

## Critical slowing down in dynamical systems driven by nonstationary correlated noise

Christopher Boettner<sup>1</sup> and Niklas Boers

*Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany;*

*Technical University of Munich, School of Engineering and Design, Earth System Modelling, 85748 Garching, Germany;*

*and University of Exeter, Global Systems Institute and Department of Mathematics, EX4 4PY Exeter, United Kingdom*



(Received 23 June 2021; accepted 20 January 2022; published 28 March 2022)

Precursor signals for bifurcation-induced critical transitions have recently gained interest across many research fields. Common indicators, including variance and autocorrelation increases, rely on the dynamical system being driven by white noise. Here, we show that these metrics raise false alarms for systems driven by time-correlated noise, if the autocorrelation of the noise process increases with time. We introduce an indicator for systems driven by nonstationary short-term memory noise, and show that this indicator performs well in situations where the classical methods fail.

DOI: [10.1103/PhysRevResearch.4.013230](https://doi.org/10.1103/PhysRevResearch.4.013230)

### I. INTRODUCTION

Many theoretical, experimental, and real-world systems undergo abrupt shifts from one dynamical regime to another. Anticipating or predicting such regime shifts is of vital importance in many instances, such as for major Earth system components or ecosystems, but also in physiology, neurology, or finance [1–6]. Such regime shifts can often be described in terms of *critical transitions* of low-dimensional nonlinear dynamical systems, induced by codimension-1 bifurcations.

The theory behind this type of transition goes back to the fluctuation-dissipation theorem [7,8] and is understood well [9]. Transitions triggered by codimension-1 bifurcations in stochastic dynamical systems with additive noise forcing are associated with a characteristic widening of the underlying basin of attraction on the way to the bifurcation [9,10]. This leads to a weakening of the restoring force to perturbations, which causes the phenomenon of *critical slowing down* (CSD) [11–13]. CSD changes the statistical properties of the system and, in particular, leads to an increase of the variance and lag-1 autocorrelation (AC1) before the critical threshold is reached. A data-driven approach to identify such statistical precursor signals for forthcoming critical transitions thus relies on measuring the variance and AC1 of a time series encoding the dynamics of the system in sliding windows, and searching for consistent increases in both indicators. This specific method has already found a number of applications to models and real-world data [14–17].

Bifurcation-induced transitions are also a widely studied mechanism for sudden climate transitions [18–21].

Precursor signals for critical transitions in terms of rising variance and AC1 have been identified in ecological and geological time series, controlled biological experiments, and a multitude of model simulations of different systems [10,12,22]. Besides these classical CSD indicators, which are most widely used in practical applications, a number of additional indicators have been proposed, including changes in skewness [9], diverging susceptibility [23], mean exit time [24], and increases in complexity of ARMA models [25]. In recent years there have been considerable efforts to improve the performance of these statistical methods by reducing false and raising true positive detection rates. Existing work has mainly focused on systems driven by uncorrelated (white) noise, which makes the methods prone to failure if this assumption is not fulfilled—which is likely in the climate system [26]—although some work has been done to construct methods for systems with correlated noise using detrended fluctuation analysis [27,28].

Here, we propose an analytical early-warning indicator for systems driven by noise with exponentially decaying autocorrelation structure. We show that the classical CSD indicators produce false alarms if the autocorrelation of the driving noise increases, while the indicator proposed here remains reliable in such cases. We thereby provide the analytical basis and confirmation for recent work that followed a purely data-driven approach to account for changing temporal correlation in the driving noise [29]. The method introduced here also suits itself to generalizations of the noise model and constitutes a flexible approach for situations when the classical CSD indicators are not applicable.

The structure of this paper is as follows: We first describe the occurrence of the classical CSD indicators for traditional one-dimensional systems driven by white noise. Thereafter, we show that the classical indicators cannot accurately predict transitions in cases of correlated noise with increasing autocorrelation, as described by a nonstationary AR(1) process. We then propose an early-warning indicator based on estimating the varying AR(1) parameters. This estimator reduces

\*boettner@pik-potsdam.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

to the AC1 in the white-noise case. We then demonstrate the performance of our method, and compare it to existing alternatives, on simple mathematical models.

## II. THEORY

### A. Classical critical slowing down

The occurrence of CSD can be formally explained on the basis of stochastic dynamics. Assume that the system is well described by a deterministic function  $f$  and driven by high-frequency forcing that can be represented by additive stochastic noise. The typical noise model assumed in the context of CSD is additive, stationary white noise with variance  $\sigma^2$ , so that the dynamical equation reads

$$dx = f(x, r(t), t)dt + \sigma dW_t, \quad (1)$$

for system states  $x$ , control parameter  $r$ , and a Wiener process  $W_t$ . If the system is close to equilibrium, locally it suffices to consider the linearized dynamics

$$dx \approx f'(x^*, r(t), t)(x - x^*)dt + \sigma dW_t, \quad (2)$$

since  $f(x^*) = 0$  by definition. Of course, the equilibrium position  $x^*$  can itself change in an open system, hence  $x^* = x^*(r(t), t)$ . By performing a coordinate change  $x \rightarrow (x - x^*)$  this part of the dynamics can be absorbed. Under the assumption that the control parameter changes sufficiently slowly, so that the system state can track the equilibrium position (quasistatic assumption), the linear approximation holds after the transformation. With  $\lambda := -f'(x^*, r(t), t)$ , Eq. (2) takes the form

$$\dot{x} = -\lambda(t)x + \sigma dW_t, \quad (3)$$

where the negative sign causes  $\lambda$  to be a positive number for stable solutions. For constant (or slowly varying)  $\lambda$ , this equation defines the Ornstein-Uhlenbeck process; in its stationary state, the parameter  $\lambda$  is related to the variance  $\text{Var}(x_t)$  and the autocorrelation  $\text{AC}(x_s, x_t)$  by [30]

$$\text{Var}(x_t) = \frac{\sigma^2}{2\lambda}, \quad \text{AC}(x_s, x_t) = e^{-\lambda|t-s|}. \quad (4)$$

In this framework, the loss of (linear) stability when approaching the bifurcation manifests as a decrease in  $\lambda$ ; by Eq. (4) this leads to the increases in variance and AC1 that are widely used to identify CSD in a given time series. Note that it has also previously been proposed to use the restoring rate  $\lambda$  as a parametric indicator for an approaching bifurcation [29,31,32], while some nonparametric approaches have been devised to analyze the full nonlinear system [33,34]. In practice, the statistical properties are usually calculated on discrete time series data. Sampling a realization of the Ornstein-Uhlenbeck process  $X_t := x_{n\Delta t}$ ,  $n \in \mathbb{Z}$ , in time steps of size  $\Delta t = \text{const}$ , leads to a discrete-time process

$$X_{t+1} = \varphi X_t + \epsilon_t, \quad X_0 = x_0, \quad \varphi := e^{-\lambda\Delta t}, \\ \epsilon_t \sim \mathcal{N}(0, \tilde{\sigma}^2), \quad \tilde{\sigma}^2 := \frac{1}{2\lambda}\sigma^2(1 - e^{-2\lambda|\Delta t|}), \quad (5)$$

which is a first-order autoregressive (AR1) model. In the stationary case, Eq. (4) implies that

$$\text{Var}(X_t) = \frac{\tilde{\sigma}^2}{1 - \varphi(\lambda)^2}, \quad \text{AC}(X_{t+1}, X_t) = \varphi(\lambda) \quad (6)$$

(see [35] and Supplemental Material [36]). Therefore, a decrease in  $\lambda$ , which leads to an increase of  $\varphi$ , in turn leads to an increase in variance and AC1 also when estimated from a discrete time series. The above derivation already reveals some limitations of the classical approach based on increasing variance and AC1. In particular, an increase of the variance may not only be caused by a decrease in the restoring rate  $\lambda$ , but also by an increase in  $\sigma$ . An increase in both variance and AC1 is therefore necessary for reliable detection of critical slowing down, even if the assumptions on white additive noise are justified; this has been prominently discussed in [13]. In cases where the assumption of white noise cannot be justified, however, further theoretical development is needed, as we will show in the following.

### B. Critical slowing down in systems driven by correlated noise

In addition to assuming  $\sigma = \text{const}$ , the classical CSD indicators in terms of rising variance and AC1 rely on the assumption that the autocorrelation of the process only increases due to a weakening in the restoring force. A natural extension to the process is a model in which the autocorrelation may change independently of critical slowing down. In fact, for any real-world natural system, it would be very hard to justify that the driving noise is white, and correlated noise structures should be considered much more realistic. A straightforward generalization of Eq. (3) is thus

$$dx_t = -\lambda x_t dt + dV_t, \quad dV_t = -\mu V_t dt + \sigma dW_t, \quad (7)$$

an Ornstein-Uhlenbeck process driven by another Ornstein-Uhlenbeck process [37]. By increasing the autocorrelation parameter  $\mu$  of the noise term, one directly manipulates both the autocorrelation and the variance of the process  $x_t$ . Since the noise process is described in itself by an Ornstein-Uhlenbeck process, the covariance of the noise term decreases exponentially in time. This model can thus be considered as a linear model driven by a short-term memory process, a type of noise that is very often considered for (paleo-)climate records [38]. A discrete analog for this model is

$$X_t = \varphi X_{t-1} + V_t, \quad V_t = \rho V_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \tilde{\sigma}^2), \quad (8)$$

which has the following statistical properties (see Supplemental Material [36]):

$$\text{Var}(X_t) = \frac{\sigma^2}{(1 - \varphi^2)(1 - \rho^2)} \left[ \frac{1 + \varphi\rho}{1 - \varphi\rho} \right], \quad (9)$$

$$\text{AC}(X_{t+1}, X_t) = \varphi + \frac{\rho(1 - \varphi^2)}{1 + \varphi\rho}. \quad (10)$$

From these equations, it is clear that a loss of stability (and increase in  $\varphi$ ) still leads to the increases in variance and AC1, just as for the case of white-noise forcing. However, in this model an increase in  $\rho$  also increases both quantities. This implies that classical CSD indicators, i.e., rising variance and AC1, are prone to false positives if there are no strong reasons

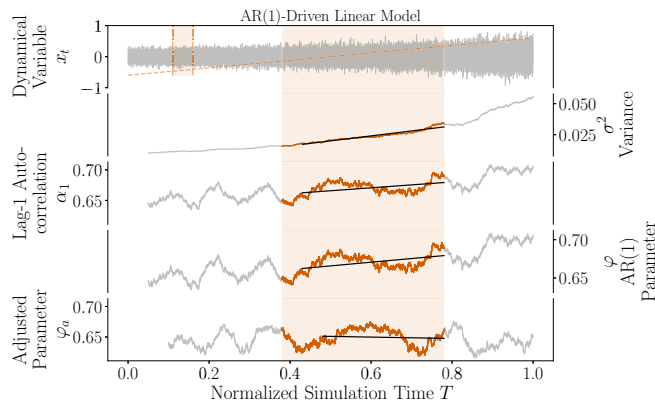


FIG. 1. Example realization for a linear model driven by AR(1) noise, with increasing noise autocorrelation  $\rho$  (orange dashed line). The CSD indicators are calculated in the area marked by the orange shading. The classical methods  $\sigma^2$  and  $\alpha_1$  as well as  $\varphi$  exhibit significant linear trends (solid black). The adjusted method  $\varphi_a$  correctly identifies no significant trend.

to assume that the noise forcing is indeed white, since in that case the autocorrelation of the noise term may increase. An example is given in Fig. 1.

### 1. Estimating $\varphi$ as an indicator of critical slowing down

Casting CSD into the above framework permits the definition of an early-warning indicator. The increase of the AR(1) parameter  $\varphi$  in Eq. (8) is a direct measure of the system’s stability, and the motivation to use the variance and AC1 relies on the behavior of  $\varphi$ . We therefore propose to estimate  $\varphi$  directly, and to use it as a robust indicator for CSD for which the likelihood of false positives is reduced.

One can estimate  $\varphi$  from time series data using an ordinary least-squares (OLS) regression. Given the time series data  $\underline{y} = (y_0, y_1, \dots, y_n)$ , the OLS fit function is

$$S(\varphi) = \sum_{i=1}^n (y_i - \varphi y_{i-1})^2. \tag{11}$$

The best estimator for  $\varphi$  is given by the value  $\hat{\varphi}$  minimizing  $S$ , given by

$$\hat{\varphi} = \frac{\sum_{i=1}^n y_i y_{i-1}}{\sum_{i=1}^n y_{i-1}^2}. \tag{12}$$

The OLS estimator is unbiased if

$$\text{cov}(\epsilon_i, \epsilon_j) = \sigma^2 \delta_{ij}, \tag{13}$$

i.e., if the noise is uncorrelated.

### 2. White noise

In the classical situation with white-noise forcing, Eq. (13) holds and the estimator is unbiased. Since  $E(X_t) = 0$  for an AR(1) process, Eq. (12) coincides with the sample estimation procedure for the linear (Pearson) correlation coefficient  $r(X_{t+1}, X_t)$  and is therefore identical to the AC1 estimation, consistent with Eq. (6). Thus, direct estimation of  $\varphi$  is equivalent with the classical CSD indicators if the time series itself

can be described well by an AR(1) process driven by white noise.

### 3. AR(1) noise

For the model described by Eq. (8), direct estimation of  $\varphi$  using Eq. (12) yields a biased result. It is, however, possible to construct an unbiased estimator [39–41]. One begins with the biased least-squares estimator for a time series  $\underline{y}$ :

$$\hat{\varphi}_b = \frac{\sum_{i=1}^n y_i y_{i-1}}{\sum_{i=1}^n y_{i-1}^2}. \tag{14}$$

For large  $n$  this quantity converges to a biased estimator [39]:

$$\varphi_b = \frac{\varphi + \rho}{1 + \varphi \rho}, \tag{15}$$

which only coincides with  $\varphi$  for the white-noise case  $\rho = 0$ . One can see that an increase in  $\rho$  alone leads to an increase in  $\varphi_b$  (for  $\varphi, \rho \in (0, 1)$ ), which makes  $\varphi_b$  unsuited as an indicator of CSD. However, the convergence properties of the least-squares estimator for  $\rho$ ,

$$\hat{\rho}_b = \frac{\sum_{i=1}^n \hat{V}_i \hat{V}_{i-1}}{\sum_{i=1}^n \hat{V}_{i-1}^2}, \quad \hat{V}_i = y_i - \hat{\varphi}_b y_{i-1}, \tag{16}$$

are also known to be [39]

$$\rho_b = \varphi \rho \varphi_b. \tag{17}$$

Using both of these properties, one can construct an adjusted, unbiased estimator [40] for  $\varphi > \rho$ , given by

$$\hat{\varphi}_a = \frac{\hat{a} + \sqrt{\hat{a} - 4\hat{b}}}{2}, \tag{18}$$

or alternatively for  $\rho > \varphi$  by

$$\hat{\varphi}_a = \frac{\hat{a} - \sqrt{\hat{a} - 4\hat{b}}}{2}, \tag{19}$$

where

$$\hat{a} = \hat{\varphi}_b + \hat{\rho}_b, \quad \hat{b} = \frac{\hat{\rho}_b}{\hat{\varphi}_b}. \tag{20}$$

This estimation procedure works as long as the processes are stationary ( $-1 < \varphi, \rho < 1$ ). For a detailed discussion on the convergence properties of these estimators, see [40,42].

## III. RESULTS

To demonstrate the benefit of our method, the performance of the classical CSD indicators, variance and AC1, are compared to the simple parameter  $\varphi$  and the adjusted parameter  $\varphi_a$  for two models. The first model is a linear model without the potential for abrupt transitions, and the second is a nonlinear model with a double-well potential, which thus allows for a critical transition from one fixed point to the other. Both are driven by white noise and nonstationary AR(1) noise for comparison. The equations are discretized using the Euler-Maruyama method and the different CSD indicators are calculated in rolling windows (see Supplemental Material for details on the employed parameters [36]). For the AR(1) noise driving the systems, the autocorrelation parameter  $\rho$  is

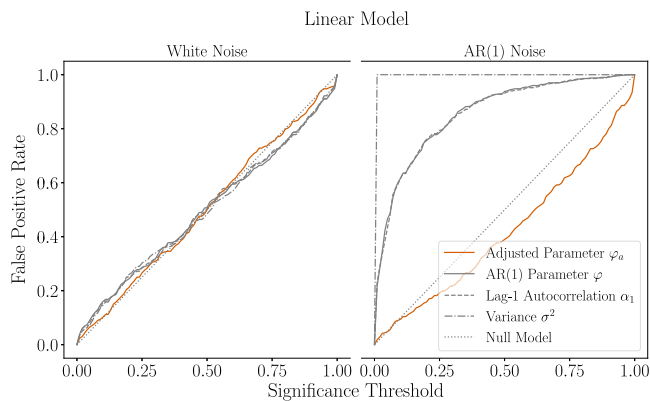


FIG. 2. Performance of the classical CSD indicators compared to the  $\varphi$  indicators proposed here, for a linear dynamical system with single-well potential, driven by white (left) and nonstationary AR(1) noise [right, the parameter  $\rho$  in Eq. (8) increases from 0 to 0.6 over the course of the simulations].

increased linearly from 0 to 0.6 over the course of the simulation. The linear trend of the quantities is calculated using a linear least-squares regression (for the double-well model only data prior to the transition are considered).

An increase in one of the CSD indicators might of course also happen due to random fluctuations. To evaluate the significance of these trends we use a surrogate test based on Fourier surrogates with randomized phases. For a given time series, 1000 surrogates are created. These surrogates coincide with the original time series in their relevant statistical properties; in particular, the total variance and total lag-1 autocorrelation are the same. However, any true trend will be removed in the surrogates, as desired for such a surrogate null model. The likelihood that a trend in the original time series is real and not caused by random fluctuations can then be estimated by comparing the magnitude of the trend in the original time series to the trends in the surrogates. This is done by calculating the proportion of trends that have a larger magnitude in the surrogates than the original time series. The test can be turned into a binary classifier by choosing a significance threshold for this test statistic—the proportion of surrogates that have a larger trend than the original time series. For example, a threshold of 0.05 would mean a maximum of 5% of the surrogate trend magnitudes are allowed to be larger than the trend in the original time series for the trend to be considered statistically significant. For more information on this procedure see [43] as well as Appendix, and for application to CSD estimation see [15,44,45].

The performance of the different CSD indicators is tested for varying values of this threshold. To quantify the performance of the indicators, the number of significant positives trends is calculated for ensembles of 500 realizations. The proportions of these ensembles that yield significant trends—according to the test prescribed above—as a function of the threshold are shown in Figs. 2 and 3.

### A. Linear model

First, we analyze the performance of the classical CSD indicators based on a linear model described by Eq. (3) for

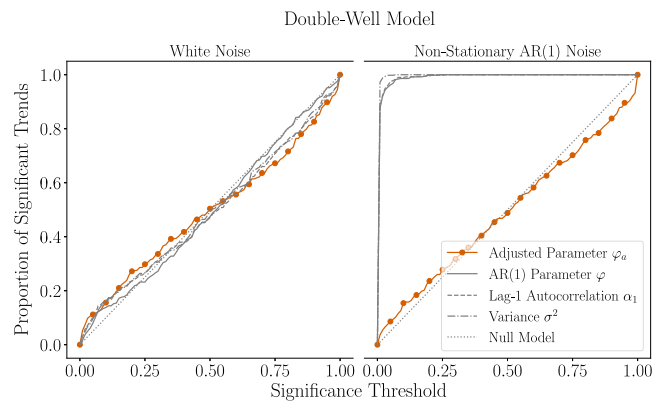


FIG. 3. Performance of the classical CSD indicators compared to the  $\varphi$  indicators proposed here, for a dynamical system with double-well potential, exhibiting a critical transition. Simulations are shown for the model driven by white (left) and nonstationary AR(1) noise [right, the parameter  $\rho$  in Eq. (8) increases from 0 to 0.6 over the course of the simulation].

the white-noise case and Eq. (7) for the nonstationary AR(1) noise case. This linear model is used as a baseline for the performance of the methods, for which no indication of CSD should occur. The trends in the estimator are calculated for  $10^4$  and  $10^5$  total data points in the time series (see Fig. 2 and Supplemental Material [36]). For the white-noise case, the classical indicators as well as  $\varphi$  and  $\varphi_a$  method all detect a number of significant (false) positive trends that matches the significance threshold (Fig. 2). This is the expected behavior for this null model in case of a correctly working CSD indicator. In contrast, based on variance, AC1, or  $\varphi$ , a significant excess of false positive detections occurs in the case of nonstationary AR(1) noise. As predicted, the changes in the noise term are misinterpreted by these indicators as signs of CSD; increases in the data availability only increase the false positive rate since these effects are more clearly resolved. In contrast, the adjusted  $\varphi_a$  does not respond to the changes in the noise forcing and hence performs exactly as expected for the AR(1) noise, and similarly as the classical methods in the white-noise case.

### B. Double-well model

Next, we compare the CSD indicator proposed here, i.e.,  $\varphi_a$ , to the previously proposed indicators for time series obtained with the nonlinear double-well model, given in the continuous case by

$$\dot{x}_t = -x_t^3 + x_t + r(t) + \sigma dW_t \quad (21)$$

( $r$  evolves linearly from  $-1$  to  $1$ ; see Supplemental Material [36] for further details). For a varying control parameter  $r$ , this model undergoes a saddle-node bifurcation preceded by a widening of the basin of attraction. To calculate the CSD indicators in this case, the resulting time series has to be detrended beforehand in order to avoid biases induced by any underlying nonlinear trends. In the case at hand, we detrend the individual stochastic realizations by subtracting the corresponding deterministic realization ( $\sigma = 0$ ). For this model, significantly positive trends constitute a true positive



detection of CSD and hence a justified alarm of a forthcoming critical transition. The classical indicators perform well in the case of white-noise forcing and exhibit a high number of true positives prior to the transition. Their performance increases with the size of the time series (see Supplemental Material [36]). For the AR(1) noise model, the number of significant trends increases faster than for the white-noise model, indicating changes in the noise term are misattributed as signs of CSD. In contrast,  $\varphi_a$  performs similarly in the white-noise case and in the AR(1) noise case, indicating that  $\varphi_a$  is able to disentangle effects in the noise term from actual CSD on the way to a bifurcation-induced transition. Its performance in the white-noise case is however slightly lower than for the classical methods by sometimes misattributing some of the effects of CSD to noise effects. Therefore, using just this method slightly increases the chance for false negative detections.

#### IV. DISCUSSION

Investigating the performance of the classical CSD indicators on simple models reconfirms that CSD provides powerful tools for the prediction of bifurcation-induced transitions. Under the right circumstances, these tools yield near 100% effective precursor signals.

However, we also demonstrated that the classical CSD indicators fail in cases where correlated noise forcing cannot be excluded, which is likely the case in many real-world situations, such as when analyzing climate data; blindly applying the classical CSD indicators, i.e., variance and AC1, likely leads to incorrect conclusions. We demonstrated this by the performance of variance and AC1 on a linear model driven by nonstationary AR(1) noise; in this case the two classical CSD indicators lead to a near 100% false positive rate. These results call attention to the need for further research into indicators for CSD in the presence of correlated noise.

Directly estimating the restoring rate of the system, or equivalently the closely related AR(1) parameter  $\varphi$ , presents strong candidates for improvement. We have demonstrated that one can readily adjust the estimation to obtain unbiased values  $\varphi_a$  for different and especially for nonstationary noise models. Using this unbiased estimate of the AR(1) parameter, i.e.,  $\varphi_a$ , directly as a CSD indicator works robustly in cases of short-memory noise with varying variance and autocorrelation. We emphasize that using the unbiased estimate  $\varphi_a$  as introduced above efficiently decreases the probability of false positive detections, which is very high if the classical indicators are used in the—in practical cases very likely—situation of noise forcing with changing temporal correlation strength. At the same time, our results indicate that using this method might slightly increase the false negative detection rate for actual transitions, by misattributing trends in the autocorrelation to changes in the noise properties. For real-world time series data, the choice of method should therefore be informed by additional information about the most probable noise model.

Since the estimation procedure for  $\varphi_a$  is more involved than for the classical methods, it is likely that more data points will be needed in practical applications of this method compared to the classical methods. This limits the usability of this method in situations where only sparse data are available, e.g., when analyzing (paleo-)climate time series. We

emphasize that the generalization made here is also applicable to other noise models. For example, Brouste and Iacus [46] propose a method to estimate the restoring rate for time series data derived from an Ornstein-Uhlenbeck process driven by a fractional Gaussian noise process. This quantity can be used as a CSD indicator in cases where the noise of the system is better modeled by a long-term memory noise process, and can be complementary to the method proposed here.

#### ACKNOWLEDGMENTS

N.B. acknowledges funding by the Volkswagen Foundation. This is TiPES contribution #143; the TiPES (Tipping Points in the Earth System) project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 820970.

#### APPENDIX: RANDOM PHASE FOURIER SURROGATE TEST

To test the significance of a (linear) trend within a time series, a surrogate data approach is taken. For a given time series  $\underline{y}$ , a number of pseudodatasets is used. These surrogate sets are created using the method of *random phase Fourier surrogates*. For a dataset  $\underline{y} = (y_0, y_1, \dots, y_n)$ , the algorithm works as follows.

- (1) Calculate the discrete Fourier transform,  $\mathcal{FT}(\underline{y})_k = Y_k = \sum_{j=0}^{n-1} y_j e^{-\frac{i2\pi}{n}kj}$ .
- (2) Create a set of uniformly distributed random phases,  $e^{i\phi_k}$ .
- (3) Create the surrogate Fourier transform as  $\tilde{Y}_k = Y_k e^{i\phi_k}$ .
- (4) Create surrogate time series by inverse transform,  $\tilde{y}_k = \mathcal{FT}^{-1}(\tilde{Y})_k$ .

This method is used since the surrogate dataset has the same variance and (linear) correlation structure as the original time series, but since the components are randomly shuffled any real (linear) trend is removed. The linear trend's significance is estimated by fitting a linear function  $f(x) = ax + b$  to the original and surrogate data using an ordinary least-squares regression; the trend is indicated by the magnitude and sign of the parameter  $a$ . If there is a significant (positive) trend in the original time series this parameter will be much larger for the real dataset than for most of the surrogate set, since any trend in the surrogate data is a purely random effect stemming from the variance and autocorrelation of the time series. Consequently, if the trends in the surrogate data are of similar or larger magnitude than for the original time series, it is an indicator that the trend is insignificant and can be explained purely by stochastic effects.

Following this logic, a hypothesis test can be established: Create a large number of surrogates (1000 in practice) and calculate the trend parameter  $a$  for the real ( $a_{\text{ref}}$ ) and all surrogate datasets ( $a_s$ ). The probability that the trend in the original time series can be explained by stochastic effects alone is then

$$p = 1 - \frac{\text{number of } a_s \text{ for which } a_s \geq a_{\text{ref}}}{\text{total number of } a_s}. \quad (\text{A1})$$

This can be turned into a binary classifier by choosing an appropriate significance threshold value for  $p$  (which is varied for Figs. 2 and 3 in the main text).

- [1] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, Tipping elements in the Earth's climate system, *Proc. Natl. Acad. Sci. USA* **105**, 1786 (2008).
- [2] S. R. Carpenter, D. Ludwig, and W. A. Brock, Management of eutrophication for lakes subject to potentially irreversible change, *Ecological Appl.* **9**, 751 (1999).
- [3] D. Angeli, J. E. Ferrell, and E. D. Sontag, Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems, *Proc. Natl. Acad. Sci. USA* **101**, 1822 (2004).
- [4] M. Breakspear, J. A. Roberts, J. R. Terry, S. Rodrigues, N. Mahant, and P. A. Robinson, A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis, *Cereb. Cortex* **16**, 1296 (2006).
- [5] S. Gualdi, M. Tarzia, F. Zamponi, and J.-P. Bouchaud, Tipping points in macroeconomic Agent-Based models, *J. Econ. Dyn. Control* **50**, 29 (2015).
- [6] V. Dakos, B. Matthews, A. P. Hendry, J. Levine, N. Loeuille, J. Norberg, P. Nosil, M. Scheffer, and L. De Meester, Ecosystem tipping points in an evolving world, *Nature Ecology Evolution* **3**, 355 (2019).
- [7] T. A. Welton and H. B. Callen, Irreversibility and generalized noise, *Phys. Rev.* **83**, 34 (1951).
- [8] R. Kubo, The fluctuation-dissipation theorem, *Rep. Prog. Phys.* **29**, 255 (1966).
- [9] C. Kuehn, A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics, *Physica D* **240**, 1020 (2011).
- [10] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara, Early-warning signals for critical transitions, *Nature (London)* **461**, 53 (2009).
- [11] S. R. Carpenter and W. A. Brock, Rising variance: A leading indicator of ecological transition, *Ecology Letters* **9**, 311 (2006).
- [12] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, Slowing down as an early warning signal for abrupt climate change, *Proc. Natl. Acad. Sci. USA* **105**, 14308 (2008).
- [13] P. D. Ditlevsen and S. J. Johnsen, Tipping points: Early warning and wishful thinking, *Geophys. Res. Lett.* **37**, L19703 (2010).
- [14] R. Wang, J. A. Dearing, P. G. Langdon, E. Zhang, X. Yang, V. Dakos, and M. Scheffer, Flickering gives early warning signals of a critical transition to a eutrophic lake state, *Nature (London)* **492**, 419 (2012).
- [15] N. Boers, Early-warning signals for Dansgaard-Oeschger events in a high-resolution ice core record, *Nat. Commun.* **9**, 2556 (2018).
- [16] M. I. Maturana, C. Meisel, K. Dell, P. J. Karoly, W. D'Souza, D. B. Grayden, A. N. Burkitt, P. Jiruska, J. Kudlacek, J. Hlinka, M. J. Cook, L. Kuhlmann, and D. R. Freestone, Critical slowing down as a biomarker for seizure susceptibility, *Nat. Commun.* **11**, 2172 (2020).
- [17] C. D. Rye and T. Jackson, Using critical slowing down indicators to understand economic growth rate variability and secular stagnation, *Sci. Rep.* **10**, 10481 (2020).
- [18] H. Stommel, Thermohaline convection with two stable regimes of flow, *Tellus* **13**, 224 (1961).
- [19] M. I. Budyko, The effect of solar radiation variations on the climate of the Earth, *Tellus* **21**, 611 (1969).
- [20] W. D. Sellers, A global climatic model based on the energy balance of the earth-atmosphere system, *J Appl Meteorol. Climatol.* **8**, 392 (1969).
- [21] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox, Tipping points in open systems: Bifurcation, noise-induced and rate-dependent examples in the climate system, *Phil. Trans. R. Soc. A* **370**, 1166 (2012).
- [22] E.-H. M. El-Hacem, T. J. Bouma, G. S. Fivash, A. A. Sall, T. Piersma, H. Olf, and L. L. Govers, Evidence for “critical slowing down” in seagrass: A stress gradient experiment at the southern limit of its range, *Sci. Rep.* **8**, 17263 (2018).
- [23] M. Berhanu, B. Gallet, R. Monchaux, M. Bourgoin, P. Odier, J.-F. Pinton, N. Plihon, R. Volk, S. Fauve, N. Mordant, F. Pétrélis, S. Aumaître, A. Chiffaudel, F. Daviaud, B. Dubrulle, and F. Ravelet, Bistability between a stationary and an oscillatory dynamo in a turbulent flow of liquid sodium, *J. Fluid Mech.* **641**, 217 (2009).
- [24] B. M. S. Arani, S. R. Carpenter, L. Lahti, E. H. van Nes, and M. Scheffer, Exit time as a measure of ecological resilience, *Science* **372**, aay4895 (2021).
- [25] D. Faranda, F. M. E. Pons, and B. Dubrulle, Statistical early-warning indicators based on Auto-Regressive Moving-Average processes, *J. Phys. A: Math. Theor.* **47**, 252001 (2014).
- [26] B. van der Bolt, E. H. van Nes, S. Bathiany, M. E. Vollebregt, and M. Scheffer, Climate reddening increases the chance of critical transitions, *Nat. Clim. Change* **8**, 478 (2018).
- [27] V. N. Livina and T. M. Lenton, A modified method for detecting incipient bifurcations in a dynamical system, *Geophys. Res. Lett.* **34**, L03712 (2007).
- [28] J. Prettyman, T. Kuna, and V. Livina, A novel scaling indicator of early warning signals helps anticipate tropical cyclones, *Europhys. Lett.* **121**, 10002 (2018).
- [29] N. Boers, Observation-based early-warning signals for a collapse of the atlantic meridional overturning circulation, *Nat. Clim. Change* **11**, 680 (2021).
- [30] B. Øksendal, *Stochastic Differential Equations* (Springer-Verlag, Berlin, 2003).
- [31] H. Held and T. Kleinen, Detection of climate system bifurcations by degenerate fingerprinting, *Geophys. Res. Lett.* **31**, L23207 (2004).
- [32] M. Heßler and O. Kamps, Bayesian on-line anticipation of critical transitions, *New J. Phys.* (2021).
- [33] B. Lehle and J. Peinke, Analyzing a stochastic process driven by Ornstein-Uhlenbeck noise, *Phys. Rev. E* **97**, 012113 (2018).
- [34] C. Willers and O. Kamps, Non-parametric estimation of a langevin model driven by correlated noise, *Eur. Phys. J. B* **94**, 149 (2021).
- [35] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and its Applications* (Springer, New York, 2011).
- [36] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.013230> for further information on simulation parameters and derivations.
- [37] B. Bercu, F. Proïa, and N. Savy, On ornstein–uhlenbeck driven by Ornstein-Uhlenbeck processes, *Statistics and Probability Letters* **85**, 36 (2014).
- [38] M. Schulz and M. Mudelsee, REDFIT: Estimating red-noise spectra directly from unevenly spaced paleoclimatic time series, *Comput. Geosci.* **28**, 421 (2002).
- [39] E. Malinvaud, Estimation et prevision dans les modeles economiques autoregressifs, *Rev. Inst. Stat. Int.* **29**, 1 (1961).

- [40] B. Bercu and F. Proïa, A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process, *ESAIM Probab. Stat.* **17**, 500 (2013).
- [41] F. Proïa, Further results on the H-Test of Durbin for stable autoregressive processes, [arXiv:1203.1871](https://arxiv.org/abs/1203.1871).
- [42] F. Proïa, Further results on the H-Test of Durbin for stable autoregressive processes, *J. Multivariate Anal.* **118**, 77 (2013).
- [43] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska, Surrogate data for hypothesis testing of physical systems, *Phys. Rep.* **748**, 1 (2018).
- [44] M. Rypdal, Early-warning signals for the onsets of greenland interstadials and the younger dryas-preboreal transition, *J. Climate* **29**, 4047 (2016).
- [45] N. Boers and M. Rypdal, Critical slowing down suggests that the western greenland ice sheet is close to a tipping point, *Proc. Natl. Acad. Sci. USA* **118**, e2024192118 (2021).
- [46] A. Brouste and S. M. Iacus, Parameter estimation for the discretely observed fractional Ornstein-Uhlenbeck process and the yuima r package, *Comput. Stat.* **28**, 1529 (2012).