

## Analysis of time series from stochastic processes

Janez Gradišek,<sup>1,\*</sup> Silke Siegert,<sup>2</sup> Rudolf Friedrich,<sup>2</sup> and Igor Grabec<sup>1</sup>

<sup>1</sup>*Faculty of Mechanical Engineering, University of Ljubljana, Aškerčeva 6, SI-1000 Ljubljana, Slovenia*

<sup>2</sup>*Institute for Theoretical Physics 3, University of Stuttgart, Pfaffenwaldring 57/5, D-70550 Stuttgart, Germany*

(Received 5 April 2000)

Analysis of time series from stochastic processes governed by a Langevin equation is discussed. Several applications for the analysis are proposed based on estimates of drift and diffusion coefficients of the Fokker-Planck equation. The coefficients are estimated directly from a time series. The applications are illustrated by examples employing various synthetic time series and experimental time series from metal cutting.

PACS number(s): 02.50.Ey, 02.50.Fz, 05.45.-a

### I. INTRODUCTION

All experimental data are to a certain extent contaminated by noise. Generally, one can distinguish between measurement noise and dynamic noise. Measurement noise is caused by measurement procedures; it is superimposed on the measured signal, and it can not influence the dynamics of the process. Dynamic noise is part of the process dynamics and can play an important role when the process is close to an instability point [1]. Time series from various processes in nature are stochastic, and it is often reasonable to assume that they contain both types of noise. However, most nonlinear techniques for time series analysis, especially those inspired by chaos theory [2], require the time series to be generated by a deterministic process, and only allow for negligible measurement noise. Consequently, the applicability of these methods to the analysis of stochastic time series is limited.

A method for analysis of stochastic data sets was recently proposed [3,4] which assumes that the data contain the additive type of dynamic noise. The evolution law of a process which generates such signals can be written as  $\dot{\mathbf{X}}(t) = \mathcal{F}[\mathbf{X}(t)] + \boldsymbol{\eta}(t)$ , where  $\mathbf{X}(t)$  denotes the process state at time  $t$ ,  $\mathcal{F}(\cdot)$  is a nonlinear function, and  $\boldsymbol{\eta}(t)$  denotes noise. The dynamics of this type of processes can also be described by the Fokker-Planck equation, which is determined by drift and diffusion coefficients. Using the method proposed in Refs. [3,4], one can estimate these coefficients directly from noisy data for a certain class of processes. Thus, a complete description of a stochastic process can be found, and the deterministic laws of the process dynamics, as well as the form and the strength of the noise, can be determined.

By extracting drift and diffusion coefficients separately from a time series, the dynamics of a stochastic process are decomposed into a deterministic and a random component. Such decomposition of the dynamics offers several new possibilities for analysis of stochastic processes. The aim of this paper is to demonstrate the following: (a) When the deterministic dynamics' component is represented as a vector field, it provides qualitative information about the local stability properties of the process in the phase space. (b) The

extracted dynamics' components can be used to reconstruct the process evolution. Deterministic evolution can be reconstructed based only on the deterministic component. (c) When both the deterministic and the random component are employed for the reconstruction, one can generate typical stochastic trajectories of the process. These trajectories possess the same deterministic and random properties as the original time series. (d) Typical trajectories can be applied to estimate the mean first passage time, i.e., the mean time elapsed between successive visits of the process trajectory to selected locations in the phase space.

We first illustrate these applications by examples employing synthetic time series generated by (1) a stochastic system exhibiting the pitchfork bifurcation, (2) the stochastic van der Pol oscillator, and (3) the stochastic Lorenz system in a chaotic regime. Next, we use the applications to analyze experimental time series obtained under different regimes of metal cutting.

### II. METHOD FOR ANALYSIS OF STOCHASTIC PROCESSES

The method proposed in Refs. [3,4] is a general method for the estimation of the drift and diffusion coefficients of the Fokker-Planck equation for stationary continuous Markovian stochastic processes. Let the evolution of a continuous  $m$ -dimensional stochastic variable  $\mathbf{X}(t)$  in phase space  $s_{\mathbf{x}}$  be governed by the Langevin equation:

$$\frac{dX_i(t)}{dt} = h_i(\mathbf{X}(t)) + \sum_j g_{ij}(\mathbf{X}(t))\Gamma_j(t), \quad (1)$$

where the fluctuating Langevin forces  $\Gamma_j(t)$  represent random noise, which is assumed to be uncorrelated,  $\langle \Gamma_i(t)\Gamma_j(t') \rangle = Q\delta_{ij}\delta(t-t')$ , with vanishing mean,  $\langle \Gamma_i(t) \rangle = 0$ , for each  $i, j$ . Due to the random term, Eq. (1) can only rarely be solved. Alternatively, the dynamic behavior of the underlying stochastic process can be described by the Fokker-Planck equation, which describes the evolution of the conditional probability density distribution of the stochastic variable  $\mathbf{X}$  in the phase space  $s_{\mathbf{x}}$ :

\*Electronic address: janez.gradisek@fs.uni-lj.si

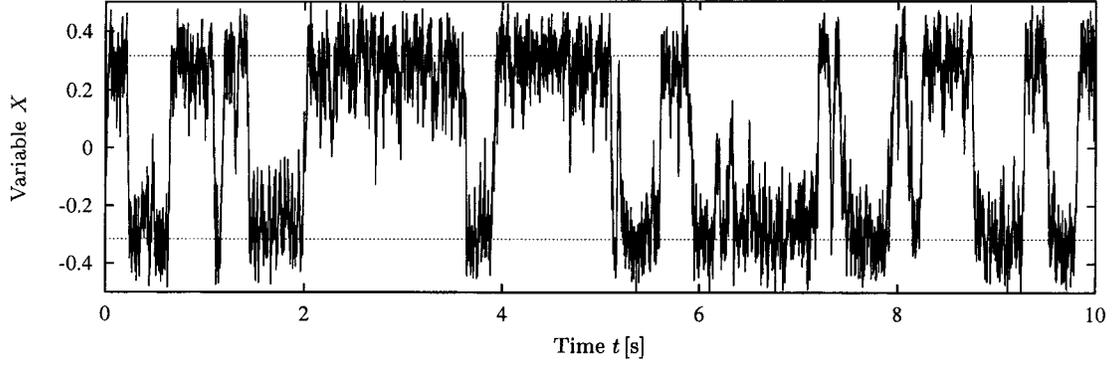


FIG. 1. Time series of a variable  $X$  from Eq. (A1). The dotted lines at  $\pm\sqrt{0.1}$  denote the two stable fixed points.

$$\begin{aligned} \frac{\partial p(\mathbf{x}, t + \tau | \mathbf{x}, t)}{\partial t} = & \left( - \sum_i \frac{\partial}{\partial x_i} D_i^{(1)}(\mathbf{x}, t) \right. \\ & \left. + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} D_{ij}^{(2)}(\mathbf{x}, t) \right) p(\mathbf{x}, t + \tau | \mathbf{x}, t). \end{aligned} \quad (2)$$

$D_i^{(1)}$  and  $D_{ij}^{(2)}$  are called drift and diffusion coefficients, respectively. If Itô's definitions of stochastic integrals are applied [5], the coefficients  $\mathbf{D}^{(1)}$  and  $\mathbf{D}^{(2)}$  can be related to the deterministic and random terms of the Langevin equation as [6]

$$D_i^{(1)}(\mathbf{x}, t) = h_i(\mathbf{x}, t), \quad (3a)$$

$$D_{ij}^{(2)}(\mathbf{x}, t) = Q \sum_k g_{ik}(\mathbf{x}, t) g_{jk}(\mathbf{x}, t). \quad (3b)$$

If coefficients  $\mathbf{D}^{(1)}$  and  $\mathbf{D}^{(2)}$  are estimated from the time series generated by the stochastic process, the dynamics of the process are in fact completely determined. Using Eqs. (3), one can further determine the deterministic term, as well as the form and strength of the fluctuating term in the Langevin equation.

For the class of stationary continuous Markovian processes with uncorrelated dynamical noise, it is always possible to determine drift and diffusion coefficients directly from given data [3,4] by using their statistical definition [6]:

$$D_i^{(1)}(\mathbf{x}, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle X_i(t + \tau) - x_i \rangle_{X(t) = \mathbf{x}}, \quad (4a)$$

$$D_{ij}^{(2)}(\mathbf{x}, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle (X_i(t + \tau) - x_i)(X_j(t + \tau) - x_j) \rangle_{X(t) = \mathbf{x}}. \quad (4b)$$

Here  $\mathbf{X}(t + \tau)$  is a solution of Eq. (1) which starts at  $\mathbf{X}(t) = \mathbf{x}$  at time  $t$ . In practice,  $\mathbf{D}^{(i)}$  can be determined from the following relations for a small time step  $\tau$ :

$$D_i^{(1)}(\mathbf{x}) = T_i^{(1)}(\mathbf{x}, \tau), \quad (5a)$$

$$D_{ij}^{(2)}(\mathbf{x}) = T_{ij}^{(2)}(\mathbf{x}, \tau) - \tau T_i^{(1)}(\mathbf{x}, \tau) T_j^{(1)}(\mathbf{x}, \tau). \quad (5b)$$

The terms

$$T_i^{(1)}(\mathbf{x}, \tau) = \frac{1}{\tau} \int_{-\infty}^{\infty} (y_i - x_i) p(\mathbf{y}, t + \tau | \mathbf{x}, t) \prod_k dy_k, \quad (6a)$$

$$T_{ij}^{(2)}(\mathbf{x}, \tau) = \frac{1}{\tau} \int_{-\infty}^{\infty} (y_i - x_i)(y_j - x_j) p(\mathbf{y}, t + \tau | \mathbf{x}, t) \prod_k dy_k \quad (6b)$$

denote the conditional moments which can be determined by numerical integration. For a stationary process,  $p(\mathbf{y}, t + \tau | \mathbf{x}, t)$  and, consequently,  $\mathbf{T}^{(i)}$  and  $\mathbf{D}^{(i)}$  are independent of time.

We illustrate the method using a one-dimensional time series generated by the Langevin equation valid for systems which exhibit noisy pitchfork bifurcation. A typical segment of a time series is shown in Fig. 1. The corresponding equation and its parameters are described in the Appendix. The estimated drift and diffusion coefficients are shown in Fig. 2 together with their theoretical dependence on  $x$ . The dependences  $D^{(i)}(x)$  estimated from the time series closely follow the theoretical ones. The largest discrepancies between the theoretical and the estimated dependences are observed in the vicinity of  $x=0$ , and at the edges of the phase space. The reason for this lies in the infrequent visits of the process trajectory to these intervals of  $x$  (Fig. 1), which results in less reliable estimates of conditional probability density in these intervals.

Obviously, the better the estimate of conditional probability density, the better the estimates of the coefficients  $\mathbf{D}^{(i)}$  obtained by the method. For the results presented in this paper we estimated the probability density using histograms with equidistant bins.

An important issue related to the estimation of coefficients  $\mathbf{D}^{(i)}$  is their dependence on the time step  $\tau$ . This dependence is discussed in detail elsewhere [7], and only the main results are given below.

*The effect of time step  $\tau$ .* For the sake of simplicity, let us consider a one-dimensional process for which the Fokker-Planck operator  $L$ ,

$$L = - \frac{\partial}{\partial x} D^{(1)}(x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} D^{(2)}(x), \quad (7)$$

possesses real eigenvalues  $\lambda_j \leq 0$  and eigenvectors  $\Phi_j(x)$ . The conditional probability density can be expressed in the form of an infinite series as

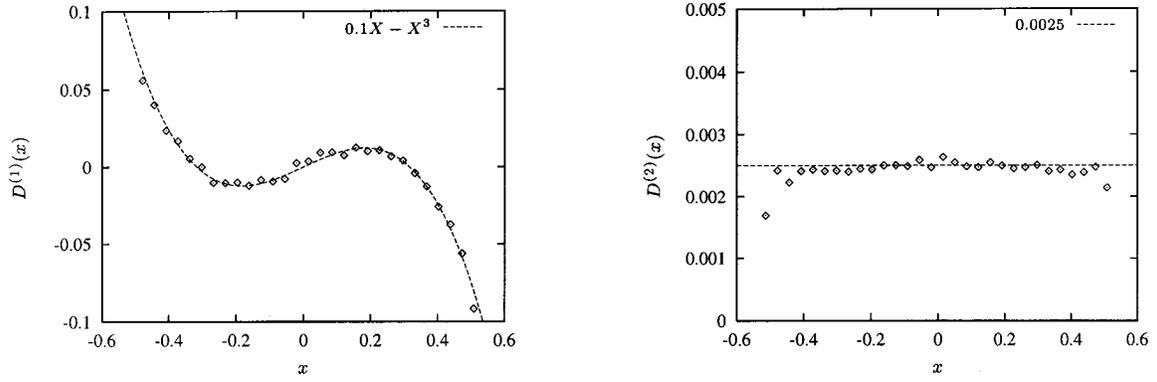


FIG. 2. The estimated drift and diffusion coefficients (diamonds) compared to their theoretical dependence on  $x$  (dashed line).

$$p(y, t' | x, t) = \sum_j e^{\lambda_j(t'-t)} \Phi_j^\dagger(x) \Phi_j(y), \quad (8)$$

where  $t' - t = \tau$ . The conditional moments are given by

$$\begin{aligned} & \langle (y(t') - x(t))^k \rangle_{y(t)=x(t)} \\ &= \sum_j e^{\lambda_j(t'-t)} \Phi_j^\dagger(x) \int_{-\infty}^{\infty} (y-x)^k \Phi_j(y) dy \end{aligned} \quad (9)$$

and the following expansions for drift and diffusion coefficients are obtained [7]:

$$D^{(1)}(x) = \sum_j \lambda_j \Phi_j^\dagger(x) \int_{-\infty}^{\infty} (y-x) \Phi_j(y) dy, \quad (10a)$$

$$D^{(2)}(x) = \sum_j \lambda_j \Phi_j^\dagger(x) \int_{-\infty}^{\infty} (y-x)^2 \Phi_j(y) dy. \quad (10b)$$

When coefficients  $D^{(i)}$  are estimated from data, the infinite series (8)–(10) are approximated by finite ones which contain  $N$  eigenfunctions. In order to obtain accurate information about the conditional probability distribution and coefficients  $D^{(i)}$ , one has to choose a time step  $\tau$  such that  $e^{\lambda_N \tau} > 0$ , for sufficiently large  $N$  [7].

As an example, let us use a time series generated by Eq. (A1). Coefficients  $D^{(i)}$  estimated from the data using different values of  $\tau$  are shown in Fig. 3. At  $\tau = 0.01, 0.1, 1$ , and  $10$ , the estimated drift coefficients match the theoretical values well, whereas for the diffusion coefficient the agreement

with the theoretical values is already lost at  $\tau = 1$ . Indeed, the dependence of root-mean-squared error  $\mathcal{E}$  of the estimates on the time step  $\tau$  confirms that the error of  $D^{(2)}$  estimates grows with increasing  $\tau$  faster than the error of  $D^{(1)}$  estimates (Fig. 4). Our experience suggests that the maximum acceptable time step  $\tau_{\max}$  for accurate estimate of the diffusion coefficient is approximately one order of magnitude smaller than  $\tau_{\max}$  which is acceptable for the estimate of drift coefficient. In terms of Eqs. (10) this means that, if the series are to approximate both coefficients equally well, the number of terms  $N$  in the series which approximates  $D^{(2)}$  must be larger.

In general, the maximum acceptable time step  $\tau_{\max}$  or, equivalently, the number of terms  $N$  required in a series depend on the properties of the process. Based on our experience,  $\tau_{\max}$  for the drift coefficient is usually not much shorter than the time step  $t_i$  required for the integration of the corresponding differential equations. For example, in the case of the van der Pol oscillator [Eqs. (A2)] we found  $\tau_{\max} \approx 0.1$ , which results in approximately 80 points per oscillation cycle, whereas the longest reasonable integration step for which the limit cycle is not yet too distorted is approximately  $t_i \approx 0.3$ .

However, in practice one is often faced with experimental data recorded with a time step for which coefficients with decreasing  $\tau$  do not converge to a limit value. In this case, the estimated coefficients should be considered a crude approximation and treated with caution. The lack of convergence of  $D^{(i)}$  with decreasing  $\tau$  can be regarded as an indication of the non-Markovian properties of the process [7].

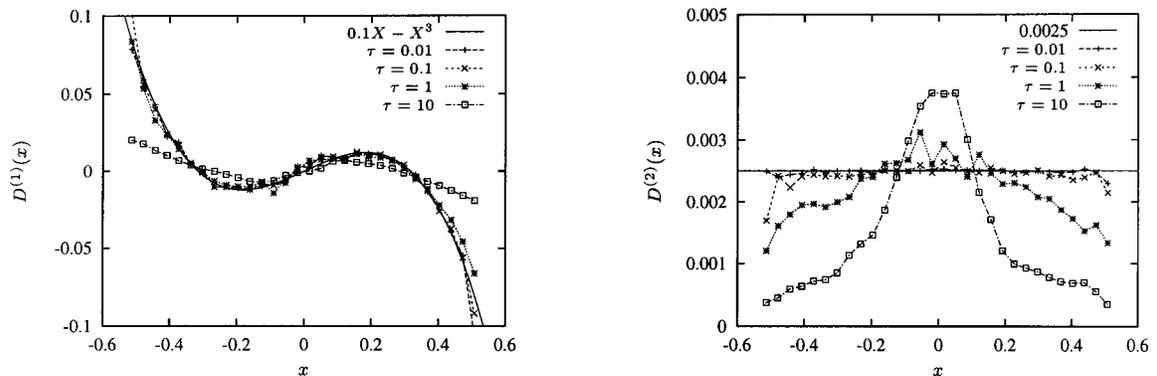


FIG. 3. Coefficients  $D^{(i)}$  estimated using four different time steps  $\tau$ .

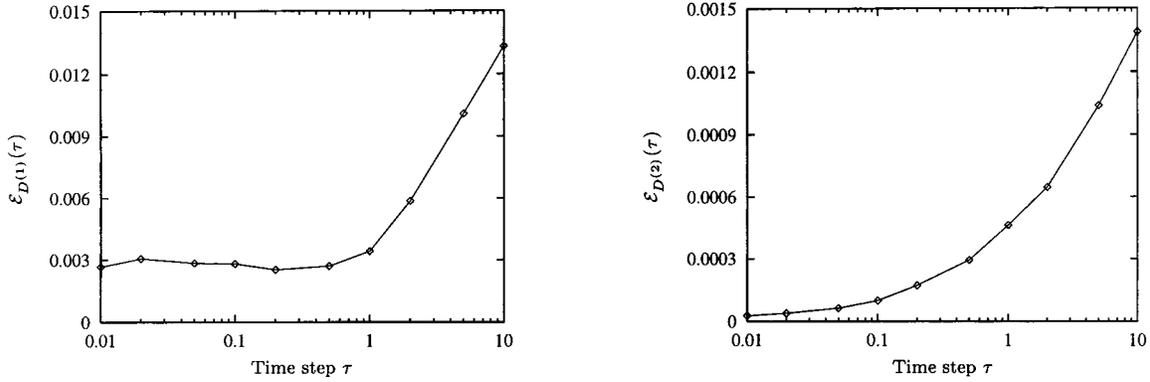


FIG. 4. Root-mean-squared error  $\mathcal{E}$  of the coefficients  $D^{(i)}$  versus the time step  $\tau$ .

III. APPLICATIONS OF THE METHOD

Applications discussed in this section are based mainly on the decomposition of stochastic process dynamics into a deterministic and a random component. For stationary continuous Markovian stochastic processes, such decomposition is performed by estimating the drift and diffusion coefficients of the Fokker–Planck equation as described in Sec. II. According to the relations in Eqs. (3), the drift coefficient  $\mathbf{D}^{(1)}$  corresponds to the deterministic component [Eq. (3a)], while the diffusion coefficient  $\mathbf{D}^{(2)}$  is related to the random component of process dynamics [Eq. (3b)].

Applications are illustrated by examples employing data sets from (a) a stochastic system exhibiting the pitchfork bifurcation, (b) a stochastic van der Pol oscillator, and (c) a stochastic Lorenz system. The corresponding Langevin equations and their parameters are given in the Appendix.

A. Deterministic vector field

When studying the dynamics of a stochastic process it is interesting to explore how the process would evolve if it was subject to no random influences. This evolution is governed by the deterministic component of the stochastic process dynamics, which in our case is determined by the drift coefficient of the Fokker–Planck equation.

Having extracted the drift coefficient  $\mathbf{D}^{(1)}$  from a stochastic time series, one can plot its components  $D_i^{(1)}$  separately versus the phase space coordinates. Such a plot is shown in Fig. 5 for a stochastic van der Pol oscillator [Eqs. (A2)].  $\mathbf{D}^{(1)}$  was estimated from a vector time series, a portion of which is

shown in the left-hand panel of Fig. 6. Except for the regions where there is no data, the estimates agree well with the theoretical values. However, for an experimentally studied process, for which the dynamics equations are not known, coefficients presented in this way are not easy to interpret.

A more informative presentation of the drift coefficient  $\mathbf{D}^{(1)}$  is achieved by plotting it as a vector field (the right-hand panel of Fig. 6). An arrow in the field graph represents a value of the drift coefficient  $\mathbf{D}^{(1)}(\mathbf{x})$  estimated at location  $\mathbf{x}$  in phase space. The orientation of the arrow indicates the average direction of deterministic motion at  $\mathbf{x}$ . In the case of a van der Pol oscillator (the right-hand panel of Fig. 6), arrows in the vector field point on average in the clockwise direction, suggesting motion on a stable nonsymmetric limit cycle in that direction. Arrows outside the limit cycle run approximately parallel to the cycle around most of the cycle, except at the upper left and at the bottom right corners, where they point towards the cycle. This indicates that the dissipation is close to zero around most of the cycle, and strongly negative at the two corners. The arrows inside the limit cycle point out towards the cycle, suggesting that an unstable fixed point is located in the center of phase space.

The example shows that examination of the drift coefficient presented as a field yields information about the local stability properties of the process. However, the information obtained from the drift coefficient is limited to the region visited by the trajectory during experiments. In order to get information about process properties in other regions of phase space, one should randomly disturb the process during experiments to make it explore a greater portion of its phase space.

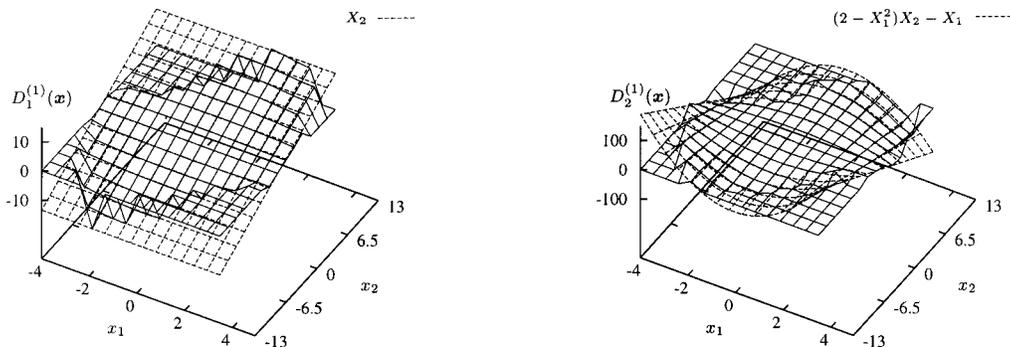


FIG. 5. Components of drift coefficient  $\mathbf{D}^{(1)}(\mathbf{x})$  for a stochastic van der Pol oscillator. Solid grid, estimated values. Dashed grid, theoretical values.

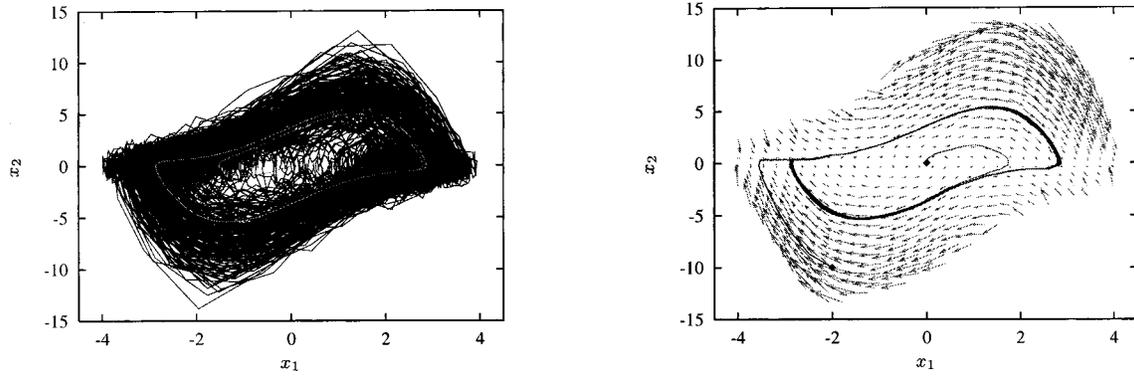


FIG. 6. Left: Phase portraits of a stochastic and a deterministic van der Pol oscillator. Right: The estimated drift coefficient is shown as a vector field with two trajectories superimposed. The trajectories were integrated using the estimated drift coefficient.

### B. Reconstruction of deterministic dynamics

The relationship between the drift coefficient of the Fokker–Planck equation and the deterministic term of the Langevin equation can also be exploited to reconstruct the deterministic dynamics of the process. If we drop the random term of the Langevin equation (1), we can solve the truncated equation numerically using the estimated  $\mathbf{D}^{(1)}$  as

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \mathbf{D}^{(1)}(\mathbf{X}(t))\Delta t. \quad (11)$$

The solution represents a deterministic trajectory of the process. Two such solutions are shown superimposed on the vector field in the right-hand panel of Fig. 6. The initial conditions of these trajectories (marked by a diamond) lie outside or inside the limit cycle, respectively. Both trajectories terminate on the stable limit cycle after a transient period. For the purpose of comparison, a theoretical deterministic trajectory is shown superimposed on the stochastic phase portrait in the left-hand panel of Fig. 6. This trajectory is a numerical solution of the deterministic term of the equations governing the dynamics of the oscillator [Eqs. (A2)]. The deterministic trajectories obtained by integrating Eq. (11) agree closely with the theoretical deterministic trajectory.

To illustrate that the method is not restricted to trivial attractors, such as fixed points and limit cycles, a vector time series from the stochastic Lorenz system in a chaotic regime was analyzed [Eqs. (A3)]. A portion of the stochastic time series and the extracted drift coefficient are shown in the left-hand and the right-hand panels of Fig. 7. The drift coef-

ficient was used to generate a deterministic trajectory, which is plotted superimposed on the vector field. A trajectory obtained by solving numerically the deterministic equations of the Lorenz system [Eqs. (A3)] is shown superimposed on the stochastic trajectory. Although both deterministic trajectories start at the same point in phase space, they do not follow the same path (the bottom traces in Fig. 8). This discrepancy stems from the imperfect estimation of conditional probability density and from the chaotic nature of the Lorenz system. However, both deterministic trajectories possess similar characteristic patterns, and the phase portraits formed by the trajectories are qualitatively similar (Fig. 7). The main difference can be seen in the inner regions of the two lobes, where the probability density was presumably poorly estimated.

In order to gain an impression of the influence of dynamic noise on the Lorenz system dynamics, compare the upper two traces in Fig. 8 which correspond to the original stochastic and deterministic trajectories. In the stochastic case, noise drives the trajectory from one lobe of the attractor to the other more frequently, and thus prevents the trajectory spiraling in a particular lobe for a protracted period, as is the case in the deterministic system. Nevertheless, the spiraling typical of the Lorenz attractor is clearly observed in the reconstructed deterministic trajectory (the bottom trace). Moreover, note the difference between the extent of the deterministic attractor and the vector field in the right-hand panel of Fig. 7. Dynamic noise increases the extent of the attractor from 45% in the  $x_1$  direction up to 83% in the  $x_3$  direction. Although the noise repeatedly drives the trajectory away

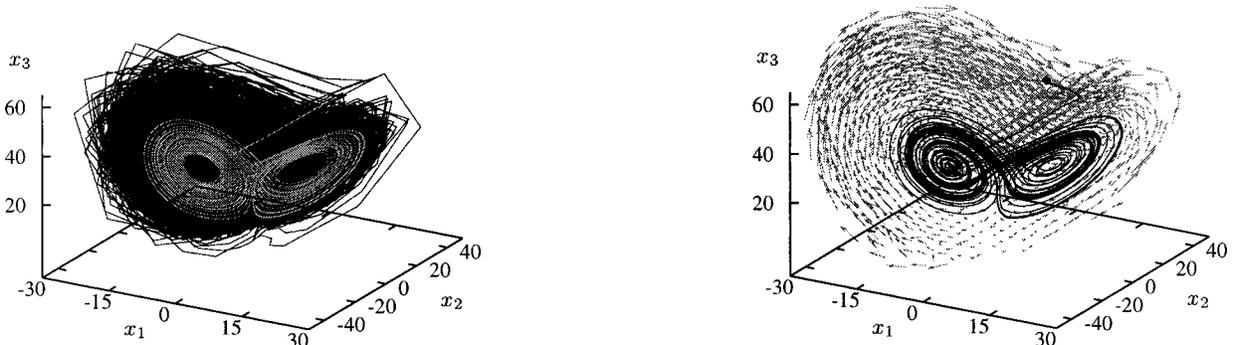


FIG. 7. Left: Phase portraits of a stochastic and a deterministic Lorenz system. Right: The estimated drift coefficient is shown as a vector field with a trajectory superimposed. The trajectory was integrated using the estimated drift coefficient.

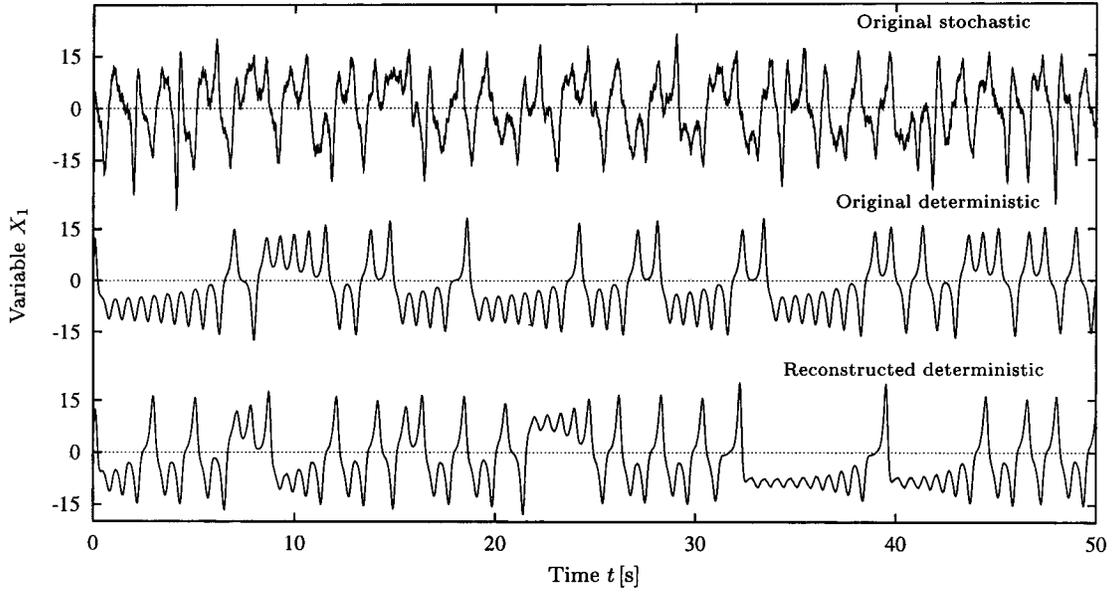


FIG. 8. Comparison of the original stochastic trajectory with the original and reconstructed deterministic trajectories of the Lorenz system.

from the deterministic attractor, the trajectory generated by the extracted drift coefficient forms an attractor which resembles the shape and the extent of the original deterministic attractor. Hence using our method, we obtain information about the deterministic properties of the process from stochastic data, although the dynamics of the process are significantly altered by the noise.

### C. Reconstruction of stochastic dynamics

In order to reconstruct a stochastic trajectory which possesses deterministic and random properties similar to those of the studied process, drift and diffusion coefficients must be estimated from the stochastic data set. A stochastic trajectory can be obtained from the Langevin equation as

$$X_i(t + \Delta t) = X_i(t) + D_i^{(1)}(\mathbf{X}(t))\Delta t + \sqrt{\Delta t} \sum_{j=1}^i g_{ij}(\mathbf{X}(t))\Gamma_j(t), \quad (12)$$

where  $\Gamma_j(t)$  represents uncorrelated noise,  $\langle \Gamma_i(t)\Gamma_j(t') \rangle$

$= Q\delta_{ij}\delta(t-t')$ , with vanishing mean,  $\langle \Gamma_i(t) \rangle = 0$ , for each  $i, j$ . Following Eq. (3b), the noise amplitudes  $g_{ij}$  can be calculated from

$$Q\mathbf{g}\mathbf{g}^\dagger = \mathbf{D}^{(2)} \quad (13)$$

using the Cholesky decomposition [8], p. 96, iff. (i)  $\mathbf{D}^{(2)}$  is positive definite and symmetric, and (ii)  $\mathbf{g}$  is a lower triangular matrix. Diffusion coefficient fulfills the condition (i), and we assume that  $\mathbf{g}$  is of a lower triangular form. The sum in Eq. (12) therefore includes only the first  $i$  terms.

For the van der Pol oscillator, the original and the reconstructed stochastic trajectories are shown in Fig. 9. As expected, the trajectories are not the same because the noise time series used to generate them are different. However, both stochastic trajectories are qualitatively similar and they possess the same deterministic and random properties.

For the Lorenz system, the original and the reconstructed stochastic trajectories are compared in Fig. 10. Again, both trajectories match qualitatively.

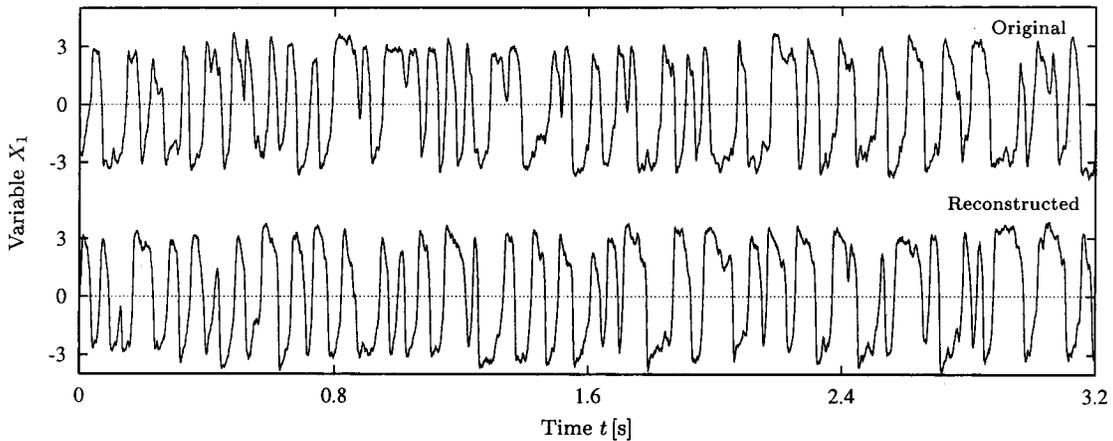


FIG. 9. Comparison of the original and reconstructed stochastic trajectories of the van der Pol oscillator.

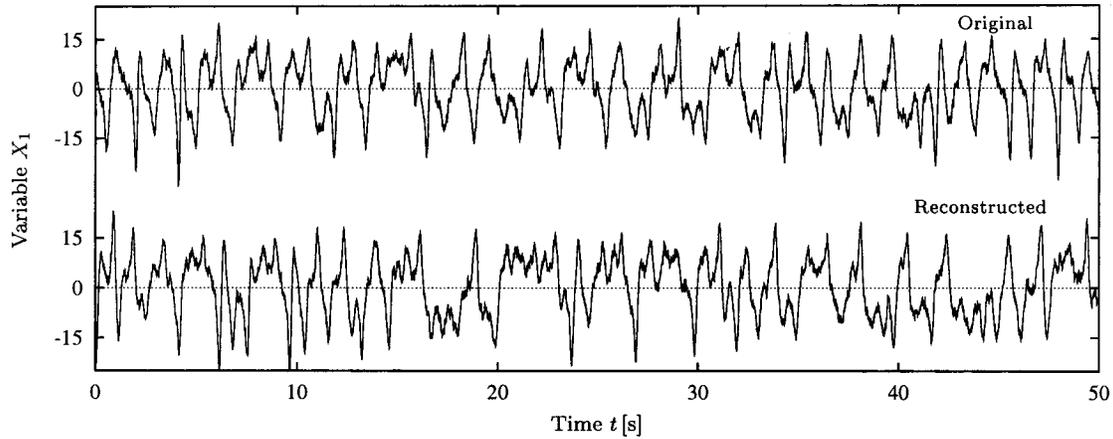


FIG. 10. Comparison of the original and reconstructed stochastic trajectories of the Lorenz system.

The possibility of generating a stochastic trajectory which resembles the deterministic as well as the random properties of the original trajectory can be applied effectively for various purposes. One of them is discussed below.

*Estimation of the mean first passage time.* The mean first passage time denotes the mean period between successive visits of the process trajectory to a selected location in phase space. Suppose that we had measured a trajectory in which only a few passages of interest were observed. To estimate reliably the mean first passage time, we would need many such passages. One solution to this problem which does not involve additional measurements is to extract drift and diffusion coefficients from the measured trajectory, and to reconstruct a stochastic trajectory which is sufficiently long to contain enough passages.

As an example, we again use a stochastic process which exhibits a pitchfork bifurcation [Eq. (A1)]. The time series generated by the process (Fig. 1) can be considered as a trace of a randomly disturbed particle in a one-dimensional two-well potential. Our goal is to estimate the mean time needed by the particle to pass from the well at  $x = -\sqrt{\epsilon}$  to the well at  $x = \sqrt{\epsilon}$ . Based on a time series containing only 3 passages, a stochastic trajectory was reconstructed containing 200 passages. In Fig. 11(a), the distribution of the estimated passage

times is compared to the distribution obtained from the original trajectory containing 200 passages. The overall shapes of both distributions agree quite well, although the curves differ in details. Still, the mean first passage times, marked by the vertical lines, are similar.

The passage times depend on the noise amplitude  $g$ . In the case of a particle in a two-well potential, the higher the noise amplitude, the shorter the mean passage time. The dependences of the mean first passage times on the noise amplitude obtained from the original and reconstructed trajectories are compared in Fig. 11(b). The agreement between the two dependences is very good.

#### IV. ANALYSIS OF EXPERIMENTAL DATA

In this section we apply the method to analyze experimental time series measured in different regimes of metal cutting [9]. The dynamics of metal cutting involve several nonlinear dynamic phenomena, such as material flow and fracture, friction between the tool and the workpiece, coupled vibrations of a machine–tool–workpiece assembly, etc. By varying the cutting parameters, dynamically different cutting regimes can be achieved. For example, if the cutting depth is increased over a certain critical value, self-excited large-

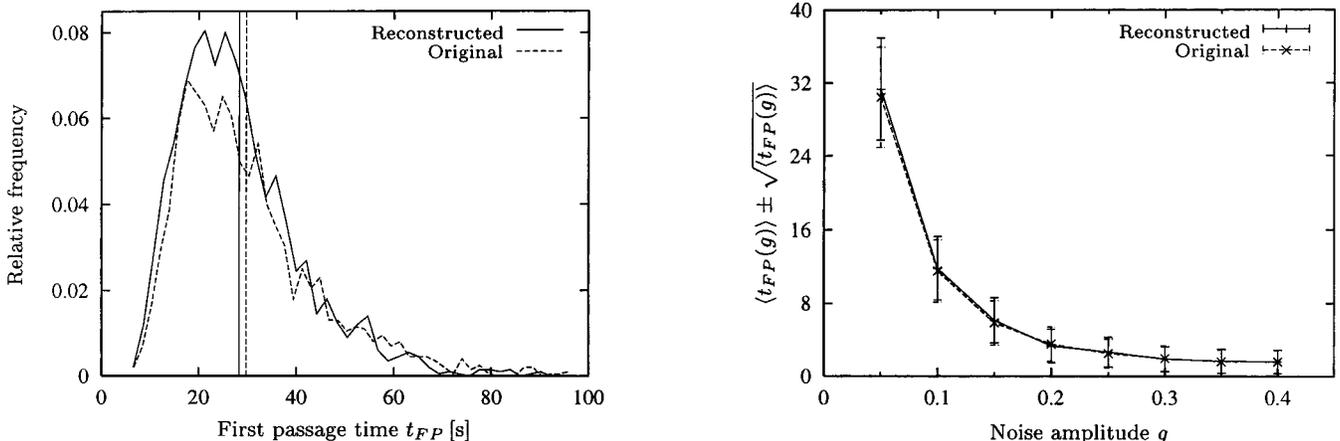


FIG. 11. (a) Comparison of distributions of first passage times from  $x = -\sqrt{\epsilon}$  to  $x = \sqrt{\epsilon}$  at noise amplitude of  $g = 0.05$  for the original and reconstructed trajectories. The vertical lines denote the mean values. (b) A comparison of dependences of the mean first passage times on the noise amplitude  $g$  for the original and the reconstructed trajectories.

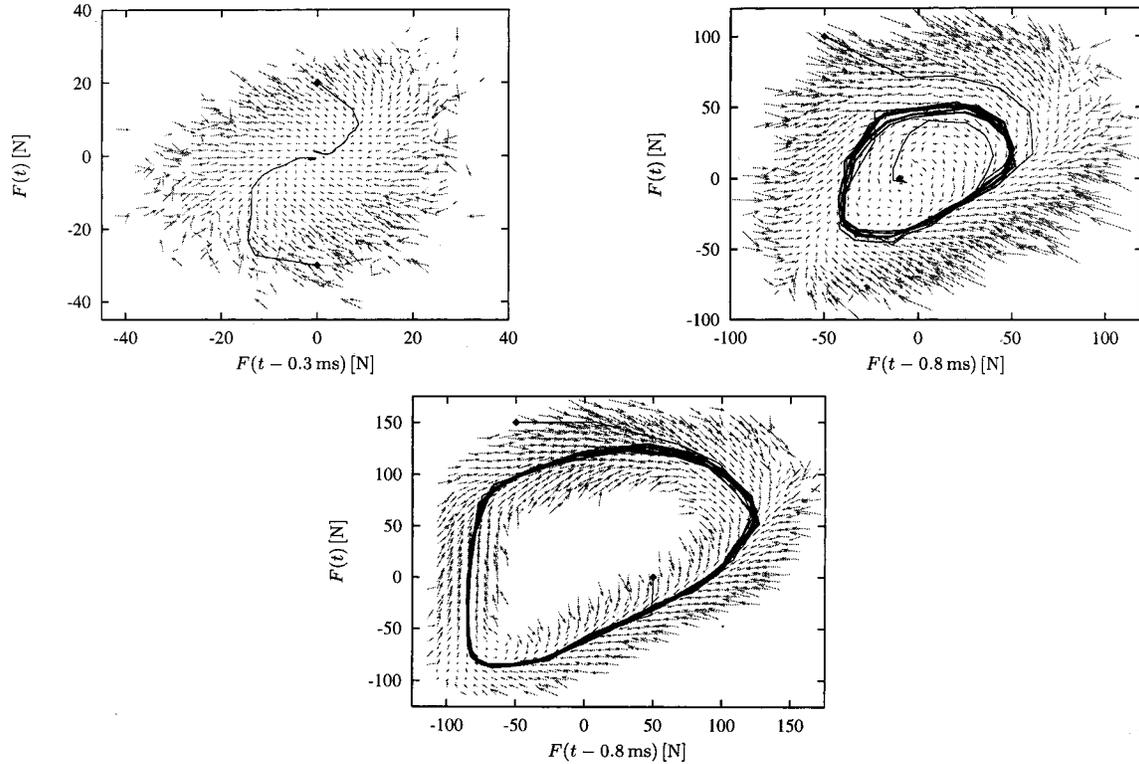


FIG. 12. Two-dimensional projections of drift coefficients for chatter-free (left), weak (right), and strong chatter regimes (bottom). The superimposed trajectories were integrated using the corresponding drift coefficients.

amplitude vibrations of the machine–tool–workpiece assembly may develop. Cutting accompanied by such vibrations is known as chatter. With its detrimental effect on the workpiece, tool, and machine, chatter has been studied intensively in recent decades. Analysis of simple nonlinear models of the cutting process has revealed that the onset of chatter can be described as a subcritical Hopf bifurcation [10,11,9]. Such a description of chatter onset has been confirmed by bifurcation diagrams obtained experimentally [12–14]. However, evidence for such a description based on analysis of measured time series has been lacking, presumably due to the stochastic nature of the process.

We chose turning on a lathe as an example of the cutting process, where a rotating workpiece is cut by a fixed tool. We analyzed the time series recorded during three cutting regimes denoted as (a) chatter-free cutting, (b) cutting accompanied by weak chatter, and (c) cutting accompanied by strong chatter. Drift coefficients (Fig. 12) were estimated in a three-dimensional phase space reconstructed from the scalar time series using the delay coordinates [15]. The trajectories, which are shown superimposed on the fields, were integrated according to Eq. (11). In the case of the chatter-free regime (the top left panel), the field arrows point towards the center of the phase space. This indicates the existence of a stable fixed point at  $(0,0,0)$ . Both trajectories which start at the edge of the phase portrait end up close to the assumed stable fixed point. In chatter regimes, a stable nonsymmetric limit cycle is present. In the weak chatter regime (the top right panel), the field arrows inside the limit cycle point out towards the cycle, which indicates that an unstable fixed point might be located in the center of the phase space. This is also shown by the path of the trajectory starting in the center of the phase

space. The unstable fixed point cannot be observed in the case of strong chatter (bottom panel), because the process trajectory does not visit the center region of the phase space.

Closer examination of drift coefficients in the chatter regimes reveals interesting changes to the local dissipation in the phase space. The average inclination of the arrows towards the limit cycle is much greater in the lower right portion of the limit cycle than in the upper left portion. The two portions of the limit cycle correspond to the tool motion towards and away from the workpiece, respectively. Different average inclinations presumably result from the dependence of damping on the relative direction of the tool motion with respect to the workpiece.

To check the stability of the observed fixed point quantitatively, the drift coefficients were approximated using third order polynomials. Eigenvalues  $\lambda_i$  of the Jacobian matrix evaluated at the fixed point  $(0,0,0)$  are listed in Table I. The real part of the pair of the largest  $\lambda_i$  is negative in chatter-free cutting and positive in both chatter regimes. This confirms that the fixed point is stable during chatter-free cutting and unstable during chatter. Such dependence of the real part of  $\lambda_{1,2}$  is typical of the Hopf bifurcation from a stable fixed point to a stable limit cycle [16].

TABLE I. Stability coefficients of the fixed point  $(0,0,0)$ , calculated from the equations which approximate  $\mathbf{D}^{(1)}$ .

Cutting regime	$\lambda_1$	$\lambda_2$	$\lambda_3$
Chatter-free	$-2.19 + i3.58$	$-2.19 - i3.58$	$-2.95$
Weak chatter	$0.11 + i1.57$	$0.11 - i1.57$	$-7.62$
Strong chatter	$0.48 + i1.86$	$0.48 - i1.86$	$0$

## V. DISCUSSION AND CONCLUSIONS

Dynamics noise often represents a serious obstacle to analysis of process dynamics. However, for stochastic processes that obey the Fokker–Planck equation, the laws of deterministic dynamics, and the form and strength of the fluctuations can be determined by estimating the drift and diffusion coefficients of the Fokker–Planck equation directly from a time series [3,4]. The estimates of the coefficients are independent of the time step applied for the estimation of the conditional probability distribution, provided that the time step is sufficiently short. Numerical study has revealed that drift coefficient estimates converge to a limit value at time steps one order of magnitude longer than estimates of diffusion coefficient. The time step, which is necessary for the estimates to converge, depends on the properties of the process. We found the necessary time step to be shorter than the time step required for the integration of the corresponding differential equations. When analyzing experimental data, the convergence of the estimates with decreasing time step should be established. Such a convergence test may also serve as a criterion for selection of an appropriate sampling time. A lack of convergence of the estimated coefficients with decreasing time step indicates non-Markovian properties of the process [7].

Since drift and diffusion coefficients correspond to the deterministic and random parts of the process dynamics, respectively, their separate estimation from data is equivalent to decomposition of the process dynamics into a deterministic and a random component. This decomposition served as a basis for the applications of stochastic data analysis presented in this paper. (a) The drift coefficient plotted as a vector field represents the deterministic dynamics of the process. Closer examination of the field can reveal additional information on the local stability properties of the process in the phase space. (b) Drift and diffusion coefficients can be employed to reconstruct the process dynamics governed by the Langevin equation. When only the drift coefficient is used to solve the truncated Langevin equation, a deterministic solution is obtained. This solution represents the trajectory of the process which would be observed in the absence of random fluctuations. Although remotely similar, such a deterministic solution should not be mistaken for filtering of dynamic noise, since the process would evolve differently under the same deterministic laws if the random fluctuations were present. (c) If both drift and diffusion coefficients are used for the integration of the Langevin equation, a representative stochastic trajectory of the process can be obtained. Stochastic trajectories can be applied as surrogate process trajectories, since they possess the same deterministic and random properties as the original process. (d) We have used the representative trajectories to estimate the mean first passage time in the case where the original trajectory contained only a few passages of interest.

Finally, experimental data sets acquired in three typical regimes of metal cutting were analyzed. Based on estimated drift coefficients, it was shown that cutting dynamics in the chatter-free regime could be described as random fluctuations around a stable fixed point, while in the two chatter regimes the fluctuations occur around a stable limit cycle. The drift coefficients were approximated by third order poly-

nomials, and the stability of the fixed point was assessed quantitatively. It was found that the transition from a chatter-free to a chatter regime corresponds to the Hopf bifurcation from a stable fixed point to a stable limit cycle. Such a description of the transition is in accordance with the analytical and qualitative experimental results.

Time series from metal cutting have often been analyzed using the methods of nonlinear time series analysis [17–21]. The dynamics of cutting with chatter were mostly described as low dimensional, while descriptions of chatter-free cutting dynamics ranged from linearly correlated random to low-dimensional chaotic. Such differences in descriptions can be attributed either to different experimental setups or to incautious use of the analysis methods. However, most researchers have reported a substantial level of noise in their data. Based on our experience [20], we suspect that different descriptions of the cutting dynamics might also have resulted due to dynamic noise in the measured time series. We have shown in this paper that dynamic noise can significantly broaden a limit cycle attractor or dramatically distort a chaotic attractor (see left panels of Figs. 6 and 7). Disregarding the effect of dynamic noise in these two cases, one might be led to search for a complicated structure in the noisy limit cycle attractor, and to overlook the underlying structure of the noisy chaotic attractor.

Only dynamic noise was considered in our study, while the influence of measurement noise has been neglected. However, by evaluating drift and diffusion coefficients the influence of measurement noise can be estimated for systems, where the dynamic noise level exceeds the measurement noise level [22].

In summary, we have proposed applications for analysis of time series from stochastic processes governed by a Langevin equation. Since the measured time series are very likely to be stochastic, the applications could complement other nonlinear time series analysis techniques, especially in the cases where dynamic noise in the process cannot be neglected.

## ACKNOWLEDGMENTS

The authors thank Tomaž Klinc for making them aware of the Cholesky decomposition [8, p. 96]. J.G. and I.G. gratefully acknowledge the support of the Volkswagen Foundation, EU COST Action P4, and the Ministry of Science and Technology of Slovenia.

## APPENDIX: SYNTHETIC DATA SETS

In the following, the uncorrelated noise  $\Gamma(t)$  is Gaussian distributed with zero mean and variance equal to one.

### 1. Pitchfork bifurcation

The dynamics of a stochastic process exhibiting the pitchfork bifurcation is governed by the following Langevin equation:

$$\dot{X}(t) = \epsilon X(t) - X(t)^3 + g\Gamma(t). \quad (\text{A1})$$

We chose  $\epsilon = 0.1$  and  $g = 0.05$ .

## 2. van der Pol oscillator

The dynamics of the stochastic van der Pol oscillator are governed by

$$\dot{X}_1 = X_2, \quad (\text{A2a})$$

$$\dot{X}_2 = (\epsilon - X_1^2)X_2 - X_1 + g\Gamma(t). \quad (\text{A2b})$$

We chose  $\epsilon=2$  and  $g=3$ .

## 3. Lorenz system

The stochastic Lorenz system is governed by the following system of equations:

$$\dot{X}_1 = \sigma(X_2 - X_1) + \sum_j g_{1j}\Gamma_j(t), \quad (\text{A3a})$$

$$\dot{X}_2 = X_1(r - X_3) - X_2 + \sum_j q_{2j}\Gamma_j(t), \quad (\text{A3b})$$

$$\dot{X}_3 = X_1X_2 - bX_3 + \sum_j g_{3j}\Gamma_j(t), \quad (\text{A3c})$$

where the parameters  $\sigma=10$ ,  $r=28$ , and  $b=8/3$  were chosen so as to assure chaotic regime of the deterministic process. The matrix  $\mathbf{g}$  was

$$\mathbf{g} = \begin{bmatrix} 4 & 5 & 3 \\ 5 & 5 & 6 \\ 3 & 6 & 10 \end{bmatrix}. \quad (\text{A4})$$

- 
- [1] H. Haken, *Advanced Synergetics*, 3rd ed. (Springer-Verlag, Berlin, 1993).
- [2] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Vol. 7 of *Cambridge Nonlinear Science Series* (Cambridge University Press, Cambridge, England, 1997).
- [3] R. Friedrich and J. Peinke, *Physica D* **102**, 147 (1997).
- [4] S. Siegert, R. Friedrich, and J. Peinke, *Phys. Lett. A* **243**, 275 (1998).
- [5] K. Itô, *Nagoya Math. J.* **1**, 35 (1950).
- [6] H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, Berlin, 1989).
- [7] J. Gradišek, S. Siegert, R. Friedrich, and I. Grabec (in preparation).
- [8] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C, The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, England, 1994).
- [9] J. Gradišek, Ph.D. thesis Dr/227, Faculty of Mechanical Engineering, University of Ljubljana, Ljubljana, Slovenia, 1999 (unpublished) (in Slovene).
- [10] G. Stépán and T. Kalmár-Nagy, in *Proceedings of the 16th ASME Biennial Conference on Mechanical Vibration and Noise*, 1997 ASME Design and Technical Conferences, Sacramento, California, 1997, pp. 1–11.
- [11] A. H. Nayfeh, C.-M. Chin, and J. R. Pratt, *ASME J. Manuf. Sci. Eng.* **119**, 485 (1997).
- [12] C. J. Hooke and S. A. Tobias, in *Proceedings of the 4th International Machine Tools Design and Research Conference* (Pergamon, Manchester, 1963), pp. 97–109.
- [13] N. H. Hanna and S. A. Tobias, *ASME J. Eng. Ind.* **96**, 247 (1974).
- [14] J. R. Pratt, M. A. Davies, C. J. Evans, and M. D. Kennedy, *Ann. CIRP* **48**, 39 (1999).
- [15] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, *Phys. Rev. Lett.* **45**, 712 (1980).
- [16] S. H. Strogatz, *Nonlinear Dynamics and Chaos* (Addison-Wesley, Reading, PA, 1994).
- [17] E. Govekar and I. Grabec, in *Progress in Acoustic Emission* (The Japanese Society for NDI, 1994), Vol. VII, pp. 613–618.
- [18] S. T. S. Bukkapatnam, A. Lakhtakia, and S. R. T. Kumara, *Phys. Rev. E* **52**, 2375 (1995).
- [19] B. S. Berger, I. Minis, Y. H. Chen, A. Chavali, and M. Rokni, *J. Sound Vib.* **184**, 936 (1995).
- [20] J. Gradišek, E. Govekar, and I. Grabec, *Mech. Syst. Signal Process.* **12**, 839 (1998).
- [21] M. A. Johnson, Ph.D. thesis, Cornell University, Ithaca, New York, 1996 (unpublished).
- [22] J. Peinke (private communication).