be regarded as a necessary step when models implying non-Gaussian statistics are involved.

## AUTHORS' CONTRIBUTIONS

All authors contributed equally to this work.

## ACKNOWLEDGMENTS

Discussions about Bayes theorem with Walter Del Pozzo are warmly acknowledged. The authors thank the referees for their many valuable comments and recommendations that helped improve the quality of the paper.

## APPENDIX: NUMERICAL DETAILS

We describe the numerical procedure in more detail. In the following, we assume that  $\theta^* = \{D/\gamma = 0.6696, \beta = 0.1637, \gamma = 0.0622\}$ , shown as cyan dots in Fig. 7.

We first worked out the region of the model parameters  $\hat{\theta}$ , which we need to sample. Via numerical integration of the SDE, we established the maximal values of  $D/\gamma$ ,  $\beta$ , and  $\gamma$  which could yield parameters not too far from  $\theta^*$  after fitting. Then, we discretized this region, taking steps of size 0.03 along  $D/\gamma$ , steps of size 0.015 along  $\beta$ , and  $2 \times 10^{-3}$  along  $\gamma$ . This means that we covered this region with  $48 \times 42 \times 60$  small boxes, of side  $0.03 \times 0.015 \times 0.002$ . To apply Eq. (11), however, we should have a distribution of model parameters for our stochastic trajectories. So, within each of the small boxes, we generated  $4 \times 10^4$  different model parameters

according to a flat distribution for  $D/\gamma$ ,  $\beta$ , and  $\gamma$ . This defines  $P(\hat{\theta})$ . For each model parameter, a stochastic trajectory, 830 months long, was numerically integrated (the integration time step was taken to be  $10^{-2}$  months); it was detrended, and the drift and the distribution  $p_s(T)$  (Fig. 2) were built and fitted to obtain the value of  $\theta$  for that particular trajectory. For each small box, we ended up with  $4 \times 10^4$   $\theta$  values (in total, we generated  $48 \times 42 \times 60 \times 4 \times 10^4$  stochastic trajectories). We selected a small region around  $\theta^*$  and counted for how many trajectories the fitted parameters “hit” this region: the fraction of hits gave  $P(\theta^*|\hat{\theta})$ . The number of hits depends on how wide is the region around  $\theta^*$  we select to define when we have a hit. Given that the fit of Niño3 anomalies yields parameters with some errors, the region was assumed to be equal to the error in the fit, and it is shown as a cyan rectangle in figure. The last unknown quantity on the rhs in Eq. (11) is  $P(\theta^*)$ : this quantity is proportional to the total number of trajectories hitting the target rectangle divided by the total number of trajectories used in the simulations. Given that we are only interested in relatively values for  $P(\hat{\theta}|\theta^*)$ , we disregarded it.

## DATA AVAILABILITY

Raw ENSO data for ENSO used throughout this work were generated at the NOAA large scale facility at <https://psl.noaa.gov/data/climateindices/>, Ref. 63. Derived data supporting the findings of this study are available from the corresponding authors upon reasonable request.

## REFERENCES

- <sup>1</sup>D. S. Lemons and A. Gythiel, “Paul Langevin’s 1908 paper ‘On the theory of Brownian motion’ [sur la théorie du mouvement brownien,” *C. R. Acad. Sci. (Paris)* **146**, 530–533 (1908)],” *Am. J. Phys.* **65**, 1079–1081 (1997).
- <sup>2</sup>S. Curtis, R. F. Adler, G. J. Huffman, G. Gu, D. T. Bolvin, and E. J. Nelkin, “Comments on ‘El Niño: Catastrophe or opportunity’,” *J. Clim.* **19**, 6439–6442 (2006).
- <sup>3</sup>F. J. Sevilla and M. Sandoval, “Smoluchowski diffusion equation for active Brownian swimmers,” *Phys. Rev. E* **91**, 052150 (2015).
- <sup>4</sup>M. L. Berre, Y.-J. Liu, J. Hu, P. Maiuri, O. Bénichou, R. Voituriez, Y. Chen, and M. Piel, “Geometric friction directs cell migration,” *Phys. Rev. Lett.* **111**, 198101 (2013).
- <sup>5</sup>O. Chepizhko and F. Peruani, “Diffusion, subdiffusion, and trapping of active particles in heterogeneous media,” *Phys. Rev. Lett.* **111**, 160604 (2013).
- <sup>6</sup>H. Wada, D. Nakane, and H.-Y. Chen, “Bidirectional bacterial gliding motility powered by the collective transport of cell surface proteins,” *Phys. Rev. Lett.* **111**, 248102 (2013).
- <sup>7</sup>K.-J.-B. Lee, M. Ann, and P. J. Park, “A modified active Brownian dynamics model using asymmetric energy conversion and its application to cargo transport by multi-motors,” *Biophys. J.* **104**, 323A (2013).
- <sup>8</sup>L. Schimansky-Geier, M. Mieth, H. Rosé, and H. Malchow, “Structure formation by active Brownian particles,” *Phys. Lett. A* **207**, 140–146 (1995).
- <sup>9</sup>O. Steuernagel, W. Ebeling, and V. Calenbuhr, “An elementary model for directed active motion,” *Chaos Solitons Fractals* **4**, 1917–1930 (1994).
- <sup>10</sup>C. Zeng, J. Zeng, F. Liu, and H. Wang, “Impact of correlated noise in an energy depot model,” *Sci. Rep.* **6**, 19591 (2016).
- <sup>11</sup>S. Zhu, A. W. Yu, and R. Roy, “Statistical fluctuations in laser transients,” *Phys. Rev. A* **34**, 4333–4347 (1986).
- <sup>12</sup>M. Lax, “Theory of laser noise,” *Proc. SPIE* **1376**, 2–20 (1991).
- <sup>13</sup>F.-F. Jin, “An equatorial ocean recharge paradigm for ENSO. Part II: A stripped-down coupled model,” *J. Atmos. Sci.* **54**, 830–847 (1997).
- <sup>14</sup>F.-F. Jin, “An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model,” *J. Atmos. Sci.* **54**, 811–829 (1997).
- <sup>15</sup>S.-H. Im, S.-I. An, S. T. Kim, and F.-F. Jin, “Feedback processes responsible for El Niño-La Niña amplitude asymmetry,” *Geophys. Res. Lett.* **42**, 5556–5563, <https://doi.org/10.1002/2015GL064853> (2015).
- <sup>16</sup>H.-L. Ren and F.-F. Jin, “Recharge oscillator mechanisms in two types of ENSO,” *J. Clim.* **26**, 6506–6523 (2013).
- <sup>17</sup>G. Burgers, F.-F. Jin, and G. J. van Oldenborgh, “The simplest ENSO recharge oscillator,” *Geophys. Res. Lett.* **32**, L13706, <https://doi.org/10.1029/2005GL022951> (2005).
- <sup>18</sup>M. Bianucci, A. Capotondi, R. Mannella, and S. Merlino, “Linear or nonlinear modeling for ENSO dynamics?,” *Atmosphere* **9**, 435 (2018).
- <sup>19</sup>M. Bianucci, A. Capotondi, S. Merlino, and R. Mannella, “Estimate of the average timing for strong El Niño events using the recharge oscillator model with a multiplicative perturbation,” *Chaos* **28**, 103118 (2018).
- <sup>20</sup>M. Bianucci, “Analytical probability density function for the statistics of the ENSO phenomenon: Asymmetry and power law tail,” *Geophys. Res. Lett.* **43**, 386–394, <https://doi.org/10.1002/2015GL066772> (2016).
- <sup>21</sup>A. Capotondi, “El Niño southern oscillation ocean dynamics: Simulation by coupled general circulation models,” in *Climate Dynamics: Why Does Climate Vary?* (American Geophysical Union, 2013), pp. 105–122.
- <sup>22</sup>F. Oliveira, “Reaction rate theory for non-Markovian systems,” *Physica A* **257**, 128–135 (1998).
- <sup>23</sup>T. Fonseca, P. Grigolini, and D. Pareo, “Classical dynamics of a coupled double well oscillator in condensed media. III. The constraint of detailed balance and its effects on chemical reaction process,” *J. Chem. Phys.* **83**, 1039–1048 (1985).
- <sup>24</sup>M. Bianucci and P. Grigolini, “Nonlinear and non-Markovian fluctuation-dissipation processes: A Fokker-Planck treatment,” *J. Chem. Phys.* **96**, 6138–6148 (1992).
- <sup>25</sup>M. Bianucci, P. Grigolini, and V. Palleschi, “Beyond the linear approximations of the conventional approaches to the theory of chemical relaxation,” *J. Chem. Phys.* **92**, 3427–3441 (1990).
- <sup>26</sup>P. Sura, “On non-Gaussian SST variability in the Gulf Stream and other strong currents,” *Ocean Dyn.* **60**, 155–170 (2010).
- <sup>27</sup>P. Sura, “A general perspective of extreme events in weather and climate,” *Atmos. Res.* **101**, 1–21 (2011).
- <sup>28</sup>S. Zhu, “Steady-state analysis of a single-mode laser with correlations between additive and multiplicative noise,” *Phys. Rev. A* **47**, 2405–2408 (1993).
- <sup>29</sup>W. Da-jin, C. Li, and Y. Bo, “Probability evolution and mean first-passage time for multidimensional non-Markovian processes,” *Commun. Theor. Phys.* **11**, 379 (1989).
- <sup>30</sup>M. Dong-cheng, X. Guang-zhong, C. Li, and W. Da-jin, “Effects of cross-correlated noises on a single-mode laser model: Steady state analysis,” *Acta Phys. Sin.* **8**, 174 (1999).
- <sup>31</sup>J. Wang and S. Zhu, “Dynamical properties of a laser with correlations between additive and multiplicative noise,” *Phys. Lett. A* **207**, 47–52 (1995).
- <sup>32</sup>H.-X. Fu, L. Cao, and D.-J. Wu, “Effect of correlation between additive and multiplicative noises on the activation from a double well,” *Phys. Rev. E* **59**, R6235–R6238 (1999).
- <sup>33</sup>P. Zhu and Y. J. Zhu, “Statistical properties of intensity fluctuation of saturation laser model driven by cross-correlated additive and multiplicative noises,” *Int. J. Mod. Phys. B* **24**, 2175–2188 (2010).
- <sup>34</sup>M. F. Daqaq, “Transduction of a bistable inductive generator driven by white and exponentially correlated Gaussian noise,” *J. Sound Vib.* **330**, 2554–2564 (2011).
- <sup>35</sup>C. Martínez-Villalobos, D. J. Vimont, C. Penland, M. Newman, and J. D. Neelin, “Calculating state-dependent noise in a linear inverse model framework,” *J. Atmos. Sci.* **75**, 479–496 (2018).
- <sup>36</sup>P. Sura and P. D. Sardeshmukh, “A global view of air-sea thermal coupling and related non-Gaussian SST variability,” *Atmos. Res.* **94**, 140–149 (2009).
- <sup>37</sup>P. D. Sardeshmukh, G. P. Compo, and C. Penland, “Need for caution in interpreting extreme weather statistics,” *J. Clim.* **28**, 9166–9187 (2015).
- <sup>38</sup>P. D. Sardeshmukh and C. Penland, “Understanding the distinctively skewed and heavy tailed character of atmospheric and oceanic probability distributions,” *Chaos* **25**, 036410 (2015).

- <sup>39</sup>P. Sura and P. D. Sardeshmukh, "A global view of non-Gaussian SST variability," *J. Phys. Oceanogr.* **38**, 639–647 (2008).
- <sup>40</sup>C. Martinez-Villalobos, M. Newman, D. J. Vimont, C. Penland, and J. David Neelin, "Observed El Niño-La Niña asymmetry in a linear model," *Geophys. Res. Lett.* **46**, 9909–9919, <https://doi.org/10.1002/2015GL066772> (2019).
- <sup>41</sup>J. Berner, P. D. Sardeshmukh, and H. M. Christensen, "On the dynamical mechanisms governing El Niño–southern oscillation irregularity," *J. Clim.* **31**, 8401–8419 (2018).
- <sup>42</sup>M. Newman, A. T. Wittenberg, L. Cheng, G. P. Compo, and C. A. Smith, "The extreme 2015/16 El Niño, in the context of historical climate variability and change," *Bull. Am. Meteorol. Soc.* **99**, S16–S20 (2018).
- <sup>43</sup>A. J. Majda, C. Franzke, and D. Crommelin, "Normal forms for reduced stochastic climate models," *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3649–3653 (2009).
- <sup>44</sup>A. H. Monahan, "Bispectral unfolding of the skewness of correlated additive and multiplicative noise processes," *Chaos* **30**, 023126 (2020).
- <sup>45</sup>S. D. Souza-Machado, A. Tangborn, P. Sura, C. Hepplewhite, and L. L. Strow, "Non-Gaussian analysis of observations from the atmospheric infrared sounder compared with era and merra reanalyses," *J. Appl. Meteorol. Climatol.* **56**, 1463–1481 (2017).
- <sup>46</sup>D. Castellana, H. A. Dijkstra, and F. W. Wubs, "A statistical significance test for sea-level variability," *Dyn. Stat. Clim. Syst.* **3**, dzy008 (2018).
- <sup>47</sup>I. A. Lubashevsky, A. Heuer, R. Friedrich, and R. Usmanov, "Continuous Markovian model for Lévy random walks with superdiffusive and superballistic regimes," *Eur. Phys. J. B* **78**, 207–216 (2010).
- <sup>48</sup>C. Penland and P. D. Sardeshmukh, "Alternative interpretations of power-law distributions found in nature," *Chaos* **22**, 023119 (2012).
- <sup>49</sup>C. R. Vogel, *Computational Methods for Inverse Problems*, Frontiers in Applied Mathematics (Society for Industrial and Applied Mathematics, 2002), pp. i–xvi.
- <sup>50</sup>F. J. Samaniego, *A Comparison of the Bayesian and Frequentist Approaches to Estimation* (Springer, New York, 2010).
- <sup>51</sup>H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking* (Wiley-IEEE Press, 2007).
- <sup>52</sup>G. B. Smith, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," in *Readings in Computer Vision* (Elsevier, 1987), pp. 562–563.
- <sup>53</sup>P. Tichavsky and K. T. Wong, "Quasi-fluid-mechanics-based quasi-Bayesian Cramer/spl acute/r-Rao bounds for deformed towed-array direction finding," *IEEE Trans. Signal Process.* **52**, 36–47 (2004).
- <sup>54</sup>D. Foster, D. Comeau, and N. M. Urban, "A Bayesian approach to regional decadal predictability: Sparse parameter estimation in high-dimensional linear inverse models of high-latitude sea surface temperature variability," *J. Clim.* **33**, 6065–6081 (2020).
- <sup>55</sup>H. L. Van Trees and K. L. Bell, "Bayesian bounds for matched field parameter estimation," in *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking* (John Wiley and Sons Inc., 2007), pp. 516–527.
- <sup>56</sup>J. Tabrikian and J. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Trans. Signal Process.* **47**, 2917–2927 (1999).
- <sup>57</sup>T. Stankovski, A. Duggento, P. V. E. McClintock, and A. Stefanovska, "A tutorial on time-evolving dynamical Bayesian inference," *Eur. Phys. J. Spec. Top.* **223**, 2685–2703 (2014).
- <sup>58</sup>F. James and M. Winkler, *Minuit User's Guide*, 2004, <https://root.cern.ch/download/minuit.pdf>.
- <sup>59</sup>J. E. Dennis, D. M. Gay, and R. E. Welsch, "An adaptive nonlinear least-squares algorithm," *ACM Trans. Math. Softw.* **7**, 348–368 (1981).
- <sup>60</sup>J. E. Dennis, D. M. Gay, and R. E. Welsch, "Algorithm 573: NL2SOL—An adaptive nonlinear least-squares algorithm [E4]," *ACM Trans. Math. Softw.* **7**, 369–383 (1981).
- <sup>61</sup>R. Mannella, "Integration of stochastic differential equations on a computer," *Int. J. Mod. Phys. C* **13**, 1177–1194 (2002).
- <sup>62</sup>T. Haszpra, M. Herein, and T. Bódai, "Investigating ENSO and its teleconnections under climate change in an ensemble view—A new perspective," *Earth Syst. Dyn.* **11**, 267–280 (2020).
- <sup>63</sup>See the NOAA "Climate Indices: Monthly Atmospheric and Ocean Time Series," at <https://psl.noaa.gov/data/climateindices>, where we have selected, from the database, the monthly averaged Nino3 data set.