Volterra-series approach to stochastic nonlinear dynamics: The Duffing oscillator driven by white noise

Roman Belousov,* Florian Berger, and A. J. Hudspeth
Howard Hughes Medical Institute, Laboratory of Sensory Neuroscience, The Rockefeller University, New York, New York 10065, USA

(Received 25 January 2019; revised manuscript received 13 March 2019; published 5 April 2019)

The Duffing oscillator is a paradigm of bistable oscillatory motion in physics, engineering, and biology. Time series of such oscillations are often observed experimentally in a nonlinear system excited by a spontaneously fluctuating force. One is then interested in estimating effective parameter values of the stochastic Duffing model from these observations—a task that has not yielded to simple means of analysis. To this end we derive theoretical formulas for the statistics of the Duffing oscillator's time series. Expanding on our analytical results, we introduce methods of statistical inference for the parameter values of the stochastic Duffing model. By applying our method to time series from stochastic simulations, we accurately reconstruct the underlying Duffing oscillator. This approach is quite straightforward—similar techniques are used with linear Langevin models—and can be applied to time series of bistable oscillations that are frequently observed in experiments.

DOI: 10.1103/PhysRevE.99.042204

I. INTRODUCTION

Some of the most interesting and complex behaviors in nature emerge from coupled systems with nonlinearities embedded in an environment. Depending on the relevant time and length scales, influences from the environment can be described effectively as fluctuating forces driving such systems. A fundamental example of such a system is the stochastic Duffing oscillator, which, together with its generalizations, has various applications in engineering and biophysics [1–8]. The Duffing equation offers the simplest nonlinear model that describes bistable oscillatory motion [9, Sec. 7.6]. Under certain physical conditions the equation represents a power-series approximation for a general class of Lienard systems [9, Sec. 7.4].

The Duffing model extends the harmonic oscillator by adding a cubic nonlinear term:

$$\ddot{x} + a\dot{x} + bx + cx^3 = f \tag{1}$$

for an unknown function of time $x(t)$ and an external force $f(t)$. The constants a, b, and c are the damping coefficient, the linear stiffness, and the cubic Duffing parameter, respectively. Because the above equation is of second order in time, the phase of this system is specified by two degrees of freedom $(x, \dot{x})$.

In various situations the form of the relevant driving force is $f(t) = A\dot{w}(t)$, in which $A > 0$ is a constant and $\dot{w}(t)$ is Gaussian white noise of zero mean and unit intensity. Equation (1) describes a stable dynamical system when the coefficients $a > 0$ and $c > 0$ are strictly positive. Unlike the harmonic oscillator, for which $c = 0$, the Duffing model admits a negative linear stiffness b.

The Duffing oscillator is bistable when $b < 0$. Its phase space is symmetric about the origin $(x, \dot{x}) = (0, 0)$, which represents an unstable fixed point in absence of external force. Two stable equilibria occur at $(\xi, 0)$ and $(-\xi, 0)$, in which $x = \pm\xi = \pm\sqrt{-b/c}$ correspond to the minima of the Duffing double-well potential $U(x) = \text{const} + bx^2/2 + cx^4/4$. In the monostable regime, for which $b \geqslant 0$, the origin is the only fixed point.

A problem that arises often in quantitative studies of bistable nonlinear systems is the determination of a model's parameter values. In experiments one usually observes time series of noisy oscillations. The model parameters may then be adjusted empirically to reproduce the measurements as closely as possible. This method is rather arbitrary and imprecise, whereas other available approaches require additional experimental data [2,10–13].

Although a time series of oscillations may in principle contain enough information to infer the parameter values of the Duffing oscillator, this approach has not been duly pursued. In the present paper we derive statistical formulas for the time series $x(t)$ in the regime of bistable oscillations. These expressions rely on the Volterra expansion of functionals [14, Chaps. 1–3], which provide the mathematical framework of nonlinear response theory [15]. Expanding on our analytical results, we then develop statistical methods to estimate the parameter values of the stochastic Duffing Eq. (1) from the time series $x(t)$.

II. GENERAL THEORY

The functional series of Volterra generalize the Taylor-Maclaurin expansion of functions in calculus [14, Sec. 1.5]. In particular, we can represent the solution of Eq. (1) as a functional of the force $f(t)$:

$$\begin{aligned} x(t|f) &= x_0(t) + \int_0^t dt_1\, g_1(t - t_1) f(t_1) \\ &\quad + \iint_0^t dt_1 dt_2\, g_2(t - t_1, t - t_2) f(t_1) f(t_2) + \cdots \\ &= x_0(t) + \gamma_1(t) + \gamma_2(t) + \cdots . \end{aligned} \tag{2}$$

*belousov.roman@gmail.com

Here g_1 and g_2 are the Volterra kernels of the linear and quadratic terms in f, $\gamma_1(t)$, and $\gamma_2(t)$, respectively.

Provided that the series (2) converges, a truncated Volterra expansion approximates the solutions of Eq. (1). We find the unknown kernels $g_{i=1,2,\ldots}$ by using the variational approach [14, Sec. 3.4]: we replace the external force $f(t)$ by a constant $f_c \equiv$ const and substitute Eq. (2) into (1). Then, by collecting terms with coefficients of equal powers in f_c, we obtain the following system of equations:

$$0 = \ddot{x}_0 + a\dot{x}_0 + bx_0 + cx_0^3, \tag{3}$$

$$f = \ddot{\gamma}_1 + a\dot{\gamma}_1 + b\gamma_1 + 3cx_0^2\gamma_1, \tag{4}$$

$$0 = \ddot{\gamma}_2 + a\dot{\gamma}_2 + b\gamma_2 + 3c\left(x_0\gamma_1^2 + x_0^2\gamma_2\right), \quad \ldots \tag{5}$$

Equation (3), which defines $x_0(t)$, is equivalent to the homogeneous Duffing problem (1) with $f \equiv 0$. The Volterra kernels can be found in successively increasing orders from the linear Eqs. (4), (5), etc.

The equilibrium solution $x_0(t) \equiv 0$ of Eq. (3) spawns a particularly convenient set of Eqs. (4) and (5) for the monostable Duffing oscillator [1]. In the bistable case, the kernels of the Volterra series at $x_0(t) \equiv 0$ diverge with $t \to \infty$ (Appendices A and B). In fact, this expansion may even fail to exist [16]. Therefore, we develop the Volterra series at the stable equilibria $x_0(t) \equiv \pm\xi$.

As a generalization of the Taylor-Maclaurin series, the Volterra expansion may be limited by a convergence region. Moreover, the accuracy of the truncated expression deteriorates as $f(t)$ becomes progressively greater: the relevant physical scales are introduced later. Due to the symmetry of the Duffing oscillator, $x_0(t) \equiv \xi$ and $x_0(t) \equiv -\xi$ lead to identical odd-order terms in Eq. (2), whereas the even-order terms differ by a factor of -1 (Appendix B). As we show shortly, these two series are accurate in the neighborhood of the expansion points as long as the system's trajectory $x(t)$ does not cross the special point $x = 0$.

If the amplitude A of the external force $f(t)$ is small, the Duffing oscillator remains in one of the two potential wells at $x = \pm\xi$. The truncated Volterra expansions then describe the solutions of Eq. (1) accurately around the respective equilibrium points. The linear response of $x(t)$ is harmonic in the first order of the parameter A,

$$\ddot{\gamma}_1 + a\gamma_1 - 2b\gamma_1 = f, \tag{6}$$

which can also be obtained by linearization of Eq. (1) at the minima of the Duffing potential.

When A is sufficiently large, the Duffing oscillator undergoes stochastic transitions between the two potential wells. The statistical average $\langle x \rangle = 0$ vanishes due to the symmetry of the problem. Although truncated Volterra expansions of $x(t)$ are inaccurate in this case, Eq. (2) may still be applied to describe pieces of the oscillator's trajectory in an ε neighborhood of each potential well $|x(t) \pm \xi| \leqslant \varepsilon < \xi$. This approximation relies on a physical assumption that the external force does not perturb the system's energy much while the oscillator remains in one of the wells. In this sense the argument f of the functional Eq. (2) is small. Statistically the selected pieces of the oscillator's trajectory belong to two ensembles of *conditional probability distributions* $p(x_\xi) = p(x\,|\,|x - \xi| \leqslant \varepsilon)$ and $p(x_{-\xi}) = p(x\,|\,|x + \xi| \leqslant \varepsilon)$ [17].

The energy barrier that separates the two wells of the Duffing potential becomes negligible for external forces of extreme amplitudes A. The oscillations then resemble those of a monostable regime.

III. STATISTICAL ANALYSIS

The time-invariant probability density of a bistable Duffing oscillator driven by white noise is generally bimodal. Two Gaussian-like peaks correspond to the minima of the double-well potential $U(x)$, for which the harmonic oscillator Eq. (6) describes the local dynamics of $x(t)$. We can construct the time-invariant probability density of x by using an exponential form:

$$p(x) \approx Z^{-1} \exp[-P(x) + O(x^5)], \tag{7}$$

in which Z is a normalization constant and $P(x) \geqslant 0$ is a polynomial of fourth order in x [17,18]. At the two global minima of Eq. (7) the following conditions must be satisfied:

$$P(\pm\xi) = 0, \quad \frac{\partial^2 P}{\partial x^2}(\pm\xi) = -4ab/A^2.$$

The last equality ensures that the Laplace approximation of $p(x)$ at $x = \pm\xi$ [19] obeys the statistics of Eq. (6) [20]. Owing to the symmetry of the bistable Duffing system $p(x) = p(-x)$, the general form of $P(x)$ is given by

$$P(x) = \frac{ac(x^2 - \xi^2)^2}{2A^2},$$

which leads to

$$p(x) \approx Z^{-1} \exp\left[-\frac{ac(x^2 - \xi^2)^2}{2A^2}\right] \propto \exp\left[-\frac{2aU(x)}{A^2}\right]. \tag{8}$$

The normalization constant Z can be found by integration of the exponential factor in the above equation:

$$Z = \frac{\pi\xi}{2} \exp(-z)[I_{-1/4}(z) + I_{1/4}(z)],$$

in which $z = ac\xi^4/(4A^2)$ and $I_i(\cdot)$ is the ith-order modified Bessel function of the first kind.

In addition the autocorrelation function $\chi(t)$ can be approximately calculated for $x_{\pm\xi}(t)$ from Eq. (6) [21]:

$$\chi(t) \approx \chi_1(t) = \frac{\langle \gamma_1(0)\gamma_1(t)\rangle}{\langle \gamma_1^2 \rangle} = \exp\left(-\frac{a\,|t|}{2}\right)\left[\cos(\Omega\,|t|) + \frac{a}{2\Omega}\sin(\Omega\,|t|)\right], \tag{9}$$

in which[1]

$$\Omega = \sqrt{-2b - a^2/4}. \tag{10}$$

[1] If $\Omega^2 < 0$ one should use its absolute value instead and replace the trigonometric functions of cosine and sine in Eq. (9) by the hyperbolic ones [20,21].

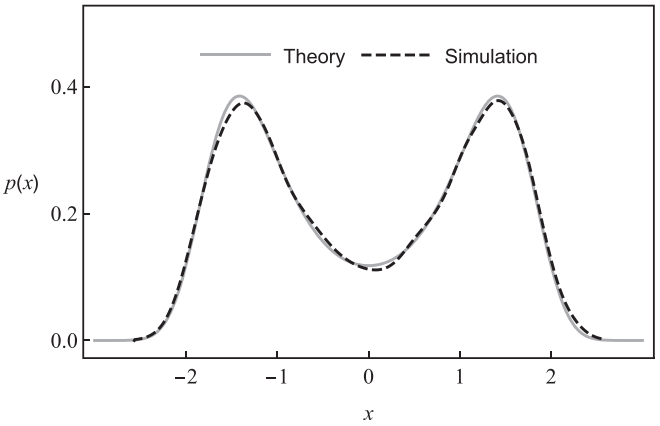

FIG. 1. Steady-state probability density $p(x)$ for the time series simulated with $A = 1.3$. The theoretical expression is given by Eq. (8), whereas the computational results are represented by a smooth histogram [25].

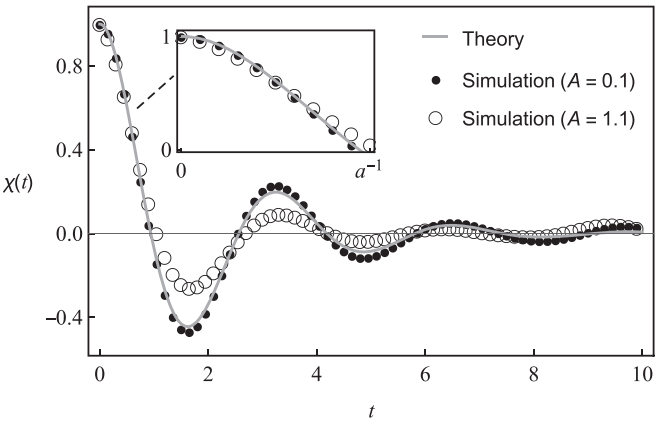

FIG. 2. Time autocorrelation function $\chi(t)$ for time series $x(t)$ driven by a weak external force ($A = 0.1$) and for the time series $x_\varepsilon(t)$ ($A = 1.1$). Error bars, which are comparable in size to the plot markers, are omitted. The theoretical curve corresponds to Eq. (9). The inset magnifies the exponential decay for $t \lesssim a^{-1}$.

IV. DIMENSIONAL ANALYSIS

Equations (1), (8), and (9) characterize physical scales of the Duffing oscillator. First we adopt the constants a^{-1} and $a/\sqrt{c}$ as the units of t and x, respectively. The energy scales are then determined by the height of the barrier between the wells of the Duffing potential $\epsilon = b^2/(4c)$ (per unit mass). It suffices therefore to consider $b = -2\sqrt{\epsilon c}$ as a typical value for a unit energy barrier ϵ. Note that the parameter A is measured in units of $a^{5/2}/\sqrt{c}$, because the units of white noise are $\sqrt{a}$.

The Boltzmann-like factor $\exp[-2aU(x)/A^2]$ in Eq. (8) relates the level of energy fluctuations $A^2/(2a)$ in the system to the Duffing potential $U(x)$. This helps us identify the amplitudes of the external force $A \lesssim \sqrt{2a\epsilon}$, for which the truncated Volterra series might be useful to describe the trajectories $x_{\pm\xi}(t)$.

Because Eq. (9) was derived from the linear-response approximation, it is independent of A. Although this formula is very convenient, its accuracy is limited to small time and energy scales, as shown below. The autocorrelation function Eq. (9) decays exponentially with a relaxation time $\tau \approx a^{-1}$ [22]. In general, we may expect the formula (9) to hold for $0 \leqslant t \lesssim \tau$.

V. PARAMETRIC INFERENCE

To test our theoretical results we simulated Eq. (1) (Appendix D) by using an operator-splitting algorithm [23,24, Appendix C]. The results are reported in the system of units reduced by the time, length, and energy constants a^{-1}, $a/\sqrt{c}$, and ϵ, respectively. As justified earlier, the constant $b = -2$ is fixed. Our simulations differ only by values of the parameter A.

Histograms of the time series $x(t)$ agree with Eq. (8) for all values of the parameter A that we explored (Fig. 1). When a simulation does not last long enough to observe sufficiently many transitions over the energy barrier ϵ, the sample of x may be biased toward the Duffing potential well in which the oscillator spends more time. The symmetry $p(-x) = p(x)$ may therefore appear imperfect in histograms of x. Another issue may emerge if the time resolution of the sample $x(t)$ is not sufficient to observe the trajectory of fast transitions between the two wells ($-\xi < x < \xi$). In this case the histogram's peaks overestimate the probability density at $x \approx \pm\xi$ and underestimate it at $x \approx 0$ with respect to Eq. (8).

By using the maximum-likelihood fitting of Eq. (8) to the time series, we can determine the values of the parameters ξ and $\sigma = A/\sqrt{ac}$. The density $p(x)$ estimated in this way is graphically indistinguishable from the theoretical prediction plotted in Fig. 1. Because Eq. (8) is not sensitive to the sample biases that are discussed above, the curve fitting of $p(x)$ yields very reliable results.

The trajectories $x_{\pm\xi}(t)$ were selected from the time series $x(t)$ with $\varepsilon = 3\sigma/(2\xi)$.[2] Profiting from the symmetry of Eq. (1), we combined these two samples: $\{x_\varepsilon(t)\} = \{x_\xi(t)\} \cup \{|x_{-\xi}(t)|\}$. As expected, the local time autocorrelation function of $x_\varepsilon(t)$ agrees well with Eq. (9) in the interval $0 \leqslant t \lesssim a^{-1}$ (Fig. 2). When the external force is too weak to drive transitions between the potential wells of the Duffing oscillator, the same theoretical expression matches perfectly the autocorrelation function of $x(t)$.

The local time autocorrelation function of $x_\varepsilon(t)$ has an undulatory shape. Note that Eq. (9) predicts quite accurately the frequency of these undulations even when $t \gg a^{-1}$; the discrepancy is due to their amplitude. This observation can be explained by considering higher-order contributions $\gamma_{i=2,3,\ldots}$ (Appendix C).

Good estimates of the frequency Ω can be obtained by fitting Eq. (9) to the time-series autocorrelations at small $t \lesssim a^{-1}$. The parameter a that controls the decay of the amplitude of $\chi(t)$, however, is very sensitive to small errors introduced by the approximate expression $\chi_1(t)$. Because the second-order correction from Eq. (2) already leads to an unwieldy

[2]We set ε to three standard deviations of the Gaussian approximation for Eq. (8) at the most likely values $x = \pm\xi$.

TABLE I. Statistical inference of parameter values for the Duffing oscillator Eq. (1), which was simulated with fixed values $a=1$, $b=-2$, and $c=1$; A varied in the range [0.6,1.3]. The estimated parameter values are denoted by $\hat{a}$, $\hat{b}$, $\hat{c}$, and $\hat{A}$, respectively. The uncertainties are given by one standard deviation calculated as described in Appendix C.

A	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{A}$
0.6	1.06 ± 0.13	-2.02 ± 0.59	1.01 ± 0.30	0.62 ± 0.16
0.7	1.13 ± 0.12	-1.99 ± 0.59	1.00 ± 0.30	0.75 ± 0.17
0.8	1.17 ± 0.12	-2.10 ± 0.56	1.06 ± 0.28	0.91 ± 0.16
0.9	1.13 ± 0.15	-2.06 ± 0.59	1.04 ± 0.30	0.99 ± 0.17
1.0	1.13 ± 0.17	-2.19 ± 0.53	1.10 ± 0.27	1.12 ± 0.16
1.1	1.15 ± 0.19	-2.14 ± 0.55	1.07 ± 0.28	1.22 ± 0.17
1.2	1.14 ± 0.27	-2.28 ± 0.53	1.14 ± 0.27	1.37 ± 0.19
1.3	1.07 ± 0.42	-2.33 ± 0.50	1.16 ± 0.25	1.44 ± 0.25

expression for $\chi(t)$, we propose instead a phenomenological equation motivated by the general form of higher-order Volterra kernels (Appendix C):

$$\chi(t) \approx \hat{\chi}(t) = (1-\alpha)\exp(-a|t|) + \exp(-a|t|/2) \times [\alpha \cos(\Omega|t|) + \beta \sin(\Omega|t|)], \tag{11}$$

in which α and β are unknown parameters, whereas the explicit expression can be substituted for $\Omega = \sqrt{-2b - a^2/4}$.

Curve fitting of Eq. (11) on the interval $0 \leqslant t \lesssim a^{-1}$ with four unknown parameters—a, b, α, and β—yields quite accurate values for a and b. Technical details of this procedure are available in Appendix C. The numerical values of c and A can be found from the estimates of a, b, ξ, and σ (Table I).

As shown above, the values of all four parameters of Eq. (1)—a, b, c, and A—can be inferred from the bistable time series $x(t)$. Our approach is limited to moderate noise intensities $A \sim \sqrt{a\epsilon}$, for which the first term of the Volterra expansion provides a tenable approximation (Table I). In a sense this method extends the statistical techniques that were developed for the harmonic oscillator driven by white noise [21,24].

A precise quantitative description of stochastic nonlinear systems is necessary to advance our understanding of complex behaviors observed in physics, engineering, and biology. The Volterra expansion offers important insights into the statistical theory of such systems. In a future communication we will present analysis of another classical model—the Van der Pol oscillator. Unlike approaches that explore ensemble and phase-space properties of a dynamical system, e.g., the Fokker-Planck equation, the perspective of the Volterra series is closer to the original Langevin picture, which regards realizations of trajectories and their local dynamics.

The convergence issues of Eq. (2) may also stimulate interest in the Wiener theory of orthogonal functional series [26, Chap. 9]. This development might even lead to more advanced theoretical results for the time autocorrelation function of a nonlinear oscillator. As we demonstrated above, the analysis of autocorrelations may provide a reliable estimation of a model's parameter values from experimental measurements. The Wiener approach can also be related to the Fokker-Planck and path-integral formalism, e.g., [27].

ACKNOWLEDGMENTS

R.B. acknowledges support as a Research Associate and A.J.H. as an Investigator of Howard Hughes Medical Institute.

APPENDIX A: MONOSTABLE DUFFING OSCILLATOR

In this section we review the Volterra-series representation of solutions for the Duffing Eq. (1) in the monostable regime of oscillations ($b \geqslant 0$) [1]. Because we study a stationary problem endowed with a time-invariant probability density, Eq. (2) can be recast as

$$\begin{aligned} x(t|f) &= x_0 + \int_{-\infty}^{\infty} dt_1\, g_1(t-t_1) f(t_1) \\ &\quad + \iint_{-\infty}^{\infty} dt_1 dt_2\, g_2(t-t_1, t-t_2) f(t_1) f(t_2) + \cdots \\ &= x_0(t) + \gamma_1(t) + \gamma_2(t) + \cdots, \end{aligned} \tag{A1}$$

in which the integration limits are extended to infinities by invoking the causality of the kernels $g_i(\ldots, t_j, \ldots) = 0$ when $t_j \leqslant 0$, and by assuming the initial condition $x(-\infty) = x_0 \equiv$ const. We will omit indication of the infinite integration limits in the following.

The trivial equilibrium of Eq. (3) is the most convenient expansion point $x_0 \equiv 0$ for Eq. (A1). Equations (4), (5), etc. then become

$$\ddot{\gamma}_1 + a\dot{\gamma}_1 + b\gamma_1 = f, \tag{A2}$$

$$\ddot{\gamma}_2 + a\dot{\gamma}_2 + b\gamma_2 = 0, \tag{A3}$$

$$\ddot{\gamma}_3 + a\dot{\gamma}_3 + b\gamma_3 = -c\gamma_1^3, \quad \cdots \tag{A4}$$

The above equations describe essentially the same harmonic oscillator

$$\ddot{\gamma}_i + a\dot{\gamma}_i + b\gamma_i = f_i \tag{A5}$$

subject to different forcing terms $f_i = f, 0, -c\gamma_1^3 \ldots$ The left-hand side of Eq. (A5) corresponds to the linearized Duffing system that can be obtained by neglecting the nonlinear cubic term in Eq. (1).

By definition the Green function of Eq. (A5) is the linear Volterra kernel g_1:

$$\gamma_1(t) = \int ds\, g_1(t-s) f(s), \tag{A6}$$

$$g_1(t) = \frac{2H(t)}{\sqrt{4b-a^2}} \exp\left(-\frac{at}{2}\right) \sin\left(\frac{\sqrt{4b-a^2}\,t}{2}\right), \tag{A7}$$

in which $H(t)$ is the Heaviside step function. We immediately see that $\gamma_2(t) \equiv 0$, which implies that the quadratic Volterra kernel vanishes identically. The cubic term $\gamma_3(t)$ is given by

$$\gamma_3(t) = -c \int ds\, g_1(t-s) \gamma_1(s)^3.$$

Owing to our choice of $x_0 = 0$, all the Volterra kernels of even order (g_i, $i = 2, 4, \ldots$) vanish. This result reflects the symmetry of the Duffing Eq. (1). The even-order kernels give rise to the statistical moments $\langle x^j \rangle$, $j = 2, 4, \ldots$, which must also vanish in the symmetric system with a time-invariant probability density $p(x) = p(-x)$. The monostable Duffing oscillator may therefore be described quite accurately by

linear-response theory. Indeed, the error of such a representation is of the order f^3:

$$x(t) = \gamma_1(t) + O(f^3). \tag{A8}$$

The stationary solutions of Eq. (A5) are described by a bounded Green function [$\int dt\,|g_1(t)| < \infty$] when $a > 0$ and $b > 0$. Because the higher-order kernels have the same property, the series Eq. (A1) converge with $x_0 = 0$ [16]. One may then estimate statistical properties of the stationary solutions from Eq. (A8).

APPENDIX B: BISTABLE DUFFING OSCILLATOR

The Volterra kernels found from Eqs. (A2), (A3), etc. diverge for $t \to \infty$ when $a < 0$ or $b < 0$. The latter case corresponds to the bistable regime of the Duffing oscillator. Equation (A5) then describes an unstable system and the statistical properties of the stationary solution $x(t)$ can no longer be calculated from Eq. (A8). Because the ensuing Volterra kernels are unbounded, even the existence of the expansion Eq. (A1) cannot be ascertained.

Equation (A2) fails when $b < 0$, because it represents a linearization of Eq. (1) around an unstable equilibrium point—the local maximum of the Duffing potential at $x = 0$. One may however construct a convergent series Eq. (A1) at the minima of the potential wells $x_0 = \pm\xi$. Instead of Eqs. (A2)–(A4) we then obtain [cf. Eq. (6)]

$$\ddot{\gamma}_1 + a\gamma_1 - 2b\gamma_1 = f, \tag{B1}$$

$$\ddot{\gamma}_2 + a\dot{\gamma}_2 - 2b\gamma_2 = \mp 3c\xi\gamma_1^2, \tag{B2}$$

$$\ddot{\gamma}_3 + a\dot{\gamma}_3 - 2b\gamma_3 = -c\big(\gamma_1^3 \pm 6\xi\gamma_1\gamma_2\big). \tag{B3}$$

The linear Volterra kernel is given by the Green function of Eq. (B1) Sec. A:

$$\gamma_1(t) = \int ds\, g_1(t-s) f(s), \tag{B4}$$

$$g_1(t) = \frac{H(t)}{\Omega} \exp\left(-\frac{at}{2}\right) \sin(\Omega t), \tag{B5}$$

in which Ω is given by Eq. (10). From Eqs. (B2) and (B3) we also find the quadratic and cubic response terms in the form

$$\gamma_2(t) = \mp 3c\xi \int ds\, g_1(t-s)\gamma_1^2 = \mp 3c\xi \int ds\, g_1(t-s) \times \iint ds_1\, ds_2\, g_1(s-s_1) g_1(s-s_2) f(s_1) f(s_2), \tag{B6}$$

$$\begin{aligned}\gamma_3(t) &= -c \int ds\, g_1(t-s)[\gamma_1(s)^3 \pm 6\xi\gamma_1(s)\gamma_2(s)] \\ &= \int ds\, g_1(t-s) \iiint ds_1\, ds_2\, ds_3\, g_1(s-s_1) \\ &\quad \times [g_1(s-s_2) g_1(s-s_3) \pm 6\xi g_2(s-s_2, s-s_3)] \\ &\quad \times f(s_1) f(s_2) f(s_3).\end{aligned} \tag{B7}$$

To extract the quadratic and cubic kernels $g_{i=2,3}$ from the above equations we rely on a simplified growing-exponential approach [14, Sec. 3.5]. We use a substitution rule for the product of the forcing terms in the form

$$f(s_1) f(s_2) \cdots \to \exp(-i\omega_1 s_1 - i\omega_2 s_2 - \cdots). \tag{B8}$$

The results that are obtained for arbitrary ω_1 and ω_2 hold also in the special case $\omega_1 = \omega_2$. Like the sum of growing exponentials [14, Sec. 3.5], our approach also renders the symmetric form of the Volterra kernels [$g_i(\ldots s_j, s_k \ldots) = g_i(\ldots s_k, s_j \ldots)$]. In general, we have

$$\begin{aligned}\gamma_i(t) &= \int \cdots \int ds_1 \cdots ds_i\, g_i(t-s_1, \ldots, t-s_i) \\ &\quad \times \exp(-i\omega_1 s_1 \cdots - i\omega_i s_i) \\ &= G_i(\omega_1, \ldots, \omega_i) \exp[it(\omega_1 + \cdots + \omega_i)],\end{aligned} \tag{B9}$$

in which G_i is the Fourier transform of the kernel g_i.

By substituting Eq. (B8) into (B6) we obtain

$$\begin{aligned}\gamma_2(t) &= \mp 3c\xi\, G_1(\omega_1) G_1(\omega_2) G_1(\omega_1+\omega_2) \\ &\quad \times \exp[it(\omega_1+\omega_2)],\end{aligned} \tag{B10}$$

and by comparing the above equation with Eq. (B9) we identify the Fourier image of the quadratic kernel

$$G_2(\omega_1, \omega_2) = \mp 3c\xi\, G_1(\omega_1) G_1(\omega_2) G_1(\omega_1+\omega_2). \tag{B11}$$

We likewise find the Fourier transform of the cubic kernel

$$\begin{aligned}&G_3(\omega_1, \omega_2, \omega_3) \\ &\quad = -cG_1(\omega_1+\omega_2+\omega_3) G_1(\omega_1) G_1(\omega_2) G_1(\omega_3) \\ &\qquad \times [1 - 18bG_1(\omega_2+\omega_3)].\end{aligned} \tag{B12}$$

Because the expressions for g_2 and g_3 are unwieldy, further calculations are more convenient in Fourier space.

The terms of even orders do not vanish in the Volterra series about $x_0 = \pm\xi$, because the potential wells surrounding these points are asymmetric. The global symmetry of the Duffing potential $U(x) = U(-x)$ ensures, however, that the even-order Volterra kernels for $x_0 = \pm\xi$ have opposite signs, whereas the odd-order kernels coincide [cf. Eqs. (B11) and (B12)].

APPENDIX C: LOCAL AUTOCORRELATIONS OF THE BISTABLE DUFFING OSCILLATOR

Equation (9), which approximates the autocorrelation function $\chi(t)$, has been derived from the Volterra expansion truncated at the linear term $\gamma_1(t)$ given by Eq. (B4). To find a second-order correction, one may include a quadratic contribution in $x(t) \approx \gamma_1(t) + \gamma_2(t)$:

$$\chi(t) \approx \chi_2(t) = \frac{\theta_1(t) + \theta_2(t)}{\theta_1(0) + \theta_2(0)}, \tag{C1}$$

in which

$$\theta_1(t) = -\frac{A^2}{4ab}\chi_1(t), \tag{C2}$$

$$\theta_2(t) = \langle\gamma_2(0)\gamma_2(t)\rangle - \langle\gamma_2\rangle^2. \tag{C3}$$

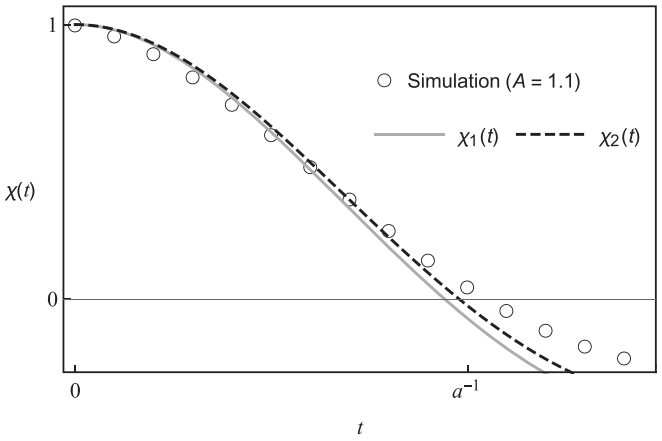

FIG. 3. Local time autocorrelation function $\chi(t)$ of the time series $x_\varepsilon(t)$: comparison of theoretical predictions $\chi_1(t)$ and $\chi_2(t)$ [Eqs. (9) and (C1)] with the simulation results for $A = 1.1$. The corrections that are introduced in $\chi_2(t)$ by including the second-order response term in Eq. (A1) provide a subtle improvement over $\chi_1(t)$.

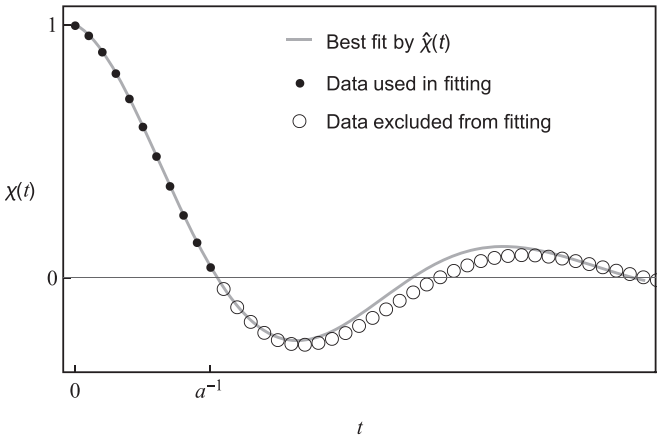

FIG. 4. Curve fitting of $\hat{\chi}(t)$ to the time autocorrelation function of $x_\varepsilon(t)$ observed in simulations with $A = 1.1$. The expression optimized in the interval $0 \leqslant t \lesssim a^{-1}$ extrapolates well up to $t \lesssim 2a^{-1}$.

From [26, Eq. (11.3-14)] one can find the Fourier transform of θ_2

$$\Theta_2(\omega) = \frac{A^4}{\pi}\int d\lambda\, |G_2(\lambda, \omega-\lambda)|^2 = -\frac{9bcA^4}{\pi}|G_1(\omega)|^2\int d\lambda\, |G_1(\omega)|^2\, |G_1(\omega-\lambda)|^2. \tag{C4}$$

Note that

$$\theta_1(t) = \int \frac{d\omega}{2\pi A^2}|G_1(\omega)|^2 \exp(i\omega t).$$

By virtue of the convolution theorem, the inverse Fourier transform of Eq. (C4) then yields

$$\begin{aligned}\theta_2(t) &\propto \int ds\, \chi_1(t-s)\chi_1(s)^2 \\ &= \mathrm{const}_0 \exp(-a\,|t|) + \exp(-a\,|t|/2)\,[\mathrm{const}_1 \sin(\Omega\,|t|) \\ &\quad + \mathrm{const}_2 \cos(\Omega\,|t|)] + \exp(-a\,|t|/2)\,[\mathrm{const}_3 \sin(2\Omega\,|t|) \\ &\quad + \mathrm{const}_4 \cos(2\Omega\,|t|)],\end{aligned} \tag{C5}$$

in which the unwieldy constants $\mathrm{const}_{i=0,1,2,3,4}$ are not spelled out for clarity. These coefficients, which can be readily found with the help of a symbolic computational software [25], depend in a complicated manner on all four parameters of the Duffing oscillator.

The complete expression of $\chi_2(t)$ reveals higher-order harmonics $j\Omega$, $j = 2, 3, \ldots$ in the autocorrelation function. Additional oscillations at these frequencies are commensurate with the undulations of $\chi_1(t)$. For this reason, as noted in the main text, Eq. (11) predicts correctly the undulatory component of $\chi(t)$.

Due to the complexity of the explicit expression for $\chi_2(t)$, using Eq. (C1) in calculations and curve fitting is problematic. The improvement that is achieved over Eq. (9) is also modest (Fig. 3). Higher-order expressions that take into account more terms from Eq. (A1) might be formidably long. In curve fitting we therefore use a phenomenological Eq. (11) as justified below.

Contemplating the development of the Volterra terms in Eqs. (B1)–(B12), one may expect that $\chi(t) \approx \chi_n(t)$, calculated from Eq. (A1) with n response terms, contains only convolution and power products of $\chi_1(t)$ [cf. Eq. (C5)]. The resulting expression would be a composition of time-dependent factors in the form $\propto\exp(iat/2)$, $\propto\sin(j\Omega t)$, and $\propto\cos(k\Omega t)$ with integers $i \geqslant 2$, $j \geqslant 1$, and $k \geqslant 1$. Whereas we retain the fundamental harmonic terms $\propto\sin(\Omega t)$ and $\propto\cos(\Omega t)$, as well as the slowly decaying exponential $\propto\exp(at)$ in $\chi_2(t)$, we introduce unknown coefficients α and β to account for higher-order corrections. By imposing an additional constraint $\chi(0) = 1$, we arrive at Eq. (11).

Finally, we discuss briefly the procedure of curve fitting for Eq. (11) (Fig. 4). By using the criterion of Lagarkov and Sergeev [21,28], we select from the sample autocorrelation data the observations in the interval $0 \leqslant t \leqslant t_0$, in which t_0 is the instant when the autocorrelation function reaches the value of zero for the first time [$\chi(t_0) = 0$]. This choice of t_0 corresponds approximately to the relaxation time $t_0 \approx \tau$, for which Eq. (11) should give accurate results. As an initial guess we recommend setting $\alpha = 0$ and $\beta = 0$. Otherwise, the least-square fitting of Eq. (11), which is a flexible expression with four unknown parameters, may return suboptimal results.

A simple least-square fitting of the phenomenological Eq. (11) underestimates standard errors of the parameter values. We report more realistic estimates, which are recalculated by using the optimized parameter values and the uncertainty of the time autocorrelation data [29]:

$$\Delta\chi \approx \frac{1-\chi^2}{\sqrt{n-2}}. \tag{C6}$$

To use the above uncertainty in the weighted least-squares fitting, we discarded the point $\chi(0)$, which equals 1 by definition. The parameter error bounds were then recalculated from the Fisher information matrix.

APPENDIX D: SIMULATION ALGORITHM

For computational experiments we convert Eq. (1) into an equivalent two-dimensional dynamical system $\boldsymbol{X} = (x, y) = (x, \dot{x})$:

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -ay - bx - cx^3 + f(t). \end{aligned} \tag{D1}$$

We adopt a second-order operator-splitting approach for stochastic systems [24, Appendix C] by decomposing the time-evolution operator $\mathcal{T}$ as

$$\dot{\boldsymbol{X}} = \mathcal{T}\boldsymbol{X} = (\mathcal{T}_{y,x} + \mathcal{T}_{x,y} + \mathcal{T}_{y,y})\boldsymbol{X}, \tag{D2}$$

in which

$$\begin{aligned} \mathcal{T}_{y,x} &= y\partial_x, \quad \mathcal{T}_{y,y} = -ay\partial_y, \\ \mathcal{T}_{x,y} &= (f - bx - cx^3)\partial_y. \end{aligned}$$

The formal solution of Eq. (D2) for a time step Δt is

$$\boldsymbol{X}(t + \Delta t) = \exp(\mathcal{T}\Delta t)\boldsymbol{X}(t),$$

in which the time-evolution operator can be approximated by

$$\begin{aligned} &\exp[\mathcal{T}\Delta t + O(\Delta t^2)] \\ &\quad = \exp\left(\frac{\mathcal{T}_{y,x}\Delta t}{2}\right)\exp\left(\frac{\mathcal{T}_{y,y}\Delta t}{2}\right) \\ &\qquad \times \exp(\mathcal{T}_{x,y}\Delta t)\exp\left(\frac{\mathcal{T}_{y,y}\Delta t}{2}\right)\exp\left(\frac{\mathcal{T}_{y,x}\Delta t}{2}\right). \end{aligned} \tag{D3}$$

The action of an individual operator of the form $\exp(\mathcal{L}\Delta t)$ can be inferred by solving the simplified dynamics

$$\dot{\boldsymbol{X}}(t) = \mathcal{L}\boldsymbol{X}(t) \Rightarrow \boldsymbol{X}(t + \Delta t) = \exp(\mathcal{L}\Delta t)\boldsymbol{X}(t). \tag{D4}$$

The composite operator (D3) then leads to the following algorithm for the numerical integration of Eq. (D2):

$$x(t + \Delta t/2) = x(t) + y(t)\Delta t/2, \tag{D5}$$

$$\begin{aligned} &y(t + \Delta t) \\ &\quad = y(t)\exp(-a\Delta t) \\ &\qquad - \exp(-a\Delta t/2)[b + cx(t + \Delta t/2)^2]x(t + \Delta t/2)\Delta t \\ &\qquad + \exp(-a\Delta t/2)\int_t^{t+\Delta t} dt\, f(t), \end{aligned} \tag{D6}$$

$$x(t + \Delta t) = x(t + \Delta t/2) + y(t + \Delta t)\Delta t/2. \tag{D7}$$

The simulations reported in Sec. V lasted 5×10^5 time steps of a size $\Delta t = 0.01$ (in the reduced units) for each value of the parameter A. Statistics were calculated from single trajectories sampled at time intervals 0.05, which deliver 10^5 observations.

[1] A. A. Khan and N. S. Vyas, Nonlin. Dyn. **24**, 285 (2001).
[2] A. Chatterjee and N. S. Vyas, J. Vib. Acoust. **125**, 299 (2003).
[3] R. Alonso, F. Goller, and G. B. Mindlin, Phys. Rev. E **89**, 032706 (2014).
[4] G. B. Mindlin, Chaos **27**, 092101 (2017).
[5] A. A. Cherevko, A. V. Mikhaylova, A. P. Chupakhin, I. V. Ufimtseva, A. L. Krivoshapkin, and K. Y. Orlov, J. Phys.: Conf. Ser. **722**, 012045 (2016).
[6] D. V. Parshin, I. V. Ufimtseva, A. A. Cherevko, A. K. Khe, K. Y. Orlov, A. L. Krivoshapkin, and A. P. Chupakhin, J. Phys.: Conf. Ser. **722**, 012030 (2016).
[7] A. A. Cherevko, E. E. Bord, A. K. Khe, V. A. Panarin, and K. J. Orlov, J. Phys.: Conf. Ser. **894**, 012012 (2017).
[8] E. Izhikevich and R. FitzHugh, Scholarpedia **1**, 1349 (2006).
[9] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd ed. (Avalon Publishing, New York, 2014).
[10] A. Chatterjee, Int. J. Mech. Sci. **52**, 1716 (2010).
[11] V. N. Smelyanskiy, D. G. Luchinsky, D. A. Timuçin, and A. Bandrivskyy, Phys. Rev. E **72**, 026202 (2005).
[12] Q. He, L. Wang, and B. Liu, Chaos Solitons Fractals **34**, 654 (2007).
[13] G. Quaranta, G. Monti, and G. C. Marano, Mech. Syst. Signal Process. **24**, 2076 (2010).
[14] W. J. Rugh, *Nonlinear System Theory: The Volterra/Wiener Approach* (Johns Hopkins University Press, Baltimore, MD, 1981).
[15] R. L. Peterson, Rev. Mod. Phys. **39**, 69 (1967).
[16] Y. H. Ku and A. A. Wolf, J. Franklin Inst. **281**, 9 (1966).
[17] A similar approach has already been applied in R. Belousov, P. De Gregorio, L. Rondoni, and L. Conti, Physica A **412**, 19 (2014).
[18] R. Belousov, E. G. D. Cohen, C.-S. Wong, J. A. Goree, and Y. Feng, Phys. Rev. E **93**, 042125 (2016).
[19] H. Touchette, Phys. Rep. **478**, 1 (2009).
[20] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[21] R. Belousov and E. G. D. Cohen, Phys. Rev. E **94**, 062124 (2016).
[22] R. Zwanzig and N. K. Ailawadi, Phys. Rev. **182**, 280 (1969).
[23] M. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).
[24] R. Belousov, E. G. D. Cohen, and L. Rondoni, Phys. Rev. E **96**, 022125 (2017).
[25] Stephen Wolfram, Mathematica (Wolfram Research, Inc., Champaign, IL, 2018), http://support.wolfram.com/kb/472.
[26] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems* (Krieger Publishing, Malabar, FL, 2006).
[27] E. Isobe and S. Sato, J. Appl. Prob. **20**, 754 (1983).
[28] A. N. Lagar'kov and V. M. Sergeev, Sov. Phys. Usp. **21**, 566 (1978).
[29] A. L. Bowley, J. Am. Stat. Assoc. **23**, 31 (1928).