Detecting physical laws from data of stochastic dynamical systems perturbed by non-Gaussian α-stable Lévy noise

Linghongzhi Lu(陆凌弘志)[1], Yang Li(李扬)[2], and Xianbin Liu(刘先斌)[1,†]

[1] *State Key Laboratory of Mechanics and Control for Mechanical Structures, College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing* 210016, *China*

[2] *School of Automation, Nanjing University of Science and Technology, Nanjing* 210094, *China*

(Received 9 October 2022; revised manuscript received 17 November 2022; accepted manuscript online 2 December 2022)

Massive data from observations, experiments and simulations of dynamical models in scientific and engineering fields make it desirable for data-driven methods to extract basic laws of these models. We present a novel method to identify such high dimensional stochastic dynamical systems that are perturbed by a non-Gaussian α-stable Lévy noise. More explicitly, firstly a machine learning framework to solve the sparse regression problem is established to grasp the drift terms through one of nonlocal Kramers–Moyal formulas. Then the jump measure and intensity of the noise are disposed by the relationship with statistical characteristics of the process. Three examples are then given to demonstrate the feasibility. This approach proposes an effective way to understand the complex phenomena of systems under non-Gaussian fluctuations and illuminates some insights into the exploration for further typical dynamical indicators such as the maximum likelihood transition path or mean exit time of these stochastic systems.

Keywords: data-driven modelling, noise-induced transitions, Lévy noise, Kramers–Moyal formuas

PACS: 05.10.–a, 05.10.Gg, 05.40.–a **DOI:** 10.1088/1674-1056/aca7ee

1. Introduction

A variety of practical phenomena are modeled by complex systems[1–3] in the form of ordinary or partial differential equations plus the stochastic perturbations in climate science, neuroscience, and ecology, etc. These models oftentimes remain unknown in governing laws but only big observational or experimental data. In decades, the developments in machine learning[4] and data science[5] have made the investigation of these systems through massive noisy data possible. Williams *et al.*[6] identified several statistical characteristics (e.g., the leading eigenvalues) in either deterministic or stochastic systems with the employment of the dynamic mode decomposition (DMD).[7] Boninsegna *et al.*[8] extracted the laws of certain stochastic dynamical systems with further research on extended sparse identification of nonlinear dynamics (SINDy).[9–11] Other data-driven approaches to extract governing equations involve neural fields[12,13] and Bayesian inference,[14,15] etc.

Almost all methods mentioned above have been applied to dispose deterministic or Gaussian diffusion process. Typically, one assumes that the overall force acting on a particle is the Gaussian thermal noise, which is usually applied with an exponentially falling probability density tail. Lévy motion, an appropriate model for non-Gaussian fluctuations, is widely used in climate change,[16] gene regulation,[17] and ecology[18] to mimic action of external random forcing. Moreover, the α-stable Lévy noise is associated with systems that are out of thermal equilibrium with a polynomial "heavy" tail leading to "jump" events, and it can occur when conditions leading to the standard fluctuation-dissipation theorem are violated.[19] In biology, many models show the generation of Lévy motions by biological organisms.[20–22] Li and Duan[23,24] managed to distill key variables in stochastic systems induced by Lévy noise through a novel data-driven method. The relevant exit problem of this system has also been discussed.[25]

Much devotion has been made throughout recent decades for statistics and statistical physics.[26–28] This two fields are relevant. The key factor is the probability theory, which helps to construct many interdisciplinary methods. For example, machine learning (ML) as a statistical tool has really benefitted the study of statistical physics with its own advantages in particle physics,[29] cosmology,[30] chemistry, and materials,[31] etc. ML concentrates on prediction by using general-purpose learning algorithms to find patterns frequently in rich and unwieldy data, and makes minimal assumptions about the data-generating systems. It can be effective even when the data are gathered without a carefully controlled experimental design and in the presence of complicated nonlinear interactions.

In this paper, we devise a method to investigate the laws of long-term dynamics of a class of complex systems, which is perturbed by α-stable Lévy noise, from the easily accessible short-term data. The nonlocal Kramers–Moyal formulas is applied to extract the drift terms in the corresponding stochastic differential equations. The parameters of the Lévy noise characterizing its intensity and stability can be approximated

†Corresponding author. E-mail: xbliu@nuaa.edu.cn

© 2023 Chinese Physical Society and IOP Publishing Ltd

by the first two moments of the amplitude of the sample paths' increments. Three examples are shown to discuss details.

2. Theory and method

Consider an n-dimensional system $x(t)$ restricted by the stochastic differential equation

$$\mathrm{d}x(t) = b(x)\,\mathrm{d}t + \sigma\,\mathrm{d}L_t, \quad x(0) = z, \tag{1}$$

where $b(x)$ is the drift vector of dimension n, L_t denotes the α-stable Lévy noise with components independent in every direction, the diagonal matrix σ represents the noise intensity, and vector α is the stability parameter. The jump measure ν of L_t is subjected to $\int_{\mathbb{R}^n\setminus\{0\}}(|y|^2\wedge 1)\nu(\mathrm{d}y) < \infty$. This motion distinguishes itself from the Brownian motion by the last term σdL_t, which characterizes a pure jump process involved with certain jump discountinuities in its sample paths. The jump is controlled by a measure ν_α satisfying $\int_{\mathbb{R}^n\setminus\{0\}}(|y|^2\wedge 1)\nu(\mathrm{d}y) < \infty$. In this paper, we concentrate on the small jump $1 < \alpha < 2$. For more details about the Lévy noise, one can see Appendix A.

2.1. Identification of the drift

For a drift term b and the sample path x of the system stated above, we have the following relationship from the nonlocal Kramers–Moyal formulas[24]

$$\begin{aligned}&\lim_{t\to 0}\ t^{-1}\mathbb{P}\{x(t)-z\in\Gamma|x(0)=z\}\cdot\mathbb{E}[(x_i(t)-z_i)|x(0)=z;\\ &x(t)-z\in\Gamma] = b_i(z),\end{aligned} \tag{2}$$

where $i = 1,2,\ldots,n$ and Γ is an n-dimensional cube $[-\delta,\delta]^n$. The formula links the intrinsic laws of the system with the statistical characteristics. The limit of the ratio of the probability of process starting at z and falling into Γ with the corresponding conditional expectation at certain direction i, to the time t, can be represented as the ith component of the drift b at initial location z. Thus, it provides the cornerstone for numerical methods to access the drift by short-term integration.

Then we propose a learning framework: In a small-step integration of length h for the process $x(t)$, we discretize the initial point as vector $Z = [z_1, z_2, \ldots, z_M]$ uniformly distributed in an interested domain with totally M starting points and its image $X = [x_1, x_2, \ldots, x_M]$ at time h. Let $\hat{M}$ denote the number of initial points in Z which falls into Γ after integration, and $\hat{Z}, \hat{X}$ be the corresponding sets of initial and terminal points in Z and X. Let $\Theta = (f_1, \ldots, f_K)$ be a dictionary of basis functions on $\mathbb{R}^n$. Assume that the system has ergodic property and $b_i(x) \approx \sum_{k=1}^K c_{i,k} f_k(x)$. Furthermore, take

$$\Theta(\hat{Z}) = \begin{bmatrix} f_1(\hat{z}_1) & \cdots & f_K(\hat{z}_1) \\ \vdots & \ddots & \vdots \\ f_1(\hat{z}_{\hat{M}}) & \cdots & f_K(\hat{z}_{\hat{M}}) \end{bmatrix}$$

and

$$E_i = [\hat{x}_{i,1} - \hat{z}_{i,1}, \ldots, \hat{x}_{i,\hat{M}} - \hat{z}_{i,\hat{M}}].$$

According to Eq. (2), we have

$$\Theta(\hat{Z})c_i = \hat{M}M^{-1}h^{-1}E_i^{\mathrm{T}}. \tag{3}$$

This regression problem needs to be tackled delicately. That is to say, c_i must be targeted exactly not only for its enough accuracy but also for sparsity. Further simplified Eq. (3) as $\mathbb{X}c = \mathbb{Y}$, it comes to the iterative sparsity enforcement:

(i) Solve the standard least square problem Eq. (3) as $\tilde{c} = (\mathbb{X}^{\mathrm{T}}\mathbb{X})^{-1}(\mathbb{X}^{\mathrm{T}}\mathbb{Y})$ for a preliminary non-sparse solution $\tilde{c}$.

(ii) Set the coefficient of $\tilde{c}$ with the smallest absolute value, say $\tilde{c}_i$, to zero.

(iii) Delete the i-th column of $\mathbb{X}$ and $\tilde{c}$ to $\mathbb{X}^*$ and $\tilde{c}^*$ and reach the one-dimension less least square problem $\mathbb{Y} = \mathbb{X}^*\tilde{c}^*$.

(iv) Iteratively repeated (i)–(iii) until cross validation indicates an optimal sparsity level of $\tilde{c}$.

The cross validation is set to balance accuracy and predictivity in the model. For instance, choose a moderate predefined threshold, its magnitude is usually about 0.1%–10% of the largest coefficient of c_i. After each sparsification step, we set the coefficients below the threshold as zero and delete the corresponding basis functions. Then the regression problem is carried out again on the remaining coefficients. The procedure is iterated until no coefficient below the threshold exists.

Remark that the choice of the dictionary does affect the performance of this algorithm, usually the polynomial basis functions are enough to fit polynomial drift or others easy to Taylor expand. Due to the uncertainty of the system's govern laws in preliminary, for complex drifts a set of basis functions abundant in forms are necessary. Greedy search[8] is an ideal but consuming solution. Sometimes we use a set of empirical functions because this step is not the main point.

2.2. Identification of the Lévy motion

Because h is a small quantity, Eq. (1) can be rewritten as

$$x(t) - z \approx b(z)t + \sigma L_t. \tag{4}$$

For the one-dimensional case, due to the characteristic function of L_t,

$$\mathbb{E}\exp(\mathrm{i}\langle u, L_t\rangle) = \exp(-t|u|^\alpha),$$

it arrives at

$$t^{-1/\alpha}L_t \overset{d}{=} L_1.$$

Substituting it back to Eq. (4) yields

$$\sigma^{-1}t^{-1/\alpha}(x(t)-z) \approx \sigma^{-1}b(z)t^{1-1/\alpha} + L_1.$$

When $\alpha \in (1,2)$, we have

$$\lim_{t\to 0}\sigma^{-1}t^{-1/\alpha}(x(t)-z) \overset{d}{=} L_1, \tag{5}$$

where $\stackrel{d}{=}$ denotes the same distribution.

Now we introduce a lemma.[32] Let $A \sim S_{\alpha/2}\left(2\gamma^2(\cos(\pi\alpha/4))^{2/\alpha},1,0\right)$ be a stable random variable and $G \sim N(0,I)$ be an standard normal random variable. A and G are independent of each other. Then the random vector $X = A^{1/2}G$ is rotationally symmetric with the characteristic function

$$\mathbb{E}\,\exp(i\langle u,X\rangle) = \exp\{-\gamma^{\alpha}|u|^{\alpha}\}.$$

Denote the amplitude of x as R

$$R = |X| = \sqrt{X_1^2+\cdots+X_n^2}.$$

Then

$$R^2 \stackrel{d}{=} A(G_1^2+\cdots+G_n^2).$$

As a result, the fractional moments of R is expressed as

$$\begin{aligned}\mathbb{E}[R^p] &= \mathbb{E}\left[A^{p/2}(G_1^2+\cdots+G_n^2)^{p/2}\right]\\ &= \mathbb{E}[A^{p/2}]\mathbb{E}\left[(G_1^2+\cdots+G_n^2)^{p/2}\right]\\ &= (2\gamma)^p\frac{\Gamma(1-p/\alpha)}{\Gamma(1-p/2)}\frac{\Gamma((n+p)/2)}{\Gamma(n/2)},\end{aligned} \tag{6}$$

in which the Gamma function $\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}\mathrm{d}t$ and $-n < p < \alpha$. Accordingly, we have the following lemma:

Lemma 1 The $\ln R$ has moment generating function $\mathbb{E}\exp(u\ln R) = \mathbb{E}R^u$ given by Eq. (6), and its first two moments are

$$\mathbb{E}(\ln R) = \ln(2\gamma) + \gamma_{\text{Euler}}\left(\frac{1}{\alpha}-\frac{1}{2}\right)+\frac{1}{2}\psi(n/2), \tag{7}$$

$$\mathrm{Var}(\ln R) = \frac{\pi^2}{6}\left(\frac{1}{\alpha^2}-\frac{1}{4}\right)+\frac{1}{4}\psi'(n/2), \tag{8}$$

where the digamma function $\psi(z) = \Gamma'(z)/\Gamma(z)$ and Euler's constant $\gamma_{\text{Euler}} = -\Gamma'(1) \approx 0.57721$.

Combined with Eq. (5) (L_1 can be seen as a one-dimensional rotationally symmetric stable random vector), we obtain

$$\begin{aligned}&\lim_{t\to 0}\mathbb{E}\left(\ln|x(t)-z|-\frac{1}{\alpha}\ln t\right)\\ &= \ln(2\sigma)+\gamma_{\text{Euler}}\left(\frac{1}{\alpha}-\frac{1}{2}\right)+\frac{1}{2}\psi(n/2),\end{aligned} \tag{9}$$

$$\begin{aligned}&\lim_{t\to 0}\mathrm{Var}\left(\ln|x(t)-z|-\frac{1}{\alpha}\ln t\right)\\ &= \frac{\pi^2}{6}\left(\frac{1}{\alpha^2}-\frac{1}{4}\right)+\frac{1}{4}\psi'(n/2).\end{aligned} \tag{10}$$

Due to the independence of stochastic components in each direction, the noise parameters of system 1 can be calculated out individually along each direction. Consider the integration of small step h in the previous section, denote Z_i and X_i for the points in the ith direction of system 1 in Z and X. According to these points, we can compute the first two moments as

$$m_i = \mathbb{E}_i[(\ln|x_i(h)-z_i|)], \tag{11}$$

$$V_i = \mathrm{Var}_i[(\ln|x_i(h)-z_i|)]. \tag{12}$$

Next we drop the subscript i for briefness. In Eq. (10), the constant $-\frac{1}{\alpha}\ln h$ does not affect the variance and the expression is independent of σ. Thus, the stability index can be calculated as

$$\alpha = \left[\frac{6}{\pi^2}\left(V-\frac{1}{4}\psi'(n/2)\right)+\frac{1}{4}\right]^{-1/2}. \tag{13}$$

Substituting it back to Eq. (9) yields

$$\sigma = \frac{1}{2}\exp\left[m-\gamma_{\text{Euler}}\left(\frac{1}{\alpha}-\frac{1}{2}\right)-\frac{1}{2}\psi(n/2)-\frac{1}{\alpha}\ln h\right]. \tag{14}$$

Remark that this method does not suffer from curse of dimensionality due to the transformation to the amplitude R of the considered data for an arbitrary-dimension system. That is to say, the increase of dimensionality does not necessarily require more data.

3. Results

3.1. A one-dimensional system

Consider a one-dimensional stochastic differential equation with the α-stable Lévy motion

$$\mathrm{d}x(t) = (4x(t)-x^3(t))\mathrm{d}t + \mathrm{d}L_t^{\alpha}. \tag{15}$$

The drift coefficient $b(x) = 4x-x^3$, and the noise intensity $\sigma = 1$. L_t^{α} is a scalar, real-value α-stable Lévy motion with triple $(0,0,\nu_{\alpha})$. The stability parameter α is set as 1.5 as well as the jump measure given by

$$\nu_{\alpha}(\mathrm{d}x) = c(1,\alpha)|y|^{-1-\alpha}\mathrm{d}x$$

with

$$c(1,\alpha) = \frac{\alpha\Gamma((1+\alpha)/2)}{2^{1-\alpha}\pi^{1/2}\Gamma(1-\alpha/2)},\quad \alpha\in(1,2).$$

The time step is chosen as $h = 0.001$. For the initial vector Z, set $M = 5000$ initial points uniformly distributed during $[-2.5,2.5]$. The Euler scheme for numerical integral to access X is

$$x_i = z_i + \sigma h^{1/\alpha}L_1,\quad i = 1,2,\ldots,M,$$

where $L_1 \sim S_{\alpha}(1,0,0)$ generated by the program of Veillette.[33] The cross validation (CV) marks the divergence of each iterative step's prediction with the trained model from the actual test set to choose the optimal sparsity. The parameter called CV score δ reflects the internal balance between accuracy and predictivity of each sparsity. It is defined as an average of all data groups' deviations from real data to fitting data within each iterative step in the cross validation procedure. As a function of the solution size p, it can monitor whether the learning framework achieves an optimal fitting solution at certain step and should quit timely. Figure 1 shows the relationship with δ and sparsity index p. The over-fitting and under-fitting regimes are obvious to easily choose optimal sparsity at the peak. One can see more details in Ref. [8].

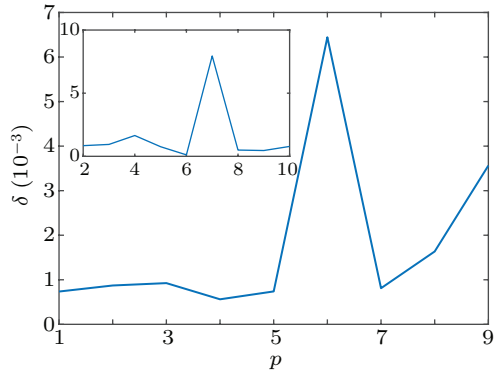


Fig. 1. The CV score δ to the solution size p for the identification of the drift of example 1. The inset shows the ratio of two adjacent δ's to p, which peaks sharply at the critical point $p = 7$.

The learning result is shown in Fig. 2. The true drift coefficient matches well with the learning one with estimated values of stability parameter and noise intensity $\alpha = 1.4954$, $\sigma = 0.9987$ in comparison to the true values $\alpha = 1.5$, $\sigma = 1$.

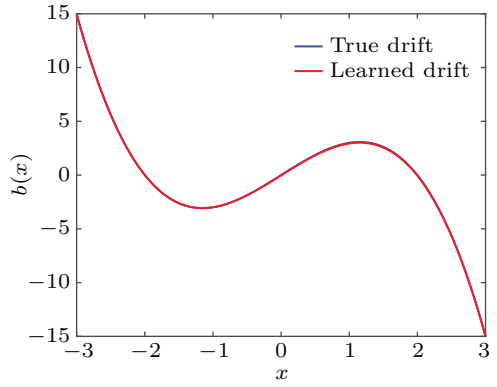


Fig. 2. Learned drift of the one-dimensional system.

3.2. A two-dimensional system

Consider a two-dimensional genetic regulatory network[34] in biology

$$\begin{aligned}\frac{\mathrm{d}k}{\mathrm{d}t} &= a_k + \frac{b_k k^n}{k_0^n + k^n} - \frac{k}{1+k+s} + \sigma_1 \mathrm{d}L_t^1,\\ \frac{\mathrm{d}s}{\mathrm{d}t} &= \frac{b_s}{1+(k/k_1)^p} - \frac{s}{1+k+s} + \sigma_2 \mathrm{d}L_t^2,\end{aligned} \tag{16}$$

where (L_t^1, L_t^2) is the two-dimensional Lévy motion with independent scalar symmetric α-stable Lévy motions which have the same generating triplet $(0,0,\nu_\alpha)$. Set the parameters as $a_k = 0.004$, $b_k = 0.14$, $b_s = 0.68$, $k_0 = 0.2$, $k_l = 0.222$, $n = 2$, $p = 5$, $\sigma = \sigma_1 = \sigma_2$, and make the scale transformation with $k' = 10k$, $s' = 2s$. Thus there are two stable locations for the deterministic part located at $(0.15262, 4.3148)$ and $(1.5732, 3.1562)$, and a saddle on the boundary of attraction basins. For the stochastic part, set $\sigma = 0.25$, $\alpha = 1.5$.

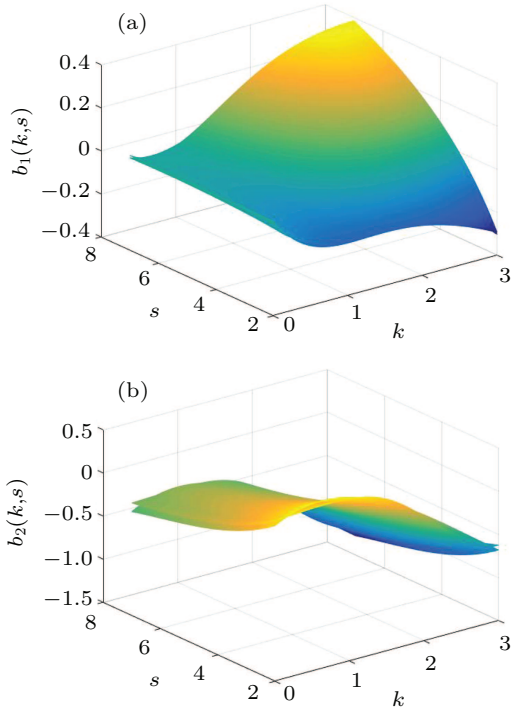


Fig. 3. The three-dimensional learning results of (a) $b_1(k,s)$ and (b) $b_2(k,s)$ in comparison to the original drift terms.

Fixing the time step $h = 0.001$ and $M = 20000 \times 20000$ mesh points uniformly distributed within $[0,3] \times [2,7]$, the one-step numerical integration is taken as

$$\begin{aligned} k_i &= x_i(1) = z_i(1) + b_1(z_i)h + \sigma_1 h^{1/\alpha_k} L_1^1,\\ s_i &= x_i(2) = z_i(2) + b_2(z_i)h + \sigma_2 h^{1/\alpha_s} L_1^2,\\ i &= 1,2,\ldots,M. \end{aligned}$$

The three-dimensional learning drifts are shown in Fig. 3 in comparison to the original b_1 and b_2. In order to see the fitting results more clearly, cross sections are made from different directions in Fig. 4.

The basis functions and their corresponding learning coefficients are listed in Table 1. The learning results for different α's are listed in Table 2.

Table 1. Learning results for the drifts of the 2D biological model.

Basis	1	k	s	k^2	ks	s^2	k^3	ks^2	k^2s	s^3
b_1[a]	0.0848	0.2299	0.1701	0	0.1607	–0.0622	–0.1740	0	–0.0524	0
b_2[b]	–0.1719	–0.6779	–0.2431	–0.4175	0	0.1010	0.1862	0	0	–0.0559

[a]The original expression is $b_1(k,s) = 10 \times \left(a_k + \frac{b_k(k/10)^n}{k_0^n + (k/10)^n} - \frac{k/10}{1+k/10+s/2}\right)$.

[b]The original expression is $b_2(k,s) = 2 \times \left(\frac{b_s}{1+(k/(10k_1))^p} - \frac{s/2}{1+k/10+s/2}\right)$.

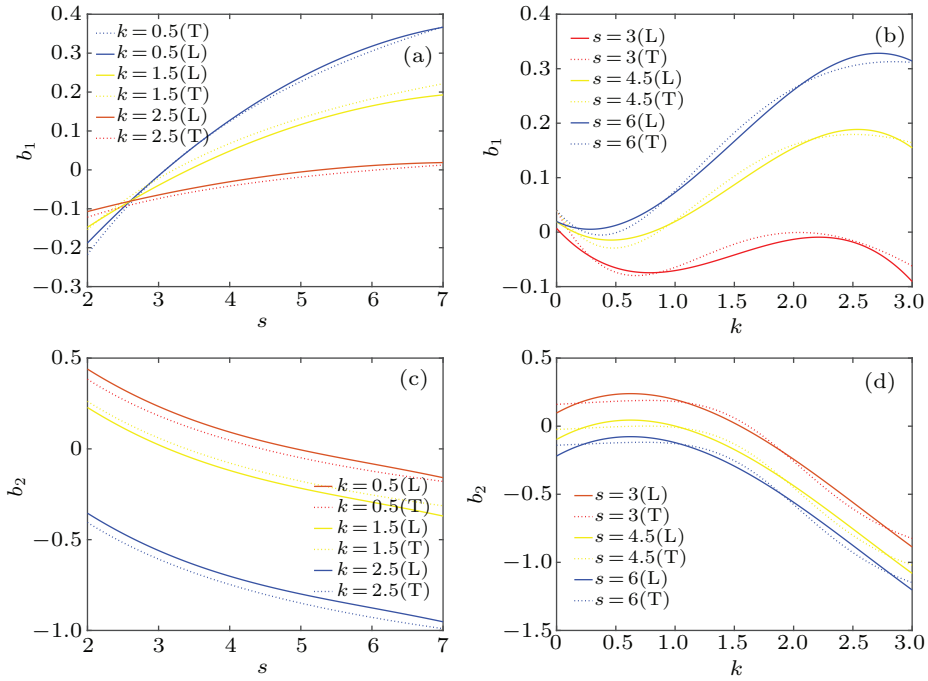


Fig. 4. The cross sections of Fig. 3 for different k's and s's. T is short for a true result and L for a learning result.

Table 2. Learning results of the Lévy noise in the 2D system from different α's.

True \ Learning	α_k	α_s	σ_1	σ_2
$\alpha=1.2, \sigma=0.25$	1.1993	1.2012	0.2517	0.2512
$\alpha=1.4$	1.3977	1.4006	0.2487	0.2524
$\alpha=1.5$	1.5167	1.4862	0.2572	0.2517
$\alpha=1.6$	1.6183	1.6231	0.2532	0.2488
$\alpha=1.8$	1.8233	1.7943	0.2506	0.2483

3.3. A three-dimensional system

Consider the three-dimensional classic Lorenz system

$$\mathrm{d}x_1 = 10(-x_1+x_2)\mathrm{d}t + 2\mathrm{d}L_{1,t}, \tag{17}$$

$$\mathrm{d}x_2 = (4x_1 - x_2 - x_1x_3)\mathrm{d}t + 2\mathrm{d}L_{2,t}, \tag{18}$$

$$\mathrm{d}x_3 = (-8/3x_3 + x_1x_2)\mathrm{d}t + \mathrm{d}L_{3,t}. \tag{19}$$

Then $\sigma_1=2$, $\sigma_2=2$, $\sigma_3=1$. Set $\alpha_1=1.2$, $\alpha_2=1.4$, $\alpha_3=1.5$. For the identification, fix $h=0.001$ and choose Z distributed uniformly within $[-2,2]\times[-2,2]\times[-2,2]$ with a mesh $400\times400\times400$. The results are listed in Tables 3 and 4, which are comparable with the original data in Eqs. (17)–(19).

Table 3. Learning results for the drifts of the 3D Lorenz system.

Basis	1	x_1	x_2	x_1^2	x_1x_2	x_2^2	x_1^3	$x_1^2x_2$	$x_1x_2^2$	x_2^3
b_1	0	–9.9567	9.9412	0	0	0	0	0	0	0
b_2	0	3.9869	–1.0012	0	0	0	–1.0023	0	0	0
b_3	0	0	0	–2.6642	0	0.9988	0	0	0	0

Table 4. Learning results of the Lévy noise in the 3D system.

Parameters	α_1	α_2	α_3	σ_1	σ_2	σ_3
True	1.2	1.4	1.5	2	2	1
Learning	1.1934	1.4036	1.4965	1.9958	2.0128	0.9983

4. Summary

In this paper, we have proposed a framework to extract the physical laws from data of n-dimensional stochastic dynamical systems perturbed by a sort of non-Gaussian noise called α-stable Lévy noise.

Specifically, the nonlocal Kramers–Moyal formulas provide a theoretical basis to approximate the drift term via the short-term evolution data and then a machine learning method is performed to solve the subsequent regression problem. Secondly, the link between the first two moments of the amplitude of increment of the sample paths and the noise parameters (stability parameter and noise intensity) is explored. Thus we have established a bridge between the system's short-term evolution data and the long-term governing factors. Several examples are shown to illustrate the feasibility and some details. In a one-dimensional double-well potential system, we conduct the fitting process and briefly analyze the avoidance of the under-fitting and over-fitting situations. A 2D genetic regulatory network and the classic 3D Lorenz system are then investigated with several different sets of noise parameters. This data-driven method are rather simple and robust. However, it still remains some challenges to be further dealt with such as situations where the stability parameter ranges between $0<\alpha\leq1$ or the category of the Lévy noise alters to n-dimensional rotationally symmetric case.

Overall, this data-driven method is rather rapid and convenient for both avoidance of the "curse of dimensionality" and direct expressions in terms of the first two moments of the amplitude of the sample paths' increments. It does not need much data and the run time is short. For the Gaussian white

noise case, there are also corresponding well-known Kramers–Moyal formulas to construct the identification algorithm. We believe its promising applications in physical fields.

Appendix A: Lévy process

An n-dimensional Lévy motion L_t is a stochastic process which satisfies:

(i) $L_0 = 0$, almost sure.

(ii) Independent increments: For any $t_1 < t_2 < \cdots < t_{n-1} < t_n$ and any n, the random variables $L_{t_2} - L_{t_1}, \ldots, L_{t_n} - L_{t_{n-1}}$ are independent.

(iii) Stationary increments: $L_t - L_s$ and L_{t-s} have the same distribution.

(iv) Stochastically continuous sample paths: $L_t \to L_s$ in probability as $t \to s$ for any $s \geq 0$.

The characteristic function for a pure jump n-dimensional Lévy process is defined as

$$E\exp(\mathrm{i}u^{\mathrm{T}}L_t) = \exp\left[t\int_{\mathbb{R}^n\setminus\{0\}}\left(\mathrm{e}^{\mathrm{i}u^{\mathrm{T}}y} - 1 - \mathrm{i}u^{\mathrm{T}}y\chi_{|y|\leq 1}\right)\nu(\mathrm{d}y)\right],$$

where χ_S is the indicator function of S. The jump measure ν defined on $\mathbb{R}^n\setminus\{0\}$ is subject to

$$\int_{\mathbb{R}^n\setminus\{0\}}(|y|^2 \wedge 1)\nu(\mathrm{d}y) < \infty,$$

which actually characterizes the jump frequency and size for the sample paths of this process.

A popular type of the Lévy motion is the α-stable Lévy motion, which is defined by imposing an extra stability restriction on increments of the process in the original definition. Usually its distribution is denoted as $S_\alpha(\delta,\beta,\lambda)$, where $\alpha \in (0,2]$ is called the stability parameter, $\delta \in (0,\infty)$ the scaling parameter, $\beta \in [-1,1]$ the skewness parameter and $\lambda \in (-\infty,\infty)$ the shift parameter.

In this paper, we consider an α-stable Lévy motion that is symmetric ($\beta = 0$) and independent in every direction with the kernel $\nu_i^\alpha(y)$ in the i-th direction written as

$$\nu_i^\alpha(y) = \frac{C_{\alpha_i}}{|y|^{1+\alpha_i}}, \tag{A1}$$

where α_i is the i-th argument of α and the intensity constant

$$C_{\alpha_i} = \frac{\alpha_i\Gamma((1+\alpha_i)/2)}{2^{1-\alpha_i}\pi^{1/2}\Gamma(1-\alpha_i/2)}, \tag{A2}$$

where Γ is the Gamma function.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant No. 12172167), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

[1] Sun J C 2016 *Chin. Phys. Lett.* **33** 100503
[2] Yan X and Wu Y 2012 *Chin. Phys. Lett.* **29** 128901
[3] Xu W, Wang L, Feng J, Qiao Y and Han P 2018 *Chin. Phys. B* **27** 110503
[4] Jordan M I and Mitchell T M 2015 *Science* **349** 255
[5] Marx V 2013 *Nature* **498** 255
[6] Williams M O, Kevrekidis I G and Rowley C W 2015 *J. Nonlinear Sci.* **25** 1307
[7] Schmid P J 2010 *J. Fluid Mech.* **656** 5
[8] Boninsegna L, Nüske F and Clementi C 2018 *J. Chem. Phys.* **148** 241723
[9] Brunton S L, Proctor J L and Kutz J N 2016 *Proc. Natl. Acad. Sci. USA* **113** 3932
[10] Schaeffer H, Caisch R, Hauck C D and Osher S 2013 *Proc. Natl. Acad. Sci. USA* **110** 6634
[11] Rudy S, Alla A, Brunton S L and Kutz J N 2019 *SIAM J. Appl. Dyn. Syst.* **18** 643
[12] Chen R T Q, Rubanova Y, Bettencourt J and Duvenaud D K 2018 *Advances in Neural Information Processing Systems* (Curran Associates, Inc.) Vol. 31
[13] Li X, Wong T K L, Chen R T and Duvenaud D 2020 *Proceedings of Machine Learning Research* (PMLR) Vol. 118, pp. 1–28
[14] Garcia C A, Otero A, Felix P, Presedo J and Marquez D G 2017 *Phys. Rev. E* **96** 022104
[15] Ruttor A, Batz P and Opper M 2013 *Advances in Neural Information Processing Systems* (Curran Associates, Inc.) Vol. 26
[16] Ditlevsen P D 1999 *Geophys. Res. Lett.* **26** 1441
[17] Raser J M and O'shea E K 2005 *Science* **309** 2010
[18] Jourdain B, Méléard S and Woyczynski W A 2012 *J. Math. Biol.* **65** 677
[19] Zoia A, Rosso A and Kardar M 2007 *Phys. Rev. E* **76** 021116
[20] Matthäus F, Mommer M S, Curk T and Dobnikar J 2011 *PloS One* **6** e18623
[21] Humphries N E, Queiroz N, Dyer J R, *et al.* 2010 *Nature* **465** 1066
[22] Ramos-Fernández G, Mateos J L, Miramontes O, Cocho G, Larralde H and Ayala-Orozco B 2004 *Behav. Ecol. Sociobiol.* **55** 223
[23] Li Y and Duan J 2021 *Physica D* **417** 132830
[24] Li Y and Duan J 2022 *J. Stat. Phys.* **186** 1
[25] Li Y, Duan J, Liu X and Zhang Y 2020 *Chaos* **30** 063142
[26] Rosenkrantz R D 2012 *ET Jaynes: Papers on probability, statistics and statistical physics* (Berlin: Springer Science & Business Media) Vol. 158
[27] Huang R Z, Liao H J, Liu Z Y, Xie H D, Xie Z Y, Zhao H H, Chen J and Xiang T 2018 *Chin. Phys. B* **27** 070501
[28] Zhang J P, Guo H M, Jing W J and Jin Z 2019 *Acta Phys. Sin.* **68** 150501 (in Chinese)
[29] Radovic A, Williams M, Rousseau D, Kagan M, Bonacorsi D, Himmel A, Aurisano A, Terao K and Wongjirad T 2018 *Nature* **560** 41
[30] Carrasco Kind M and Brunner R J 2013 *Mon. Not. R. Astron. Soc.* **432** 1483
[31] Butler K T, Davies D W, Cartwright H, Isayev O and Walsh A 2018 *Nature* **559** 547
[32] Lindsey J 1999 *Biometrics* **55** 1277
[33] Veillette M 2014 https://github.com/markveillette/stbl
[34] Süuel G M, Garcia-Ojalvo J, Liberman L M and Elowitz M B 2006 *Nature* **440** 545