



An Efficient Algorithm to Estimate the Potential Barrier Height from Noise-Induced Escape Time Data

Tamás Bódai^{1,2,3,4} 

Received: 10 March 2019 / Accepted: 21 May 2020 / Published online: 11 June 2020
© The Author(s) 2020

Abstract

An algorithm is developed for determining the potential barrier height experimentally, provided that we have control over the noise strength σ . We are concerned with the situation when the laboratory or numerical experiment requires large resources of time or computational power, respectively, and wish to find a protocol that provides the best estimate in a given amount of time. The optimal noise strength σ^* to use is found to be very simply related to the potential barrier height $\Delta\Phi$ as: $y^* = \Delta\Phi^{-1}$, $y = \sigma^{-2} - \sigma_a^{-2}$, with some “anchor point” σ_a^{-2} ; and, as a second ingredient, an iterative method is proposed for the estimation. For a numerical verification of the *optimality*, we apply the algorithm to a simple system of an over-damped particle confined to a double-well potential, when it is feasible to evaluate statistics of the estimator. Subsequently, we also apply it to a high-dimensional case of a diffusive energy balance model, when the potential barrier height—concerning e.g. the warm-to-snowball-climate transition—cannot be determined analytically, but we would have to resort to more sophisticated numerical methods.

Keywords Freidlin-Wentzell theory · Quasi-potential · Exit rate · Bistability · Potential barrier · Energy balance model

1 Introduction

It is a common phenomenon in nature and technology that a system under noise perturbations exits a regime of its usual dynamics [1–6]. Often it is possible to define a potential function, or nonequilibrium potential [7,8], whereby a potential well can be associated with a usual or persistent dynamics, and a saddle of the potential adjacent to the potential well is a

Communicated by Valerio Lucarini.

✉ Tamás Bódai
bodai@pusan.ac.kr

¹ Centre for the Mathematics of Planet Earth, University of Reading, Reading, UK

² Department of Mathematics and Statistics, University of Reading, Reading, UK

³ Center for Climate Physics, Institute for Basic Science, 46241 Busan, Republic of Korea

⁴ Pusan National University, 46241 Busan, Republic of Korea

feature through which the noise-driven exit takes place [9]. One of the two most important conditions for the possibility of defining a potential appears to be a time-scale separation in the deterministic dissipative dynamics, when the fast processes can be modeled as noise. This is a very common modeling approach in the climate sciences, known as stochastic parametrization; see [10] for a technical reference which also details many caveats. In another scenario, fast perturbations to a system may be truly exogenous. The other condition is that we are in the weak-noise limit [11], when the nonequilibrium potential turns out to be analogous with the classical mechanical action (for which reason it is also called the Freidlin-Wentzell action). The action is a solution of a Hamilton-Jacobi equation, which equation arises from a series expansion of the Fokker-Planck equation with respect to the noise strength, and is associated with the leading-order terms. The potential difference between the bottom of the potential well and the saddle is often termed a potential barrier. The expected exit time then depends on the height of this potential barrier and the (small) noise strength (as expressed by Eq. (4) in Sect. 2). Therefore, knowing the potential barrier height is often of strong interest, because then one can predict—for a given or applied noise strength—the expected escape time.

We develop an algorithm to determine the potential barrier height experimentally, provided that we have control over the noise strength. We are concerned with the situation when the experiment or numerical simulation requires large resources of time or computational power, respectively, and wish to find a protocol that provides the best estimate in a given amount of time. We encountered such a situation when wanted to determine expected transition times to the cold climate for the noisy version of a climate model; see Fig. 2 of [12].

We can envisage another application scenario as follows. Consider that there is a bistability in a deterministically parametrized version of an Earth System Model (ESM), say, a desert-forest bistability, and in a stochastically parametrized version rare transitions can occur, i.e., the system becomes transitive. The key is that one can control the noise strength associated with the stochastic parametrization of the model, and, so, one can apply stronger than realistic noise. Although it is crucial that we remain in the said weak-noise limit. Thereby one can generate transitions more frequently, such that it can become feasible even in an expensive-to-simulate ESM to determine the potential barrier height using the proposed algorithm. Subsequently, it will be just a back-on-the-envelope calculation to estimate the mean transition/exit time for realistic, much smaller, noise strengths.

When the noise strength cannot be controlled, one can apply the adaptive multilevel splitting algorithm [13] in order to determine the mean exit time. In this case there is, of course, no limitation imposed on the noise strength.

Next, in Sect. 2, we detail the derivation of the optimal noise strength for the algorithm, and also detail the iterative procedure to estimate the potential difference. Then, in Sect. 3 we provide a numerical proof-of-concept via applying the algorithm to two example systems. Finally, in Sect. 4, expanding on the above, we provide some remarks on the application and scope of the algorithm.

2 The Efficient Algorithm

First we recap the main results of the large deviation theory of noise-perturbed dynamical systems, including the exponential scaling law of mean exit times, Eq. (4), which is a key ingredient of our algorithm. A precursor to this is Kramers' seminal work [14] on the 'escape

rate¹ of a particle from a “deep” (1D) potential well over a barrier, a concise presentation of which can be found in Sect. 5.10.1 of [9]. We consider the rather generic situation when the dynamics is governed by the following Langevin stochastic differential equation (SDE):

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \sigma \mathbf{D}\xi(t), \tag{1}$$

$\mathbf{x}, \mathbf{F}, \xi \in \mathbb{R}^n$, and the diffusion matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$ is independent of \mathbf{x} , i.e., the white noise ξ is additive. A lack of time-scale separation between resolved (\mathbf{x}) and unresolved variables in the context of stochastic parametrization leads to a memory effect, something that is not included in Eq. (1) [16]; however, this form can be recovered already by a moderate time scale separation [17]. The vector field $\mathbf{F}(\mathbf{x})$ is such that it gives rise, under the deterministic dynamics, $\sigma = 0$, to the coexistence of multiple attractors (including the possibility of an attractor at infinity) and at least one nonattracting invariant set, often called a saddle set. The saddle set is embedded in the boundary of some basins of attraction. Based on a well-established theory due to Freidlin and Wentzell [18], and its extension [19,20], the steady state probability distribution in the weak-noise limit, $\sigma \ll 1$, can be written as

$$W(\mathbf{x}) \sim Z(\mathbf{x}) \exp(-2\Phi(\mathbf{x})/\sigma^2), \tag{2}$$

in which $\Phi(\mathbf{x})$ is called the nonequilibrium- or quasi-potential. In gradient systems where $\mathbf{F}(\mathbf{x}) = -\nabla V(\mathbf{x})$ we have that $\Phi(\mathbf{x}) = V(\mathbf{x})$, provided that $\mathbf{D} = \mathbf{I}$. If \mathbf{D} does depend on \mathbf{x} , then $W(\mathbf{x})$ might not satisfy a large deviation law $\lim_{\sigma \rightarrow 0} \sigma^2 \ln W(\mathbf{x}) = -2\Phi(\mathbf{x})$. See e.g. [21] for an example of multiplicative noise where $\lim_{\sigma \rightarrow 0} \sigma^2 \ln W(\mathbf{x} = x)$ does not exist for some parameter setting and $W(x)$ has a fat tail.

The probability that a noise-perturbed trajectory does not escape the basin of attraction over a time span of t_t decays exponentially:

$$P(t_t) \sim \frac{1}{\tau} \exp(-t_t/\tau). \tag{3}$$

The approximation is in fact quite good already for times $t_t \approx E[t_t] = \tau$ or even smaller. The reciprocal of the expectation value τ can be written as an integral of the probability current through the basin boundary, whose leading component as $\sigma \rightarrow 0$ comes from a point \mathbf{x}_e where $\Phi(\mathbf{x})$ is minimal on the boundary. It is, of course, the global minimum when multiple local minima, corresponding to multiple invariant saddle sets embedded in the basin boundary, are present. The proportionality of the probability current to $W(\mathbf{x})$ leads [8,11,15,22] to:

$$\tau \propto \exp(2\Delta\Phi/\sigma^2), \tag{4}$$

where

$$\Delta\Phi = \Phi(\mathbf{x}_e) - \Phi(A) \tag{5}$$

is what we call the potential barrier height. Both the saddle and the attractor can be chaotic, in which cases $\Phi(\mathbf{x}_e)$ and $\Phi(A)$ have been shown [19,20] to be constant over the saddle [19] and attractor [20], respectively.

Considering (4), the expected transition times increase “explosively” as the noise strength σ decreases. From the point of view of estimating $\Delta\Phi$, say, in a setting of linear regression with $\ln \tau$ as the dependent-and $2\sigma^{-2}$ as the independent variable, there seems to be a trade-off between an improving accuracy of the estimation and an increasing demand of resources as the range of the independent variable is expanded towards smaller values of σ . Accordingly, if we fix the amount of resources that we are willing to commit, an improvement of accuracy

¹ Or, we can call it the ‘exit rate’, in order to avoid confusion with the global stability property of saddle-type invariant sets, also called ‘escape rate’ [15].

is not guaranteed, because we can register fewer transitions with smaller values of σ on average. On the other hand, aiming for boosting the sample number by restricting σ to larger values might not improve accuracy either for the reason that Eq. (7) will illuminate below. We assume that for some σ_a we can estimate $\tau = \tau_a$ arbitrarily accurately because a large number of transitions can be achieved relatively inexpensively. We also assume that in this “anchor point” (4) applies accurately:

$$\tau \approx \tau_a \exp(2\Delta\Phi(\sigma^{-2} - \sigma_a^{-2})), \quad \sigma < \sigma_a. \tag{6}$$

We note in passing that in gradient systems the prefactor $\tau_a \exp(-2\Delta\Phi\sigma_a^{-2})$ in the above is given by the Eyring-Kramers formula [23]; see a generalisation of that for irreversible diffusion processes in [24], where it is assumed, however, that the saddle is a fixed-point. Subsequently, as a departure from the regression-type estimation framework, we will identify the accuracy of estimation by

$$\delta\Delta\Phi = \frac{\sqrt{\text{Var}[\ln \bar{t}_t]}}{y}, \tag{7}$$

where we introduced $y = \sigma^{-2} - \sigma_a^{-2}$, and $\bar{t}_t = \frac{1}{N} \sum_{i=1}^N t_i$ is our finite- N estimate of τ for a fixed σ . Now we can see that as $\sigma \rightarrow \sigma_a$, the inaccuracy explodes. That is, in the present specific setting of estimation there should exist an optimal value σ^* of σ . This is what we determine next.

The sum of the exponentially distributed random variables, $N\bar{t}_t$, does in fact follow an Erlang distribution [25], and so:

$$P(\bar{t}_t) \sim \frac{1}{\tau^N} \frac{(N\bar{t}_t)^{N-1}}{(N-1)!} \exp(-N\bar{t}_t/\tau)N. \tag{8}$$

Note that since $E[\bar{t}_t] = E[t_t] = \tau$, our estimator \bar{t}_t is unbiased. Furthermore, $\text{Var}[\bar{t}_t] = \text{Var}[t_t]/N = \tau^2/N$ in accordance with the Central Limit Theorem. It can be shown that (8) implies that

$$\text{Var}[\ln \bar{t}_t] = \Psi^{(1)}(N), \tag{9}$$

where $\Psi^{(1)}(N)$ is the first derivative of the digamma function [26]. We can make the interesting observation that $\text{Var}[\ln \bar{t}_t]$ does not depend on τ , only on N . Next, we make use of the approximation [26]

$$\Psi^{(1)}(N) \sim 1/N \tag{10}$$

and, upon substitution in (7), write that

$$\delta\Delta\Phi \sim \sqrt{\frac{\tau_a}{T}} \frac{\exp(\Delta\Phi y)}{y}, \tag{11}$$

where, furthermore, we assumed a certain fixed commitment of resources, which can be expressed simply by $T = N\tau$, and also made use of (6). We look for a $\sigma = \sigma^*$ or $y = y^*$ that minimizes $\delta\Delta\Phi$, for which we need to solve $d\delta\Delta\Phi/dy = 0$. This yields our main result:

$$y^* = \Delta\Phi^{-1}. \tag{12}$$

We can make the interesting observation that it is independent of τ_a and T , which we comment on shortly. Rather, y^* depends only on $\Delta\Phi$ (in a very simple way), the unknown that we wanted to determine in the first place, and, so, the result may seem irrelevant to practice for the first sight. However, one can simply start out with an initial guess value, $\Delta\hat{\Phi}_0$, and *iteratively* update the estimate as $\Delta\hat{\Phi}_i$ by performing a maximum likelihood estimation (MLE) [27]

each time a new value of $t_{t,i}$ is acquired. This way, for the acquisition of $t_{t,i+1}$, one continues the experiment/simulation with an updated noise strength $y_{i+1}^* = \Delta \hat{\Phi}_i^{-1}$, $i = 1, \dots, N$, according to (12). The MLE of $\Delta \Phi$ is based on the probability distribution (3) jointly with (6). This is an analogous procedure to the well-established nonstationary extreme value statistics when one or more parameters of the Generalised Extreme Value (GEV) distribution is a function of a covariate that could depend on time (see Chapter 6 of [27]). In our case τ and σ correspond to the parameter of the GEV distribution and the covariate, respectively. We recall that as σ^* does not depend on T , at any time into the experiment/run (for large enough N , though, such that (10) is a good approximation) our estimate of $\Delta \Phi$ is done most efficiently, and, so, we can revise our commitment: either stopping the experiment/run early or extending it.

3 Numerical Verification

Next, we demonstrate the use of our algorithm on two examples; in a single- as well as a multi-dimensional system.

3.1 Example 1: Overdamped Particle in a Symmetrical 1D Double-Well Potential

The governing equation of this system is the following SDE:

$$dx = -V' dt + \sigma dW. \tag{13}$$

We specify our example as: $V = x^4/4 - x^2$. The two minima are at $x_{\pm} = \pm\sqrt{2}$, and the local maximum in between x_{\pm} is at $x_0 = 0$. These are fixed points of the deterministic case ($\sigma = 0$). A numerical solution of the SDE (13) is obtained by using an Euler-Maruyama integrator [28] with a time step size of $h = 0.02$. Examples of time series realisations are shown in Fig. 1, indicating the regime behaviour with transitions between the two regimes. The time series clearly evidence bimodal marginal distributions—corresponding to the two regimes—whose maxima, and the local minimum in between (not shown), are exactly at $x_{\pm} = \pm\sqrt{2}$ and $x_0 = 0$, respectively. With substituting these into (5) we obtain that $\Delta V = \Delta \Phi = 1$. This shows up as the slope of the curve in Fig. 2. The green coloring indicates that (4) is satisfied well even with so strong noise that the time spent in a regime is not so clear cut any more, as seen in Fig. 1 (b). The result of applying our algorithm is shown in Fig. 3, indicating that it serves its purpose, i.e., $\Delta V = \Delta \Phi$ is correctly estimated to be about 1, and that the convergence is rather fast. Finally, Fig. 4 verifies the corner stone of the algorithm, given by Eq. (12), showing the sample standard deviation of a number of estimates. For the purpose of comparison, results with the new algorithm (horizontal red line) and, as a reference, results with different *fixed* sample values of σ (blue circle markers) are shown in one diagram, indicating that the accuracy of the estimate by our algorithm is just about the best accuracy achievable by the *same* amount of computation using the optimal fixed σ^* (vertical gold line). Note that we chose $N = 30$ for our algorithm, resulting in some computational time T , and then we realised $N = \lceil T/\tau(\sigma) \rceil$ transitions using the different fixed σ 's.

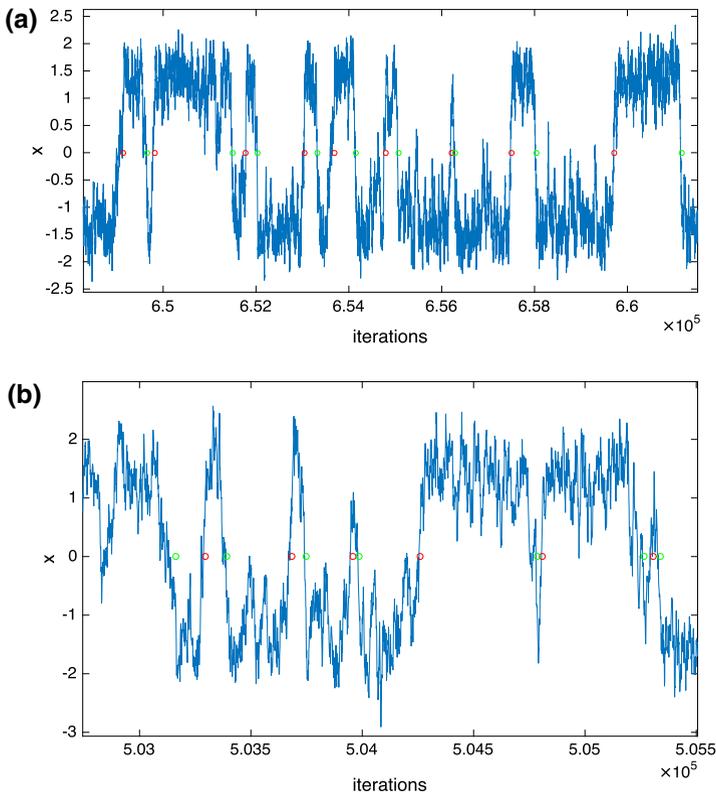
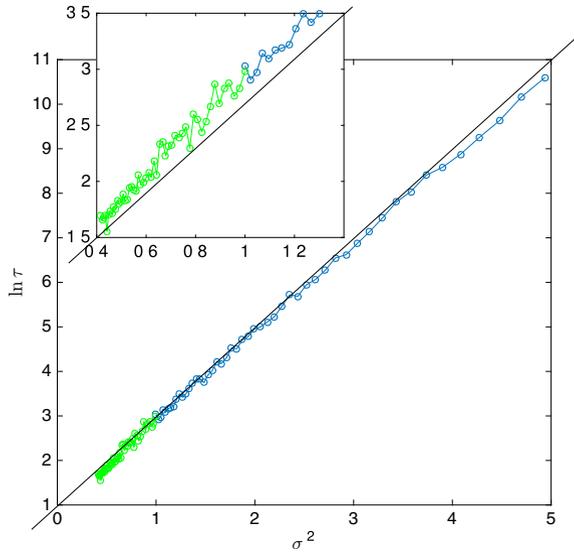


Fig. 1 Numerical solutions of (13). **a** $\sigma = 1.0$, **b** $\sigma = 1.55$. Red and green circle markers indicate transition times defined as a first crossing to the bottom of the upper (lower) potential well since a crossing to the lower (upper) well (color figure online)

3.2 Example 2: The Ghil-Sellers Energy Balance Climate Model (GSEBM)

One of the most striking facts about Earth's climate is its global-scale bistability: beside the relatively warm climate that we live in, under the present astronomical conditions a very cold climate featuring a fully glaciated so-called snowball Earth is also possible, and this state might have been experienced a number of times by Earth in the past few hundred million years [29]. Different hypotheses of transitioning from the warm to the cold climate and the other way round involve external forcings, but in principle it is possible that the climate system in itself—in its autonomous form, without external effects—is transitive, at least in the warm-to-cold direction. This would be a somewhat counterintuitive scenario of no bistability in a strict sense, but the coexistence of an attractor corresponding to the cold climate and a nonattracting set [15] corresponding to the warm climate. Escape from the nonattracting set is then modeled as a noise-induced transition (when the concepts of escape rate and exit rate coalesce), where the noise models some unresolved internal, say, atmospheric and/or oceanic dynamics. Without a requirement for physical realism, we consider additive noise perturbations of the Ghil-Sellers model [30] written for the long time average surface air temperature $T(\phi, t)$ as a function of latitude $\phi \in [-\pi/2, \pi/2]$ or $x = 2\phi/\pi \in [-1, 1]$. The

Fig. 2 Demonstration of the validity of (4) in (13). A straight line of slope $\Delta V = 1$ is included in the diagram for reference. The inset provides a blow-up focusing on rather strong noise strengths—the part of the curve coloured green; see also the main text. To estimate τ we averaged $N = 200$ transition times each sample value of σ corresponding to a circle marker (color figure online)



deterministic GSEBM stands in the form of a diffusive heat equation:

$$\begin{aligned}
 C(x)\partial_t T(t, x) &= \mu Q(x)(1 - \alpha(x, T)) - O(T) \\
 &+ M(x)\partial_x[(D_1(x) + D_2(x)g(T))\partial_x T].
 \end{aligned}
 \tag{14}$$

See [30,31] for the concrete form of the equation, and [32] for a numerical implementation. The model expresses an energy budget, namely, that the tendency of internal energy in latitudinal bands on the left hand side is equal to the sum of—following the order of terms on the right—the incoming short-wave solar radiation modulated by the albedo α , the outgoing long-wave thermal radiation O and the rate of diffusive heat transport from neighbouring latitudinal bands. $M(x)$ is to do with the spherical geometry, and D_1 and $D_2g(T)$ are sensible and latent heat diffusivities, respectively. A global mean albedo can be derived from the dynamics, whose temperature-dependence features a cross-over between approximately constant values for very cold and warm conditions, corresponding to a completely snow-covered ‘snowball’ and completely snow/ice-free conditions, respectively. It is this temperature dependence—giving rise also to the so-called ice-albedo positive feedback—that is responsible for the global scale bistability [31], which is a robust feature of the climate model hierarchy.

Unlike [32] that integrates the deterministic model using Matlab’s `pdepe`, this time we have a stochastic system at hand and, so, we take advantage of Matlab’s SDE simulation suite `simulate`. For this we discretize the eq. with respect to T by the ‘method of lines’, converting the PDE into an ODE, i.e., Eq. (1). The particular difference schemes that we apply, using a regular grid, are:

$$\begin{aligned}
 \partial_x[D_1(x)\partial_x T] &\approx [(T_{j+1} - T_j)D_{1,j+1/2} - (T_j - T_{j-1})D_{1,j-1/2}]/\Delta x^2, \\
 \partial_x[D_2(x)g(T)\partial_x T] &= \partial_x[D_2(x)\partial_x G] \approx [(G(T_{j+1}) - G(T_j))D_{2,j+1/2} \\
 &\quad - (G(T_j) - G(T_{j-1}))D_{2,j-1/2}]/\Delta x^2,
 \end{aligned}
 \tag{15}$$

$j = 1, \dots, J$, where $T_j \approx T(x_j)$, $x_j = (j - 1/2)\Delta x - 1$, $\Delta x = 2/J$, and $D_{1,j\pm 1/2} = D_1(x_{j\pm 1/2})$, $x_{j\pm 1/2} = x_j \pm \Delta x/2$ (see p. 1046 of [33] regarding the x -dependent diffusivity).

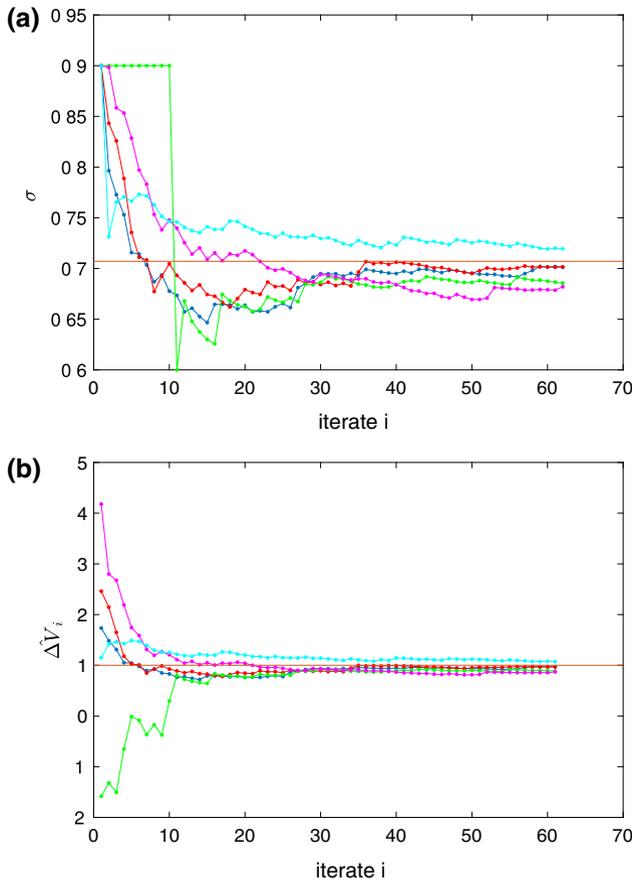


Fig. 3 Proof of concept, I: convergence of estimates $\hat{\Delta}\Phi_i$. The anchor τ_a was established with $\sigma_a = 1$ using $N = N_0 = 400$. Five different realizations of the experiment are shown. The initial value for each was $\sigma_0^* = 0.9 < \sigma_a$. A “safeguarding” of the procedure is facilitated by overriding (12) such that $\sigma_{i+1}^* = \sigma_a$ if $\Delta \hat{\Phi}_i < 0$ and $\sigma_i^* = \sigma_{min}^* = 0.6$ when (12) dictates smaller (color figure online)

Fig. 4 Proof of concept, II: sample standard deviation of 200 estimates $\Delta \hat{V}_N$. For the efficient algorithm we chose $N = 30$, which implies (see the main text) the different N 's for the different fixed sample values of σ . The vertical line marks the prediction of (12) (color figure online)

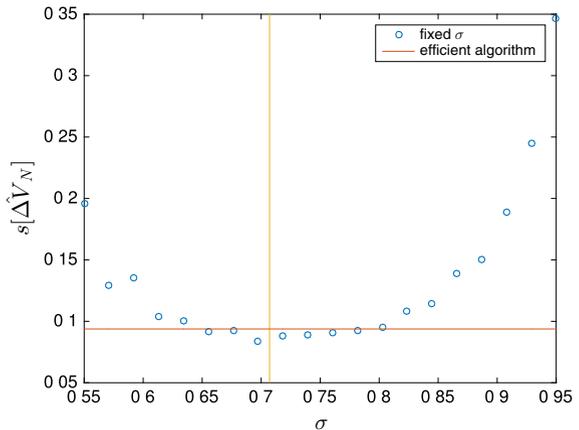
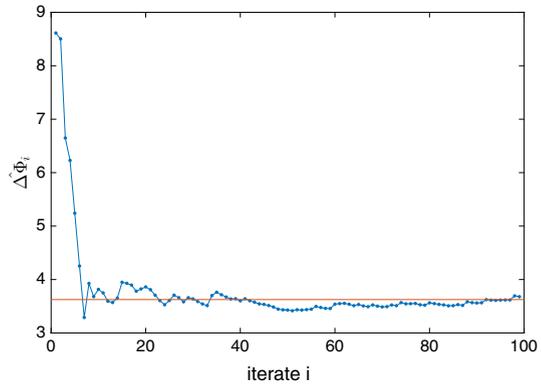


Fig. 5 Same as Fig. 3b but for the discretized GSEBM, $J = 10$, and $\sigma_a = 1.7$, $N_0 = 400$, $\sigma_0^* = 1.5$, $\sigma_{min}^* = 1.0$. We considered only warm-to-cold transitions given the present day solar strength $\mu = 1$. The red horizontal line marks the potential difference $\Delta\Phi$ between the warm climate and the saddle computed by a completely independent method; see the main text. Note that here we measure time in 10^6 seconds [Ms], as a departure from [31,32] (color figure online)



The boundary conditions are eliminated by the ‘method of reflection’, setting $T_0 = T_J$ and $T_{J+1} = T_1$. Such a grid deals effectively with the singularity of $M(x)$ at the poles, but the resulting ODE can be somewhat stiff.

Figure 5 suggests that our new algorithm works also in a multi-dimensional setting: the estimates $\hat{\Delta\Phi}_i$ (blue markers) do convergence to the reference value (horizontal red line). It is key that the reference value is obtained by a completely *independent* method. In the high-dimensional setting, even in gradient systems, in general $\Delta\Phi$ cannot be calculated analytically. To provide a reference, we calculated $\Delta\Phi$ for the discretized system, the SDE, by an action-minimizing procedure [3], using the computer code publicly available as supplementary material of that paper. We realised the time-discretization of the instanton by breaking it down into 100 segments over a span of 2000 [Ms]. We note that for the feasibility of the minimization, already with $J = 10$, it is crucial to provide *symbolically* the gradient of the action with respect to the displacement of the discrete sample points of the instanton. Without this symbolic expression the minimization takes orders-of-magnitude more time (as we estimated by examining toy examples). The symbolic expression—an extremely long expression, practically impossible to keep track of manually—is generated by the code of [3], which makes use of Matlab’s symbolic algebra toolbox. For the code to be successful in this, we need to be able to obtain an explicit expression for the tendencies \dot{T}_j . This could not be achieved by the sophisticated method implemented in Matlab’s `pdepe`, which is why we developed the discretization scheme (15) using the method of lines.

We note that an expression for the potential functional $\Phi(T)$ of the PDE was given in [30]. However, it does not seem to be possible to evaluate this expression for the particular model at hand, not even numerically, because Eq. (7d) of [30] would need to be *first* integrated *analytically*, which does not seem feasible using symbolic algebra software. The reason for this may be that the compact expression featuring multiple applications of the absolute value function to represent the piece-wise conditional expression (arising from the cutoff values for the albedo) may be too complicated.

4 Closing Remarks

It is important to emphasize that we do not seek to determine an effective 1D potential, the scenario that [34] is concerned with, but the deterministic multistable part of the dynamics can be a system of any high dimensionality. In particular, an asymmetry of the potential function

does not affect the proposed procedure. The condition for the existence of a quasi-potential is that the noise is weak and that, when the noise is a representation of unresolved small-scale dynamics, there is a time-scale separation between the resolved and unresolved dynamics. When the noise is weak, the probability of any escape path is exponentially small compared with the most probable exit path through the saddle where the potential over the deterministic basin boundary has a minimum. This includes the possibility of multiple saddle sets embedded in the basin boundary. An example of this may be the reduced stochastic Duffing oscillator-like system underlying the turbulent swirling flow studied in [35]. However, as in this experiment we can actually see multiple well-separated clustering of transition paths that prompt the multiplicity of local potential minima over the basin boundary, we can say that the noise is not weak in the above sense. In this case the potential barrier height should not be expected to predict the mean transition times. Nevertheless, it is not clear what errors to expect as e.g. Fig. 2 indicates that the theoretical scaling applies when the noise is rather strong.

Our algorithm relies on an anchor point in which a large number of transitions are generated so that the mean transition time can be determined highly accurately. For this to be affordable, we look to set the noise as large as possible. However, it should be weak enough such that with the variation of the noise strength the mean transition times scale according to our assumption (as in Eq. (4)), based on the existence of a quasi-potential. There seems to be no a priori indicator what noise strength is already small enough. To check that it is, perhaps the most efficient procedure is an iterative one: (1) One uses a noise strength for the anchor that is affordable by far; (2) uses our iterative algorithm to determine the optimal noise strength belonging to this anchor—alongside the estimate of the potential difference; (3) in a new secondary iterate one associates this optimal noise strength with a new anchor, and (4) determines a new estimate of the potential difference; and so on. If, as a result of the secondary iteration, two subsequent estimates of the potential difference are close, it should mean that the noise is weak enough.

The primary iteration of our algorithm, as described in Sect. 2, dictates that simulation (or experimental) runs that end up in a transition take place sequentially. If simulations are run in parallel, and a new one is started with an updated noise strength right away, then a bias arises towards shorter transition times, which implies a biased estimate of the potential difference. However, it would be a worthwhile future investigation to determine the dependence of the bias on the number of parallel simulations, as perhaps a moderate number of parallel simulations would result in a reasonably small bias, while accelerate the procedure already considerably.

A limitation of our algorithm is that it is applicable only to systems whose deterministic part is autonomous. In climate research applications—what motivated the present work—at least one periodic external forcing is always present: the seasonal cycle. There are two ways to possibly remove the issue with this external forcing. First, it might be possible to achieve dimension reduction such that the deterministic part represents the evolution of a *long-term* average. An example of this is the GSEBM. Second, a periodic forcing is special in that a discrete-time sampling of the dynamics by using the period of the forcing yields an *autonomous discrete-time* dynamical system, the so-called stroboscopic Poincaré return map [36,37]. The quasi-potential can be defined also for such discrete-time maps [15,20], and our algorithm is applicable to maps too. Then, what remains to ensure is that the noise is weak enough, as proposed above via a secondary iteration. With an aperiodic external forcing perhaps it is possible to generalise the Freidlin-Wentzell large-deviation theory of stochastic dynamical systems and define an instantaneous so-called snapshot or pullback [37] quasi-potential, as well as snapshot instantonic paths analogous to manifolds of snapshot saddles [15,38]. However, our algorithm is based on temporal statistics and so not applicable,

even in the case of quasistatically slow forcing, when snapshot objects closely trace their stationary counterparts in autonomous variants of the system. We remark that the latter setting can provide a generalised theory of stochastic resonance [39], which is relevant to approaches dealing with long paleoclimatic time series [34].

Acknowledgements I would like to thank Ying Tang for his extremely helpful support for using their code [3], and Tamás Tél for providing valuable feedback on a draft of the manuscript. I would like to acknowledge Valerio Lucarini for the many inspiring exchanges on the topic of critical transitions, and also Michael Ghil and Freddy Bouchet for helpful comments on this work. I would like to thank three anonymous referees for their suggestions to improve this paper, and one of them for drawing my attention to the references [34,35]. This work received funding from the EU Blue-Action project (under Grant No. 727852), and also from the Institute for Basic Science (IBS), South Korea, under IBS-R028-D1.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Kuehn, C., Martens, E.A., Romero, D.M.: Critical transitions in social network activity. *J. Complex Netw.* **2**(2), 141–152 (2014)
2. Dakos, V., Bascompte, J.: Critical slowing down as early warning for the onset of collapse in mutualistic communities. *Proc. Natl. Acad. Sci.* **111**(49), 17546–17551 (2014)
3. Tang, Y., Yuan, R., Wang, G., Zhu, X., Ao, P.: Potential landscape of high dimensional nonlinear stochastic dynamics with large noise. *Sci. Rep.* **7**(1), 15762 (2017)
4. Faranda, D., Lucarini, V., Manneville, P., Wouters, J.: On using extreme values to detect global stability thresholds in multi-stable systems: The case of transitional plane Couette flow. *Chaos Solitons Fractals* **64**, 26–35 (2014)
5. Takács, D., Stépán, G., John Hogan, S.: Isolated large amplitude periodic motions of towed rigid wheels. *Nonlinear Dyn.* **52**(1), 27–34 (2008)
6. Feudel, U., Pisarchik, A.N., Showalter, K.: Multistability and tipping: from mathematics and physics to climate and brain—minireview and preface to the focus issue. *Chaos: An Interdiscip. J. Nonlinear Sci.* **28**(3), 033501 (2018)
7. Graham, R., Tél, T.: Existence of a potential for dissipative dynamical systems. *Phys. Rev. Lett.* **52**, 9–12 (1984)
8. Grassberger, P.: Noise-induced escape from attractors. *J. Phys. A* **22**(16), 3283–3290 (1989)
9. Risken, H.: *The Fokker-Planck Equation*. Springer, Berlin (1996)
10. Nicolis, C., Nicolis, G.: On the stochastic parametrization of short-scale processes. *Q. J. R. Meteorol. Soc.* **145**(718), 243–257 (2019)
11. Graham, R., Tél, T.: On the weak-noise limit of Fokker-Planck models. *J. Stat. Phys.* **35**(5), 729–748 (1984)
12. Lucarini, V., Bódai, T.: Transitions across melancholia states in a climate model: reconciling the deterministic and stochastic points of view. *Phys. Rev. Lett.* **122**, 158701 (2019)
13. Rolland, J., Bouchet, F., Simonnet, E.: Computing transition rates for the 1-D stochastic Ginzburg-Landau-Allen-Cahn equation for finite-amplitude noise with a rare event algorithm. *J. Stat. Phys.* **162**(2), 277–311 (2016)
14. Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**(4), 284–304 (1940)
15. Lai, Y.-C., Tél, T.: *Transient Chaos*. Springer, New York (2011)
16. Wouters, J., Lucarini, V.: Disentangling multi-level systems: averaging, correlations and memory. *J. Stat. Mech.* **2012**(03), P03003 (2012)

17. Wouters, J., Gottwald, G.A.: Stochastic model reduction for slow-fast systems with moderate time scale separation. *Multiscale Modeling Simul.* **17**(4), 1172–1188 (2019)
18. Freidlin, M.I., Wentzell, A.D.: *Random Perturbations of Dynamical Systems*. Springer, New York (1984)
19. Hamm, A., Tél, T., Graham, R.: Noise-induced attractor explosions near tangent bifurcations. *Phys. Lett. A* **185**(3), 313–320 (1994)
20. Graham, R., Hamm, A., Tél, T.: Nonequilibrium potentials for dynamical systems with fractal attractors or repellers. *Phys. Rev. Lett.* **66**, 3089–3092 (1991)
21. Bóдай, T., Franzke, C.: Predictability of fat-tailed extremes. *Phys. Rev. E* **96**, 032120 (2017)
22. Polovinkin, A.V., Pankratova, E.V., Luchinsky, D.G., McClintock, P.V.E.: Resonant activation in single and coupled stochastic FitzHugh-Nagumo elements. In: *Proceedings of SPIE*, 5467:5467–5467–10 (2004)
23. Bovier, A., Eckhoff, M., Gayard, V., Klein, M.: Metastability in reversible diffusion processes I: sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6**(4), 399–424 (2004)
24. Bouchet, F., Reygner, J.: Generalisation of the Eyring-Kramers transition rate formula to irreversible diffusion processes. *Ann. Henri Poincaré* **17**(12), 3499–3532 (2016)
25. Evans, M., Hastings, N., Peacock, B.: *Statistical Distributions*. Wiley, New York (2000)
26. Abramowitz, M., Stegun, I.A.: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York (1972)
27. Coles, S.: *An Introduction to Statistical Modeling of Extreme Values*. Springer, Berlin (2001)
28. Kloeden, P.E., Platen, E.: *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin (1995)
29. Hoffman, P.F., Schrag, D.P.: The snowball Earth hypothesis: testing the limits of global change. *Terra Nova* **14**(3), 129–155 (2002)
30. Ghil, M.: Climate stability for a Sellers-type model. *J. Atmos. Sci.* **33**, 3–20 (1976)
31. Bóдай, T., Lucarini, V., Lunkeit, F., Boschi, R.: Global instability in the Ghil-Sellers model. *Clim. Dyn.* **44**(11), 3361–3381 (2015)
32. <https://uk.mathworks.com/matlabcentral/fileexchange/46391-gsebm>
33. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: *Numerical Recipes: The Art of Scientific Computing*, 3rd edn. Cambridge University Press, New York (2007)
34. Lucarini, V., Faranda, D., Willeit, M.: Bistable systems with stochastic noise: virtues and limits of effective one-dimensional Langevin equations. *Nonlinear Process. Geophys.* **19**(1), 9–22 (2012)
35. Faranda, D., Sato, Y., Saint-Michel, B., Wiertel, C., Padilla, V., Dubrulle, B., Daviaud, F.: Stochastic chaos in a turbulent swirling flow. *Phys. Rev. Lett.* **119**, 014502 (2017)
36. Bóдай, T., Tél, T.: Annual variability in a conceptual climate model: Snapshot attractors, hysteresis in extreme events, and climate sensitivity. *Chaos: An Interdiscip. J. Nonlinear Sci.* **22**(2), 023110 (2012)
37. Drótos, G., Bóдай, T., Tél, T.: Probabilistic concepts in a changing climate: a snapshot attractor picture. *J. Clim.* **28**(8), 3275–3288 (2015)
38. Bóдай, T., Altmann, E.G., Endler, A.: Stochastic perturbations in open chaotic systems: random versus noisy maps. *Phys. Rev. E* **87**, 042902 (2013)
39. Lucarini, V.: Stochastic resonance for nonequilibrium systems. *Phys. Rev. E* **100**, 062124 (2019)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.