Analysis of stochastic time series in the presence of strong measurement noise

B. Lehle

1Hohewartstr. 44, D-70469 Stuttgart, Germany

A new approach for the analysis of Langevin-type stochastic processes in the presence of strong measurement noise is presented. For the case of Gaussian distributed, exponentially correlated measurement noise it is possible to extract the strength and the correlation time of the noise as well as polynomial approximations of the drift and diffusion functions from the underlying Langevin equation.

PACS numbers: 05.40.Ca, 02.50.Ey
Keywords: Measurement noise, Stochastic processes

I. INTRODUCTION

In the last years there has been significant progress in the analysis and characterization of the dynamics of processes underlying the time series of complex dynamical systems [1–3]. If the temporal evolution of a quantity \( X_t \) can be described by a Langevin equation, it is possible to extract drift and diffusion functions of the underlying stochastic process from a given time series. This can be done because the moments of the conditional probability densities of \( X_{t+\tau} | X_t = x \) can be related to these functions.

Since this approach was introduced [4–8] it has been successfully carried out in a broad range of fields. For example for data from financial markets [9], traffic flow [10], chaotic electrical circuits [11–12], human heart beat [13], climate indices [14–15], turbulent fluid dynamics [16], and for electroencephalographic data from epilepsy patients [17–18].

Real-world data, however, also give rise to some problems. One of them is, that experimental data are only given with a finite sampling rate. So methods had to be proposed to deal with the effects arising from this fact [19–21]. Another problem is the virtually unavoidable measurement noise [12–23]. In the presence of measurement noise \( Y_t \) the values of \( X_t \) or any of its probability densities are no longer accessible, but only \( X_t^* = X_t + Y_t \) and its density distributions.

Recently an approach has been presented which allows the estimation of drift and diffusion functions in the presence of strong delta-correlated, Gaussian noise [24–25]. Starting with initial estimates for the noise strength and the drift and diffusion functions a functional of these unknowns is iteratively minimized.

The aim of this paper is to introduce an alternative approach to the treatment of strong measurement noise. This approach is able to deal also with exponentially correlated, Gaussian noise. The basic idea is, not to look at the conditional moments in the first place but at the joint probability density \( \rho(x_1, x_2, \tau) \) of pairs \((X_t, X_{t+\tau})\). If the measurement noise is independent of \( X_t \), then \((X_t, X_{t+\tau})\) and \((Y_t, Y_{t+\tau})\) are independent random variables and the joint probability density \( \rho^*(x_1, x_2, \tau) \) of their sum \((X_t^*, X_{t+\tau}^*)\) is given by the convolution of \( \rho \) and \( \rho_Y \), where \( \rho_Y(x_1, x_2, \tau) \) is the joint probability density of \((Y_t, Y_{t+\tau})\).

The noise is assumed to be Gaussian and the Gauss function has some special properties with regard to convolution and Fourier transform. This makes it possible to extract the noise parameters from the moments of \( \rho^* \). Furthermore the abovementioned relation between the conditional moments and the unknown functions can be transformed into a relation in Fourier space. This allows polynomial approximations of the drift and diffusion functions to be extracted using purely algebraic relations between quantities that can be calculated directly from a given, noisy time series.

This paper is organized as follows: Section [II] is devoted to the noise-free stochastic process, the definition of its joint probability density and expressions for the moments of this density in terms of a Taylor-\( \text{It}^o \) expansion. Section [III] provides the properties of the measurement noise under consideration and in section [IV] expressions for the moments of a noisy process will be derived. In section [V] these expressions will be used to extract the parameters of the measurement noise and in section [VI] to extract polynomial approximations for drift and diffusion functions. Finally in section [VII] the results are applied to some synthetic time series. The used properties of the Gauss function and further computational details are given in appendices [A] to [C].

II. STOCHASTIC PROCESS

Let \( X_t \) be a stochastic process that can be described by a time-independent \( \text{It}^o \) Langevin equation

\[
dX_t = D^{(1)}(X_t)dt + \sqrt{D^{(2)}(X_t)}dW_t, \tag{1}
\]

where \( D^{(1)} \) and \( D^{(2)} \) are the Kramers-Moyal coefficients of the corresponding Fokker-Planck equation and \( dW \) is the increment of a standard Wiener process \( < dW dW > = dt \).

The above equation cannot be solved analytically in general. To get finite-time step approximations for the temporal evolution of \( X_t \), a Taylor-\( \text{It}^o \) expansion can be applied.

\[
X_{t+\tau} = X_t + h(X_t, \tau) + R \tag{2}
\]

The function \( h \) contains the expansion terms and \( R \) is a remainder that will be dropped when using the approximation. The above equation allows the numerical integration of Eq. (1) but will not be needed in the following. What will be needed are approximations for the first and second moment of
moments which usually are defined as

\[ m_j = \langle X_{t+\tau} X_t \rangle |_{X_t = x} \]

(3a)

\[ m_j = \langle (X_{t+\tau} - X_t)^2 \rangle |_{X_t = x} \]

(3b)

A Taylor-Itô expansion also provides expressions for these moments. For an expansion of order \( k \) the functions \( h_1 \) and \( h_2 \) are polynomials of order \( k \) in \( \tau \) (without a constant part) and the remainders \( R_1 \) and \( R_2 \) are of order \( \tau^{k+1} \).

\[ h_1(x, \tau) = e_1(x) \tau + \ldots + e_k(x) \tau^k \]

(4a)

\[ h_2(x, \tau) = d_1(x) \tau + \ldots + d_k(x) \tau^k \]

(4b)

The polynomial coefficients generally are functions of \( D^{(1)} \) and \( D^{(2)} \) and their derivatives. The first order approximations that will be used later are given by

\[ h_1(x, \tau) = D^{(1)}(x) \tau \]

(5a)

\[ h_2(x, \tau) = D^{(2)}(x) \tau \]

(5b)

A detailed description of the Taylor-Itô expansion and its moments can be found in [26].

The probability density of \( X_t \) and the joint probability density of \( (X_t, X_{t+\tau}) \) are formally given by

\[ \rho(x_1) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \delta(X_t - x_1) dt \]

(6a)

\[ \rho(x_1, x_2, \tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \delta(X_t - x_1) \times \delta(X_{t+\tau} - x_2) dt \]

(6b)

Let the conditioned moments \( m_j \) be defined as

\[ m_j(x_1, \tau) = \int_{-\infty}^\infty (x_2 - x_1)^j \rho(x_1, x_2, \tau) dx_2. \]

(7)

(These moments are not to be confused with the conditional moments which usually are defined as \( m_j/m_0 \).) Inserting Eq. (6b) the moments can be expressed as

\[ m_j(x_1, \tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \delta(X_t - x_1) \times (X_{t+\tau} - x_1)^j dt. \]

(8)

Using

\[ \lim_{T \to \infty} \frac{1}{T} \int_0^T \delta(X_t - x_1) f(X_t) dt = \begin{cases} \int f(X_t) |_{X_t = x_1} > \rho(x_1) \\ \end{cases} \]

(9)

leads to

\[ m_j(x_1, \tau) = \langle (X_{t+\tau} - X_t)^j \rangle |_{X_t = x_1} > \rho(x_1) \]

(10)

Inserting Eqs. (3) and omitting the remainders the final approximations of the moments up to order two are therefore given by

\[ m_0(x_1) = \rho(x_1) \]

(11a)

\[ m_1(x_1, \tau) = h_1(x_1, \tau) m_0(x_1) \]

(11b)

\[ m_2(x_1, \tau) = h_2(x_1, \tau) m_0(x_1) \]

(11c)

These are the equations that establish the relation between the observable quantities \( m_j \) and the unknown coefficients contained in \( h_1 \) and \( h_2 \) within the precision of the chosen Taylor-Itô expansion.

In the absence of noise Eqs. (11) can directly be used to determine \( D^{(1)} \) and \( D^{(2)} \). If e.g. the above mentioned first order scheme was chosen, the equations read

\[ m_1(x_1, \tau) = \tau D^{(1)}(x_1) + O(\tau^2) \]

\[ m_2(x_1, \tau) = \tau D^{(2)}(x_1) + O(\tau^2). \]

### III. MEASUREMENT NOISE

Let \( Y_t \) be exponentially correlated Gaussian noise as produced by an Ornstein-Uhlenbeck process.

\[ dY_t = -aY_t dt + b dW_t \]

(12)

This equation can be solved analytically (see e.g. [27]). The probability densities of \( Y_t \) can be expressed either in terms of \( a \) and \( b \) or, equivalently, in terms of the 'macroscopic' parameters \( T \) (characteristic time scale) and \( \sigma^2 \) (variance) with

\[ T = \frac{1}{a} \quad \text{and} \quad \sigma^2 = \frac{b^2}{2a}. \]

(13)

The unconditioned distribution of \( Y_t \) will be denoted by \( K \) because it will serve as a convolution kernel later. It is given by

\[ K(x) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{1}{2} \left( \frac{x}{\sigma} \right)^2}. \]

(14)

The joint probability density of \( (Y_t, Y_{t+\tau}) \) can then be expressed as

\[ \rho_Y(x_1, x_2, \tau) = K(x_1) \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{1}{2} \left( \frac{x_2 - x_1 - \mu(\tau)\tau}{\sigma} \right)^2}, \]

(15)

where the decay-function

\[ \mu(\tau) = e^{-\frac{\tau}{T}} \]

(16)

and the auxiliary quantity

\[ s^2 = \sigma^2(1 - \mu^2(\tau)) \]

(17)

have been introduced for notational simplicity. The parameter \( T \) will also be called ‘the’ correlation time. This is motivated by

\[ \frac{\langle Y_t Y_{t+\tau} \rangle}{\sigma^2} = e^{-\frac{\tau}{T}}. \]

(18)
IV. NOISY STOCHASTIC PROCESS

Let \( X^*_t = X_t + Y_t \) be the sum of a stochastic signal \( X_t \) and measurement noise \( Y_t \) as introduced in sections III and III respectively. Then, because \( X \) and \( Y \) are independent, the joint probability density \( \rho^* \) of the pairs of noisy variables, \((X^*_t, X^*_{t+\tau})\), is given by the convolution of \( \rho \) and \( \rho_Y \).

\[
\rho^*(x_1, x_2, \tau) = \rho_Y(x_1, x_2, \tau) * \rho(x_1, x_2, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_Y(x_1 - x_1', x_2 - x_2', \tau) \times \rho(x_1', x_2', \tau) dx_1' dx_2'
\]

Instead of the conditioned moments \( m_j \), only their noisy counterparts \( m_j^* \) can be determined.

\[
m_j^*(x_1, \tau) = \int_{-\infty}^{\infty} (x_2 - x_1)^j \rho^*(x_1, x_2, \tau) dx_2 \tag{20}
\]

These can be expressed as follows (see appendix B). The index of the variable \( x_1 \) and the function argument of \( \mu(\tau) \) are omitted for notational simplicity.

\[
m_0^*(x) = K(x) * m_0(x) \tag{21a}
\]

\[
m_1^*(x, \tau) = K(x) * m_1(x, \tau) - (1 - \mu)(xK(x)) * m_0(x) \tag{21b}
\]

\[
m_2^*(x, \tau) = K(x) * m_2(x, \tau) - 2(1 - \mu)(xK(x)) * m_1(x, \tau) + (1 - \mu)^2(x^2K(x)) * m_0(x) + 2(1 - \mu)^2K(x) * m_0(x). \tag{21c}
\]

These relations have already been derived earlier for the case of delta-correlated noise and have been formulated with conditional probabilities \([24]\). The above convolutional notation makes it more obvious how to proceed further. Because \( K \) is a Gaussian function with variance \( \sigma^2 \), the terms \( x^jK(x) \) can be expressed as (see appendix A):

\[
xK(x) = -\sigma^2 \partial_x K(x)
\]

\[
x^2K(x) = (\sigma^2 + \sigma^4 \partial_x^2)K(x)
\]

Because of \((\partial_x f) * g = \partial_x (f * g)\) this leads to (using Eq. (17) and omitting arguments now completely):

\[
m_0^* = K * m_0 \tag{22a}
\]

\[
m_1^* = K * m_1 + (1 - \mu)\sigma^2 \partial_x (K * m_0) \tag{22b}
\]

\[
m_2^* = K * m_2 + 2(1 - \mu)\sigma^2 \partial_x (K * m_1) + (1 - \mu)^2 \sigma^4 \partial_x^2 (K * m_0)
\]

Substituting \( K * m_0 \) and \( K * m_1 \) and using Eqs. (22a) and (22b) yields

\[
m_0^* = K * m_0 \tag{23a}
\]

\[
m_1^* = K * m_1 + (1 - \mu)\sigma^2 \partial_x m_0^* \tag{23b}
\]

\[
m_2^* = K * m_2 + 2(1 - \mu)\sigma^2 \partial_x m_1^* - (1 - \mu)^2 \sigma^4 \partial_x^2 m_0^* + 2(1 - \mu)\sigma^2 m_0^*. \tag{23c}
\]

Using Eqs. (11) to express \( m_1 \) and \( m_2 \) finally gives

\[
m_0^* = K * m_0 \tag{24a}
\]

\[
m_1^* = K * (h_1 m_0) + (1 - \mu)\sigma^2 \partial_x m_0^* \tag{24b}
\]

\[
m_2^* = K * (h_2 m_0) + 2(1 - \mu)\sigma^2 \partial_x m_1^* - (1 - \mu)^2 \sigma^4 \partial_x^2 m_0^* + 2(1 - \mu)\sigma^2 m_0^*. \tag{24c}
\]

Eq. (24b) can be used to extract the parameters of the measurement noise without determining the Kramers-Moyal coefficients \( D^{(1)} \) and \( D^{(2)} \). This will be done in the next section.

V. EXTRACTING MEASUREMENT NOISE PARAMETERS

Multiplying Eq. (24b) by a weight function \( \Psi(x) \) and subsequently performing an integration with respect to \( x \) leads to

\[
A(\tau) = \int_{-\infty}^{\infty} (K(x) * (h_1(x, \tau)m_0(x)) \Psi(x) dx + (1 - \mu(\tau))\sigma^2 B \tag{25}
\]

with

\[
A(\tau) = \int_{-\infty}^{\infty} m_1^*(x, \tau) \Psi(x) dx \tag{26}
\]

and

\[
B = \int_{-\infty}^{\infty} (\partial_x m_1^*(x)) \Psi(x) dx = -\int_{-\infty}^{\infty} m_0^*(x)(\partial_x \Psi(x)) dx \tag{27}
\]

In the last step an integration by parts has been applied, assuming \( \Psi m_0^* \to 0 \) when \( |x| \to \infty \). Using Eq. (14) the remaining integral in Eq. (25) can be expressed as a polynomial in \( \tau \).

\[
\int_{-\infty}^{\infty} \{K(x) * (h_1(x, \tau)m_0(x)) \Psi(x) dx = \sum_{j=1}^{k} \tau^j \int_{-\infty}^{\infty} \{K(x) * (c_j(x)m_0(x)) \} \Psi(x) dx
\]
Finaly, after a division by $B$, Eq. (25) reads
\[
z(\tau) = (1 - \mu(\tau))\sigma^2 + C_1\tau + C_2\tau^2 + \ldots
\] (29)
with
\[
z(\tau) = \frac{A(\tau)}{B} \quad \text{and} \quad C_j = \frac{\tilde{C}_j}{B}.
\] (30)

Once a weight function $\Psi$ has been chosen, $z(\tau)$ can be derived from experimental data. Eq. (29) can be used to fit the unknown parameters $\sigma^2$, $T$ (contained in $\mu = e^{-\tau}$) and $C_j$ on the righthand side. This can be done for example by an iterated least-square fit.

The simplest choice $\Psi = x$ is not the most accurate one, especially in the presence of heavy tails. It puts a high weight on the tails of $m_1^*$, where density is low and fluctuations of experimental data thus high. Choosing $\partial_x \Psi$ to be a rough, piecewise steady approximation of the density $\rho^*$, the temporal evolution of $\tilde{X}_t^*$ itself can be approximated by a piecewise constant function:
\[
\tilde{X}_t^* = x_i \quad (t_i \leq t < t_{i+1})
\] (31)

The density distribution of $(\tilde{X}_t^*, \tilde{X}_{t+\tau_k}^*)$ with $\tau_k = k\Delta t$ can be evaluated using Eq. (6b) and is given by a sum of Dirac-distributions (the coordinates $x_i'$ and $x_i''$ are used to avoid confusion with the values $x_i$ of the given time series).
\[
\tilde{\rho}^*(x_1', x_2', \tau_k) = \frac{1}{N} \sum_{i=1}^N \delta(x_1 - x_1') \delta(x_2 - x_2')
\] (32)

The moments of $\tilde{\rho}^*$ evaluate to
\[
\tilde{m}_j^*(x_1', \tau_k) = \frac{1}{N} \sum_{i=1}^N (x_{i+k} - x_i') \delta(x_1 - x_1')
\] (33)
what results in
\[
\tilde{A}(\tau_k) = \frac{1}{N} \sum_{i=1}^N (x_{i+k} - x_i') \Psi(x_i')
\]
\[
\tilde{B} = -\frac{1}{N} \sum_{i=1}^N \Psi'(x_i)
\] (34)

where $\Psi'$ denotes the derivative with respect to $x$. The values $\tilde{z}(\tau_k)$ derived from a given time series $x_i$ can therefore be expressed as
\[
\tilde{z}(\tau_k) = -\sum_i (x_{i+k} - x_i') \Psi(x_i') / \sum_i \Psi'(x_i')
\] (35)

To be able to successfully perform a fit it must be possible to distinguish between the polynomial $\sum C_j \tau^j$ and the function $1 - e^{-\tau}$ within the given range of increments $\tau$. This means that a good polynomial approximation for $e^{-\tau}$ should require a higher order than the polynomial defined by the $C_j$. The proposed method is therefore limited to measurement noise with a correlation time $T$ considerably smaller than the time scale of the underlying stochastic process. An illustrative example is given in Fig. 1.
coefficients. Let $D^{(1)}$ and $D^{(2)}$ be given by

$$D^{(1)} = \sum_{j=0}^{N_1} a_j x^j$$

(41a)

$$D^{(2)} = \sum_{j=0}^{N_2} b_j x^j.$$  

(41b)

Choosing the first order approximation Eqs. (38) for $h_1$ and $h_2$, the right-hand sides of Eqs. (38) read

$$\gamma_1 = \tau F \left\{ K \ast (D^{(1)} m_0) \right\}$$

$$= \tau \sum_{j=0}^{N_1} a_j F \left( K \ast (x^j m_0) \right)$$

(42a)

$$\gamma_2 = \tau F \left\{ K \ast (D^{(2)} m_0) \right\}.$$  

(42b)

Again because $K$ is a Gauss function with variance $\sigma^2$, the Fourier transform of $K \ast (x^j m_0)$ can be expressed as linear combination of the transforms of $x^j (K \ast m_0)$ (see appendix A). Since $K \ast m_0 = m_0^\tau$ these transforms can be derived from the noisy data. Using the shortcuts

$$\Phi_j = F \left( x^j (K \ast m_0) \right) = F \left( x^j m_0^\tau \right)$$

(43)

$$F_j = F \left( K \ast (x^j m_0) \right)$$

(44)

the relation between $\Phi$ and $F$ is given by

$$F_j = \sum_{k=0}^{j} \binom{j}{k} \varphi_{j-k} \Phi_k$$

(45)

with

$$\varphi_j = i^j \sum_{k=0}^{j} |a_{jk}| \sigma^j k \omega^k$$

(46)

and

$$a_{jk} = -a_{(j-1)(k-1)} + (k+1) a_{(j-1)(k+1)}$$

$$a_{00} = 1$$

$$a_{jk} = 0 \quad (j < 0, k < 0, j < k).$$

(47)

Eqs. (38) now read

$$\hat{m}_1^* - M i \omega \hat{m}_0^* = \tau \sum_{j=0}^{N_1} a_j F_j$$

(48a)

$$\hat{m}_2^* - 2M (\hat{m}_0^* + i \omega \hat{m}_1^*) - M^2 \omega^2 \hat{m}_0^* = \tau \sum_{j=0}^{N_2} b_j F_j.$$  

(48b)

If the noise parameters have been extracted according to section VII then $M = (1-\mu)\sigma^2$ and $F_j$ are known and the coefficients $a_j$ and $b_j$ can be extracted by a least square fit. It is also possible to use Eqs. (48) to fit noise parameters and polynomial coefficients simultaneously by an iterated least square fit. But this has turned out to be less accurate.

Adding an increment dependency to the polynomial coefficients however allows some of the approximation errors to be absorbed by the additional parameters and has shown to increase accuracy.

$$a_j \rightarrow a_j + \tau a_j^1$$

(49a)

$$b_j \rightarrow b_j + \tau b_j^1$$

(49b)

With this extension the final equations for parameter fitting are given by

$$\hat{m}_1^* - M i \omega \hat{m}_0^* = \tau \sum_{j=0}^{N_1} (a_j + \tau a_j^1) F_j$$

(50a)

$$\hat{m}_2^* - 2M (\hat{m}_0^* + i \omega \hat{m}_1^*) - M^2 \omega^2 \hat{m}_0^* = \tau \sum_{j=0}^{N_2} (b_j + \tau b_j^1) F_j.$$  

(50b)

To be able to actually perform a fit, a finite number of values for $\omega$ and $\tau$ has to be chosen. It is sufficient to restrict the choice of $\omega$ to positive values because the above equations are symmetric (conjugate complex) in $\omega$. The following heuristic approach has shown to work quite well:

- Define an upper bound $\omega^*$ by

$$\int_0^{\omega^*} |\hat{m}_1|^2 d\omega = 0.99 \int_0^{\infty} |\hat{m}_1|^2 d\omega.$$  

(51)

- Chose $N_\omega$ values $\omega_j$ equally distributed in $[0, \omega^*]$. The actual number of values is of minor importance but it should be high enough to properly sample $\hat{m}_1^\tau$.

VII. APPLICATION TO NUMERICAL DATA

In order to test the accuracy of the proposed method, synthetic stochastic signals have been generated by numerical integration. As a first test-case an Ornstein-Uhlenbeck process has been chosen. Its drift and diffusion functions have been defined as

$$D^{(1)}(x) = -x$$  

(52a)

$$D^{(2)}(x) = 2.$$  

(52b)

The generated time series will be referred to as data set A. It consists of $10^6$ data points at time increments $\Delta t = 10^{-2}$. The integration has been performed with a timestep size of...
\(\sigma\) is the standard deviation of the superimposed noise. Drift and diffusion functions are given by \(D^{(1)} = -x\) and \(D^{(2)} = 2\).

The noise parameters have been extracted according to section \(V\). The maximum increment has been \(\tau = 60\). The absolute error \(\Delta\sigma\) in the estimation of the strength of the measurement noise is presented in Fig. 3 as a function of the true measurement noise \(\sigma\). The results are accurate within the achievable precision of the finite time series. This means that the measurement noise alone, without any stochastic process, cannot be estimated much more accurate: For a sample of \(N\) data points of Gaussian, delta-correlated noise the estimation \(S\) of the true standard deviation, \(\sigma\), obeys a variance of approximately \(\sigma^2 / 2N\) (see appendix \(C\)). A lower bound for the standard deviation of \(\Delta\sigma\) is therefore given by \(\sigma \sqrt{2 / N}\). These limits are indicated in Fig. 3 as dashed lines.

Next the drift and diffusion functions have been fitted choosing the polynomial ansatz \(D^{(1)} = a_0 + a_1 x\) and \(D^{(2)} = b_0\). As maximum increment a value of \(\tau = 25\) (again in units of increments of data points) has been chosen. The results are shown in Figs. 4 and 5 and are in good agreement with the true values \((a_0 = 0, a_1 = -1, b_0 = 2)\).

After this basic test-case a more general case has been chosen. It includes multiplicative noise, heavy tails and a density distribution which is not symmetric. Also the superposition of
FIG. 6: Probability density function of data set B. \( \sigma \) is the standard deviation of the superimposed noise. Drift and diffusion functions are given by \( D^{(1)} = 1 - x \) and \( D^{(2)} = 2 - 2x + 2x^2 \).

noise with finite correlation time has been investigated in this setup. Drift and diffusion functions have been defined as

\[
D^{(1)}(x) = 1 - x \\
D^{(2)}(x) = 2 - 2x + 2x^2.
\]  

The generated time series will be referred to as data set B. Again it consists of \( 10^6 \) data points at time increments \( \Delta t = 10^{-2} \). The integration has been performed with a timestep size of \( 2.0 \cdot 10^{-5} \) using the Euler scheme. The standard deviation of the data points is approximately 1.83 and the timescale of the process has been estimated via the autocorrelation function to be about \( \tau = 100 \) in units of increments of data points.

In a first step again delta-correlated measurement noise with standard deviation in the range \( \sigma = 0.0 \ldots 2.0 \) has been superimposed to the process. Some of the density histograms of the resulting signals are shown in Fig. 6.

As maximum increment for noise-fitting \( \tau = 60 \) and for coefficient-fitting \( \tau = 25 \) has been chosen. The polynomial ansatz for drift and diffusion functions is given by \( D^{(1)} = a_0 + a_1x \) and \( D^{(2)} = b_0 + b_1x + b_2x^2 \). The results are summarized in Fig. 7.

In a second step measurement noise with a correlation time \( T = 2.0 \) (in units of data point increments) has been superimposed. The results are given in Fig. 7 and obey higher fluctuations as in the delta-correlated case. The correlation time of the measurement noise is estimated nicely except for small amplitudes of \( \sigma \). To check the assumption that this is caused by the finite-sample fluctuations of \( z(\tau) \) a larger data set, B1, has been generated. It consists of \( 10^7 \) data points and its analysis should show less variability in all estimated quantities. This is indeed the case. The results are shown in Figs. 9 and 10.

FIG. 7: Results for data set B. Estimations of the polynomial coefficients \( a_j \) and \( b_j \) and absolute error \( \Delta \sigma \) in the estimation of the measurement noise. The superimposed noise is delta-correlated.

VIII. CONCLUSIONS

A new procedure has been described to analyze stochastic time series that are superimposed by strong measurement noise. The algorithm is able to cope with exponentially correlated noise and accurately extracts strength and correlation time of the measurement noise as well as the parameters defining the drift and diffusion functions of the underlying stochastic process. This has been shown by the analysis of synthetically generated time series. The chosen stochastic processes include the cases of multiplicative noise and heavy tailed density distribution.

The computational costs of the algorithm are quite low. It takes less than a minute to analyze a data set of size \( 10^7 \) data points on a usual desktop PC.

A first order Taylor-Itô expansion for the moments of the finite time differences of the process variable is used in the current implementation. It is straightforward to extend the algorithm to also take higher order terms into account. This should extend the range of time increments that can be used for the analysis and thus should increase the accuracy of the results. This has to be done in the future.

Another future task is the extension to higher dimensional processes. Up to now it is an open question if a general multidimensional Ornstein-Uhlenbeck process as a model for the measurement noise can be used or if some restrictions need to be imposed on it.
FIG. 8: Results for data set B. Estimations of the polynomial coefficients $a_j$ and $b_j$, correlation time $T$ and absolute error $\Delta \sigma$ in the estimation of the measurement noise. The superimposed noise has a correlation time of $T = 2.0$.

IX. ACKNOWLEDGMENTS

The author especially wants to thank Joachim Peinke, Rudolf Friedrich, David Kleinhans, Pedro G. Lind and Maria Haase for many useful discussions.

Appendix A: Gauss functions

1. Definitions and basic properties

Let Fourier transform and convolution be defined as below. For notational simplicity the `hat` syntax will be used to denote the transform of single functions. For more complex expressions the functional form $F(...)$ will usually be the better choice.

\[
F(f(x)) = \hat{f}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega x} f(x)dx \tag{A1}
\]

\[
f(x) * g(x) = \int_{-\infty}^{+\infty} f(x-x')g(x')dx \tag{A2}
\]

Above definitions imply the following properties:

\[
F(\partial_x f(x)) = (i\omega)^j \hat{f}(\omega) \tag{A3a}
\]

\[
F(x^j f(x)) = (i\partial_x)^j \hat{f}(\omega) \tag{A3b}
\]

\[
F(f(\frac{x}{a})) = |a| \hat{f}(a\omega) \tag{A3c}
\]

\[
F(f(x) * g(x)) = \hat{f}(\omega)\hat{g}(\omega) \tag{A3d}
\]

\[
\partial_x^j (f(x) * g(x)) = (\partial_x^j f(x)) * g(x)
= f(x) * (\partial_x^j g(x)) \tag{A3e}
\]
FIG. 10: Results for data set B1. Estimations of the polynomial coefficients \( a_j \) and \( b_j \), correlation time \( T \) and absolute error \( \Delta \sigma \) in the estimation of the measurement noise. The superimposed noise has a correlation time of \( T = 2.0 \).

2. Derivatives and polynomial products of Gauss functions

Let the Gauss function \( G \) be defined as below.

\[
G(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{A4a}
\]

\[
\hat{G}(\omega) = e^{-\frac{1}{2}\omega^2} = \sqrt{2\pi}G(\omega) \tag{A4b}
\]

All derivatives \( \partial_x^j G(x) \) have the form \( P_j(x)G(x) \), where \( P_j \) is a polynomial of order \( j \) in \( x \).

\[
\partial_x G(x) = \{-x\} G(x)
\]

\[
\partial_x^2 G(x) = \{-1 + x^2\} G(x)
\]

\[
\vdots
\]

\[
\partial_x^j G(x) = \left\{ \sum_{k=0}^{j} a_{jk} x^k \right\} G(x) \tag{A5}
\]

The values of the polynomial coefficients can be determined by a recursion formula:

\[
a_{jk} = -a_{(j-1)(k-1)} + (k+1)a_{(j-1)(k+1)} \quad \text{for } j < k
\]

\[
a_{00} = 1
\]

\[
a_{jk} = 0 \quad \text{for } j < 0, k < 0, j < k
\]

(A6)

It can also be shown, that the coefficients are zero, if \( j + k \) or \( j - k \) is odd. Otherwise, within the restrictions of formula (A6), \( a_{jk} \) is non-zero and its sign is given by \( i^{|j+k|} \).

It will be useful later, to express the derivatives of \( G \) in a similar way. The proceeding is the same as for \( G \) and leads to the recursion formula (A8).

\[
\partial_x^j \frac{1}{G(x)} = \left\{ \sum_{k=0}^{j} \tilde{a}_{jk} x^k \right\} \frac{1}{G(x)} \tag{A7}
\]

\[
\tilde{a}_{jk} = \tilde{a}_{(j-1)(k-1)} + (k+1)\tilde{a}_{(j-1)(k+1)}
\]

\[
\tilde{a}_{00} = 1
\]

\[
\tilde{a}_{jk} = 0 \quad \text{for } j < 0, k < 0, j < k
\]

(A8)

Having a closer look at the coefficients, it turns out, that \( \tilde{a}_{jk} \) and \( a_{jk} \) only differ in sign.

\[
\tilde{a}_{jk} = |a_{jk}| \tag{A9}
\]

To get explicit expressions for \( x^j G(x) \) in terms of derivatives of \( G \), a Fourier transformation is applied to Eq. (A5).

\[
(i\omega)^j \hat{G}(\omega) = \left\{ \sum_{k=0}^{j} a_{jk} (i\omega)^k \right\} \hat{G}(\omega). \tag{A10}
\]

Using Eq. (A4b) and substituting \( \omega \) by \( x \), finally yields:

\[
x^j G(x) = \left\{ \sum_{k=0}^{j} x^{k-j}a_{jk} \partial_x^k \right\} G(x). \tag{A11}
\]

3. Scaled Gauss functions

This can also be expressed for scaled Gauss functions having standard deviation \( \sigma \). The Fourier transform of such a function is given by

\[
F(G(\frac{x}{\sigma})) = \sigma \hat{G}(\sigma \omega). \tag{A12}
\]
The relations between derivatives and polynomial products then read:

\[
\partial_x^j G\left(\frac{x}{\sigma}\right) = \left\{ \sum_{k=0}^{j} a_{jk} \sigma^{-j-k} x^k \right\} G\left(\frac{x}{\sigma}\right) \quad (A13)
\]

\[
\partial_x^j \frac{1}{G\left(\frac{x}{\sigma}\right)} = \left\{ \sum_{k=0}^{j} |a_{jk}| \sigma^{-j-k} x^k \right\} \frac{1}{G\left(\frac{x}{\sigma}\right)} \quad (A14)
\]

\[
x^j G\left(\frac{x}{\sigma}\right) = \left\{ \sum_{k=0}^{j} i^{k-j} a_{jk} \sigma^{j+k} \partial_x^k \right\} G\left(\frac{x}{\sigma}\right) \quad (A15)
\]

4. Convolutions

The above properties can be used to express convolutions of the form \((x^j G) \ast f\) and \(G \ast (x^j f)\) in terms of the ‘raw’ convolution \(G \ast f\) and its polynomial products \(x^j (G \ast f)\). Using Eqs. (A3e) and (A15) immediately leads to Eq. (A16).

\[
(x^j G\left(\frac{x}{\sigma}\right)) \ast f(x) = \left\{ \sum_{k=0}^{j} i^{k-j} a_{jk} \sigma^{j+k} \partial_x^k \right\} G\left(\frac{x}{\sigma}\right) \ast f(x)
\]

\[
= \left\{ \sum_{k=0}^{j} i^{k-j} a_{jk} \sigma^{j+k} \partial_x^k \right\} \left( G\left(\frac{x}{\sigma}\right) \ast f(x) \right)
\]

(A16)

To get an expression for \(G \ast (x^j f)\) it is useful to look at the term \(G \cdot \partial^i f\) first.

\[
G\left(\frac{x}{\sigma}\right) \partial_x^i f(x) = G\left(\frac{x}{\sigma}\right) \partial_x^i \left\{ \frac{1}{G\left(\frac{x}{\sigma}\right)} G\left(\frac{x}{\sigma}\right) f(x) \right\}
\]

\[
= G\left(\frac{x}{\sigma}\right) \sum_{k=0}^{j} \left( \begin{array}{c} j \\ k \end{array} \right) \partial_x^{j-k} \frac{1}{G\left(\frac{x}{\sigma}\right)} \partial_x^k \left( G\left(\frac{x}{\sigma}\right) f(x) \right)
\]

\[
= G\left(\frac{x}{\sigma}\right) \sum_{k=0}^{j} \left( \begin{array}{c} j \\ k \end{array} \right) \sum_{l=0}^{j-k} |a_{j-k,l}| \sigma^{-j-k-l} x^l \frac{1}{G\left(\frac{x}{\sigma}\right)} \partial_x^k \left( G\left(\frac{x}{\sigma}\right) f(x) \right)
\]

\[
= \sum_{k=0}^{j} \left( \begin{array}{c} j \\ k \end{array} \right) \sum_{l=0}^{j-k} |a_{j-k,l}| \sigma^{-j-k-l} x^l \partial_x^k \left( G\left(\frac{x}{\sigma}\right) f(x) \right)
\]

This allows the transform of \(G \ast (x^j f)\) to be written as follows:

\[
\mathbf{F} \left( G\left(\frac{x}{\sigma}\right) \ast (x^j f(x)) \right) = \sigma \hat{G}(\sigma \omega) (i \partial_\omega)^j \hat{f}(\omega) = i^j \sigma \sqrt{2\pi} \cdot G(\sigma \omega) \partial_x^j \hat{f}(\omega)
\]

\[
= i^j \sigma \sqrt{2\pi} \sum_{k=0}^{j} \left( \begin{array}{c} j \\ k \end{array} \right) \sum_{l=0}^{j-k} |a_{j-k,l}| \sigma^{-j-k-l} \omega^l \partial_x^k \left( G(\sigma \omega) \hat{f}(\omega) \right)
\]

\[
= \sum_{k=0}^{j} \left( \begin{array}{c} j \\ k \end{array} \right) \sum_{l=0}^{j-k} i^{j-k} |a_{j-k,l}| \sigma^{j-k-l} \omega^l (i \partial_\omega)^k \left( \sigma \hat{G}(\sigma \omega) \hat{f}(\omega) \right)
\]

\[
= \sum_{k=0}^{j} \left( \begin{array}{c} j \\ k \end{array} \right) \sum_{l=0}^{j-k} i^{j-k} |a_{j-k,l}| \sigma^{j-k-l} \omega^l \mathbf{F} \left( x^k (G\left(\frac{x}{\sigma}\right) \ast f(x)) \right)
\]

(A17)

Omitting arguments and introducing the auxiliary quantity \(\varphi_j\) this can be expressed as

\[
\mathbf{F} \left( G \ast (x^j f) \right) = \sum_{k=0}^{j} \left( \begin{array}{c} j \\ k \end{array} \right) \varphi_{j-k} \mathbf{F} \left( x^k (G \ast f) \right), \quad \varphi_j = i^j \sum_{k=0}^{j} |a_{jk}| \sigma^{j+k} \omega^k.
\]

(A18)
Appendix B: Conditioned moments \( m^*_j \)

After inserting Eqs. (15) and (19) and interchanging the order of integration Eq. (20) reads

\[
m^*_j(x_1, \tau) = \int_{-\infty}^{\infty} K(x_1 - x'_1) \int_{-\infty}^{\infty} \rho(x'_1, x'_2, \tau) I_j dx'_2 dx'_1,
\]

where

\[
I_j = \int_{-\infty}^{\infty} (x_2 - x_1)^2 \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{1}{2} \left( \frac{x_2 - \mu(\tau)(x_1 - x'_1) + s^2}{s} \right)^2} dx_2.
\]

Substituting

\[
z = \frac{x_2 - \mu(\tau)(x_1 - x'_1) - x'^2}{s}
\]
gives

\[
I_j = \int_{-\infty}^{\infty} (sz + (x'_2 - x'_1) - (1 - \mu(\tau))(x_1 - x'_1))^2 \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{1}{2} z^2} dz
\]

and therefore

\[
I_0 = 1
\]

\[
I_1 = (x'_2 - x'_1) - (1 - \mu(\tau))(x_1 - x'_1)
\]

\[
I_2 = s^2 + (x'_2 - x'_1)^2 - 2(1 - \mu(\tau))(x_1 - x'_1)(x'_2 - x'_1) + (1 - \mu(\tau))^2(x_1 - x'_1)^2
\]

Now

\[
I_j = \int_{-\infty}^{\infty} \rho(x'_1, x'_2, \tau) I_j dx'_2
\]

can be evaluated using the definition of \( m_j \) (Eq. (B3)).

\[
I_0 = m_0(x'_1)
\]

\[
I'_1 = m_1(x'_1, \tau) - (1 - \mu(\tau))(x_1 - x'_1)m_0(x'_1)
\]

\[
I'_2 = m_2(x'_1, \tau) - 2(1 - \mu(\tau))(x_1 - x'_1)m_1(x'_1, \tau) + \{(1 - \mu(\tau))^2(x_1 - x'_1)^2 + s^2\} m_0(x'_1)
\]

Using \( \int_{-\infty}^{\infty} (x - x')^k f(x - x')g(x')dx' = (x^k f(x)) * g(x) \) the noisy conditioned moments read

\[
m^*_0(x_1) = K(x_1) * m_0(x_1)
\]

\[
m^*_1(x_1, \tau) = K(x_1) * m_1(x_1, \tau) - (1 - \mu(\tau))(x_1 K(x_1)) * m_0(x_1)
\]

\[
m^*_2(x_1, \tau) = K(x_1) * m_2(x_1, \tau) - 2(1 - \mu(\tau))(x_1 K(x_1)) * m_1(x_1, \tau)
\]

\[
+ (1 - \mu(\tau))^2(x_1^2 K(x_1)) * m_0(x_1) + s^2 K(x_1) * m_0(x_1).
\]

Appendix C: Error in the estimation of \( \sigma \)

Let \( \xi \) be a Gaussian random variable with mean zero and standard deviation \( \sigma \). Further let \( \xi_1, \ldots, \xi_N \) be a sample of \( N \) independent realizations of \( \xi \). Then the expectation values of \( \xi, \xi \xi_j \) and \( \xi_j \xi_j \) are given by

\[
< \xi_i > = 0
\]

\[
< \xi_i \xi_j > = \sigma^2 \delta_{ij}
\]

\[
< \xi_i^2 \xi_j^2 > = \sigma^4 + 2\sigma^4 \delta_{ij}
\]

The variance of \( \xi \) can be estimated from the sample as

\[
V = \frac{1}{N} \sum_{i=1}^{N} \xi_i^2
\]

with an expectation value of \( \sigma^2 \) due to Eq. (C2).

\[
< V > = \sigma^2
\]

Therefore \( V \) can be written as

\[
V = \sigma^2 + \Delta V \quad \text{with} \quad < \Delta V > = 0.
\]

The variance of \( \Delta V \) can be evaluated using Eq. (C3).

\[
< (\Delta V)^2 > = \frac{1}{N} \sum_{i=1}^{N} \xi_i^2 - \sigma^2)^2
\]

\[
= \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \xi_i^2 \xi_j^2 - \sigma^4
\]
\[
\frac{1}{N^2}(N^2 \sigma^4 + 2N \sigma^4) - \sigma^4 = \frac{1}{2} \sigma^4
\] (C7)

Up to now only an estimation of the variance of \(\xi\) has been made. The estimation \(S\) of its standard deviation is given by

\[ S = \sqrt{\frac{1}{N^2}(N^2 \sigma^4 + 2N \sigma^4) - \sigma^4} = \frac{1}{N} \sigma^4 \] (C8)

which, for large \(N\), can be approximated as

\[ S \approx \sigma \cdot (1 + \frac{1}{2} \frac{\Delta V}{\sigma^2}) = \sigma + \frac{\Delta V}{2 \sigma}. \] (C9)

The standard deviation of the error in \(S\) is therefore given by

\[ \Delta \sigma = \sqrt{\frac{1}{N^2}(N^2 \sigma^4 + 2N \sigma^4) - \sigma^4} = \frac{\sigma}{\sqrt{2N}} \] (C10)

[1] H. Kantz and T. Schreiber, Nonlinear Time Series Analysis, (Cambridge University Press, Cambridge, England, 1997)
[2] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring, Rev. Mod. Phys. 65, 1313 (1993)
[3] R. Friedrich, J. Peinke, and M.R.R. Tabar, Complexity in the view of stochastic processes in Springer Encyclopedia of Complexity and Systems Science (Springer, Berlin, 2008)
[4] R. Friedrich and J. Peinke, Phys. Rev. Lett. 78, 863 (1997)
[5] G. Ryskin, Phys. Rev. E 56, 5123 (1997)
[6] S. Siegert, R. Friedrich, and J. Peinke, Phys. Lett. A 243, 275 (1998)
[7] R. Friedrich et al., Phys. Lett. A 271, 217 (2000)
[8] J. Gradisek, S. Siegert, R. Friedrich, and I. Grabec, Phys. Rev. E 62, 3146 (2000)
[9] R. Friedrich, J. Peinke, and Ch. Renner, Phys. Rev. Lett. 84, 5224 (2000)
[10] S. Kriso, J. Peinke, R. Friedrich, and P. Wagner, Phys. Lett. A 299, 287 (2002)
[11] M. Siefert, A. Kittel, R. Friedrich, and J. Peinke, Europhys. Lett. 61, 466 (2003)
[12] M. Siefert and J. Peinke, Int. J. Bifurcation Chaos Appl. Sci. Eng. 14, 2005 (2004)
[13] T. Kuusela, Phys. Rev. E 69, 031916 (2004)
[14] C. Collette and M. Ausloos, Int. J. Mod. Phys. C 15, 1353 (2004)
[15] P.G. Lind, A. Mora, J.A.C. Gallas and M. Haase, Phys. Rev. E 72, 056706 (2005)
[16] A.P. Nawroth, J. Peinke, D. Kleinhans, and R. Friedrich, Phys. Rev. E 76, 056102 (2007)
[17] J. Prusseit and K. Lehnert, Phys. Rev. E 77, 041914 (2008)
[18] D. Lamoureux, and K. Lehnert, Phys. Lett. A 373, 3507 (2009)
[19] P. Sura and J. Barsugli, Phys. Lett. A 305, 304 (2002)
[20] D. Kleinhans, R. Friedrich, A. Nawroth, and J. Peinke, Phys. Lett. A 346, 42 (2005)
[21] J. Gottschall and J. Peinke, Drift and diffusion estimates for disturbed Langevin-type processes New J. Phys. 10, 20 (2008)
[22] E. J. Kostelich and T. Schreiber, Phys. Rev. E 48, 1752 (1993)
[23] J. P. M. Heald and J. Stark, Phys. Rev. Lett. 84, 2366 (2000)
[24] F. Boettcher, J. Peinke, D. Kleinhans, R. Friedrich, P.G. Lind, and M. Haase, Phys. Rev. Lett. 97, 090603 (2006)
[25] P.G. Lind, M. Haase, F. Boettcher, J. Peinke, D. Kleinhans, R. Friedrich, Phys. Rev. E 81, 041125 (2010)
[26] P. E. Kloeden and E. Platen, Numerical Solution of Stochastic Differential Equations (Springer, New York, 1999)
[27] H. Risken, The Fokker-Planck Equation (Springer, New York, 1989)