Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2024-0001

Vol. **37**, No. 5, pp. 1277-1304
May 2025

Weak Collocation Regression for Inferring Stochastic Dynamics with Lévy Noise

Liya Guo[1], Liwei Lu[1], Zhijun Zeng[1], Pipi Hu[2] and Yi Zhu[1,3,*]

[1] *Yau Mathematical Sciences Center, Tsinghua University, Beijing, 100084, China.*
[2] *Microsoft Research AI4Science, Beijing, 100080, China.*
[3] *Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing, 101408, China.*

Received 2 January 2024; Accepted (in revised version) 12 June 2024

Abstract. With the rapid increase of observational, experimental and simulated data for stochastic systems, tremendous efforts have been devoted to identifying governing laws underlying the evolution of these systems. Despite the broad applications of non-Gaussian fluctuations in numerous physical phenomena, the data-driven approaches to extracting stochastic dynamics with Lévy noise are relatively few. In this work, we propose a Weak Collocation Regression (WCR) to explicitly reveal unknown stochastic dynamical systems, i.e., the Stochastic Differential Equation (SDE) with both α-stable Lévy noise and Gaussian noise, from discrete aggregate data. This method utilizes the evolution equation of the probability distribution function, i.e., the Fokker-Planck (FP) equation. With the weak form of the FP equation, the WCR constructs a linear system of unknown parameters where all integrals are evaluated by Monte Carlo method with the observations. Then, the unknown parameters are obtained by a sparse linear regression. For a SDE with Lévy noise, the corresponding FP equation is a partial integro-differential equation (PIDE), which contains nonlocal terms, and is difficult to deal with. The weak form can avoid complicated multiple integrals. Our approach can simultaneously distinguish mixed noise types, even in multi-dimensional problems. Numerical experiments demonstrate that our method is accurate and computationally efficient.

AMS subject classifications: 70F17, 65N35, 45Q05, 65M70, 35S10

Key words: Weak collocation regression, learning stochastic dynamics, Lévy process, Fokker-Planck equations, weak SINDy.

---

*Corresponding author. *Email addresses:* `gly22@mails.tsinghua.edu.cn` (L. Guo), `llw20@mails.tsinghua.edu.cn` (L. Lu), `zengzj22@mails.tsinghua.edu.cn` (Z. Zeng), `pisquare@microsoft.com` (P. Hu), `yizhu@tsinghua.edu.cn` (Y. Zhu)

1 Introduction

Learning the laws behind time-series data has been a very hot topic. These laws are often described by ordinary differential equations (ODEs) and partial differential equations (PDEs) [22,29,46]. Depending on the physical problems, differential equations with stochastic terms are also used [10–12]. In this context, we consider the data modeled by stochastic differential equations (SDEs), in line with prior studies [10,30,31]. These SDEs typically encompass drift and diffusion terms, both being functions of state variables, while the diffusion terms are often considered to be arisen from Gaussian noise. Common methodologies employed to reveal unknown systems driven by Gaussian noise include parameter or non-parametric inference [5,11,12,19,37,39,40], and neural network-based approaches [24,31,36].

Phenomena such as stock price fluctuations and abnormal diffusion [2, 15] display heavy-tailed distributions and jumps, which cannot be precisely characterized by SDEs driven solely by Gaussian noise. Consequently, there is a shift in focus to consider SDEs with Lévy noises, exhibiting stationary and independent increments that depend solely on the time interval. Recent studies have revealed the coexistence of Gaussian and Lévy transports in various complex systems [8, 47, 48]. However, there is relatively little research on the inverse problem of dynamic systems containing Lévy noise.

We consider that the stochastic system follows SDEs driven by both Gaussian and Lévy noises, characterized by sufficient parameters to capture its dynamics. Challenges posed by Lévy noise, due to its jump properties which lead to an infinite computational domain, necessitate the consideration of a bounded area [18,44,45]. Some previous works apply Kramers-Moyal formula to directly represent the unknown terms [27, 28]. Leveraging the fact that the data distribution from an SDE adheres to a specific Fokker-Planck (FP) equation [14, 41], [10, 44] to infer the dynamic systems by neural networks. Motivated by [10], we reveal the SDEs based on the FP equation. Therefore, estimating the unknown parameters of the SDEs is transformed into inferring the corresponding parameters of the PDEs. However, PDEs arising from dynamics with Lévy noise, known as Partial Integro-Differential Equations (PIDEs), incorporate nonlocal integral terms. They are also classified as fractional partial differential equations (FPDEs) due to the presence of fractional-order derivatives. Although numerical methods such as the finite difference method are employed to solve low-dimensional PIDEs [18], extending these methods to high dimensions may result in the curse of dimensionality. Recent studies [20, 38] demonstrate the application of machine learning methods for the inversion of FPDEs in high-dimensional problems, and [21] theoretically analyzes that to what extent one can recover PDE's parameters.

In many practical scenarios, trajectory data may not be readily accessible, necessitating the inference of a stochastic dynamical system from discrete observations. Models such as physics-informed neural networks (PINNs) and Wasserstein generative adversarial networks (WGANs) have been used to learn inversions of dynamical systems from discrete sparse samples [10,31,44]. These methods facilitate the estimation of probability