

Lévy Backward SDE Filter for Jump Diffusion Processes and Its Applications in Material Sciences

Feng Bao^{1,*}, Richard Archibald² and Peter Maksymovych³

¹ *Department of Mathematics, Florida State University, Tallahassee, Florida, 32306, USA.*

² *Computer Science and Mathematics Division, Oak Ridge, Tennessee, 37831, USA.*

³ *Center for Nanophase Material Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831, USA.*

Received 15 September 2018; Accepted (in revised version) 5 March 2019

Abstract. We develop a novel numerical method for solving the nonlinear filtering problem of jump diffusion processes. The methodology is based on numerical approximation of backward stochastic differential equation systems driven by jump diffusion processes and we apply adaptive meshfree approximation to improve the efficiency of numerical algorithms. We then use the developed method to solve atom tracking problems in material science applications. Numerical experiments are carried out for both classic nonlinear filtering of jump diffusion processes and the application of nonlinear filtering problems in tracking atoms in material science problems.

AMS subject classifications: 65C30, 65K10

Key words: Nonlinear filtering problem, backward SDEs, jump diffusion processes, material sciences.

1 Introduction

The nonlinear filtering problem is one of the key missions in data assimilation, in which observations of a system are incorporated into the state of a numerical model of that system. Mathematically, the nonlinear filtering problem is to obtain, recursively in time, the best estimate of the state of unobservable stochastic dynamics $S = \{S_t : t \geq 0\}$, based on an associated observation process, $M = \{M_t : t \geq 0\}$, whose values are a function of S after corruption by noises. This suggests the optimal filtering problem of obtaining the conditional distribution of the state S_t from the observations up until time t , which

*Corresponding author. *Email addresses:* bao@math.fsu.edu (F. Bao), archibaldrk@ornl.gov (R. Archibald), maksymovychp@ornl.gov (P. Maksymovych)

achieves the best estimate of this distribution, in the squared error sense, based on the available observations.

The nonlinear filtering theory finds its applications in numerous scientific and engineering research areas, such as target tracking [17, 34], signal processing [26, 38], image processing [51, 52], biology [16, 37, 53], or mathematical finance [10, 19, 21]. Some of the pioneer contributions to the development of nonlinear filters are due to Kushner [36] and Stratonovich [54]. Later, Zakai [59] introduced an alternative approach to the computation of the nonlinear filter by developing the so-called Zakai equation, which is a stochastic partial differential equation (SPDE), and the best estimate of the nonlinear filter, i.e. the conditional distribution, is represented by the solution of the Zakai equation. Although the Zakai's approach produces the "exact" solution of the nonlinear filtering problem in theory, solving the SPDEs numerically can be extraordinarily difficult, especially when the state processes are in high dimensions [5, 23, 32, 60]. A more widely accepted method by practitioners to solve the nonlinear filtering problem is the sequential Monte Carlo approach, which is also known as the particle filter method [4, 11, 13, 18, 24, 33, 40, 41]. The particle filter method uses a number of independent random variables, called particles, sampled directly from the state space to represent the prior probability, and updates the prior by including the new observation to get the posterior. This particle system is properly located, weighted and propagated recursively according to Bayes' theorem. As a Monte Carlo approach, with sufficient large number of samples the particle filter provides an accurate representation of the state probability density function (pdf) as desired in the nonlinear filtering problem. Convergence of a particle filter to the optimal filter was shown under certain conditions [14, 15, 27]. In addition to the Zakai's approach and Monte Carlo type approach, the authors have developed an alternative method, which solves the nonlinear filtering problem through a forward backward doubly stochastic differential equations (BDSDEs) system. The theoretical basis of the BDSDEs approach is the fact that the BDSDEs system is equivalent to a parabolic type SPDE and the solution of that system is the conditional distribution of the state as required in the nonlinear filtering problem [3, 6–8]. In this connection, it produces the exact solution of the nonlinear filtering problem, just like the Zakai's approach. In the meantime, as a stochastic *ordinary* differential equation (SDE) approach, it also relies on stochastic sampling, just like the particle filter method. Therefore, the BDSDEs approach builds the bridge between the Zakai's approach and the Monte Carlo type approach.

In this paper, we consider a more general nonlinear filtering problem – the nonlinear filtering problem for jump diffusion processes, in which the state process S_t is a jump diffusion process and the state dynamic is perturbed by both traditional Gaussian noises and other kinds of Lévy type noises. Different from classical nonlinear filtering problems, numerical methods to solve the nonlinear filtering problem for jump diffusion processes are not well developed. The existing methods for solving this type of problems focus on numerical approximation for its corresponding Zakai equation [43, 49, 50]. However, due to the nonlocal behavior of the state dynamics as a jump diffusion process, the corresponding Zakai equation contains fractional derivatives in spatial dimension, which is a

stochastic partial-integral differential equation (PIDE). The current numerical approaches to solve this type of stochastic PIDEs are extensions of existing methods for local partial differential equations, such like finite element, finite difference, etc.. It is well known that the partial-integral operator is a nonlocal operator, which may result in severe computational difficulties coming from the dramatic deterioration of the sparsity of the stiffness matrices required by the underlying linear systems. Although the particle filter framework could be extended to solve the nonlinear filtering problem for jump diffusion processes, as a typical drawback of Monte Carlo type approach, the particle filter method has very poor tail approximation for the state distribution especially when the state is perturbed by Lévy type noises, which causes severe degeneracy problem for long term simulations [12, 18].

In this work, we aim to develop an efficient numerical method to solve the nonlinear filtering problem for jump diffusion processes. The main theme of our approach is established on numerical approximations for a special kind of SDEs system, which is combined by a forward SDE and a backward SDE. We call this SDEs system a *backward SDEs system*, and name the general backward SDEs approach for solving nonlinear filtering problems the *Backward SDE filter* [8]. Instead of using the equivalent relation between BDSDEs and the standard Zakai equation as introduced in [8], in our novel approach for the nonlinear filtering problem for jump diffusion processes, we derive a backward SDEs system driven by jump diffusion processes to describe the time evolution of the state dynamics S_t . The theoretical validation of this effort is based on the probabilistic interpretation for PIDEs, which is discussed in [2, 9], and our methodology in this paper simplifies the BDSDEs approach by avoiding simulations of the doubly stochastic term in BDSDEs. When receiving the observation measurements, we incorporate the data with the backward SDEs system by using the Bayes' theorem. In this connection, our method also produces a representation of the conditional state distribution as required in the nonlinear filtering problem and it maintains the accuracy property of the Zakai's approach. On the other hand, different from the Zakai's approach, the Backward SDE filter for jump-diffusion processes is still an *SDE* based approach and it allows us to approximate the solution of nonlinear filtering problem on any selection of space points. Taking this advantage, we introduce a stochastic space points generation algorithm which generates meshfree space points adaptively according to the state dynamics and the observations. The central idea of this meshfree space points generation is to build dynamic random space points according to the approximate state distribution. Specifically, we generate a set of random samples from the initial distribution of the state and choose these samples as our random space points. Then, we propagate these space points through the dynamic system of the state. In this way, the space points move randomly according to the state model and are more concentrated around the state of the dynamic model in the nonlinear filtering problem. It is important to point out that the numerical approximation of the backward SDEs system for different selection of space points is independent. Therefore the Backward SDE filter has the same scalability as the Monte Carlo type approach. However, there is an essential difference with respect to the selection of random space

points in the Backward SDE filter and the generation of random samples in the Monte Carlo type approaches. In the Backward SDE filter we approximate the value of state pdf on each space point instead of using the number of samples to describe the empirical distribution. In this way, the backward SDE filter requires much fewer space points compared to the sample-size in the Monte Carlo method. With the numerical approximation of the solution of the backward SDE on the dynamic space points, we apply the Shepard's method which is an effective meshfree interpolation method to construct the approximation of the entire conditional pdf in the state space. To prevent space points degeneracy for long term simulations, we introduce a Markov Chain Monte Carlo based resampling method to update the space points according to the observation information. In order to distinguish the Backward SDE filter introduced in [8], in this paper, we name the Backward SDE filter for jump diffusion processes as the "Lévy Backward SDE filter". It is also worthy to mention that although the original Backward SDE filter would solve the classic nonlinear filtering problem effectively and efficiently, since its lack of the assumption for the jumps in the state model, it is not suitable for estimating the state of stochastic dynamics in the form of jump diffusion processes. With the appropriate Lévy terms introduced in the backward SDE system, to our best knowledge, the Lévy Backward SDE filter discussed in this paper is the first attempt to use backward SDE systems to solve the nonlinear filtering problem for jump diffusion processes.

Taking the aforementioned advantages of the Lévy Backward SDE filter, we then apply our developed method in solving problems in material science applications. Specifically, we study the application of Lévy Backward SDE filter in tracking atom trajectories moving on material surfaces. The so-called atom-tracking method provides a general approach to study diffusion of single atoms and molecules on a flat surface (typically a noble metal), and it also underlies the mechanism of single-atom manipulation [29, 31, 44]. In both cases an atomically sharp probe (made from a sharp metal needle) is scanned over the surface with sub-angstrom accuracy, typically producing the images of atomic-scale corrugation of the electronic density of states (in the case of scanning tunneling microscopy) or total electron density (in the case of atomic force microscopy [22]). The regimes of observation and manipulation are delineated in these methods of scanning probe microscopy by the strength of probe-atom (or probe-molecule) interaction, which in turn can be controlled by the degree of proximity between the probe and the observed atom [39, 44]. Many interesting examples of this method have been demonstrated – from creation of artificial atomic structures (such as quantum corrals [29]) to non-trivial diffusion mechanisms, including quantum tunneling of hydrogen. The tracking efficiency translates into both success and relative speed of manipulation, and is therefore of utmost importance. The frequency of both manipulation and diffusion of single atoms generally follow Poisson statistics, which makes the analysis of corresponding time-series a great fit to our Lévy Backward SDE filter.

The major contribution of this paper is in twofolds: to develop an efficient Lévy Backward SDE filter for jump diffusion processes, and to introduce the potential application of the developed algorithms in material sciences. The rest of this paper is organized as

following. In Section 2, we introduce some preliminary theories for the nonlinear filtering problem and the backward SDEs system. In Section 3, we develop the methodology and numerical algorithms for our Lévy Backward SDE filter. Numerical experiments are carried out in this section to examine the performance of our algorithms. In Section 4, we discuss the application of the Lévy Backward SDE filter in material sciences and demonstrate the effectiveness of the Lévy Backward SDE filter in tracking atom trajectories on two different material surfaces.

2 Preliminaries

In this section, we introduce preliminary theories for the methodology of this paper. We first give a brief discussion to nonlinear filtering problems for jump diffusion processes. Then we introduce the backward SDEs system for jump diffusion processes and its relation to integral-partial differential equations.

2.1 Nonlinear filtering problems for jump diffusion processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and denote $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ to be a filtration possesses right continuity, i.e. $\mathcal{F}_t = \mathcal{F}_{t+}$, and \mathcal{F}_0 is the σ -algebra contains all the \mathbb{P} zero measure zero sets. In this paper, the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ is assumed to be generated by two mutually independent processes: a d -dimensional Brownian motion W_t and a d -dimensional Poisson random measure $\mu(t, A)$ on $[0, T] \times \mathbb{E}$ where $\mathbb{E} = \mathbb{R}^k \setminus \{0\}$ is equipped with its Borel field \mathcal{E} , with compensator $\nu(dt, de) = dt\lambda(de)$, such that $\{\tilde{\mu}([0, t] \times A) = (\mu - \nu)([0, t] \times A)\}_{t \geq 0}$ is a martingale for all $A \in \mathcal{E}$ satisfying $\lambda(A) < \infty$ and $\lambda(de)$ is assumed to be a σ -finite measure on $(\mathbb{E}, \mathcal{E})$ satisfying

$$\int_{\mathbb{E}} (1 \wedge |e|^2) \lambda(de) < +\infty,$$

where $|\cdot|$ denotes the standard Euclidean norm in Euclidean spaces.

In this paper, we consider the nonlinear filtering problem of jump diffusion processes in its state-space form on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and introduce the following stochastic processes

$$\begin{aligned} dS_t &= b(S_t)dt + \sigma_t dW_t + \int_{\mathbb{E}} \beta_t(e) \tilde{\mu}(dt, de), \\ M_t &= h(S_t, B_t), \end{aligned} \tag{2.1}$$

where $b: \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $h: \mathbb{R}^d \times \mathbb{R}^l \rightarrow \mathbb{R}^l$ are nonlinear functions, $\beta_t: \mathbb{E} \rightarrow \mathbb{R}^d$ is a d -dimensional process on \mathbb{E} , $W_t \in \mathbb{R}^d$ and $B_t \in \mathbb{R}^l$ are two independent Brownian motions, $\tilde{\mu}$ is the compensated Poisson random measure describes jumps in the model, $\sigma_t \in \mathbb{R}^{d \times d}$ is the coefficient matrix for Brownian motion W_t . The given initial value S_0 of (2.1) is independent of W_t , $\tilde{\mu}$ and B_t , and has a probability distribution with density function Pr . The stochastic process $\{S_t\}_{t \geq 0}$ defined by the stochastic differential equation (SDE) describes the dynamics of a jump diffusion process, which is named the *state process*, and $\{M_t\}_{t \geq 0}$ is the

noise perturbed measurement of the state $\{S_t\}_{t \geq 0}$ with observation function $h(S_t, B_t)$. When the coefficient $\beta \equiv 0$ in the state process, the state-space form (2.1) gives a nonlinear filtering problem of standard Brownian motion driven diffusion processes. In most applications, the measurement M_t is received at discrete time and the noise from measurement can be assumed to be additive. In this way, the measurement is formulated in a discrete manner as

$$M_{t_n} = h(S_{t_n}) + R\dot{B}_{t_n}, \quad n = 1, 2, \dots, \quad (2.2)$$

where $R \in \mathbb{R}^l \times \mathbb{R}^l$ is the covariance matrix of the noise. The goal of the nonlinear filtering problem is to derive the least squares estimate of a functional $\psi(S_t)$ of the state process S_t , given all the observation information of M_t up to time t . In other words, we aim to find the optimal filter $\tilde{\psi}(S_t)$, such that

$$\tilde{\psi}(S_t) := E[\psi(S_t) | \mathcal{M}_t] = \inf\{E[|\psi(S_t) - K_t|^2]; K_t \in \mathcal{K}_t\},$$

where $\mathcal{M}_t := \sigma\{M_s, 0 \leq s \leq t\}$ is the σ -algebra containing all the observation information up to time t , and \mathcal{K}_t is the space of all \mathcal{M}_t measurable and square integrable random variables.

2.2 Backward SDEs driven by jump diffusion processes

As preliminaries for our methodology, we introduce a system of backward SDEs on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ driven by the state process S_t defined in (2.1). Let \mathcal{S}^2 denote the set of \mathcal{F}_t -adapted càdlàg one dimensional processes such that for stochastic process $\{Y_t, 0 \leq t \leq T\} \in \mathcal{S}^2$, we have

$$\|Y\|_{\mathcal{S}^2} = \left\| \sup_{0 \leq t \leq T} |Y_t| \right\|_{L^2(\Omega)} < \infty.$$

Let L_W^2 be the set of \mathcal{F}_t - progressively measurable d -dimensional processes such that for stochastic process $\{Z_t, 0 \leq t \leq T\} \in L_W^2$, we have

$$\|Z\|_{\mathcal{S}^2} = \left(E \int_0^T |Z_t|^2 dt \right)^{1/2} < \infty.$$

In addition, we use $L_{\tilde{\mu}}^2$ to denote the set of mappings $U: \Omega \times [0, T] \times E \rightarrow \mathbb{R}$ which are $\mathcal{P} \otimes \mathcal{E}$ measurable, where \mathcal{P} denotes the σ -algebra of \mathcal{F}_t -predictable subsets of $\Omega \times [0, T]$, and such that

$$\|U\|_{L^2(\tilde{\mu})} := \left(E \int_0^T \int_{\mathbb{E}} U_t(e)^2 \lambda(de) dt \right)^{1/2} < \infty.$$

Finally, we define $\mathcal{B}^2 = \mathcal{S}^2 \times L_W^2 \times L_{\tilde{\mu}}^2$.

We consider the following forward backward stochastic differential equations system

$$\bar{X}_t = \bar{X}_0 + \int_0^t b(\bar{X}_s) ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{E}} \beta_s(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \tag{2.3a}$$

$$Y_t = \psi(\bar{X}_T) + \int_t^T f(s, \bar{X}_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s - \int_t^T \int_{\mathbb{E}} U_s(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \tag{2.3b}$$

where $W_t \in \mathbb{R}^d$ is a d -dimensional standard Brownian motion, $\tilde{\mu}$ is a compensated Poisson random measure, $b: \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma_t \in \mathbb{R}^{d \times d}$, $\beta_t \in \mathbb{R}^d$ and $f: [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a measurable function. For the convenience of presentation, we call the stochastic system (2.3) a backward SDEs system. Eq. (2.3a) in the above backward SDEs system is a forward SDE which describes the same diffusion of the state process (2.1) and Eq. (2.3b) is a backward SDE with a given initial condition $Y_T = \psi(\bar{X}_T)$. The solution of the backward SDEs system (2.3) is a quadruplet $(\bar{X}_t, Y_t, Z_t, U_t)$ with the triple $(Y_t, Z_t, U_t) \in \mathcal{B}^2$. The existence and uniqueness of the solution quadruplet are provided by [9]. It is known that Y_t and Z_t are both functions of \bar{X}_t (see [45]) and we also known from [9] that under the condition $\bar{X}_0 = x \in \mathbb{R}^d$,

$$Z_t = \sigma_t \nabla Y_t, \quad U_t = Y_t(\bar{X}_{t-} + \beta_t(e)) - Y_t(\bar{X}_{t-}), \tag{2.4}$$

where ∇ is the gradient and the minus sign in X_{t-} is the left limit.

An important property of the backward SDEs system is its equivalence to the following parabolic integro-partial differential equation (PIDE)

$$\begin{aligned} -\frac{\partial u_t}{\partial t}(x) &= \mathcal{L}_t u_t(x) + f(t, x, u_t, \sigma_t \nabla u_t(x)), \quad (t, x) \in [0, T] \times \mathbb{R}^d, \\ u_T(x) &= \psi(x), \quad x \in \mathbb{R}^d, \end{aligned} \tag{2.5}$$

where the second order integral-differential operator \mathcal{L} is of the form $\mathcal{L}_t = A_t + K_t$, with

$$A_t \phi(x) = \frac{1}{2} \sum_{i,j=1}^d \left((\sigma_t \sigma_t^*)_{ij} \frac{\partial^2 \phi}{\partial x^2}(x) \right) + \sum_{i=1}^d (b_t)_i(x) \frac{\partial \phi}{\partial x_i}(x), \quad \phi \in C^2(\mathbb{R}^d),$$

$$K_t \phi(x) = \int_{\mathbb{E}} [\phi(x + \beta_t(e)) - \phi(x) - \beta_t(e) \nabla \phi(x)] \lambda(de), \quad \phi \in C^2(\mathbb{R}^d).$$

The side condition of the above equation is given at time T and the propagation direction is backward, i.e. from T to 0. It has been shown in [9] that the solution Y of the backward SDEs system (2.3) is the unique viscosity solution of (2.5), i.e. for any given point $x \in \mathbb{R}^d$ we have

$$Y_t(x) = u_t(x), \quad (t, x) \in [0, T] \times \mathbb{R}^d. \tag{2.6}$$

3 Lévy Backward SDE filter for jump diffusion processes

Now, we introduce the backward SDEs approach for the nonlinear filtering problem of jump diffusion processes. In Subsection 3.1, we first derive the methodology of solving the nonlinear filtering problem through a backward SDEs system and we name this

methodology the *Lévy Backward SDE filter*. In Subsection 3.2, we develop an effective and efficient numerical scheme to solve the Lévy Backward SDE filter. Then, we present numerical examples to show the performance of our Lévy Backward SDE filter in solving the nonlinear filtering problem (2.1) in Subsection 3.3.

3.1 Methodology

The framework of the Lévy Backward SDE filter is composed by two major tasks: (i) estimate the state process without observation information; (ii) incorporate observation with the state estimation. The second task is typically achieved by applying Bayes' theorem, which we also adopt in the Lévy Backward SDE filter for jump diffusion processes. To achieve the first task, we introduce a backward SDEs system, which describes the probability density function (pdf) of the state, and solve the corresponding backward SDEs system.

To proceed, we first consider the following special case of backward SDEs system (2.3)

$$S_t = S_0 + \int_0^t b(S_s) ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{E}} \beta_s(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \quad (3.1a)$$

$$Y_t = \psi(S_T) - \int_t^T Z_s dW_s - \int_t^T \int_{\mathbb{E}} U_s(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \quad (3.1b)$$

where the forward SDE (3.1a) is a standard jump diffusion process which has the same definition as the state process (2.1) in the nonlinear filtering problem. Taking conditional expectation to both sides of (3.1b), from the martingale property of $\int_t^T Z_s dW_s$ and $\int_t^T \int_{\mathbb{E}} U_s(e) \tilde{\mu}(ds, de)$, we obtain

$$E[Y_t] = E[\psi(S_T)],$$

which is a simplified Feynman-Kac formula. From the equivalence relation between backward SDEs systems and PIDEs, we know that the backward SDEs system (3.1) is equivalent to the following PIDE

$$\begin{aligned} -\frac{\partial u_t}{\partial t}(x) &= \frac{1}{2} \sum_{i,j=1}^d \left((\sigma_t \sigma_t^*)_{ij} \frac{\partial^2 u_t}{\partial x^2}(x) \right) + \sum_{i=1}^d (b_t)_i(x) \frac{\partial u_t}{\partial x_i}(x) \\ &\quad + \int_{\mathbb{E}} [u_t(x + \beta_t(e)) - u_t(x) - \beta_t(e) \nabla u_t(x)] \lambda(de) \quad (t, x) \in [0, T] \times \mathbb{R}^d, \\ u_T(x) &= \psi(x), \quad x \in \mathbb{R}^d, \end{aligned} \quad (3.2)$$

which is also known as a Kolmogorov backward equation of jump diffusion processes. By applying integration by parts formula and Taylor expansion, one can prove that the Kolmogorov backward equation (3.2) is adjoint to the following Fokker Planck type equation

tion [50,55]

$$\begin{aligned} \frac{\partial p_t}{\partial t}(x) &= \sum_{i,j=1}^d \frac{1}{2} (\sigma_t \sigma_t^*)_{ij} \frac{\partial^2 p_t}{\partial x^2}(x) - \sum_{i=1}^d \left((b_t)_i(x) \frac{\partial p_t}{\partial x_i}(x) + \frac{\partial (b_t)_i}{\partial x_i}(x) p_t(x) \right) \\ &\quad + \int_{\mathbb{E}} \left(p_t(x - \beta_t(e)) - p_t(x) + \beta_t(e) \nabla p_t \right) \lambda(de), \quad (t, x) \in [0, T] \times \mathbb{R}^d, \\ p_0(x) &= Pr(x), \end{aligned} \tag{3.3}$$

where the initial condition p_0 is the given pdf Pr of the initial state random variable S_0 , and the propagation direction of the equation is forward, i.e. from 0 to T . Actually, it is well known that the solution p_t of the Fokker-Planck type equation (3.3) describes the probability evolution of the stochastic process S_t defined by (2.1), i.e. p_t is the pdf for S_t .

From the equivalence condition between backward SDEs and PIDEs, we know that there's a backward SDEs system corresponding to the Fokker-Planck type equation (3.3). Since the Fokker-Planck type equation (3.3) has opposite time propagation direction to the Kolmogorov backward equation (3.2), the backward SDEs system equivalent to (3.3) also has opposite propagation direction to the backward SDEs system (3.1). From the general equivalence relation between (2.3) and (2.5), and the expression of the Fokker-Planck type equation (3.3), we obtain that the following backward SDEs system is equivalent to equation (3.3)

$$X_0 = X_t - \int_0^t b(X_s) ds - \int_0^t \sigma_s d\overleftarrow{W}_s - \int_0^t \int_{\mathbb{E}} \beta_s(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \tag{3.4a}$$

$$P_t = Pr(X_0) - \int_0^t \sum_{i=1}^d \frac{\partial (b_s)_i}{\partial x_i}(X_s) P_s ds - \int_0^t Q_s d\overleftarrow{W}_s - \int_0^t \int_{\mathbb{E}} V_s(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \tag{3.4b}$$

where $\int_0^t \cdot d\overleftarrow{W}_t$ is an Itô integral integrated backwardly, i.e. from t to 0, adapted to a backward filtration of the Brownian motion W_t and is named *backward Itô integral* (see [46] for details). With any given state X_t as the side condition and the backward Itô integral in the equation, (3.4a) can be considered as an SDE with backward propagation direction. We also want to mention that although we have inverted the propagation direction for the SDE (3.4a), it doesn't matter what sign we put in front of the backward Itô integral $\int_0^t \sigma_s d\overleftarrow{W}_s$ due to the symmetric property of Itô integrals. Similarly, we can observe that the second equation in (3.4) is also a time inverse backward SDE with initial condition $P_0 = Pr$ and has the same structure to (2.3b). As a result, (3.4) is a time inverse backward SDEs system with solution quadruplet (X_t, P_t, Q_t, V_t) . From the equivalent relation (2.6), we know that the solution P_t of the backward SDEs system (3.4) is equivalent to the solution p_t of the Fokker-Planck type equation (3.3), i.e.

$$P_t(x) = p_t(x). \tag{3.5}$$

Therefore, the numerical approximation of the solution P_t in (3.4) is also the approximation for the pdf of state S_t .

In this connection, we introduce the methodology of the Lévy Backward SDE filter, which solves for the conditional pdf of the state process in two steps: a *Prediction Step* and an *Update Step*. In the Prediction Step, we solve the backward SDE system (3.4) numerically, and use the approximate solution as the predicted pdf for the state without using observation information. In the Update Step, we incorporate the observed measurement with the predicted pdf by using the Bayes' theorem. In what follows, we provide our efficient numerical algorithms for the Lévy Backward SDE filter.

3.2 Numerical algorithms

To introduce a recursive discretized numerical scheme, we consider a temporal partition over the time period $[0, T]$ as $\mathcal{R}_t := \{t_n | 0 = t_0 < t_1 < t_2 < \dots < t_{N_T-1} < t_{N_T} = T\}$ and let $\Delta t_n = t_{n+1} - t_n$, $\Delta W_{t_n} = W_{t_{n+1}} - W_{t_n}$. We also assume that the observation measurements are received at time t_n , $n = 1, 2, \dots, N_T$.

Lévy Backward SDE filter framework

Prediction Step. The major task in the Prediction Step is to solve the backward SDEs system (3.4) on time interval $[t_n, t_{n+1}]$, $n = 0, 1, \dots, N_T - 1$, i.e.

$$X_{t_n} = X_{t_{n+1}} - \int_{t_n}^{t_{n+1}} b(X_s) ds - \int_{t_n}^{t_{n+1}} \sigma_s d\overleftarrow{W}_s - \int_0^t \int_{\mathbb{E}} \beta_s(e) \tilde{\mu}(ds, de), \quad (3.6a)$$

$$P_{t_{n+1}} = P_{t_n} - \int_{t_n}^{t_{n+1}} \sum_{i=1}^d \frac{\partial (b_s)_i}{\partial x_i} (X_s) P_s ds - \int_{t_n}^{t_{n+1}} Q_s d\overleftarrow{W}_s - \int_{t_n}^{t_{n+1}} \int_{\mathbb{E}} V_s(e) \tilde{\mu}(ds, de), \quad (3.6b)$$

assuming P_{t_n} is known. It's worthy to point out that in the nonlinear filtering problem, the initial condition P_{t_n} for the above backward SDEs system is chosen to be the conditional pdf of the state S_{t_n} given observation information \mathcal{M}_{t_n} , i.e. $P_{t_n} = p(S_{t_n} | \mathcal{M}_{t_n})$. In what follows, we first derive a temporal discretization scheme for (3.6), and then discuss the spatial discretization later in this subsection.

Consider Eq. (3.6a). We use Euler-Maruyama scheme to discretize the integrals and approximate the solution X_{t_n} by

$$X_{t_n} = X_{t_{n+1}} - b(X_{t_{n+1}}) \Delta t_n - \sigma_{t_{n+1}} \Delta W_{t_n} - \int_{\mathbb{E}} \beta_{t_{n+1}}(e) \tilde{\mu}(\Delta t_n, de) + R_X^n, \quad (3.7)$$

where R_X^n is the approximation error of the equation.

To derive a numerical scheme for the backward SDE in (3.6), we take conditional expectation $E_{n+1}[\cdot]$ on both sides of (3.6b), where $E_{n+1}[\cdot] := E[\cdot | X_{t_{n+1}}]$. It follows from the facts

$$E_{n+1} \left[\int_{t_n}^{t_{n+1}} Q_s d\overleftarrow{W}_s \right] = 0,$$

$$E_{n+1} \left[\int_{t_n}^{t_{n+1}} \int_{\mathbb{E}} V_s(e) \tilde{\mu}(ds, de) \right] = 0$$

and $E_{n+1}[P_{t_{n+1}}] = P_{t_{n+1}}$ that Eq. (3.6b) becomes

$$P_{t_{n+1}} = E_{n+1}[P_{t_n}] - \int_{t_n}^{t_{n+1}} E_{n+1} \left[\sum_{i=1}^d \frac{\partial (b_s)_i}{\partial x_i} (X_s) P_s \right] ds. \tag{3.8}$$

Then, we use the left point formula to discretize the deterministic integral in (3.8) to get

$$P_{t_{n+1}} = E_{n+1}[P_{t_n}] - E_{n+1} \left[\sum_{i=1}^d \frac{\partial (b_{t_n})_i}{\partial x_i} (X_{t_n}) P_{t_n} \right] \Delta t_n + R_p^n, \tag{3.9}$$

where R_p^n is the approximation error of the equation.

Next, we drop the errors terms R_X^n and R_p^n in approximations (3.7) and (3.9), respectively, and obtain the numerical schemes for solving X and P as following

$$X_n = X_{n+1} - b(X_{n+1})\Delta t_n - \sigma_{t_{n+1}} \Delta W_{t_n} - \int_{\mathbb{E}} \beta_{t_{n+1}}(e) \tilde{\mu}(\Delta t_n, de), \tag{3.10a}$$

$$\tilde{P}_{n+1} = E_{n+1}[P_n] - E_{n+1} \left[\sum_{i=1}^d \frac{\partial (b_{t_n})_i}{\partial x_i} (X_n) P_n \right] \Delta t_n, \tag{3.10b}$$

where X_n is an approximation for solution X_{t_n} and \tilde{P}_{n+1} is an approximation for solution $P_{t_{n+1}}$. From the equivalence relation (3.5), we know that the numerical solution \tilde{P}_{n+1} is an approximation for conditional pdf of the state S_t at time instant t_{n+1} before receiving measurement data, i.e. $\tilde{P}_{n+1} \approx p(S_{n+1} | \mathcal{M}_{t_n})$.

Remark 3.1. Although the solution of the backward SDE system (3.6) is a quadruplet (X, P, Q, V) , we do not need numerical approximations for solutions Q and V in neither the nonlinear filtering applications, nor the numerical scheme (3.10). Therefore, in this approach, we do not discuss numerical schemes for Q and V .

Update Step. To derive an approximation for the conditional pdf $p(S_{t_{n+1}} | \mathcal{M}_{t_{n+1}})$ and to incorporate the measurement data at time t_{n+1} , we apply Bayes' theorem to combine the estimate pdf \tilde{P}_{n+1} obtained in the Prediction Step with the data $M_{t_{n+1}}$. Specifically, we let

$$\Pi_{n+1}(x) := \frac{\tilde{P}_{n+1}(x) \Psi_{n+1}(x)}{C_{n+1}}, \tag{3.11}$$

where $\Psi_{n+1}(x) := \exp(-\frac{1}{2R} \|M_{t_{n+1}} - \psi(x)\|^2)$ is proportional to the Gaussian likelihood function, C_{n+1} is a normalization factor. Apparently, Π_{n+1} is an approximation for the conditional pdf of the state given observation information $M_{t_{n+1}}$, i.e. $\Pi_{n+1} \approx p(S_{t_{n+1}} | \mathcal{M}_{t_{n+1}})$, and we let $P_{n+1} = \Pi_{n+1}$.

With the Prediction Step and the Update Step introduced above, we establish the basic framework of our Lévy Backward SDE filter: At each recursive time stage $t_n \rightarrow t_{n+1}$, $n = 0, 1, 2, \dots, N_T - 1$, we let $P_n = \Pi_n$ be the initial condition of the time inverse backward SDE system and use numerical scheme (3.10) to calculate predicted state pdf \tilde{P}_{n+1} . Then, we update the state pdf through (3.11) to get the approximate conditional pdf Π_{n+1} for the state $S_{t_{n+1}}$.

Adaptive meshfree approximations

The numerical scheme (3.10) for the backward SDE system (3.6) can be considered as a temporal discretization scheme. In order to provide an effective numerical approximation for the solution $P_{t_{n+1}}$ as a function of the random variable $X_{t_{n+1}}$, we need to approximate the conditional expectation $E_{n+1}[\cdot]$, which is a functional of $X_{t_{n+1}}$. This could also be considered as a spatial discretization method. Since $X_{t_{n+1}}$ is a continuous random variable in the state space \mathbb{R}^d , it's impossible to approximate $\mathbb{E}_{n+1}[\cdot]$ on all possible values of $X_{t_{n+1}}$ and a representation of $X_{t_{n+1}}$ is required. In this work, we choose a set of space points, denoted by $\mathcal{D}_{n+1} := \{x_{n+1}^1, \dots, x_{n+1}^N\} \in \mathbb{R}^d$, to be a representation of $X_{t_{n+1}}$, and the corresponding conditional expectation values, i.e. $\{E[\cdot | X_{t_{n+1}} = x_{n+1}^i]\}_{x_{n+1}^i \in \mathcal{D}_{n+1}}$, is the approximation of the conditional expectation $E_{n+1}[\cdot]$. Although the standard tensor product mesh is a straightforward option for the representation of the random variable $X_{t_{n+1}}$, the numerical approximation for $E_{n+1}[\cdot]$ over a tensor product mesh is not feasible for two reasons. First of all, the tensor product mesh suffers the so-called "curse of dimensionality" – the number of grid points increases exponentially as the dimension d increases. Secondly, the pdf of the random variable $X_{t_{n+1}}$ has unbounded support and the domain of the tensor product mesh needs to be sufficiently large to cover the true target state.

To address the aforementioned difficulties of tensor product mesh, we use a stochastic meshfree construction of \mathcal{D}_{n+1} , which could also be considered as an adaptive space points generation method. To proceed, we first generate a set of N random samples, denoted by $\{\xi^i\}_{i=1}^N$, from the pdf Pr of the initial state S_0 . Apparently, the random samples $\{\xi^i\}_{i=1}^N$ are more concentrated in the high density region of Pr and we let $\mathcal{D}_0 := \{x_0^i\} = \{\xi^i\}$, i.e. $x_0^i = \xi^i$ for $i = 1, 2, \dots, N$. For time step $n = 0, 1, 2, \dots, N_T - 1$, we propagate space points $\{x_n^i\}$ to $\{x_{n+1}^i\}$ through the state dynamic (2.1), i.e.

$$x_{n+1}^i = x_n^i + b(x_n^i)\Delta t_n + \sigma_{t_n} w_{t_n}^i + L_{t_n}^i(\beta, \tilde{\mu}), \quad i = 1, 2, \dots, N, \quad (3.12)$$

to get our space points set $\mathcal{D}_{n+1} := \{x_{n+1}^i\}$ at time stage t_{n+1} , where $w_{t_n}^i$ is the i -th random sample according to the normal distribution $N(0, \Delta t_n)$, and we denote $L_{t_n}^i(\beta, \tilde{\mu})$ as the numerical approximation for the Lévy term $\int_{\mathbb{E}} \beta_{t_n}(e) \tilde{\mu}(\Delta t_n, de)$ corresponding to the i -th sample. There are different numerical approximation schemes for $L_{t_n}^i(\beta, \tilde{\mu})$ based on the choice of Lévy characteristic function in the problem. The discussion of numerical approximations for Lévy processes is out of scope of this paper and we refer to [30, 48, 61] for details. We can see from our construction of space points that the points in \mathcal{D}_{n+1} , $n = 0, 1, 2, \dots, N_T - 1$, move dynamically according to the state model (2.1) in a stochastic manner and the points are more concentrated in the high probability density region of the state $S_{t_{n+1}}$.

With a set of points \mathcal{D}_{n+1} , we approximate the conditional expectation $E_{n+1}[\cdot]$ on each point in \mathcal{D}_{n+1} , i.e. compute $\{E_{n+1}^{x^i}[\cdot]\}_{i=1}^N$, where $E_{n+1}^{x^i}[\cdot] := E[\cdot | X_{t_{n+1}} = x_{n+1}^i]$ is the value of conditional expectation on the space point x_{n+1}^i . In this research, we use the Monte

Carlo method to approximate conditional expectations [1, 42]. To be specific, for each given space point x_{n+1}^i , we approximate the expectation $E_{n+1}^{x_i} [P_n]$ on the right hand side of (3.10b) as

$$E_{n+1}^{x_i} [P_n] \approx \frac{1}{M} \sum_{m=1}^M \hat{P}_n(x_n^{i,m}), \tag{3.13}$$

where M is the number of samples in the Monte Carlo simulation, and the space point $x_n^{i,m}$ is obtained by solving the forward SDE in the backward SDE system, which is given by Eq. (3.10a) as $x_n^{i,m} = x_{n+1}^i - b(x_{n+1}^i)\Delta t_n + \sigma_{t_{n+1}} \bar{w}_{t_n}^m - \bar{L}_{t_{n+1}}^m(\beta, \tilde{\mu})$, where $\bar{w}_{t_n}^m$ is the m -th sample of the distribution $N(0, \Delta t_n)$ independent from $\{w_{t_n}^i\}$ and $\bar{L}_{t_{n+1}}^m$ is the m -th sample of the numerical approximation for $\int_{\mathbb{E}} \beta_{t_{n+1}}(e) \tilde{\mu}(\Delta t_n, de)$, which is also independent from $\{L_{t_n}^i\}$. The approximation term $\hat{P}_n(x_n^{i,m})$ in Eq. (3.13) is an interpolatory approximation of P_n at $x_n^{i,m}$ based on values $\{P_n(x_n^i)\}_{x_n^i \in \mathcal{D}_n}$ with the scheme

$$\hat{P}_n(x_n^{i,m}) = \sum_{x_n^i \in \mathcal{D}_n} P_n(x_n^i) \Gamma^i(x_n^{i,m}),$$

where $\{\Gamma^i\}_{i=1}^N$ is a set of basis functions. The Monte Carlo approximation for the expectation $E_{n+1} \left[\sum_{i=1}^d \frac{\partial (b_{t_n})_i}{\partial x_i} (X_n) P_n \right]$ in (3.10b) is followed directly by a scheme similar to (3.13).

In order to approximate \hat{P}_n given values of P_n on randomly generated meshfree space points \mathcal{D}_n , an effective interpolation method is needed. However, the standard polynomial interpolation methods are not applicable to approximate \hat{P}_n due to uncontrollable approximation errors [57]. To overcome this challenge, we use *radial basis approximation* to construct the interpolant \hat{P}_n . Specifically, in this work we apply Shepard’s method [20] as our radial basis approximation to compute $\hat{P}_n(x)$ for any point $x \in \mathbb{R}^d$. The Shepard’s method is also known as the “Inverse Distance Weighting” method. It uses the weighted average of values $\{P_n(x_n^i)\}_{x_n^i \in \mathcal{D}_n}$ based on the distance between x and $\{x_n^i\}$ to construct the interpolant. For a given space point $x \in \mathbb{R}^d$, we reorder the points in \mathcal{D}_n by the distance to x from short to long to get a sequence $\{x_n(j)\}_{j=1}^N$ such that $\|x_n(j) - x\| \leq \|x_n(k) - x\|$ if $j < k$, where $\|\cdot\|$ is the Euclidean norm in \mathbb{R}^d . Then, we choose a proper integer $J \leq N$ and use the weighted average of the first J values in $\{P_n(x_n(j))\}_{j=1}^N$ to approximate $\hat{P}_n(x)$, i.e.

$$\hat{P}_n(x) = \sum_{j=1}^J P_n(x_n^j) \bar{\Gamma}^j(x),$$

where $\bar{\Gamma}^j(x)$ is the inverse distance weight and is defined by

$$\bar{\Gamma}^j(x) := \frac{\|x_n(j) - x\|}{\sum_{j=1}^J \|x_n(j) - x\|}, \quad x_n(j) \in \{x_n(j)\}_{j=1}^N.$$

Note that $\sum_{j=1}^J \bar{\Gamma}^j(x) = 1$. As a result of our stochastic space points generation method and the meshfree approximation in Monte Carlo simulations, we obtain numerical approximations for conditional expectations, and therefore numerical approximations for the state pdf on adaptively selected meshfree space points.

It is worthy to point out that there are essential differences between the adaptive meshfree approximation in the Lévy Backward SDE filter and the empirical approximation in the Monte Carlo type approaches. Although both the stochastic space points generation in the Lévy Backward SDE filter and the traditional Monte Carlo method generate random samples from a given probability distribution, the Lévy Backward SDE filter could provide more accurate approximation for the solution of nonlinear filtering problems with fewer sample points for the following reasons. First of all, the Lévy Backward SDE filter approximates the value of the state pdf, which is the solution P_t in the backward SDE system, at each space point instead of using the number of samples to describe an empirical distribution in the Monte Carlo type approach. In this way, the Lévy Backward SDE filter requires much fewer space points compare to the sample-size of the Monte Carlo method. Secondly, in the Lévy Backward SDE filter we use meshfree interpolation to construct a smooth approximation of the state pdf. This is unlike the empirical distribution obtained in the Monte Carlo method, which is equivalent to a piece-wise constant approximation. Therefore the approximation that we obtain in the Lévy Backward SDE filter is smoother and more accurate than the empirical distribution simulation in the Monte Carlo method. In addition, for nonlinear filtering problems of jump diffusion processes, the state distribution is more likely to have heavy tails due to the nonlocal behavior of the state process. In this case, people need sufficient large number of samples in the Monte Carlo method to describe the heavy tails in a relatively large region. On the other hand, the adaptive meshfree approximation in the Lévy Backward SDE filter would provide smooth and more accurate tail distributions with much fewer space points in the tails.

To demonstrate the performance of the adaptive meshfree approximation in describing heavy-tailed distributions, we approximate an α stable distribution by using adaptive meshfree approximation and classic Monte Carlo method. To proceed, we define an α stable distribution $\Phi(\alpha, \gamma, \beta, \delta)$ of a random variable X with characteristic function given by

$$E[\exp(itX)] = \exp\left(-\gamma^\alpha |t|^\alpha \left[1 + i\beta \text{sign}(t) \tan \frac{\pi\alpha}{2} ((\gamma|t|)^{1-\alpha} - 1)\right] + i\delta t\right), \quad \alpha \neq 1,$$

where the parameters are chosen to be $\alpha = 0.75$, $\gamma = 1$, $\beta = 0$ and $\delta = 0$. In Fig. 1, we compare the adaptive meshfree approximation with the empirical distribution of Monte Carlo method in approximating the α stable distribution Φ . The original α stable distribution is represented by the red curve in each subplot. For the adaptive meshfree approximation, we approximate the distribution with 50 space points and use the black curve and black cross marks to describe the approximated distribution and the adaptive space points, respectively. The histogram in each subplot is the empirical distribution

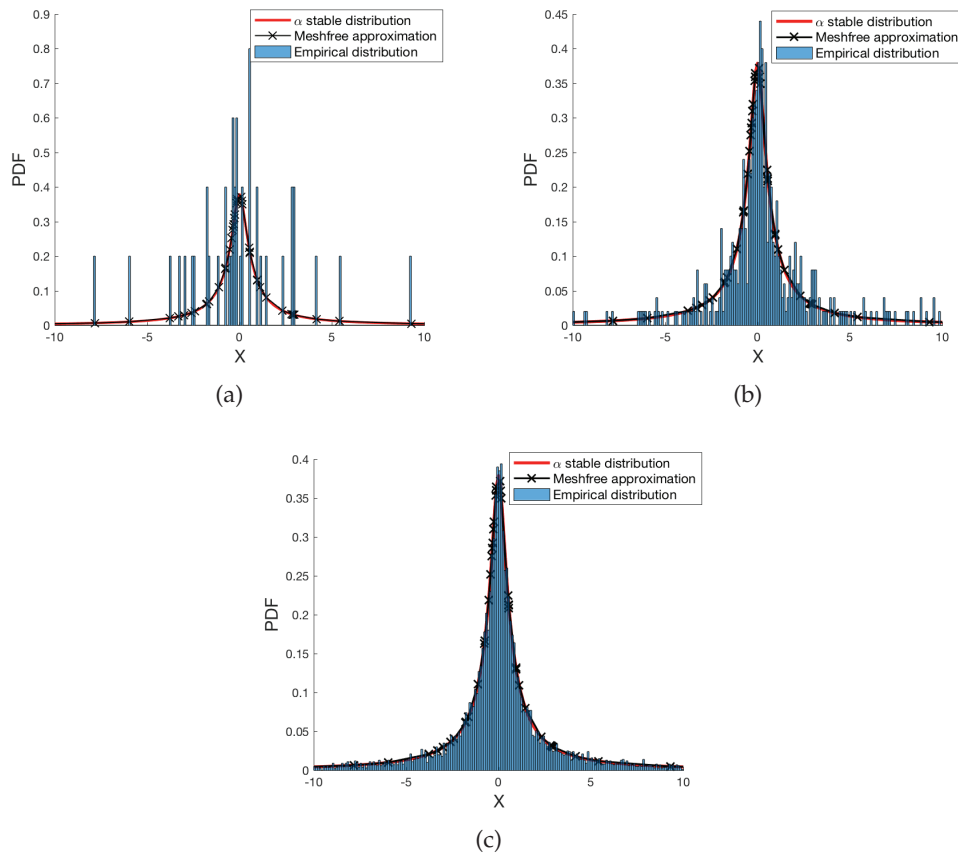


Figure 1: Comparison of adaptive meshfree approximation with the empirical distribution of Monte Carlo method in approximating the α stable distribution Φ : (a) empirical distribution with 50 samples; (b) empirical distribution with 500 samples; (c) empirical distribution with 10,000 samples. The adaptive meshfree approximation uses 50 space points.

of Monte Carlo method and the number of samples increases from 50 to 500, and then to 10,000 in subplots (a), (b) and (c), respectively. Since the adaptive space points in the meshfree approximation are chosen randomly according to the target distribution, in this demonstration we use the same 50 samples as the Monte Carlo simulation presented in subplot (a). We can see from the figure that as the sample-size increases, the Monte Carlo simulation becomes smoother and more accurate, and the empirical distribution with 10,000 samples is comparable to adaptive meshfree approximation with 50 space points. Moreover, we can see from subplots (a) and (b) that with smaller sample-size in Monte Carlo simulations, very few samples lie in the tails of distribution which makes the tail description very unreliable. On the other hand, although only 50 space points are used to approximate the distribution in the adaptive meshfree approximation and limited space points lie in the tail regions, the interpolatory approximation still makes the tail distribution smooth and accurate.

Markov Chain Monte Carlo resampling

In the stochastic space points generation method, the space points \mathcal{D}_n move according to the state model (2.1), which is a diffusion process. Therefore, the space points cloud diverge for long term simulations and less space points are located in the high probability density region of the state pdf after several simulation steps. This would make the space points very sparse in the state space and the probability density tends to concentrate on a few points, which dramatically reduces the meshfree interpolation accuracy.

To avoid the divergence problem of stochastic space points, inspired by the resampling procedure in the particle filter method [1,24], we introduce a Markov Chain Monte Carlo (MCMC) method based resampling step. MCMC method is a class of algorithms for sampling from a probability distribution based on constructing an aperiodic and irreducible Markov chain that has the desired distribution as its equilibrium distribution [58]. The state of the chain after sufficient large number of simulation steps can be treated as an independent sample of the desired distribution. It is well known that MCMC method is an effective sampling method for complicated distributions in high-dimensional spaces. Taking the advantage of MCMC sampling, in this work we combine the solution of the backward SDE system with the observation data and apply the MCMC method to remove the stochastic space points away from statistically insignificant regions of the state pdf.

Specifically, in the Lévy Backward SDE filter framework, when we get the approximate solution Π_n on \mathcal{D}_n ($n = 1, 2, \dots, N_T - 1$) and initiate the recursive stage $t_n \rightarrow t_{n+1}$ by setting $P_n = \Pi_n$, instead of propagating \mathcal{D}_n to \mathcal{D}_{n+1} directly to construct the space point set for Π_{n+1} , we use MCMC sampling to create an observation informed intermediate point set $\mathcal{D}_{n+\frac{1}{2}}$, and then propagate $\mathcal{D}_{n+\frac{1}{2}}$ through the state model to get \mathcal{D}_{n+1} . To create the point set $\mathcal{D}_{n+\frac{1}{2}}$, we generate a Markov chain for each space point $x_n^i \in \mathcal{D}_n$ to move it away from the statistically insignificant regions. Since the interpolatory approximation \hat{P}_n in the Lévy Backward SDE filter is a point-wise numerical approximation of the state pdf, we use \hat{P}_n as the stationary distribution of the Markov chain for each space point. It is worthy to point out that the Markov chain for each space point is based on the global approximation of the state pdf which is also incorporated with the observation information up to time level t_n . In this way, our MCMC resampling procedure is incorporated with the observation data, the data informed space points $\mathcal{D}_{n+\frac{1}{2}}$ uses the observation information sufficiently and construct adaptive space points more effectively. There are many sampling algorithms for the MCMC method and we use Metropolis-Hastings algorithm [28] as an example to demonstrate our MCMC resampling method.

Algorithm summary

We summarize the recursive algorithms for Lévy Backward SDE filter for Jump Diffusion Processes as follows:

Define a pdf Pr as the initial guess for the state S_0 by setting $\Pi_0 = Pr$, generate space points $\mathcal{D}_0 \sim \Pi_0$, and choose parameters N as number of space points, M as number of

Monte Carlo samples, J as number of meshfree interpolation points, L as MCMC iteration number. The space point set \mathcal{D}_1 is propagated from \mathcal{D}_0 directly without resampling since there's no observation information at time $t=0$.

For the recursive stage $t_{n-1} \rightarrow t_n$, $n=1, 2, \dots, N_T$, we implement

- *Prediction Step*: Solve the backward SDEs system with scheme (3.10) to get the predict state pdf \hat{P}_n ;
- *Update Step*: Update the state pdf with scheme (3.11) to get Π_n ;
- *Adaptive Meshfree Approximation*: Let $P_n = \Pi_n$ and expand P_n on \mathcal{D}_n to \hat{P}_n through meshfree interpolation;
- *Resampling Step*: Use MCMC resampling procedure to construct observation informed space points $\mathcal{D}_{n+\frac{1}{2}}$. Then propagate $\mathcal{D}_{n+\frac{1}{2}}$ to \mathcal{D}_{n+1} through the state model (2.1).

3.3 Numerical examples

We present two numerical examples to demonstrate the performance of the Lévy Backward SDE filter in solving the nonlinear filtering problem (2.1). Both examples that we discuss in this subsection are benchmark problems with Lévy noise added to the state equation as indicated in (2.1). In the first example, we estimate the trajectory of a target moves along a one dimensional periodic potential curve. In Example 2, we solve a two dimensional bearing-only tracking problem [60]. We examine the performance of the Lévy Backward SDE filter by comparing our Lévy Backward SDE filter with auxiliary particle filter [47], which is one of the most widely accepted nonlinear filtering methods by practitioners. The numerical experiments are carried out on an *Intel Core i5 2.7 GHz* CPU.

Example 1

We first consider a periodic energy potential, denoted by U with $U = -\frac{10}{3} \cos\left(\frac{3x}{10}\right)$, and assume that there's a target moves along the potential curve U . The curve is plotted in Fig. 2. We can see from this figure that the potential U has some wells. If a particle moves on this potential curve U and the potential difference forms the force to move the particle, the trajectory of the particle satisfies the ordinary differential equation $dS_t = \sin\left(\frac{3S_t}{10}\right)dt$. Without the external perturbation, the particle wanders around the bottom of one of the wells. In this example, we assume that the particle is influenced by a Lévy noise, which is the external perturbation to excite the particle from one potential well to another. The derived state process of the nonlinear filtering problem is given as follows,

$$dS_t = \sin\left(\frac{3S_t}{10}\right)dt + 4dW_t + \int_{\mathbb{E}} 10e\tilde{\mu}(dt, de),$$

where W_t is a standard Brownian motion and $\tilde{\mu}$ is the compensated Poisson measure. The dynamic system of this example is similar to the classic double well potential problem.

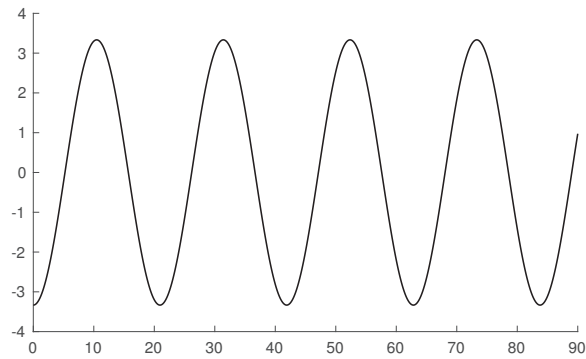


Figure 2: One dimensional periodic potential curve U

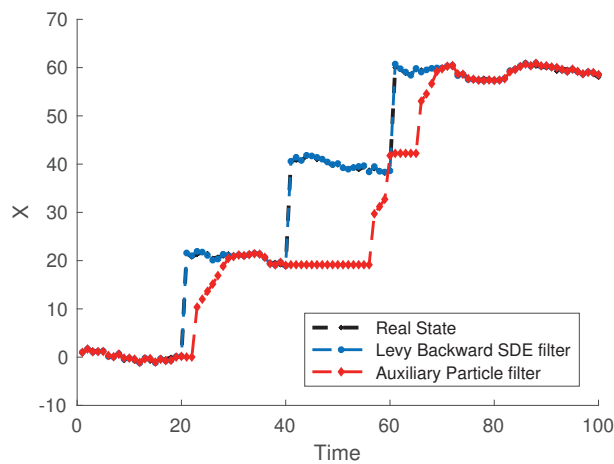


Figure 3: Comparison for one dimensional potential tracking.

But instead of two wells, it has multiple wells that allows a particle switch between. The transitions of the particle and the evolution of the system is observed via data given by the following observation process

$$M_{t_n} = S_{t_n} + R\dot{B}_{t_n}, \quad n = 1 \cdots, N_t$$

where B_t is a standard Brownian motion independent from W_t . In this problem, we track the particle as the target from time $t = 0$ to $t = 2$ with uniform time step $\Delta t = 0.02$, i.e. $N_t = 100$, and use a compound Poisson process to generate jumps in the state. For the observation, we choose $R = 0.1$. In Fig. 3, we compare the performance of our Lévy Backward SDE filter with Auxiliary Particle filter (APF). In the Lévy Backward SDE filter, we use 200 space points and in the APF, we use 800 particles. We can see from Fig. 3 that the Lévy Backward SDE filter could capture the change of state very accurately while

Table 1: Example 1: Efficiency comparison.

Numerical methods	CPU time (seconds)	err_G
Lévy Backward SDE filter (200 space points)	9.01	0.3302
Auxiliary particle filter (400 particles)	10.97	34.1392
Auxiliary particle filter (800 particles)	23.07	27.5474
Auxiliary particle filter (1600 particles)	50.49	16.7304
Auxiliary particle filter (3200 particles)	95.40	10.8136

the auxiliary particle filter takes many more steps to capture the changes. To further compare the performance of the Lévy Backward SDE filter and the APF, we repeat the above experiment 50 times and compute overall root mean square errors (RMSE). In Table 1, we present the RMSE of the Lévy Backward SDE filter with 200 space points as well as APF with 400, 800, 1600 and 3200 particles. From this table, we can see that with 200 space points, the Lévy Backward SDE filter takes 9 seconds to track the target trajectory and the RMSE is 0.3302. Although the APF with 400 particles could track the target with similar CPU time to the Lévy Backward SDE filter, the RMSE of APF is 34.1392, which is much higher than RMSE of the Lévy Backward SDE filter. When using more and more particles in APF, the RMSE reduces and the computing cost increases significantly. However, even using 3200 particles to track the target in the APF, which consumes more than 10 times of CPU time of the Lévy Backward SDE filter, the RMSE of APF is still much higher than the Lévy Backward SDE filter.

Example 2

In this example, we solve a bearing-range tracking problem, in which a target is moving on a two dimensional plane with a near constant velocity and the state of the target is perturbed by Lévy noise. The state equation is given as follows

$$dS_t = AS_t dt + \sigma dW_t + \int_{\mathbb{E}} \beta(e) \tilde{\mu}(dt, de), \quad (3.14)$$

where $S_t = (X_t, Y_t, \dot{X}_t, \dot{Y}_t)^T$ is a 4 dimensional vector, (X_t, \dot{X}_t) and (Y_t, \dot{Y}_t) are the position and velocity of the target at time t corresponding to X and Y co-ordinate, respectively. W_t is a 4 dimensional standard Brownian motion and $\tilde{\mu}$ is the compensated Poisson measure. The matrices A , σ and β are given by

$$A = \begin{pmatrix} I_2 & I_2 \\ 0 & I_2 \end{pmatrix}, \quad \sigma = \text{diag}(0.1, 0.1, 0.05, 0.05), \quad \beta(e) = e(2, 2, 0.2, 0.2)^T,$$

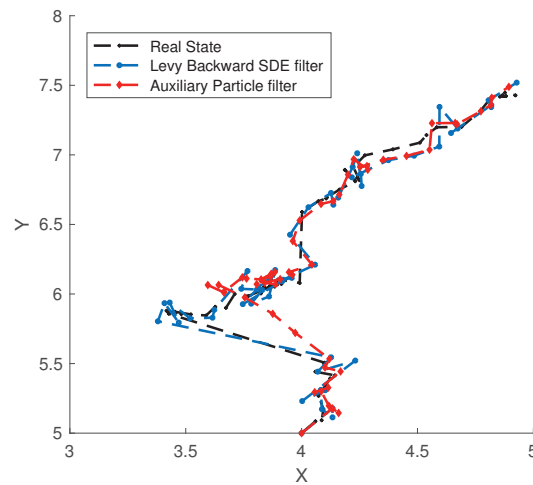


Figure 4: Comparison of target tracking for $\alpha = 1$.

where I_2 is 2×2 identity matrix. The target is observed by a bearing-range director located at (X_{obs}, Y_{obs}) , i.e. the measurement M_t is given by

$$Y_t = \left(\begin{array}{c} \arctan\left(\frac{Y_t - Y_{obs}}{X_t - X_{obs}}\right) \\ \sqrt{(X_{t_n} - X_{obs})^2 + (Y_t - Y_{obs})^2} \end{array} \right) + R\dot{B}_t, \quad n = 1, 2, \dots, N_t,$$

where B_t is a standard two dimensional Brownian motion independent from W_t and $R := \text{diag}(0.01, 0.1)$ is a two dimensional matrix.

In this example, we use α -stable process to generate jumps in the noise and track the target for the time period $t = 0$ to $t = 2$ with uniform time step $\Delta t = 0.04$, i.e. $N_t = 50$. In Fig. 4, we compare the tracking performance of the Lévy Backward SDE filter with APF and choose $\alpha = 1$. For the Lévy Backward SDE filter, we use 1,500 space points. For the APF, we use 6,000 particles to describe the conditional pdf of the target state. The black curve is the real target trajectory in the XY-plane. We can see there are several jumps in the trajectory. The red curve in the figure is the estimate trajectory obtained by the APF, and the blue curve is the estimate trajectory obtained by the Lévy Backward SDE filter. We can see from Fig. 4 that the Lévy Backward SDE filter could capture the change of state much faster than the auxiliary particle filter.

To better demonstrate the effectiveness of our Lévy Backward SDE filter in tracking a target with jumps, we choose $\alpha = 0.5$ for the α -state Lévy process and solve the same bearing-only tracking problem. Since we choose a small α in this experiment, the α -stable process has thicker tail in the distribution, which results higher probability for big jumps. In Fig. 5, we present the tracking performance for the case $\alpha = 0.5$ and we can observe two significant jumps in the target trajectory. From this figure, when choosing a smaller α , we can see that estimate of APF lost target for several steps and the irregular

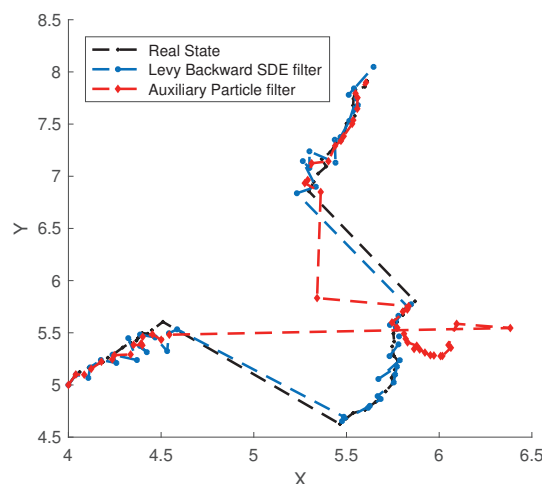


Figure 5: Comparison of target tracking for $\alpha=0.5$.

behavior when trying to catch up with the target is typically caused by its degeneracy problem. However, our Lévy Backward SDE filter still provides reliable estimate for the target state even with two big jumps.

4 Application of Lévy Backward SDE filter in material science

In this section, we discuss the application of the Lévy Backward SDE filter in nano-phase material sciences, in the context of development of algorithmic control over single atom processes that will enable large-scale and automated manipulation and synthesis of single atoms and molecules [31, 35, 44]. The majority of studies on atomic-scale manipulation to date have been carried out largely in the open-loop regime, where there is no direct feedback between excitation (electric field, electric current, direct chemical interaction) and the excited object, or in rare but intriguing cases – with human control [25]. However, automated control methods will be needed to achieve two major goals: dramatically increase the manipulation speed, well beyond the human capacity, and equally importantly to enable deterministic selectivity over the reaction steps – a specific kind of reaction, specific direction of motion etc. [39].

The automated synthesis algorithm can be generally described by three procedures: the stochastic optimization procedure, which designs the optimal material potential surface; the tracking procedure, which tracks the movement of a target atom based on the designed material potential surface and the observation data received from scanning probe (or related electron microscopy); the optimal control procedure, which controls the material condition to minimize the cost of synthesis based on the estimation of the molecule state. In this atomic level material synthesis, the role of the Lévy Backward SDE filter

is to track the atom trajectory in the single-atom manipulation and single-molecule reaction, both of which are registered as abrupt events in the relevant observables, such as tunneling current and/or interaction force between the manipulating probe and the manipulated entity.

To examine the applicability of our algorithm in the aforementioned framework, we use Lévy Backward SDE filter to simulate tracking atoms moving on the potential surface of some well-known materials. Single atoms on the surface experience atomic-scale interaction potential, due to preferential atomic bonding toward maximally coordinated sites [44]. In this work, we approximate the corresponding potential using a simple triangular lattice of potential energy wells as a sum of sinusoids, where the amplitude was calibrated to match the experimentally observed diffusion potential. This would coarsely resemble the potential of the 111-terminated noble-metal surface (such as Ag(111), Au(111) and others) toward interaction with a single atom. The depth of the wells was chosen to mimic a recent work on atomic motion by Giessibl et al. [56]. In what follows, we present the performance of the Lévy Backward SDE filter in tracking the atom trajectories on approximated potential surfaces.

Experiment 1

We first depict our approximated potential in Fig. 6 (a), which represents the potential energy landscape for diffusion of atomic scale species (such as atoms or molecules), and denote this potential by F_1 . From this image, we can see that there are several deep wells. Once the atom falls in the bottom of one well, it will be trapped in this well unless some excitation occurs. In Fig. 6 (b), we present the 2D plan view of the energy potential F_1 . The dark blue disks represent low potential regions which are bottoms of wells and the connected light blue region represent high potential area.

The position of an atom would follow the force caused by the potential of material surface presented in Fig. 6. Different from the synthetic examples we presented in Section 3.3, in which we have analytic drift function b in the nonlinear filtering problem (2.1), the drift term of the state equation in this experiment is calculated by the simulated physical potential force F_1 . As a result, the state in this experiment is given by

$$dS_t = -\nabla \hat{F}_1(S_t)dt + \sigma dW_t + \int_{\mathbb{E}} \beta(e) \tilde{\mu}(dt, de), \quad (4.1)$$

where $S_t = (X_t, Y_t)^T$ is a two dimensional vector describing the position of the atom in the XY plane. In this experiment, the function $\nabla \hat{F}_1$ is the drift term of the state equation given by the force caused by the potential surface. Since the energy potential F_1 is obtained from simulation on a given grid mesh, it is different from the synthetic examples in Section 3.3, in which the drift term b is an explicit function. In order to derive the gradient of F_1 , we use polynomial approximation to construct a smooth surface of F_1 and then calculate the gradient of the surface, i.e. $\nabla \hat{F}_1$ on the XY plane. To simulate the trajectory, we choose $\sigma = \text{diag}(0.1, 0.1)$, $\beta(e) = e(10, 10)^T$ and use a compound Poisson process to generate jumps in the state. In Fig. 7, we present a sample atom trajectory of the simulated state S_t over

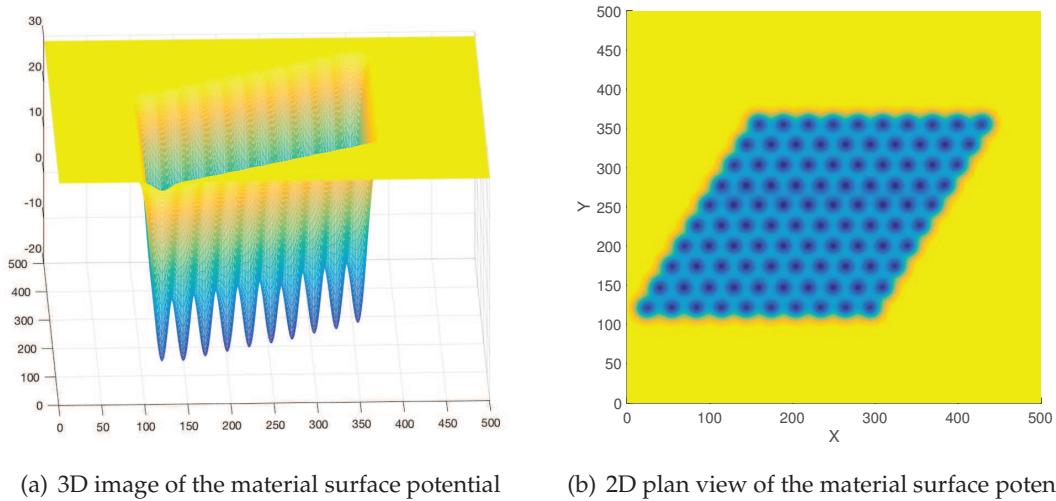
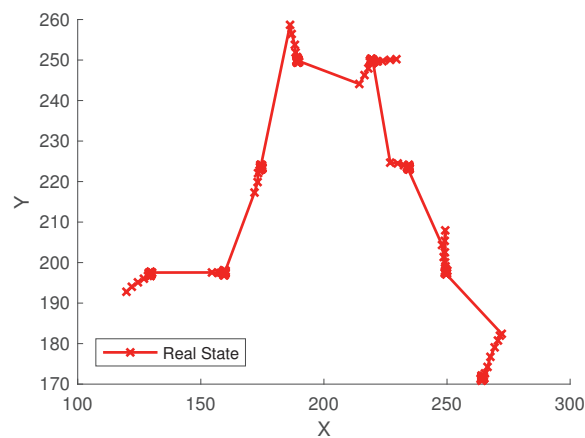
Figure 6: The material surface potential F_1 .

Figure 7: A sample atomic trajectory.

time interval $[0,10]$ with uniform time step $\Delta t = 0.02$, i.e. $N_t = 500$. From this figure, we can see that the atom has some jumps with several stable states. Once the atom arrives at a stable state, it remains at this state until excited by some force and jumps to another random state. To demonstrate the details of the atom trajectory, we plot the atom position in each direction in Fig. 8. From this figure, we can see that there are 8 stable positions, in which the atom remains in the similar $X-Y$ position for a period, and the atom has some random smaller scale jumps which is not enough to escape from the bottom of the potential well and is dragged back by the potential force. In Fig. 9, we plot the sample atom trajectory depicted in Fig. 7 in the energy potential which has been shown in Fig. 6

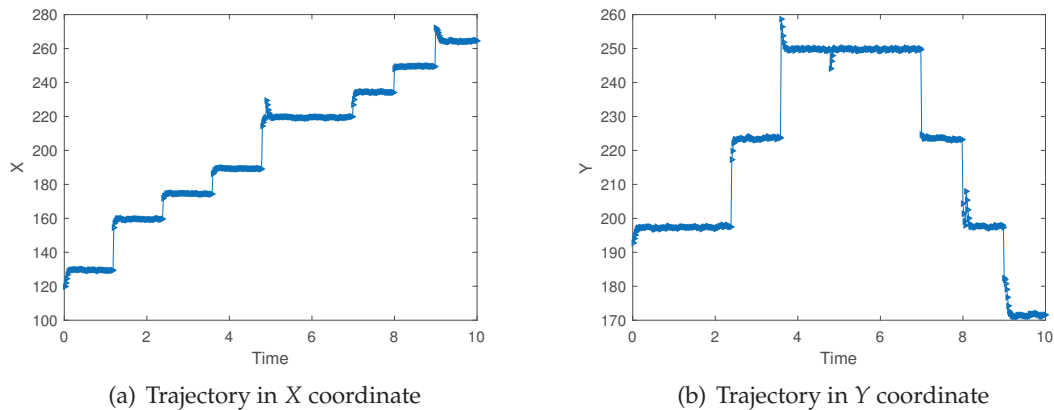
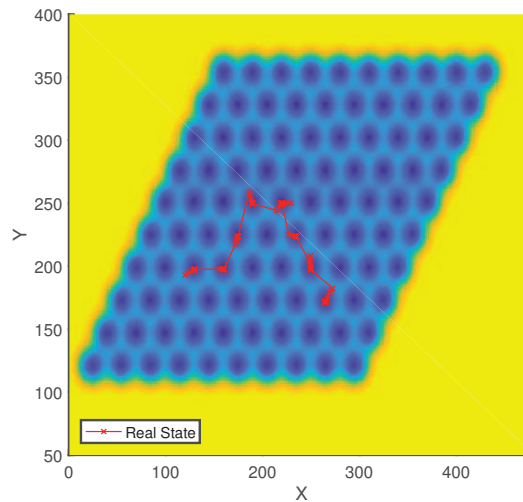


Figure 8: Sample trajectory in X, Y directions.

Figure 9: The sample atom trajectory with respect to the energy potential F_1 .

(b). From this figure, we can see that all the stable positions of the atom trajectory is around the bottom of a potential well.

In order to track this atom, we use tunnel electron microscope to receive observation of the atom. In this experiment, we assume that the observation is noise perturbed atom position, i.e.

$$M_{t_n} = S_{t_n} + R\dot{B}_{t_n}, \quad n = 1, 2, \dots, 500,$$

where B_t is a two dimensional Brownian motion independent from W_t and we choose $R = \text{diag}(0.05, 0.05)$. In Fig. 10, we present the performance of our Lévy Backward SDE filter in tracking this atom. The red trajectory in this figure is the real target atom trajectory

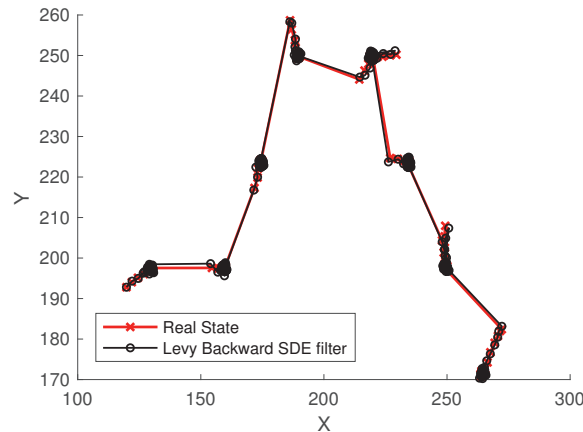
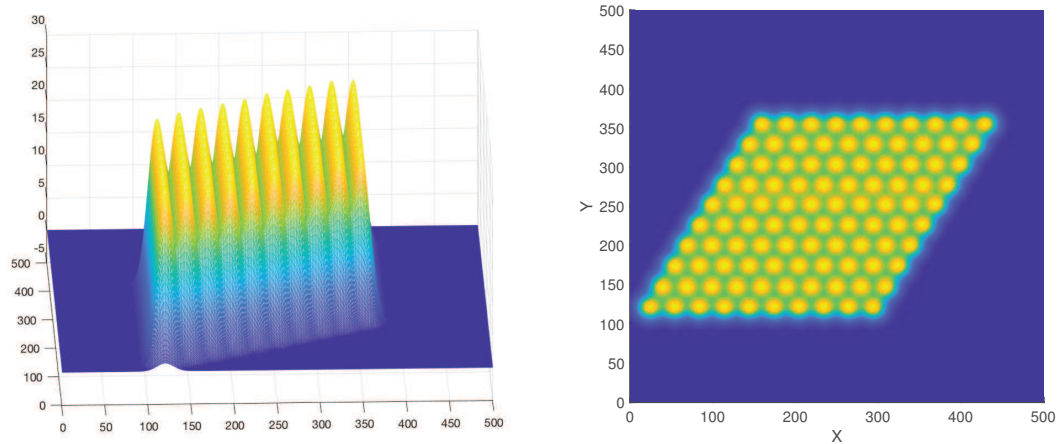


Figure 10: Tracking performance of Lévy Backward SDE filter in the real energy potential problem.

and the black path marked by circles is the estimated target obtained by using the Lévy Backward SDE filter. We can see from this figure that the Lévy Backward SDE filter could track the atom trajectory accurately.

Experiment 2

In this experiment, we approximate a different energy potential surface F_2 and use the Lévy Backward SDE filter to track the atom trajectory based on observation M_{t_n} . In Fig. 11 (a), we present the 3D potential surface that we approximated. On the contrary to energy



(a) 3D image of the material surface potential

(b) 2D plan view of the material surface potential

Figure 11: The material surface potential F_2 .

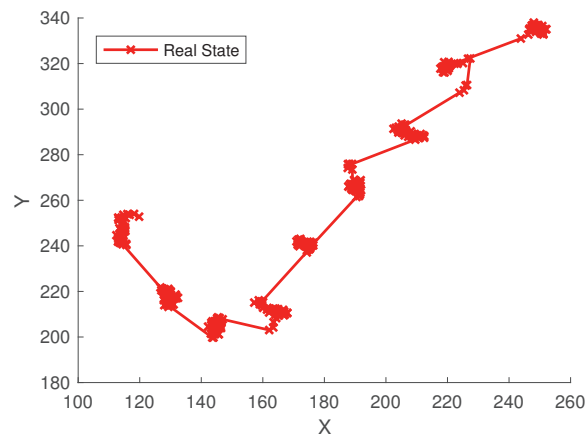


Figure 12: A sample atomic trajectory.

F_1 , we can see there are several peaks for the energy potential. If the atom gets on the peak of potential by any excitation, it would slide down and remains in the bottoms of different peaks which has lower energy with more stable state. In Fig. 11 (b), we plot the 2D plan view image of the energy potential F_2 . The yellow disks represent the energy peaks and the green region has lower energy in which the atom is in stable state. Given the simulated energy potential F_2 , we derive the state equation of the nonlinear filtering problem (2.1) as

$$dS_t = -\nabla \hat{F}_2(S_t)dt + \sigma dW_t + \int_{\mathbb{E}} \beta(e) \tilde{\mu}(dt, de),$$

where $\nabla \hat{F}_2$ is the drift term of the state equation, which is the approximate gradient of the energy potential surface F_2 .

Similar to Experiment 1, we simulate the state equation over time interval $[0, 10]$ with uniform time step $\Delta t = 0.02$, i.e. $N_t = 500$ with $\sigma = \text{diag}(0.1, 0.1)$ and $\beta(e) = e(10, 10)^T$. In Fig. 12, we plot a sample atom trajectory and put this trajectory on the material surface in the XY plane with the 2D plan view of energy potential F_2 in Fig. 13. From Fig. 13, we can see that for most of time, the atom remains stable in the lower energy region. The position of the atom is perturbed by both Gaussian noises, which makes it linger around its current position, and Poisson noise, which causes some jumps. If the atom jumps onto one of the peak, as we can see in the figure, the potential force pushes it down and keep it remain in the bottom of all the peaks.

Finally, we present the performance of our Lévy Backward SDE filter in tracking this atom in Fig. 14. The red trajectory in this figure is also the real target atom trajectory and the black path marked by circles is the estimated target obtained by using the Lévy Backward SDE filter. We can see from this figure that under the energy potential F_2 , the Lévy Backward SDE filter could also track the atom trajectory accurately.

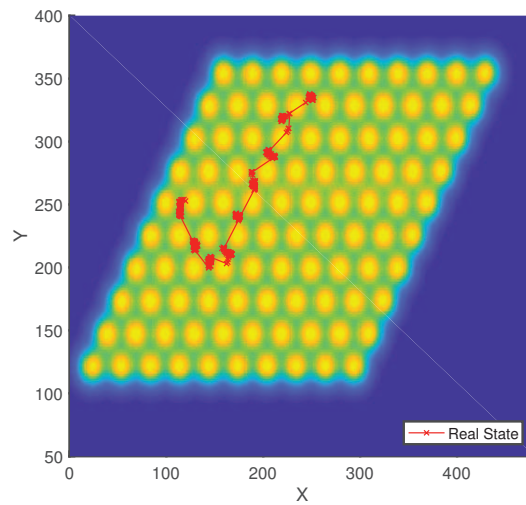


Figure 13: The sample atomic trajectory with respect to the energy potential F_2 .

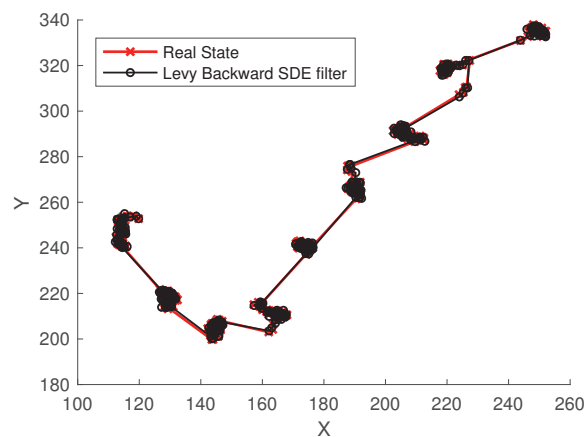


Figure 14: Tracking performance of Lévy Backward SDE filter in the real energy potential problem.

Acknowledgments

This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program through the FASTMath Institute (under Lawrence Livermore National Laboratory Subcontract B626484). The first author acknowledges support by U.S. National Science Foundation under grant number DMS-1720222 and support by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, SciDAC program through the CompFUSE project. The third author also

acknowledges support by the Center for Nanophase Materials Sciences, sponsored by the Division of User Facilities, Basic Energy Sciences, U.S. Department of Energy.

References

- [1] C. Andrieu, A. Doucet, and R. Holenstein. Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(3):269–342, 2010.
- [2] F. Bao and Y. Cao. Adjoint forward backward stochastic differential equations driven by jump diffusion processes and its application to nonlinear filtering problems. *International Journal for Uncertainty Quantification*, in press.
- [3] F. Bao, Y. Cao, and X. Han. Forward backward doubly stochastic differential equations and optimal filtering of diffusion processes, arxiv: 1509.06352. 2016.
- [4] F. Bao, Y. Cao, X. Han, and J. Li. Efficient particle filtering for stochastic korteweg-de vries equations. *Stochastics and Dynamics*, 17(2):1750008, 2017.
- [5] F. Bao, Y. Cao, C. Webster, and G. Zhang. A hybrid sparse-grid approach for nonlinear filtering problems based on adaptive-domain of the Zakai equation approximations. *SIAM/ASA J. Uncertain. Quantif.*, 2(1):784–804, 2014.
- [6] F. Bao, Y. Cao, and W. Zhao. Numerical solutions for forward backward doubly stochastic differential equations and Zakai equations. *Int. J. Uncertain. Quantif.*, 1(4):351–367, 2011.
- [7] F. Bao, Y. Cao, and W. Zhao. A backward doubly stochastic differential equation approach for nonlinear filtering problems. *Communications in Computational Physics*, 23(5):1573–1601, 2018.
- [8] F. Bao and V. Maroulas. Adaptive meshfree backward sde filter. *SIAM Journal on Scientific Computing*, 39(6):A2664–A2683, 2017.
- [9] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics Stochastics Rep.*, 60(1-2):57–83, 1997.
- [10] A. Bensoussan, J. Keppo, and S. Sethi. Optimal consumption and portfolio decisions with partially observed real prices. *Math. Finance*, 19(2):215–236, 2009.
- [11] M. Bolić, P. Djurić, and S Hong. Resampling algorithms and architectures for distributed particle filters. *IEEE Trans. Signal Process.*, 53(7):2442–2450, 2005.
- [12] Z. Chen. Bayesian filtering: From kalman filters to particle filters, and beyond. *Statistics*, pages 1–69, 2003.
- [13] A. Chorin, M. Morzfeld, and X. Tu. A survey of implicit particle filters for data assimilation. In *State-space models*, Stat. Econom. Finance, pages 63–88. Springer, New York, 2013.
- [14] D. Crisan. Particle filters—a theoretical perspective. In *Sequential Monte Carlo methods in practice*, Stat. Eng. Inf. Sci., pages 17–41. Springer, New York, 2001.
- [15] D. Crisan and A. Doucet. A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans. Signal Process.*, 50(3):736–746, 2002.
- [16] P. Dawson, R. Gailis, and A. Meehan. Detecting disease outbreaks using a combined Bayesian network and particle filter approach. *J. Theoret. Biol.*, 370:171–183, 2015.
- [17] M. Djogatović, M. Stanojević, and N. Mladenović. A variable neighborhood search particle filter for bearings-only target tracking. *Comput. Oper. Res.*, 52(part B):192–202, 2014.
- [18] A. Doucet and A. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. http://www.cs.ubc.ca/arnaud/doucet_johansen_tutorialPF.pdf, 2008.
- [19] R. Elliott and T. Siu. Option pricing and filtering with hidden Markov-modulated pure-jump processes. *Appl. Math. Finance*, 20(1):1–25, 2013.

- [20] G. F. Fasshauer. *Meshfree Approximation Methods with MATLAB*. Interdisciplinary Mathematical Sciences (Book 6). World Scientific Publishing Company, 2007.
- [21] R. Frey and T. Schmidt. Pricing and hedging of credit derivatives via the innovations approach to nonlinear filtering. *Finance Stoch.*, 16(1):105–133, 2012.
- [22] F. J. Giessibl. Advances in atomic force microscopy. *Rev. Mod. Phys.*, 75:949, 2003.
- [23] E. Gobet, G. Pagès, H. Pham, and J. Printems. Discretization and simulation of the Zakai equation. *SIAM J. Numer. Anal.*, 44(6):2505–2538 (electronic), 2006.
- [24] N.J Gordon, D.J Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE PROCEEDING-F*, 140(2):107–113, 1993.
- [25] M.F. Green, T. Esat, C. Wagner, P. Leinen, A. Grotsch, F.S. Tautz, and R. Temirov. Patterning a hydrogen-bonded molecular monolayer with a hand-controlled scanning probe microscope. *Beilstein J Nanotechnol*, pages 1926–1932, 2014.
- [26] M. Hairer, A. Stuart, and J. Voss. Signal processing problems on function space: Bayesian formulation, stochastic PDEs and effective MCMC methods. *The Oxford handbook of nonlinear filtering*, pages 833–873, 2011.
- [27] X. Han, J. Li, and D. Xiu. Error analysis for numerical formulation of particle filter. *Discrete Contin. Dyn. Syst. Ser. B*, 20(5):1337–1354, 2015.
- [28] W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97109.
- [29] E.J. Heller, M.F. Crommie, C.P. Lutz, and D.M. Bigler. Scattering and absorption of surface electron waves in quantum corrals. *Nature*, 369:464–466, 2013.
- [30] D. Higham and P. Kloeden. Numerical methods for nonlinear stochastic differential equations with jumps. *Numerische Mathematik*, 101:101119.
- [31] S.W. Hla. Atom-by-atom assembly. *Reports on Progress in Physics*, 77(5):056502, 2014.
- [32] Y. Hu, G. Kallianpur, and J. Xiong. An approximation for the Zakai equation. *Appl. Math. Optim.*, 45(1):23–44, 2002.
- [33] K. Kang, V. Maroulas, I. Schizas, and F. Bao. Improved distributed particle filters for tracking in a wireless sensor network. *Computational Statistics and Data Analysis*, 117(Supplement C):90 – 108, 2018.
- [34] D. Kim, T. Song, and D. Musicki. Highest probability data association for multi-target particle filtering with nonlinear measurements. *IEICE TRANSACTIONS on Communications*, E96-B (1):281–290, 2013.
- [35] Y Kim, K. Motobayashi, T. Frederiksen, H. Ueba, and M. Kawai. Action spectroscopy for single-molecule reactions experiments and theory. *Progress in Surface Science*, 90:85–143, 2015.
- [36] H. Kushner. On the differential equations satisfied by conditional probability densities of Markov processes, with applications. *J. Soc. Indust. Appl. Math. Ser. A Control*, 2:106–119, 1964.
- [37] Y. Lee. An asymmetric information modeling framework for ultra-high frequency transaction data: a nonlinear filtering approach. In *State-space models*, Stat. Econom. Finance, pages 279–309. Springer, New York, 2013.
- [38] M. Little and N. Jones. Signal processing for molecular and cellular biological physics: an emerging field. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 371(1984):20110546, 18, 2013.
- [39] P. Maksymovych, S. V. Kalinin, and A. Gruverman. *Scanning probe microscopy of functional materials nanoscale imaging and spectroscopy*. Springer, 2011.
- [40] Vasileios Maroulas and Andreas Nebenführ. Tracking rapid intracellular movements: a

- Bayesian random set approach. *Ann. Appl. Stat.*, 9(2):926–949, 2015.
- [41] Vasileios Maroulas and Panos Stinis. Improved particle filters for multi-target tracking. *J. Comput. Phys.*, 231(2):602–611, 2012.
- [42] N. Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44:335–341.
- [43] T. Meyer-Brandis and F. Proske. Explicit solution of a non-linear filtering problem for Lévy processes with application to finance. *Appl. Math. Optim.*, 50(2):119–134, 2004.
- [44] N. Morgenstern, N. Lorente, and K.-H Rieder. Controlled manipulation of single atoms and small molecules using the scanning tunnelling microscope. *Phys. Status Solidi B*, 250:16711751, 2013.
- [45] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.
- [46] É. Pardoux and P. Protter. A two-sided stochastic integral and its calculus. *Probab. Theory Related Fields*, 76(1):15–49, 1987.
- [47] M. Pitt and N. Shephard. Filtering via simulation: auxiliary particle filters. *J. Amer. Statist. Assoc.*, 94(446):590–599, 1999.
- [48] E. Platen and N. Bruti-Liberati. *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*. Springer, 2010.
- [49] S. Popa and S. S. Sriitharan. Nonlinear filtering of Itô-Lévy stochastic differential equations with continuous observations. *Commun. Stoch. Anal.*, 3(3):313–330, 2009.
- [50] H. Qiao and J. Duan. Nonlinear filtering of stochastic dynamical systems with lvy noises. *Advances in Applied Probability*, 47(3):902918, 2015.
- [51] Ioannis Sgouralis, Andreas Nebenführ, and Vasileios Maroulas. A Bayesian topological framework for the identification and reconstruction of subcellular motion. *SIAM J. Imaging Sci.*, 10(2):871–899, 2017.
- [52] S. Singh, S. Digumarthy, A. Back, J. Shepard, and Kalra M. Radiation dose reduction for chest ct with non-linear adaptive filters. *Acta Radiologica*, 55(2):169–174, 2013.
- [53] O. A. Stepanov, D. A. Koshaev, and A. V. Motorin. Identification of gravity anomaly model parameters in airborne gravimetry problems using nonlinear filtering methods. *Gyroscopy and Navigation*, 6(4):318–323, 2015.
- [54] R. Stratonovich. Conditional markov processes. *Theoretical Probability and Its Applications*, 5:156–178, 1960.
- [55] X. Sun and J. Duan. Fokker-Planck equations for nonlinear dynamical systems driven by non-Gaussian Lévy processes. *J. Math. Phys.*, 53(7):072701, 10, 2012.
- [56] M. Ternes, C.P. Lutz, F.J. Hirjibehedin, F.J. Giessibl, and A.J. Heinrich. The force needed to move an atom on a surface. *Science*, 319:1066–1069, 2008.
- [57] G. Walter. Norm estimate for inverses of vandermonde matrices. *Numerische Mathematik*, 23(4):337347, 1975.
- [58] J. Weare. Particle filtering with path sampling and an application to a bimodal ocean current model. *J. Comput. Phys.*, 228(12):4312–4331, 2009.
- [59] M. Zakai. On the optimal filtering of diffusion processes. *Ztschr. Wahrscheinlichkeitstheor und verw. Geb.*, 11:230–243, 1969.
- [60] H. Zhang and D. Laneuville. Grid based solution of zakai equation with adaptive local refinement for bearing-only tracking. *IEEE Aerospace Conference*, 2008.
- [61] M. Zheng and G. Karniadakis. Numerical methods for spdes with tempered stable processes. *SIAM Journal of Scientific Computing*, 37(3):A1197A1217, 2015.