Some Studies on Parameter Estimations

By

Chen Su

Submitted to the Department of Mathematics and the
Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Dr. Yaozhong Hu, Chairperson

Dr. Xuemin Tu, Chairperson

Committee members

Dr. David Nualart

Dr. Terry Soo

Dr. Jianbo Zhang

Date defended: May 4, 2016

The Dissertation Committee for Chen Su certifies
that this is the approved version of the following dissertation :

Some Studies on Parameter Estimations

Dr. Yaozhong Hu, Chairperson

Dr. Xuemin Tu, Chairperson

Date approved: May 4, 2016

Abstract

Parameter estimation has wide applications in such fields as finance, biological science, weather prediction, oil deposit detection, etc. Researchers are particularly interested in reconstructing some unknown parameters from the observed data set which may be sparse and noisy. This is a typical inverse problem which tends to be ill-conditioned in many cases. A plethora of literature has been devoted to this area and there has been a concrete progress recently in the design of more efficient estimation techniques.

Depending on the model (usually an equation or a system of equations) we choose, we divide the estimation into two categories: stochastic and deterministic parameter estimations. The former involves a stochastic system, usually a stochastic differential equation (SDE) or system of SDEs where unknown parameters are present. A standard estimator to use for stochastic parameter estimation problems is the maximum likelihood estimator (MLE), since it, in many situations, enjoys desirable properties such as consistency and asymptotic normality. One major obstacle in obtaining the MLE is that the transition density of SDE, which is essential for deriving MLE, is often not available. To address this issue, various approximations of transition density was introduced in the past decades. In chapter 1, we will present several popular density approximation schemes including Euler-Maruyama methods and Hermite expansions. We will also introduce the parametrix approximation in which we derive a point-wise approximation of the transition density that is uniform in the parameter. As a consequence, the approximated MLE from parametrix method will eventually converge to the true MLE so that those desired properties of MLE can be preserved. We will see

some applications of the parametrix approximation and some necessary preliminaries regarding the ergodicity of SDEs and the consistency of the estimators will also be presented.

The deterministic parameter estimation involves a partial differential equation (PDE) or a system of PDEs. A key feature for this type of estimation is the high level of uncertainty for recovering the parameters, i.e. different choices of parameters may all yield reasonable explanation of the data. This is a typical feature for many ill-conditioned inverse problems. The Bayesian inference formulation provides a systematic way to characterize this uncertainty. It incorporates a prior, which is from the historical data before any experiment is done, and a likelihood, which measures how likely the data will be provided that certain parameter value is chosen, to form a posterior density. It generates a neat solution which takes the form of a posterior probability density. However, how to interpret this posterior density is a non-trivial task since the forward model may be very expensive and the discretized parameter field may result in a high dimensional density. As a consequence, efficient sampling techniques are called for to better characterize the posterior. In chapter 2, we will introduce some traditional sampling methods such as Gaussian approximations, MCMC and importance sampling. We also introduce our implicit sampling methods together with its sequential implementation. We will apply these methods to a seismic wave inversion problem where a detailed comparison among other methods demonstrates a clear superiority of our implicit sampling method.

Finally in chapter 3, we will give some concluding remarks and point out possible future work.

Acknowledgements

I would like to express my gratitude towards all the people who made this thesis possible and who made my graduate study at the University of Kansas a lifetime memory.

I would like to give my deepest gratitude Professor Yaozhong Hu and Professor Xuemin Tu. They opened the gate of mathematics to me and patiently taught me on how to do scientific research. They also helped me a lot for developing my career plan. Without their help and supports, I will never make my achievements possible.

I would like to thank Professor David Nualart who extended my knowledge on stochastic calculus. I am also grateful to Professor Terry Soo and Professor Jianbo Zhang for serving as my committee members.

I am grateful to Professor Atanas Stefanov and Professor Mathew Johnson for introducing me to the would of advanced PDEs. I am also grateful to Professor Erik Van Vleck for his helpful suggestions on numerical methods.

I would like to thank Professor Weishi Liu, Professor Bozenna Pasik Duncan, Professor Judith Roitman and Professor Jack Porter for giving me valuable advice and suggestions on my teaching style.

Many thanks to the department stuff: Kerrie Brecheisen, Debbie Garcia, Gloria Prothe, Lori Springs and Samantha Reinblatt for their generous help with my life at the department of mathematics. The thesis can not be made easily without their support.

I would also like to thank Jun Fu for giving me valuable interview tips and for her great support of my life at KU. Many thanks to my friends, Wenjun Ma, Zheng Han,

Yanru Su, Yiying Cheng and Hongjuan Zhou for their useful discussions that helped me extend my knowledge.

Finally, I would like to thank my parents and my girlfriend Fan Yang for their love and encouragement that support me all the time.

Contents

List of Figures

List of Tables

Chapter 1

Parameter Estimation for Stochastic Systems

In this chapter, we introduce the problem of estimating parameters for a stochastic differential equation (SDE). Depending on the nature of observations, this problem is divided into two categories: continuous time observations and discrete time observations. In both cases, estimation relies heavily on the maximum likelihood estimator (MLE). However, the techniques used for these two categories are quite different. The method for continuous observations is based on Girsanov change of measure formula, see [34] for a thorough study on this topic. When the observations are discrete, we resort to the transition probability density of the SDE which is often unavailable. As a consequence, various approximations of the transition density are studied to deal with this issue. We will focus on discrete-observation estimation problem in this thesis so that we seek schemes which approximate the transition densities. In particular, we will present several existing methods such as Euler-Maruyama approximation which was thoroughly studied in [45] [46] and Hermite polynomial approximation first introduced in [2]. We will also present our method which is based on the parametrix approximation. In particular, we will show its uniform convergence with respect to the unknown parameter so that the approximated MLE possesses the desired property.

1.1 Introduction

Diffusion processes have been a popular tool and are widely used in such fields as mathematical finance and biological science. Most commonly used models in practice are parametric multi-dimensional diffusion processes which take the form of a stochastic differential equation. For simplicity, we will consider time-homogeneous SDE defined on a finite interval $[0,T]$ in this thesis:

$$\begin{cases} dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t, \ \ 0 \le t \le T \\ X_0 = x_0 \end{cases} \tag{1.1}$$

where $b : \mathbb{R}^d \times \Theta \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \Theta \to \mathbb{R}^{d\times d}$ are measurable functions and W_t is a d-dimensional Brownian Motion on a complete probability space $(\Omega, \mathscr{F}, P)$ with a filtration $(\mathscr{F}_t)_{0\le t\le T}$ that satisfies the usual conditions [33]. Moreover, θ is an unknown parameter that belongs to a compact set Θ. For a given θ, (1.1) describes the evolution of the d-dimensional state variable X_t. Under this framework, we have a parametric estimation problem and what we are interested in is the estimation for θ based on the discrete observations of the process X_t.

One popular technique in estimation theory is based on the likelihood function which can be constructed via the transition probability density of the SDE (1.1). If the transition probability density is available, it is natural to choose maximum likelihood estimator (MLE) since it enjoys many desirable asymptotic properties such as consistency and asymptotic normality. To be precise, let us assume that b and σ of (1.1) satisfy the conditions that guarantee the existence of a unique weak solution (see [33] for detailed discussions). Here we list the common sufficient conditions, which we assume to hold uniformly in $\theta \in \Theta$:

Assumption 1. - *(Local Lipchitz) For all $N > 0$, there exists some $K_N > 0$ such that*

$$|b(x;\theta) - b(y;\theta)| \le K_N|x-y|,$$

$$|\sigma(x;\theta) - \sigma(y;\theta)| \le K_N|x-y|$$

for all x and y such that $|x|, |y| \leq N$. *Here* $|\cdot|$ *is the Euclidean norm on* $\mathbb{R}^d$.

Assumption 2. - *(Linear Growth) There exists some* $K > 0$ *such that*

$$|b(x;\theta)| + |\sigma(x;\theta)| \leq K(1+|x|)$$

for all $x \in \mathbb{R}^d$.

Denote by $p(t,x,y;\theta)$ the time-homogeneous transition density associated with X_t, i.e.

$$\int_B p(t,x,y;\theta)dy = \mathbb{E}_\theta[X_t \in B | X_0 = x],$$

where $\mathbb{E}_\theta$ is the expectation under P_θ with θ being the true parameter. Let's assume that the process (1.1) is observed at equal time instants $t_i = i\Delta$, $i = 1,2,\cdots$ with the corresponding observations $X_0 = x_0$, X_Δ, $X_{2\Delta}$, $X_{3\Delta}, \cdots$, then the log-likelihood function is

$$\begin{aligned} l_n(\theta) &= \ln[p(\Delta, x_0, X_\Delta;\theta)p(\Delta, X_\Delta, X_{2\Delta};\theta)\cdots p(\Delta, X_{(n-1)\Delta}, X_{n\Delta};\theta)] \\ &= \sum_{i=1}^{n} \ln p(\Delta, X_{(i-1)\Delta} X_{i\Delta};\theta). \end{aligned} \tag{1.2}$$

Define $\hat{\theta}_n$ to be the maximizer of $l_n(\theta)$:

$$\hat{\theta}_n = \underset{\theta\in\Theta}{\operatorname{argmax}}[l_n(\theta)].$$

For simplicity, let us assume that the maximizer is unique. When it is not unique, we can take any one of the maximizers, which will not affect the following paragraphs. Thus $\hat{\theta}_n$ is the maximum likelihood estimator (MLE).

When the exact transition probability density $p(t,x,y;\theta)$ is known, it has been shown in [20] that under some mild conditions on the coefficients b and σ, $\hat{\theta}_n$ is consistent and asymptotically normal. As a consequence, the knowledge on transition probability density is essential in the construction of MLE. Unfortunately, only in a few cases can we find the closed form transition den-

sity. As a result, various approximations of MLE are developed to solve this issue. For example, D.Dacunha and D.Florens [20] derived an approximation of transition density through Girsanov theorem and Brownian bridge. They also studied the asymptotic property of MLE under the ergodic case of (1.1). In [45], A.R.Perdersen proposed a method based on the Euler-Maruyama scheme and he also showed the consistency and asymototic normality in [46]. A recent paper by Yacine Aït-Sahalia [2] utilized the Hermite polynomial to obtain the approximated MLE and the author also established its asymptotic behavior when the sample is large. In the following subsections, we will briefly introduce these schemes. We need to point out that there are other schemes available. However, to the author's knowledge, these two approaches are among the most popular techniques.

1.2 Euler-Maruyama Scheme

For the given stochastic differential equation (1.1), the simplest way to solve it is the Euler-Maruyama (EM) scheme with fine time discretization. In [45], A.R.Perdersen constructed an approximated transition density that was based on this discretization. To see how it works, write (1.1) in its integral form

$$X_t = x_0 + \int_0^t b(X_u;\theta)du + \int_0^t \sigma(X_u;\theta)dW_u, \quad t \geq 0. \tag{1.3}$$

Denote by P_θ the probability law associated with the process (1.3), see [33]. We assume P_θ has a density

$$P_\theta(X_t \in dy | X_0 = x) = p(t,x,y;\theta)dy.$$

The corresponding Euler-Maruyama scheme is the stochastic analogue of the Euler's scheme for ordinary differential equations and the N-th EM approximation goes as follows: for fixed $t > 0$, divide the interval $(0,t)$ into N subintervals with equal length (although equal length is not a

necessary condition for the scheme to work, we will assume it throughout this section). Define

$$\Delta = \frac{t}{N},$$

$$X_0^{(N)} = x,$$

$$u_i = i\Delta,\ i = 0,1,\ldots,N,$$

and

$$X_{u_{i+1}}^{(N)} = X_{u_i}^{(N)} + b(X_{u_i}^{(N)};\theta)\Delta + \sigma(X_{u_i}^{(N)};\theta)\sqrt{\Delta}Z_i,$$

where Z_i are i.i.d. standard normal distribution $N(0,1)$. It has been shown in [45] that under Assumptions 1 and 2,

$$X_t^{(N)} \to X_t$$

in $L^1(P_\theta)$ as $N \to \infty$. The idea of A.R.Perdersen's approximation is based on the transition density of discretized process $X_t^{(N)}$. To be precise, when $N = 1$, we define the one-step transition probability density to be

$$\begin{aligned} p^{(1)}(t,x,y;\theta) &= |a^{-1}(x;\theta)|^{\frac{1}{2}}(2\pi t)^{-\frac{d}{2}} \\ &\quad \times \exp\left(\frac{1}{2t}[y-x-tb(x;\theta)]^T a^{-1}(x;\theta)[y-x-tb(x;\theta)]\right). \end{aligned} \tag{1.4}$$

Here $a(x;\theta) = \sigma(x;\theta)\sigma(x;\theta)^T$ is a $d \times d$ matrix and σ^T is the transpose of σ. It is easy to see that $p^{(1)}$ is the transition density of $X^{(1)}$. This is a coarse approximation of the true transition density when t is large, and this relationship only gives the transition density between adjacent time points. For finer approximations, we define the N-step ($N \geq 2$) approximated density to be

$$p^{(N)}(t,x,y;\theta) = \int_{\mathbb{R}^{(N-1)d}} \prod_{i=1}^{N} p^{(1)}(\Delta, z_{i-1}, z_i;\theta)\, dz_1 \ldots dz_{N-1} \tag{1.5}$$

where $\Delta = t/N$, $z_0 = x$ and $z_N = y$. Obviously, by Chapman-Kolmogorov equation, $P^{(N)}(t,x,y;\theta)$ gives the transition density of $X_t^{(N)}$. This construction is the most natural one. However, there is a minor drawback in this approximation, that is, we can only obtain the L^1 convergence of the approximated transition density, rather than a point-wise convergence. Notice that in Euler scheme, we generate the standard normal Z for different realizations of the trajectories. As a consequence, $p^{(N)}(t,x,y;\theta)$ depends on the version of X_t we choose and in most cases we can not get the point-wise convergence of $p^{(N)}(t,x,y;\theta)$ to $p(t,x,y;\theta)$. This is stated in the following theorem:

Theorem 1. *- (A.R.Perdersen 1995) Let Assumption 1 and 2 hold. In addition, assume that $b(x;\theta)$ and $\sigma(x;\theta)$ are bounded with bounded derivatives, uniform in $\theta \in \Theta$. Also assume that $a(x;\theta) = \sigma(x;\theta)\sigma(x;\theta)^T$ is uniformly elliptic, i.e. there exist $m, M > 0$ such that*

$$m\xi^T\xi \leq \xi^T a(x;\theta)\xi \leq M\xi^T\xi,$$

for all $\xi \in \mathbb{R}^d$. Then

$$p^{(N)}(t,x,y;\theta) \to p(t,x,y;\theta)$$

in $L^1(P_\theta)$. If in addition, we have that $p^{(N)}$ converges point-wise in (t,x,y), then the limit must be $p(t,x,y;\theta)$.

As pointed out by the author, only in a few cases where the the true transition density is known, pointwise convergence holds. However, this drawback is partially fixed when it comes to the approximated log-likelihood function and the approximated MLE.

Now let

$$\begin{aligned} l_n^{(N)}(\theta) &= \ln[p^{(N)}(\Delta, x_0, X_\Delta;\theta)p^{(N)}(\Delta, X_\Delta, X_{2\Delta};\theta)\cdots p^{(N)}(\Delta, X_{(n-1)\Delta}, X_{n\Delta};\theta)] \qquad (1.6) \\ &= \sum_{i=1}^{n} \ln p^{(N)}(\Delta, X_{(i-1)\Delta}X_{i\Delta};\theta) \end{aligned}$$

be the approximated log-likelihood function. Then

$$l_n^{(N)}(\theta) \to l_n(\theta)$$

in P_θ probability as $N \to \infty$, although we can only derive L^1 convergence of $p^{(N)}(t,x,y;\theta)$. Consequently, the approximated MLE will have the usual consistency and asymptotic normality under regularity conditions.

Note that the definition of $p^{(N)}$ involves the $(N-1)$-fold integral on d-dimensional space $\mathbb{R}^d$. One way to implement this method is to use Monte Carlo integration. The idea is as follows: since

$$\begin{aligned} p^{(N)}(t,x,y;\theta) &= \int_{\mathbb{R}^{(N-1)d}} \prod_{i=1}^{N} p^{(1)}(\Delta, z_{i-1}, z_i;\theta) dz_1 \dots dz_{N-1} \\ &= \int_{\mathbb{R}^d} p^{(N)}((N-1)\Delta, x, z;\theta) p^{(1)}(\Delta, z, y;\theta) \\ &= E_Z(p^{(1)}(\Delta, Z, y;\theta)). \end{aligned}$$

where E_Z is the expectation under $Z = X_{N-1}^{(N)}$. As a consequence, we can generate a large number of random draws of Z_n, $n = 1,2,\dots,N_e$ from $X_{N-1}^{(N)}$ using the EM discretization, and then take the average of $p^{(1)}(\Delta, Z, y;\theta)$ with different values of Z, i.e.

$$p^{(N)}(t,x,y;\theta) \approx \frac{1}{N_e} \sum_{n=1}^{N_e} p^{(1)}(\Delta, Z_n, y;\theta).$$

We will give a detailed discussion on Monte Carlo integration in Chapter 2. The major drawback of using Pedersen's approximation is its computational cost. In a recent paper by G.B.Durham and A.R.Gallant [22], the authors reviewed various improvements on the Pedersen's EM approximation. The paper focused on a scalar and time-homogeneous SDE with a single parameter and it introduced many bias-reduction techniques and variance-reduction techniques that seek to improve the algorithm's performance. We refer the interested reader to their paper.

1.3 Hermite Expansion Scheme

Aït Yacine [2] proposed an approximation in the scalar case which was different from the previous Euler-Maruyama simulation based method. Instead, he seeks a way to expand the transition density as a series. In the real computations, he truncates the series to certain orders. The series is chosen in such a way that for a small enough but fixed Δ, the normal density is served as the leading term. However, this is not achievable in most cases since the diffusion X_t will be far away from normal due to the drift and diffusion terms, and the resulting series starting from a normal distribution will diverge.

To find a series that eventually converges to the true transition density of X_t, he performed two transformations in which the first one transforms the diffusion into Y_t with a unit diffusion coefficient and the second one normalizes Y_t to Z_t. The aim of these two transformations is to make the diffusion closer to normal with a unit diffusion coefficient so that the series which takes $N(0,1)$ as leading term will converge for the normalized diffusion Z_t. Due to the explicit formulas for both transformations, we can easily derive the transition density of X_t by the inverse formula.

Now we make the above description explicit. Let

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dW_t$$

be a scalar diffusion and we assume that $\sigma(x,\theta) > 0$ for all x and $\theta \in \Theta$. Since we also have two other diffusions Y and Z, we denote by $p_X(t,x_0,x;\theta)$, $p_Y(t,y_0,y;\theta)$ and $p_Z(t,z_0,z;\theta)$ their transition densities respectively. Note that this notation is only for this section. In order to transform X_t into a diffusion with unit dW_t coefficient, we introduce $Y_\cdot$ as

$$Y_\cdot = c(X_\cdot;\theta) \triangleq \int^{X_\cdot} \frac{du}{\sigma(u;\theta)}. \tag{1.7}$$

Since $\sigma(x;\theta) > 0$, Y is an increasing function of X and it has a unique inverse. This is helpful when we need to obtain the density of X_t from that of Y_t.

By Ito's formula, we have

$$dY_t = b_Y(Y_t;\theta) + dW_t, \tag{1.8}$$

where

$$b_Y(y;\theta) = \frac{b(c^{-1}(y;\theta);\theta)}{\sigma(c^{-1}(y;\theta);\theta)} - \frac{1}{2}\frac{\partial\sigma}{\partial x}(c^{-1}(y;\theta);\theta).$$

Notice that the transformation from X to Y may introduce singularities in the drift term, which may lead to undesired properties. To exclude those singularities, Aït Yacine specified some restrictions on $b_Y(y;\theta)$. Most of them concern the growth conditions on the boundaries so that the resulting diffusion Y_t will behave nicely to perform further analysis. We skip the detailed restrictions and refer the reader to [2] for the assumptions.

It can be proved that Y_t admits a smooth transition density $p_Y(t,x,y;\theta)$ that is continuously differentiable in t and infinitely differentiable in x and y under the conditions given by Yacine. Moreover, the following two inequalities hold under the same conditions:

$$\begin{aligned} 0 < \quad & p_Y(t,y_0,y;\theta) \\ \leq \quad & C_1 t^{-\frac{1}{2}} \exp\left[-\frac{3(y-y_0)^2}{8t}\right] \exp\left(C_2|y-y_0||y_0| + C_3|y-y_0| + C_4|y_0| + C_5|y_0^2|\right), \end{aligned} \tag{1.9}$$

and

$$\begin{aligned} & \frac{\partial p_Y(t,y_0,y;\theta)}{\partial y} \\ \leq \quad & D_1 t^{-\frac{1}{2}} \exp\left[-\frac{3(y-y_0)^2}{8t}\right] Q(|y_0|,|y|) \exp\left(C_2|y-y_0||y_0| + C_3|y-y_0| + C_4|y_0| + C_5|y_0^2|\right), \end{aligned} \tag{1.10}$$

where Q is certain polynomial in y_0 and y.

What these upper bounds tell us is that the tail of p_Y looks like that of Gaussian due to the exponential decay. As indicated by the author, this is essential for an Hermite expansion to converge. However, this is still not enough to obtain a desired convergence in many cases since p_Y behaves like the dirac-delta function when t is small. In order to deal with this issue, we use the so-called

pseudo-normalizing transformation:

$$Z_t = t^{-\frac{1}{2}}(Y_t - y_0). \tag{1.11}$$

Just as the standard normalization for normal random variables, this pseudo-normalization makes Z closer to $N(0,1)$ so that we can use it as the leading term. Since we have explicit formulas for both transformations, we can readily derive the transition density of Y_t and X_t once we get that of Z_t.

Since Z is close to $N(0,1)$, this suggests that we use Hermite expansions. Recall that the n-th Hermite polynomial is defined as

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n}\left(e^{-\frac{x^2}{2}}\right).$$

They have the well known orthogonality property:

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)\phi(x)dx = \begin{cases} n! & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases}$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ is the standard normal density. As a consequence, we write $p_Z(t, z_0, z; \theta)$ as

$$p_Z(t, z_0, z; \theta) = \phi(z) \sum_{n=0}^{\infty} \mu_Z^{(n)}(t, z_0; \theta) H_n(z), \tag{1.12}$$

with $\mu_Z^{(n)}(t, z_0; \theta)$ being the coefficient, which can be derived using the orthogonality:

$$\mu_Z^{(n)}(t, z_0; \theta) = \frac{1}{n!} \int_{-\infty}^{\infty} H_n(z) p_Z(t, z_0, z; \theta) dz. \tag{1.13}$$

Usually in practice, we truncate the series to have N terms (N is small, say, $N = 1$ or 2):

$$p_Z^{(N)}(t, z_0, z; \theta) = \phi(z) \sum_{n=0}^{N} \mu_Z^{(n)}(t, z_0; \theta) H_n(z).$$

To calculate the integral in (1.13), we can exploit the Monte Carlo methods again, i.e. generate the random path of Z via simulations and calculate the integral by taking averages. The author also indicated ways to calculate this integral by using Taylor expansions. Finally, to obtain the densities of Y and X, we have

$$p_Y^{(N)}(t, y_0, y; \theta) = t^{-\frac{1}{2}} p_Z^{(N)}(t, y_0, t^{-\frac{1}{2}}(y - y_0); \theta), \tag{1.14}$$

and

$$p_X^{(N)}(t, x_0, x; \theta) = \sigma(x; \theta)^{-1} p_Y^{(N)}(t, c(x_0; \theta), c(x; \theta); \theta). \tag{1.15}$$

Yacine also showed that under the assumptions mentioned above, there exists some Δ which is small enough, such that

$$p_X^{(N)}(t, x_0, x; \theta) \to p_X(t, x_0, x; \theta)$$

whenever $t \in (0, \Delta)$. Moreover, the convergence is uniform in $\theta \in \Theta$, which is a desired property for statistical inference.

1.4 Parametrix Approximation

In this section, we introduce the parametrix approximation for the transition density of X_t, which is the third method considered in this thesis. Unlike the EM scheme and Hermite expansions, which were only given a brief description, we show a detailed process of the parametrix construction. In contrast to the previous two schemes, parametrix approximation does not rely on the simulations of the process X_t, nor does it only exhibit L^1 convergence as in EM scheme. It is completely a deterministic procedure. We will show its point-wise convergence to the true density uniformly in $\theta \in \Theta$. As in Hermite expansion scheme, parametrix method represents the transition density in the form of a series. However, we do not need to perform any transformations to obtain the desired convergence.

1.4.1 Construction of Parametrix: Sub-linear Growth

Now we turn to the parametrix construction. We will do this in multivariate case to its full generality. This is an advantage to the Hermite expansion scheme which was applied to scalar case.

Let us now consider multivariate SDE (1.1). As before, let

$$b(x;\theta) = (b_1(x;\theta), b_2(x;\theta), \ldots, b_d(x;\theta))$$

and

$$\sigma(x;\theta) = \{\sigma_{ij}(x;\theta)\}_{i,j=1}^{d}$$

be the coefficients in (1.1). In this section, we assume that both b and σ are defined on $\mathbb{R}^d \times \Theta$. Define $a(x;\theta) = \sigma(x;\theta) \cdot \sigma^T(x;\theta)$ so that $a(x;\theta)$ is a $d \times d$ non-negative definite matrix.

It is well known that under the Assumptions 1 and 2 in Section 1.1, the transition density of (1.1) is the fundamental solution of the Kolmogorov backward equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x;\theta)\frac{\partial^2 p}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x;\theta)\frac{\partial p}{\partial x_i} \tag{1.16}$$

By fundamental solution of (1.16) on $[0,T]$, we mean a function $p(t,x,y;\theta)$ which is defined on the set

$$\tau = [(t,x,y;\theta) \in \mathbb{R}^{2d+1} \times \Theta, t > 0]$$

and satisfies the following conditions:

Condition 1. *For fixed y and θ, p satisfies (1.16) as a function of (t,x) where $0 < t \leq T$ and $x \in \mathbb{R}^d$.*

Condition 2. *For every continuous and bounded f on $\mathbb{R}^d$, we have*

$$\lim_{t \downarrow 0} \int_{\mathbb{R}} p(t,x,y;\theta) f(y) dy = f(x). \tag{1.17}$$

i.e. p satisfies (1.16) with initial condition the dirac-delta function at $y = x$.

If the initial condition is regular enough, say, a continuous function on a bounded domain, classical functional analysis theory already established the existence of the solution. However, with the initial condition so ill-behaved like the dirac-delta function, we need to resort to other method. Here we will employ the Levi's parametrix method [28] to derive the existence of the fundamental solution to (1.16) and to give a concrete construction. This construction starts directly from the PDE (1.16), meaning that it does not involve any randomness as in the schemes which start from the SDE.

Fundamental solutions for (1.16) with continuous coefficients on bounded domain or with bounded continuous coefficients on unbounded domains have been established via the parametrix method. See [28] for a detailed construction process. Now we are trying to generalize it on the unbounded domain $\mathbb{R}^d$ with possibly unbounded coefficients so that more models can be incorporated in our framework.

The essence of parametrix method is that we first write down the fundamental solution of a reduced PDE where we freeze the coefficient $a(x;\theta)$ at $x = y$ and remove the first order terms of (1.16). By using this known fundamental solution as a starting kernel, we can construct the fundamental solution of (1.16) through a Neumann series which eventually converges point-wise to it. Before we actually start the construction, we list the following assumptions for parametrix to work.

Besides Assumption 1 and 2 in Section 1.1, we have

Assumption 3. *All coefficients $a_{ij}(x;\theta)$ and $b_i(x;\theta)$ are jointly continuous functions of (x,θ) in $\mathbb{R}^d \times \Theta$.*

Assumption 4. *The operator $A \triangleq \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x;\theta)\partial_{x_i}\partial_{x_j}$ is uniformly elliptic, uniformly in θ, i.e. there exist positive numbers m and M, not depending on x and θ, such that*

$$m|\xi|^2 \leq \sum_{i,j=1}^{d} a_{ij}(x;\theta)\xi_i\xi_j \leq M|\xi|^2, \ \textit{for all } \theta \in \Theta,$$

for all $x, \xi \in \mathbb{R}^d$. Here $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$.

Assumption 5. *The coefficient $a(x;\theta)$ is Lipchitz in x, uniformly in θ in the sense that*

$$\sum_{i,j=1}^{d} \left|a_{ij}(x;\theta) - a_{ij}(y;\theta)\right| \leq C_a|x-y|,$$

for any $\theta \in \Theta$ and C_a is independent from x, y and θ.

Assumption 6. *The coefficient $b(x;\theta)$ has the following property: there exists some $0 < \beta < 1$, such that*

$$|b(x;\theta)| \leq C_b(|x|^\beta + 1) \text{ for all } \theta \in \Theta$$

where C_b is independent from x and θ.

Note that we need the drift term to possess a growth condition which is slower than linear. We will deal with the linear case in the next section. Define

$$Z(t,x,y;\theta) = \frac{|a^{-1}(y;\theta)|^{\frac{1}{2}}}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{\langle x-y, a^{-1}(y;\theta)(x-y)\rangle}{2t}}, \tag{1.18}$$

where $a(y,\theta) = \{a_{i,j}(y;\theta)\}_{i,j=1}^{d}$, $\langle\cdot,\cdot\rangle$ is the standard inner product on $\mathbb{R}^d$ and $|a^{-1}(y;\theta)|$ is the determinant of $a^{-1}(y;\theta)$. Then it is easy to check that Z satisfies the following heat equation:

$$\frac{\partial Z}{\partial t} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(y;\theta)\frac{\partial^2 Z}{\partial x_i \partial x_j}. \tag{1.19}$$

The function Z also has the following property:

$$\lim_{s \nearrow t} \int_{\mathbb{R}^d} Z(t-s,x,y;\theta) f(s,y)dy = f(t,x),$$

for any $f(t,x)$ that is jointly continuous and bounded, i.e. the function $Z(t,x,y;\theta)$ is the fundamental solution of (1.19). Notice that Z satisfies the equation with "constant" coefficients if we regard x as the variable and y as a "constant", i.e. we freeze the coefficient a at a non-variable point y. For notational simplicity, we define the volume potential of $f(t,x,y)$ with respect to the

kernel $T = T(t,x,y;\theta)$ as

$$T_f(t,x,y;\theta) = \int_0^t \int_{\mathbb{R}^d} T(t-s,x,u;\theta) f(s,u,y) ds dy \triangleq T * f(t,x,y;\theta).$$

It plays the central role in the construction of fundamental solution, as we shall see later. In fact, we are looking for fundamental solution of the form

$$\begin{aligned} \Gamma(t,x,y;\theta) &= Z(t,x,y;\theta) + \int_0^t \int_{\mathbb{R}^d} Z(t-s,x,u;\theta)\psi(s,u,y;\theta) du ds \\ &= Z(t,x,y;\theta) + Z * \psi(t,x,y;\theta). \end{aligned} \tag{1.20}$$

Thus we are taking Z, the fundamental solution of (1.19), whose coefficients are frozen at y, as the principal part of that of (1.16) and we are searching such an ψ that the corresponding Γ will satisfy Conditions 1 and 2 stated above. Our goal is to construct such a ψ via Neumann series.

To this end, we first consider the following form of ψ:

$$\psi(t,x,y;\theta) = LZ(t,x,y;\theta) + LZ * \psi(t,x,y;\theta), \tag{1.21}$$

where the operator

$$\begin{aligned} LZ &= \sum_{i,j=1}^{d} \left[a_{ij}(x;\theta) - a_{ij}(y;\theta)\right] \frac{\partial^2 Z(t,x,y;\theta)}{\partial x_i \partial x_j} \\ &\quad + \sum_{i=1}^{d} b_i(x;\theta) \frac{\partial Z(t,x,y;\theta)}{\partial x_i}. \end{aligned} \tag{1.22}$$

This is a Volterra integral equation with a singular kernel $LZ(t,x,y;\theta)$. In fact we are going to prove that this singularity is removable so that the classical infinite series $\psi = \sum_{n=1}^{\infty} (LZ)_n(t,x,y;\theta)$ yields the solution. Here

$$(LZ)_{n+1} = LZ * (LZ)_n. \tag{1.23}$$

The reason to first solve this equation may seem to be unclear now, but we will clarify it in

Theorem 2. The proof of existence of the fundamental solution is rather long and technical and we break it into several lemmas. Our plan is:

1. Prove the convergence of $\sum_{n=1}^{\infty}(LZ)_n(t,x,y;\theta)$ by a key estimate.

2. Prove two continuity properties of $\psi(t,x,y;\theta)$ so that the change of orders between derivatives and integrals is feasible.

3. Show that the function $\Gamma(t,x,y;\theta)$ is the fundamental solution.

Lemma 1. *Let Assumptions 1, 2 and Assumptions 3–6 hold, then the series* $\sum_{n=1}^{\infty}(LZ)_n(t,x,y;\theta)$ *converges on compact sets of* $(0,T]\times\mathbb{R}^{2d}$. *Moreover, the convergence is uniform in* θ. *As a consequence,* $\psi(t,x,y;\theta)=\sum_{n=1}^{\infty}(LZ)_n(t,x,y;\theta)$ *is the solution to (1.21).*

Proof. We first derive an estimate which is essential in proving the convergence of the series. For fixed ε, there exists a sequence $\{\varepsilon_n\}$ with $\varepsilon_1=\varepsilon$ and $\{\lambda_n\}$ such that

$$|(LZ)_n|(t,x,y;\theta)\leq \quad C_{\varepsilon}^{n}[F_{\varepsilon_n\lambda^*}(y)]^n\frac{1}{\Gamma(\frac{n}{2})}\frac{1}{t^{\frac{2+d-n}{2}}}e^{-\frac{\lambda^*|x-y|^2}{2t}}. \tag{1.24}$$

where λ^* and C_ε are suitable constants and $F_{\varepsilon\lambda}(y)=\left(|y|+\sqrt{\frac{\beta T}{\varepsilon\lambda}}\right)^{\beta}+1$.

Indeed, notice that under Assumption 4 , we have the well-known estimate for the heat kernel: uniformly in θ,

$$\left|\frac{\partial Z}{\partial x_i}\right| \quad\leq\quad C_{\lambda,1}t^{-\frac{d+1}{2}}e^{-\frac{\lambda|x-y|^2}{2t}}, \tag{1.25}$$

and

$$\left|\frac{\partial^2 Z}{\partial x_i\partial x_j}\right| \quad\leq\quad C_{\lambda,2}t^{-\frac{d+2}{2}}e^{-\frac{\lambda|x-y|^2}{2t}}, \tag{1.26}$$

for any $\lambda<M^{-1}$ and any $i,j=1,2,\ldots,d$. Here $C_{\lambda,1}$ and $C_{\lambda,2}$ are generic constants that only

depends on λ. We fix λ in this proof. Since $a(x;\theta)$ is Lipchitz continuous, we obtain

$$\begin{aligned}
&\left|\sum_{i,j=1}^{d}[a_{ij}(x;\theta)-a_{ij}(y;\theta)]\frac{\partial^2 Z}{\partial x_i \partial x_j}\right| \\
\leq\ & C_a C_{\lambda,2}|x-y|\cdot\frac{1}{t^{\frac{d+2}{2}}}\cdot e^{-\frac{\varepsilon\lambda|x-y|^2}{2t}}e^{-\frac{(1-\varepsilon)\lambda|x-y|^2}{2t}} \\
=\ & \sqrt{2}C_a C_{\lambda,2}\left(\frac{|x-y|^2}{2t}\right)^{\frac{1}{2}}e^{-\frac{\varepsilon\lambda|x-y|^2}{2t}}\cdot\frac{1}{t^{\frac{d+1}{2}}}\cdot e^{-\frac{(1-\varepsilon)\lambda(x-y)^2}{2t}} \\
\leq\ & C_1\cdot\frac{1}{t^{\frac{d+1}{2}}}\cdot e^{-\frac{(1-\varepsilon)\lambda|x-y|^2}{2t}},
\end{aligned}$$

where ε is a small number to be specified later. Notice that C_1 depends only on C_a and λ. Due to Assumption 6,

$$\left|\sum_{i=1}^{d}b_i(x;\theta)\frac{\partial Z}{\partial x_i}\right| \leq C_b(|x|^\beta+1)e^{-\frac{\varepsilon\lambda|x-y|^2}{2t}}e^{-\frac{(1-\varepsilon)\lambda|x-y|^2}{2t}}\cdot\frac{1}{t^{\frac{d+1}{2}}},$$

for some constant C_b. To deal with this term, let

$$F(x)=|x|^\beta e^{-\frac{\varepsilon\lambda|x-y|^2}{2t}}.$$

Simple extreme value calculation shows that

$$F(x)\leq\left(|y|+\sqrt{\frac{\beta T}{\varepsilon\lambda}}\right)^\beta.$$

Therefore,

$$\left|\sum_{i=1}^{d}b_i(x;\theta)\frac{\partial Z}{\partial x_i}\right|\leq C_2\left[\left(|y|+\sqrt{\frac{\beta T}{\varepsilon\lambda}}\right)^\beta+1\right]e^{-\frac{(1-\varepsilon)\lambda|x-y|^2}{2t}}\cdot\frac{1}{t^{\frac{d+1}{2}}}.$$

Adding together the estimates for both first and second order terms, we see that

$$\begin{aligned} |LZ(t,x,y;\theta)| &\leq C_3 \left[\left(|y| + \sqrt{\frac{\beta T}{\varepsilon\lambda}} \right)^{\beta} + 1 \right] e^{-\frac{(1-\varepsilon)\lambda|x-y|^2}{2t}} \cdot \frac{1}{t^{\frac{d+1}{2}}} \\ &= C_3 F_{\varepsilon\lambda}(y) \frac{1}{t^{\frac{d+1}{2}}} e^{-\frac{(1-\varepsilon)\lambda|x-y|^2}{2t}}. \end{aligned}$$

Let us define $\varepsilon_1 = \varepsilon < 1$, fixed, and choose $\{\varepsilon_i\}_{i=1}^{\infty}$ to be such that $\sum \varepsilon_i \leq 1$. Then the infinite product $\prod(1-\varepsilon_i)$ will be strictly greater than 0. We next define $\lambda_1 = (1-\varepsilon_1)\lambda$ and $\lambda_n = (1-\varepsilon_n)\lambda_{n-1}$. We will show the following estimate:

$$\begin{aligned} &|(LZ)_n|(t,x,y;\theta) \\ &\leq C_{\varepsilon}^n \prod_{i=1}^{n} F_{\varepsilon_i\lambda_i}(y) \left(\frac{2\pi}{\lambda_i} \right)^{\frac{d}{2}} \frac{1}{\Gamma(\frac{n}{2})} \frac{1}{t^{\frac{2+d-n}{2}}} e^{-\frac{\lambda_n|x-y|^2}{2t}}, \end{aligned} \tag{1.27}$$

for appropriate C_{ε}. Assuming the above inequality is true for n, we estimate

$$\begin{aligned} &|(LZ)_{n+1}|(t,x,y;\theta) = |LZ * (LZ)_n(t,x,y;\theta)| \\ \leq\ & C^n C_3 \prod_{i=1}^{n} F_{\varepsilon_i\lambda_i}(y) \left(\frac{2\pi}{\lambda_i} \right)^{\frac{d}{2}} \frac{1}{\Gamma(\frac{n}{2})} \\ & \cdot \int_0^t \int_{\mathbb{R}^d} F_{\varepsilon_1\lambda_1}(u) \cdot \frac{1}{(t-s)^{\frac{d+1}{2}}} \cdot e^{-\frac{\lambda_1|x-u|^2}{2(t-s)}} \frac{1}{s^{\frac{2+d-n}{2}}} e^{-\frac{\lambda_n|u-y|^2}{2s}} du ds \end{aligned}$$

Observe that

$$\begin{aligned} F_{\varepsilon_1\lambda_1}(u) &= \left| \left(|u| + \sqrt{\frac{\beta T}{\varepsilon_1\lambda_1}} \right)^{\beta} + 1 \right| \\ &\leq |u|^{\beta} + \left(\sqrt{\frac{\beta T}{\varepsilon_1\lambda_1}} \right)^{\beta} + 1 \\ &\leq C_4(|u|^{\beta} + 1), \end{aligned}$$

for some fixed C_4. As a consequence,

$$\begin{aligned}
|(LZ)_{n+1}|(t,x,y;\theta) &\leq C^n C_3 C_4 \prod_{i=1}^{n} F_{\varepsilon_i\lambda_i}(y)\left(\frac{2\pi}{\lambda_i}\right)^{\frac{d}{2}}\frac{1}{\Gamma(\frac{n}{2})} \\
&\quad \cdot \int_0^t \int_{\mathbb{R}^d} (|u|^\beta + 1)\cdot e^{-\frac{\varepsilon_{n+1}\lambda_n|u-y|^2}{2(t-s)}}\cdot \frac{1}{(t-s)^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda_1|x-u|^2}{2(t-s)}} \\
&\quad \cdot \frac{1}{s^{\frac{2+d-n}{2}}} e^{-\frac{(1-\varepsilon_{n+1})\lambda_n|u-y|^2}{2s}} du ds \\
&\leq C^n C_3 C_4 \prod_{i=1}^{n} F_{\varepsilon_i\lambda_i}(y)\left(\frac{2\pi}{\lambda_i}\right)^{\frac{d}{2}}\frac{1}{\Gamma(\frac{n}{2})} F_{\varepsilon_{n+1}\lambda_{n+1}}(y) \\
&\quad \cdot \int_0^t \int_{\mathbb{R}^d} \frac{1}{(t-s)^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda_1|x-u|^2}{2(t-s)}}\cdot \frac{1}{s^{\frac{2+d-n}{2}}}\cdot e^{-\frac{(1-\varepsilon_{n+1})\lambda_n|u-y|^2}{2s}} du ds \\
&\leq C^n C_3 C_4 \prod_{i=1}^{n} F_{\varepsilon_i\lambda_i}(y)\left(\frac{2\pi}{\lambda_i}\right)^{\frac{d}{2}} F_{\varepsilon_{n+1}\lambda_{n+1}}(y)\frac{1}{\Gamma(\frac{n}{2})} \\
&\quad \cdot \int_0^t \int_{\mathbb{R}^d} \frac{1}{(t-s)^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda_{n+1}|x-u|^2}{2(t-s)}}\cdot \frac{1}{s^{\frac{2+d-n}{2}}}\cdot e^{-\frac{\lambda_{n+1}|u-y|^2}{2s}} du ds \\
&= C^n C_3 C_4 \prod_{i=1}^{n+1} F_{\varepsilon_i\lambda_i}(y)\left(\frac{2\pi}{\lambda_i}\right)^{\frac{d}{2}}\frac{1}{\Gamma(\frac{n}{2})}\frac{1}{t^{\frac{d+1-n}{2}}}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{1+n}{2})}\cdot e^{-\frac{\lambda_{n+1}|x-y|^2}{2t}} \\
&= C^n C_3 C_4 \Gamma\left(\frac{1}{2}\right)\prod_{i=1}^{n+1} F_{\varepsilon_i\lambda_i}(y)\left(\frac{2\pi}{\lambda_i}\right)^{\frac{d}{2}}\frac{1}{\Gamma(\frac{1+n}{2})}\frac{1}{t^{\frac{d+1-n}{2}}}\cdot e^{-\frac{\lambda_{n+1}|x-y|^2}{2t}}.
\end{aligned}$$

By taking $C_\varepsilon = \max\{C, C_3, C_4, \Gamma\left(\frac{1}{2}\right)\}$ (notice that they do not change through the induction steps), we see that (1.27) holds for $n+1$. Now let us define $\lambda^* = \prod_{i=1}^{\infty}(1-\varepsilon_i)$, then by the choice of ε_i, we obtain $\lambda_i \downarrow \lambda_* > 0$ for any i. Finally, by noting that the functions $F_{\varepsilon\lambda}(y)$, $e^{-\frac{\lambda|x-y|^2}{2t}}$ and $\left(\frac{2\pi}{\lambda}\right)^{\frac{d}{2}}$ are all decreasing with respect to λ, we see that (1.24) holds for all $n \geq 1$.

Now we turn to the proof that $\sum_{n=1}^{\infty}(LZ)_n$ converges on compact subsets of $(0,T]\times\mathbb{R}^{2d}$. For

this end, we first consider the sum $\sum_{n=2+d}^{\infty}(LZ)_n(t,x,y;\theta)$. We have

$$
\begin{aligned}
&\sum_{n=2+d}^{\infty}(LZ)_n(t,x,y;\theta)\\
\leq\ &\sum_{n=2+d}^{\infty}\frac{C_\varepsilon^n\left[\left(|y|+\sqrt{\frac{\beta T}{\varepsilon_n\lambda^*}}\right)^\beta+1\right]^n t^{\frac{n-2-d}{2}}}{\Gamma(\frac{n}{2})}\cdot e^{-\frac{\lambda^*|x-y|^2}{2t}}\\
\leq\ &\sum_{n=2+d}^{\infty}\frac{C_\varepsilon^n\left[\left(|y|+\sqrt{\frac{\beta T}{\varepsilon_n\lambda^*}}\right)^\beta+1\right]^n T^{\frac{n-2-d}{2}}}{\Gamma(\frac{n}{2})}\cdot e^{-\frac{\lambda^*|x-y|^2}{2t}}\\
\leq\ &\sum_{n=2+d}^{\infty}\frac{C_5^n\left[\left(|y|+\sqrt{\frac{\beta T}{\varepsilon_n\lambda^*}}\right)^\beta+1\right]^n}{\Gamma(\frac{n}{2})}e^{-\frac{\lambda^*|x-y|^2}{2t}}\\
\leq\ &\sum_{n=2+d}^{\infty}\frac{C_6^n(|y|^\beta+1)^n}{\Gamma(\frac{n}{2})}\cdot e^{-\frac{\lambda^*|x-y|^2}{2t}}+\sum_{n=2+d}^{\infty}\frac{C_6^n\left(\sqrt{\frac{\beta T}{\varepsilon_n\lambda^*}}\right)^{n\beta}}{\Gamma(\frac{n}{2})}\cdot e^{-\frac{\lambda^*|x-y|^2}{2t}}\\
\leq\ &\sum_{n=2+d}^{\infty}\frac{C_7^n|y|^{\beta n}}{\Gamma(\frac{n}{2})}e^{-\frac{\lambda^*|x-y|^2}{2t}}+\sum_{n=2+d}^{\infty}\frac{C_7^n}{\Gamma(\frac{n}{2})}e^{-\frac{\lambda^*|x-y|^2}{2t}}\\
&+\sum_{n=2+d}^{\infty}\frac{C_7^n\left(\sqrt{\frac{\beta T}{\varepsilon_n\lambda^*}}\right)^{n\beta}}{\Gamma(\frac{n}{2})}e^{-\frac{\lambda^*|x-y|^2}{2t}}\\
\triangleq\ &F_1+F_2+F_3.
\end{aligned}
$$

Using the expansion

$$\Gamma\left(\frac{n}{2}\right)\sim(n!)^{\frac{1}{2}}n^{-\frac{1}{4}}\left(\frac{1}{2}\right)^{\frac{n-1}{2}},$$

we see that $F_2<\infty$ on compact subsets of $(0,T]\times\mathbb{R}^{2d}$ and for any $\theta\in\Theta$. Moreover, if we choose δ small enough and $\varepsilon_n=\varepsilon_1\cdot n^{-(\frac{1}{\beta}-\delta)}$ with $(\frac{1}{\beta}-\delta)>1$ (Note that $\beta<1$),

$$
\begin{aligned}
F_3\ &\leq\ \sum_{n=2+d}^{\infty}\frac{C_8^n\varepsilon^{-\frac{\beta n}{2}}n^{(\frac{1}{2}-\frac{\delta\beta}{2})n}}{\Gamma(\frac{n}{2})}e^{-\frac{\lambda^*|x-y|^2}{2t}}\\
&\sim\ \sum_{n=2+d}^{\infty}\frac{C_8^n\varepsilon^{-\frac{\beta n}{2}}n^{\frac{1}{4}}n^{\frac{n}{2}-\frac{\delta\beta n}{2}}}{(n!)^{\frac{1}{2}}(\frac{1}{2})^{\frac{n-1}{2}}}e^{-\frac{\lambda^*|x-y|^2}{2t}}<\infty.
\end{aligned}
$$

Finally let us treat F_1. Note that by Hölder inequality

$$\begin{aligned} F_1 &\leq \sum_{n=2+d}^{\infty} \frac{C_9^n n^{\frac{1}{4}} |y|^{\beta n}}{(n!)^{\frac{1}{2}}} \\ &\leq \left(\sum_{n=2+d}^{\infty} \left(\frac{n^{\frac{1}{4}}}{2^n} \right)^2 \right)^{\frac{1}{2}} \left(\sum_{n=2+d}^{\infty} \left(\frac{|2C_9 y^{\beta}|^n}{(n!)^{\frac{1}{2}}} \right)^2 \right)^{\frac{1}{2}} \\ &\leq C_{10} e^{2C_9 |y|^{\beta}}. \end{aligned}$$

Since $\beta < 1$, we can find for any h a suitable constant C_h such that $C_{10} e^{2C_9 |y|^{\beta}} \leq C_h e^{h|y|}$. Therefore, we have proved that $\sum (LZ)_n$ converges on compact subsets of $(0,T] \times \mathbb{R}^{2d}$ for any $\theta \in \Theta$. We also have proved the estimate: for any $h > 0$, there exists a constant C_h such that

$$|\psi(t,x,y;\theta)| = |\sum_{n=1}^{\infty} (LZ)_n(t,x,y;\theta)| \leq C_h \frac{1}{t^{\frac{d+1}{2}}} e^{-\frac{\lambda^* |x-y|^2}{2t}} e^{h|y|}. \tag{1.28}$$

Here the $1/t^{\frac{d+1}{2}}$ term comes from the summation of $|\sum_{n=1}^{d+1} (LZ)_n(t,x,y;\theta)|$. The uniformity in θ is clear from the proof. □

Now we prove two more lemmas that will allow us to interchange the derivative with integrals.

Lemma 2. *Assume the same conditions as Lemma 1. For any $\alpha < 1$, $\gamma = 1 - \alpha$ and $\lambda^{**} < \lambda^*$, the following inequality holds:*

$$\begin{aligned} &|\psi(t,x_1,y;\theta) - \psi(t,x_2,y;\theta)| \\ \leq\ & C|x_1 - x_2|^{\alpha} \cdot \frac{1}{t^{\frac{d+2-\gamma}{2}}} \cdot \left[e^{-\frac{\lambda^{**}|x_1-y|^2}{2t}} + e^{-\frac{\lambda^{**}|x_2-y|^2}{2t}} \right] e^{h|y|}, \end{aligned} \tag{1.29}$$

where C is a suitable constant. That is, $\psi(t,x,y;\theta)$ is Hölder continuous in x for fixed t and y, uniformly in θ.

Proof. Notice that $\psi = \psi_1 + \psi_1 * \psi$ where $\psi_1 = LZ(t,x,y;\theta)$. Thus we only need to show (1.29) holds true for both ψ_1 and $\psi_1 * \psi$. For ψ_1, we distinguish two cases:

Case 1. $|x_1 - x_2|^2 \le t$. In this case,

$$\begin{aligned}
&\psi_1(t,x_1,y;\theta) - \psi_1(t,x_2,y;\theta) \\
= &\sum_{i,j=1}^{d} [a_{ij}(x_1;\theta) - a_{ij}(y;\theta)] \frac{\partial^2 Z}{\partial x_i \partial x_j}(t,x_1,y;\theta) + \sum_{i=1}^{d} b_i(x_1;\theta) \frac{\partial Z}{\partial x_i}(t,x_1,y;\theta) \\
&- \sum_{i,j=1}^{d} [a_{ij}(x_2;\theta) - a_{ij}(y;\theta)] \frac{\partial^2 Z}{\partial x_i \partial x_j}(t,x_2,y;\theta) - \sum_{i=1}^{d} b_i(x_2;\theta) \frac{\partial Z}{\partial x_i}(t,x_2,y;\theta).
\end{aligned}$$

Let us look at the second order terms first. Write the difference as

$$\begin{aligned}
&\sum_{i,j=1}^{d} [a_{ij}(x_1;\theta) - a_{ij}(y;\theta)] \frac{\partial^2 Z}{\partial x_i \partial x_j}(t,x_1,y;\theta) \\
&- \sum_{i,j=1}^{d} [a_{ij}(x_2;\theta) - a_{ij}(y;\theta)] \frac{\partial^2 Z}{\partial x_i \partial x_j}(t,x_2,y;\theta) \\
= &\sum_{i,j=1}^{d} [a_{ij}(x_1;\theta) - a_{ij}(x_2;\theta)] \frac{\partial^2 Z}{\partial x_i \partial x_j}(t,x_1,y;\theta) \\
&+ \sum_{i,j=1}^{d} [a_{ij}(x_2;\theta) - a_{ij}(y;\theta)] \left[\frac{\partial^2 Z}{\partial x_i \partial x_j}(t,x_1,y;\theta) - \frac{\partial^2 Z}{\partial x_i \partial x_j}(t,x_2,y;\theta) \right] \\
\triangleq &\ F_1 + F_2.
\end{aligned}$$

Now we estimate

$$\begin{aligned}
|F_1| &\le C_1 |x_1 - x_2| \cdot \frac{1}{t^{\frac{d+2}{2}}} \cdot e^{-\frac{\lambda^* |x_1 - y|^2}{2t}} \\
&= C_1 |x_1 - x_2|^{\alpha} \cdot \frac{1}{t^{\frac{d+2}{2}}} \cdot |x_1 - x_2|^{1-\alpha} \cdot e^{-\frac{\lambda^* |x_1 - y|^2}{2t}} \\
&\le C_1 |x_1 - x_2|^{\alpha} \cdot \frac{1}{t^{\frac{2+d-\gamma}{2}}} \cdot e^{-\frac{\lambda^* |x_1 - y|^2}{2t}}.
\end{aligned}$$

By the mean value theorem and the estimate for $\frac{\partial^3 Z}{\partial x^3}$, (here we abuse the notation by taking $\frac{\partial^n Z}{\partial x^n}$ the

n-th order derivative of Z), we obtain for some ξ on the line connecting x_1 and x_2,

$$\begin{aligned}
&\left|\frac{\partial^2 Z}{\partial x^2}(t,x_1,y;\theta)-\frac{\partial^2 Z}{\partial x^2}(t,x_2,y;\theta)\right| \\
\leq\ & |x_1-x_2|\cdot\frac{\partial^3 Z}{\partial x^3}(t,\xi,y;\theta) \\
\leq\ & C_2|x_1-x_2|\cdot\frac{1}{t^{\frac{d+3}{2}}}\cdot e^{-\frac{\lambda^*|\xi-y|^2}{2t}} \\
\leq\ & C_2|x_1-x_2|\cdot\frac{1}{t^{\frac{d+3}{2}}}\cdot e^{-\frac{k|x_2-y|^2}{2t}},
\end{aligned}$$

for some k. Therefore,

$$\begin{aligned}
|F_2| &\leq C_3|x_2-y||x_1-x_2|^{\alpha}|x_1-x_2|^{1-\alpha}\cdot\frac{1}{t^{\frac{d+3}{2}}}\cdot e^{-\frac{k|x_2-y|^2}{2t}} \\
&\leq C_3|x_2-y||x_1-x_2|^{\alpha}\cdot\frac{1}{t^{\frac{d+3-\gamma}{2}}}\cdot e^{-\frac{k|x_2-y|^2}{2t}} \\
&\leq C_4|x_1-x_2|^{\alpha}\cdot\frac{1}{t^{\frac{d+2-\gamma}{2}}}\cdot e^{-\frac{\lambda'|x_2-y|^2}{2t}},
\end{aligned}$$

for some $\lambda' < k$ and λ^*. Lower order terms are treated similarly. Notice that $e^{h|y|}\geq 1$, we conclude that ψ satisfies (1.29) with $\lambda^{**}=\lambda'$ in this case.

Case 2. $|x_1-x_2|^2 > t$. We have by the key estimate (1.24)

$$\begin{aligned}
\psi_1(t,x_1,y;\theta) &\leq C_4 F_{\varepsilon_1\lambda^*}(y)\cdot\frac{1}{t^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda^*|x_1-y|^2}{2t}} \\
&= C_4 F_{\varepsilon_1\lambda^*}(y)\cdot\frac{1}{t^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda^*|x_1-y|^2}{2t}}\left(\frac{t^{\frac{1-\gamma}{2}}}{t^{\frac{1-\gamma}{2}}}\right) \\
&\leq C_4 F_{\varepsilon_1\lambda^*}(y)\cdot\frac{1}{t^{\frac{d+2-\gamma}{2}}}\cdot e^{-\frac{\lambda^*|x_1-y|^2}{2t}}\cdot|x_1-x_2|^{\alpha} \\
&\leq C_5|x_1-x_2|^{\alpha}\cdot\frac{1}{t^{\frac{d+2-\gamma}{2}}}\cdot e^{-\frac{\lambda'|x_1-y|^2}{2t}}\cdot e^{h|y|},
\end{aligned}$$

where the last inequality is due to the polynomial growth of $F_{\varepsilon\lambda}(y)$. Similarly,

$$\psi_1(t,x_2,y;\theta) \le C_5|x_1-x_2|^{\alpha} \cdot \frac{1}{t^{\frac{d+2-\gamma}{2}}} \cdot e^{-\frac{\lambda'|x_2-y|^2}{2t}} e^{h|y|}.$$

Thus we derived that $\psi_1(t,x,y;\theta)$ satisfies (1.29).

Next we show that (1.29) also holds for $\psi_1 * \psi$. First observe that

$$\begin{aligned} & |\psi_1(t,x_1,y;\theta) - \psi_1(t,x_2,y;\theta)| \\ \le\ & C|x_1-x_2|^{\alpha} \cdot \frac{1}{t^{\frac{d+2-\gamma}{2}}} \cdot \left[e^{-\frac{\lambda'|x_1-y|^2}{2t}} + e^{-\frac{\lambda'|x_2-y|^2}{2t}} \right] F_{\varepsilon_1\lambda^*}(y) \end{aligned}$$

which can be seen from the previous proof, since $F_{\varepsilon_1\lambda^*}(y) \ge 1$. As a consequence,

$$\begin{aligned} & |\psi_1 * \psi(t,x_1,y;\theta) - \psi_1 * \psi(t,x_2,y;\theta)| \\ \le\ & \int_0^t \int_{\mathbb{R}^d} |\psi_1(t-s,x_1,u;\theta) - \psi_1(t-s,x_2,u;\theta)| \psi(s,u,y;\theta) du ds \\ \le\ & C_6|x_1-x_2|^{\alpha} e^{h|y|} \int_0^t \int_{\mathbb{R}^d} F_{\varepsilon_1\lambda^*}(u) \cdot \frac{1}{(t-s)^{\frac{d+2-\gamma}{2}}} \cdot \left[e^{-\frac{\lambda'|x_1-u|^2}{2(t-s)}} + e^{-\frac{\lambda'|x_2-u|^2}{2(t-s)}} \right] \\ & \cdot \frac{1}{s^{\frac{d+1}{2}}} \cdot e^{-\frac{\lambda^*|u-y|^2}{2s}} \\ \le\ & C_7|x_1-x_2|^{\alpha} e^{h|y|} F_{\varepsilon_1\lambda^*}(y) \int_0^t \int_{\mathbb{R}^d} \frac{1}{(t-s)^{\frac{d+2-\gamma}{2}}} \cdot \left[e^{-\frac{\lambda^{**}|x_1-u|^2}{2(t-s)}} + e^{-\frac{\lambda^{**}|x_2-u|^2}{2(t-s)}} \right] \\ & \cdot \frac{1}{s^{\frac{d+1}{2}}} \cdot e^{-\frac{\lambda^{**}|u-y|^2}{2s}} \\ =\ & C_7|x_1-x_2|^{\alpha} e^{h|y|} F_{\varepsilon_1\lambda^*}(y) \cdot \frac{1}{t^{\frac{d+1-\gamma}{2}}} \cdot \left[e^{-\frac{\lambda^{**}|x_1-y|^2}{2t}} + e^{-\frac{\lambda^{**}|x_2-y|^2}{2t}} \right] \\ \le\ & C_8|x_1-x_2|^{\alpha} e^{2h|y|} \cdot \frac{1}{t^{\frac{d+2-\gamma}{2}}} \cdot \left[e^{-\frac{\lambda^{**}|x_1-y|^2}{2t}} + e^{-\frac{\lambda^{**}|x_2-y|^2}{2t}} \right]. \end{aligned}$$

where $\lambda^{**} < \lambda'$. Here in the third inequality, we used the same trick as in Lemma 1. The proof is complete. □

Lemma 3. *With the same assumptions as Lemma 1, $\psi(t,x,y;\theta)$ is jointly continuous in (t,x) on*

$(0,T]\times\mathbb{R}^{2d}$ *for fixed* y*, uniformly in* θ.

Proof. Since $\psi=\sum_{n=1}^{\infty}(LZ)_n(t,x,y;\theta)$ converges on compact subsets of $(0,T]\times\mathbb{R}^{2d}$, we only need to show that $\psi_{n+1}=\psi_1*\psi_n$ is jointly continuous for fixed y, uniformly in θ. Write

$$\begin{aligned}
&|\psi_{n+1}(t_1,x_1,y;\theta)-\psi_{n+1}(t_2,x_2,y;\theta)|\\
=\;&\Big|\int_0^{t_1-\delta}\int_{\mathbb{R}^d}\psi_1(t_1-s,x_1,u;\theta)\psi_n(s,u,y;\theta)duds\\
&+\int_{t_1-\delta}^{t_1}\int_{\mathbb{R}^d}\psi_1(t_1-s,x_1,u;\theta)\psi_n(s,u,y;\theta)duds\\
&-\int_0^{t_2-\delta}\int_{\mathbb{R}^d}\psi_1(t_2-s,x_2,u;\theta)\psi_n(s,u,y;\theta)duds\\
&-\int_{t_2-\delta}^{t_2}\int_{\mathbb{R}}\psi_1(t_2-s,x_2,u;\theta)\psi_n(s,u,y;\theta)duds\Big|\\
\triangleq\;&|A_1+A_2-A_3-A_4|.
\end{aligned}$$

Now if we choose δ so small that $t_1-\delta<t_1/2$, then by (1.24),

$$\begin{aligned}
&\left|\int_{t_1-\delta}^{t_1}\int_{\mathbb{R}^d}\psi_1(t_1-s,x_1,u;\theta)\psi_n(s,u,y;\theta)\right|\\
\le\;&C_1\prod_{i=1}^{n}F_{\varepsilon_i\lambda_i}(y)\int_{t_1-\delta}^{t_1}\int_{\mathbb{R}^d}F_{\varepsilon_1\lambda_1}(u)\cdot\frac{1}{(t_1-s)^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda_1|x_1-u|^2}{2(t_1-s)}}\cdot\frac{1}{s^{\frac{d+2-n}{2}}}\cdot e^{-\frac{\lambda_n|u-y|^2}{2s}}duds\\
\le\;&C_2\prod_{i=1}^{n+1}F_{\varepsilon_i\lambda_i}(y)\int_{t_1-\delta}^{t_1}\int_{\mathbb{R}^d}\frac{1}{(t_1-s)^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda^*|x_1-u|^2}{2(t_1-s)}}\cdot\frac{1}{s^{\frac{d+2-n}{2}}}\cdot e^{-\frac{\lambda^*|u-y|^2}{2s}}duds\\
\le\;&C_3\prod_{i=1}^{n+1}F_{\varepsilon_i\lambda_i}(y)\left(\frac{t_1}{2}\right)^{-\frac{d+2-n}{2}}\cdot\int_{t_1-\delta}^{t_1}\int_{\mathbb{R}^d}\frac{1}{(t_1-s)^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda^*|x_1-u|^2}{2(t_1-s)}}duds\\
=\;&C(y)\sqrt{\delta}.
\end{aligned}$$

where we used the following fact for the last equality:

$$\int_{t_1-\delta}^{t_1}\int_{\mathbb{R}^d}\frac{1}{(t_1-s)^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda^*|x_1-u|^2}{2(t_1-s)}}duds=C_4\sqrt{\delta},$$

for some C_4. This can be seen after a change of variable. Thus we derived that $|A_2|\le C(y)\sqrt{\delta}$ for

some constant $C(y)$ depending on y. Similarly, we can prove that

$$|A_4| \leq C_1(y)\sqrt{\delta}.$$

The standard dominated convergence theorem implies that $|A_1 - A_3|$ converges to 0 when (t_2, x_2) tends to (t_1, x_1) since no singularity is involved there. We omit the detailed proof here. Therefore we derived the joint continuity of $\psi(t,x,y;\theta)$ in (t,x). □

Remark: Lemma 2 and Lemma 3 enable us to invoke Theorems 2-5 in [28] to change the order of integrals and derivatives (up to the second order derivative in x) in the volume potential. Note that in [28], these theorems are proved for bounded domain. However, a careful examination of these proofs indicates that we can derive the same conclusion for unbounded domain through an almost word by word repetition of the proofs there. Now we are ready to state the main theorem:

Theorem 2. *Under Assumptions 1, 2 and Assumptions 3–6, the function* $\Gamma(t,x,y;\theta)$ *with*

$$\psi = \sum_{n=1}^{\infty} (LZ)_n(t,x,y;\theta)$$

is the fundamental solution of (1.16) so that we can express it in terms of infinite series.

Proof. To show Condition 1 holds, note that

$$\psi(t,x,y;\theta) = LZ(t,x,y;\theta) + LZ * \psi(t,x,y;\theta),$$

where

$$LZ = \sum_{i,j=1}^{d} \left[a_{ij}(x;\theta) - a_{ij}(y;\theta)\right] \frac{\partial^2 Z(t,x,y;\theta)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x;\theta) \frac{\partial Z(t,x,y;\theta)}{\partial x_i}$$

Due to Lemma 2 and 3, we have

$$\frac{\partial \Gamma}{\partial t} = \frac{\partial Z}{\partial t} + \int_0^t \int_{\mathbb{R}^d} \frac{\partial Z(t-s,x,u;\theta)}{\partial t} \psi(s,u,y;\theta) du ds + \psi(t,x,y;\theta),$$

$$\frac{\partial \Gamma}{\partial x_i} = \frac{\partial Z}{\partial x_i} + \int_0^t \int_{\mathbb{R}^d} \frac{\partial Z(t-s,x,u;\theta)}{\partial x_i} \psi(s,u,y;\theta) du ds,$$

and

$$\frac{\partial^2 \Gamma}{\partial x_i \partial x_j} = \frac{\partial^2 Z}{\partial x_i \partial x_j} + \int_0^t \int_{\mathbb{R}^d} \frac{\partial^2 Z(t-s,x,u;\theta)}{\partial x_i \partial x_j} \psi(s,u,y;\theta) du ds.$$

where $i, j = 1, 2, \ldots, d$.

Now Condition 1 can be easily seen by using the identity

$$\begin{aligned}
& \sum_{i,j=1}^{d} a_{ij}(x;\theta) \frac{\partial^2 \Gamma}{\partial x_i x_j}(t,x,y;\theta) \\
= & \sum_{i,j=1}^{d} \left[a_{ij}(x;\theta) - a_{ij}(y;\theta) \right] \frac{\partial^2 Z(t,x,y;\theta)}{\partial x_i \partial x_j} + \sum_{i,j=1}^{d} a_{ij}(y;\theta) \frac{\partial^2 Z(t,x,y;\theta)}{\partial x_i \partial x_j} \\
& + \int_0^t \int_{\mathbb{R}^d} \sum_{i,j=1}^{d} \left[a_{ij}(x;\theta) - a_{ij}(y;\theta) \right] \frac{\partial^2 Z(t-s,x,u;\theta)}{\partial x_i \partial x_j} \psi(s,u,y;\theta) du ds \\
& + \int_0^t \int_{\mathbb{R}^d} \sum_{i,j=1}^{d} a_{ij}(y;\theta) \frac{\partial^2 Z(t-s,x,u;\theta)}{\partial x_i \partial x_j} \psi(s,u,y;\theta) du ds.
\end{aligned}$$

and the fact that Z verifies equation (1.18).

Now we only need to show that (1.17) in Condition 2 is met by $\Gamma(t,x,y;\theta)$. Note that

$$\begin{aligned}
& \lim_{t \downarrow 0} \int_{\mathbb{R}^d} \Gamma(t,x,y;\theta) f(y) dy \\
= & \lim_{t \downarrow 0} \int_{\mathbb{R}^d} Z(t,x,y;\theta) f(y) dy + \lim_{t \downarrow 0} \int_{\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} Z(t-s,x,u;\theta) \psi(s,u,y;\theta) f(y) du ds dy \\
= & f(x) + \lim_{t \downarrow 0} \int_{\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} Z(t-s,x,u;\theta) \psi(s,u,y;\theta) f(y) du ds dy
\end{aligned}$$

But if $||f||_\infty = K$, we have by (1.24)

$$\begin{aligned}
& \lim_{t\downarrow 0}\left|\int_{\mathbb{R}^d}\int_0^t\int_{\mathbb{R}^d} Z(t-s,x,u;\theta)\psi(s,u,y;\theta)f(y)dudsdy\right| \\
\leq\ & \lim_{t\downarrow 0} K\int_{\mathbb{R}^d}\int_0^t\int_{\mathbb{R}^d} Z(t-s,x,u;\theta)\psi(s,u,y;\theta)dudsdy \\
\leq\ & \lim_{t\downarrow 0} K_h\int_{\mathbb{R}^d}\int_0^t\int_{\mathbb{R}^d} \frac{1}{(t-s)^{\frac{d}{2}}}\cdot e^{-\frac{\lambda^*|x-u|^2}{2(t-s)}}\cdot\frac{1}{s^{\frac{d+1}{2}}}\cdot e^{-\frac{\lambda^*|u-y|^2}{2t}}e^{h|y|}dudsdy \\
=\ & \lim_{t\downarrow 0} K_h'\int_{\mathbb{R}^d}\cdot e^{-\frac{\lambda^*|x-y|^2}{2s}}e^{h|y|}dy \\
=\ & 0
\end{aligned}$$

where K_h and K_h' are appropriate constants. □

1.4.2 Generalization to Linear Growth

Notice that the parametrix method can be used not only in the situations where we start from a Gaussian transition probability density, but it can be generalized to other cases provided that a close-by transition density is known. In the following two subsections, we will generalize the parametrix method to two general cases where this technique works.

In the last section, we demonstrated the parametrix approximation for drift term $b(x;\theta) = O(|x|^\beta + 1)$ with $\beta < 1$. We now extend it to the linear case. For simplicity, we assume that $d = 1$. Most of the previous proofs are valid in this case and we only point out necessary modifications. We need the following assumptions:

Assumption 7. *The diffusion coefficient $\sigma(x;\theta) = \sigma > 0$ is a constant (so that $a{=}\sigma^2 > 0$ is also constant) and $b(x;\theta)$ is jointly continuous in $(x,\theta) \in \mathbb{R}\times\Theta$.*

Assumption 8. *The coefficient $b(x;\theta)$ has the following property: for all $\theta\in\Theta$, there exists some constant $\alpha\in\mathbb{R}$ and $0<\beta<1$ such that*

$$\lim_{x\to\pm\infty}\frac{|b(x;\theta)-\alpha x|}{|x|^\beta+1} \leq C_b.$$

Or equivalently,

$$b(x;\theta) = \alpha x + b'(x;\theta),$$

where $b'(x;\theta) = O(|x|^{\beta}+1)$ *as* $x \to \pm\infty$.

To extend parametrix to linear case, we define

$$Y(t,x,y) = \frac{1}{(2\pi a f(t))^{\frac{1}{2}}} e^{-\frac{(y-xe^{\alpha t})^2}{2af(t)}},$$

where $f(t) = \frac{1}{2\alpha}(e^{2\alpha t}-1)$. Note that we have $f(t) = O(t)$ as $t \to 0$ and this is the main reason why most of the previous proofs carry over to this case. Now Y is the fundamental solution of

$$\frac{\partial Y}{\partial t} = \frac{1}{2} a \frac{\partial^2 Y}{\partial x^2} + \alpha x \frac{\partial Y}{\partial x}. \tag{1.30}$$

It is not hard to see that the corresponding stochastic process which generates this PDE is the Ornstein-Uhlenbeck process with constant diffusion coefficient:

$$dX_t = \alpha X_t dt + \sigma dB_t.$$

We also seek the fundamental solution in the form of (1.20), with Z replaced by Y. A similar calculation shows that the function $\psi(t,x,y;\theta)$ in (1.20) satisfies:

$$\psi(t,x,y;\theta) = L'Y(t,x,y;\theta) + L'Y * \psi(t,x,y;\theta),$$

where (notice a is a constant)

$$L'Y = b'(x;\theta)\frac{\partial Y(t,x,y)}{\partial x}.$$

As above, $b'(x;\theta)$ satisfies that for all $\theta \in \Theta$,

$$|b'(x,\theta)| \le C_{b'}(|x|^{\beta}+1).$$

Now the proof follows the same pattern in the last section and we just point out some modifications. First, the estimates for the derivatives of Y are:

$$\begin{aligned}\left|\frac{\partial Y}{\partial x}\right| &\leq C_{\lambda,1} f(t)^{-1} e^{-\frac{\lambda(y-xe^{\alpha t})^2}{2f(t)}}, &(1.31)\\ \left|\frac{\partial^2 Y}{\partial x^2}\right| &\leq C_{\lambda,2} f(t)^{-\frac{3}{2}} e^{-\frac{\lambda(y-xe^{\alpha t})^2}{2f(t)}}, &(1.32)\end{aligned}$$

for some $\lambda < a$. Next, in the proof of convergence of the series $\sum_{n=1}^{\infty}(L'Y)_n(t,x,y;\theta)$, we need to define $F(x)$ as

$$F(x) = |x|^{\beta} e^{-\frac{\varepsilon\lambda(y-xe^{\alpha t})^2}{2f(t)}}.$$

Therefore the extreme value analysis yields, for $t \in [0,T]$,

$$\begin{aligned}|F(x)| &\leq \left(\frac{|y|}{e^{\alpha t}} + \left(\frac{f(t)\beta}{\varepsilon\lambda e^{2\alpha t}}\right)^{\frac{1}{2}}\right)^{\beta}\\ &\leq \left(\frac{|y|}{e^{\alpha t}} + \left(\frac{f(T)\beta}{\varepsilon\lambda e^{2\alpha T}}\right)^{\frac{1}{2}}\right)^{\beta}\\ &\leq C\left(|y| + \left(\frac{f(T)\beta}{\varepsilon\lambda}\right)^{\frac{1}{2}}\right)^{\beta},\end{aligned}$$

for some constant C. In the second inequality we used the fact that $f(t)/e^{2\alpha t}$ achieves maximum at $t = T$. Thus

$$\begin{aligned}L'Y &\leq C\frac{1}{f(t)}\left[\left(|y| + \left(\frac{f(T)\beta}{\varepsilon\lambda}\right)^{\frac{1}{2}}\right)^{\beta} + 1\right] e^{-\frac{(1-\varepsilon)\lambda(y-xe^{\alpha t})^2}{2f(t)}}\\ &= C\frac{1}{f(t)} F_{\varepsilon\lambda}(y) e^{-\frac{(1-\varepsilon)\lambda(y-xe^{\alpha t})^2}{2f(t)}},\end{aligned}$$

where we define

$$F_{\varepsilon\lambda}(y) = \left(|y| + \left(\frac{f(T)\beta}{\varepsilon\lambda}\right)^{\frac{1}{2}}\right)^{\beta} + 1. \tag{1.33}$$

The key estimate for the convergence of ψ becomes, after an almost line by line repetition,

$$|(L'Y)_n|(t,x,y;\theta) \leq \quad C_\varepsilon^n [F_{\varepsilon_n \lambda^*}(y)]^n \frac{1}{\Gamma(\frac{n}{2})} \frac{1}{f(t)^{\frac{3-n}{2}}} e^{-\frac{\lambda^*(y-xe^{\alpha t})^2}{2f(t)}} \tag{1.34}$$

where $F_{\varepsilon_n \lambda^*}(y)$ is given by (1.33). The other estimate (1.28) takes the form

$$|\psi(t,x,y;\theta)| = \left| \sum_{n=1}^{\infty} (LZ)_n(t,x,y;\theta) \right| \leq C_h \frac{1}{f(t)} e^{-\frac{\lambda^*(y-xe^{\alpha t})^2}{2f(t)}} e^{h|y|} \tag{1.35}$$

The rest of the proofs also follows.

1.4.3 Generalization to a Singular Case

Consider the equation

$$dR_t = \frac{1}{R_t} dt + dW_t, \tag{1.36}$$

where $W(t)$ is one-dimensional Brownian Motion. This is a special case of Bessel process with index $\nu = \frac{1}{2}$. Bessel processes have important applications in mathematical finance, especially in modeling the term structure of interest rate. It is known that Bessel process (1.36) has transition density which is given by

$$p(t,x,y) = \frac{y}{t} \left(\frac{y}{x}\right)^{\frac{1}{2}} \exp\left(-\frac{x^2+y^2}{2t}\right) I_{\frac{1}{2}}\left(\frac{xy}{t}\right), \tag{1.37}$$

where I_ν is the modified Bessel function with index ν. If the starting point $R(0) = r_0 > 0$, then the process never reaches $r = 0$ for $t > 0$ so that the above transition density is defined for $t > 0, x > 0, y > 0$. In this section we only consider a positive starting point. We will investigate the transition density when we add a perturbation to this Bessel process in the drift term by the amount of $b(x;\theta)$, i.e. we are interested in the transition density of

$$dX_t = \left[\frac{1}{X_t} + b(X_t;\theta)\right] dt + dW_t. \tag{1.38}$$

with $X(0) = x_0 > 0$. Here $b(x;\theta)$ is a function which satisfies:

Assumption 9. *$b(x;\theta)$ is continuous on $[0,\infty)$ and $b(x;\theta) \geq 0$. We also have:*

1. *Growth condition near the boundary $x = 0$: as $x \downarrow 0$*

$$b(x;\theta) = O(x^{\frac{1}{2}}).$$

2. *Growth condition at infinity: there exists $\varepsilon \in (0, \frac{1}{2})$, such that when $x \to \infty$,*

$$b(x;\theta) = O(x^{\frac{1}{2}-\varepsilon}).$$

Due to Assumption 9, the comparison theorem of stochastic differential equation implies that the solution to (1.38) is strictly positive. Now the transition density of (1.38) is the fundamental solution of

$$\frac{\partial q}{\partial t} = \frac{1}{2}\frac{\partial^2 q}{\partial x^2} + \left[\frac{1}{x} + b(x;\theta)\right]\frac{\partial q}{\partial x}. \tag{1.39}$$

Let us use the expression $I_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}}\sinh(x)$ to rewrite (1.37) as

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left[e^{-\frac{(x-y)^2}{2t}} - e^{-\frac{(x+y)^2}{2t}}\right].$$

This is the density on which we will base our parametrix approximation. As before, we are attempting to obtain a transition density in the form

$$\Gamma(t,x,y;\theta) = p(t,x,y) + \int_0^t \int_0^\infty p(t-s,x,u)\psi(s,u,y;\theta)duds.$$

Similar to the calculations in previous sections, we have that the unknown $\psi(s,u,y;\theta)$ takes the form

$$\psi = \sum_{n=1}^{\infty}(Lp)_n(t,x,y;\theta),$$

provided that the series converges on compact subsets of $(0,T]\times\ (0,\infty)\times(0,\infty)$. Here

$$Lp(t,x,y;\theta) = b(x;\theta)\frac{\partial p}{\partial x}(t,x,y),$$

and

$$(Lp)_n(t,x,y;\theta) = \int_0^t \int_0^\infty Lp(t-s,x,u;\theta)(Lp)_{n-1}(s,u,y;\theta)duds.$$

These expressions suggest that we need an estimation of the derivative of $p(t,x,y)$ in order to prove the convergence. Now

$$\begin{aligned}\frac{\partial p}{\partial x}(t,x,y) &= -\frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left(\frac{1}{x}\right)\left[e^{-\frac{(x-y)^2}{2t}} - e^{-\frac{(x+y)^2}{2t}}\right] \\ &\quad \frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left[-\frac{x-y}{t}e^{-\frac{(x-y)^2}{2t}} + \frac{x+y}{t}e^{-\frac{(x+y)^2}{2t}}\right].\end{aligned}$$

Notice that the elementary inequality $e^{-a}-e^{-b} \le C_\alpha e^{-a}(b-a)^\alpha$ holds for $a<b$ when $0<\alpha<1$. Consequently, the first term

$$\begin{aligned}&\left|-\frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left(\frac{1}{x}\right)\left[e^{-\frac{(x-y)^2}{2t}} - e^{-\frac{(x+y)^2}{2t}}\right]\right| \\ \le\ & C_\alpha \frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left(\frac{1}{x}\right)e^{-\frac{(x-y)^2}{2t}}\left(\frac{2xy}{t}\right)^\alpha.\end{aligned}$$

Taking $\alpha = \frac{1}{2}$, we have

$$\left|-\frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left(\frac{1}{x}\right)\left[e^{-\frac{(x-y)^2}{2t}} - e^{-\frac{(x+y)^2}{2t}}\right]\right| \le C_1 \frac{1}{t}\left(\frac{y}{x}\right)^{\frac{3}{2}} e^{-\frac{(x-y)^2}{2t}}.$$

For the second term in $\frac{\partial p}{\partial x}$, it is easy to see that

$$\begin{aligned}&\left|\frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left[-\frac{x-y}{t}e^{-\frac{(x-y)^2}{2t}} + \frac{x+y}{t}e^{-\frac{(x+y)^2}{2t}}\right]\right| \\ \le\ & C_\lambda \frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left[e^{-\frac{\lambda(x-y)^2}{2t}} + e^{-\frac{\lambda(x+y)^2}{2t}}\right] \\ \le\ & C_\lambda \frac{2}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)e^{-\frac{\lambda(x-y)^2}{2t}},\end{aligned}$$

where $\lambda \in (0,1)$. Combining these calculations, we obtain that with $\lambda \in (0,1)$,

$$\left|\frac{\partial p}{\partial x}(t,x,y)\right| \leq C\frac{1}{t}\left[\left(\frac{y}{x}\right)^{\frac{3}{2}} + \left(\frac{y}{x}\right)\right] e^{-\frac{\lambda(x-y)^2}{2t}}, \tag{1.40}$$

for some appropriate constant C. Having this compact estimation of $\partial p/\partial x$, we next show that

Lemma 4. *Under Assumption 9, the following estimation holds:*

$$\begin{aligned} & (Lp)_k(t,x,y;\theta) \\ \leq\ & (C^*)^k b(x;\theta)\frac{\Gamma(\frac{1}{2})^k}{\Gamma(\frac{k}{2})} t^{\frac{k-3}{2}} e^{-\frac{\lambda^*(x-y)^2}{2t}} \left[\left(\frac{y}{x}\right) + \left(\frac{y}{x}\right)^{\frac{3}{2}} + \left(\frac{y}{x^{\frac{3}{2}}}\right) + \left(\frac{y^{\frac{3}{2}}}{x}\right)\right] \\ & \cdot [F_{\varepsilon_k\lambda^*}(y)]^k, \end{aligned}$$

where C^*, λ^* and ε_k are suitable constants and $F_{\varepsilon\lambda}(y) = \left(y + \sqrt{\frac{(1-\varepsilon)T}{\varepsilon\lambda}}\right)^{1-\varepsilon} + 1$.

Proof. We need some initial setup for this proof. Let us define $\varepsilon_1 = \varepsilon$ and λ, fixed. Choose $\{\varepsilon_i\}_{i=1}^{\infty}$ to be such that $\sum \varepsilon_i \leq 1$ from which the infinite product $\prod(1-\varepsilon_i)$ will be strictly greater than 0. We next define $\lambda_1 = (1-\varepsilon_1)\lambda$ and $\lambda_k = (1-\varepsilon_k)\lambda_{k-1}$. We also extend $b(x,\theta)$ for $x \in (-\infty, 0)$ in such a way that $b(x;\theta) = b(-x;\theta)$ when $x < 0$ and we call this function $\bar{b}(x;\theta)$.

Apparently the above inequality holds for $k = 1$. We assume that we already have

$$\begin{aligned} & (Lp)_k(t,x,y;\theta) \\ \leq\ & C_1^k b(x;\theta)\frac{\Gamma(\frac{1}{2})^k}{\Gamma(\frac{k}{2})} t^{\frac{k-3}{2}} e^{-\frac{\lambda_k(x-y)^2}{2t}} \left[\left(\frac{y}{x}\right) + \left(\frac{y}{x}\right)^{\frac{3}{2}} + \left(\frac{y}{x^{\frac{3}{2}}}\right) + \left(\frac{y^{\frac{3}{2}}}{x}\right)\right] \\ & \prod_{i=1}^{k} F_{\varepsilon_i\lambda_i}(y), \end{aligned}$$

then

$$
\begin{aligned}
|(Lp)_{k+1}(t,x,y;\theta)| &= \left|\int_0^t\int_0^\infty (Lp)(t-s,x,u;\theta)(Lp)_k(s,u,y;\theta)duds\right| \\
&\le |b(x;\theta)|\prod_{i=1}^k F_{\varepsilon_i\lambda_i}(y)\int_0^t\int_0^\infty C\frac{1}{t-s}\left[\left(\frac{u}{x}\right)^{\frac{3}{2}}+\left(\frac{u}{x}\right)\right]e^{-\frac{\lambda_1(x-y)^2}{2(t-s)}}b(u;\theta) \\
&\quad \cdot C_1^k b(u;\theta)\frac{\Gamma(\frac{1}{2})^k}{\Gamma(\frac{k}{2})}s^{\frac{k-3}{2}}e^{-\frac{\lambda_k(u-y)^2}{2s}}\left[\left(\frac{y}{u}\right)+\left(\frac{y}{u}\right)^{\frac{3}{2}}+\left(\frac{y}{u^{\frac{3}{2}}}\right)+\left(\frac{y^{\frac{3}{2}}}{u}\right)\right]duds \\
&= F_1+F_2+F_3+F_4+F_5+F_6+F_7+F_8,
\end{aligned}
$$

where F_i are the corresponding terms after expanding the expression in the square bracket, with eight terms in total. Let us compute each term.

$$
\begin{aligned}
F_1 &= CC_1^k b(x;\theta)\prod_{i=1}^k F_{\varepsilon_i\lambda_i}(y)\frac{\Gamma(\frac{1}{2})^k}{\Gamma(\frac{k}{2})}\int_0^t\int_0^\infty \frac{1}{t-s}\left(\frac{u}{x}\right)^{\frac{3}{2}}e^{-\frac{\lambda_1(x-y)^2}{2(t-s)}}b(u;\theta) \\
&\quad \cdot\frac{y}{u}s^{\frac{k-3}{2}}e^{-\frac{\lambda_k(u-y)^2}{2s}}duds \\
&\le CC_1^k b(x;\theta)\prod_{i=1}^k F_{\varepsilon_i\lambda_i}(y)\frac{\Gamma(\frac{1}{2})^k}{\Gamma(\frac{k}{2})}\int_0^t\int_{-\infty}^\infty \frac{1}{t-s}\left|\frac{u}{x}\right|^{\frac{3}{2}}e^{-\frac{\lambda_1(x-y)^2}{2(t-s)}}\bar{b}(u;\theta) \\
&\quad \cdot\left|\frac{y}{u}\right|s^{\frac{k-3}{2}}e^{-\frac{\lambda_k(u-y)^2}{2s}}duds \\
&\le CC_1^k b(x;\theta)\frac{\Gamma(\frac{1}{2})^k}{\Gamma(\frac{k}{2})}\left(\frac{y}{x^{\frac{3}{2}}}\right)\frac{\Gamma(\frac{1}{2})\Gamma(\frac{k}{2})}{\Gamma(\frac{k+1}{2})}t^{\frac{k-2}{2}}e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}}\left(\frac{2\pi}{\lambda}\right)^{\frac{1}{2}}\prod_{i=1}^{k+1}F_{\varepsilon_i\lambda_i}(y) \\
&= CC_1^k b(x;\theta)\left(\frac{2\pi}{\lambda}\right)^{\frac{1}{2}}\frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})}\left(\frac{y}{x^{\frac{3}{2}}}\right)t^{\frac{k-2}{2}}e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}}\prod_{i=1}^{k+1}F_{\varepsilon_i\lambda_i}(y) \\
&\le C_2^k b(x;\theta)\frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})}\left(\frac{y}{x^{\frac{3}{2}}}\right)t^{\frac{k-2}{2}}e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}}\prod_{i=1}^{k+1}F_{\varepsilon_i\lambda_i}(y),
\end{aligned}
$$

where C_2 is some suitable constant. In the second inequality above, we used a result in Lemma 1

by noticing that $|u|^{\frac{1}{2}}\bar{b}(u;\theta) = O(|u|^{1-\varepsilon^*})$ as $u \to \infty$. Similar computation implies:

$$
\begin{aligned}
F_2 &\leq C_2^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} \left(\frac{y}{x}\right)^{\frac{3}{2}} t^{\frac{k-2}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y), \\
F_3 &\leq C_2^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} \left(\frac{y}{x}\right)^{\frac{3}{2}} t^{\frac{k-2}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y), \\
F_4 &\leq C_2^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} \left(\frac{y^{\frac{3}{2}}}{x}\right) t^{\frac{k-2}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y), \\
F_5 &\leq C_2^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} \left(\frac{y}{x}\right) t^{\frac{k-2}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y), \\
F_6 &\leq C_2^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} \left(\frac{y}{x^{\frac{3}{2}}}\right) t^{\frac{k-2}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y), \\
F_7 &\leq C_2^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} \left(\frac{y}{x^{\frac{3}{2}}}\right) t^{\frac{k-2}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y), \\
F_8 &\leq C_2^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} \left(\frac{y}{x}\right) t^{\frac{k-2}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y).
\end{aligned}
$$

The reason that we need as $b(u;\theta) = O(u^{\frac{1}{2}})$ when $u \downarrow 0$ is that we need to remove the singularity from $1/u^{\frac{1}{2}}$ term in the integral. After adjusting the constant C_2 and absorbing a factor 2 (i.e. two identical terms), we see that

$$
\begin{aligned}
& (Lp)_{k+1}(t,x,y;\theta) \\
\leq\ & (C^*)^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} t^{\frac{k+1-3}{2}} e^{-\frac{\lambda_{k+1}(x-y)^2}{2t}} \left[\left(\frac{y}{x}\right) + \left(\frac{y}{x}\right)^{\frac{3}{2}} + \left(\frac{y}{x^{\frac{3}{2}}}\right) + \left(\frac{y^{\frac{3}{2}}}{x}\right)\right] \\
& \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda_i}(y) \\
\leq\ & (C^*)^{k+1} b(x;\theta) \frac{\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} t^{\frac{k+1-3}{2}} e^{-\frac{\lambda^*(x-y)^2}{2t}} \left[\left(\frac{y}{x}\right) + \left(\frac{y}{x}\right)^{\frac{3}{2}} + \left(\frac{y}{x^{\frac{3}{2}}}\right) + \left(\frac{y^{\frac{3}{2}}}{x}\right)\right] \\
& \prod_{i=1}^{k+1} F_{\varepsilon_i\lambda^*}(y).
\end{aligned}
$$

where $\lambda^* = \lambda \prod_{i=1}^{\infty}(1-\varepsilon_i) > 0$. □

This lemma enables us to use the argument in Lemma 1 to derive the convergence of the

series $\psi = \sum_{n=1}^{\infty}(Lp)_n(t,x,y;\theta)$ on compact subsets of $(0,T]\times\ (0,\infty)\times(0,\infty)$. Moreover, this convergence is uniform in θ.

We next look at the case $\nu = \frac{3}{2}$. We need the following expression:

$$I_{\frac{3}{2}}(x) = \sqrt{\frac{2}{\pi x}}\left(\cosh x - \frac{1}{x}\sinh x\right).$$

As a consequence,

$$\begin{aligned} p(t,x,y) &= \frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left[e^{-\frac{(x-y)^2}{2t}} + e^{-\frac{(x+y)^2}{2t}}\right] \\ &\quad + \frac{\sqrt{t}}{\sqrt{2\pi}}\left(\frac{1}{x^2}\right)\left[e^{-\frac{(x+y)^2}{2t}} - e^{-\frac{(x-y)^2}{2t}}\right]. \end{aligned}$$

According as the previous arguments, we compute the partial derivative:

$$\begin{aligned} \frac{\partial p}{\partial x} &= -\frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x^2}\right)\left[e^{-\frac{(x-y)^2}{2t}} + e^{-\frac{(x+y)^2}{2t}}\right] \\ &\quad + \frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)\left[e^{-\frac{(x-y)^2}{2t}}\left(-\frac{x-y}{t}\right) - e^{-\frac{(x+y)^2}{2t}}\left(\frac{x+y}{t}\right)\right] \\ &\quad + \frac{\sqrt{t}}{\sqrt{2\pi}}\left(\frac{-2}{x^3}\right)\left[e^{-\frac{(x+y)^2}{2t}} - e^{-\frac{(x-y)^2}{2t}}\right] \\ &\quad + \frac{\sqrt{t}}{\sqrt{2\pi}}\left(\frac{1}{x^2}\right)\left[e^{-\frac{(x+y)^2}{2t}}\left(-\frac{x+y}{t}\right) + e^{-\frac{(x-y)^2}{2t}}\left(\frac{x-y}{t}\right)\right]. \end{aligned}$$

Therefore, we can derive an estimate:

$$\begin{aligned} \left|\frac{\partial p}{\partial x}\right| &\leq \frac{2}{\sqrt{2\pi t}}\left(\frac{y}{x^2}\right)e^{-\frac{(x-y)^2}{2t}} + \frac{1}{\sqrt{2\pi t}}\left(\frac{y}{x}\right)e^{-\frac{\lambda(x-y)^2}{2t}} \\ &\quad + \frac{\sqrt{t}}{\sqrt{2\pi}}\left(\frac{4}{x^3}\right)e^{-\frac{(x-y)^2}{2t}} + \frac{\sqrt{t}}{\sqrt{2\pi}}\left(\frac{1}{x^2}\right)e^{-\frac{\lambda(x-y)^2}{2t}} \\ &\leq Ce^{-\frac{\lambda(x-y)^2}{2t}}\left[\frac{1}{\sqrt{t}}\left(\frac{y}{x^2}\right) + \frac{1}{\sqrt{t}}\left(\frac{y}{x}\right) + \sqrt{t}\left(\frac{1}{x^3}\right) + \sqrt{t}\left(\frac{1}{x^2}\right)\right]. \end{aligned}$$

Notice that this estimate is similar to the case $\nu = \frac{1}{2}$. A tedious computation will show that the

series

$$\psi = \sum_{n=1}^{\infty} (Lp)_n(t,x,y;\theta),$$

where

$$Lp(t,x,y;\theta) = b(x;\theta)\frac{\partial p}{\partial x}(t,x,y),$$

converges on compact subsets of $(0,T] \times \ (0,\infty) \times (0,\infty)$, uniformly in θ provided that $b(x;\theta)$ satisfies the following:

1. $b(x;\theta) \geq 0$;

2. Growth condition near the boundary $x = 0$: when $x \to 0$,

$$b(x;\theta) = O(x^3).$$

3. Growth condition at infinity: there exists $\beta \in (0,1)$, such that when $x \to \infty$,

$$b(x;\theta) = O(x^{\beta}).$$

1.5 Application of Parametrix Approximation

In this section, we will apply the parametrix approximation to the parameter estimation of SDE (1.1). We will prove the results for $d = 1$. i.e. the scaler SDE, although these results can be generalized into higher dimensions without difficulty.

1.5.1 Preliminaries

Under Assumptions 1 and 2, there exists a time-homogeneous transition kernel $P(t,x,A;\theta)$ for (1.1). It has a density with respect to the Lebesgue measure λ, which is denoted by p:

$$P(t,x,A;\theta) = \int_A p(t,x,y;\theta)dy,$$

i.e. p is the fundamental solution of (1.16). In this subsection, we assume that the exact expression for p is known for the statistical inference on θ.

Let $(x_0, x_1, \ldots, x_n)$ be an observation of $(X_0, X_\Delta, \ldots, X_{n\Delta})$ from the process (1.1) with an initial distribution μ. Here we require the observations be equidistant. We denote the skeleton process $(X_0, X_\Delta, \ldots, X_{n\Delta}, \ldots)$ by $(X_0, X_1, \ldots, X_n, \ldots)$ for notational simplicity. Note that (1.1) requires the starting point to be fixed x_0. However it can be replaced by any initial distribution μ without affecting the convergence results.

Define $\mathscr{F}_n = \sigma(X_0, X_1, \ldots, X_n)$, the σ-algebra generated by $(X_0, X_1, \ldots, X_n)$ and let $P_{\theta,\mu}$ be the law of (1.1) with the parameter θ and initial distribution μ. That is,

$$P_{\theta,\mu}(X_0 \in A_0, \ldots, X_n \in A_n) = \int_{A_0 \times \cdots \times A_n} \mu(dx_0) \prod_{i=1}^{n} p(\Delta, x_{i-1}, x_i; \theta) dx_1 \ldots dx_n.$$

Then we have a dominated statistical model in the sense of [19]. We can define the likelihood function to be

$$L_n(\theta) = \mu(X_0) \prod_{i=1}^{n} p(\Delta, X_{i-1}, X_i; \theta).$$

We assume that μ has a density ϕ with respect to Lebesgue measure so that

$$d\mu(x) = \phi(x)dx.$$

Then we have the corresponding log-likelihood

$$l_n(\theta) = \ln[L_n(\theta)] = \ln \phi(x_0) + \sum_{i=1}^{n} \ln p(\Delta, X_{i-1}, X_i; \theta).$$

It can be seen that the first term will be dominated by the summation as n approaches infinity. Since we always consider large samples for statistical inference, it is convenient to drop the first term and denote

$$l_n(\theta) = \ln[L_n(\theta)] = \sum_{i=1}^{n} \ln p(\Delta, X_{i-1}, X_i; \theta). \tag{1.41}$$

The maximum likelihood estimator is defined to be the maximizer of $l_n(\theta)$:

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmax}}[l_n(\theta)].$$

It is the solution of the equation

$$\frac{\partial}{\partial \theta_i} l_n(\theta) = 0, \quad i = 1, 2, \ldots, N_\theta \tag{1.42}$$

where N_θ is the number of parameters. It may happen that the above equation have more than one solution. Moreover, it could also be that the solution is only a local maximum, rather than absolute maximum. In these cases, things become more complicated. However, in many applications the solution is unique and we will not consider these matters in this thesis.

We notice that when the observations are discrete, we obtain the skeleton of the process X_t as a Markov chain. The properties of X_t is also inherited by this chain. Most statistical inferences on a Markov process rely on the Markov chain which is Harris recurrent and has an invariant probability measure. For completeness, we give these definitions.

Definition 1. *A σ-finite measure μ is called an invariant measure for a Markov chain $\{X_n\}$ with transition kernel P if*

$$\mu(A) = \int_{\mathbb{R}} \mu(dx) P(x, A).$$

Definition 2. *A Markov Chain X_n defined on the canonical space $(\mathbb{R}^n, \mathscr{B}^n)$ is called Harris chain if there exists a σ-finite, invariant measure μ such that $\mu(A) > 0$ implies:*

$$P_x \left[\sum_{n=1}^{\infty} I_A(X_n) = \infty \right] = 1.$$

for any $x \in \mathbb{R}$. Here P_x denote the probability law when the chain starts from x. The chain is called positive Harris if the measure μ can be made a probability measure (i.e. μ is a finite measure so that it can be normalized).

One striking fact about the positive Harris chain is that it has the following law of large numbers (LLN) and central limit theorem (CLT), even though the random variables in the chain are correlated.

Theorem 3. *(Law of Large Numbers) Let the Markov chain $\{X_n\}$ with transition kernel P be positive Harris. Let F be a function on $\mathbb{R}^2$ such that*

$$\int_{\mathbb{R}^2} F(x_1,x_2)\mu(dx_1)P(x_1,dx_2) < \infty,$$

where μ is the invariant probability measure, then

$$\frac{1}{n}\sum_{i=1}^{n} F(X_{i-1},X_i) \longrightarrow \int_{\mathbb{R}^2} F(x_1,x_2)\mu(dx_1)P(x_1,dx_2), \quad P_{\theta,\mu}-a.s.$$

The convergence also holds in $L^1(P_{\theta,\mu})$.

Theorem 4. *(Central Limit Theorm) Assume the function F in Theorem 3 satisfies*

$$\int_{\mathbb{R}^2} F^2(x_1,x_2)\mu(dx_1)P(x_1,dx_2) < \infty.$$

Then

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(F(X_{i-1},X_i) - PF(X_{i-1})) \longrightarrow N(0,\sigma^2(F)),$$

where

$$PF(x) = \int_{\mathbb{R}} P(x,dy)F(x,y),$$

and

$$\sigma^2(F) = \int_{\mathbb{R}^2} F^2(x_1,x_2)\mu(dx_1)P(x_1,dx_2) - \int_{\mathbb{R}} (PF)^2 d\mu.$$

Since the observations are dependent, the proofs for these theorems require special treatment unlike those for the classic LLN and CLT. In Appendix 1, we will give a self-contained proof of Theorem 3. The result there is more general than Theorem 3. For a proof of Theorem 4, we refer

the readers to [19] and will omit it here. It turns out that the law of large numbers and central limit theorem are key to proving the consistency and convergence rate of the maximum likelihood estimator. These properties of MLE make it our primary choice for making statistical estimations.

1.5.2 Asymptotic Behaviors of Approximated Estimators

Since the exact transition probability density is often unknown for most diffusions, we will employ the approximated transition density as a substitute. As a consequence, the maximizer from the approximation is also a substitute for the true MLE. In this section, we will adopt the parametrix approximation for the fundamental solutions of (1.16). More precisely, suppose that we have the approximated transition density $p^{(N)}(t,x,y;\theta)$ of true $p(t,x,y;\theta)$, then we define the approximated likelihood function

$$L_n^{(N)}(\theta) = \prod_{i=1}^{n} p^{(N)}(\Delta, X_{i-1}, X_i; \theta).$$

and approximated log-likelihood

$$l_n^{(N)}(\theta) = \ln[L_n^{(N)}(\theta)] = \sum_{i=1}^{n} \ln p^{(N)}(\Delta, X_{i-1}, X_i; \theta). \tag{1.43}$$

The approximated maximum likelihood estimator is defined as

$$\hat{\theta}_n^{(N)} = \underset{\theta\in\Theta}{\operatorname{argmax}}[l_n^{(N)}(\theta)]. \tag{1.44}$$

We would like to show that when n and N approach infinity, we can find subsequences n_m and N_m such that $\hat{\theta}_{n_m}^{(N_m)}$ has the desired consistency when $m \to \infty$ under certain conditions and we would also be interested in its convergence rate.

To show these results, we need to know the properties of $\hat{\theta}_n$, the true MLE. We will return to the Harris chain $\{X_n\}$ with transition density $p(x,y;\theta)$ and invariant distribution μ, as discussed in the previous section and assume the following conditions:

Assumption 10. *Given the Harris chain $\{X_n\}$. For any x, the set of y such that $p(x,y;\theta) > 0$ is*

independent of θ. Assume that

$$f(x,y;\theta) = \ln p(x,y;\theta)$$

is well defined, and $\partial f/\partial\theta_i, \partial^2 f/(\partial\theta_i\partial\theta_j)$, $\partial^3 f/(\partial\theta_i\partial\theta_j\partial\theta_k)$ are well defined and continuous in θ. For $\theta\in\Theta$, there exists a neighborhood N_b of θ such that

$$E_{\theta,\mu}\left[\sup_{\theta'\in N_b}\left|\frac{\partial^3 f}{\partial\theta_i\partial\theta_j\partial\theta_k}(x,y;\theta')\right| < \infty\right]$$

where $E_{\theta,\mu}$ is the expectation under θ when starting from the invariant distribution μ.

Assumption 11. *We assume for $i = 1,2\ldots,N_\theta$*

$$E_{\theta,\mu}\left[\frac{\partial f}{\partial\theta_i}(X_1,X_2;\theta)\right]^2 < \infty,$$

and if

$$I_{ij} = E_{\theta,\mu}\left[\frac{\partial f}{\partial\theta_i}(X_1,X_2;\theta)\frac{\partial f}{\partial\theta_j}(X_1,X_2;\theta)\right],$$

then the Fisher's information matrix $I = \{I_{ij}\}$ is positive definite.

We need the following lemma. The proof is a simple application of fixed point theorem, which can be found in [1].

Lemma 5. *Let f be continuous and maps Θ into Θ. Suppose that it satisfies the following property: for any θ with $|\theta|\le\delta$, we have $\theta\cdot f(\theta) < 0$, then there exists $\hat\theta$ such that $|\hat\theta| < \delta$ and $f(\hat\theta) = 0$*

We have the following proposition:

Proposition 1. *Assume that the Harris chain $\{X_n\}$ with transition density $p(x,y;\theta)$ and invariant probability measure μ satisfies Assumption 10 and 11, then there exists a sequence of maximum likelihood estimators $\hat\theta_n$, which is a solution to (1.42), such that $\hat\theta_n$ converges to the true parameter θ_0 under $E_{\theta_0,\mu}$, i.e. $\hat\theta_n$ is a consistent estimator.*

Proof. The proof we give here is only sketchy. For simplicity, we denote $\partial f/\partial\theta_i$ by f_i. Analogously we have f_{ij} and f_{ijk}. We also drop the invariant measure μ and denote the expectation by

E_θ. By Assumption 10, the interchange of differentiation and integral is permissible. As a result, we have

$$E_\theta \left[f_i(X_1, X_2; \theta) \right] = 0. \tag{1.45}$$

The Fisher's information matrix $\{I_{ij}\}$ has ij-th entry

$$I_{ij} = E_\theta \left[f_i(X_1, X_2; \theta) f_j(X_1, X_2; \theta) \right] = -E_\theta \left[f_{ij}(X_1, X_2; \theta) \right].$$

By Taylor expansion, for $\theta \in N_b$,

$$\begin{aligned} f_i(x_m, x_{m+1}; \theta) &= f_i(x_m, x_{m+1}; \theta_0) + \sum_{j=1}^{N_\theta} (\theta_j - \theta_{0,j}) f_{ij}(x_m, x_{m+1}; \theta_0) \\ &\quad + \frac{1}{2} \sum_{j,k} (\theta_{jk} - \theta_{0,jk})^T f_{ijk}(x_m, x_{m+1}; \bar{\theta}_0)(\theta_{jk} - \theta_{0,jk}) \\ &= f_i(x_m, x_{m+1}; \theta_0) + \sum_{j=1}^{N_\theta} (\theta_j - \theta_{0,j}) f_{ij}(x_m, x_{m+1}; \theta_0) \\ &\quad + \frac{1}{2} \rho |\theta - \theta_0|^2 F(x_m, x_{m+1}), \end{aligned}$$

where

$$F(x_m, x_{m+1}) = \sup_{\theta' \in N_b} |f_{ijk}(x_m, x_{m+1}; \theta')|,$$

and $\bar{\theta}_0$ is some point on the segment connecting θ and θ_0. Moreover, ρ is a positive number with $\rho \leq 1$. It follows that

$$\begin{aligned} \frac{1}{n} \frac{\partial}{\partial \theta_i} l_n(\theta) &= \frac{1}{n} \sum_{m=1}^{n} f_i(x_m, x_{m+1}; \theta_0) + \frac{1}{n} \sum_{m=1}^{n} \sum_{j=1}^{N_\theta} (\theta_j - \theta_{0,j}) f_{ij}(x_m, x_{m+1}; \theta_0) \\ &\quad + \frac{1}{2n} \rho |\theta - \theta_0|^2 \sum_{m=1}^{n} F(x_m, x_{m+1}) \end{aligned} \tag{1.46}$$

Notice that by Theorem 3, Assumption 10 and (1.45), we have for $i = 1, 2 \ldots, N_\theta$,

$$\lim_{n \to \infty} (P_{\theta_0}) \frac{1}{n} \sum_{m=1}^{n} f_i(x_m, x_{m+1}; \theta_0) = 0,$$

$$\lim_{n\to\infty}(P_{\theta_0})\frac{1}{2n}\sum_{m=1}^{n}|F(x_m,x_{m+1})| = M,$$

and

$$\lim_{n\to\infty}(P_{\theta_0})\frac{1}{n}\sum_{m=1}^{n} f_{ij}(x_m,x_{m+1};\theta_0) = -I_{ij}.$$

where M is suitable constant. Since $\{I_{ij}\}$ is strictly positive definite, we can find some K such that

$$\sum I_{ij}u_iu_j \geq K,$$

for any u with $|u| = 1$.

As a consequence, for any ε, we can choose n large enough so that with probability P_{θ_0} greater than $1-\varepsilon$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{m=1}^{n} f_i(x_m,x_{m+1};\theta_0) < \delta^2,$$

$$\lim_{n\to\infty}\frac{1}{2n}\sum_{m=1}^{n}|F(x_m,x_{m+1})| < M+1,$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{m=1}^{n} f_{ij}(x_m,x_{m+1};\theta_0) + I_{ij} < \delta,$$

where δ is chosen to be such that

$$\delta < K/3N_\theta^2(M+1).$$

Thus by (1.46), when $|\theta-\theta_0| \leq \delta$,

$$\begin{aligned}
& \left|\frac{1}{n}\sum_{m=1}^{n} f_i(x_m,x_{m+1};\theta) + \sum_{j=1}^{n} I_{ij}(\theta_j-\theta_{0,j})\right| \\
\leq\ & \left|\frac{1}{n}\sum_{m=1}^{n} f_i(x_m,x_{m+1};\theta_0)\right| + \left|\frac{1}{n}\sum_{m=1}^{n}\sum_{j=1}^{N_\theta}(\theta_j-\theta_{0,j})f_{ij}(x_m,x_{m+1};\theta_0) + \sum_{j=1}^{n} I_{ij}(\theta_j-\theta_{0,j})\right| \\
& + \left|\frac{1}{2n}\rho|\theta-\theta_0|^2\sum_{m=1}^{n} F(x_m,x_{m+1})\right| \\
\leq\ & \delta^2 + \delta N_\theta|\theta-\theta_0| + N_\theta^2|\theta-\theta_0|^2(M+1) \\
\leq\ & 3N_\theta^2\delta^2(M+1).
\end{aligned}$$

It follows that when $|\theta-\theta_0|=\delta$,

$$\begin{aligned}
&\sum_{i=1}^{N_\theta}\left[\frac{1}{n}\sum_{m=1}^{n} f_i(x_m,x_{m+1};\theta)\right](\theta_i-\theta_{0,i})\\
\leq\ & -\sum_{i,j=1}^{N_\theta} I_{ij}(\theta_i-\theta_{0,i})(\theta_j-\theta_{0,j})+3N_\theta^2\delta^2(M+1)\\
\leq\ & -K|\theta-\theta_0|+3N_\theta^2\delta^2(M+1)\\
=\ & -K\delta+3N_\theta^2\delta^2(M+1)\\
<\ & 0,
\end{aligned}$$

by the choice of δ. The proof is finished by a direct application of Lemma 5. □

We can now discuss the convergence rate of $\hat{\theta}_n$ after deriving its consistency. We have the following:

Proposition 2. *Assuming the same conditions as Proposition 1, then the sequence $\sqrt{n}(\hat{\theta}_n-\theta_0)$ converges in law under P_{θ_0} to the Gaussian distribution $N(0,I^{-1}(\theta_0))$ and the estimator $\hat{\theta}_n$ is asymptotically efficient.*

Proof. Since $\hat{\theta}_n$ is the solution to (1.42), we have

$$0=\frac{\partial l_n}{\partial\theta_i}(\hat{\theta}_n) \quad = \quad \frac{\partial l_n}{\partial\theta_i}(\theta_0)+\sum_{j=1}^{N_\theta}\frac{\partial^2 l_n}{\partial\theta_i\partial\theta_j}(\bar{\theta}_n)(\hat{\theta}_{n,j}-\theta_{0,j}).$$

where $\bar{\theta}_n$ is a point on the segment connecting θ_0 and $\hat{\theta}_n$. Denote $\partial l_n/\partial\theta_i(\theta_0)$ by T_i^n and let $T^n(\theta_0)=(T_1^n,\ldots,T_{N_\theta}^n)$, then

$$0=\frac{1}{\sqrt{n}}T_i^n+\sum_{j=1}^{N_\theta}\sqrt{n}(\hat{\theta}_{n,j}-\theta_{0,j})\frac{1}{n}\frac{\partial^2 l_n}{\partial\theta_i\partial\theta_j}(\bar{\theta}_n).$$

By (1.45) and Theorem 4, we have under P_{θ_0},

$$T^n(\theta_0) \to N(0, I(\theta_0)).$$

where the convergence is in law under P_{θ_0}. Due to the fact that $\hat{\theta}_n \to \theta_0$ in probability and Assumptions 10 and 11,

$$\frac{1}{n}\frac{\partial^2 l_n}{\partial \theta_i \theta_j}(\bar{\theta}_n) \to -I_{ij}(\theta_0), \quad P_{\theta_0} - \text{a.s.}$$

As a consequence, by Slutsky's theorem we derive

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \longrightarrow N(0, I^{-1}(\theta_0) I(\theta_0) I^{-1}(\theta_0)) = N(0, I^{-1}(\theta_0)),$$

where the convergence is under P_{θ_0}. The asymptotic efficiency is obvious. □

We will go back to our estimation problem for equation (1.1). It has been proved in [20] that the ergodic diffusion satisfies Assumptions 10 and 11, provided the following assumption holds for $b(x;\theta)$:

Assumption 12. *b is three times continuously differentiable in θ. There exists $\gamma > 0$ such that for $i = 1,2,3$ and $j = 1,2$,*

$$\frac{\partial^i}{\partial x^i}\left(\frac{\partial^j b}{\partial \theta^j}\right) = O(|b|^\gamma(x;\theta_0)), \quad as\ |x| \to \infty.$$

It should be pointed out that the above assumption is not restrictive and it is usually satisfied by an ergodic diffusion with multiplicative parameters. We remind the reader that the ergodicity of (1.1) is equivalent to

$$\lim_{|x|\to\infty} \int_0^x e^{-\frac{2\int_0^u b(v;\theta)dv}{\sigma^2(u;\theta)}} du = \pm\infty \tag{1.47}$$

and

$$\lim_{|x|\to\infty} \int_0^x e^{\frac{2\int_0^u b(v;\theta)dv}{\sigma^2(u;\theta)}} du < \infty \tag{1.48}$$

Common ergodic diffusions include OU process and CIR process. Thus we arrive at the conclusion:

Theorem 5. *The maximum likelihood estimator $\hat{\theta}_n$ for (1.1) is consistent and asymptotically normal, provided that (1.1) is an ergodic diffusion which satisfies Assumption 12.*

We now study the asymptotic behavior for the approximated MLE (1.44). For simplicity, we study the parametrix approximation under sub-linear growth. Recall that we define

$$p^{(N)}(t,x,y;\theta) = Z(t,x,y;\theta) + \int_0^t \int_{\mathbb{R}} Z(t-s,x,u;\theta)\psi^{(N)}(s,u,y;\theta),$$

where

$$\psi^{(N)} = \sum_{i=1}^{N} (LZ)_i(t,x,y;\theta),$$

with L defined in (1.22). It follows that $p^{(N)}(t,x,y;\theta)$ converges point-wise to the true (but unknown) transition density $p(t,x,y;\theta)$, uniformly in θ. The pointwise convergence with uniformity in θ is crucial to us since it enables us prove the convergence of the maximizer, see Lemma 6 below. Recall that we define $\hat{\theta}_n^{(N)}$ as the maximizer of the approximated log-likelihood (1.43). Our goal is to show that $\hat{\theta}_n^{(N)}$ shares the same asymptotic property as $\hat{\theta}_n$ at least for particularly chosen subsequences n_m and N_m. That is, if $\hat{\theta}_n$ is consistent and asymptotically normal, so is $\hat{\theta}_{n_m}^{(N_m)}$. For this, we need the following lemma.

Lemma 6. *Assume that a sequence of functions $\{f_n(x)\}_{n=1}^{\infty}$ converges point-wise to $f(x)$ on a compact set A. In addition, we assume the following conditions hold:*
(1) $f_n(x)$ and $f(x)$ are continuous on A;
(2) Define

$$w_n(\alpha) = sup\{|f_n(x) - f_n(y)|; |x-y| < \alpha\}.$$

There exist sequences (α_k) and (ε_k) such that $\alpha_k \to 0$, $\varepsilon_k \to 0$ and for all k, we have

$$w_n(\alpha_k) \leq \varepsilon_k, \quad as \quad n \to \infty.$$

If x_n^* and x^* are the unique minimizers of f_n and f on A, then we have $\lim_{n\to\infty} x_n^* = x^*$

Proof. Without loss of generality, we assume $f(x^*) = 0$. For any open ball B centered at x^*, we can find an ε such that $f(x) \geq 2\varepsilon$ on A/B by continuity. Since $\varepsilon_k \to 0$, we can assume $\varepsilon_k < \varepsilon$ for all k. Now for the α_k given above, we can cover A/B by finite number of open balls B_i, $1 \leq i \leq M$, whose center is x_i and radius is α_k.

Let us take $x \in B_i$, then

$$f_n(x) \geq f_n(x_i) - |f_n(x) - f_n(x_i)|.$$

This implies

$$\inf_{x\in B_i} f_n(x) \geq f_n(x_i) - \sup_{x\in B_i} |f_n(x) - f_n(x_i)| = f_n(x_i) - w_n(\alpha_k).$$

Therefore, by combining all the open balls B_i together,

$$\inf_{x\in A/B} f_n(x) \geq \inf_{1\leq i\leq M} f_n(x_i) - w_n(\alpha_k).$$

Suppose that to the contrary, $x_n^* \in A/B$ when $n \geq N$ where N is some large enough integer, then we have

$$\inf_{x\in A/B} f_n(x) < f_n(x^*).$$

This yields

$$\inf_{1\leq i\leq M} f_n(x_i) - w_n(\alpha_k) < f_n(x^*).$$

Thus

$$\inf_{1\leq i\leq M} f_n(x_i) - f_n(x^*) < w_n(\alpha_k) \leq \varepsilon_k < \varepsilon.$$

Letting $n \to \infty$ and noticing that $f(x^*) = 0$, we derive

$$\inf_{1\leq i\leq M} f(x_i) < \varepsilon,$$

which is a contradiction since $f(x) \geq 2\varepsilon$ on A/B. □

Corollary 1. *Suppose that f_n is a sequence of functions that are continuous on the compact set A and $f_n \to f$ uniformly. If x_n^* and x^* are the unique minimizers of f_n and f on A, then we have* $\lim_{n\to\infty} x_n^* = x^*$.

Applying the lemma to the approximated MLE, we have

Corollary 2. *Let (1.1) be an ergodic diffusion which satisfies Assumptions 3-6 and Assumption 12. Suppose for fixed n, $\hat{\theta}_n^{(N)}$ and $\hat{\theta}_n$ are the unique minimizers of (1.41) and (1.43) respectively. Then for any $\theta_0 \in \Theta$, $\hat{\theta}_n^{(N)}$ converges to $\hat{\theta}_n$ in P_{θ_0} probability as $N \to \infty$*

Proof. It has been shown that $p^{(N)}(t,x,y;\theta)$ converges to $p(t,x,y;\theta)$ point-wise in (t,x,y) and uniformly in θ. Since by the conditions of the corollary, $p^{(N)}(t,x,y;\theta)$ are continuous in θ, the same is true for $p(t,x,y;\theta)$. As a consequence, $l_n^{(N)}(\theta)$ converges to $l_n(\theta)$ uniformly. Thus the conclusion is immediate from Corollary 1. □

In order to obtain the asymptotic property of $\hat{\theta}_n^{(N)}$ we invoke a simple lemma regarding the convergence in probability and distribution. This is proved in [46].

Lemma 7. *Let $\{\xi_n^{(N)}\}_{n,N=1}^{\infty}$ and $\{\xi_n\}_{n=1}^{\infty}$ be two sequences of random variables taking values in a metric space X. Suppose for fixed $n \in \mathbb{N}$, $\xi_n^{(N)} \to \xi_n$ in probability and as $n \to \infty$, $\xi_n \to \xi$ in probability or distribution. Then there exists subsequences $n_m, N_m \to \infty$ such that $\xi_{n_m}^{(N_m)} \to \xi$ in probability or distribution.*

Finally, a direct application of this lemma yields:

Theorem 6. *Let (1.1) be an ergodic diffusion which satisfies Assumptions 3-6 and Assumption 12. Then there exists subsequences $n_m, N_m \to \infty$ such that the approximated maximum likelihood estimator $\hat{\theta}_{n_m}^{(N_m)}$ is consistent and asymptotically normal with mean θ_0 and asymptotic covariance matrix $I^{-1}(\theta_0)$ where θ_0 is the true parameter. Moreover, $\hat{\theta}_{n_m}^{(N_m)}$ is asymptotically efficient.*

1.5.3 An Example

As an illustration, we consider a simple example where the one-dimensional diffusion is given by

$$\begin{cases} dX_t = -\theta X_t^{\frac{1}{3}} dt + dW_t, \ \ 0 \le t \le T \\ X_0 = x_0 \end{cases} \tag{1.49}$$

For simplicity, the only unknown parameter appears in the drift term b. It can be verified that (1.49) is an ergodic diffusion process. Notice that the drift term $b(x;\theta) = -\theta x^{\frac{1}{3}}$ exhibits sub-linear growth so that the previous parametrix approximation is applicable. The goal is to estimate θ based on observed values of X_t.

For the parametrix approximation, we truncate the series to the first order so that we have

$$p(t,x,y;\theta) \approx p^{(1)}(t,x,y;\theta) = Z(t,x,y) + \int_0^t \int_{\mathbb{R}} Z(t-s,x,u) LZ(s,u,y;\theta) du ds,$$

where

$$Z(t,x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$$

and

$$LZ(t,x,y;\theta) = -\theta x^{\frac{1}{3}} \frac{\partial Z}{\partial x}(t,x,y).$$

The integral in the second terms is evaluated using quadrature methods. Before we perform the MLE, let us first examine the parametrix approximation of transition density and compare it with Euler-Maruyama scheme.

In Figure 1.1, we plot the probability density from both the Euler-Maruyama scheme and parametrix approximation. We fix the parameter $\theta = 0.5$ and $x = 1$. We then draw the density as a function of y for different time $T = 0.1$ and $T = 0.4$, i.e. we plot the densities $p(0.1,1,\cdot;0.5)$ and $p(0.4,1,\cdot;0.5)$. It can be seen that for small time instant $T = 0.1$, both scheme yield similar approximation of the transition density. When t is larger, deviation of parametrix from EM scheme is more noticeable. Since the EM scheme is implemented with very fine mesh on time

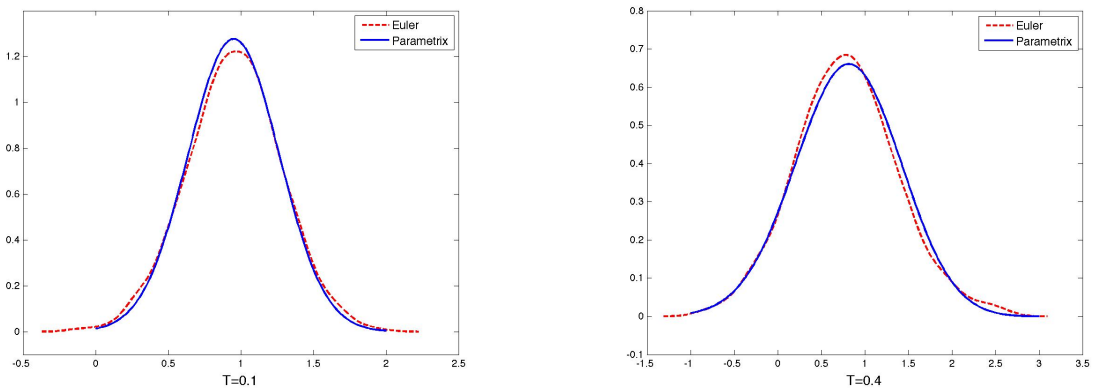


Figure 1.1: Comparison of parametrix approximation and Euler-Maruyama scheme for SDE (1.49). We fix $\theta = 0.5$ and $x = 1$. Left: $T = 0.1$; Right: $T = 0.4$.

Number of data	100	400	1000	10000
MLE $\hat{\theta}_n^{(1)}$	0.8894	0.3907	0.4552	0.4919

Table 1.1: Approximated MLE with different number of data points. True value: $\theta_0 = 0.5$.

domain and then averaged with large number of trajectories, it is relatively accurate. This suggests that more accurate quadrature method is needed or more terms in the series should be added for the parametrix approximation. However, in real applications of parameter estimations, t is usually kept small. For example, in the estimation problems in finance, $t = 1/250$ with 250 being the working days in a year, or $t = 1/12$ with 12 being the number of months. This implies that parametrix method suffices in these cases.

We next examine the performance of the approximated MLE. We fix the true value $\theta_0 = 0.5$ and generate synthetic data points with this value. We perform the maximization of the approximated log-likelihood with the standard optimization routine in Matlab with initial guess $\theta = 2$ and list the result in Table 1.1. It can be seen that as the number of data increases, $\hat{\theta}_n^{(1)}$ approaches the true value 0.5. This is in accordance with the theorem in the previous subsection. However, this convergence is rather slow. One possible reason is due to the errors introduced when truncating the parametrix series. Seeking more accurate and more efficient ways to calculate the parametrix approximation is a possible future research.

Chapter 2

Estimation for Deterministic Systems

2.1 Bayesian Framework

Estimating the parameter which is involved in a deterministic system, such as a partial differential equation (PDE), is a central problem in many applications. It can be viewed as an inverse problem. For instance, we have a PDE model that describes the wave propagation after an explosion occurs on the surface. The parameter which represents the unknown subsurface media property may be involved in the PDE and, based on the observations on the ground, we need to reconstruct the subsurface media property. A sudden change in this property is an indication of possible oil deposits, etc. Another typical example is a PDE which describes the motion of air pollution. A parameter that represents the locations of sources of pollution is present in the system. Once again we need to reconstruct the location of pollution sources, based on the observed density of pollutants.

There is uncertainty which is inherent in this type of problem even though the system itself is deterministic. Uncertainty is a typical feature for all inverse problems. It may come from the the measurement error, model errors, missing observations or the uncertainty from prior information. Moreover, the inverse problem is often ill-posed, meaning that different choices of parameters could give reasonable explanations of data. As a consequence, characterizing this uncertainty is

the core of solving inverse problems.

The Bayesian approach provides a systematic framework for modeling the uncertainty in the parameter estimation for inverse problems. This approach forms a posterior density by combining a prior probability density function (pdf), which incorporates all available information we already know before the experiment, and a likelihood function, which measures how likely the data would be for a given parameter. To give a mathematical formulation, we denote by θ the unknown parameter. Let $p_0(\theta)$ be the prior density of the parameter and $p(b|\theta)$ be the likelihood function where b denotes the observations. The Bayesian rule simply states that the posterior density of the parameters is given by

$$p(\theta|b) \propto p_0(\theta)p(b|\theta). \tag{2.1}$$

Note that the posterior is defined up to an (unknown) constant. This poses some difficulties for sampling, which will be addressed in the next few sections.

Let us make this even clearer by introducing some notations. Suppose b, the observation in our inverse problem, can be written as

$$b = h(\theta, e),$$

where θ is an unknown N_θ dimensional parameter, e represents the observation error, and h is a function which maps the parameter space into observations. In applications, it can be derived from the discretization of PDEs. Certainly the space of parameter θ may be an infinitely dimensional space, such as a coefficient or the initial condition in the PDE. Here we will restrict ourselves to the discrete finite dimensional parameter space, which is consistent with the the discretization of the forward model. In this chapter, we consider time-dependent PDEs so that the solution is derived sequentially. We denote the discrete forward model as, for $n = 0, \cdots, N-1$,

$$X^{n+1} = M^{n+1}(X^n, \theta), \tag{2.2}$$

where M is the forward model solution operator which, for a given parameter θ, solves the state variables from time t_n to t_{n+1}. Moreover, X^{n+1} is an N_x dimensional state variable. The M

observations are collected sequentially and we assume that the errors are additive. That is, for $m = 1, \cdots, M$,

$$b^m = h^m(X^{n_m}(\theta)) + e^m, \tag{2.3}$$

where e^m are independent Gaussian noise with zero mean and covariance matrix Γ_e, denoted as $e^m \sim N(0, \Gamma_e)$. Let $p_0(\theta)$ be the prior distribution of θ. Using the fact that the observation errors e^m are independent and the state variables X^n are deterministic functions of θ, we can write the posterior density of θ, by the Bayesian rule, as

$$p(\theta|b^{1:M}) \propto p_0(\theta)\Pi_{m=1}^{M} p(b^m|X^{n_m}(\theta)), \tag{2.4}$$

where $p(b^m|X^{n_m}(\theta))$ is the likelihood function which can be derived from (2.3).

For example, assume that the prior is Gaussian with mean θ_p and covariance matrix C_p and the errors are centered gaussian with covariance matrix C_e, then the posterior can be written as

$$\begin{aligned} p(\theta|b^{1:M}) \quad \propto \quad & \exp\left[-\frac{1}{2}|\theta-\theta_p|^T C_p^{-1}|\theta-\theta_p| \right. \\ & \left. -\frac{1}{2}|b^{1:M} - h^m(X^{n_m}(\theta))|^T C_e^{-1}|b^{1:M} - h^m(X^{n_m}(\theta))|\right]. \end{aligned}$$

Gaussian assumptions are reasonably good in many applications. However, we point out that even if both the prior and the observation errors e^m are Gaussian, the posterior may still be non-Gaussian due to the possible nonlinear dependence of the forward solution operator M on parameters or the nonlinear observation function h^m.

The posterior density incorporates information from both the historical information and the observations. It provides a complete solution to the inverse problem under the Bayesian framework, [51]. We do not need a full description of the posterior. Instead, we are more interested in a functional description in terms of its moments. For example, one can estimate and quantify the uncertainty in the parameters by using the mean and variance of the posterior density. However, the forward model is discretized on fine meshes for numerical solutions in most cases. As

a consequence, the parameter space should also be discretized accordingly and this process will lead to an extremely high dimensional inverse problem. In practice, due to the large scale and high nonlinearity of the involved system, direct characterizing the posterior using numerical integration is computationally prohibitive. Therefore, Monte Carlo methods are widely used. One needs to obtain enough samples from the posterior density to analyze its statistical features, such as the mean and covariance. Therefore, sampling the posterior is fundamental to the solution of inverse problems.

In the following sections we will briefly introduce various sampling techniques including Gaussian approximations, MCMC and our implicit sampling/sequential implicit sampling method. We will give a detailed derivation of implicit sampling method while keeping others succinct. If not otherwise indicated, we will sample the posterior (2.4).

2.2 Gaussian Approximations

The most popular Gauss Approximations are the Kalman Filter method. In fact, it is a family of methods including the classic Kalman Filter (KF), Extended Kalman Filter (EKF) and the most popular Ensemble Kalman Filter (EnKF). The original Kalman Filter method is designed to update the state variables in a linear system with linear observations, rather than updating the parameters in our case. Recall that we assume that parameter is time-invariant for simplicity. We will first demonstrate the classical Kalman Filter and Extended Kalman Filter, after which we adapt it to our parameter estimations.

2.2.1 Kalman Filter

We begin with the evolution equation

$$X_{n+1} = AX_n + U_n, \tag{2.5}$$

where X_n is an N_x-dimensional variable and U_n is the associated $N_x \times N_x$ centered Gaussian process. It is assumed that U_n has a known,common covariance matrix Q. Moreover, A is the evolution matrix which is assumed to be stationary over time. For example, A can be derived from the discretization of a linear PDE. Note that when the PDE is non-linear, we can not directly use the classical Kalman Filter framework.

Besides the evolution equation, we have the following observation equation:

$$Y_{n+1} = BX_{n+1} + V_{n+1}, \tag{2.6}$$

where Y_n is the $N_y \times 1$ vector of observations, V_n is another centered Gaussian error process with known and common covariance R and B is the $N_y \times N_x$ observation matrix. Also, both U_n and V_n are independent from X and Y.

Suppose that at stage n, we have obtained an optimal estimation $\hat{X}_n$ of X_n together with the covariance matrix P_n of $X_n - \hat{X}_n$ (see below), we are seeking an optimal estimation $\hat{X}_{n+1}$ of X_{n+1} in terms of certain optimal criterion. There are two steps in deriving the Kalman formulas. The first step is to forecast a new state using the evolution equation (2.5). Since

$$E(\hat{X}_{n+1}|Y_n) = E(AX_n + U_n|Y_n) = E(AX_n|Y_n) = A\hat{X}_n,$$

due to independence and optimality of $\hat{X}_n$, we have the state projection

$$\hat{X}'_{n+1} = A\hat{X}_n. \tag{2.7}$$

This is the best estimation conditioned on the previous observation, in the absence of any new data. The formula is intuitively clear since we simply move the best estimation forward by the evolution equation. Next we step into the second stage, the analysis step, when the new data Y_{n+1} is available. We do not choose the form of optimal estimate at random. Instead, we are looking for

the optimal estimation of X_{k+1} in the following form:

$$\hat{X}_{n+1} = \hat{X}_{n+1}^{'} + K_{n+1}(Y_{n+1} - B\hat{X}_{n+1}^{'}), \tag{2.8}$$

where K_{n+1} is the Kalman Gain matrix which is to be determined. Notice that this is a linear combination of prediction $\hat{X}_{n+1}^{'}$ and the measurement residue $Y_{n+1} - B\hat{X}_{n+1}^{'}$. Now K_{n+1} is chosen so that the mean square error:

$$P_{n+1} = E[(X_{n+1} - \hat{X}_{n+1})(X_{n+1} - \hat{X}_{n+1})^T]$$

is minimized. This is the standard for optimality. Denote by P_n the covariance matrix of $X_n - \hat{X}_n$ and $P_n^{'}$ the covariance matrix of $X_n - \hat{X}_n^{'}$. An application of vector calculus shows that K_{n+1} can be written as

$$K_{n+1} = P_{n+1}^{'} B^T (BP_{n+1}^{'} B^T + R)^{-1}. \tag{2.9}$$

The relationship between $P_{n+1}^{'}$ and P_n is

$$P_{n+1}^{'} = AP_nA^T + Q.$$

Up to this point, we have finished constructing the best estimation in the sense of mean square error if the current estimation is available. There is one last formula needed to complete the Kalman Filter loop, i.e. the formula for updating P_n. Direct calculations yields

$$P_{n+1} = P_{n+1}^{'} - K_{n+1}BP_{n+1}^{'}.$$

The above derivation is only sketchy and more details can be found in [26]. We summarize the Kalman Filter loop in Algorithm 1.

Algorithm 1 The Kalman Filter Algorithm

1: Initialization: Choose the optimal X_0 and optimal covariance matrix P_0.
2: The Kalman Filter iteration: for $n = 1, \cdots, N$ (N is the total evolution steps.)

1. Forecast step: propagate the state with the evolution equation, giving

$$\hat{X}_{n+1}^{'} = A\hat{X}_n.$$

Calculate its covariance as

$$P_{n+1}^{'} = AP_nA^T + Q.$$

2. Analysis step: Update the propagated state with new data, yielding

$$\hat{X}_{n+1} = \hat{X}_{n+1}^{'} + K_{n+1}(Y_{n+1} - B\hat{X}_{n+1}^{'}),$$

where the Kalman Gain matrix is given by

$$K_{n+1} = P_{n+1}^{'}B^T(BP_{n+1}^{'}B^T + R)^{-1}.$$

Calulate the updated covariance as

$$P_{n+1} = P_{n+1}^{'} - K_{n+1}BP_{n+1}^{'}.$$

2.2.2 Extended Kalman Filter

When we have nonlinear evolutions and observations, classic Kalman Filter does not apply. One way to solve the problem is to resort to Extended Kalman Filter. Suppose that we have evolution equation:

$$X_{n+1} = A(X_n) + U_n \tag{2.10}$$

and observation equation:

$$Y_{n+1} = B(X_{n+1}) + V_{n+1} \tag{2.11}$$

where, unlike Kalman Filter, A and B are possible nonlinear functions of its arguments. Specifically, A maps $\mathbb{R}^{N_x}$ into $\mathbb{R}^{N_x}$ and B maps $\mathbb{R}^{N_x}$ into $\mathbb{R}^{N_y}$. For simplicity, we still assume additive white noise process U_n and error process V_n. The algorithm of Extended Kalman Filter is quite similar to that of classic Kalman Filter, except that we first linearize A and B. Denote by J_A and J_B the

Jacobian matrix of A and B, i.e.

$$J_A = \begin{bmatrix} \frac{\partial A_1}{\partial X_1} & \frac{\partial A_1}{\partial X_2} & \cdots & \frac{\partial A_1}{\partial X_{N_x}} \\ \vdots & & \ddots & \vdots \\ \frac{\partial A_{N_x}}{\partial X_1} & \frac{\partial A_{N_x}}{\partial X_2} & \cdots & \frac{\partial A_{N_x}}{\partial X_{N_x}} \end{bmatrix}$$

and

$$J_B = \begin{bmatrix} \frac{\partial B_1}{\partial X_1} & \frac{\partial B_1}{\partial X_2} & \cdots & \frac{\partial B_1}{\partial X_{N_x}} \\ \vdots & & \ddots & \vdots \\ \frac{\partial B_{N_y}}{\partial X_1} & \frac{\partial B_{N_y}}{\partial X_2} & \cdots & \frac{\partial B_{N_y}}{\partial X_{N_x}} \end{bmatrix}$$

The Extended Kalman Filter has the following algorithm:

Algorithm 2 The EKF Algorithm

1: Initialization: Choose the optimal X_0 and optimal covariance matrix P_0.
2: The Extended Kalman Filter iteration: for $n = 1, \cdots, N$ (N is the total evolution steps.)

1. Forecast step: propagate the state with the evolution equation, giving

$$\hat{X}^{'}_{n+1} = A(\hat{X}_n).$$

Calculate its covariance as

$$P^{'}_{n+1} = J_A(\hat{X}_n) P_n J_A^T(\hat{X}_n) + Q.$$

2. Analysis step: Update the propagated state with new data, yielding

$$\hat{X}_{n+1} \approx \hat{X}^{'}_{n+1} + K_{n+1}(Y_{n+1} - B(\hat{X}^{'}_{n+1}))$$

where the Kalman Gain matrix is given by

$$K_{n+1} = P^{'}_{n+1} J_B^T(\hat{X}^{'}_{n+1})[J_B(\hat{X}^{'}_{n+1}) P^{'}_{n+1} J_B^T(\hat{X}^{'}_{n+1}) + R]^{-1}.$$

Calulate the updated covariance as

$$P_{n+1} = P^{'}_{n+1} - K_{n+1} J_B(\hat{X}^{'}_{n+1}) P^{'}_{n+1}$$

It can be seen that the Extended Kalman Filter algorithm is the analogue of Kalman Filter algorithm with the matrices A and B replaced by their linearization (the Jacobian). The Extended Kalman Filter satisfies first order optimality, although higher order optimality can be achieved with

more terms added into the Taylor expansions of nonlinear functions A and B. We refer the reader to [49] for detailed derivation of these facts.

2.2.3 Ensemble Kalman Filter for Parameter Estimations

In practice, Extended Kalman Filter is not widely used due to the calculation of the Jacobian J_A and J_B, which can be computationally expensive for large scale problems. Instead, Ensemble Kalman filter is widely used. Specifically, what we have at time k is a bunch of samples $\hat{X}_{n,i},\ i = 1,2,\ldots,N_e$ from the optimal Gaussian distribution $N(\hat{X}_n, P_n)$. Here N_e is the total number of samples. In the prediction step, we propagate the samples via

$$X_{n+1,i} = M(X_n, i) + U_n, \quad i = 1,2,\ldots,N_e$$

to generate the prediction samples. We also calculate the sample mean and sample covariance. In the analysis step, we still use the Kalman Gain formula with the only difference being that all covariance matrices are replaced by sample covariance which can be easily calculated.

It turns out that Ensemble Kalman Filter (EnKF) method is the ideal one for parameter estimations in the Kalman family. Now we generalize the EnKF method to fit our parameter estimation. This generalization is adapted to [31]. Note that we assume the parameter is time-invariant while the EnKF is updating the state variable at different time step. As we see later, the update in the parameters will be done through the update of state variable. To explain how it works, we will return to the notations in section 2.1. Recall that we have the state equation

$$X^{n+1} = M^{n+1}(X^n, \theta),$$

and the observation equation:

$$b^m = h^m(X^{n_m}(\theta)) + e^m.$$

Finally, the posterior density is given by

$$p(\theta|b^{1:M}) \propto p_0(\theta)\Pi_{m=1}^{M} p(b^m|X^{n_m}(\theta)),$$

We define a new state of the assimilation algorithm as

$$Q^n = \begin{pmatrix} \theta \\ X^n \\ h^n(X^n) \end{pmatrix}, \quad \Phi^{n+1}(Q^n) = \begin{pmatrix} \theta \\ M^{n+1}\left(X^n, \theta\right) \\ h^{n+1}\left(M^{n+1}\left(X^n, \theta\right)\right) \end{pmatrix},$$

Thus, the new system and the observation equation can be written as

$$Q^{n+1} = \Phi^{n+1}(Q^n) \tag{2.12}$$

and

$$b^{n+1} = H^{n+1}(Q^{n+1}) + e^{n+1}, \tag{2.13}$$

where $H^{n+1} = (0;0;I)$. Here I denotes the identity matrix.

In each EnKF assimilation cycle (forecast and analysis), we need to update the ensemble at time t^{n+1} so that the ensemble are the samples from the distribution

$$p(Q^{n+1}|b^{1:n+1}) \propto p(Q^{n+1}|b^{1:n})p(b^{n+1}|Q^{n+1}).$$

At time t^n, we have the ensemble members from previous analysis step, which are samples from $p(Q^n|b^{1:n})$. In the forecast step, the ensemble are propagated using the forward model to obtain the samples to represent the distribution $p(Q^{n+1}|b^{1:n})$. In the analysis step, $p(Q^{n+1}|b^{1:n})$ is approximated by a Gaussian distribution, whose mean and variance are given by the sample mean and variance from the ensemble obtained in the forecast step. Then the ensembles are updated using Kalman formula in a way to make sure the sample mean and variance converge to the true ones

when the size of the ensemble goes to infinity. Detailed implementation of perturbed observation EnKF is provided in the following Algorithm 3. There are several techniques such as localization to make the EnKF work better with a small number of ensemble for large data set, see [31]. In our numerical experiments, we use the localization techniques for the EnKF.

Algorithm 3 The EnKF Algorithm for Parameter Estimate

1: Initialization

$$Q^{0,a,\{i\}} = \begin{pmatrix} \theta^{\{i\}} \\ X^0 \\ h^0(X^0) \end{pmatrix},$$

where $\{\theta^{\{i\}}\}_{i=1}^{N_e}$ are the samples from the prior distribution $p_0(\theta)$ and X^0 is the initial condition for the state variables.

2: the EnKF iteration: for $n = 1, \cdots, N$ (N is the total data-assimilation steps.)

1. Forecast step: propagate the ensemble with (2.12), giving

$$Q^{n,f,\{i\}} = \Phi^n(Q^{n,a,\{i\}}), \quad i = 1, \cdots, N_e.$$

2. Analysis step: The sample mean and covariance of the ensemble are calculated as follows

$$\bar{Q}^{n,f} = \frac{1}{N_e} \sum_{i=1}^{N_e} Q^{n,f,\{i\}},$$

$$P^{n,f} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(Q^{n,f,\{i\}} - \bar{Q}^{n,f} \right) \left(Q^{n,f,\{i\}} - \bar{Q}^{n,f} \right)^T,$$

$$Q^{n,a,\{i\}} = Q^{n,f,\{i\}} + K^n \left(b^{n,\{i\}} - H^n \left(Q^{n,f,\{i\}} \right) \right),$$

where the Kalman gain matrix K is given by

$$K^n = P^{n,f} (H^n)^T \left(H^n P^{n,f} (H^n)^T + \Gamma_e \right)^{-1},$$

and $b^{n,\{i\}}$ are perturbed observations, given by

$$b^{n,\{i\}} = b^n + e^{n,\{i\}}, \quad e^{n,\{i\}} \sim N(0, \Gamma_e), \quad i = 1, \cdots, N_e. \tag{2.14}$$

2.3 Markov Chain Monte Carlo

The Kalman Filter methods introduced in the previous section are essentially Gaussian approximations. The final sampling results are Gaussian random variables which are the best approximation to the posterior in mean square sense. However, the posterior is often non-Gaussian in practice, especially when the discretization of equation generates a non-linear function of the parameters. In such cases, Gaussian approximations often introduce biases, which will be demonstrated in Section 2.7. Thus, it will be desirable to design a method which does not involve any Gaussian assumptions.

Among these methods of PDE-based inverse problems, Markov Chain Monte Carlo (MCMC) is the most popular one and has long been serving as a golden standard for sampling any given densities. It has such generality that it can be performed for large scale PDEs with any discretization (mesh invariance) [9] [14] [38]. It could even be conceived on a function space and then be implemented under numerical discretization. In this section, we will introduce the basics of MCMC and pick up several popular algorithms which are widely used in practice. For a review of the state-of-the-art MCMC methods together with their properties, we refer the reader to [14]

The idea of MCMC is to build up a Markov chain with the posterior as its invariant density. If this is possible and we run this chain long enough, the chain will take the posterior as its invariant distribution and the samples from the chain can be viewed as samples from the posterior. We need to point out that the samples we obtain from MCMC are not independent since the chain has memories. This is a drawback of MCMC. However, it does not affect the description of the posterior much if we only care about its moments.

We have a quick review of Markov Chains and their properties in Appendix 1 so that we can move to the essence of MCMC quickly. We begin with a transition probablity kernel $Q(\theta, A)$ on continous (or discrete, depending on the contex) space. Here $\theta \in \mathbb{R}^{N_\theta}$ and $A \in \mathbb{B}(\mathbb{R}^{N_\theta})$, where $\mathbb{B}(\mathbb{R}^{N_\theta})$ is the Borel σ-algebra on $\mathbb{R}^{N_\theta}$. If the kernel has a density, which is commonly seen in applications, we denote it by $q(\theta, \gamma)$, i.e.

$$Q(\theta, A) = \int_A q(\theta, \gamma) d\gamma$$

One useful feature of a Markov chain is that under certain conditions, the chain has an invariant distribution that it converges to. By invariant distribution we mean a distribution $\pi(d\gamma)$ on $\mathbb{R}^{N_\theta}$ such that

$$\pi(A) = \int_{\mathbb{R}^{N_\theta}} Q(\theta, A) \pi(d\theta)$$

for any $A \in \mathbb{B}(\mathbb{R}^{N_\theta})$. Note that for a given chain, we do not know whether there is an invariant distribution or not, nor do we know the form the invariant distribution unless we check those conditions carefully and do some explicit computations.

The design of MCMC for parameter estimation is the inverse of the above process. It takes the posterior $p(\theta|b^{1:M})$ as its invariant distribution and seek a chain that eventually converges to it. Since the general theory of Markov chains indicates that the transition kernel Q determines the distribution of the chain [1], this is equivalent to saying that we are seeking a transition probability kernel $Q(\theta, A)$ that takes $p(\theta|b^{1:M})$ as its invariant distribution. One sufficient condition to guarantee this is the so called reversibility condition:

$$p(\theta|b^{1:M}) Q(\theta, d\gamma) = p(\gamma|b^{1:M}) Q(\gamma, d\theta). \tag{2.15}$$

In the case where Q have densities, we have

$$p(\theta|b^{1:M}) q(\theta, \gamma) = p(\gamma|b^{1:M}) q(\gamma, \theta). \tag{2.16}$$

We stick to the kernel with densities from this point if not otherwise stated. It is easy to see that the reversibility condition guarantees that the probability of going from θ to γ is the same as the probability of going from γ to θ, see [9].

[1] The initial distribution also determines the distribution. However, the asymptotic theory guarantees the convergence to invariant distribution regardless of the choice of initial distribution.

In practice, we can choose any transition density $q(\theta,\gamma)$. However, it is likely that the reversibility condition (2.16) will not be satisfied by this random choice. The Metropolis-Hastings MCMC will compensate this by introducing $\alpha(\theta,\gamma)$ so that

$$q_{MC}(\theta,\gamma) = q(\theta,\gamma)\alpha(\theta,\gamma)$$

will make (2.16) valid. That is,

$$p(\theta|b^{1:M})q(\theta,\gamma)\alpha(\theta,\gamma) = p(\gamma|b^{1:M})q(\gamma,\theta)\alpha(\gamma,\theta)$$

To rule out the possibility of a transition density that is greater than one, we choose

$$\alpha(\theta,\gamma) = \begin{cases} \min\left[\frac{p(\gamma|b^{1:M})q(\gamma,\theta)}{p(\theta|b^{1:M})q(\theta,\gamma)},1\right] & \text{if } p(\theta|b^{1:M})q(\theta,\gamma) > 0 \\ 1 & \text{otherwise} \end{cases} \tag{2.17}$$

The above process completes the design of Metropolis-Hastings MCMC algorithm. Intuitively, if the chain is moving towards the maximum of the posterior density, we will take this sample for sure since it is moving to a higher probability region [9]. If the chain is moving along an opposite direction, then there is certain probability of rejecting it. This is the analogue of acceptance-rejection (AR) algorithm. In fact, we will use the AR algorithm in Metropolis-Hastings MCMC, which is shown in Algorithm 4

There is a practical issue of choosing the appropriate transition kernel $q(\theta,\gamma)$ since the samples are coming directly from it. We need to choose it in such a way that it is easy to sample and the convergence associated with this kernel is fast. The first requirement is essential since we definitely do not want a density which is even harder to sample than the original posterior. It is also easy to be achieved. For example, one common choice of $q(\theta,\gamma)$ is the isotropic Gaussian kernel

$$q(\theta,\gamma) = \phi(\theta-\gamma),$$

Algorithm 4 The Metropolis-Hastings Algorithm

1: Generate a random initialized state θ_0.
2: The Metropolis-Hastings iteration: for $n = 1, \cdots, N$.
   1. Generate γ from $q(\theta_n, \gamma)$ and u from $U(0,1)$.
   2. calculate $\alpha(\theta_n, \gamma)$ as in (2.17).
   3. If $u < \alpha(\theta_n, \gamma)$, set $\theta_{n+1} = \gamma$;
   4. else, set $\theta_{n+1} = \theta_n$.
3: Return $\{\theta_1, \theta_2, \ldots, \theta_N\}$.
4: Take the samples $\{\theta_K, \theta_{K+1}, \ldots, \theta_N\}$ where K is a large enough integer so that the chain has reached its steady state.

where ϕ is the multivariate Gaussian density on $\mathbb{R}^{N_\theta}$. The MCMC derived from this kernel is called random walk MCMC since its samples are generated in a way that mimics the random walk. It is easy to see that random walk MCMC is easy to design and implement. The convergence is always guaranteed provided we make the step size small enough [9] [14] and it is usually used as a golden standard for sampling when comparing with other methods. However, the convergence of random walk MCMC is usually slow and the efficiency, which is mentioned above, is not desirable for large scale problems. Choosing a big step size may cause convergence problems while a small step size may lead to extremely slow convergence.

Much effort has been made to improve the efficiency of classic Metropolis-Hastings MCMC algorithm. Here we pick up one of them, namely the stochastic Newton MCMC, which was originated from the paper by [38]. For simplicity, let us look at the posterior composed of Gaussian prior and Gaussian likelihood as in Section 2.1:

$$\begin{aligned} p(\theta|b^{1:M}) \quad \propto \quad & \exp\left[-\frac{1}{2}(\theta-\theta_p)^T C_p^{-1}(\theta-\theta_p)\right. \\ & \left.-\frac{1}{2}(b^{1:M}-h^m(X^{n_m}(\theta)))^T C_e^{-1}(b^{1:M}-h^m(X^{n_m}(\theta)))\right] \end{aligned}$$

The essence of random walk MCMC is to move the chain towards the high probability region of the posterior. This can be slow for large scale problems. To facilitate this process, the stochastic Newton MCMC exploits the local Gaussian structure of posterior and uses a pseudo-maximization

process. To be precise, let

$$\begin{aligned} F(\theta) &= \frac{1}{2}(\theta-\theta_p)^T C_p^{-1}(\theta-\theta_p) \\ &\quad +\frac{1}{2}(b^{1:M}-h^m(X^{n_m}(\theta)))^T C_e^{-1}(b^{1:M}-h^m(X^{n_m}(\theta))). \end{aligned} \tag{2.18}$$

Thus $F(\theta)$ is the negative logarithm of $p(\theta|b^{1:M})$. Notice that maximizing $p(\theta|b^{1:M})$ is equivalent to minimizing $F(\theta)$. Suppose at step k, we have the sample θ_k, then we define $g(\theta_k)=\nabla F(\theta_k)$ and $H(\theta_k)=\Delta F(\theta_k)$ where Δ is the Hessian operator. Since θ_k is usually not the minimizer of F, $H(\theta_k)$ may not be positive definite. In this case, we will have trouble constructing an approximated Gaussian proposal with $H(\theta_k)$. To make it positive definite, we can modify it while keeping it close to the original $H(\theta_k)$. One simple choice is to replace all negative eigenvalues of $H(\theta_k)$ with small positive thresholds [38]. We denote the modified hessian by $\hat{H}(\theta_k)$. Then the proposal density is chosen to be

$$q(\theta_k,\gamma)\propto \exp\left[-\frac{1}{2}(\gamma-\theta_k+\hat{H}^{-1}(\theta_k)g(\theta_k))^T\hat{H}(\theta_k)(\gamma-\theta_k+\hat{H}^{-1}(\theta_k)g(\theta_k))\right] \tag{2.19}$$

That is, we use the Gaussian density $N(\theta_k-\hat{H}^{-1}(\theta_k),\hat{H}(\theta_k))$.

Note that the sample generated in the k-th step

$$\theta_{k+1}=\theta_k-\hat{H}^{-1}(\theta_k)$$

is essentially a Newton step. It moves the sample to the local minimizer of F, which in turn is the maximizer of $p(\theta|b^{1:M})$. In this way, the chain reaches stationary state much faster than the random walk MCMC. We present the stochastic Newton MCMC method in Algorithm 5.

It can be easily seen that the only difference between Newton MC and random walk MC is the proposal density. We need to point out that in the stochastic Newton MC, we need to calculate the Hessian H which is derived from the PDE system. It can be very expensive for large scale problems, which in turn lowers the efficiency of the algorithm. One way to remedy this issue is to

Algorithm 5 The Stochastic Newton MCMC Algorithm

1: Generate a random initialized state θ_0.
2: Compute $p(\theta_0|b^{1:M})$,$g(\theta_0)$ and $\hat{H}(\theta_0)$.
3: The Stochastic Newton MC iteration: for $n = 1, \cdots, N$,
   1. Generate γ from (2.19) and u from $U(0,1)$.
   2. Calculate $p(\theta_n|b^{1:M})$,$g(\theta_n)$ and $H(\theta_n)$.
   3. Calculate α as in (2.17).
   4. If $u < \alpha(\theta_n, \gamma)$, set $\theta_{n+1} = \gamma$;
   5. else, set $\theta_{n+1} = \theta_n$.
4: Return $\{\theta_1, \theta_2, \ldots, \theta_N\}$.
5: Take the samples $\{\theta_K, \theta_{K+1}, \ldots, \theta_N\}$ where K is a large enough integer so that the chain has reached its steady state.

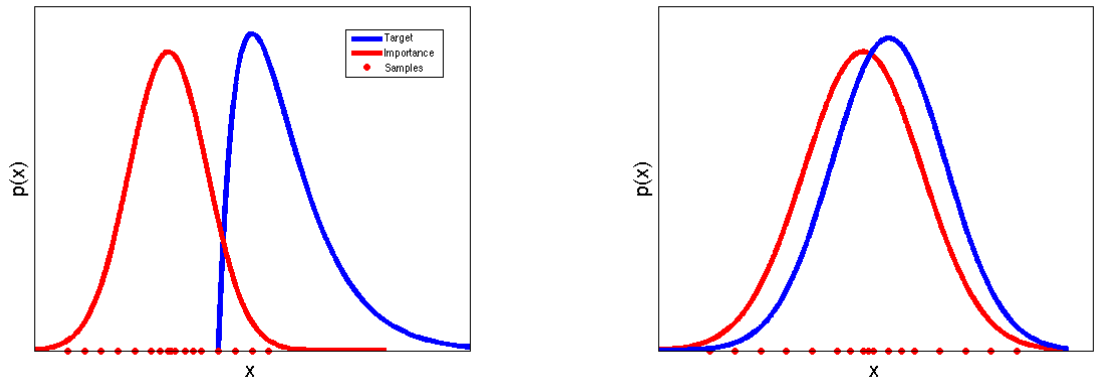
use the adjoint method as in [38].

2.4 Implicit Sampling and Sequential Implicit Sampling

Implicit sampling method [3] [12] can be viewed as an updated importance sampling that significantly improves the sampling efficiency. The importance sampling method is another popular Monte Carlo method [10] which generates independent samples without any Gaussian assumptions. In the importance sampling method, we first generate samples from another easy-to-sample density, called the importance density, and weigh those samples using the ratio of the target density to the importance density. The weighted samples can empirically approximate the target posterior density. Suppose we need to find the expectation of $g(\theta)$, where g is a scalar function on $\mathbb{R}^{N_\theta}$, then by law of large numbers,

$$\begin{aligned} E[g(\theta)] = \int g(\theta) p(\theta|b^{1:M}) d\theta &= \int g(\theta) \frac{p(\theta|b^{1:M})}{q(\theta)} q(\theta) d\theta \\ &\approx \frac{1}{\sum_{i=1}^N \omega_i} \sum_{i=1}^N g(\theta_i)\omega_i \end{aligned}$$

where we abuse the notation by using θ as both the random variable and its realization. Here $q(\theta)$ is the importance density, $\theta_i \sim q(\theta)$ are the importance samples and $\omega_i = \frac{p(\theta_i|b^{1:M})}{q(\theta_i)}$ are correspond-



(a) Bad importance density

(b) Good importance density

Figure 2.1: Different choices of importance densities.

ing weights. Since the samples are not directly from $p(\theta|b^{1:M})$, but from the importance density $q(\theta)$, different samples are assigned different weights to reflect their contributions to the posterior.

Choosing importance density appropriately is crucial to the successful implementation of importance sampling. A good choice of the importance density should

1. be easy to sample.

2. be close to the posterior density.

3. be of small variances in the weights.

While item 1 is easy to achieve, item 2 and 3 need special consideration. Poor choices may lead to huge amount of computational waste on samples that contribute little or even nothing (e.g. q and p are singular to each other) to the posterior density. See figure 2.1 for a simple visualization. In particular, most of our endeavor is devoted to the search for importance density that is large where the target density is large so that the samples drawn from the importance density can make adequate contributions to the posterior.

The implicit sampling method, [16, 12, 41], provides a general framework to choose the importance density for the importance sampling method. It is convenient to denote

$$F(\theta) = -\log p(\theta|b^{1:M}) = -\log\left[p_0(\theta)\Pi_{m=1}^{M} p(b^m|X^{n_m}(\theta))\right], \tag{2.20}$$

the negative logarithm of the posterior density (2.4).

The first step of implicit sampling is to locate the high probability region of the posterior density $p(\theta|b^{1:M})$ by minimizing $F(\theta)$. We define ϕ_F as the minimum value of $F(\theta)$ and μ as the minimizer, i.e. $\phi_F = \min F$ and $\mu = \text{argmin} F$. This is the most expensive step in implicit sampling and it is the same as finding the MAP (maximum a posteriori) point, i.e. μ is the MAP point of $p(\theta|b^{1:M})$.

The second step in implicit sampling is to generate samples around the MAP point. We pick up a reference random variable ξ with probability density function $g(\xi) \propto e^{-G(\xi)}$ and define $\phi_G = \min G$. The samples for θ are generated by first drawing samples from ξ and then solving the following equation:

$$F(\theta) - \phi_F = G(\xi) - \phi_G. \tag{2.21}$$

By an application of the change of variable formula, we derive that the associated weights for the samples are given by

$$w_j \propto J(\theta), \tag{2.22}$$

where J is the Jacobian of the one-to-one and onto map $\theta \to \xi$. Since the samples from ξ are independent and located around its MAP point, θ_i's are also independent from each other and close to μ, the the MAP point of (2.4). In this way, we are able to efficiently focus on samples from the high probability region which make significant contributions to the target density.

As usual, the reference density in this thesis is taken as standard multivariate Gaussian so that $G(\xi) = \frac{1}{2}\xi^T\xi$. We note that this will not lead to any Gaussian assumption of the sampling algorithms since ξ is only a reference. The only thing left is to solve equation (2.21) for a given ξ_i from the reference density. The solution to (2.21) is not unique. Two approaches for the solutions are popular: linear and random map, [39].

Random map: We seek the solution to (2.21) in the form

$$\theta = \mu + \lambda L^T \xi, \tag{2.23}$$

Here λ is the scalar to be determined and L is a preconditioning matrix which helps to avoid the weight degeneracy. A common choice of L is the Cholesky factorization of the inverse of the Hessian at the MAP point μ. The only unknown in (2.23) is the scalar λ which can be solved by Newton's method. This is easy to implement and relatively cheap. Finally, what we care about is the weight associated with each sample. Here we give a sketchy derivation. More details can be found in [41]. Write

$$\theta = \mu + \hat{\lambda} L^T \xi,$$

where $\hat{\lambda} = \lambda / \sqrt{\xi^T \xi}$. Then we can compute

$$\frac{\partial \theta}{\partial \xi} = L^T \left(\xi \frac{\partial \hat{\lambda}}{\partial \xi} \right) + \hat{\lambda} L^T$$

We have

$$\frac{\partial \hat{\lambda}}{\partial \xi} = \frac{\partial \hat{\lambda}}{\partial \rho} \frac{\partial \rho}{\partial \xi} = 2 \frac{\partial \hat{\lambda}}{\partial \rho} \xi^T$$

where $\rho = \xi^T \xi$. Thus we have

$$\frac{\partial \theta}{\partial \xi} = L^T \left(2 \frac{\partial \hat{\lambda}}{\partial \rho} \xi \xi^T + \hat{\lambda} I \right)$$

A direct application of the determinant formula yields

$$J(\theta) = |\det L| \left| \hat{\lambda}^{N_\theta} \left(\hat{\lambda} + 2 \frac{\partial \hat{\lambda}}{\partial \rho} \rho \right) \right|$$

Converting back to λ, we have

$$\begin{aligned} J(\theta) &= |\det L| \rho^{1-N_\theta/2} \left| \lambda^{N_\theta - 1} \frac{\partial \lambda}{\partial \rho} \right| \\ &= |\det L| (\xi^T \xi)^{1-N_\theta/2} \left| \lambda^{N_\theta - 1} \frac{\partial \lambda}{\partial \rho} \right| \end{aligned}$$

Finally, $\partial \lambda / \partial \rho$ can be computed numerically, or using the aforementioned adjoint method, yield-

ing the following weights formula:

$$w \propto e^{\phi_F}(\xi^T\xi)^{1-N_\theta/2}\left|\frac{\lambda^{N_\theta-1}}{\nabla F\cdot(L^T\eta)}\right|, \tag{2.24}$$

where $\eta = \xi/|\xi|$. We summarize the random map implementation as in Algorithm 6.

Algorithm 6 Implicit Sampling with Random Map

1: Minimize $F(\theta)$;
2: Generate samples from the reference density $\xi_i \sim e^{-G(\xi)}$;
3: Derive samples from importance density via $\theta_i = \mu + \lambda_i L^T\eta$ by solving λ_i in

$$F(\mu+\lambda_i L^T\eta) - \phi_F = G(\xi_i) - \phi_G;$$

4: Assign weights according to

$$w_i \propto e^{\phi_F}(\xi_i^T\xi_i)^{1-N_\theta/2}\left|\frac{\lambda_i^{N_\theta-1}}{\nabla F\cdot(L^T\eta_i)}\right|;$$

5: Resampling.

Linear map: we first expand $F(\theta)$ to the second order:

$$F(\theta) \approx \phi_F + \frac{1}{2}(\theta-\mu)^T H(\theta-\mu) \triangleq F_0(\theta),$$

where H is the Hessian at μ. We solve the equation

$$F_0(\theta) - \phi_F = \frac{1}{2}\xi^T\xi,$$

which has the solution

$$\theta = \mu + L\xi, \tag{2.25}$$

where L is from the Cholesky factorization of H, i.e. $H = LL^T$. The weights associated with this

choice of importance density are given by

$$w \propto e^{F_0(\theta)-F(\theta)}, \tag{2.26}$$

since the Jacobian is constant for each sample. We summarize the linear map implementation in Algorithm 7.

Algorithm 7 Implicit Sampling with Linear Map

1: Minimize the $F(\theta)$ which appears in $e^{-F(\theta)}$;
2: Generate samples from the reference Gaussian density $\xi_i \sim e^{-G(\theta)}$;
3: Derive samples from the importance density via $\theta_i = \mu + L\xi_i$;
4: Assign weights according to $w_i \propto e^{F_0(\theta_i)-F(\theta_i)}$;
5: Resampling.

2.4.1 The Sequential Implicit Sampling Method

When the data are coming in a sequence, it is more practical to analyze data sequentially rather than waiting until all data are gathered. For example, in weather forecast, we have to forecast the weather in a given time period and cannot afford to wait for all data collected. The EnKF method is a sequential method and can update the estimate of the parameter sequentially when the new data come in. However, the analysis step in the EnKF method uses a Gaussian approximation and may not work well for nonlinear non-Gaussian problems. The implicit sampling method does not rely on any Gaussian assumption but it is not a sequential method. Here, we combine these two methods and propose a sequential implicit sampling method to make the parameter estimation sequentially.

We divide the total M observations into K observation windows and each window contains M_k data points such that

$$M_1 + M_2 + \cdots + M_K = M.$$

Instead of using all M observations for the posterior density function, we can write the posterior

density based on the first M_1 data as

$$p_1(\theta|b^{1:M_1}) \propto p_0(\theta)\Pi_{m=1}^{M_1} p(b^m|X^{n_m}(\theta)). \tag{2.27}$$

With the implicit sampling method, we can generate weighted samples θ_i^1 with weights w_i^1 of the density defined in (2.27). We employ the idea of the EnKF to approximate (2.27) as a Gaussian with mean $\theta_{(1)}$ and variance $V_{\theta_{(1)}}$ obtained from those samples. Because the observation errors e^n are independent from each other and the state variable X^n is a deterministic function of θ, we have

$$\begin{aligned}
& p_2(\theta|b^{1:M_2}) \propto p_0(\theta)\Pi_{m=1}^{M_2} p(b^m|X^{n_m}(\theta)) \\
= \;& p_0(\theta)\Pi_{m=1}^{M_1} p(b^m|X^{n_m}(\theta))\Pi_{m=M_1+1}^{M_2} p(b^m|X^{n_m}(\theta)) \\
\propto \;& p_1(\theta|b^{1:M_1})\Pi_{m=M_1+1}^{M_2} p(b^m|X^{n_m}(\theta)) \\
\approx \;& \exp\left[-\frac{1}{2}(\theta-\theta_{(1)})^T V_{\theta_{(1)}}^{-1}(\theta-\theta_{(1)})\right]\Pi_{m=M_1+1}^{M_2} p(b^m|X^{n_m}(\theta)).
\end{aligned}$$

Namely we are using a Gaussian prior with the updated information. In this way, we will incorporate the information we already obtained from the first batch of data and use that to guide the next sampling stage.

As we point out in the previous section, in the implicit sampling method, the most expensive part is to locate the high probability region, i.e. the optimization. In the sequential implicit sampling, the sample mean $\theta_{(1)}$ from the previous time step provides a "good" initial guess for the optimization. It provides a nature way to perform the optimization sequentially. In [42], the sequential optimization is done by approximating the object functions on multiple grids. Here we avoid additional discretization of the forward model PDE.

Similarly, we can sample the subsequent posteriors

$$\begin{aligned} & p_l(\theta|b^{1:M_l}) \propto p_0(\theta)\Pi_{m=1}^{M_l} p(b^m|X^{n_m}(\theta)) \\ = \;& p_0(\theta)\Pi_{m=1}^{M_{l-1}} p(b^m|X^{n_m}(\theta))\Pi_{m=M_{l-1}+1}^{M_l} p(b^m|X^{n_m}(\theta)) \\ \propto \;& p_{l-1}(\theta|b^{1:M_{l-1}})\Pi_{m=M_{l-1}+1}^{M_l} p(b^m|X^{n_m}(\theta)) \\ \approx \;& \exp\left[-\frac{1}{2}(\theta-\theta_{(l-1)})^T V_{\theta_{(l-1)}}^{-1}(\theta-\theta_{(l-1)})\right]\Pi_{m=M_{l-1}+1}^{M_l} p(b^m|X^{n_m}(\theta)), \end{aligned} \tag{2.28}$$

where $\theta_{(l-1)}$ and $V_{\theta_{(l-1)}}$ are the $(l-1)$-th sample mean and the $(l-1)$-th sample covariance matrix respectively. This process in completed until all windows are sampled or certain convergence has been observed, see Section 2.7. We present the algorithm for sequential implicit sampling in Algorithm 8.

Algorithm 8 Sequential Implicit Sampling

1: Generate samples of (2.27) and assign weights with implicit sampling method;
2: **for** each integer $l = 2,3,\ldots,K$ **do**
3: Resampling with weights;
4: Form the sample mean θ_{l-1} and sample variance $V_{\theta_{l-1}}$;
5: Generate samples of (2.28) and assign weights with implicit sampling method.
6: **end for**

Our sequential implicit sampling method is different from the particle filter, where the state variables have their forward dynamics. Given the particles at the previous time step, the prior density at current time step can be obtained from the forward model. Here the parameter θ is a set of static parameters and the prior density at current time step is only given by a set of samples (particles). Our algorithm proposed here is a hybrid of the EnKF and the implicit sampling method. We use a Gaussian to approximate the prior density given by the samples but this is the only Gaussian approximation used here. If, based on the statistics of those samples, some other density is more suitable to represent the prior distribution, we can easily change the prior to that density and the rest of the algorithm does not make any essential changes. This is different from the EnKF, where the Kalman formula is used in the analysis step and therefore some further Gaussian assumption is necessary.

The sequential implicit sampling method has some similarity with some hybrid algorithms, which combine the EnKF and four-dimensional Variational method (4DVAR) [52]. In the 4DEnKF [27] [30] or the En4DVAR [36], the same idea of breaking all data into observation windows has been exploited. The background covariance (B matrix) in 4DVAR is constructed from the ensemble forecasts, which is similar to our prior distribution that is constructed from the samples obtained from the previous step. In those algorithms, the optimization in the 4DVAR is realized in the subspace spanned by the ensemble to avoid the adjoint models in the implementation. Here, to compare them with our sequential implicit sampling method, we denote the hybrid algorithm as the En-4DVAR, which does the optimization over the full space in our numerical experiments, see Section 2.7. The exploration of the sequential implicit sampling method, with possible optimization in a subspace, will be given in our future study.

2.5 Application to an Inverse Seismic Wave Problem

As a test of the performance of various sampling algorithms introduced in this chapter, we apply these methods to an inverse seismic wave problem in which the logarithm of the stiffness parameter needs to be estimated. The following forward model is taken from [38], where the PDE describes the seismic wave propagation into the subsurface medium after an explosion hits the the surface. The stiffness parameter we estimate could be viewed as a description of the subsurface medium property and could be used to detect possible underground deposits. We will mainly focus on the performance of implicit sampling and sequential implicit sampling while comparing them with other traditional methods.

2.5.1 Problem Setup

We consider the following one dimensional wave equation on $\Omega_s \times \Omega_t = [0,L] \times [0,T]$:

$$\begin{cases} \rho u_{tt}(x,t) - (\lambda(x)u_x(x,t))_x = F(t)\delta(x-0), \\ \lambda(L)u_x(L,t) = -\sqrt{\rho\lambda(L)}u_t(L,t), \\ \lambda(0)u_x(0,t) = 0, \\ u(x,0) = 0, \\ u_t(x,0) = 0, \end{cases} \tag{2.29}$$

where L is the maximum depth. The value of $u(x,t)$ is wave intensity at the spatial point x and time t. To mimic the surface explosion, we have a pulse term $F(t)$ which is chosen as the Richer wavelet, also known as Mexican hat wavelet:

$$F(t) = \frac{2}{\sqrt{3\sigma}\pi^{\frac{1}{4}}}(1 - \frac{t^2}{\sigma_f^2})\exp(-\frac{t^2}{2\sigma_f^2}). \tag{2.30}$$

The function $\delta(x-0)$ is the dirac-delta function to ensure that the explosion is only limited on the surface.

The parameter ρ represents the density of the medium and λ is the stiffness parameter that we need to estimate. We assume that λ is time-invariant but different points in the subsurface medium have different stiffness. Namely, the stiffness parameter λ is a function of the depth x only. Since the stiffness parameter is positive, we assume that $\lambda(x) = e^{\theta(x)}$ and we will estimate the logarithm of stiffness, i.e. $\theta(x)$.

In our numerical experiment, we take $\rho = 1$ in (2.29) and $\sigma_f = 2$ in (2.30). The total depth of measurement L is 4 and the total observation time T is 6, i.e. the PDE is defined on $\Omega_s \times \Omega_t = [0,4] \times [0,6]$. The space domain and the time domain are divided equally into 128 intervals and 120 intervals, respectively. We assume that θ is a constant on each spacial interval and the dimension for the parameter $N_\theta = 128$. The forward model (2.29) is discretized using a piecewise linear finite

element in space and backward finite difference in time. The observation is the measurements of the surface intensity $u(0,t)$, which are perturbed with a Gaussian random variable ε. We assume that we have observations at every time step:

$$b_i = u(0,t_i) + \varepsilon_i = h_i(\theta) + \varepsilon_i, \quad i = 1,2,\ldots,120, \tag{2.31}$$

where ε_i's are i.i.d. Gaussian random variables $\sim N(0,\sigma^2)$ and $h_i(\theta) = u(0,t_i)$. In this paper, we choose $\sigma = 1 \times 10^{-3}$ so that the average error of measurements is approximately 2%.

2.5.2 The Prior, Likelihood Function, and Posterior

The prior density demonstrates how much information we have for the parameters before any experiment is performed. In our experiment, the prior is chosen as $p_0(\theta) \sim N(0, I_{N_\theta})$, where I_{N_θ} is the $N_\theta \times N_\theta$ identity matrix.

By (2.31), the likelihood function is

$$p(b|\theta) \quad \propto \quad \exp\left[\frac{1}{2\sigma^2}\sum_{i=1}^{120}(b_i - h_i(\theta))^2\right],$$

and by the Bayesian rule, the posterior density of θ is:

$$\begin{aligned} p(\theta|b) &\propto p_0(\theta)p(b|\theta) \\ &= \exp\left[-\frac{1}{2}\theta^T\theta - \frac{1}{2\sigma^2}\sum_{i=1}^{120}(b_i - h_i(\theta))^2\right]. \end{aligned} \tag{2.32}$$

We will investigate this posterior density Gaussian approximates, MCMC and implicit sampling as discussed above. Numerical results will be provided in the next section.

Even the prior distribution and the observation errors are Gaussian, the posterior density is not Gaussian due to the nonlinearity of the observation function $h_i(\theta)$. For example, in Figure 2.2, we give the contour plot of the posterior probability density function for a 2D problem. Here we divide the $[0,L]$ into two equal-length subintervals and assume that θ is a constant on each

subinterval. The space discretization is $N_x = 128$. If the posterior density were Gaussian, the contour plot would be composed of ellipses. However, in Figure 2.2, we can observe a substantial deviation from ellipses and the posterior density shows a clear skewness. This is an indication that the posterior is non-Gaussian.

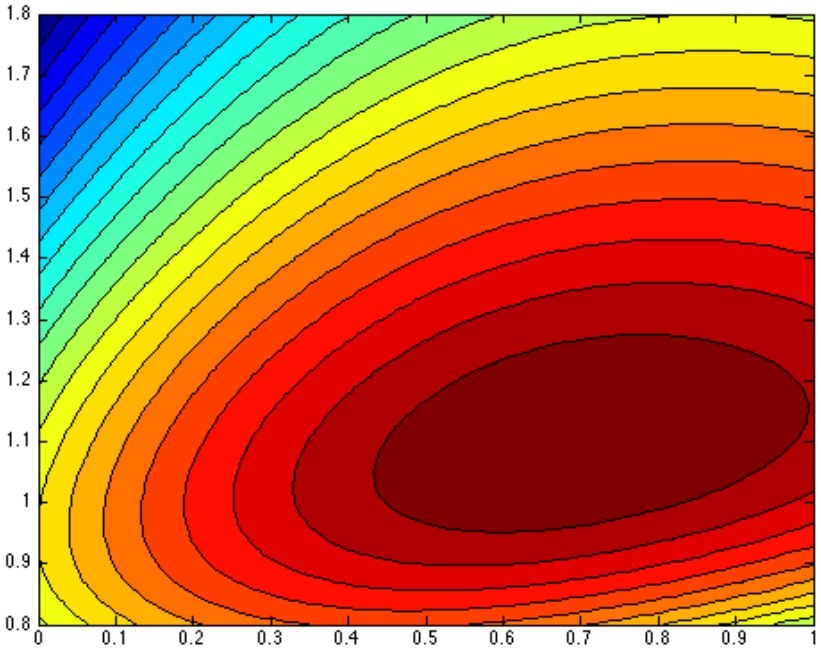


Figure 2.2: A contour visualization of the posterior probability density function for a 2D problem. Notice that the contours significantly deviate from ellipses, which is an indication of non-Gaussianity of the posterior.

2.6 Implementation of Implicit Sampling and Sequential Implicit Sampling

While the implementations of Gaussian approximations and MCMC are relatively straightforward, the implementation of implicit sampling and sequential implicit sampling needs special considerations. The costly and difficult part of implicit sampling lies in the fact that we need to find the MAP point, the minimizer of the negative logarithms of the posterior density defined in (2.32).

There are different optimization methods which can be used to perform the optimization, see [43]. Here we use the quasi-Newton method BFGS [43, Chapter 3], which only requires the gradient of the object function.

By (2.32), the negative logarithms of posterior is

$$F(\theta) = -\log p(\theta|b) = \frac{1}{2}\theta^T\theta + \frac{1}{2\sigma^2}\sum_{i=1}^{120}(b_i - h_i(\theta))^2 \tag{2.33}$$

and the gradient is

$$\nabla_\theta F(\theta) = \theta + \frac{1}{\sigma^2}\sum_{i=1}^{120}(b_i - h_i(\theta))\nabla_\theta h_i(\theta).$$

To evaluate the gradient, we need to calculate $\nabla_\theta h_i(\theta)$, which involves the derivatives of the forward model solution with respect to the parameters. This could be derived either from finite difference approximation or from an adjoint method [38]. The former can be used in low dimension case. For large scale problems, we choose the adjoint method which is a consequence of the Lagrange multiplier. The adjoint equation, after integration by parts, is

$$\begin{cases} \rho q_{tt}(x,t) - (\lambda(x)q_x(x,t))_x = -\dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{120}(q(x,t)-b_i)\delta(x-0)\delta(t-t_i) \\ \lambda(L)q_x(L,t) = \sqrt{\rho\lambda(L)}q_t(L,t) \\ \lambda(0)q_x(0,t) = 0 \\ q(x,T) = 0 \\ q_t(x,T) = 0 \end{cases} \tag{2.34}$$

The gradient is computed as the following integration:

$$\begin{aligned} g &= \mu + \int_0^T\int_0^L [\nabla\mu(x)]q_x(x,t)u_x(x,t)dxdt \\ &\quad + \int_0^T \frac{1}{2}\sqrt{\frac{\rho}{\mu(L)}}[\nabla\mu(L)]q(L,t)u_t(L,t)dt \end{aligned} \tag{2.35}$$

For the details we refer to [38]. As a consequence, we only need to run two PDE solvers (i.e. one

Methods	1st	2nd	3rd	4th	5th	6th	Total
Window: 120							4.0×10^3
Window: 40	810	2262	638				3.7×10^3
Window: 30	536	784	614	560			2.5×10^3
Window: 20	382	504	688	692	268	240	2.8×10^3

Table 2.1: The cost comparison for optimization among different sampling windows. The numbers are the forward runs, which have been converted to the full time runs.

forward model and one backward model) to obtain the gradient, which reduces the computation substantially.

In order to further improve the efficiency of the BFGS optimization, [42] uses multiple grids. The forward model is discretized at different coarse to fine level grids. The optimization is first done with the forward model discretized at the coarse grid and the result is used as an initial guess for the optimization based on the finer grid. In the sequential implicit sampling method, the sampling result from previous step can be used as an initial guess for the current step and no additional discretization of the forward model is needed.

To test the performance of the optimization in implicit sampling and sequential implicit sampling methods, we performed four tests with different sampling windows, namely 20, 30, 40 and 120. We note that the 120-window takes all data and it is the original implicit sampling method. In the test, we generate 200 samples for each choice of the sampling window. The initial guess, for the optimization in the next time window, is the sample mean from the previous time window calculation. Table 2.1 lists the forward model runs in each step for the optimization in sequential implicit sampling with different sampling window selections. All numbers have been converted to the forward model solvers with the full time window. We observed that the total number of the forward model runs for the 30-window is about 60% of the non-sequential implementation.

After the optimization, the second step in implicit sampling and sequential implicit sampling methods is to generate samples. We can use either the linear or the random map. Both methods require the Hessian of F, defined in (2.33), at the minimizer. The finite difference approximation of the Hessian will need $N_\theta(N_\theta+1)$ forward model runs and it is very expensive when N_θ is large.

In our numerical experiment, $N_\theta = 128$ and it needs 16512 forward model runs. Here we adopt an approximated Hessian which is commonly used in the LMAP method [31],

$$H \approx \bar{H} = I - Q^T (QQ^T + \Gamma_e)^{-1} Q, \tag{2.36}$$

where Q is the gradient of the observation map with respect to the parameter and Γ_e is the covariance matrix of the observation errors. In our numerical experiment, Γ_e is the diagonal matrix and we use the finite difference method to approximate Q, which requires $N_\theta + 1$ forward model runs. After finding the approximated Hessian, the linear map generates samples by (2.25) with weights calculated by (2.26). Each sample requires only one forward model run. The random map generates samples by solving (2.21) with the ansatz (2.23). We use the Newton method to solve those nonlinear equations. The main cost in the Newton's iteration is the calculation of the gradient and it needs two forward model runs by the adjoint method. In our numerical experiment, we observed that, with the initial guess $0.1 \times ||\xi||$ of λ, we can obtain a fast convergence with approximately 1 to 3 Newton steps for each sample. Therefore it needs 2 to 6 forward model runs for each sample using the random map. Recall that ξ is the sample from the reference density, see Algorithms 6 and 7.

2.7 Numerical Results

In this section, we will present the sampling results for the inverse wave posterior (2.32). We consider various sampling techniques for comparison, including the implicit sampling, sequential implicit sampling, the Ensemble Kalman Filter, Metropolis-Hastings MCMC and En-4DVAR method. The cost of all algorithms will be measured by the number of full time window forward model runs, since they are the most expensive part among all costs. In the numerical experiment, we take $N_\theta = 128$ and present the results for one of the parameters, which is the furthest from the surface, i.e. θ_{128}. We set the true value $\theta_{128} = 2$. We first study our sequential implicit sampling method with different setup in detail and then compare the optimal choice for the sequential

Methods	Mean	Std	Optimization Cost	Sampling Cost
RM	2.004	1.30×10^{-3}	4.0×10^3	1.0×10^3
LM	2.005	1.40×10^{-3}	4.0×10^3	3.3×10^2

Table 2.2: The cost comparison for generating 200 samples using the implicit sampling with the random map (RM) and the linear map (LM). The true value: $\theta_{128} = 2$

implicit sampling method with other sampling methods.

2.7.1 The Sequential Implicit Sampling

We first test the performance of the linear and random maps in the original implicit sampling method. These two methods are used to generate the samples around the high probability regions of the posterior density. Each method generates 200 samples and the sample mean and standard deviation of θ_{128} are provided in Table 2.2. According to Table 2.2, both methods provide similar results and the random map gives slightly smaller variance.

The quality of the samples generated by the linear or random maps can be evaluated using the effective sample size [4], the number of the sample size divided by R,

$$R = \frac{E(w^2)}{(E(w))^2}.$$

The variance of the weights is equal to $R-1$ and a good ensemble has a small variance of the weights. We repeat the experiments 10 times and obtain that R is 12.5 ± 0.5 for the random map and 14.3 ± 0.6 for the linear map. Both methods demonstrated similar sample qualities. However, to generate a sample, the cost of the random map is about three times as that for the linear map. Therefore we will use the linear map to generate the samples in our implicit sampling and sequential implicit sampling.

Secondly, we study the effect of the size of the sampling window in the sequential implicit sampling method. The shorter the window, the more frequently the observation data will be used to update the prior, and therefore the prior can be more accurate. On the other hand, with shorter windows, the samples need to be generated more frequently and this leads to more forward model

Methods	Mean	Standard Deviation	Cost
Window: 120	2.005	1.40×10^{-3}	4.3×10^3
Window: 40	1.995	1.70×10^{-3}	4.3×10^3
Window: 30	2.003	1.40×10^{-3}	3.3×10^3
Window: 20	2.010	1.80×10^{-3}	3.9×10^3

Table 2.3: Comparison among the sequential implicit sampling method with different sampling windows. The 120-window corresponds to all data. The true value: $\theta_{128} = 2$

Statistics	1st	2nd	3rd	4th
Mean	6.45×10^{-4}	2.411	2.010	2.003
Std	6.15×10^{-3}	5.02×10^{-3}	1.99×10^{-3}	1.40×10^{-3}

Table 2.4: Convergence of the mean and standard deviation for the sequential implicit sampling method with 30 sampling window.

runs. Moreover, the length of the sampling window will affect the optimization and result in different optimization cost, as we discussed in Section 3.3. We performed four different sampling windows: 20, 30, 40 and 120-window. Again, the 120-window corresponds to the original implicit sampling method. We generate 200 samples for each test case and the results are provided in Table 2.3. The results show that all window sizes give consistent statistics in terms of the mean and the standard deviation. Figure 2.3 gives the estimates of the mean with the number of the data for different sampling windows. Here all methods yield good estimations after approximately 90 data points are collected.

In Table 2.4, we give the details of the mean and the standard deviation for each step in the 30-window case, which has the minimum cost among all. From the results, we can see that the third step already yields satisfactory estimate compared with the last step. Using the sequential methods, we could stop the collection of data if the estimate converges already. This is another advantage for the sequential implementation.

2.7.2 Comparison with Other Methods

In this subsection, we compare the results from different sampling methods, including the sequential implicit sampling (SIS-30), Markov Chain Monte Carlo (MCMC), our En-4DVAR, the hybrid

Methods	Mean	Standard Deviation	Cost
SIS-30	2.003	1.40×10^{-3}	3.3×10^3
MCMC	2.002	1.60×10^{-3}	2.5×10^4
En-4DVar	2.005	2.20×10^{-3}	2.8×10^3
EnKF 1000	2.423	3.21×10^{-2}	1.0×10^3
EnKF 10000	2.126	2.54×10^{-2}	1.0×10^4
EnKF 20000	2.122	2.40×10^{-2}	2.0×10^4

Table 2.5: The estimates and the cost of θ_{128} with different methods. SIS-30: Sequential Implicit Sampling with 30-window; MCMC: Markov Chain Monte Carlo; En-4DVar: our hybrid EnKF and 4DVAR; EnKF 1000, 10000, 20000: Ensemble Kalman Filter with 1000, 10000, 20000 ensembles. The true value: $\theta_{128} = 2$.

version of the EnKF and 4DVAR, and the Ensemble Kalman Filter (EnKF). Here the results for the sequential implicit sampling (SIS-30) and En-4DVAR are given by 200 samples with 30-window. The results are given in Table 2.5.

A large class of sampling methods for Bayesian inverse problems is the MCMC, where one constructs a Markov chain so that the posterior is its invariant density. We move the chain forward and draw samples after the chain reaches its steady state. We adopt the Metropolis-Hastings MCMC and an isotropic Gaussian proposal density is used. This method requires one forward model run per step. We tuned the step size so that the acceptance rate is approximately 30%. We start the chain at the MAP point. While we observed a relatively fast burn-in time of 2.5×10^4 steps for θ_{128}, other parameters did not show a settled state even after 5×10^4 steps. See Figure 2.4 for the convergence of the sample means for θ_1, θ_5, θ_{125} and θ_{128} . This is an indication of the complexity of the posterior density. We did not run the chain until all parameters are settled due to time consideration.

Table 2.5 shows that the sampling results for θ_{128} from the MCMC and SIS-30. Those statistics are very close and both methods cover the true value $\theta_{128} = 2$ within two standard deviations. However the latter is much faster than the former. Notice that the number of samples needed for the MCMC is $O(10^4)$ while 200 samples from SIS-30 provides similar accuracy. The samples from the SIS-30 are all independent while those of the MCMC can be correlated. This is another advantage of the SIS-30 over the MCMC. Moreover, it is worth mentioning that we can exploit

parallelizing computation to generate samples for the SIS-30 while it is quite hard for the MCMC.

We also compare the SIS-30 with our En-4DVar. Recall that our En-4DVar is also a Gaussian method where we simply drop the resampling step in Algorithm 8. Notice that the number of forward runs is saved by about 5×10^2 since no weight is needed for samples so that we do not need run any forward model for samples. The standard deviation for the SIS-30 is smaller than that of the En-4DVar due to resampling. Although both methods provide similar sample means, the En-4DVar fails to capture such non-Gaussian features as skewness and kurtosis.

In Figure 2.5, we present the kernel density functions of the θ_{128} from 10^4 samples of the MCMC, sequential implicit sampling method (SIS-30), and our En-4DVAR. We can easily see the posterior is not a Gaussian. For the MCMC, we captured a kurtosis of 3.07 and a skewness of -0.12 while in SIS-30, the kurtosis and skewness are 3.2 and -0.10 respectively. These two sampling methods correctly capture the non-Gaussian feature with the SIS-30 demonstrates a more peaked distribution.

Another popular method is the Ensemble Kalman filter for parameter estimation. As indicated in Table 2.5, EnKF does not work as good as implicit sampling and MCMC, especially when the sample size is small (e.g.1000). The mean from EnKF does not give accurate approximation of the true value, which is a consequence of the nonlinear structure of the problem. Even with large samples, the variance from EnKF is much larger than that of SIS-30 and MCMC. This is because the EnKF is making Gaussian assumptions and may overestimate the uncertainty. The EnKF has the advantage of implementation speed, but it has a lack of accuracy since the errors from non-Gaussian posterior accumulate over time.

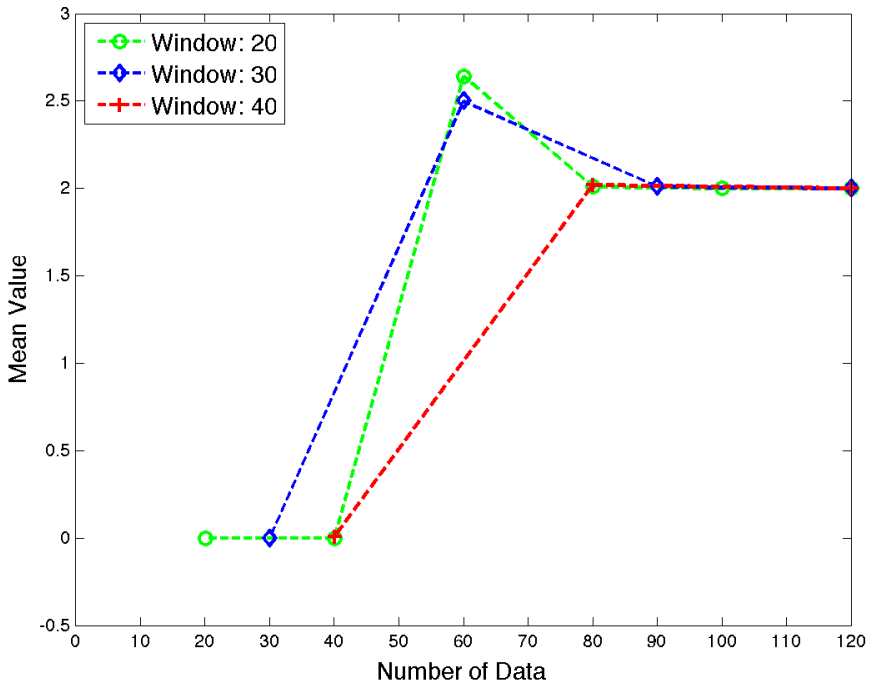


Figure 2.3: Convergence of mean value for different sampling windows. The 20,30,40-window schemes achieve desired convergence after the first 90 data are collected.

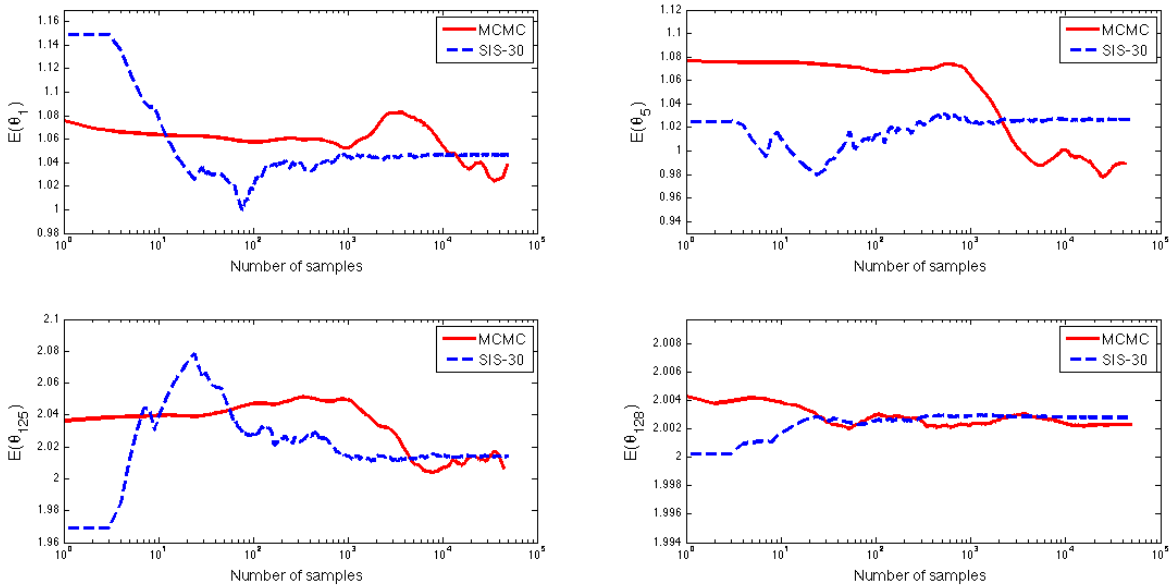


Figure 2.4: Convergence of the sample means of θ_1, θ_5, θ_{125} and θ_{128} with the number of samples obtained from the MCMC and SIS-30.

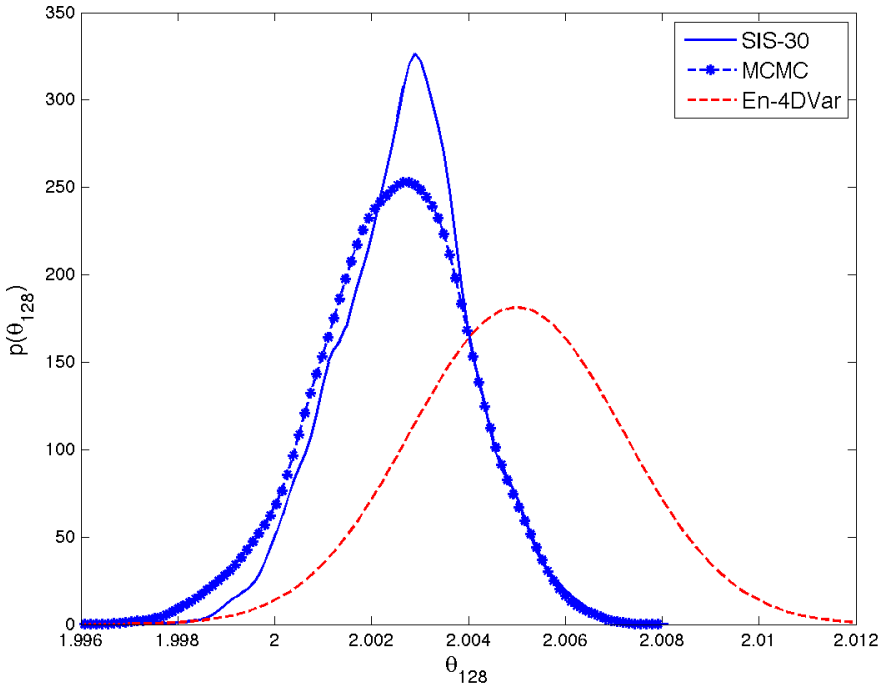


Figure 2.5: The kernel density estimation (KDE) for the marginals of θ_{128} of the posterior computed with the MCMC, SIS-30, and our En-4DVAR.

Chapter 3

Concluding Remarks

In this thesis, we studied parameter estimation problem for both stochastic differential equations (SDE) and partial differential equations (PDE). The approaches for the two types of equations are different.

For estimations in SDEs, we exploit the maximum likelihood estimator (MLE) which has desired properties such as consistency and asymptotic normality when the diffusion process is ergodic. A major obstacle for MLE is that the transition density can not be expressed in closed form for most equations. As a consequence, approximation of transition density is need in practice. We studied two existing approximations and introduced our method based on parametrix method. We showed the convergence of parametrix approximation to the true density and established the asymptotic behavior of approximated MLE under the ergodic assumption. There exists numerical test for parametrix approximation [13]. However, as the author indicated, some of the equations they considered did not fit into the framework of original parametrix, although the numerical solutions yielded satisfactory results. Here we generalized the original parametrix method and theoretically proved why their numerical solutions were working. However, more work on efficient ways to calculate the parametrix approximation is needed in the future.

For estimations in PDEs, we adopt the Bayesian framework which expresses the uncertainty in the parameters as a posterior probability density. Since numerical integration of the posterior is

computationally intractable for large scale problems, we instead resort to Monte Carlo integration which involves efficiently sampling the posterior. We reviewed some classic Monte Carlo methods and introduced our implicit sampling and sequential implicit sampling. We listed the algorithms of all methods for reader's convenience. We applied these methods in a seismic wave inversion problem where the numerical solutions showed a clear evidence that the sequential implicit sampling is superior to other traditional methods in terms of efficiency and accuracy. However, the optimizations step, which is part of implicit sampling, is still computationally expensive. The implicit sampling with optimization in a possible subspace will be studied in the future.

References

[1] J.Aitchison and S.D.Silvey. Maximum likelihood estimation of parameters subject to restraints *Ann.Math.Stat.*, 29:813–828, 1958.

[2] Y. Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70:223–262, 2002.

[3] C. Alexandre and X. Tu. An iterative implementation of the implicit nonlinear filter. *M2AN*, 46:535–543, 2012.

[4] M.Arulampalam, S.Maskell, N.Gordon and T.Clapp. A tutorial on particle filters for online nonlinear/nongaussian Bayesia tracking. *IEEE Trans. Sig. Proc.*, 50:174–188, 2002.

[5] E. Atkins, M. Morzfeld and A.J. Chorin. Implicit particle methods and their connection with variational data assimilation. *Monthly Weather Rev.*, 141:1786–1803, 2013.

[6] A. Beskos, A. Jasra, E.A. Muzaffer and A.Stuart. Sequential Monte Carlo methods for Bayesian elliptic inverse. *arXiv*:1412.4459, 2015.

[7] P. Bickel, B. Li and T. Bengtsson. Sharp failure rates for the bootstrap particle filter in high dimensions. *IMS Collections: Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh*, 3:318–329, 2008.

[8] T.Bui-Thanh, O.Ghattas, J.Martin and G.Stadler. A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion. *SIAM J. Sci. Comput.*, 35(6):A2494–A2523, 2013.

[9] S.Chib and E.Greenburg. Understanding the Metropolis-Hastings Algorithm. *The American Statistician,*, 49(4):327–335, 1995.

[10] A.J. Chorin and O.H. Hald. Monte Carlo Strategies for Scientific Computing. *Springer, New York*, 2008.

[11] A.J. Chorin and M. Morzfeld. Conditions for successful data assimilation. *J. Geophys. Res.:Atmospheres*, 118:11522–11533, 2013.

[12] A.J. Chorin, M. Morzfeld and X. Tu. Implicit particle filters for data assimilation. *Statist. Sci*, 28(3):424–446.

[13] F. Corielli, P. Foschi and A. Pascucci. Parametrix approximation of diffusion transition densities. *Siam J. Financial Math* , 1:833–867, 2010.

[14] S. L. Cotter, G. O. Roberts, A. M. Stuart and D. White. MCMC methods for functions: modifying old algorithms to make them faster. , 34(1):89–98, 2013.

[15] A.J. Chorin, M. Morzfeld, and X. Tu. Implicit sampling, with application to data assimilation. *Chin. Ann. Math. Ser. B*, 34(1):89–98, 2013.

[16] A.J. Chorin and X. Tu. Implicit sampling for particle filters. *Proc. Nat. Acad. Sc. USA*, 106:17249–17254, 2009.

[17] A.J. Chorin and X. Tu. Interpolation and iteration for nonlinear filters. *M2AN Math. Model. Numer. Anal.*, 46:535–543, 2012.

[18] Y. Cui, Y. Hu, C. Su and J. Tong. Pointwise ergodic theorems for Markov chains and skeletons of diffusions. *In progress*.

[19] D.Dacunha-Castelle and M. Duflo. Probability and statistics II. *Springer-Verlag*, 1986.

[20] D.Dacunha-Castelle and D.Florens-Zmirou. Estimation of the coefficients of a diffusion from discrete observations. *Stochastics*, 19:263–284, 1986.

[21] T.Deck and S.Kruse. Parabolic differential equations with unbounded coefficients-a generalization of the parametrix method. *Acta Applicandae Mathematicae*, 74:71–91, 2002.

[22] G.B.Durham and A.R.Gallant. Numerical techniques for maximum likelihood estimation of continuous-time diffusions processes. *Journal of Business and Economic Statistics*, 20:297–316, 2002.

[23] G.Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99:10143–10162, 1994.

[24] G.Evensen. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53: 343–367, 2003.

[25] G.Evensen. Data assimilation: the ensemble Kalman filter. *Springer*, 2006.

[26] R. Faragher. Understanding the basis of the kalman filter via a simple and intuitive derivation. *IEEE Signal Processing Magazine*, 29(5):128–132, 2012.

[27] E.J. Fertig, J.Harlim and B.R. Hunt. A comparative study of 4D-Var and a 4D ensemble Kalman filter: perfect model simulations with Lorenz-96. *Tellus*, 59(A):96–100, 2007.

[28] A.Friedman. Partial differential equations of parabolic type. *Prentice-Hall, Englewood Cliffs, NJ*, 1964.

[29] Y. Hu and S. Watanabe. Donsker's delta functions and approximation of heat kernels by time discretization methods. *J. Math. Kyoto University*, 36:499–518, 1996.

[30] B.R. Hunt, E.Kalnay, E.J. Kostelich, E.Ott, D.J. Patil, T.Sauer, I.Szunyogh, J.A. Yorke and A.V. Zimin. Four-dimensional ensemble Kalman filtering. *Tellus*, 56(A):273–277, 2004.

[31] M.Iglesias, Law K., and A.M. Stuart. Evalutaion of Gaussian approximations for data assimilation in reservoir models. *Comput. Geosci.*, 17:851–885, 2013.

[32] N.Kantas, A.Beskos and A.Jasra. Sequential Monte Carlo methods for high-dimensional inverse problems: a case study for the Navier-Stokes equations. *SIAM/ASA J. Uncertain. Quantif.*, 2(1):464–489, 2014.

[33] I.Karatzas and S.E.Shreve. Brownian motion and stochastic calculus. *Springer-Verlag.*, 1991.

[34] Y.Kutoyants. Statistical inference for ergodic diffusion processes. *Springer-Verlag.*, 2004.

[35] O. A. Ladyzenskaya. Linear and quasi-linear equations of parabolic type. *American Mathematical Society*, 1995

[36] C.Liu, Q.Xiao and B.Wang. An ensemble-based four-dimensional variational data assimilation scheme. Part I: technical formulation and preliminary test. *Monthly Weather Review*, 136:3363–3373, 2008.

[37] J.S. Liu. Stochastic tools in mathematics and science. *Springer, New York*, 2013.

[38] J.Martin, L.C. Wilcox, C.Burstedde, and O.Ghattas. A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. *SIAM J. Sci. Comput.*, 34(3):A1460–A1487, 2012.

[39] M.Morzfeld, X.Tu, E.Atkins and A.J. Chorin. A random map implementation of implicit filters. *Journal of Computational Physics*, 231(4):2049–2066, 2012.

[40] M. Morzfeld and A.J. Chorin. Implicit particle filtering for models with partial noise, and an application to geomagnetic data assimilation. *Nonlin. Processes Geophys.*, 19:365–382, 2012.

[41] M. Morzfeld, X. Tu, E. Atkins and A.J. Chorin. A random map implementation of implicit filters. *Journal of Computational Physics*, 231(4):2049–2066, 2012.

[42] M. Morzfeld, X. Tu, J. Wilkening and A.J. Chorin. Parameter estimation by implicit sampling. *Commun. Appl. Math. Comput. Sci.*, 10:205–225, 2015.

[43] J.Nocedal and S.T. Wright. Numerical optimization. *Springer, second edition*, 2006.

[44] D. Oliver, A. Reynolds and N. Liu. Inverse theory for petroleum reservoir characterization and history matching. *Cambridge university press*, 2008.

[45] A.R.Pedersen. A new approach to maximum likelihood estimation for stochastic differential equations based on discreet observations. *Scandinavian Journal of Statistics*, 22:55–71, 1995.

[46] A.R.Pedersen. Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes. *Bernoulli*, 1(3):257–279, 1995.

[47] N.Petra, J.Martin, G.Stadler and O.Ghattas. A computational framework for infinite-dimensional Bayesian inverse problems, Part II: Stochastic Newton MCMC with application to ice sheet flow inverse problems. *SIAM J. Sci. Comput.*, 36(4):A1525–A1555, 2014.

[48] D.Revuz. Markov chains. *North-Holland Pub.*, 1975.

[49] M. Ribeiro. Kalman and extended kalman filters: concept, derivation and properties. *Institute for Systems and Robotics, IST.*, 2004.

[50] C. Snyder, T. Bengtsson, P. Bickel and J. Anderson. Obstacles to high-dimensional particle filtering. *Mon. Wea. Rev.*, 136:4629–4640, 2008.

[51] A.M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010.

[52] O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478):1311–1328, 1987.

[53] X.Tu and C. Su. Sequential implicit sampling methods for bayesian inverse problems. *submitted.*

Appendix A

Markov Chain Basics and Convergence Theorems

Let $(E,\mathscr{E})$ be a measurable space and let $P(x,B) : E \otimes \mathscr{E} \to \mathbb{R}$ be a Markovian kernel, namely, $P(x,B)$ satisfies the following conditions:

1. For any fixed $B \in \mathscr{E}$, $P(\cdot,B)$ is measurable with respect to $\mathscr{E}$.

2. For fixed $x \in E$, $P(x,\cdot)$ is a probability measure on $(E,\mathscr{E})$.

Denote by $\mathrm{b}\mathscr{E}$ the set of all bounded measurable functions on $(E,\mathscr{E})$ and denote by $\mathrm{b}.\mathscr{M}$ the set of all bounded measures on $(E,\mathscr{E})$. Then we can define $P : \mathrm{b}\mathscr{E} \to \mathrm{b}\mathscr{E}$ and/or $P : \mathrm{b}.\mathscr{M} \to \mathrm{b}.\mathscr{M}$ in the following ways:

$$Pf(x) = \int_E p(x,dy) f(y)\,, \quad \forall\, f \in \mathrm{b}\mathscr{E}$$

and

$$\nu P(A) = \int_E \nu(dx) P(x,A)\,, \quad \forall\, \nu \in \mathrm{b}.\mathscr{M}\,.$$

Obviously, P can be extended more general measurable functions and more general signed measures.

Remark 7. *We use the same letter P to denote the Markov kernel and its two induced (linear)*

operators on $b\mathscr{E}$ and on $b\mathscr{M}$. This is because it will not cause confusion while it makes the notation concise.

Let μ be a σ-finite positive measure on $(E,\mathscr{E})$. Then we can identify the space of bounded positive measures that are absolute continuous with respect to μ as the space $L^1_+(E,\mathscr{E},\mu)$. Moreover, for any measurable functions $f \in L^1(\mu)$ and $g \in L^\infty(\mu)$, we denote

$$\langle f,g\rangle_\mu = \int_E f(x)g(x)\mu(dx)\,.$$

This notation is also valid when $f \in L^p(\mu)$ and $g \in L^q(\mu)$ with p and q being conjugate each other ($1/p+1/q=1$, $1 \le p,q \le \infty$).

Lemma 8. *The space $b\mathscr{M}_\mu$ of bounded measures that are absolutely continuous with respect to μ is invariant by the above operator P (namely $b\mathscr{M}_\mu P \subseteq b\mathscr{M}_\mu$) if and only if $\mu P \ll \mu$.*

Proof. First, we show the "if" part. Assume $\mu P \ll \mu$. Then $\mu(A)=0$ implies $\mu P(A)=0$, which in turn implies that $P(\cdot,A)=0$, μ-a.e. Now if $\nu = f\mu$ is an element in $b\mathscr{M}_\mu$ with $f \in L^1_+(E,\mathscr{E},\mu)$, then

$$\nu P(A) = \int_E \nu(dx)P(x,A) = \int_E f(x)\mu(dx)P(x,A) = 0\,.$$

This shows that $\nu P \ll \mu$.

Now we show the "only if" part. Let $\mu(A)=0$. Then $\nu(A)=0$ for any $\nu \in b\mathscr{M}_\mu$. Since μ is σ-finite, there exists a sequence $f_n \nearrow 1$ with $f_n \in L^1(\mu)$. It is easy to see that $\nu_n = f_n(x)\mu$ is in $b\mathscr{M}_\mu$, which is invariant by P. Thus, we have

$$\int_E f_n(x)\mu(dx)P(x,A)dx = (\nu_n P)(A) = 0\,.$$

Letting $n \to \infty$ yields $\mu P(A)=0$. □

Definition 3. *Let $\mu P \ll \mu$. For $f \in L^1(\mu)$, we define the (linear) operator T induced by P as*

follows

$$Tf = \frac{d[(f\mu)P]}{d\mu}.$$

Proposition 3. *T is a positive contraction on $L^1(\mu)$ (positivity means that $f \geq 0$ implies $Tf \geq 0$).*

Proof. The positivity is clear. The contraction follows from the fact that $f \to f\mu$ is an isometry from $L^1(\mu)$ to b$\mathscr{M}_\mu$ and P is a contraction from b$\mathscr{M}_\mu$ to b$\mathscr{M}_\mu$. □

It is known that if T is an operator on $L^1(\mu)$ (when we say operator we always mean the linear operator in this paper), then its adjoint, T^*, is an operator on $L^\infty(\mu)$. In particular, we have

Proposition 4. *Let T be induced by P. Then $T^* = P$ in the sense that $T^*g = Pg$ μ-a.e. for all $g \in L^\infty(\mu)$.*

Proof. For any $f \in L^1(\mu)$, $g \in L^\infty(\mu)$,

$$\begin{aligned}\int_E f(x)(T^*g)(x)\mu(dx) &= \int_E (Tf)(x)g(x)\mu(dx)\\ &= \int_E \frac{d[(f\mu)P]}{d\mu}(x)g(x)\mu(dx) = \int_E g(x)\,[(f\mu)P]\,(dx)\\ &= \int_E\int_E f(y)\mu(dy)P(y,dx)g(x) = \int_E f(x)\left[\int_E P(x,dy)g(y)\right]\mu(dx)\end{aligned}$$

which yields $T^*g = Pg$ μ-a.e. □

Definition 4. *A function $\varphi \in L^\infty_+(\mu)$ is called superharmonic if $T^*\varphi \leq \varphi$ μ-a.e.*

Using the above lemma, we see that $\varphi \in L^\infty_+(\mu)$ is superharmonic iff $P\varphi \leq \varphi$ μ-a.e. It is apparent that the constant function 1 is superharmonic.

Theorem 8 (Maximal Ergodic Lemma)**.** *Let $f \in L^1(\mu)$ and let*

$$E_f = \bigcup_{n=1}^{\infty}\left\{\sum_{k=1}^{n} T^k f > 0\right\}. \tag{A.1}$$

Then for any superharmonic function φ,

$$\int_{E_f} \varphi(x) f(x) \mu(dx) \geq 0 . \tag{A.2}$$

Proof. Denote the following function from E to $\mathbb{R}$:

$$h_N = \sup\{0, f, f+Tf, f+Tf+T^2 f, \ldots, f+Tf+T^2 f+\ldots, +T^N f\}, \quad N = 0,1,2,\cdots$$

Then h_N is an increasing family of functions (increases in N) and $E_N \triangleq \{h_N > 0\} \uparrow E_f$. Next, we see easily

$$Th_N \geq \sup\{0, Tf, \ldots, Tf+T^2 f+\cdots+T^{N+1} f\}$$

which immediately implies

$$f + Th_N \geq \sup\{f, f+Tf, \ldots, f+Tf+T^2 f+\cdots+T^{N+1} f\} .$$

Note that the right hand side of above inequality is h_{N+1} on E_{N+1}. This means that $f I_{E_{N+1}} \geq h_{N+1} I_{E_{N+1}} - Th_N I_{E_{N+1}}$. Therefore, we have

$$\begin{aligned}
\int_{E_{N+1}} \varphi f d\mu &\geq \int_E \varphi h_{N+1} d\mu - \int_{E_{N+1}} \varphi T h_N d\mu \\
&\geq \int_E \varphi h_{N+1} d\mu - \int_E \varphi T h_N d\mu \\
&\geq \int_E \varphi h_{N+1} d\mu - \int_E (T^* \varphi) h_N d\mu \\
&\geq \int_E \varphi h_{N+1} d\mu - \int_E \varphi h_N d\mu \geq 0 .
\end{aligned}$$

Letting $N \to \infty$ yields the desired inequality. □

In ergodic theory, we are particularly interested in the following inequality:

$$D_n(f,g) = \frac{\sum_{k=0}^n T^k f}{\sum_{k=0}^n T^k g} \tag{A.3}$$

Later on we will show that under certain conditions on T, this ratio will converge P_μ-a.e. We still need a few preliminaries before we get to this point. To begin with, we have

Lemma 9. *Let $f,g \in L^1_+(\mu)$. Then $g=0$ on $F=\left\{\sup_{n\geq 1} D_n(f,g)=\infty\right\}$.*

Proof. For any $c>0$, applying the maximum ergodic lemma with $\varphi=1$ and $f-cg$, we obtain $\int_{E_{f-cg}}(f-cg)d\mu \geq 0$ so that

$$\int_F g d\mu \leq \int_{E_{f-cg}} g d\mu \leq c^{-1}||f||_{L^1} < \infty$$

since $F \subset E_{f-cg}$ for any $c>0$. Letting $c\to\infty$, we obtain $\int_F g d\mu = 0$. This means that $g=0$ on F. □

We next introduce the Hopf decomposition of E with respect to the positive contraction T:

Theorem 9 (Hopf decomposition theorem)**.** *There exists a set $C \subseteq E$, unique up to μ-equivalence class, such that for any $f \in L^1_+$,*

(1) $\sum_{n=0}^\infty T^n f = 0$ or ∞ on C;

(2) $\sum_{n=0}^\infty T^n f < \infty$ on $D = C^c$.

Proof. Let $f,g \in L_1(\mu)$ and let f,g be strictly positive. First, we claim

$$\sum_{n=0}^\infty T^n f < \infty \quad \text{on} \quad \left\{\sum_{n=0}^\infty T^n g < \infty\right\}.$$

Indeed, if $\sum_{n=0}^\infty T^n f = \infty$, then on $\{\sum_{n=0}^\infty T^n g < \infty\}$, $\sup_{n\geq 1} D_n(f,g)=\infty$, which, by Lemma 9, implies that $g=0$ on $\{\sum_{n=0}^\infty T^n g<\infty\}$. This is a contradiction since g is strictly positive. Changing the role of f and g, we have

$$\sum_{n=0}^\infty T^n g < \infty \quad \text{on} \quad \left\{\sum_{n=0}^\infty T^n f < \infty\right\}.$$

Now let $C = \{\sum_{n=0}^{\infty} T^n f = \infty\}$ for $f \in L^1(\mu)$ and f strictly positive. From the above argument we see that C does not depend on the choice of f provided that $f > 0$.

We next show that if $h \in L^1(\mu)$ and $\sum_{n=0}^{\infty} T^n h < \infty$, then $\sum_{n=0}^{\infty} T^n h = 0$. Indeed, if $\sum_{n=0}^{\infty} T^n h < \infty$ on $B \subset C$, then for any integer $k > 0$, $\sum_{n=0}^{\infty} T^n (T^k h) < \infty$ on B. Therefore by the definition of C, for a strictly positive g, we have $\sup_{n\geq 1} D_n(g, T^k h) = \infty$ on B. But then by Lemma 9 we have $T^k h = 0$ μ-a.e. on B. Sine k is arbitrary, we see $\sum_{n=0}^{\infty} T^n h = 0$ μ-a.e. on B. □

Definition 5. *We call C and D the conservative part and the dissipative part of E with respect to the positive contraction T, respectively. If $C = E$ μ-a.e., then T is called conservative.*

Now we can introduce the concept of invariant σ-algebra for a conservative positive contraction T:

Proposition 5. *Let T be conservative and let the class $\mathscr{C}$ be defined as follows:*

$$\mathscr{C} = \left\{ C_f \,,\ f \text{ runs through } L^1_+(\mu) \text{ is a sub-}\sigma\text{-algebra of } \mathscr{E} \right\},$$

where

$$C_f = \left\{ \sum_{n=1}^{\infty} T^n f = \infty \right\}.$$

Then $\mathscr{C}$ is a sub-σ-algebra of $\mathscr{E}$. Moreover, for $h \in L^\infty_+$, the following statements are equivalent:

(1) $T^* h \leq h$, μ-*a.e.*

(2) $T^* h = h$, μ-*a.e.*

(3) $h \in \mathscr{C}$.

In particular, $T^* I_B = I_B$ on B implies $B \in \mathscr{C}$ and if $B \in \mathscr{C}$, then $T^* I_B = I_B$ μ-a.e.

Proof. We first prove the equivalence between (1) and (2). Clearly (2) implies (1). For the con-

verse, take $h \in L^\infty_+$ such that $T^*h \leq h$, μ-a.e. Let's also pick a strictly positive $f \in L^1(\mu)$, then

$$\begin{aligned}\langle \sum_{n=0}^{N-1} T^n f, h - T^*h \rangle &= \langle f, \sum_{n=0}^{N-1} (T^*)^n (h - T^*h) \rangle \\ &= \langle f, h - (T^*)^N h \rangle \leq ||f||_{L^1} ||h||_{L^\infty} < \infty.\end{aligned}$$

Letting $N \to \infty$, we have

$$\langle \sum_{n=0}^{\infty} T^n f, h - T^*h \rangle < \infty$$

Since T is conservative and $f > 0$, we have $\sum_{n=0}^{\infty} T^n f = \infty$ μ-a.e. This implies that $T^*h = h$, μ-a.e. In particular, we have $T^*1 = 1$.

Next, let H be the subspace of $L^\infty(\mu)$ such that $T^*h = h$ μ-a.e. Then $1 \in H$. By continuity of T^*, we see that $h_n \nearrow h$ with $h_n \in H$ implies $h \in H$. Moreover, $h, h' \in H$ implies $h \wedge h' \in H$. Indeed, take a such that $a + h \geq 0$ and $a + h' \geq 0$. Then

$$a + T*(h \wedge h') = T^*(a + h \wedge h') \leq (a + h) \wedge (a + h') = a + h \wedge h'.$$

Since $a + h \wedge h' \geq 0$, we have $T*(h \wedge h') = h \wedge h'$ by the equivalence of (1) and (2). By the monotone class theorem for functions, there exists a sub-σ-algebra $\mathscr{G}$ of $\mathscr{E}$ such that $H = b\mathscr{G}$, where $b\mathscr{G}$ is the space of bounded $\mathscr{G}$-measurable functions.

Before we identify $\mathscr{G}$, we see that if $B \in \mathscr{G}$, then $T^*I_B = I_B$ μ-a.e. If $T^*I_B = I_B$ on B, then $T^*I_{B^c} = 1 - T^*I_B = 0$ on B. Thus $T^*I_{B^c} \leq I_{B^c}$ and hence $T^*I_{B^c} = I_{B^c}$. As a consequence, B^c belongs to $\mathscr{G}$.

Finally, we show

$$\mathscr{G} = \mathscr{C} = \{C_f : f \in L^1_+\}.$$

First, from the Hopf decomposition theorem (Theorem 9) we have

$$0 = \langle I_{C_f^c}, \sum_{n=1}^{\infty} T^n f \rangle = \langle T^* I_{C_f^c}, \sum_{n=0}^{\infty} T^n f \rangle$$

Therefore $T^*I_{C_f^c} = 0$ on C_f which implies that $T^*I_{C_f^c} \le I_{C_f^c}$. This implies $C_f^c \in \mathscr{G}$. Thus, we have $C_f \in \mathscr{G}$. Conversely, take $B \in \mathscr{G}$ and choose $f \in L^1_+(\mu)$ such that $B = \{f > 0\}$. We show $B = C_f$. If $x \in B$, then $\sum_{n=0}^{\infty} T^n f \ge f(x) > 0$. By Theorem 9 we have $\sum_{n=0}^{\infty} T^n f = \infty$. Thus $B \subseteq C_f$. On the other hand, from the equivalence (1) and (2) and the fact that $B \in \mathscr{G}$, we have

$$\langle I_{B^c}, T^n f\rangle = \langle (T^*)^n I_{B^c}, f\rangle = \langle I_{B^c}, f\rangle = 0\,.$$

As a consequence, we have $\sum_{n=0}^{\infty} T^n f = 0$ on B^c. This implies $C_f \subseteq B$. The proof is then completed.

□

Definition 6. *The σ-algebra $\mathscr{C}$ defined in Proposition 5 is called the invariant σ-algebra of T. A sets in $\mathscr{C}$ is called invariant set and a $\mathscr{C}$-measurable function is called invariant function (denote also $h \in C\mathscr{C}$. If $\mathscr{C}$ is μ-trivial, namely, if for any $A \in \mathscr{C}$, $\mu(A) = 0$ or $\mu(A) = 1$, then the conservative positive contraction T is called ergodic under μ.*

Proposition 6. *If h is a bounded invariant function, then $T(hf) = hT(f)$ for any $f \in L^1(\mu)$.*

Proof. We first prove that if $g \in L^\infty(\mu)$, then $T^*(hg) = hT^*(g)$. We need to prove the above identity for $h = I_B$ and $g = I_A$ where $B \in \mathscr{C}$ and $A \in \mathscr{E}$. In this case we have

$$T^*(I_A I_B) \le \inf\{T^*I_A, T^*I_B\} = \inf\{T^*I_A, I_B\} = I_B(T^*I_A)\,.$$

Similarly, $T^*(I_{A^c}I_B) \le I_B(T^*I_{A^c})$. Adding these two inequalities gives $T^*I_B \le I_B$. The Proposition 5 implies $T^*I_B = I_B$. Thus

$$\begin{aligned} T^*I_B - T^*(I_{A^c}I_B) = T^*(I_A I_B) &\le I_B(T^*I_A) = (T^*I_B)(1 - T^*I_{A^c}) \\ &= T^*I_B - (T^*I_B)(T^*I_{A^c}) = T^*I_B - I_B T^*I_{A^c}\,. \end{aligned}$$

This implies $T^*(I_{A^c}I_B) \ge I_B(T^*I_{A^c})$ and then $T^*(I_{A^c}I_B) = I_B(T^*I_{A^c})$. Similarly, we have $T^*(I_A I_B) =$

$I_B(T^*I_A)$. By the monotone class theorem, we have

$$T^*(hg) = hT^*(g) \quad \forall \text{ bounded } h \in \mathscr{C} \text{ and } \forall\, g \in L^\infty(\mu)\,.$$

Again let $f \in L^1(\mu)$ and $g \in L^\infty(\mu)$ and let $h \in \mathscr{C}$ be bounded. Then, we have

$$\begin{aligned}\langle h(Tf), g\rangle &= \langle Tf, hg\rangle = \langle f, T^*(hg)\rangle = \langle f, h(T^*g)\rangle \\ &= \langle hf, T^*g\rangle = \langle T(hf), g\rangle\,.\end{aligned}$$

As a consequence, we have $T(hf) = hT(f)$. □

The following lemma is needed in the proof of Chacon-Ornstein Theorm:

Lemma 10. *Let $g \in L^1(\mu)$ be strictly positive. Then for any $f \in L^1(\mu)$ and for any $k \in \mathbb{N}$, we have*

$$\lim_{n\to\infty}\left(\frac{T^{n+k}f}{\sum_{m=0}^{\infty} T^m g}\right) = 0 \quad \mu\text{-almost everywhere as } n \to \infty.$$

Proof. Without loss of generality, we assume $k = 0$. Take a measure ν such that $d\nu/d\mu = g$. Since $g > 0$, we know $\nu \sim \mu$. For any $\varepsilon > 0$, set

$$f_0 = f - \varepsilon g, \ldots, f_n = T^n f - \varepsilon \sum_{m=0}^{n} T^m g$$

so that $f_n = Tf_{n-1} - \varepsilon g$. Set $A_n = \{f_n > 0\}$. We need only to show $\nu(\overline{\lim_n} A_n) = 0$ and this implies the desired convergence. To this end, notice $f_n^+ = f_n I_{A_n} \le Tf_{n-1}^+ - \varepsilon I_{A_n} g$. This means $\varepsilon I_{A_n} g \le Tf_{n-1}^+ - f_n^+$. Therefore

$$\begin{aligned}\varepsilon\nu(A_n) &\le \langle Tf_{n-1}^+, 1\rangle - \langle f_n^+, 1\rangle \\ &\le \langle f_{n-1}^+, 1\rangle - \langle f_n^+, 1\rangle\,.\end{aligned}$$

Thus

$$\varepsilon \sum_n \nu(A_n) \leq \int f_0^+ d\mu \leq f^+ d\mu < \infty$$

which implies that $\nu(\overline{\lim_n} A_n) = 0$ by the Borel-Cantelli lemma. □

Theorem 10 (Chacon-Ornstein theorem)**.** *Let T be a conservative positive contraction. Let $f, g \in L^1(\mu)$ with g strictly positive. Then*

$$D_n(f,g) \to \frac{E[f|\mathscr{C}]}{E[g|\mathscr{C}]} \quad \mu\text{-almost everywhere as } n \to \infty,$$

where $\mathscr{C}$ is the invariant σ-algebra.

Proof. First, we assume that μ is bounded. Let g be strictly positive and fixed. We divide the proof into several steps.

Step 1 If $f = hg$ with h being bounded and invariant, then $T^n f = hT^n g$ by proposition 6. Thus

$$D_n(f,g) = h \to h = \frac{E[f|\mathscr{C}]}{E[g|\mathscr{C}]}.$$

Step 2 If $f = (I-T)k$ with $k \in L^1(\mu)$, then for any $h \in \mathscr{C}$, we have

$$\langle h, f\rangle = \langle h, (I-T)k\rangle = \langle h,k\rangle - \langle h, Tk\rangle = \langle h,k\rangle - \langle T^* h, k\rangle = 0.$$

This yields $E[f|\mathscr{C}] = 0$ and hence $\frac{E[f|\mathscr{C}]}{E[g|\mathscr{C}]} = 0$. On the other hand, we have $\sum_{m=0}^n T^m f = k - T^{n+1}k$. An application of lemma 10 gives

$$\frac{\sum_{m=0}^n T^m f}{\sum_{m=0}^n T^m g} = \frac{k - T^{n+1}k}{\sum_{m=0}^n T^m g} \to 0.$$

Step 3 The subspace space $L_g = \left\{hg + (I-T)k, h \in \mathscr{C}, k \in L^1(\mu)\right\}$ of $L^1(\mu)$ is dense in $L^1(\mu)$. Indeed, if f is such that $f \perp L_g$, then $\langle f, (I-T)k\rangle = 0$ for all $k \in L^1(\mu)$. Thus $\langle f,k\rangle = \langle f, Tk\rangle = \langle T^* f, k\rangle$. Or $T^* f = f$. From Proposition 5, the equivalence between (2) and (3), we have $f \in \mathscr{C}$. But then $f \perp hg$ for all $h \in \mathscr{C}$. Therefore, $f = 0$ μ-a.s.

Step 4 Next, we show that if $f_p \in L^1(\mu)$ and $\sum_{p=1}^{\infty} ||f_p||_{L^1} < \infty$, then $\sup_n D_n(f_p, g) \to 0$ as $p \to \infty$. To see this, take any $\varepsilon > 0$ and apply the maximal ergodic lemma to $|f_p| - \varepsilon g$ and $\varphi = 1$. We obtain

$$\varepsilon \sum_{p=1}^{\infty} \int_{E_{|f_p|-\varepsilon g}} g d\mu \leq \sum_{p=1}^{\infty} \int |f_p| d\mu < \infty$$

With $d\nu = g d\mu$, the above inequality can be rewritten as $\sum_{p=1}^{\infty} \nu \left\{ E_{|f_p|-\varepsilon g} \right\} < \infty$. The Borel-Cantelli lemma yields

$$\nu\{\overline{\lim_p}(E_{|f_p|-\varepsilon g})\} = 0 \quad \text{and hence} \quad \mu\{\overline{\lim_p}(E_{|f_p|-\varepsilon g})\} = 0\,.$$

This implies $\sup_n D_n(f_p, g) \to 0$ as $p \to \infty$.

Step 5 Finally, let $f \in L^{\infty}(\mu)$ and take $f_p \in L_g$ such that $f_p \to f$ in $L^1(\mu)$ and $\sum_{p=1}^{\infty} ||f - f_p||_{L^1} < \infty$. Then

$$\begin{aligned}
\limsup_{n\to\infty} \left| D_n(f,g) - \frac{E[f|\mathscr{C}]}{E[g|\mathscr{C}]} \right| &\leq \limsup_{n\to\infty} \left| D_n(f,g) - D_n(f_p,g) \right| \\
&\quad + \sup_{n\to\infty} \left| D_n(f_p,g) - \frac{E[f_p|\mathscr{C}]}{E[g|\mathscr{C}]} \right| + \left| \frac{E[f_p|\mathscr{C}]}{E[g|\mathscr{C}]} - \frac{E[f|\mathscr{C}]}{E[g|\mathscr{C}]} \right| \\
&\leq \sup_{n\to\infty} \left| D_n(f_p,g) - \frac{E[f_p|\mathscr{C}]}{E[g|\mathscr{C}]} \right| + \frac{\left| E[f_p|\mathscr{C}] - E[f|\mathscr{C}] \right|}{E[g|\mathscr{C}]}\,.
\end{aligned}$$

Now let $p \to \infty$. The above first terms goes to 0 by Step 4. The above second term goes to 0 since $\left| E[f_p|\mathscr{C}] - E[f|\mathscr{C}] \right| \leq E\left| f_p - f \right| \to 0$.

The proof of the theorem is complete. □

Remark 11. *There are various extensions of the Chacon-Ornstein theorem. See e.g. [48].*

As an application, we prove the well-known Birkhoff's theorem:

Theorem 12. *Let $\theta : E \to E$ be a measure preserving point transformation of a probability space $(E, \mathscr{E}, m)$. Let $\mathscr{D} = \left\{ A \in \mathscr{E}, A = \theta^{-1}(A) \right\}$ be the σ-algebra of invariant sets under θ. Then for*

any $f \in L^1(m)$,

$$\lim_n \frac{1}{n+1} \sum_{k=0}^{n} f \circ \theta^k = \mathbb{E}[f|\mathscr{D}] .$$

where the convergence holds both m-a.s. and in $L^1(m)$ *sense.*

Proof. (1) Let us define on $L^1(m)$ the operator $Tf = f \circ \theta$. Then clearly T is positive. By the measure preserving property of T we see that T is also a contraction.

(2) We show that T is conservative. In fact, since $1 \in L^1(m)$, we see that the conservative part in the Hopf's decomposition

$$\mathscr{C} = \{\sum_{k=0}^{\infty} 1 \circ \theta^k = \infty\} = E .$$

(3) $\mathscr{D}$ is the invariant σ-algebra $\mathscr{C}$ in the sense of Definition 6. To see that, if $B = \{\sum_{k=0}^{\infty} f \circ \theta^k = \infty\}$ for some $f \in L^1_+(m)$, then clearly $B = \theta^{-1}B$. Conversely, if $B = \theta^{-1}B$, then $B = \{\sum_{k=0}^{\infty} I_B \circ \theta^k = \infty\}$ m-a.s.

(4) Applying the Chacon-Ornstein theorem with $g = 1$, we have m-a.s.

$$\lim_n \frac{1}{n+1} \sum_{k=0}^{n} f \circ \theta^k = \mathbb{E}[f|\mathscr{D}] .$$

(5) For the $L^1(m)$ convergence, we define

$$f_n = \frac{1}{n+1} \sum_{k=0}^{n} f \circ \theta^k .$$

It is easy to verify

$$||f_n||_{L^1} \leq \frac{1}{n+1} \sum_{k=0}^{n} ||f \circ \theta^k||_{L^1} = ||f||_{L^1} .$$

Therefore, the dominated convergence theorem finishes the proof. □

Finally let us prove a useful result for the Harris chain. Harris chain is defined in Definition 2 of Chapter 1.

A remarkable fact involving the Harris chain is that the θ-invariant σ-algebra $\mathscr{C}$ (the σ-algebra of all sets A such that $\theta^{-1}(A) = A$) is P_x trivial for any P_x. We now prove this fact here. First, let

us introduce

Definition 7. *A measurable function $f(x)$ defined on $(E,\mathscr{E})$ is called harmonic if $Pf = f$ for any $x \in E$. It is called superharmonic if $f \geq 0$ and $Pf \leq f$.*

Superharmonic functions have a significant probabilistic meaning that relates to martingale theory. That is, if f is superharmonic, then $f(X_n)$ is a supermartingale for the filtration $\mathscr{F}$ generated by the Markov process X_n. This is due to a simple calculation as follows:

$$E_\nu[f(X_m)|\mathscr{F}_n] = E_{X_n}[f(X_{m-n})] = P_{m-n}f(X_n) \leq f(X_n), \quad P_\nu\text{-a.s.}$$

It is also apparent that $f(X_n)$ is a martingale if f is harmonic. Now we are ready to prove the following lemma:

Lemma 11. *The bounded superharmonic and harmonic functions with respect to the transition kernel P of a Harris chain X_n are m-a.e. constants.*

Proof. Let us prove the superharmonic part first. Suppose that f is bounded harmonic and is not a constant, then there exist two real numbers a and b such that $m(f < a) > 0$ and $m(f > b) > 0$. Since $f(X_n)$ is a bounded positive supermartingale, we have

$$\lim_n f(X_n) = Y$$

for some random variable Y. Since X_n is Harris and $A := m(f < a) > 0$ and $B := m(f > b) > 0$, it is seen that X_n will visit both sets A and B infinitely often. That implies that $Z < a$, P_m-a.s. and $Z > b$, P_m-a.s. , which is an obvious contradiction.

To prove the result for harmonic functions, simply apply the previous arguments to the positive superharmonic functions $g = f + |f|$ and $h = |f| - f$. □

Now we can invoke a result which states that if all bounded harmonic functions with respect to a Markovian kernel P are constants, then the σ-algebra of all θ-invariant sets is P_x trivial. See [48] for the technical details there. This finishes proving the fact that $\mathscr{C}$ is P_x trivial.

Theorem 13. *Let X be a positive Harris chain with a unique invariant probability measure m and let f be a bounded measurable function on $(E^k, \mathscr{E}^k)$. Then for m-a.s. $x \in E$,*

$$\begin{aligned} & \lim_n \frac{1}{n} \sum_{i=0}^{n-1} f(X_i, X_{i+1}, \ldots, X_{l+k-1}) \\ = \; & \mathbb{E}_m f(X_0, X_1, \ldots, X_{k-1}) \\ = \; & \int_E m(dx_0) \int_E P(x_0, dx_1) \cdots \int_E P(x_{k-2}, x_{k-1}) f(x_0, x_2, \ldots, x_{k-1}), \end{aligned}$$

Proof. Let X_n be the canonical process on the probability space $(E^{\mathbb{N}}, \mathscr{E}^{\mathbb{N}}, P_m)$ with invariant probability P_m., i.e. X_n is positive Harris under P_m and the shift operator

$$\theta : E^{\mathbb{N}} \to E^{\mathbb{N}}, (x_1, x_2, \ldots, x_n, \ldots) \to (x_2, x_3, \ldots, x_n, \ldots)$$

is measure preserving. Since f is a bounded measurable function on $(E^k, \mathscr{E}^k)$, we have

$$f(X_l, X_{l+1}, \ldots, X_{l+k-1}) = (f \circ \theta_l)(X_0, X_1, \ldots, X_{k-1})$$

By Birkhoff theorem,

$$\begin{aligned} & \lim_n \frac{1}{n} \sum_{l=0}^{n-1} f(X_l, X_{l+1}, \ldots, X_{l+k-1}) \\ = \; & \lim_n \frac{1}{n} \sum_{l=0}^{n-1} (f \circ \theta_l)(X_0, X_1, \ldots, X_{k-1}) \\ = \; & E_m[f(X_0, X_1, \ldots, X_{k-1}) | \mathscr{C}] \end{aligned}$$

where $\mathscr{C}$ is the θ-invariant σ-algebra (the σ-algebra of all sets A such that $\theta^{-1}(A) = A$). The above

convergence holds P_m-a.s.

$$\begin{aligned} & \lim_n \frac{1}{n} \sum_{l=0}^{n-1} f(X_l, X_{l+1}, \ldots, X_{l+k-1}) \qquad \text{(A.4)} \\ = \; & E_m[f(X_0, X_1, \ldots, X_{k-1})] \\ = \; & \int_E m(dx_0) \int_E P(x_0, dx_1) \cdots \int_E P(x_{k-2}, dx_{k-1}) f(x_0, x_1, \ldots, x_{k-1}) \end{aligned}$$

Let A be the set such that (A.4) holds, then $P_m(A) = 1$. □