2017

Essays on numerical solutions to forward-backward stochastic differential equations and their applications in finance

BOSTON UNIVERSITY

QUESTROM SCHOOL OF BUSINESS

Dissertation

ESSAYS ON NUMERICAL SOLUTIONS TO FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS IN FINANCE

by

LIANGLIANG ZHANG

B.S., Fudan University, 2003
M.S., Fudan University, 2006
M.A., Boston University, 2007

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2017

Approved by

First Reader

Jérôme B. Detemple, Ph.D.
Morton H. and Charlotte Friedman Professor in Management
Professor of Finance

Second Reader

Marcel Rindisbacher, Ph.D.
Associate Professor of Finance

Third Reader

Rodolfo Prieto, Ph.D.
Assistant Professor of Finance

Where there is a will, there is a way.

Acknowledgments

I would like to deeply thank my advisers Professor Jérôme Detemple and Professor Marcel Rindisbacher for their generous support, time and the great advise given to me. Their wisdom inspires me. I have so much to say but feel wordless to describe their immense help. They are simply the best advisers in this world.

I benefit from various discussions with Professor Rodolfo Prieto. Also I would like to express my thanks to Professor Stephan Sturm of Worcester Polytechnic Institute, Department of Mathematical Sciences and Professor Matthew Lorig of University of Washington at Seattle, Department of Applied Mathematics, for their willingness to be my co-authors and their useful advice on how to refine papers.

Liangliang Zhang

Ph.D. Candidate in Mathematical Finance

Questrom School of Business

Department of Mathematical Finance

Boston University

ESSAYS ON NUMERICAL SOLUTIONS TO FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS IN FINANCE

LIANGLIANG ZHANG

Boston University, Questrom School of Business, 2017

Major Professors: Jérôme B. Detemple, Ph.D.
Morton H. and Charlotte Friedman Professor in Management
Professor of Finance

Marcel Rindisbacher, Ph.D.
Associate Professor of Finance

ABSTRACT

In this thesis, we provide convergent numerical solutions to non-linear forward-BSDEs (Backward Stochastic Differential Equations). Applications in mathematical finance, financial economics and financial econometrics are discussed. Numerical examples show the effectiveness of our methods.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 About This Thesis

In this thesis, we provide convergent numerical schemes to solve non-linear uncoupled forward-backward stochastic differential equations with jumps. Various applications in mathematical finance, financial economics and financial econometrics are discussed.

1.2 Literature Review

Bismut (1973) introduced linear BSDEs to study stochastic optimal control problems in the stochastic version of the Pontryagins maximum principle. Non-linear BSDEs were first studied theoretically by Pardoux and Peng (1990), who suggest a general stochastic maximum principle with first and second order adjoint equations. Since then, a substantial literature on theoretical developments and applications of BSDEs to various financial problems has emerged. On the theoretical side, a number of papers have examined the existence and uniqueness of solutions to BSDEs/FBSDEs, including Pardoux and Peng (1990), Antonelli (1993), the first to study Forward-BSDEs, Tang and Li (1994), who study BSDEs with random jumps, Pardoux and Tang (1999), Kobylanski (2000) for quadratic BSDEs and Zhang (2006a,b) for possibly degenerate ones. Peng (2014) proposes a comparison theorem for BSDEs. Pardoux and Peng (1992) show that the solution of the BSDE in the Markovian case corresponds to a probabilistic solution of a non-linear PDE and give a generalization of the Feynman-

Kac formula. Other representation results have been used to establish the relation between solutions of FBSDEs and quasi-linear parabolic PDEs; see Ma and Zhang (2002) and Zhang (2005) for degenerate FBSDEs. El Karoui et al. (2008) provide a comprehensive review of theoretical developments for BSDEs. We refer the interested readers for more references in that work. On the application side, early uses of BSDEs/FBSDEs in financial models appear in Detemple and Zapatero (1991, 1992), Duffie and Epstein (1992), El Karoui et al. (1997a), El Karoui et al. (1997b), Ma and Yong (2000) and Carmona (2009), among others. For example, El Karoui et al. (1997a) discuss reflected BSDEs and their relation to optimal stopping problems. El Karoui et al. (1997b) examine the valuation and hedging of European contingent claims in complete markets and markets with portfolio constraints. El Karoui and Quenez (1997) develop non-linear pricing theory using BSDEs. The portfolio choice problem of a large investor is studied in Cvitanic and Ma (1996), where a fully-coupled FBSDE system is obtained. Recent contributions, such as Bichuch et al. (2015a,b), use BSDEs to compute the XVA of a European contingent claim taking account of funding costs, counterparty risk and collateralization. Nowadays, BSDEs play a prominent role in mathematical finance, financial economics and mathematical economics.

In addition to BSDEs with Lipschitz continuous and linearly growing drivers, quadratic BSDEs appear in risk sensitive control problems, dynamic risk measures, indifference pricing and dynamic portfolio choice problems with incomplete markets. The first discussion of quadratic BSDEs can be found in Kobylanski (2000) in a Brownian filtration setting, under the assumption that the terminal conditions are bounded. Her results are extended by Briand and Hu (2006, 2008), who consider, under Brownian filtration, quadratic BSDEs with unbounded terminal conditions. Extensions can also be found in, for example, Tevzadze (2008), who proves the existence of a unique solution to a general backward stochastic differential equation with quadratic growth driven by martingales. Applications appear in Hu et al. (2005), Morlais (2009) and references therein. Recently, Fujii and Takahashi (2016a) study the existence and uniqueness of solutions to quadratic-exponential BSDEs with jumps, i.e., BSDEs whose drivers exhibit quadratic growth in some variables and exponential growth in others.

Unfortunately, the cases where BSDEs/FBSDEs have closed-form solutions are rare. One has to resort to numerical methods for practical implementations. Due to the nature of the problems considered, the dimensionality of the state vector is often high. Standard numerical approaches, such as the finite difference method for the associated PDEs, usually fail in such situations. To circumvent difficulties associated with large dimensions, Fujii and Takahashi (2012a,b) present analytical approximation methods based on perturbation theory to solve non-linear FBSDEs, but do not provide error estimates. Takahashi and Yamada (2014) and Gobet and Pagliarani (2014) study analytical expansion schemes for BSDEs based on small-diffusion and small-time expansions and derive the associated error bounds.

The references closest to the applications discussed in Chapter 4 are Liu (2007) for portfolio choice with incomplete markets under a quadratic-affine framework, Aït-Sahalia (2002, 2008), Yu (2007), Choi (2013, 2015), Filipović et al. (2013) and Li and Chen (2016) for transition density expansion.

1.3 Main Contributions

While the existence and uniqueness of solutions to the aforementioned BSDEs is by now well-understood, numerical solutions are not trivial to obtain. This thesis attempts to fill this gap: it provides general procedures to compute approximate solutions to general uncoupled FBSDEs (with jumps).

The First Expansion Scheme

The first expansion scheme (in Chapter 2) is based on Picard iteration and nested PDE expansions. It extends the parabolic PDE expansion method first developed formally in a scalar setting in Pagliarani and Pascucci (2012) and later generalized to multiple dimensions with rigorous error estimates in Lorig et al. (2015a). The method documented in Chapter 2 does not discretize the state space, nor does it apply *only* to small time settings, meaning that the expansion scheme is only accurate when time to maturity of the problem is small.

The procedure has five main features. First, the scheme converges and rigorous error estimates are available. Second, it only requires integration over the time domain, which is a one-dimensional computation. Third, it imposes standard restrictions, which are often assumed in the expansion literature, such as boundedness and continuity of derivatives up to some order, on the coefficients of the FBSDEs. Fourth, in addition to the number of time discretizations n and Picard iterations k, it provides two control parameters, m, representing the order of Taylor expansion of the terminal condition and PDE source term and l, representing the order of the PDE expansion, to control the rate of decay of the error bound from the nested PDE expansion. Larger values of m and l increase the speed of this decay as k and n go to ∞. Finally, the evaluation of expansion terms is recursive in nature and can be programmed using any software language that enables symbolic computations (mainly symbolic differentiation).

In order to provide perspective, it is useful to compare our method with Takahashi and Yamada (2014), Gobet and Pagliarani (2014) and Lorig et al. (2015a). The approach in Takahashi and Yamada (2014) is based on small-diffusion expansions. The diffusion coefficient of the forward SDE is multiplied by a small perturbation parameter ϵ which serves as the basis for the expansion of the solution. This approach works best when ϵ is small. Gobet and Pagliarani (2014) solve BSDEs with non-smooth drivers using a perturbation technique. They obtain good performance with mild non-linearity and short time in the case of non-smooth drivers. They contend that the method is more suitable than merely smoothing the driver and applying the expansion methods available for smooth coefficients. Compared to Gobet and Pagliarani (2014), our method is convergent and does not rely on a short maturity or a small perturbation coefficient. It extends Lorig et al. (2015a) from parabolic PDEs to FBSDEs with the help of Picard iteration and the non-linear Feynman-Kac formula for FBSDEs. Moreover, the approach in Lorig et al. (2015a) requires a d-dimensional integration in order to solve for the expansion. Our method only requires integration over the time domain, which considerably simplifies the computation. To summarize, we provide a numerical method, to solve FBSDEs, to the literature, which is easy to implement.

The Second Expansion Scheme

The second expansion scheme (in Chapter 3) builds on the first one. The major steps are:

- Use a sequence of FBSDEs with coefficients that are smooth and have bounded derivatives of all orders to approximate the original FBSDE.
- For every FBSDE in the sequence, apply Picard iteration to linearize it.
- Associate the linearized FBSDE to a PIDE.
- Use time discretization and Taylor expansion (at a fixed point x_0 which will be described later) to solve the PIDE analytically.

The method is based on the results of Liu and Li (2000), Jum (2015) and Lorig et al. (2013, 2015a,b,c). Liu and Li (2000) and Jum (2015) document the weak convergence of stochastic Taylor expansions to approximate the expectation of a known function, with at most polynomial growth, of a jump-diffusion process. We generalize their work in the following way. Liu and Li (2000) and Jum (2015) suggest a Monte Carlo evaluation of the conditional expectation. We use the law of iterated expectations and polynomial expansion to approximate conditional expectations and the combined method results in an analytical approximation scheme. In Lorig et al. (2015b), the authors derive small-time error bounds for their higher order PIDE approximation. We extend their results to large time and obtain convergence. The method introduced also generalizes the literature on asymptotic expansions, for example, Takahashi and Yamada (2014), Fujii and Takahashi (2016b), Fujii (2016) and Fujii and Takahashi (2016a), in that the convergence does not rely on a small perturbation parameter. Our method may be more suitable than simulation in some cases, where nested evaluations or the higher order derivatives of the solutions are needed.

Numerical experiments indicate the efficiency of the second expansion method. A numerical experiment with 6 rounds of Picard iterations, 2000 time-discretizations with Taylor expansion order 2 takes only 40 seconds on an i7 PC.

Financial Applications

As for applications, we consider several important problems in mathematical finance, financial economics and financial econometrics, which include portfolio choice with incomplete markets, optimal investment for an insurer and transition density approximation for stochastic differential equations with jumps. We provide convergent numerical algorithms to solve the aforementioned problems numerically and compare performance relative to some selected methods.

We provide numerical solutions to dynamic portfolio choice problems. The algorithm can serve as a general procedure to compute the optimal portfolios and the optimal wealth functions with complete or incomplete markets.

On the econometrics side, our transition density approximation is feasible, fast and accurate. Transition densities are needed when we want to estimate the model parameters using maximum likelihood method and they are usually not known in closed-form. Extending the methods studied in Aït-Sahalia (2002, 2008), Yu (2007) and Choi (2013, 2015), our algorithm does not require us to solve partial differential equations recursively to obtain the coefficients of the expansion. Moreover, as a theoretical extension to Aït-Sahalia (2002, 2008), Yu (2007), Choi (2013, 2015) and Li and Chen (2016), our method is convergent and the convergence does not rely on a small parameter. For some selected SDEs with jumps, We provide error plots for the transition density approximation. In the end, an MLE estimation exercise is performed on a CIR model with or without positive jumps.

1.4 Organization

The organization of the thesis is as follows. Chapter 2 describes the first expansion scheme. Chapter 3 introduces the second expansion scheme. Chapter 4 contains all the financial applications. Chapter 5 concludes. All the proofs can be found in the Appendix.

Chapter 2

The First Expansion Scheme

2.1 Outline of This Chapter

In this chapter, we develop a numerical expansion scheme to solve a general uncoupled forward-backward stochastic differential equation. We first introduce the FBSDE. Then, we describe the numerical expansion scheme. In the end, we state the assumptions required in this chapter and derive the error bounds and prove convergence of the proposed scheme.

2.2 The FBSDE Considered

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a complete, filtered probability space where a d-dimensional Brownian motion $W = (W_t)_{t\geq 0}$ is defined such that $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is the the natural filtration generated by W. Consider the following uncoupled forward-backward stochastic differential equation (FBSDE)

$$\begin{aligned} \mathrm{d}X_t &= \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t, \qquad X_0 = x \in \mathbb{R}^d, \\ \mathrm{d}Y_t &= -f(t, X_t, Y_t, Z_t)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \quad Y_T = \psi(X_T) \in \mathbb{R}, \end{aligned} \tag{2.2.1}$$

where the process $X = (X_t)_{t\in[0,T]}$ lives in $\mathbb{R}^d$, the process $Y = (Y_t)_{t\in[0,T]}$ lives in $\mathbb{R}$, the process $Z = (Z_t)_{t\in[0,T]}$ lives in $\mathbb{R}^d$ and the functions (μ, σ, f, ψ) map

$$\begin{aligned} &\mu : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d, && \sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}, \\ &f : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}, && \psi : \mathbb{R}^d \to \mathbb{R}. \end{aligned}$$

Precise conditions satisfied by the functions (μ, σ, f, ψ) will be given in Assumptions 2.4.5, 2.4.6 and 2.4.7. For now, it is assumed that the functions are sufficiently well behaved, e.g., sufficiently smooth, to validate the manipulations performed below.

We assume that the FBSDE (2.2.1) has a unique solution (Y, Z). The goal is to find an approximation to the solution of the FBSDE that, in some limit, converges to the solution (Y, Z) but is much easier to calculate numerically. The precise sense in which the algorithm converges will be stated in Theorem 2.4.8. For now, we focus on explaining how the approximate solution is constructed.

2.3 The Numerical Expansion Scheme

Step 1: Picard iteration

The first step is to write the solution (Y, Z) as the limit of a Picard iteration scheme. Specifically, define the processes $Y^{(0)} = \big(Y_t^{(0)}\big)_{t\in[0,T]}$ and $Z^{(0)} = \big(Z_t^{(0)}\big)_{t\in[0,T]}$ by

$$Y_t^{(0)} := \mathbb{E}_t \psi(X_T) + \mathbb{E}_t \int_t^T f(s, X_s, 0, 0)\,\mathrm{d}s, \qquad Z_t^{(0)} := \mathcal{D}_t Y_t^{(0)},$$

where $\mathbb{E}_t\cdot$ denotes the conditional expectation $\mathbb{E}[\,\cdot\,|\mathcal{F}_t]$ and $\mathcal{D}_t$ is the Malliavin gradient operator with respect to the d-dimensional Brownian motion W. Next, for any $k \geq 1$, define $Y^{(k)} = \big(Y_t^{(k)}\big)_{t\in[0,T]}$ and $Z^{(k)} = \big(Z_t^{(k)}\big)_{t\in[0,T]}$ as the solution $\big(Y^{(k)}, Z^{(k)}\big)$ of the following linear FBSDE

$$\mathrm{d}Y_s^{(k)} = -f\big(s, X_s, Y_s^{(k-1)}, Z_s^{(k-1)}\big)\,\mathrm{d}s + Z_s^{(k)}\,\mathrm{d}W_s, \qquad Y_T^{(k)} = \psi(X_T). \tag{2.3.1}$$

It is known (see for example (El Karoui et al., 1997b, Corollary 2.1)) that under appropriate conditions on (μ, σ, f, ψ) the sequence $\big(Y^{(k)}, Z^{(k)}\big)$ converges to (Y, Z).

Step 2: Reduction to a sequence of linear PDEs

The second step is to relate $\big(Y^{(k)}, Z^{(k)}\big)$ to the solution of a linear parabolic partial differential equation (PDE). Specifically, let the sequence of functions $\big(u^{(k)}\big)_{k\geq 0}$ be the unique classical solution (assumed to exist for now, detailed analysis will be given later) of the following sequence of nested PDEs

$$(\partial_t + \mathcal{A})u^{(k)} + f^{(k)} = 0, \qquad u^{(k)}(T,\cdot) = \psi(\cdot), \qquad k \geq 0, \tag{2.3.2}$$

where the operator $\mathcal{A}$ (the infinitesimal generator of X) and the function $f^{(k)}$ are given by

$$\mathcal{A} = \sum_{i=1}^{d} \mu_i(t,x)\partial_{x_i} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^{\mathrm{T}})_{i,j}(t,x)\partial_{x_i}\partial_{x_j},$$
$$f^{(k)}(t,x) = f\big(t,x,u^{(k-1)}(t,x), \nabla_x u^{(k-1)}(t,x)\cdot\sigma(t,x)\big), \quad k \geq 0,$$

with $u^{(-1)} := 0$. If we define

$$Y_t^{(k)} := u^{(k)}(t, X_t), \qquad Z_t^{(k)} := \nabla_x u^{(k)}(t,X_t)\cdot\sigma(t,X_t),$$

then, the pair $\big(Y^{(k)}, Z^{(k)}\big)$ solves FBSDE (2.3.1). Note, however, that for general (μ,σ,f,ψ), there is no explicit solution $\big(u^{(k)}\big)_{k\geq 0}$ to (2.3.2).

Step 3: Approximate solution of the sequence of PDEs

From (2.3.2), we see that each $u^{(k)}$ in the sequence $\big(u^{(k)}\big)_{k\geq 0}$ satisfies a linear parabolic PDE of the form

$$(\partial_t + \mathcal{A})u + f = 0, \qquad u(T,\cdot) = \psi, \qquad \mathcal{A} = \sum_{1\leq|\alpha|\leq 2} a_\alpha(t,x)\partial_x^\alpha, \tag{2.3.3}$$

where the operator $\mathcal{A}$ has been rewritten using standard multi-index notation

$$
\begin{aligned}
\alpha &= (\alpha_1, \alpha_2, \dots, \alpha_d), & |\alpha| &= \sum_{i=1}^{d} \alpha_i, & \partial_x^\alpha &= \prod_{i=1}^{d} \partial_{x_i}^{\alpha_i}, \\
x^\alpha &= \prod_{i=1}^{d} {x_i}^{\alpha_i}, & \alpha! &= \prod_{i=1}^{d} \alpha_i!. &&
\end{aligned} \tag{2.3.4}
$$

For general coefficients (a_α), there is no closed-form solution u to the PDE (2.3.3). However, an approximate solution can be obtained using the methods developed in (Lorig et al., 2014). We briefly review the key elements of their approach here.

First, we fix a point $\bar{x} \in \mathbb{R}^d$. Assuming that the coefficients of $\mathcal{A}$ are smooth enough, we can expand each of these as a Taylor series about the point $\bar{x}$. Formally, the operator $\mathcal{A}$ can then be written as

$$
\begin{aligned}
\mathcal{A} &= \sum_{i=0}^{\infty} \mathcal{A}_i^{\bar{x}}, & \mathcal{A}_i^{\bar{x}} &= \sum_{1 \le |\alpha| \le 2} a_{\alpha,i}^{\bar{x}}(t,x) \partial_x^\alpha, \\
a_{\alpha,i}^{\bar{x}}(t,x) &= \sum_{|\beta|=i} \frac{1}{\beta!} \partial_x^\beta a_\alpha(t,\bar{x})(x-\bar{x})^\beta. &&
\end{aligned} \tag{2.3.5}
$$

Note that we have explicitly indicated the dependence of $\mathcal{A}_n^{\bar{x}}$ on the expansion point $\bar{x}$. Next, we expand the function u as an infinite series

$$
u = \sum_{l=0}^{\infty} u_l^{\bar{x}},
$$

where, once again, we have explicitly indicated with a superscript the dependence of each $u_l^{\bar{x}}$ on $\bar{x}$. Inserting the expansions for $\mathcal{A}$ and u into the PDE (2.3.3), and collecting terms whose subscripts sum to like order, we obtain

$$
\begin{cases}
\big(\partial_t + \mathcal{A}_0^{\bar{x}}\big) u_0^{\bar{x}} + f = 0, \quad u_0(T,\cdot) = \psi, \\
\big(\partial_t + \mathcal{A}_0^{\bar{x}}\big) u_l^{\bar{x}} + \sum_{i=1}^{l} \mathcal{A}_i^{\bar{x}} u_{l-i}^{\bar{x}} = 0, \quad u_l^{\bar{x}}(T,\cdot) = 0, \quad l \ge 1.
\end{cases} \tag{2.3.6}
$$

Note that $\mathcal{A}_0^{\bar{x}}$ is a second-order elliptic operator. Thus, the operator $\mathcal{A}_0^{\bar{x}}$ generates a *semigroup* $\mathcal{P}_0^{\bar{x}}$. As the coefficients of $\mathcal{A}_0^{\bar{x}}$ are constant in x, the action of the semigroup can be written in closed form

$$\mathcal{P}_0^{\bar{x}}(t,T)g(x) = \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t,x;T,y)g(y), \qquad x \in \mathbb{R}^d, \qquad t \leq T, \quad (2.3.7)$$

where $\Gamma_0^{\bar{x}}$ is the transition density of a Gaussian process whose mean vector and covariance matrix are given by equation (A.1.1) in Appendix A.1. By Duhamel's principle (see, for example, Thomée and Zhang (1989)), the sequence of functions $\left(u_l^{\bar{x}}\right)_{l\geq 0}$ can be written in semi-closed form (as an integral)

$$\begin{cases} u_0^{\bar{x}}(t,x) = \mathcal{P}_0^{\bar{x}}(t,T)\psi(x) + \displaystyle\int_t^T \mathrm{d}t_1 \mathcal{P}_0^{\bar{x}}(t,t_1)f(t_1,x), \\ u_l^{\bar{x}}(t,x) = \displaystyle\int_t^T \mathrm{d}t_1 \mathcal{P}_0^{\bar{x}}(t,t_1) \sum_{i=1}^{l} \mathcal{A}_i^{\bar{x}} u_{l-i}^{\bar{x}}(t_1,x), \qquad l \geq 1. \end{cases}$$

Step 4: Taylor expansion of the PDE source terms and terminal condition

We will need to find an approximate solution to a PDE of the form (2.3.3) at every step in the Picard iteration. Because of this, it will be useful to have a closed-form approximation for each $u_l^{\bar{x}}$ in the sequence $(u_l^{\bar{x}})_{l\geq 0}$. Note that, if p is a polynomial of degree m, then $q^{\bar{x}}(t,x) := \mathcal{P}_0^{\bar{x}}(t,T)p(x)$ will be a polynomial in x of degree m as well. An explicit expression for $\mathcal{P}_0^{\bar{x}}(t,T)p(x)$ is given in equation (A.1.2) in Appendix A.1.

If the source term f and the terminal condition ψ of the Cauchy problem (2.3.6) are polynomials, then $u_l^{\bar{x}}(t,x)$ will be a polynomial in x for every l. With this in mind, we shall Taylor expand the functions (f,ψ). Let us define the *Taylor expansion operator* $\mathbf{T}_m^{\bar{x}}$, which maps any $C^m(\mathbb{R}^d)$ function to its mth-order Taylor expansion about the point $\bar{x}$

$$\mathbf{T}_m^{\bar{x}} g(x) := \sum_{i=0}^{m} \sum_{|\alpha|=i} \frac{1}{\alpha!} \partial_x^\alpha g(\bar{x})(x-\bar{x})^\alpha.$$

For a fixed m, define $\left(u_{l,m}^{\bar{x}}\right)_{l\geq 0}$ as the unique classical solutions of the following nested sequence of PDEs

$$\begin{cases} \left(\partial_t + \mathcal{A}_0^{\bar{x}}\right) u_{0,m}^{\bar{x}} + \mathbf{T}_m^{\bar{x}} f = 0, \quad u_{0,m}^{\bar{x}}(T,\cdot) = \mathbf{T}_m^{\bar{x}} \psi, & \\ \left(\partial_t + \mathcal{A}_0^{\bar{x}}\right) u_{l,m}^{\bar{x}} + \sum_{i=1}^{l} \mathcal{A}_i^{\bar{x}} u_{l-i,m}^{\bar{x}} = 0, \quad u_{l,m}^{\bar{x}}(T,\cdot) = 0, & l \geq 1. \end{cases} \tag{2.3.8}$$

Comparing (2.3.6) with (2.3.8), we see that the only change was to apply the Taylor expansion operator $\mathbf{T}_m^{\bar{x}}$ to two terms: f and ψ. The sequence of functions $\left(u_{l,m}^{\bar{x}}\right)_{l\geq 0}$, given by

$$\begin{cases} u_{0,m}^{\bar{x}}(t,x) = \mathcal{P}_0^{\bar{x}}(t,T)\mathbf{T}_m^{\bar{x}}\psi(x) + \int_t^T \mathrm{d}t' \mathcal{P}_0^{\bar{x}}(t,t')\mathbf{T}_m^{\bar{x}} f(t',x), & \\ u_{l,m}^{\bar{x}}(t,x) = \int_t^T \mathrm{d}t' \mathcal{P}_0^{\bar{x}}(t,t') \sum_{i=1}^{l} \mathcal{A}_i^{\bar{x}} u_{l-i,m}^{\bar{x}}(t',x), & l \geq 1, \end{cases}$$

can be computed explicitly (i.e., all integrals can be evaluated analytically).

We have yet to comment on the choice of $\bar{x}$. In general, the choice of $\bar{x}$ for which the partial sum $\sum_{i=1}^{l} u_{i,m}^{\bar{x}}$ most accurately approximates u at the point x is $\bar{x} = x$. Thus, for this special case, we define

$$u_{l,m}(t,x) := u_{l,m}^{\bar{x}}(t,x)\Big|_{\bar{x}=x}.$$

To give the reader a clear idea of the structure of the approximating solution, we point out that while $u_{l,m}^{\bar{x}}(t,x)$ is a polynomial in x it is *not*, in general, a polynomial in $\bar{x}$. As such, $u_{l,m}(t,x)$ will not generally be a polynomial in x.

Step 5: Time discretizing PDEs

The partial sum $\sum_{i=1}^{l} u_{i,m}$ most accurately approximates u when $T-t$ is small and loses accuracy as $T-t$ grows large. To overcome this limitation, we introduce a time discretization scheme. Let us divide the interval $[t,T]$ into n equally spaced intervals $[t_{i-1}, t_i]$ with $i = 1, 2, \ldots, n$, where

$$t_i = t + i\delta_t, \qquad \delta_t = (T-t)/n, \qquad i = 0, 1, 2, \ldots, n.$$

We define $u^{\bar{x}}_{l,m,n}$ as the solution of the following sequence of PDEs

$$\begin{cases} \big(\partial_t + \mathcal{A}^{\bar{x}}_0\big) u^{\bar{x}}_{0,m,n} + \mathbf{T}^{\bar{x}}_m f = 0, \quad u^{\bar{x}}_{0,m,n}(T,\cdot) = \mathbf{T}^{\bar{x}}_m \psi, \quad t \in [t_{n-1}, T), \\ \big(\partial_t + \mathcal{A}^{\bar{x}}_0\big) u^{\bar{x}}_{l,m,n} + \displaystyle\sum_{i=1}^{l} \mathcal{A}^{\bar{x}}_i u^{\bar{x}}_{l-i,m,n} = 0, \quad u^{\bar{x}}_{l,m,n}(T,\cdot) = 0, \qquad l \geq 1, \end{cases} \tag{2.3.9}$$

and, for $j = 1, 2, 3, \ldots, n-1$,

$$\begin{cases} \big(\partial_t + \mathcal{A}^{\bar{x}}_0\big) u^{\bar{x}}_{0,m,n} + \mathbf{T}^{\bar{x}}_m f = 0, \\ \qquad u^{\bar{x}}_{0,m,n}(t_{n-j},\cdot) = \mathbf{T}^{\bar{x}}_m u^{\mathtt{tc}}_{0,m,n}(t_{n-j},\cdot), \qquad t \in [t_{n-j-1}, t_{n-j}), \\ \big(\partial_t + \mathcal{A}^{\bar{x}}_0\big) u^{\bar{x}}_{l,m,n} + \displaystyle\sum_{i=1}^{l} \mathcal{A}^{\bar{x}}_i u^{\bar{x}}_{l-i,m,n} = 0, \\ \qquad u^{\bar{x}}_{l,m,n}(t_{n-j},\cdot) = \mathbf{T}^{\bar{x}}_{m-2l} u^{\mathtt{tc}}_{l,m,n}(t_{n-j},\cdot), \quad l \geq 1, \end{cases} \tag{2.3.10}$$

where we have defined

$$\partial^\beta_x u^{\mathtt{tc}}_{l,m,n}(t,x) := \partial^\beta_x \big(u^{\bar{x}}_{l,m,n}(t,x) \big|_{\bar{x}=x} \big), \tag{2.3.11}$$

and the superscript `tc` stands for *terminal condition*. It will be helpful to explain briefly how to construct $u_{l,m,n}(t,x)$ for any $t \in [0,T]$. First, one solves (2.3.9) using

$$\begin{cases} u^{\bar{x}}_{0,m,n}(t,x) \\ \qquad = \mathcal{P}^{\bar{x}}_0(t,T)\mathbf{T}^{\bar{x}}_m \psi(x) + \displaystyle\int_t^T \mathrm{d}t' \mathcal{P}^{\bar{x}}_0(t,t')\mathbf{T}^{\bar{x}}_m f(t',x), \quad t \in [t_{n-1},T), \\ u^{\bar{x}}_{l,m,n}(t,x) \\ \qquad = \displaystyle\int_t^T \mathrm{d}t' \mathcal{P}^{\bar{x}}_0(t,t') \sum_{i=1}^{k} \mathcal{A}^{\bar{x}}_i u^{\bar{x}}_{l-i,m,n}(t',x), \quad l \geq 1. \end{cases} \tag{2.3.12}$$

Combining (2.3.11) with (2.3.12) yields an explicit expression for $u^{\mathtt{tc}}_{l,m,n}(t,x)$, which is valid for any $t \in [t_{n-1},T]$. Note that $u^{\mathtt{tc}}_{l,m,n}$ will generally *not* be polynomial in x. Next, for every $j = 0,1,2,3,\dots,n-1$, one solves (2.3.10) using

$$\begin{cases} u^{\bar{x}}_{0,m,n}(t,x) \\ \qquad = \mathcal{P}^{\bar{x}}_0(t,t_{n-j})\mathbf{T}^{\bar{x}}_m u^{\mathtt{tc}}_{0,m,n}(t_{n-j},x) \\ \qquad\quad + \displaystyle\int_t^{t_{n-j}} \mathrm{d}t' \mathcal{P}^{\bar{x}}_0(t,t')\mathbf{T}^{\bar{x}}_m f(t',x), \qquad t \in [t_{n-j-1},t_{n-j}), \\ u^{\bar{x}}_{l,m,n}(t,x) \\ \qquad = \mathcal{P}^{\bar{x}}_0(t,t_{n-j})\mathbf{T}^{\bar{x}}_{m-2l} u^{\mathtt{tc}}_{l,m,n}(t_{n-j},x) \\ \qquad\quad + \displaystyle\int_t^{t_{n-j}} \mathrm{d}t' \mathcal{P}^{\bar{x}}_0(t,t') \sum_{i=1}^{l} \mathcal{A}^{\bar{x}}_i u^{\bar{x}}_{l-i,m,n}(t',x), \quad l \geq 1. \end{cases} \tag{2.3.13}$$

Setting $\bar{x} = x$ in (2.3.13) yields at the j^{th} step, an explicit expression for $u_{l,m,n}(t,x)$, which is valid for any $t \in [t_{n-j-1},t_{n-j})$. After $n-1$ total iterations, one obtains the value for $u_{l,m,n}(t,x)$ for any $t \in [t_0,t_1)$. Throughout this chapter, we will use the shorthand

$$\partial_x^\beta u_{l,m,n}(t,x) := \left(\partial_x^\beta u_{l,m,n}^{\bar{x}}(t,x)\right)\big|_{\bar{x}=x}. \tag{2.3.14}$$

Observe that $u_{l,m,n} = u_{l,m,n}^{\mathtt{tc}}$ by comparing equations (2.3.11) and (2.3.14), however $\partial_x^\beta u_{l,m,n} \neq \partial_x^\beta u_{l,m,n}^{\mathtt{tc}}$. Also observe that the function $u_{l,m,n}$ is not continuous in t at times $\{t_j\}_{j=1}^{n-1}$ for any (l,m,n) because

$$\lim_{t\to t_j-} u_{l,m,n}(t,x) = \mathbf{T}_{m-2l}^{\bar{x}} u_{l,m,n}^{\mathtt{tc}}(t_j,x)\Big|_{\bar{x}=x} \neq u_{l,m,n}(t_j,x),$$

in general.

2.4 The Error Bounds and the Convergence Result

We are now in a position to define our approximate solution of FBSDE (2.2.1). To simplify notation, we will always use over-bar to denote a partial sum over the first subscript of any object. For example

$$\bar{u}_l := \sum_{i=0}^{l} u_i, \qquad \bar{u}_{l,m} := \sum_{i=0}^{l} u_{i,m}, \qquad \bar{u}_{l,m,n} := \sum_{i=0}^{l} u_{i,m,n}, \tag{2.4.1}$$

and likewise for other objects. We begin with the following definition.

Definition 2.4.1. *Let $u_{l,m,n}^{(k)}$ be given by $u_{l,m,n}$ in* (2.3.12), (2.3.13) *and* (2.3.14) *with f replaced by*

$$f_{l,m,n}^{(k)}(t,x) = f\big(t,x,\bar{u}_{l,m,n}^{(k-1)}(t,x),\nabla_x \bar{u}_{l,m,n}^{(k-1)}(t,x)\cdot\sigma(t,x)\big), \quad \bar{u}_{l,m,n}^{(k)}(t,x) = \sum_{i=0}^{l} u_{i,m,n}^{(k)},$$

where, by convention, we set $\bar{u}_{l,m,n}^{(-1)} := 0$. Define the $(k,l,m,n)^{\text{th}}$ order approximation of *(Y,Z) by*

$$Y_t^{(k,l,m,n)} := \bar{u}_{l,m,n}^{(k)}(t,X_t), \qquad Z_t^{(k,l,m,n)} := \nabla_x \bar{u}_{l,m,n}^{(k)}(t,X_t)\cdot\sigma(t,X_t). \tag{2.4.2}$$

We refer to k as the degree of Picard iteration, *to l as the* degree of PDE expansion, *to* m *as the* degree of Taylor expansion *and to* n *as* the degree of time discretization.

In our main result (Theorem 2.4.8 below), we will state conditions under which and the sense in which the $(k,l,m,n)^{\text{th}}$ order approximation $\left(Y^{(k,l,m,n)}, Z^{(k,l,m,n)}\right)$ converges to (Y,Z). To do this, we shall need the following definitions and assumptions.

Definition 2.4.2. *Let* $C_b^m(\mathbb{R}^N;\mathbb{R})$ *be the space of bounded functions* $f : \mathbb{R}^N \to \mathbb{R}$ *whose derivatives up to order* m *are bounded and continuous. We will use the shorthand notation* C_b^m *wherever appropriate.*

Definition 2.4.3. *For any function* $g \in C_b^\chi(\mathbb{R}^d;\mathbb{R})$*, define its order* $\rho_g^{(m,\chi)}$ *by*

$$\rho_g^{(m,\chi)} = \inf_{\rho}\left\{\rho \geq 0 : \|\partial_x^\beta g\|_\infty \leq \left(\max_{0\leq|\gamma|\leq m}\|\partial_x^\gamma g\|_\infty\right)|\beta|^{\rho|\beta|}, \quad 0 \leq |\beta| < \chi\right\}, \quad m \geq 1,$$

where we use the usual convention $\inf\emptyset = +\infty$.

The definition of $\rho_g^{(m,\chi)}$ depends on the function g and on the choice of m and χ. One can see that if a function $g \in C_b^\chi(\mathbb{R}^d;\mathbb{R})$ has order $\rho_g^{(m_1,\chi)}$ and $f \in C_b^\chi(\mathbb{R}^d;\mathbb{R})$ has order $\rho_f^{(m_2,\chi)}$, then the functions fg and $f+g$ have at most order $\rho^{\max(m_1,m_2),\chi} = \max\left(\rho_g^{(m_1,\chi)}, \rho_f^{(m_2,\chi)}\right)$.

Remark 2.4.4. *When* $\chi = +\infty$*, the above definition of order is equivalent to the following definition*

$$\begin{aligned}\rho_g^{(m,+\infty)} &= \limsup_{|\beta|\to+\infty}\frac{\log\|\partial_x^\beta g\|_\infty}{\log\left(\max_{0\leq|\gamma|\leq m}\|\partial_x^\gamma g\|_\infty\right) + \rho|\beta|\log|\beta|} \\ &= \limsup_{|\beta|\to+\infty}\frac{\log\|\partial_x^\beta g\|_\infty}{\rho|\beta|\log|\beta|}, \quad |\beta| \geq 2.\end{aligned}$$

Below, in Assumptions 2.4.5, 2.4.6 and 2.4.7, χ is either the integer $n(m+1)+1$ or $+\infty$, which we will specify whenever needed.

Assumption 2.4.5 (on the coefficients a_α)**.** *The coefficients* a_α *are* $\mathcal{B}\big([0,T]\big) \otimes \mathcal{B}\big(\mathbb{R}^d\big)$*-Borel measurable and satisfy*

$$a_\alpha(t,\cdot) \in C_b^\chi(\mathbb{R}^d;\mathbb{R}) \quad \forall\, t \in [0,T], \quad a_\alpha(\cdot,x) \in C_b^1([0,T];\mathbb{R}) \quad \forall\, x \in \mathbb{R}^d.$$

There exists a constant $M > 0$ such that

$$M^{-1}|\xi|^2 \leq \sum_{|\alpha|=2} a_\alpha(t,x)\xi^\alpha \leq M|\xi|^2 \qquad \forall\, t \in [0,T],\ x, \xi \in \mathbb{R}^d.$$

Moreover, for some sufficiently large integer m the functions $a_\alpha(t,\cdot)$ have orders $0 \leq \rho_\alpha^{(m,\chi)} < +\infty$ for all $t \in [0,T]$.

Assumption 2.4.6 (on the terminal datum ψ)**.** *The terminal datum ψ is $\mathcal{B}\big(\mathbb{R}^d\big)$-Borel measurable and satisfies $\psi \in C_b^\chi(\mathbb{R}^d;\mathbb{R})$. There exists a sufficiently large m such that function $\psi(\cdot)$ has order $0 \leq \rho_\psi^{(m,\chi)} < +\infty$.*

Assumption 2.4.7 (on the driver f)**.** *The driver f is $\mathcal{B}\big([0,T]\big) \otimes \mathcal{B}\big(\mathbb{R}^{2d+1}\big)$-Borel measurable and satisfies*

$$\begin{aligned} &f(t,\cdot,\cdot,\cdot) \in C_b^\chi(\mathbb{R}^{2d+1};\mathbb{R}) \quad \forall\, t \in [0,T], \\ &f(\cdot,x,y,z) \in C_b^1([0,T];\mathbb{R}) \quad \forall\, (x,y,z) \in \mathbb{R}^{2d+1}. \end{aligned}$$

For some sufficiently large integer m the function $f(t,\cdot)$ has order $0 \leq \rho_f^{(m,\chi)} < +\infty$ for all $t \in [0,T]$.

Theorem 2.4.8. *Let (Y,Z) be the solution of the FBSDE (2.2.1). Let $Y_t^{(k,l,m,n)}$ and $Z_t^{(k,l,m,n)}$ be as given in equation (2.4.2). Then, under Assumptions 2.4.5, 2.4.6 and 2.4.7 for $\chi = n(m+1)+1$, we have for a fixed level l of the PDE expansion and a degree $m \geq 4l+3$ of the Taylor Expansion that*

$$\begin{aligned} &\left\| Y_\cdot - Y_\cdot^{(k,l,m,n)} \right\|_{L^2}^2 + \left\| Z_\cdot - Z_\cdot^{(k,l,m,n)} \right\|_{L^2}^2 \\ &\quad \leq K \frac{(2\delta)^k}{1-2\delta} + C\left(\frac{T-t}{n}\right)^{l+2} + Cn^{2l+2}\left(\frac{T-t}{n}\right)^{m-2l}, \end{aligned} \tag{2.4.3}$$

where the constant $\delta \in \big(0, \frac{1}{2}\big)$ is independent of the choice of k, l, m and n, the constant K is independent of δ, k, l, m and n, and the constant C is depending only on k, m, l, T and η. The L^2-norm $\|\cdot\|_{L^2}^2$ is given by $\|\xi_\cdot\|_{L^2}^2 := \mathbb{E}\int_0^T \mathrm{d}t\, |\xi_t|^2$.

Corollary 2.4.9. *Under Assumptions 2.4.5, 2.4.6 and 2.4.7 with $\chi = +\infty$ and for fixed l and $m \geq 4l+3$, we have*

$$\lim_{k\to+\infty} \lim_{n\to+\infty} \left(\left\| Y_\cdot - Y_\cdot^{(k,l,m,n)} \right\|_{L^2}^2 + \left\| Z_\cdot - Z_\cdot^{(k,l,m,n)} \right\|_{L^2}^2 \right) = 0.$$

Remark 2.4.10. *The limit in Corollary 2.4.9 is sequential. First n, then k are sent to $+\infty$. The parameters (l, m), which are kept constant, affect the error bound through the constant C and the power of $\frac{T-t}{n}$, as can be seen from* (2.4.3) *and the proof in the Appendix.*

Remark 2.4.11. *Although it is assumed that the dimension of the forward-SDE is the same as the dimension of the Brownian motion, the results can be extended to the case where these two quantities are different. Furthermore, the results hold also in the case where $Y_t \in \mathbb{R}^q$ with $q \geq 2$.*

The proofs of Theorem 2.4.8 and Corollary 2.4.9 can be found in Appendix A.2. Specifically, the proof of Theorem 2.4.8 relies on the following Proposition

Proposition 2.4.12. *Let Assumptions 2.4.5, 2.4.6 and 2.4.7 hold with $\chi = n(m+1)$. Let u be the unique classical solution of* (2.3.3) *and let $\bar{u}_{l,m,n}$ be as defined in* (2.3.14) *and* (2.4.1)*. Then, for any multi-index β with $0 \leq |\beta| \leq 1$ and n large enough, we have*

$$\begin{aligned}\sup_{x\in\mathbb{R}^d}\left|\partial_x^\beta u(t,x) - \partial_x^\beta \bar{u}_{l,m,n}(t,x)\right| \\ \leq C\left(\frac{T-t}{n}\right)^{(l+3-|\beta|)/2} + Cn^{l+1}\left(\frac{T-t}{n}\right)^{(m+1-|\beta|-2l)/2},\end{aligned} \tag{2.4.4}$$

where C is a constant that depends only on m, l and T.

Proof. The proof of Proposition 2.4.12 is given in Appendix A.3. □

Chapter 3

The Second Expansion Scheme

3.1 Outline of This Chapter

In this chapter, we propose the second numerical expansion method to solve a general uncoupled quadratic-exponential forward-backward stochastic differential equation with jumps (QEFBSDEJ). The method extends the one documented in Chapter 2. Most importantly, we do not need the coefficients of the FBSDE to be smooth or bounded. We first state the FBSDE, then introduce the general algorithm. Error bounds and convergence result follow.

3.2 The FBSDE Considered

Consider a filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P}\big)$, with $T \in \mathbb{R}^+$. The space is supporting a d-dimensional Brownian motion $W_t = (W_t^1, \cdots, W_t^d)$ and a Poisson random measure N on $\mathcal{B}([0,T]) \otimes \mathcal{E}$, where $\mathcal{B}([0,T])$ is the Borel σ-algebra on $[0,T]$ and $(E, \mathcal{E})$ is a measurable space. Define $E := \mathbb{R}^q$ and $\mathcal{E}$ as the Borel σ-algebra on E. $\mathbb{P}$ is the probability measure on $\mathcal{F}$. The filtration $(\mathcal{F}_t)_{0\le t\le T}$ is completed with all $\mathbb{P}$-null sets, right-continuous and $\mathcal{F}_t = \mathcal{F}_t^{W,N}$ is generated by $(W_t, N(\cdot, [0,t], \cdot))$ for $t \in [0,T]$. Assume that $\mathcal{F} = \mathcal{F}_T^{W,N}$ and W and N are mutually independent under $\mathbb{P}$. Suppose that the compensating measure of N is $\nu(\,\mathrm{d}t, \,\mathrm{d}e) := \nu(\,\mathrm{d}e)\,\mathrm{d}t$, where ν is a σ-finite measure on $(E, \mathcal{E})$ satisfying $\int_E (1 \wedge |e|^2)\nu(\,\mathrm{d}e) < \infty$. The corresponding compensated Poisson random measure is defined by $\tilde{N}(\omega, \,\mathrm{d}t, \,\mathrm{d}e) := N(\omega, \,\mathrm{d}t, \,\mathrm{d}e) - \nu(\,\mathrm{d}e)\,\mathrm{d}t$.

The FBSDE, with solution (X, Y, Z, U), is

$$\begin{aligned}
\mathrm{d}X_t &= \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t + \int_E \gamma(t-, X_{t-}, e)\tilde{N}(\,\mathrm{d}t,\,\mathrm{d}e) \\
X_0 &= x_0 \\
\mathrm{d}Y_t &= -f(t, X_t, Y_t, Z_t, V_t)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t + \int_E U_{t-}(e)\tilde{N}(\,\mathrm{d}t,\,\mathrm{d}e) \\
Y_T &= \phi(X_T)
\end{aligned} \tag{3.2.1}$$

where $V_t = \int_E U_t(e)\rho(e)\nu(\,\mathrm{d}e)$ for a given smooth and bounded function ρ. $X_t \in \mathfrak{F}_t$ is r-dimensional. The standard Brownian motion W is d-dimensional, $Y_t \in \mathfrak{F}_t$ is one-dimensional, $Z_t \in \mathfrak{F}_t$ is d-dimensional, $U_t(e) \in \mathfrak{F}_t$ is q-dimensional, $V_t \in \mathfrak{F}_t$ is 1-dimensional and $\tilde{N}$ is q-dimensional. We assume $r \leq d$ throughout the chapter.

3.3 The Expansion Scheme

3.3.1 Picard Iterations

The first step is to represent (Y, Z, U) as the limit of $\left(Y^{(k)}, Z^{(k)}, U^{(k)}\right)_{k=0}^{\infty}$, which is defined recursively by the following linear FBSDEJs

$$\begin{aligned}
\mathrm{d}X_t &= \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t + \int_E \gamma(t-, X_{t-}, e)\tilde{N}(\,\mathrm{d}t,\,\mathrm{d}e) \qquad X_0 = x_0 \\
\mathrm{d}Y_t^{(0)} &= -f(t, X_t, 0, 0, 0)\,\mathrm{d}t + Z_t^{(0)}\,\mathrm{d}W_t + \int_E U_{t-}^{(0)}(e)\tilde{N}(\,\mathrm{d}t,\,\mathrm{d}e) \qquad Y_T^{(0)} = \phi(X_T)
\end{aligned}$$

and for $k \geq 1$

$$\begin{aligned}
\mathrm{d}X_t &= \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t + \int_E \gamma(t-, X_{t-}, e)\tilde{N}(\,\mathrm{d}t,\,\mathrm{d}e) \qquad X_0 = x_0 \\
\mathrm{d}Y_t^{(k)} &= -f\big(t, X_t, Y_t^{(k-1)}, Z_t^{(k-1)}, V_t^{(k-1)}\big)\,\mathrm{d}t \\
&\quad + Z_t^{(k)}\,\mathrm{d}W_t + \int_E U_{t-}^{(k)}(e)\tilde{N}(\,\mathrm{d}t,\,\mathrm{d}e) \qquad Y_T^{(k)} = \phi(X_T).
\end{aligned}$$

In Section 3.4, it is shown that $\left(Y^{(k)}, Z^{(k)}, U^{(k)}\right) \to (Y, Z, U)$ as $k \to \infty$. Note that, to obtain zeroth-order solution $\left(Y^{(0)}, Z^{(0)}, U^{(0)}\right)$, we only need to evaluate conditional expectations. Suppose that

$$u^{(0)}(t, X_t) = Y_t^{(0)} = \mathbb{E}\left[\phi(X_T) + \int_t^T f(v, X_v, 0, 0, 0)\,\mathrm{d}v \,\middle|\, \mathcal{F}_t\right].$$

The non-linear Feynman-Kac formula, presented in Bouchard and Elie (2008), enables the following

$$Z_t^{(0)} = \partial_x u^{(0)}(t, X_t)\sigma(t, X_t) \qquad U_t^{(0)}(e) = u^{(0)}(t, X_t + \gamma(t, X_t, e)) - u^{(0)}(t, X_t).$$

The Markovian nature of the solution to the zero-th order FBSDEJ ensures a Markovian representation of $(Y^{(k)}, Z^{(k)}, U^{(k)})$. More precisely,

$$\begin{aligned} Y_t^{(k)} &= u^{(k)}(t, X_t) \\ Z_t^{(k)} &= \partial_x u^{(k)}(t, X_t)\sigma(t, X_t) \\ U_t^{(k)}(e) &= u^{(k)}(t, X_t + \gamma(t, X_t, e)) - u^{(k)}(t, X_t) \end{aligned}$$

where the functions $(u^{(k)})_{k=0}^{\infty}$ are sufficiently differentiable (to be verified) and satisfy the following PIDE for $k = 1, 2, \cdots$

$$\begin{aligned} 0 = {} & \partial_t u^{(k)}(t, x) + \mu(t, x)\partial_x u^{(k)}(t, x) + \frac{1}{2}\mathbf{Tr}[\sigma(t, x)\sigma(t, x)^\intercal \partial_x^2 u^{(k)}(t, x)] \\ & + \int_E (u^{(k)}(t, x + \gamma(t, x, e)) - u^{(k)}(t, x) - \partial_x u^{(k)}(t, x)\gamma(t, x, e))\nu(\,\mathrm{d}e) \\ & + f\Bigg(t, x, u^{(k-1)}(t, x), \partial_x u^{(k-1)}(t, x)\sigma(t, x), \\ & \int_E \rho(e)\big(u^{(k-1)}(t, x + \gamma(t, x, e)) - u^{(k-1)}(t, x)\big)\nu(\,\mathrm{d}e)\Bigg) \\ \phi(x) = {} & u^{(k)}(T, x). \end{aligned} \tag{3.3.1}$$

Assume that a fundamental solution to PIDE (3.3.1) exists and denote it by Γ, which solves

$$
\begin{aligned}
0 &= \partial_\tau \Gamma(t,x;\tau,y) + \mu(\tau,y)\partial_y \Gamma(t,x;\tau,y) + \frac{1}{2}\mathbf{Tr}[\sigma(\tau,y)\sigma(\tau,y)^\intercal \partial_y^2 \Gamma(t,x;\tau,y)] \\
&\quad + \int_E (\Gamma(t,x;\tau,y+\gamma(\tau,y,e)) - \Gamma(t,x;\tau,y) - \partial_y \Gamma(t,x;\tau,y)\gamma(\tau,y,e))\nu(\,\mathrm{d}e) \\
\delta_x(y) &= \Gamma(t,x;t+,y).
\end{aligned}
$$

Here $\delta_x(\cdot)$ is the Dirac-Delta function at x. By Duhamel's principle, the solution of the PIDE (3.3.1) is

$$
\begin{aligned}
u^{(k)}(t,x) &= \int_{\mathbb{R}^r} \mathrm{d}y\, \Gamma(t,x;T,y)\phi(y) + \int_t^T \mathrm{d}\tau \int_{\mathbb{R}^r} \mathrm{d}y\, \Gamma(t,x;\tau,y) f^{(k)}(\tau,y) \\
&:= \int_{\mathbb{R}^r} \mathrm{d}y\, \Gamma(t,x;T,y)\phi(y) + \int_t^T \mathrm{d}\tau \int_{\mathbb{R}^r} \mathrm{d}y \\
&\quad \Gamma(t,x;\tau,y) f\Big(\tau, y, u^{(k-1)}(\tau,y), \partial_x u^{(k-1)}(\tau,y)\sigma(\tau,y), \\
&\quad \int_E \rho(e)\big(u^{(k-1)}(\tau,y+\gamma(\tau,y,e)) - u^{(k-1)}(\tau,y)\big)\nu(\,\mathrm{d}e)\Big)
\end{aligned}
$$

under certain conditions. In addition, the Feynman-Kac formula suggests

$$
\mathbb{E}_{t,x}[\phi(X_T)] = \int_{\mathbb{R}^r} \mathrm{d}y\, \Gamma(t,x;T,y)\phi(y) \quad \partial_x \mathbb{E}_{t,x}[\phi(X_T)] = \int_{\mathbb{R}^r} \mathrm{d}y\, \partial_x \Gamma(t,x;T,y)\phi(y)
$$

where $\mathbb{E}_{t,x}[\cdot] := \mathbb{E}[\cdot | X_t = x]$.

3.3.2 Evaluating Conditional Expectations

By Picard iterations, we can *decompose* the solution to (3.2.1) into a sequence of nested conditional expectations. To be specific, we have

$$
\begin{aligned}
Y_t^{(k)} &= u^{(k)}(t, X_t) = \mathbb{E}\left[\phi(X_T) + \int_t^T f\big(v, X_v, Y_v^{(k-1)}, Z_v^{(k-1)}, V_v^{(k-1)}\big)\,\mathrm{d}v \middle| \mathcal{F}_t\right] \\
Z_t^{(k)} &= \partial_x u^{(k)}(t, X_t)\sigma(t, X_t) \\
U_t^{(k)}(e) &= u^{(k)}(t, X_t + \gamma(t, X_t, e)) - u^{(k)}(t, X_t)
\end{aligned}
$$

where $\big(Y_t^{(k-1)}, Z_t^{(k-1)}, U_t^{(k-1)}\big)_{k=1}^{\infty}$ are explicit functions of (t, X_t). It is therefore crucial to evaluate the conditional expectations at each Picard iteration. Approximate closed-form expressions facilitate computations.

In what follows, we propose a concrete algorithm. Euler discretization for the forward-SDE (FSDE) and Taylor polynomial expansion for the evaluations of conditional expectations are introduced. Suppose we are at time t and the terminal time is T. Introduce $n+1$ equally-spaced points $\{t_j\}_{j=0}^{n}$ in interval $[t, T]$ such that $t_0 = t$, $t_n = T$ and $h = \frac{T-t}{n}$ is the time increment. The Euler scheme reads

$$
\begin{aligned}
X_{t_j}^h &:= X_{t_{j-1}}^h + \mu\big(t_{j-1}, X_{t_{j-1}}^h\big)h + \sigma\big(t_{j-1}, X_{t_{j-1}}^h\big)\Delta W_{t_j} \\
&\quad + \int_E \gamma\big(t_{j-1}, X_{t_{j-1}}^h, e\big)\tilde{N}(h, \mathrm{d}e) \\
X_t^h &:= x_0
\end{aligned}
\tag{3.3.2}
$$

for $1 \le j \le n$ and $\Delta W_{t_j} = W_{t_j} - W_{t_{j-1}}$. Equation (3.3.2) becomes (3.3.3) with $(t, \bar{x})$ replacing $(t_{j-1}, X_{t_{j-1}})$

$$
\begin{aligned}
X_{t+h}^{h,x,\bar{x}} &:= x + \mu\,(t, \bar{x})\,h + \sigma(t, \bar{x})\Delta W_{t_j} + \int_E \gamma(t, \bar{x}, e)\tilde{N}(h, \mathrm{d}e) \\
X_t^{h,x,\bar{x}} &:= x
\end{aligned}
\tag{3.3.3}
$$

where we treat $\bar{x}$ as a constant. The characteristic function $\widehat{\Gamma}_0^{\bar{x}}(t,x;t+h,\xi)$ of the Lévy process $X_t^{h,x,\bar{x}}$ defined by Equation (3.3.3) is

$$\widehat{\Gamma}_0^{\bar{x}}(t,x;t+h,\xi) = \mathbb{E}\big[\exp\big(\mathtt{i}\xi X_{t+h}^{h,x,\bar{x}}\big)\big|X_t^{h,x,\bar{x}} = x\big] = \exp(\mathtt{i}x\xi + \Phi_0^{\bar{x}}(t,t+h;\xi))$$

where

$$\begin{aligned}\Phi_0^{\bar{x}}(t,t+h;\xi) = \mathtt{i}\xi\mu(t,\bar{x})h &- \frac{1}{2}\xi\Sigma(t,\bar{x})\xi^{\intercal}h \\ &+ h\int_E(\exp(\mathtt{i}\xi\gamma(t,\bar{x},e)) - 1 - \mathtt{i}\xi\gamma(t,\bar{x},e))\nu(\,\mathrm{d}e)\end{aligned}$$

with $\Sigma = \sigma\sigma^{\intercal}$. Once we compute the conditional characteristic function, we replace ξ with $-\mathtt{i}\xi$ to get the conditional moment generating function of $X_t^{h,x,\bar{x}}$ and therefore obtain the polynomial moments of $X_t^{h,x,\bar{x}}$. Let $\Gamma_0^{\bar{x}}(t,x;t+h,y)$ be the transition density of (3.3.3), where x is the backward variable and y is the forward variable.

Next step involves the Taylor expansion of the terminal condition $\phi(\cdot)$ and the intermediate solutions. Denote $\mathbf{T}_m^{\bar{x}}$ as the Taylor expansion operator

$$\mathbf{T}_m^{\bar{x}}f(x) := \sum_{|\alpha|=0}^{m}\frac{\partial_x^\alpha f(\bar{x})}{\alpha!}(x-\bar{x})^\alpha.$$

Suppose that, at each time step, we apply $\mathbf{T}_m^{x_0}$ on the intermediate solutions, where x_0 is the starting point of the FSDE at time $t = t_0$. The expansion solution is

$$\begin{aligned}u_{k,m,n}^{\bar{x}}(t,x) := &\int_{\mathbb{R}^r}\mathrm{d}y\Gamma_0^{\bar{x}}(t,x;T,y)\mathbf{T}_m^{x_0}\phi(y) \\ &+\int_t^T\mathrm{d}\tau\int_{\mathbb{R}^r}\mathrm{d}y\Gamma_0^{\bar{x}}(t,x;\tau,y)\mathbf{T}_m^{x_0}f^{(k)}(\tau,y)\end{aligned} \tag{3.3.4}$$

for $t \in [t_{n-1}, T]$, and

$$u^{\bar{x}}_{k,m,n}(t,x) := \int_{\mathbb{R}^r} \mathrm{d}y \Gamma^{\bar{x}}_0(t,x;t_{i+1},y) \mathbf{T}^{x_0}_m u^y_{k,m,n}(t_{i+1},y) + \int_t^{t_{i+1}} \mathrm{d}\tau \int_{\mathbb{R}^r} \mathrm{d}y \Gamma^{\bar{x}}_0(t,x;\tau,y) \mathbf{T}^{x_0}_m f^{(k)}(\tau,y) \tag{3.3.5}$$

for $t \in [t_i, t_{i+1})$. Note that, $u^z_{k,m,n}(t_{i+1}, y)$ is an m-th degree polynomial in y with coefficients depending on z. The Taylor expansion reads

$$\mathbf{T}^{x_0}_m u^y_{k,m,n}(t_{i+1},y) := \sum_{j=0}^{m} \sum_{|\beta|=j} \frac{1}{\beta!} \partial^\beta_x u^x_{k,m,n}(t_{i+1},x)|_{x=x_0} (y-x_0)^\beta \tag{3.3.6}$$

where $|\beta|$, $\beta!$ and $(x-x_0)^\beta$ follow the multivariate conventions

$$\beta! = \beta_1! \times \cdots \times \beta_r! \qquad (x-\bar{x})^\beta = \prod_{j=1}^{r} (x_j - \bar{x}_j)^{\beta_j} \qquad |\beta| = \sum_{j=1}^{r} \beta_j.$$

Equations (3.3.4) and (3.3.5) are our final expansion solutions and they correspond to the probabilistic representation

$$\begin{aligned} &u^{x_0}_{k,m,n}(t_0,x_0) \\ &= \mathbb{E}[\mathbf{T}^{x_0}_m \cdots \mathbb{E}[\mathbf{T}^{x_0}_m \mathbb{E}[\mathbf{T}^{x_0}_m \phi(X^h_T)|X^h_{t_{n-1}} = \bar{x}]_{\bar{x}=x_{n-1}} |X^h_{t_{n-2}} = \bar{x}]_{\bar{x}=x_{n-2}} \cdots |X^h_{t_0} = \bar{x}]_{\bar{x}=x_0} \\ &\quad + h \sum_{j=1}^{n} \mathbb{E}[\mathbf{T}^{x_0}_m \cdots \mathbb{E}[\mathbf{T}^{x_0}_m \mathbb{E}[\mathbf{T}^{x_0}_m f^{(k)}(t_j, X^h_{t_j})|X^h_{t_{j-1}} = \bar{x}]_{\bar{x}=x_{j-1}} |X^h_{t_{j-2}} = \bar{x}]_{\bar{x}=x_{j-2}} \\ &\quad \cdots |X^h_{t_0} = \bar{x}]_{\bar{x}=x_0} \end{aligned} \tag{3.3.7}$$

where $\bar{x}$ is the fixed point at each Euler discretization step at which the coefficients μ, σ and γ are evaluated. The notation $\mathbb{E}[\phi(X^h_T)|X^h_{t_{n-1}} = \bar{x}]_{\bar{x}=x_{n-1}}$ means that we first take conditional expectations with $\bar{x}$ fixed and then set $\bar{x} = x_{n-1}$. To proceed, let

$$v^{x_0}_{k,n}(t_0,x_0) \tag{3.3.8}$$

$$
\begin{aligned}
&= \mathbb{E}[\cdots \mathbb{E}[\mathbb{E}[\phi(X_T^h)|X_{t_{n-1}}^h = \bar{x}]_{\bar{x}=x_{n-1}}|X_{t_{n-2}}^h = \bar{x}]_{\bar{x}=x_{n-2}} \cdots |X_{t_0}^h = \bar{x}]_{\bar{x}=x_0} \\
&\quad + h\sum_{j=1}^{n} \mathbb{E}[\cdots \mathbb{E}[\mathbb{E}[f^{(k)}(t_j, X_{t_j}^h)|X_{t_{j-1}}^h = \bar{x}]_{\bar{x}=x_{j-1}}|X_{t_{j-2}}^h = \bar{x}]_{\bar{x}=x_{j-2}} \\
&\quad \cdots |X_{t_0}^h = \bar{x}]_{\bar{x}=x_0}.
\end{aligned}
$$

Higher order derivatives of $v_{k,n}^x(t,x)$ and $u_{k,m,n}^x(t,x)$ are defined by

$$
\partial_x^\beta v_{k,n}^x(t,x) = [\partial_x^\beta v_{k,n}^z(t,x)]_{z=x} \qquad \partial_x^\beta u_{k,m,n}^x(t,x) = [\partial_x^\beta u_{k,m,n}^z(t,x)]_{z=x}.
$$

Note that this definition is different from that of the higher order derivatives in the Taylor expansion of the intermediate solutions.

Definition 3.3.1. *Define*

$$
\begin{aligned}
\big(Y_t^{k,m,n}, Z_t^{k,m,n}, U_t^{k,m,n}(e)\big) := \big(&u_{k,m,n}^{X_t}(t,X_t), \partial_x u_{k,m,n}^{X_t}(t,X_t)\sigma(t,X_t), \\
&u_{k,m,n}^{X_t+\gamma(t,X_t,e)}(t, X_t+\gamma(t,X_t,e)) - u_{k,m,n}^{X_t}(t,X_t)\big)
\end{aligned}
$$

as the approximate solution to $(Y_t, Z_t, U_t(e))$.

Later we will establish

$$
\lim_{k\to\infty}\lim_{n\to\infty} \big(Y_t^{k,m,n}, Z_t^{k,m,n}, U_t^{k,m,n}(e)\big) \to (Y_t, Z_t, U_t(e))
$$

in some sense.

Remark 3.3.2. *The following facts should be pointed out*

1. *The order m of the Taylor expansion need not go to infinity to establish convergence. Instead, we first send n (the number of time discretizations) and then k (the order of Picard iteration) to infinity. However, a minimum order of Taylor expansion is required (see Theorem 3.4.17). The exact space in which the convergence is established will be clear in Section 3.4.*

2. *From Equation* (3.3.6)*, we know that, for a fixed m, the total number of expansion terms is $\sum_{k=0}^{m}\binom{k+r-1}{k}$, which is a polynomial in r.*

3. *We know from the description of the numerical approximation scheme that the number of terms does not increase with the number of time discretizations.*

Because we will insert the intermediate solutions $\{u^{x_0}_{k,m,n}(t,x)\}_k$ into the next round Picard iterations and boundedness is important for us to prove convergence, we will always localize them. Denote the localization index ζ. In what follows, we will often suppress this notation. Regularization arguments such as localization can be found in Section 3.4.3.

3.4 The Error Bounds and the Convergence Result

This section first describes the spaces of random variables we will use and the technical assumptions, then introduces approximations of the original FBSDE such that the approximate FBSDEs have *well-behaved* coefficients. Error bounds are given and the convergence is established.

3.4.1 Definitions

Let $\mathcal{T}_0^T$ be the set of $\mathbb{F}$-stopping times $\tau \in [0,T]$. For a $\mathbb{R}^r$-valued function $x : [0,T] \to \mathbb{R}^r$, let the sup-norm be

$$\|x\|_{[a,b]} := \sup\{|x_t|, t \in [a,b]\}.$$

We use the following spaces for stochastic processes and $p = 2$

- $\mathbb{S}^p_r[s,t]$ is the set of $\mathbb{R}^r$-valued adapted càdlàg processes X such that

$$\|X\|_{\mathbb{S}^p_r[s,t]} := \mathbb{E}\Big[\|X(\omega)\|^p_{[s,t]}\Big]^{\frac{1}{p}} < \infty.$$

  We sometimes write $\mathbb{S}^p_r[0,T]$ as $\mathbb{S}^p_r$ if doing so causes no ambiguity. The same is true for the spaces to be defined below.

- $\mathbb{S}_r^{\infty}[s,t]$ is the set of $\mathbb{R}^r$-valued essentially bounded adapted càdlàg processes X such that

$$\|X\|_{\mathbb{S}_r^{\infty}[s,t]} := \Big\| \sup_{v\in[s,t]} |X_v| \Big\|_{\infty} < \infty.$$

  Here norm $\|\beta_t\|_{\infty} := \sup_{\omega\in\Omega} |\beta_t(\omega)|$.

- $\mathbb{H}^p[s,t]$ is the set of progressively measurable $\mathbb{R}^d$-valued processes Z such that

$$\|Z\|_{\mathbb{H}^p[s,t]} := \mathbb{E}\Bigg[\Bigg(\int_s^t |Z_v|^2 \,\mathrm{d}v\Bigg)^{\frac{p}{2}}\Bigg]^{\frac{1}{p}} < \infty.$$

- $\mathbb{J}^p[s,t]$ is the set of q-dimensional functions $\psi = \{\psi^i, 1 \le i \le q\}$, $\psi^i : \Omega \times [0,T] \times E \to \mathbb{R}$ which is $\mathbb{F} \times \mathcal{B}([0,T]) \times \mathcal{B}(E)$-measurable and satisfy

$$\|\psi\|_{\mathbb{J}^p[s,t]} := \mathbb{E}\Bigg[\Bigg(\sum_{j=1}^{q} \int_s^t \int_E |\psi^i(\omega,v,e)|^2 \nu^i(\,\mathrm{d}e)\,\mathrm{d}v\Bigg)^{\frac{p}{2}}\Bigg]^{\frac{1}{p}} < \infty.$$

- $\mathbb{J}^{\infty}[s,t]$ is the space of functions which are $\mathrm{d}\mathbb{P}\otimes\nu(\,\mathrm{d}e)$ essentially bounded, i.e.

$$\|\psi\|_{\mathbb{J}^{\infty}[s,t]} := \Big\| \sup_{v\in[s,t]} \|\psi(\cdot,v,\cdot)\|_{\mathbb{L}^{\infty}(\nu)} \Big\|_{\infty} < \infty.$$

  Here $\mathbb{L}^{\infty}(\nu)$ is the space of $\mathbb{R}^q$-valued measurable functions, $\nu(\,\mathrm{d}e)$-a.e. bounded endowed with the usual essential sup-norm.

- $\mathcal{K}^p[s,t]$ is the set of functions (Y,Z,ψ) in the space $\mathbb{S}^p[s,t] \times \mathbb{H}^p[s,t] \times \mathbb{J}^p[s,t]$ with the norm

$$\|(Y,Z,\psi)\|_{\mathcal{K}^p[s,t]} := \Big(\|Y\|^p_{\mathbb{S}^p[s,t]} + \|Z\|^p_{\mathbb{H}^p[s,t]} + \|\psi\|^p_{\mathbb{J}^p[s,t]}\Big)^{\frac{1}{p}}.$$

Let $C_b^g(D)$ be the space of bounded functions that have continuous and bounded derivatives up to order g in the domain $D \subset \mathbb{R}^r$ and $C^g(D)$ the space of functions that have continuous derivatives up to order g.

Definition 3.4.1. *Let M be a square-integrable martingale. If*

$$\|M\|^2_{\mathbf{BMO}} := \sup_{\tau\in\mathcal{T}_0^T} \left\|\mathbb{E}\left[\,(M_T - M_{\tau-}\mathbf{1}_{\tau>0})^2 \,\middle|\mathcal{F}_\tau\right]\right\|_\infty < \infty$$

then M is called a BMO-martingale. The space of BMO-martingales is **BMO**.

Further introduce the following

- $\mathbb{H}^2_{\mathbf{BMO}}$ is the set of progressively measurable $\mathbb{R}^d$-valued process Z such that

$$\|Z\|^2_{\mathbb{H}^2_{\mathbf{BMO}}} := \left\|\int_0^{\cdot} Z_v \,\mathrm{d}W_v\right\|^2_{\mathbf{BMO}} = \sup_{\tau\in\mathcal{T}_0^T}\left\|\mathbb{E}\left[\int_\tau^T |Z_v|^2\,\mathrm{d}v\middle|\mathcal{F}_\tau\right]\right\|_\infty < \infty.$$

- $\mathbb{J}^2_{\mathbf{BMO}}$ and $\mathbb{J}^2_{\mathbf{B}}$ are the sets of $\mathbb{F}\times\mathcal{B}([0,T])\times\mathcal{B}(E)$ functions $\psi:\Omega\times[0,T]\times\mathbb{E}\to\mathbb{R}^q$ satisfying

$$\|\psi\|^2_{\mathbb{J}^2_{\mathbf{BMO}}} := \left\|\int_0^{\cdot}\int_E \psi(\omega,v,e)\tilde{N}(\,\mathrm{d}v,\,\mathrm{d}e)\right\|^2_{\mathbf{BMO}} < \infty$$

$$\|\psi\|^2_{\mathbb{J}^2_{\mathbf{B}}} := \sup_{\tau\in\mathcal{T}_0^T}\left\|\mathbb{E}\left[\int_\tau^T\int_E |\psi(\omega,v,e)|^2\nu(\,\mathrm{d}e)\,\mathrm{d}v\middle|\mathcal{F}_\tau\right]\right\|_\infty < \infty.$$

3.4.2 Assumptions

Assumption 3.4.2 (On ϕ and f)**.** *For every $(x,y,z,\psi)\in\mathbb{R}^r\times\mathbb{R}\times\mathbb{R}^d\times\mathbb{L}^2(E,\nu;\mathbb{R}^q)$, there exist three constants $\beta\geq 0$, $\lambda\geq 0$ and $l\geq 0$ such that*

$$-l-\beta|y|-\frac{\lambda}{2}|z|^2-\int_E j_\lambda(-\psi(e))\nu(\,\mathrm{d}e)\leq f(t,x,y,z,\psi)$$
$$\leq l+\beta|y|+\frac{\lambda}{2}|z|^2+\int_E j_\lambda(\psi(e))\nu(\,\mathrm{d}e)$$

where $j_\lambda(u) := \frac{1}{\lambda}(\exp(\lambda u)-1-\lambda u)$. Also assume that

1. *$|\phi(X_T)|$ is essentially bounded, i.e., $\|\phi(X_T)\|_\infty<\infty$. Moreover, $\phi(\cdot)$ is bounded and Lipschitz continuous w.r.t. x.*

2. *For each $M > 0$, for every $(x, y, z, \psi) \in \mathbb{R}^r \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{L}^2(E, \nu; \mathbb{R}^q)$ and $(x', y', z', \psi') \in \mathbb{R}^r \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{L}^2(E, \nu; \mathbb{R}^q)$ satisfying the relation*

$$(|x|, |x'|, |y|, |y'|, \|\psi\|_{\mathbb{L}^\infty(\nu)}, \|\psi'\|_{\mathbb{L}^\infty(\nu)}) \leq M$$

there exists a positive constant K_M possibly depending on M such that

$$\begin{aligned} &|f(t, x, y, z, \psi) - f(t, x', y', z', \psi')| \\ &\leq K_M(\rho(|x - x'|) + |y - y'| + \|\psi - \psi'\|_{\mathbb{L}^\infty(\nu)}) \\ &\quad + K_M(1 + |z| + |z'| + \|\psi\|_{\mathbb{L}^\infty(\nu)} + \|\psi'\|_{\mathbb{L}^\infty(\nu)})|z - z'|. \end{aligned}$$

Here $\rho(x)$ is bounded and globally Lipschitz continuous, w.r.t. x, with $\rho(\mathbf{0}) = 0$.

3. *$f(t, \cdots)$ is $C_b^1([0, T])$ and α-Hölder continuous with some $0 < \alpha < \frac{1}{2}$ in t uniformly for every (x, y, z, ψ).*

Assumption 3.4.3 (On (μ, σ, γ))**.** *The following conditions are satisfied*

1. *There exists a unique strong solution X to the FSDE (3.2.1) such that $X \in \mathbb{S}_r^p[0, T]$ for any $T > 0$ and $p > 1$.*

2. *$\sigma\sigma^\intercal$ is positive-definite for $(t, x) \in [0, T] \times \mathbb{R}^r$.*

3. *For $t \in [0, T]$*

$$\max(|\mu(t, x)|, |\sigma(t, x)|, |\gamma(t, x, e)|) \leq C(1 + |x|).$$

4. *The globally Lipschitz continuity condition is satisfied for a constant C independent of (t, x, x')*

$$\begin{aligned} &\max(|\mu(t, x) - \mu(t, x')|, |\sigma(t, x) - \sigma(t, x')|, |\gamma(t, x, e) - \gamma(t, x', e)|) \\ &\leq C|x - x'|. \end{aligned}$$

5. *(μ, σ, γ) are $C_b^1([0, T])$ and α-Hölder continuous with some $0 < \alpha < \frac{1}{2}$ for every (x, e) uniformly in t.*

6. *γ satisfies for $0 \leq \theta \leq 1$*

$$C^{-1}|x - x'| \leq |(x - x') + \theta\,(\gamma(t, x, e) - \gamma(t, x', e))| \leq C|x - x'|$$

where $C > 0$ is independent of (t, x, x', e) and for $0 < \alpha < 1$

$$|\gamma(t, x, e) - \gamma(t', x', e)| \leq C \left(|x - x'|^{\alpha} + |t - t'|^{\frac{\alpha}{2}}\right)$$
$$|(x - x') + \theta\left(\gamma(t', x, e) - \gamma(t', x', e)\right)| \leq M_0 \left(|x - x'| + |t - t'|^{\frac{1}{2}}\right)$$

where $M_0 > 0$ is a constant independent of (t, t', x, x').

7. $\Gamma_0^x(t, x; v, y) \leq C_1 \exp\left(-C_2 \frac{|y-x|}{(v-t)^{\alpha}}\right)$ *for positive constants (C_1, C_2) independent of (t, v, y, x), some $0 < \alpha \leq \frac{1}{2}$ and any $0 \leq t < v \leq T$, where $\Gamma_0^x(t, x; \tau, y)$ is the transition density of $X_t^{t,\tau,x}$, satisfying*

$$X_{\tau}^{t,\tau,x} = \int_t^{\tau} \mu(v, x)\,\mathrm{d}v + \int_t^{\tau} \sigma(v, x)\,\mathrm{d}W_v + \int_t^{\tau}\int_E \gamma(v, x, e)\tilde{N}(\,\mathrm{d}v,\,\mathrm{d}e) \quad X_t^{t,\tau,x} = x.$$

Also $\Gamma(t, x; v, y) \leq C_1 \exp\left(-C_2 \frac{|y-x|}{(v-t)^{\alpha}}\right)$, where $\Gamma(t, x; v, y)$ is the transition density of X. Therefore, for any $g \geq 0$, we have $l^g \mathbb{P}\big(\sup_{v\in[t,T]} |X_v^{(t,x)}| \geq l\big) \to 0$ as $l \to \infty$.

Under Assumptions 3.4.2, 3.4.3 and (Fujii and Takahashi, 2016a, Assumption 4.1), the QEFBSDEJ (3.2.1) has a unique solution (X, Y, Z, U) in the space $\mathbb{S}_r^2[0, T] \times \mathbb{S}^{\infty} \times \mathbb{H}^2_{\mathbf{BMO}} \times \mathbb{J}^2_{\mathbf{BMO}}$ (see Fujii and Takahashi (2016a) and also Appendix A.4).

Remark 3.4.4. *As long as the approximating sequence of Lipschitz coefficients implies weak convergence of the approximate Lévy process, the conditional expectation of the approximate process will also converge to the conditional expectation of the true one under suitable uniform integrability conditions. For more details on the regularization arguments, please see the following section.*

3.4.3 Approximations

Smoothing

Because our numerical approximation scheme requires higher order derivatives of the coefficients, it is necessary to have smooth approximate coefficients. First, we need the following lemma

Lemma 3.4.5. *For functions* $(\mu, \sigma, \gamma, f, \phi)$ *which are globally Lipschitz continuous in variables* (x, y, z, w, e)*, we can find sequences of* $C^{1,\chi}$ *functions* $\{\mu_h\}_{h=1}^{\infty}$*,* $\{\sigma_h\}_{h=1}^{\infty}$*,* $\{\gamma_h\}_{h=1}^{\infty}$*,* $\{f_h\}_{h=1}^{\infty}$ *and* $\{\phi_h\}_{h=1}^{\infty}$*, such that, with* $\kappa = \mu, \sigma, \gamma, f, \phi$

$$\begin{aligned} \sup_{(t,x,y,z,w,e)\in[0,T]\times\mathbb{R}^{r+d+q+1}\times E} |\kappa_h(t,x,y,z,w,e) - \kappa(t,x,y,z,w,e)| &\leq D_h \\ \lim_{h\to\infty} D_h &= 0 \\ |\kappa_h(t,x,y,z,w,e) - \kappa_h(t,x',y',z',w',e)| &\leq C|\Theta - \Theta'| \\ h &\geq 1 \end{aligned} \tag{3.4.1}$$

where $\Theta = (x, y, z, w)$*,* $C > 0$ *is independent of* $(t, x, x', y, y', z, z', w, w', e)$ *and* $D_h > 0$ *is independent of* (t, x, y, z, w, e)*.*

Then, we have the following theorem

Theorem 3.4.6. *Assume that the functions* $(\mu, \sigma, \gamma, f, \phi)$ *are globally Lipschitz continuous in spatial variables* (x, y, z, w)*. Denote by* $\left(X^{(h)}, Y^{(h)}, Z^{(h)}, U^{(h)}\right)$ *the unique solution to the L-FBSDEJ with coefficients* $(\mu_h, \sigma_h, \gamma_h, f_h, \phi_h)$ *and by* (X, Y, Z, U) *the unique solution to the L-FBSDEJ with coefficients* $(\mu, \sigma, \gamma, f, \phi)$*, then*

$$\left\|X^{(h)} - X\right\|_{\mathbb{S}_r^2[0,T]} + \left\|(Y^{(h)} - Y, Z^{(h)} - Z, U^{(h)} - U)\right\|_{\mathcal{K}^2[0,T]} \leq C_h \qquad \lim_{h\to\infty} C_h = 0.$$

The constant $C_h > 0$ *depends on* T *and* h*.*

Localization

Next, we introduce a *localization* argument to the coefficients $(\mu_h, \sigma_h, \gamma_h, f_h, \phi_h)$. Assume two sequences of compact subsets of $\mathbb{R}^r$ denoted by $\{U_s\}_{s=1}^{\infty}$ and $\{V_s\}_{s=1}^{\infty}$ with $U_s \subseteq U_{s+1}$, $V_s \subseteq V_{s+1}$, $U_s \subseteq V_s$, $\bigcup_{s=1}^{\infty} U_s = \mathbb{R}^r$ and $\bigcup_{s=1}^{\infty} V_s = \mathbb{R}^r$. Define $\mu_{h,s}$, $\gamma_{h,s}$, $\phi_{h,s}$ and $f_{h,s}$ as $C_b^{1,\chi}$ or C_b^{χ} functions which are equal to μ_h, γ_h, ϕ_h and f_h in U_s and vanish outside V_s. Define $\sigma_{h,s}(t,x) = \sigma_h(t, \Upsilon_s(x))$. Then, with the a priori estimates in Lemma A.4.3, we can prove the following theorem

Theorem 3.4.7. *Assume that the functions* $(\mu_h, \sigma_h, \gamma_h, f_h, \phi_h)$ *are globally Lipschitz continuous in spatial variables* (x, y, z, w)*. Define by* $\left(X^{(h,s)}, Y^{(h,s)}, Z^{(h,s)}, U^{(h,s)}\right)$ *the* *unique solution to the L-FBSDEJ with coefficients* $(\mu_{h,s}, \sigma_{h,s}, \gamma_{h,s}, \phi_{h,s}, f_{h,s})$*, then*

$$\left\|X^{(h,s)} - X^{(h)}\right\|_{\mathbb{S}^2_r[0,T]} + \left\|(Y^{(h,s)} - Y^{(h)}, Z^{(h,s)} - Z^{(h)}, U^{(h,s)} - U^{(h)})\right\|_{\mathcal{K}^2[0,T]} \leq C_{h,s}$$

with $\lim_{s\to\infty} C_{h,s} = 0$ *and*

$$\left\|f\big(X^{(h,s)}\big) - f\big(X^{(h)}\big)\right\|_{\mathbb{S}^2_r[0,T]} \leq C_{h,s} \qquad\qquad \lim_{s\to\infty} C_{h,s} = 0$$

for smooth function f *with bounded derivatives of all orders.* $C_{h,s} > 0$ *depends on* T*,* h *and* s*.*

Non-degeneracy Transformation

Our scheme uses Picard iteration to linearize the L-FBSDEJ and relates the linear L-FBSDEJ obtained to a linear parabolic PIDE. To validate the representation results, a uniform ellipticity condition on σ is required. However, we only assume that $\sigma\sigma^\intercal$ is positive-definite. This section seeks a solution to this problem. The following assumption is needed

Assumption 3.4.8. *The smoothed coefficients satisfy*

$$\zeta^\intercal \sigma_h(t,x)\sigma_h(t,x)^\intercal \zeta > 0 \qquad \forall x, \zeta \in \mathbb{R}^r \qquad \forall t \in [0,T] \qquad \forall h \geq 1.$$

Assume that there exists a uniformly elliptic matrix $\mathbf{I}(t,x)$ *(or* $\mathbf{I}$ *hereafter) with bounded and sufficiently smooth elements such that* $\mathbf{I}(t,x)^{-1}\sigma_h(t,x)$ *has eigenvalues that have positive real parts and are bounded and smooth in* (t,x)*.*

Then, after we localize the coefficients, the following holds when $x \in \mathbb{R}^r$

$$0 < \zeta^\intercal \sigma_{h,s}(t,x)\sigma_{h,s}(t,x)^\intercal \zeta \leq M_{h,s}|\zeta|^2 \qquad \forall x, \zeta \in \mathbb{R}^q \qquad \forall t \in [0,T] \qquad \forall (h,s).$$

Here $M_{h,s}$ is a constant depending on (h,s) only. Without loss of generality, we assume that $\sigma_{h,s}$ is a square matrix and either $\det(\sigma_{h,s}) > 0$ or $\det(\sigma_{h,s}) < 0$ holds almost everywhere under the product Lebesgue measure on $\mathcal{B}([0,T]) \otimes \mathcal{B}(\mathbb{R}^r)$. Let $\Sigma_{h,s} = \sigma_{h,s}\sigma_{h,s}^\intercal$ and $\Sigma_{h,s,i} = \Sigma_{h,s} + \frac{1}{i}\mathbf{I}$, where $\mathbf{I}$ is the matrix of appropriate dimension, then we know that $\Sigma_{h,s,i}$ satisfies the uniform ellipticity condition. For the case where $\det(\sigma_{h,s}) > 0$, then $\sigma_{h,s,i} = \sigma_{h,s} + \frac{1}{\sqrt{i}}[(\sigma_{h,s} + \frac{1}{\sqrt{i}}\mathbf{I})(\sigma_{h,s} + \frac{1}{\sqrt{i}}\mathbf{I})^{\intercal}]^{-1}(\sigma_{h,s} + \frac{1}{\sqrt{i}}\mathbf{I}) + \frac{1}{\sqrt{i}}\mathbf{I}$ is the candidate new diffusion matrix. The related $\sigma_{h,s,i}$ and $\Sigma_{h,s,i} = \sigma_{h,s,i}\sigma_{h,s,i}^{\intercal}$ are uniformly elliptic, bounded and smooth in (t,x). This is because $\big|\det\big(\sigma_{h,s} + \frac{1}{\sqrt{i}}\mathbf{I}\big)\big| \geq \big|\det(\sigma_{h,s})\big| + \frac{1}{\sqrt{i^d}}\big|\det(\mathbf{I})\big|$, see Zhan (2005). For the case where $\det(\sigma_{h,s}) < 0$, we choose $\sigma_{h,s,i} = \sigma_{h,s} + \frac{1}{\sqrt{i}}[(\sigma_{h,s} - \frac{1}{\sqrt{i}}\mathbf{I})(\sigma_{h,s} - \frac{1}{\sqrt{i}}\mathbf{I})^{\intercal}]^{-1}(\sigma_{h,s} - \frac{1}{\sqrt{i}}\mathbf{I}) - \frac{1}{\sqrt{i}}\mathbf{I}$. Then, we have the following theorem

Theorem 3.4.9. *Let Assumption 3.4.8 hold and assume that $(\mu_h, \sigma_h, \gamma_h, f_h, \phi_h)$ are globally Lipschitz continuous. Denote by $\big(X^{(h,s,i)}, Y^{(h,s,i)}, Z^{(h,s,i)}, U^{(h,s,i)}\big)$ the unique solution to the L-FBSDEJ with the coefficients $(\mu_{h,s}, \sigma_{h,s,i}, \gamma_{h,s}, f_{h,s}, \phi_{h,s})$. Then we have*

$$\begin{aligned}&\big\|X^{(h,s,i)} - X\big\|_{\mathbb{S}_r^2[0,T]} + \big\|Y^{(h,s,i)} - Y\big\|_{\mathbb{S}^\infty[0,T]}\\ &\quad + \big\|Z^{(h,s,i)} - Z\big\|_{\mathbb{H}^\infty[0,T]} + \big\|U^{(h,s,i)} - U\big\|_{\mathbb{J}^\infty[0,T]} \leq C_{h,s,i} \qquad \lim_{h,s,i\to\infty} C_{h,s,i} = 0.\end{aligned}$$

Here constant $C_{h,s,i}$ depends on (T,h,s,i).

Remark 3.4.10. *Because of Theorem 3.4.9, we will always assume that $\sigma\sigma^{\intercal}$ is uniformly elliptic.*

In what follows, we work under the following assumption

Assumption 3.4.11. *Assume that $(\mu,\sigma,\gamma,f,\phi) \in C_b^{1,\chi}$ for all $(t,x,y,z,w,e) \in [0,T] \times \mathbb{R}^{r+d+2q+1}$ and σ is uniformly elliptic. They are the result of the above mollifying operations of smoothing, localization and non-degeneracy transformation.*

3.4.4 Error Bounds and Convergence

From now on, we assume that all the coefficients are regularized as indicated previously. We then compute error bounds and establish convergence for our approximation scheme based on Euler discretization. First, we have the following well-known theorem (Theorem 3.4.12), e.g., (Menaldi and Garroni, 1992, Theorem 3.1, Chapter II).

Theorem 3.4.12 (Existence and Uniqueness Result for PIDE (3.3.1))**.** *Under Assumption 3.4.11, there exists a unique $C_b^{1,2}([0,T]\times\mathbb{R}^r)$ solution to the PIDE system* (3.3.1).

In addition, the following theorems hold.

Theorem 3.4.13 (Convergence of Picard Iteration)**.** *Under Assumption 3.4.11, we have*

$$\left\|(Y,Z,U)-\left(Y^{(k)},Z^{(k)},U^{(k)}\right)\right\|_{\mathcal{K}^2[0,T]}\leq C\epsilon^k$$

where C is independent of ϵ and C and $0<\epsilon<1$ are independent of k.

The proof of Theorem 3.4.13 is a direct consequence of the proof of the a priori estimates documented in (Halle, 2010, Theorem 3.2) and (Halle, 2010, Lemma 3.3)[1]. For non-Lipschitz case, we refer the interested readers to Fujii and Takahashi (2016a) for the method to Lipschitzianize the quadratic-exponential driver f in a convergent way. Notice that the definitions of norms in (Halle, 2010, Section 2.2) involve a parameter β, while the definitions of our norms are special cases with $\beta=0$. However, it should be understood that $\|Y\|_{\beta_1}\leq\|Y\|_{\beta_2}$, whenever $0\leq\beta_1<\beta_2$. Also, the discussions in Halle (2010) apply to the case where the Poisson random measure is 1-dimensional. Extension to q-dimensional is straightforward. Therefore the error bound in Theorem 3.4.13 follow. We also need the next three theorems to establish convergence

Theorem 3.4.14. *Under Assumption 3.4.11, we have*

$$\left\|\left(Y^{(k)},Z^{(k)},U^{(k)}\right)-\left(Y^{(k,v)},Z^{(k,v)},U^{(k,v)}\right)\right\|_{\mathcal{K}^2[t,T]}\leq C\left(\frac{T-t}{n}\right)$$

where C is independent of n and $\left(Y^{(k,v)},Z^{(k,v)},U^{(k,v)}\right)$ is the intermediate solution at each Picard iteration by plugging $v_{k,n}$, instead of the true solution $u^{(k)}$, into the driver f.

[1]Note that, Halle (2010) only considers 1-dimensional BSDEs. However, the generalization of the results to multi-dimension case is obvious.

Theorem 3.4.15. *Under the Assumption 3.4.11, we have the following error bound and convergence*

$$\left|\partial_x^\beta u^{(k)}(t_0,x) - \partial_x^\beta v_{k,n}^x(t_0,x)\right| \le C\left(\frac{T-t}{n}\right) \qquad |\beta| \le 1$$

$$\lim_{n\to+\infty} \sup_{x\in\mathbb{R}^r} \left|\partial_x^\beta u^{(k)}(t_0,x) - \partial_x^\beta v_{k,n}^x(t_0,x)\right| = 0 \qquad |\beta| \le 1.$$

Here $u^{(k)}$ is defined in Section 3.3.1 and $v_{k,n}^x(t_0,x)$ is defined in Equation (3.3.8) *and the convergence is point-wise. The constant C depends on (T,k,β).*

Theorem 3.4.16. *Under Assumption 3.4.11, we have for $m \ge 4$ and n sufficiently large*

$$|\partial_x^\beta v_{k,n}^{x_0}(t_0,x_0) - \partial_x^\beta u_{k,m,n}^{x_0}(t_0,x_0)| \le C\left(\frac{T-t}{n}\right) \qquad |\beta| \le 1$$

and

$$\lim_{n\to+\infty} |\partial_x^\beta v_{k,n}^{x_0}(t_0,x_0) - \partial_x^\beta u_{k,m,n}^{x_0}(t_0,x_0)| = 0 \qquad |\beta| \le 1.$$

The constant C depends on x_0 and (T,k,β,m), and the convergence is established at initial point (t_0,x_0).

Based on Theorems 3.4.13, 3.4.14, 3.4.15 and 3.4.16, we have the final error bound

Theorem 3.4.17. *Under Assumption 3.4.11, we have for a sufficiently large n with $m \ge 4$*

$$\begin{aligned}&\left\|(Y,Z,U) - \left(Y^{(k,m,n)}, Z^{(k,m,n)}, U^{(k,m,n)}\right)\right\|_{\mathcal{K}^2[0,T]}\\ &\le C_{h,s,i} + K\epsilon^k + kC_\zeta + C\left(\frac{T-t}{n}\right)\end{aligned}$$

where $C_{h,s,i}$ is the error introduced by smoothing, localization and non-degeneracy transformation, $\{U_\zeta\}_{\zeta\ge1}$ is the set of compact sets that serves to localize the intermediate solutions at every Picard iteration, ζ is the index of the sequence of the compact sets U_ζ and $\lim_{\zeta\to\infty} C_\zeta = 0$. The constant C depends on (T,k,ζ,m,s).

The proof of Theorem 3.4.16 can be found in Appendix A.5 and Theorem 3.4.17 follows from the fact that, if the error bound holds at every (t_0,x_0), then for general (t, X_t) it also holds. There is an issue with respect to Lemma 3.4.5. Mollifying theory is applied and a multi-dimensional integration is needed to evaluate the mollifiers. Note that, integration, especially in high-dimensions, is not easy in general. However, in the worst case we can apply Monte-Carlo simulation to evaluate the approximate smoothed functions and their higher order derivatives. Convergence can be established because of the strong law of large numbers. We leave this exercise to the interested readers.

Remark 3.4.18. *In theorems 3.4.16 and 3.4.17, we assume that* $|\beta| \leq 1$*. The results can be generalized to the case where* $|\beta| > 1$ *with* m*, the order of Taylor expansion, satisfies* $m - |\beta| \geq 3$*.*

Chapter 4

Financial Applications

4.1 Outline of This Chapter

This chapter illustrates the financial applications of our methods. We first discuss European contingent claim valuation. Then, we introduce two dynamic portfolio choice problems. Various financial econometric topics for stochastic differential equations with jumps are studied using our numerical expansion methods. Numerical examples are given with comparisons to some selected methods in the literature.

4.2 European Option Pricing

This section applies the algorithm introduced in Section 2.3 to a European option pricing problem in an incomplete market. The exponential Ornstein-Uhlenbeck model as detailed in, e.g., Fouque et al. (2011) gives a very good fit to market prices but is notoriously hard to use computationally for option pricing. The evolution of the underlying asset price $S = (S_t)_{t\in[0,T]}$ under the physical measure $\mathbb{P}$ is

$$\begin{aligned} \mathrm{d}S_t &= \big(r + \Theta\sigma^S \exp(X_t)\big) S_t \,\mathrm{d}t + \sigma^S \exp(X_t) S_t \,\mathrm{d}W_t^1, \qquad && S_0 = s \in \mathbb{R}_{\geq 0}, \\ \mathrm{d}X_t &= (\theta - \kappa X_t)\,\mathrm{d}t + \sigma^X \,\mathrm{d}W_t^2, && X_0 = x \in \mathbb{R}, \end{aligned}$$

where W^1 and W^2 are standard $\mathbb{P}$-Brownian motions with correlation coefficient ρ. Note that the volatility-driving process $X = (X_t)_{t\in[0,T]}$, which is an Ornstein-Uhlenbeck process, affects the stock returns. Consider a European derivative with terminal payoff $\psi(S_T)$ at the maturity date T. It is known that the FBSDE associated with the price Y of the derivative is

$$\begin{aligned}
\mathrm{d}S_t &= \big(r + \Theta\sigma^S \exp(X_t)\big) S_t \,\mathrm{d}t + \sigma^S \exp(X_t) S_t \,\mathrm{d}W_t^1, && S_0 = s, \\
\mathrm{d}X_t &= (\theta - \kappa X_t)\,\mathrm{d}t + \sigma^X \,\mathrm{d}W_t^2, && X_0 = x, \\
\mathrm{d}Y_t &= -(rY_t - \Theta Z_t^1)\,\mathrm{d}t - Z_t \,\mathrm{d}W_t, \quad t \in [0, T], && Y_T = \psi(S_T).
\end{aligned}$$

Here $W = (W^1, W^2)$, $Z = (Z^1, Z^2)$ and

$$Y_t = p(t, S_t, X_t) = \mathbb{E}_t\left[\exp(-r(T-t) - \frac{1}{2}\Theta^2(T-t) - \Theta(W_T^1 - W_t^1))\psi(S_T)\right],$$

for some function p by linearity and Markovianity of the BSDE.

In our numerical experiment, the payoff function of the derivative is set to

$$\begin{aligned}
\psi(S) &= \Phi(-d_-)K\exp(-r\epsilon) - \Phi(-d_+)S, \\
d_\pm &= \frac{\log(\frac{S}{K}) + (r \pm \frac{1}{2}(\sigma^{BS})^2)\epsilon}{\sigma^{BS}\sqrt{\epsilon}},
\end{aligned}$$

where Φ is the standard Gaussian CDF. Note that when $\epsilon \to 0$, the payoff function $\psi(S) \to (K-S)^+$. Hence, $\psi(S)$ serves as a smooth approximation of the non-smooth put payoff function $(K-S)^+$. Figure 4·1 contains the plots of implied volatility obtained from inverting expansion prices using Black-Scholes formula. The 95%-confidence bands are computed for a (slow but accurate) Monte Carlo simulation.

4.3 Merton's Problem with Incomplete Markets

The second application is Merton's portfolio selection problem with incomplete markets, formulated in Liu (2007). We compare our method with Bick et al. (2013) and Briand and Labart (2012). In Bick et al. (2013), the authors try to find a near op-

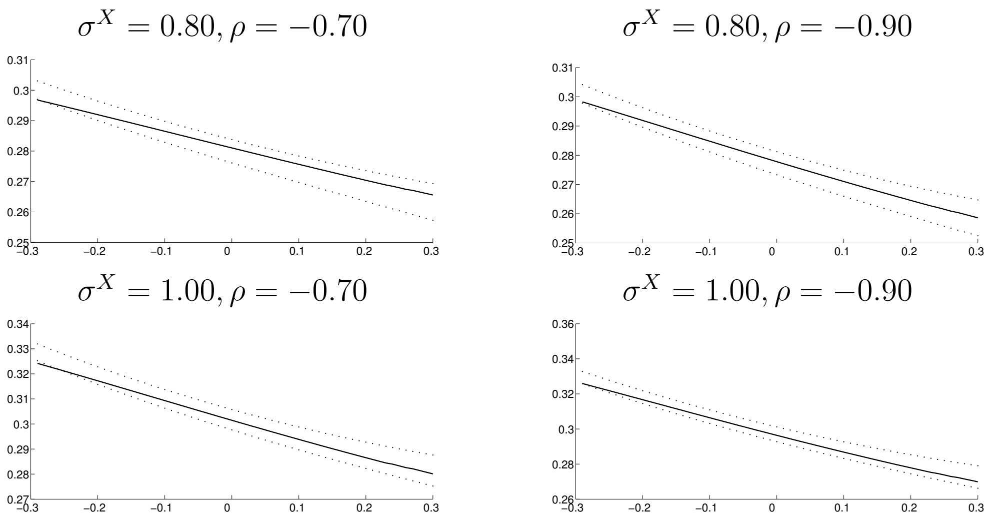


Figure 4·1: Implied volatility as a function of log-moneyness for the model considered in Section 4.2. The solid lines represent the implied volatility curves obtained from the expansion approximation of the price. The dotted lines are 95%-confidence bands of implied volatility as obtained from a Monte Carlo simulation. In all four plots, the following parameters are fixed: $T = \frac{2}{3}$, $r = 0$, $\sigma^S = 0.24$, $\kappa = 0.20$, $\theta = 0.00$, $\sigma^{BS} = 0.60$, $\epsilon = 0.50$, $S_0 = 1.00$ and $X_0 = 0.00$.

timal solution by linearizing the unspanned price of risk functions and Briand and Labart (2012) uses Wiener-Chaos expansion to construct their numerical method to solve BSDEs. The optimization problem is

$$\max_{\pi \in \mathcal{A}(x)} \mathbb{E}\left[\frac{X_T^{1-R}}{1-R}\right]$$

where $\mathcal{A}(x)$ is the set of admissible portfolio processes given initial capital x and X is the wealth process which evolves according to

$$\frac{\mathrm{d}X_t^{\pi,x}}{X_t^{\pi,x}} = r(t, Y_t)\,\mathrm{d}t + \pi_t \sigma(t, Y_t)(\Theta(t, Y_t)\,\mathrm{d}t + \mathrm{d}W_t) \qquad X_0^{\pi,x} = x.$$

We omit the integrability and adaptivity conditions for brevity. Also, $\Theta(t, y) = \sigma(t, y)^\intercal[\sigma(t, y)\sigma(t, y)^\intercal]^{-1}(\mu(t, y) - r(t, y))$ is one choice of the market price of risk function (as the market is incomplete). It is the projection of all market prices of risk on the manifold spanned by the columns of the volatility matrix. W is a standard d-dimensional Brownian motion and $Y \in \mathbb{R}^d$ is a vector of state variables satisfying the SDE

$$\mathrm{d}Y_t = \alpha(t, Y_t)\,\mathrm{d}t + \beta(t, Y_t)\,\mathrm{d}W_t \qquad\qquad Y_0 = y_0.$$

There are n risky assets with prices $S \in \mathbb{R}^n_{++}$ such that $n < d$ and

$$\mathrm{d}S_t = \mathbb{D}_{S_t}\big(\mu(t, Y_t)\,\mathrm{d}t + \sigma(t, Y_t)\,\mathrm{d}W_t\big) \qquad\qquad S_0 = s_0$$

where $\mathbb{D}_{S_t}$ is a diagonal matrix of dimension $n \times n$ with diagonal elements S_t. The quadratic FBSDE for $g(t, y, x)$, satisfying equation

$$g(t, Y_t, \xi_t) = \mathbb{E}\left[\int_t^T \xi_v I(\lambda_0^* \xi_v)\,\mathrm{d}v \,\middle|\, Y_t, \xi_t\right]$$

where λ_0^* is defined in (He and Pearson, 1991, Section 6), $I(x) = x^{-\frac{1}{R}}$, ξ_t is the state price density the agent uses in the incomplete market setting to infer his optimal portfolios. g can be represented by $g(t, y, x) = (\lambda_0^*)^{-\frac{1}{R}} \exp(h(t, y)) x^a$, where $a = 1 - \frac{1}{R}$ is a known constant and x stands for the state price density, is

$$\begin{aligned}
\mathrm{d}Y_t &= \alpha(t, Y_t)\,\mathrm{d}t + \beta(t, Y_t)\,\mathrm{d}W_t && Y_0 = y_0 \\
\mathrm{d}V_t &= -f(t, Y_t, V_t, Z_t)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t && V_T = 0 \\
f(t, y, v, z) &= -\frac{a}{2(a-1)}\|z^\intercal P(t, y)\|^2 - a\Theta(t, y)z - ar(t, y) \\
&\quad + \frac{1}{2}a(a-1)\|\Theta(t, y)\|^2 \\
\Theta(t, y) &= \sigma(t, y)^\intercal[\sigma(t, y)\sigma(t, y)^\intercal]^{-1}(\mu(t, y) - r(t, y)) \\
P(t, y) &= \mathbf{I} - \sigma(t, y)^\intercal[\sigma(t, y)\sigma(t, y)^\intercal]^{-1}\sigma(t, y).
\end{aligned}$$

Here $V_t = h(t, Y_t)$. The FBSDE is derived from the PDE in (He and Pearson, 1991, Theorem 7) using the well-known nonlinear Feynman-Kac theorem. Although lacking theoretical justification (the relation between the FBSDE and the PDE is not guaranteed to hold under our setting), it is shown that the numerical solution of the FBSDE converges to the true solution of the PDE given in (He and Pearson, 1991, Theorem 7). In this example, we take $Y = (\theta, r)$

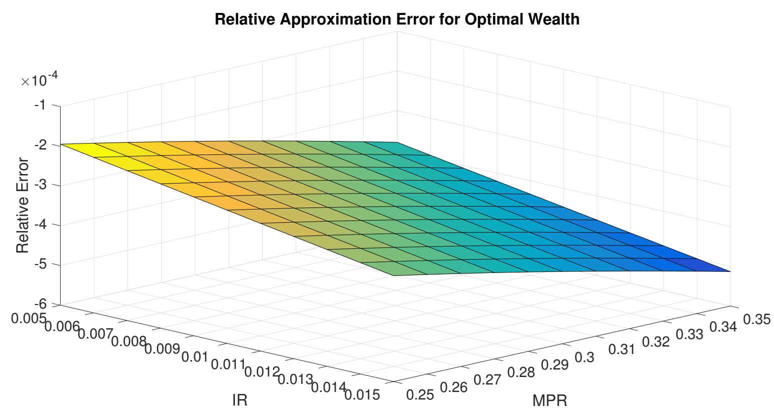
$$\begin{aligned}
\mathrm{d}\theta_t &= \kappa^\theta(\lambda^\theta - \theta_t)\,\mathrm{d}t + \sigma^\theta\,\mathrm{d}W_t^\theta & \theta_0 &= \eta \\
\mathrm{d}r_t &= \kappa^r(\lambda^r - r_t)\,\mathrm{d}t + \sigma^r\sqrt{r_t}\,\mathrm{d}(\rho W_t^\theta + \sqrt{1-\rho^2}W_t^r) & r_0 &= r \\
\frac{\mathrm{d}S_t}{S_t} &= (r_t + b\theta_t)\,\mathrm{d}t + \sigma^{S,\theta}\,\mathrm{d}W_t^\theta + \sigma^{S,r}\,\mathrm{d}W_t^r & S_0 &= s_0
\end{aligned}$$

where b is a positive constant and $b\theta_t$ is the risk premium of the stock. Then the coefficients of the PDE are

$$\begin{aligned}
\alpha(t, y) &= (\kappa^\theta(\lambda^\theta - \theta_t), \kappa^r(\lambda^r - r_t)) \\
\beta(t, y) &= \begin{pmatrix} \sigma^\theta & 0 \\ \sigma^r\rho\sqrt{r} & \sigma^r\sqrt{1-\rho^2}\sqrt{r} \end{pmatrix}.
\end{aligned}$$

A closed-form solution to this problem can be found in Liu (2007). Figure 4·2 contains the error plots for parameter values $\kappa_\theta = 2.00$, $\lambda_\theta = 0.30$, $\sigma_\theta = 0.20$, $\kappa_r = 2.00$, $\lambda_r = 0.01$, $\sigma_r = 0.10$, $b = 0.20$, $\sigma^{S,\theta} = 0.15$, $\sigma^{S,r} = 0.00$, $R = 3.00$, $T = 10.00$, $\rho = -0.50$. The running time of our approximation is 40 seconds for maturity 10 years with Taylor expansion order 2 and time discretization $2,000$, on an Intel $i7$ computer. The plots in Figure 4·2 show the surface of the errors between the approximate optimal wealth function and the true one as initial interest rate and market price of risk vary. In Tables 4.1 and 4.2, we compare the first expansion scheme with the method in Briand and Labart (2012) for a maturity of 0.20 and 1.00 years.

Relative Error



Absolute Error

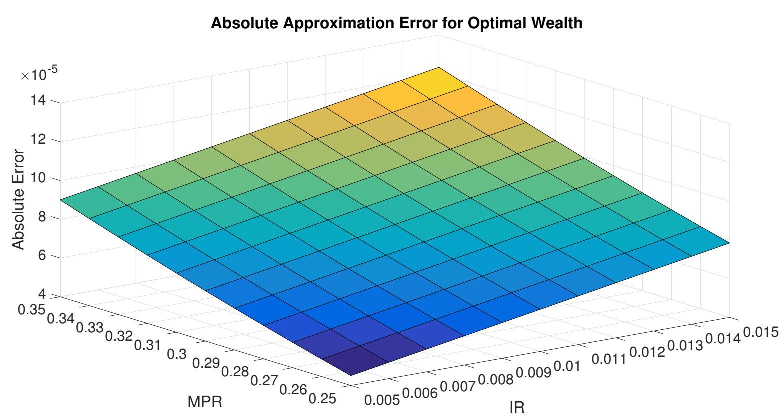


Figure 4·2: Absolute and relative errors in optimal wealth between the expansion solution and the true closed-form solution as functions of the interest rate and the market price of risk. Parameter values are $\kappa_\theta = 2.00$, $\lambda_\theta = 0.30$, $\sigma_\theta = 0.20$, $\kappa_r = 2.00$, $\lambda_r = 0.01$, $\sigma_r = 0.10$, $b = 0.20$, $\sigma^{S,\theta} = 0.15$, $\sigma^{S,r} = 0.00$, $R = 3.00$, $T = 10.00$, initial wealth $x = 1.00$ and $\rho = -0.50$. The order of Taylor expansion is 2 and the number of time discretizations is $2,000$.

$T = 0.20$ for Optimal Wealth Function					
Briand and Labart's Method			Expansion Method		
(S, C)	Time	Abs Relative Error	N	Time	Abs Relative Error
$(10000, 11)$	1.3434	6.1585%	10	0.2031	0.1344%
$(15000, 12)$	2.1094	5.9610%	20	0.4219	0.0508%
$(20000, 13)$	3.0861	5.4113%	30	0.6563	0.0245%
$(25000, 14)$	4.3176	4.9130%	40	0.8594	0.0116%
$(30000, 15)$	6.0767	4.6795%	50	1.0313	0.0041%
$(35000, 16)$	8.1127	0.0246%	60	1.2500	0.0010%

Table 4.1: Efficiency Table. Parameter values are $\kappa_\theta = 2.00$, $\lambda_\theta = 0.30$, $\sigma_\theta = 0.20$, $\kappa_r = 2.00$, $\lambda_r = 0.01$, $\sigma_r = 0.10$, $b = 0.20$, $\sigma^{S,\theta} = 0.15$, $\sigma^{S,r} = 0.00$, $R = 3.00$, $T = 0.20$ and $\rho = -0.50$.

$T = 1.00$ for Optimal Wealth Function					
Briand and Labart's Method			Expansion Method		
(S, C)	Time	Abs Relative Error	N	Time	Abs Relative Error
$(10000, 11)$	1.3794	1.6302%	10	0.2188	1.8370%
$(15000, 12)$	1.8505	1.6302%	20	0.4219	0.7994%
$(20000, 13)$	2.8979	0.8035%	30	0.6563	0.4859%
$(25000, 14)$	4.1904	0.3902%	40	0.8906	0.3350%
$(30000, 15)$	5.9301	0.1786%	150	3.8125	0.1601%
$(35000, 16)$	8.1127	0.0648%	400	9.6094	0.0541%

Table 4.2: Efficiency Table. Parameter values are the same as Table 4.1.

(S, C) is the number of simulation paths and of time discretizations. Columns 1 to 3 of Tables 4.1 and 4.2 report the results of Briand and Labart (2012), where the number of chaos is 2 and the order of Picard iteration is 5. The remaining columns contain the results from the proposed expansion method. N is the number of time discretizations. Errors are computed as absolute values of relative errors. Time to maturity is 0.20 and 1.00 years. Figure 4·2 compares our method with that of Bick et al. (2013) (strictly speaking, Bick et al. (2013) consider a more general problem with labor income, here we apply their technique to solve the problem in this numerical example) for the same parameter values as in Figure 4·2 with maturity 0.20 years. Strictly speaking, the driver of this quadratic FBSDE, although yielding a unique closed-form solution,

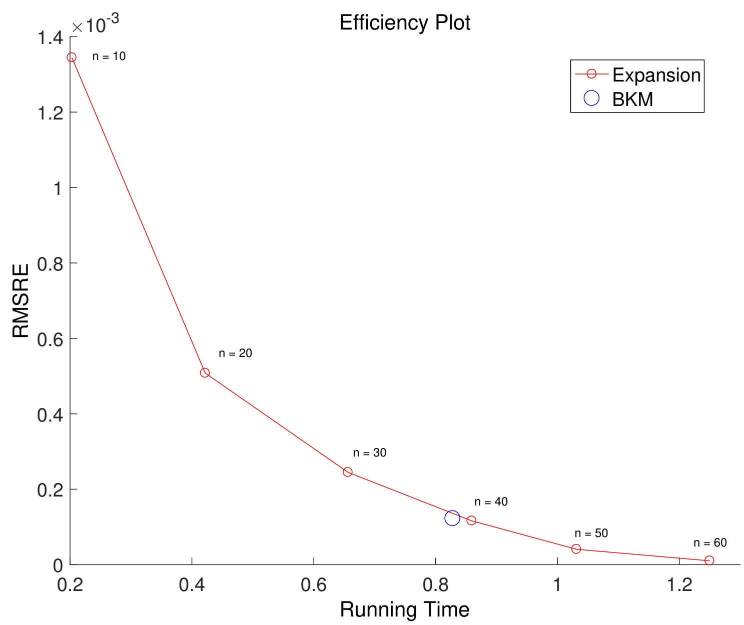


Figure 4·3: Comparison with Bick, Kraft and Munk's (BKM's, Bick et al. (2013)) Method in RMSRE-Running time space. The solid line corresponds to the expansion solution and the blue circle to BKM. The parameter values remain the same as in Figure 4·2 but time to maturity T is 0.20. The order of Taylor expansion is 2 for all the points on the red curve. Red circles correspond to time discretizations. Although accurate in our experiment, BKM's method is not convergent. Its error will therefore not go to zero as the computation time budget increases.

does not satisfy the boundedness condition on the state variables (r, θ). However, it seems that numerically our expansion solution converges. This suggests that the boundedness assumption on the coefficients of the QEFBSDEJ to guarantee that the solution exists and is unique, may not be essential for implementation of our numerical scheme.

4.4 Utility Maximization Problem for An Insurer

Next, consider the utility maximization problem in (Delong, 2013, Chapter 11) of an insurer (an investor) who can trade in a financial market with risky asset X to meet a stream of liabilities P and maximize expected utility of terminal wealth. The dynamics of the risky asset X and insurance payment P are given in Equation (4.4.1). Suppose that the insurer has an exponential utility function. The optimization problem is

$$\Phi(0, x) = \sup_{\pi \in \mathcal{A}^{\exp}(x)} \mathbb{E}\big[- \mathrm{e}^{-\alpha \Pi_T^\pi} \big]$$

where α is the absolute risk aversion parameter of the investor, Π^π is optimal wealth, π is the optimal portfolio and $\mathcal{A}^{\exp}(x)$ is the set of admissible portfolios for initial wealth x. Wealth Π^π satisfies

$$\begin{aligned}
&\mathrm{d}\Pi_t^\pi \\
&= \pi_t \frac{\mathrm{d}X_t}{X_t} + \big(\Pi_t^\pi - \pi_t\big) \frac{\mathrm{d}X_t^0}{X_t^0} - \mathrm{d}P_t \\
&= \pi_t(\mu(\nu_t)\,\mathrm{d}t + \sigma_X\,\mathrm{d}W_t) + \big(\Pi_t^\pi - \pi_t\big) r\,\mathrm{d}t - (H(P_t)\,\mathrm{d}t + G(\,\mathrm{d}N_t - \lambda\,\mathrm{d}t))
\end{aligned}$$

where $\Pi_0^\pi = x$ and X^0 is the locally riskfree asset with interest rate r. The FBSDE that characterizes the optimal solution is (see (Delong, 2013, Chapter 11))

$$
\begin{aligned}
\mathrm{d}X_t &= \mu(X_t)\,\mathrm{d}t + \sigma_X\,\mathrm{d}W_t & X_0 &= x_0\\
\mathrm{d}P_t &= H(P_t)\,\mathrm{d}t + G(\,\mathrm{d}N_t - \lambda\,\mathrm{d}t) & P_0 &= p_0\\
\mathrm{d}Y_t &= \Bigg(\frac{\mu(X_t)^2}{2\alpha\sigma_X^2} + \frac{\mu(X_t)}{\sigma_X}Z_t - H(P_t)\\
&\quad - \bigg(\frac{1}{\alpha}\big(\mathrm{e}^{\alpha(G+U_t)} - 1\big) - U_t\bigg)\lambda\Bigg)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t + U_t(\,\mathrm{d}N_t - \lambda\,\mathrm{d}t) & Y_T &= 0
\end{aligned}
\tag{4.4.1}
$$

where $(\kappa, \theta, \sigma_\nu, \sigma_X, \alpha, G, H_P, \lambda)$ are constants and the FBSDE (4.4.1) satisfies Assumptions 3.4.2, 3.4.3 and 3.4.11. Here $\mu(x) = \kappa(\theta - x)$ when $|x| < M$, $\mu(x) = 0$ when $|x| \geq M+1$. Likewise, $H(x) = H_P x$ when $|x| < M$, $H(x) = 0$ when $|x| \geq M+1$. Different pieces are concatenated smoothly and M is a large integer, for example 10^{10}, such that the functions μ and H behave like linear functions. Further, denote by λ the constant intensity of the Poisson process N_t. X is the risky asset price, P is the insurance payment, Y is related to the value function $\Phi(\cdot, x)$ and is defined such that $\pi_t^* = \frac{1}{\sigma_X}\big(Z_t + \frac{\mu(X_t)}{\alpha\sigma_X}\big)$ is the optimal portfolio.

Finding the numerical solution to this problem is challenging because this FBSDEJ has an exponentially growing driver. Although we might get numerical convergence with the schemes that only apply to Lipschitz and linearly growing drivers, theoretical convergence is not guaranteed. Second, the diffusion matrix of the FSDE is degenerate. We therefore need to use the non-degeneracy transformation argument introduced in Chapter 3.

The parameters are $\kappa = 0.20$, $\theta = 0.20$, $\sigma_\nu = 0.15$, $\sigma_X = 0.20$, $H_P = 0.20$, $G = 0.01$, $\alpha = 0.50$, $F = 0$, $X_0 = 1$, $P_0 = 0.10$, $\lambda = 0.25$ and $T = 1.00$. Table 4.3 illustrates the numerical behavior of the expansion scheme. A computation budget of 10 Picard iterations, $1,000$ time discretizations and Taylor expansion order 12 gives a value of 0.342753, which we use as the benchmark solution to compare with. Note that this example violates the Lipschitz assumption made in almost all the current references on numerical solutions to FBSDEJ.

Picard	Discretization	Expansion	Value	Time	Abs Relative Error
0	1	1	0.261494	0.0000	23.7077%
1	2	2	0.263722	0.0000	23.0577%
2	5	2	0.313304	0.0156	8.5919%
3	20	2	0.335203	0.1406	2.2028%
4	50	3	0.339808	0.7500	0.8592%
5	100	3	0.341357	2.7813	0.4073%
5	200	3	0.342134	4.7813	0.1806%
5	400	3	0.342524	9.4219	0.0668%
5	800	3	0.342719	18.5938	0.0099%

Table 4.3: Efficiency Table. The parameters are $\kappa = 0.20$, $\theta = 0.20$, $\sigma_\nu = 0.15$, $\sigma_X = 0.20$, $H_P = 0.20$, $G = 0.01$, $\alpha = 0.50$, $F = 0$, $X_0 = 1$, $P_0 = 0.10$, $\lambda = 0.25$ and $T = 1.00$.

We compare the expansion method to a recent simulation-based procedure proposed by Lejay et al. (2014). The performance of the method in Lejay et al. (2014) is summarized in Table 4.4

Time Discretization and Simulation	Value	Time	Abs Relative Error
50	0.4238	0.0029	23.6459%
100	0.3974	0.0106	15.9436%
200	0.3928	0.0220	14.6015%
500	0.3897	0.0937	13.6970%
1000	0.3810	0.4385	11.1588%
2000	0.3780	2.0820	10.2385%
5000	0.3755	16.3869	9.5541%
7000	0.3584	37.2857	4.5651%

Table 4.4: Efficiency Table. The parameters are $\kappa = 0.20$, $\theta = 0.20$, $\sigma_\nu = 0.15$, $\sigma_X = 0.20$, $H_P = 0.20$, $G = 0.01$, $\alpha = 0.50$, $F = 0$, $X_0 = 1$, $P_0 = 0.10$, $\lambda = 0.25$ and $T = 1.00$.

The method in Lejay et al. (2014) is straightforward and fast. However, it imposes a large memory requirement by letting the number of time discretizations equal the number of simulation paths. Because the running time already exceeds our expansion solution when the approximate value is still away from the *true solution*, we can claim that our method dominates in this specific example.

4.5 Density Expansion for SDEs with Jumps

In this section, we develop a transition density approximation scheme for stochastic differential equations with jumps. We first describe the SDE and the assumptions.

4.5.1 The SDE Considered

We study the following time-inhomogeneous multivariate stochastic differential equation with jumps (MSDEJ)

$$\mathrm{d}X_t = \mu(t, X_t|\theta)\,\mathrm{d}t + \sigma(t, X_t|\theta)\,\mathrm{d}W_t + \int\limits_E \gamma(t, X_t, e|\theta)\tilde{N}(\,\mathrm{d}t,\,\mathrm{d}e|\theta) \tag{4.5.1}$$

$$X_0 = x_0$$

where we explicitly state the dependence of coefficients (μ, σ, γ) on the model parameters θ, which is a g-dimensional vector taking values in a compact subset Θ of $\mathbb{R}^g$. Now, we introduce some spaces which are useful to carry out further analysis. In this chapter, we might drop θ from (μ, σ, γ) whenever doing so causes no confusions. For a $\mathbb{R}^r$-valued function $x : [0, T] \to \mathbb{R}^r$, let the sup-norm be

$$\|x\|_{[a,b]} := \sup\{|x_t|, t \in [a, b]\}.$$

- $\mathbb{S}^p_r[s, t]$ is the set of $\mathbb{R}^r$-valued adapted càdlàg processes X such that

$$\|X\|_{\mathbb{S}^p_r[s,t]} := \mathbb{E}\Big[\|X(\omega)\|^p_{[s,t]}\Big]^{\frac{1}{p}} < \infty.$$

- $\mathbb{S}_r^*[s,t]$ is the set of $\mathbb{R}^r$-valued adapted càdlàg processes X such that

$$\|X\|_{\mathbb{S}_r^*[s,t]} := \left\| \sup_{v\in[s,t]} |X_v| \right\|_{\mathbb{S}_r^2[s,t]} < \infty.$$

Let $C_b^\chi(D)$ be the space of bounded functions that have continuous and bounded derivatives up to order χ in the domain $D \subset \mathbb{R}^r$, and $C^\chi(D)$ the space of functions that have continuous derivatives up to order χ. In addition to Assumption 3.4.3, we also suppose that Assumptions $(3, 5, 6, 7)$, described in (Yu, 2007, Appendix A), hold. These assumptions are necessary for our problems to be well-defined and have solutions.

In this chapter, we will use a closed-form expansion to approximate the transition density of MSDEJ (4.5.1), prove convergence and discuss the relevant asymptotic properties of the MLE estimator based on the approximate density and score function of MSDEJ (4.5.1).

Because the error bounds and convergence results are only established for MSDEJs with $C_b^{(1,\infty)}([0,T] \times \mathbb{R}^r)$ coefficients and diffusion matrices uniformly elliptic, the following approximations are introduced: (i) smoothing, (ii) localization and (iii) non-degeneracy transformation. The idea is to approximate the original MSDEJ with a sequence of MSDEJs that have coefficients with desired properties.

The detailed mollifying arguments can be found in Chapter 3. Because of the above four approximations, we make the following assumption

Assumption 4.5.1. (μ, σ, γ) *are the result of the smoothing, localization and non-degeneracy transformation arguments.*

4.5.2 The Transition Density Approximation

Let $\Gamma(t, x; T, y)$ be the transition density. From Filipović et al. (2013), we have

$$\Gamma(t, x; T, y) := \omega\left(\frac{y-x}{\sqrt{T-t}}\right) \sum_{j=0}^{\infty} \sum_{|\alpha|=j} c_\alpha(t, T, x) p_\alpha\left(\frac{y-x}{\sqrt{T-t}}\right)$$

where ω is a probability density function and $\{p_\alpha\}_\alpha$ are the orthogonal polynomials related to ω (for details see Filipović et al. (2013)). Denote

$$\Gamma_J(t, x; T, y) := \omega\left(\frac{y-x}{\sqrt{T-t}}\right) \sum_{|\alpha|=0}^{J} c_\alpha(t, T, x) p_\alpha\left(\frac{y-x}{\sqrt{T-t}}\right)$$

where

$$c_\alpha(t, T, x) = \mathbb{E}_{t,x}\left[p_\alpha\left(\frac{X_T - x}{\sqrt{T-t}}\right)\right]$$

and conditional expectation operator $\mathbb{E}_{t,x}[\cdot] := \mathbb{E}[\cdot|X_t = x]$. We refer the readers to Chapter 3 for the computation of c_α. Denote the approximate coefficients as $c_{\alpha,m,n}$, where m denotes the order of Taylor expansion and n the number of time discretizations. We have

Theorem 4.5.2. *Denote $c_{\alpha,m,n}(t, T, x)$ as the approximate evaluation of*

$$c_\alpha(t, T, x) = \mathbb{E}_{t,x}\left[p_\alpha\left(\frac{X_T - x}{\sqrt{T-t}}\right)\right].$$

Then, we have

$$\sup_{x \in \mathbb{R}^r} |c_\alpha(t, T, x) - c_{\alpha,m,n}(t, T, x)| \leq C\left(\frac{T-t}{n}\right)$$

where the constant C is independent of x and n but might depend on m, t or T.

Proof. The proof follows from the weak convergence of the Euler discretized MSDEJ to the true MSDEJ, the localization argument and the relevant proofs in Chapter 3. □

It follows from Menaldi and Garroni (1992) that $\Gamma(t,x;T,y)$ is uniformly bounded in (x,y) under Assumption 4.5.1. In addition, we have the following convergence theorem

Theorem 4.5.3. *Given the Assumptions 3.4.3 and 4.5.1, denote the approximate transition density* $\Gamma_{J,m,n}(t,x;T,y) := \omega\left(\frac{y-x}{\sqrt{T-t}}\right)\sum_{|\alpha|=0}^{J} c_{\alpha,m,n}(t,T,x)p_\alpha\left(\frac{y-x}{\sqrt{T-t}}\right)$ *with coefficients* $c_{\alpha,m,n}(t,T,x)$*, we have*

$$\sup_{(x,y)\in\mathbb{R}^{2r}} |\Gamma_{J,m,n}(t,x;T,y) - \Gamma(t,x;T,y)| \leq C_{J,m,n}$$

where $C_{J,m,n}$ *is a constant independent of* (x,y) *and*

$$\lim_{J\to\infty}\lim_{n\to\infty} C_{J,m,n} = 0.$$

The limit above is sequential in (n,J)*.*

The proof of the above theorem can be found in the Appendix A.6.

Note that, the above results are valid under Assumption 4.5.1. The coefficients in that assumption are the result of smoothing, localization and non-degeneracy transformations. Let us, from now to the end of this section, denote the solution to the original MSDEJ as X and the mollified MSDEJ $X^{(h,s,i)}$, where (h,s,i) means smoothing, localization and non-degeneracy transformation, respectively. We then have the theorem below from the L^2 convergence of $X^{(h,s,i)}$ to X

Theorem 4.5.4. *Under Assumptions 3.4.3 and 4.5.1, we have*

$$\lim_{(h,s,i)\to 0} \Gamma_{h,s,i}(t,x;T,y) = \Gamma(t,x;T,y)$$

in a pointwise sense.

Later on, we will use the full notation $\Gamma_{h,s,i,J,m,n}(t,x;T,y)$ to denote the approximate transition density.

4.5.3 Relations to Other Methods

We compare our method, theoretically, to other methods in the literature along various dimensions. Table 4.5.3 provides a summary.

Categories	Multi-Dim	Diff	Jump	Time-Inhom	Arb-Coeffs	Convergence
AS2002	No	Yes	No	No	Yes	Asymptotic
AS2008	Yes	Yes	No	No	Yes	Asymptotic
Yu2007	Yes	Yes	Yes	No	Yes	Asymptotic
FMS2013	Yes	Yes	Yes	No	No	Global
Choi2015	Yes	Yes	No	Yes	Yes	Asymptotic
Li, Chen 2016	Yes	Yes	Yes	No	Yes	Asymptotic
This Thesis	Yes	Yes	Yes	Yes	Yes	Global

Table 4.5: Theoretical Comparisons. The references are Aït-Sahalia (2002), Aït-Sahalia (2008), Yu (2007), Filipović et al. (2013), Choi (2015) and Li and Chen (2016), respectively.

4.5.4 MLE Estimation

In this section, we discuss the MLE inference problem for the MSDEJ (4.5.1) with complete observations. Given a series of observations $\{x_{t_d}\}_{d=0}^{N}$, we are interested to find

$$\widehat{\theta}_{0,N} := \mathbf{argmax}_{\theta\in\Theta} \sum_{d=1}^{N} \log\Gamma(t_{d-1}, x_{t_{d-1}}; t_d, x_{t_d}|\theta). \quad (\mathbf{P})$$

However, as discussed previously, it is often hard to compute Γ in closed-form. Therefore, we replace Problem ($\mathbf{P}$) with the following Problem ($\mathbf{A}$)

$$\widehat{\theta}_{h,s,i,J,m,n,N} := \mathbf{argmax}_{\theta\in\Theta} \sum_{d=1}^{N} \log\Gamma_{h,s,i,J,m,n}(t_{d-1}, x_{t_{d-1}}; t_d, x_{t_d}|\theta). \quad (\mathbf{A})$$

We first define the following quantities

$$L_i(\theta) = \log\Gamma(t_{i-1}, X_{i-1}; t_i, X_i)$$

$$l_N(\theta) = \sum_{i=1}^{N} \log \Gamma(t_{i-1}, X_{i-1}; t_i, X_i)$$
$$i_N(\theta) = \sum_{i=1}^{N} \mathbb{E}_\theta[(\partial_\theta L_i(\theta))(\partial_\theta L_i(\theta))^\intercal].$$

The next theorem provides asymptotic properties for the estimator $\widehat{\theta}_{h,s,i,J,m,n,N}$

Theorem 4.5.5 (Asymptotic Properties)**.** *Under Assumption 3.4.3 and Assumptions* $(3,5,6,7)$ *in (Yu, 2007, Appendix A), we have, for a fixed sample size N*

$$\lim_{(h,s,i)\to\infty} \lim_{J\to\infty} \lim_{n\to+\infty} \widehat{\theta}_{h,s,i,J,m,n,N} =_{\mathbb{P}_{\theta_0}} \widehat{\theta}_{0,N}.$$

In particular, if we denote by $\widehat{\theta}_0$ *as the true values of the parameter, we have*

$$|\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_{0,N}| = \mathbf{o}_r(1)$$

and

$$\sqrt{Ni_N(\widehat{\theta}_0)}(\widehat{\theta}_{0,N} - \widehat{\theta}_0) = \mathcal{N}(\mathbf{0}, \mathbf{I}_{r\times r}) + \mathbf{o}_r(1)$$

as $(h,s,i,J,n,N)\to\infty$ *which makes* $\widehat{\theta}_{h,s,i,J,m,n,N}$ *and* $\widehat{\theta}_{0,N}$ *share the same asymptotic distribution described in the following*

$$\sqrt{Ni_N(\widehat{\theta}_0)}(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_0) = \mathcal{N}(\mathbf{0}, \mathbf{I}_{r\times r}) + \mathbf{o}_r(1)$$

where $\mathcal{N}(\mathbf{0}, \mathbf{I}_{r\times r})$ *denotes an* r*-dimensional Gaussian random variable of mean* $\mathbf{0}$ *and unit variance.*

The proof of this theorem can be found in Appendix A.6.

4.5.5 Numerical Experiments

A Comparison to Yu (2007)

The first numerical experiment we consider is the random-walk model

$$\mathrm{d}X_t = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t + J\,\mathrm{d}N_t \qquad\qquad X_0 = x_0.$$

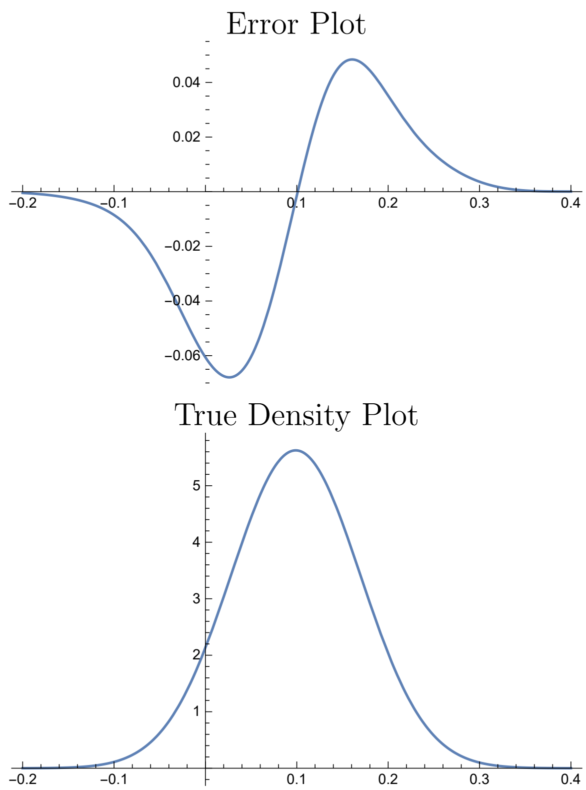


Figure 4·4: Error and Density Plots.

Here N is a Poisson process with constant intensity λ. The parameter values are $\mu = 0.10$, $\sigma = 0.50$, $J = -0.05$, $\lambda = 2.50$. Let $x_0 = 0.10$ and $T = 0.02$ representing approximately 1 week. A first order expansion using Yu (2007) and our method yields an MSE of 0.0100 and 0.0047, indicating a smaller MSE for our approximate transition density.

Next, consider the time-inhomogeneous model

$$\mathrm{d}X_t = \mu t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t + J \,\mathrm{d}N_t \qquad\qquad X_0 = x_0$$

the MSE for a 10-th order expansion approximation with our method is 0.0006. The following plot 4·4 shows the difference between the true density and the expansion approximate one. Note that Yu (2007) deals mainly with time-homogeneous jump-diffusions.

Statistical Inference of Cox-Ingersoll-Ross Model

Now, we compute the approximation to the transition density of the CIR model

$$\mathrm{d}X_t = \kappa(\theta - X_t)\,\mathrm{d}t + \sigma X_t^{\gamma}\,\mathrm{d}W_t \qquad X_0 = x_0$$

where $\kappa > 0$, $\theta > 0$, $\sigma > 0$, $\gamma = 0.50$ and $2\kappa\theta > \sigma^2$ are the constraints on the parameters. We set $(\kappa, \theta, \sigma, x_0) = (0.50, 6.00, 0.35, 6.00)$. Assume that the $\Delta = 1/50$ representing the weekly frequency and time discretization in between is set to be 100 points. Here are the plots for the true density, approximate density and the absolute error between the two densities. The MLE-inference results are listed in the following table

Parameters	True Value	Y-P	Y-P SD	Our MLE	Our SD	AS MLE	AS SD
κ	0.5000	0.7754	0.1931	0.5686	0.1566	0.5145	0.0634
θ	6.0000	5.9830	> 1.00	6.0085	0.4182	5.9846	0.4176
σ	0.3500	0.3726	0.0171	0.3563	0.0537	0.3228	0.0314
γ	0.5000	0.4966	0.0356	0.4943	0.0843	0.5478	0.0583

Table 4.6: MLE result of density expansion with $\Delta = 1/50$, Hermite order 3, Taylor order 3, time discretization 5 and 20 years of weekly data. SD means standard deviation. We take 1000 estimations and compare with Yu-Phillips (Y-P) method documented in Kawai and Maekawa (2004) and Aït-Sahalia's (AS) method in Aït-Sahalia (2008).

Statistical Inference of Cox-Ingersoll-Ross Model with Jumps

The model we consider is

$$\mathrm{d}X_t = \kappa(\theta - X_t)\,\mathrm{d}t + \sigma X_t^{\gamma}\,\mathrm{d}W_t + J(\,\mathrm{d}N_t - \lambda\,\mathrm{d}t) \qquad X_0 = x_0$$

where $\kappa > 0$, $\theta > 0$, $\sigma > 0$, $J > 0$, $\gamma = 0.50$ and $2\kappa\theta > \sigma^2$ are the constraints on the parameters. We have $\Theta = (\kappa, \theta, \sigma, J, \lambda, x_0) = (0.50, 0.18, 0.70, 0.01, 0.50, 0.18)$. We simulate a trajectory of X based on weekly frequency $T = 1/50$ and estimate the model parameters. The number of observations is 500. The result is $\widehat{\Theta} = (0.4966, 0.1795, 0.7031, 0.0100, 0.4970)$ with the robust standard deviation, computed using Ait-Sahalia's online code, $(0.0587, 0.0223, 0.0343, 0.0021, 0.1048)$. The

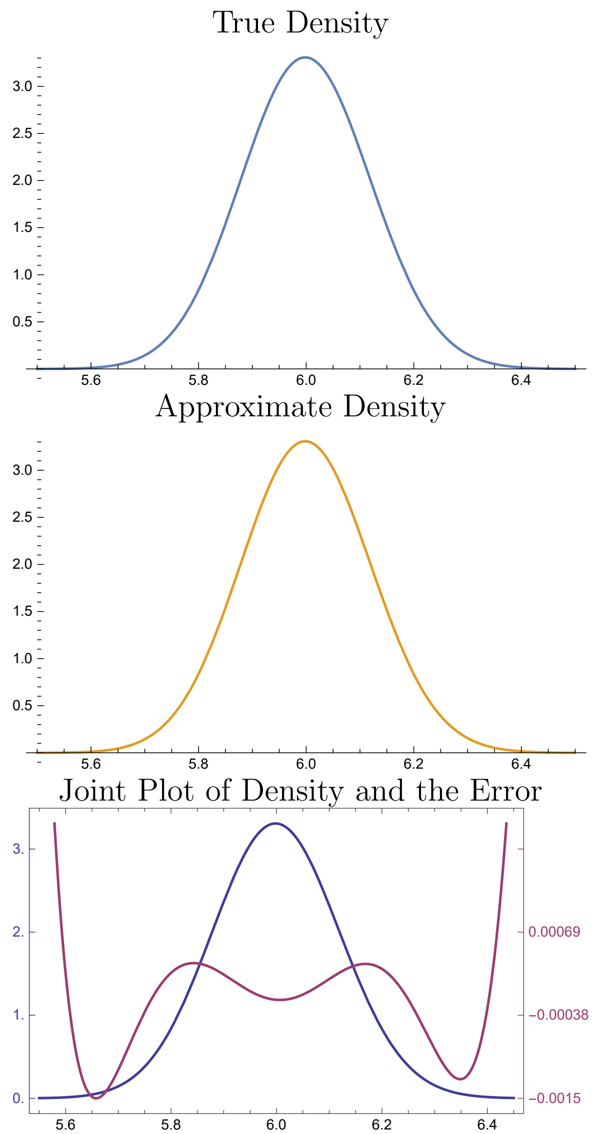


Figure 4·5: Plots for the true density, expansion density and the absolute error between two densities. Parameters are $(\kappa, \theta, \sigma, x_0) = (0.50, 6.00, 0.35, 6.00)$ with $\Delta = 1/50$ representing the weekly frequency and time discretization in between is set to be 100 points. We take Hermite order 3, Taylor order 3 and time discretization 5. For the third plot, the left-axis is the value of the approximated transition density and the right-axis is the error.

approximate density is computed using orthogonal polynomial order 3, Taylor order 3 and time discretization 6.

Chapter 5

Conclusions

This thesis proposes two convergent numerical methods to solve non-linear forward-backward stochastic differential equations, potentially with jumps. The first method is based on Picard iteration, time discretization, asymptotic PDE expansion and Taylor expansion. The approximate solution is a function of the coefficients of the FBSDE and their higher order derivatives. The second method builds on the first one. It has two additional features. First, it applies to quadratic-exponential FBSDEs with jumps. Second, it simplifies the first method, in the sense that, it Taylor-expands the intermediate solutions around the fixed-point x_0, at which the solution at time t_0 is evaluated. This results in a solution with fixed number of terms when we work backwards in time, in contrast to the first expansion method.

We apply the methods developed to various problems in finance: European derivatives pricing, dynamic portfolio choice with incomplete markets, transition density approximation for stochastic differential equations with jumps and maximum-likelihood inference problem. Numerical experiments show the effectiveness of the schemes for the selected problems, compared to some recent methods in the literature. A by-product of our schemes is an approximation method to evaluate conditional expectations of functionals of jump-diffusion processes. This approximation can be used even in the case of path-dependent functionals. In principle, our methods are applicable to all problems, financial or non-financial, that can be related to FBSDEs.

Although suitable for many financial applications, our current Taylor-expansion based methods do not work efficiently with *fast-varying functions* ($y = f(nx)$, where n is a large positive real number), which appear when we mollify non-smooth functions. This is because Taylor expansion involves higher-order derivatives of the target function and, for fast varying functions, it is only accurate in a very small neighborhood of the fixed expansion point. In order to achieve a desired accuracy, we need to assume a very small time increment (often of the order $O(\frac{1}{n^2})$) in the time discretization step and this increases the running time. To circumvent this problem, we have to seek approximation methods other than Taylor expansion, for example, Monte Carlo simulation and other polynomial expansion methods.

Future extensions of our methods include numerical solutions to more complicated equations, e.g., coupled BSDEs, (doubly) reflected BSDEs, constrained BSDEs,, 2BSDEs and Mckean-Vlasov SDEs. It would also be of interest to improve the robustness of the current methods. These extensions, which are important and non-trivial, are left for future research.

Appendix A

Appendix

A.1 The semigroup operator $\mathcal{P}_0^{\bar{x}}(t,T)$

In this appendix, we collect some basic results concerning the semigroup operator $\mathcal{P}_0^{\bar{x}}(t,T)$, defined in (2.3.7). Seen as a function of the forward variable y, the kernel $\Gamma_0^{\bar{x}}(t,x;T,y)$ associated with the semigroup operator $\mathcal{P}_0^{\bar{x}}(t,T)$ is a Gaussian density with mean vector and covariance matrix

$$\mathbf{m}^{\bar{x}}(t,T) = x + \int_t^T \mathrm{d}t' m^{\bar{x}}(t'), \qquad \mathbf{C}^{\bar{x}}(t,T) = \int_t^T \mathrm{d}t' C^{\bar{x}}(t'), \tag{A.1.1}$$

where $m^{\bar{x}} : [0,T] \to \mathbb{R}^d$ and $C^{\bar{x}}(t') : [0,T] \to \mathbb{R}^{d\times d}$ are given by

$$m^{\bar{x}}(t') = \begin{pmatrix} a^{\bar{x}}_{(1,0,\dots,0),0}(t') & a^{\bar{x}}_{(0,1,\dots,0),0}(t') & \dots & a^{\bar{x}}_{(0,0,\dots,1),0}(t') \end{pmatrix},$$

$$C^{\bar{x}}(t') = \begin{pmatrix} 2a^{\bar{x}}_{(2,0,\dots,0),0}(t') & a^{\bar{x}}_{(1,1,\dots,0),0}(t') & \dots & a^{\bar{x}}_{(1,0,\dots,1),0}(t') \\ a^{\bar{x}}_{(1,1,\dots,0),0}(t') & 2a^{\bar{x}}_{(0,2,\dots,0),0}(t') & \dots & a^{\bar{x}}_{(0,1,\dots,1),0}(t') \\ \vdots & \vdots & \ddots & \vdots \\ a^{\bar{x}}_{(1,0,\dots,1),0}(t') & a^{\bar{x}}_{(0,1,\dots,1),0}(t') & \dots & 2a^{\bar{x}}_{(0,0,\dots,2),0}(t') \end{pmatrix}.$$

Let us define the following operator

$$\mathcal{X}^{\bar{x}}(t,T) := x + \mathbf{m}^{\bar{x}}(t,T) + \mathbf{C}^{\bar{x}}(t,T)\nabla_x.$$

As shown in (Lorig et al., 2014, Theorem 2.6), the operator $\mathcal{X}^{\bar{x}}(t,T)$ has the property that, for any multi-index β, the following holds

$$\left(\mathcal{X}^{\bar{x}}(t,T)\right)^{\beta}\Gamma^{\bar{x}}(t,x;T,y) = y^{\beta}\Gamma^{\bar{x}}(t,x,T,y).$$

Using the above property, if p is a polynomial, then we have

$$\begin{aligned}\mathcal{P}_0^{\bar{x}}(t,T)p(x) &= \int\limits_{\mathbb{R}^d} \mathrm{d}y\Gamma_0^{\bar{x}}(t,x;T,y)p(y)\\ &= p(\mathcal{X}^{\bar{x}}(t,T))\int\limits_{\mathbb{R}^d} \mathrm{d}y\Gamma_0^{\bar{x}}(t,x;T,y) = p(\mathcal{X}^{\bar{x}}(t,T))1. \qquad \text{(A.1.2)}\end{aligned}$$

A.2 Proof of Theorem 2.4.8 and Corollary 2.4.9

Proof of Theorem 2.4.8. The proof of (El Karoui et al., 1997b, Corollary 2.1) gives the following error bound

$$\left\|Y_\cdot^{(k+1)} - Y_\cdot^{(k)}\right\|_\eta^2 + \left\|Z_\cdot^{(k+1)} - Z_\cdot^{(k)}\right\|_\eta^2 \leq K\delta^k,$$

for η larger than a finite constant $\bar{\eta}$, where $\delta \in (0,1)$ is independent of k, and K is a constant independent of δ and k. Here the η-norm $\|\cdot\|_\eta$ is defined by $\|\xi_\cdot\|_\eta^2 := \mathbb{E}\int_0^T \mathrm{d}t\, \mathrm{e}^{\eta t}|\xi_t|^2$. Thus, using the triangle inequality

$$\begin{aligned}&\left\|Y_\cdot - Y_\cdot^{(k)}\right\|_\eta^2 + \left\|Z_\cdot - Z_\cdot^{(k)}\right\|_\eta^2 \qquad \text{(A.2.1)}\\ &\leq \sum_{i=k}^{\infty} 2^{i+1}\left(\left\|Y_\cdot^{(i+1)} - Y_\cdot^{(i)}\right\|_\eta^2 + \left\|Z_\cdot^{(i+1)} - Z_\cdot^{(i)}\right\|_\eta^2\right) \leq 2K\frac{(2\delta)^k}{1-2\delta}.\end{aligned}$$

Moreover, it follows also from the proof of (El Karoui et al., 1997b, Corollary 2.1) that $\delta \leq \frac{\bar{\eta}}{\eta}$, thus that for all $\eta > 2\bar{\eta}$ we have $\delta \in (0,1/2)$. Next, observe that

$$\begin{aligned}&\left\|Y_\cdot - Y_\cdot^{(k,l,m,n)}\right\|_\eta^2 + \left\|Z_\cdot - Z_\cdot^{(k,l,m,n)}\right\|_\eta^2 \qquad \text{(A.2.2)}\\ &\leq 2\Bigg(\left\|Y_\cdot - Y_\cdot^{(k)}\right\|_\eta^2 + \left\|Y_\cdot^{(k)} - Y_\cdot^{(k,l,m,n)}\right\|_\eta^2\end{aligned}$$

$$+\left\|Z_\cdot - Z_\cdot^{(k)}\right\|_\eta^2 + \left\|Z_\cdot^{(k)} - Z_\cdot^{(k,l,m,n)}\right\|_\eta^2\Bigg).$$

Under Assumptions 2.4.5, 2.4.6 and 2.4.7, we have $Y_t^{(k)} = u^{(k)}(t, X_t)$ and $Z_t^{(k)} = \nabla_x u^{(k)}(t, X_t) \cdot \sigma(t, X_t)$ where $u^{(k)}$ solves (2.3.2). From (2.4.2), it therefore follows that $Y_t^{(k,l,m,n)} = \bar{u}_{l,m,n}^{(k)}(t, X_t)$, and $Z_t^{(k,l,m,n)} = \nabla_x \bar{u}_{l,m,n}^{(k)}(t, X_t) \cdot \sigma(t, X_t)$. The error bound (2.4.3) now follows for all $\eta > \bar{\eta}$ from (2.4.4), (A.2.1) and (A.2.2) by taking $|\beta| = 1$. The definition of the norm $\|\cdot\|_\eta^2$ shows that the limit (2.4.3) holds for all non-negative η once it holds for one $\bar{\eta}$. As $\|\xi_\cdot\|_{L^2}^2 \leq \|\xi_\cdot\|_\eta$ for every process ξ by the definition of the norms, the bound (2.4.3) holds in particular also for the L^2 norm. □

Proof of Corollary 2.4.9. Given (2.4.3), because the constant C is independent of n, first sending n to $+\infty$, we know that the last two terms will vanish. Then send k to $+\infty$ and the first term will also vanish. □

A.3 Proof of Proposition A.4.4

In this section, we prove Proposition A.4.4. The proof relies on a number of lemmas, which we establish below.

Lemma A.3.1. *Let Assumption 2.4.5 hold. Then, for any multi-indices $\beta, \gamma \in \mathbb{N}_0^d$, there exists a positive constant C, that depends only on T, $|\gamma|$ and $|\beta|$, such that*

$$\int_{\mathbb{R}^d} \mathrm{d}y \left|(y-x)^\gamma \partial_x^\beta \Gamma(t, x; T, y)\right| \leq C\,(T-t)^{\frac{|\gamma|-|\beta|}{2}}, \qquad 0 \leq t < T, \qquad x, y \in \mathbb{R}^d,$$

where Γ is the fundamental solution of $(\partial_t + \mathcal{A})$.

Proof. We will only prove the Lemma for $|\beta| = 1$ and $d = 1$. Higher order cases ($|\beta| \geq 2$ and $d \geq 2$) are analogous. According to (Ladyzenskaja et al., 1986, Section 4.11), we can express the fundamental solution $\Gamma(t, x; T, y)$ of the PDE (2.3.3) as

$$\Gamma(t, x; T, y) = \widehat{\Gamma}_0(t, x; T, y) + \int_t^T \mathrm{d}\tau \int_{\mathbb{R}} \mathrm{d}z[\widehat{\Gamma}_0(t, x; \tau, z) Q(\tau, z; T, y)], \quad \text{(A.3.1)}$$

where

$$\widehat{\Gamma}_0(t, x; T, y)$$

$$= \frac{1}{(4\pi(T-t))^{\frac{1}{2}}|\sigma(T,y)|} \exp\left(-\frac{1}{4(T-t)\sigma^2(T,y)}(y-x)^2\right), \qquad \text{(A.3.2)}$$

and

$$\begin{aligned}
Q(t,x;T,y) &= \sum_{m=1}^{+\infty}(-1)^m K_m(t,x;T,y), \\
K_m(t,x;T,y) &= \int_t^T \mathrm{d}t' \int_{\mathbb{R}} \mathrm{d}z K(t,x;t',z)K_{m-1}(t',z,T,y), \\
K(t,x;T,y) &= (\sigma(T,y)-\sigma(t,x))\partial_x^2\widehat{\Gamma}_0(t,x;T,y) + \mu(t,x)\partial_x\widehat{\Gamma}_0(t,x;T,y).
\end{aligned}$$

Here μ corresponds to a_1 and $\frac{1}{2}\sigma^2$ corresponds to a_2 as defined in (2.3.3). Note that the $\widehat{\Gamma}_0$ is a Gaussian kernel. It is different from the Gaussian kernel $\Gamma_0^{\bar{x}}$ introduced in the previous sections. Taking the first order derivative with respect to x on both sides of equation (A.3.1) yields

$$\partial_x^\beta\Gamma(t,x;T,y) = \partial_x^\beta\widehat{\Gamma}_0(t,x;T,y) + \int_t^T \mathrm{d}\tau \int_{\mathbb{R}} \mathrm{d}z[\partial_x^\beta\widehat{\Gamma}_0(t,x;\tau,z)Q(\tau,z;T,y)].$$

Also taking the first order derivative with respect to x on both sides of (A.3.2) and setting $T=\tau$ gives

$$\begin{aligned}
&\partial_x\widehat{\Gamma}_0(t,x;\tau,y) \\
&\quad = -\frac{1}{2(\tau-t)\sigma^2(\tau,y)}(y-x)\frac{1}{\sqrt{4\pi}|\sigma(\tau,y)|}\frac{1}{(\tau-t)^{\frac{1}{2}}} \\
&\quad \times \exp\left[-\frac{1}{4(\tau-t)\sigma^2(\tau,y)}(y-x)^2\right].
\end{aligned}$$

So we obtain the estimate

$$|\partial_x\widehat{\Gamma}_0(t,x;\tau,y)| \leq \frac{C}{(\tau-t)^{\frac{3}{2}}}(y-x)\exp\left(-C\frac{|y-x|^2}{\tau-t}\right).$$

An upper bound for Q is given in (Ladyzenskaja et al., 1986, Section 4.11)

$$|Q(t,x;\tau,y)| \leq \frac{C}{(\tau-t)} \exp\left(-C\frac{|y-x|^2}{\tau-t}\right).$$

Combining the upper bounds for $Q(t,x;\tau,y)$ and $\partial_x \widehat{\Gamma}_0(t,x;\tau,y)$, we find

$$\begin{aligned}
&\int_{\mathbb{R}} \mathrm{d}y|y-x|^\gamma |\partial_x \Gamma(t,x;T,y)| \\
&\leq \int_{\mathbb{R}} \mathrm{d}y|y-x|^\gamma |\partial_x \widehat{\Gamma}_0(t,x;T,y)| \\
&\quad + \int_{\mathbb{R}} \mathrm{d}y|y-x|^\gamma \int_t^T \mathrm{d}\tau \int_{\mathbb{R}} \mathrm{d}z |\partial_x \widehat{\Gamma}_0(t,x;\tau,z)||Q(\tau,z;T,y)| \\
&\leq \int_{\mathbb{R}} \mathrm{d}y|y-x|^{\gamma+1} \frac{C}{(T-t)^{\frac{3}{2}}} \exp\left(-\frac{C(y-x)^2}{(T-t)}\right) \\
&\quad + \int_{\mathbb{R}} \mathrm{d}y|y-x|^\gamma \int_t^T \mathrm{d}\tau \int_{\mathbb{R}} \mathrm{d}z \Bigg[|z-x| \frac{C}{(\tau-t)^{\frac{3}{2}}} \\
&\quad \times \exp\left(-\frac{C(z-x)^2}{(\tau-t)}\right) \frac{C}{(T-\tau)} \exp\left(-\frac{C(y-z)^2}{(T-\tau)}\right)\Bigg] \\
&\leq \int_{\mathbb{R}} \mathrm{d}y \frac{C}{(T-t)^{\frac{3}{2}}} |y-x|^{\gamma+1} \exp\left(-\frac{C(y-x)^2}{(T-t)}\right) \\
&\quad + \int_{\mathbb{R}} \mathrm{d}y|y-x|^\gamma \int_t^T \mathrm{d}\tau \frac{C}{(\tau-t)^{\frac{3}{2}}} \frac{1}{(T-\tau)} \\
&\quad \times \int_{\mathbb{R}} \mathrm{d}z(|z-y|+|y-x|) \exp\left(-\frac{C(y-z)^2}{(T-t)} - \frac{C(y-x)^2}{(T-t)}\right) \\
&\leq C(T-t)^{\frac{\gamma-1}{2}},
\end{aligned}$$

where in the last step we have used the triangle inequality and the higher order moments of a multivariate folded normal distribution by differentiating the moment generating function documented in (3.19) of Chakraborty and Chatterjee (2013). $\square$

Lemma A.3.2. *Define a sequence of functions* $\{\Phi_{p,m}(x)\}_{p=0}^n$ *by the following recur-*

sive relationship

$$\Phi_{p+1,m}(x) = \sum_{j=0}^{m} \sum_{|\gamma|=j} \frac{\partial_x^\gamma \Phi_{p,m}(x)}{\gamma!} v^{(\gamma)}(x) n^{-\left[\frac{j+1}{2}\right]}, \tag{A.3.3}$$

where $\{v^{(\gamma)}(x)\}_{|\gamma|=0}^m$ *and* $\Phi_{0,m}(x)$ *are bounded functions with bounded and continuous derivatives up to order* χ *and* $v^{(0,0,\cdots,0)}(x) \equiv 1$. *Here* $[x]$ *denotes the integer part of* x. *Also assume that the orders* $\rho_{\Phi_{0,m}}$ *and* $\rho_{v^{(\gamma)}}$ *are finite. Then, for sufficiently large* n, *we have*

$$\|\partial_x^\beta \Phi_{p,m}(x)\|_\infty \leq C, \quad 0 \leq p \leq n, \quad |\beta| \leq m+1, \tag{A.3.4}$$

where the constant C *does not depend on* n *or* p *and* $\|\cdot\|_\infty$ *is the sup norm over the space of bounded functions.*

Proof. We will only prove the lemma for the case $d = 1$ and coefficients in C_b^∞. The proof for the other cases is analogous. For $\Phi_{1,m}$, we have

$$\begin{aligned}
\partial_x^j \Phi_{1,m}(x) &= \partial_x^j \sum_{i=0}^{m} \frac{\partial_x^i \Phi_{0,m}(x)}{i!} v^{(i)}(x) n^{-\left[\frac{i+1}{2}\right]} \\
&= \sum_{i=0}^{m} \sum_{s=0}^{j} \binom{j}{s} \frac{\partial_x^{i+s} \Phi_{0,m}(x)}{i!} (v^{(i)}(x))^{(j-s)} n^{-\left[\frac{i+1}{2}\right]}.
\end{aligned}$$

Denote $\rho = \max(\rho_\alpha^{m+1,\chi}, \rho_\psi^{m+1,\chi}, \rho_f^{m+1,\chi})$, then we have for $1 \leq j \leq m+1$

$$\begin{aligned}
|\partial_x^j \Phi_{1,m}(x)| &\leq C\big(1 + (j+1)^{(\rho+1)(j+1)} n^{-1} + \cdots + (j+m)^{(\rho+1)(j+m)} n^{-\frac{m}{2}}\big) \\
&\leq C\big(1 + (j+m)^{(\rho+1)(j+1)} n^{-1} + \cdots + (j+m)^{(\rho+1)(j+m)} n^{-\frac{m}{2}}\big) \\
&\leq C\left(1 + (j+m)^{(\rho+1)(j+1)} n^{-1} \frac{1 - \big((j+m)^{(\rho+1)} n^{-\frac{1}{2}}\big)^m}{1 - (j+m)^{(\rho+1)} n^{-\frac{1}{2}}}\right) \\
&\leq C\big(1 + 2(2m+1)^{(\rho+1)(m+2)} n^{-1}\big).
\end{aligned}$$

Here we make use of the definition of the order in Definition 2.4.3 and set

$$C := \max_{1 \leq i \leq m+1, 1 \leq k \leq m} \big(\max(\|\partial_x^i \Phi_{0,m}\|_\infty, \|\partial_x^i v^{(k)}\|_\infty, \|\partial_x^i \Phi_{0,m}\|_\infty \times \|\partial_x^{m+1-i} v^{(k)}\|_\infty)\big).$$

We also require n to be sufficiently large such that

$$\frac{1-\left((j+m)^{(\rho+1)}n^{-\frac{1}{2}}\right)^m}{1-(j+m)^{(\rho+1)}n^{-\frac{1}{2}}} \leq 2,$$

holds true. Note that $C(1+2(2m+1)^{(\rho+1)(m+2)}n^{-1})$ is the new constant C in equation (A.3.4) for the bound of $\Phi_{1,m}$ and its derivatives. Following the same rationale, we have

$$\begin{aligned}
&|\partial_x^j \Phi_{2,m}(x)| \\
&\leq C\big(1+2(2m+1)^{(\rho+1)(m+2)}n^{-1}\big) \\
&\quad \times \big(1+(j+1)^{(\rho+1)(j+1)}n^{-1}+\cdots+(j+m)^{(\rho+1)(j+m)}n^{-\frac{m}{2}}\big) \\
&\leq C\big(1+2(2m+1)^{(\rho+1)(m+2)}n^{-1}\big) \\
&\quad \times \big(1+(j+m)^{(\rho+1)(j+1)}n^{-1}+\cdots+(j+m)^{(\rho+1)(j+m)}n^{-\frac{m}{2}}\big) \\
&\leq C\big(1+2(2m+1)^{(\rho+1)(m+2)}n^{-1}\big) \\
&\quad \times \left(1+(j+m)^{(\rho+1)(j+1)}n^{-1}\frac{1-\left((j+m)^{(\rho+1)}n^{-\frac{1}{2}}\right)^m}{1-(j+m)^{(\rho+1)}n^{-\frac{1}{2}}}\right) \\
&\leq C\big(1+2(2m+1)^{(\rho+1)(m+2)}n^{-1}\big)^2.
\end{aligned}$$

Continuing the iteration until $p=n$, we have

$$|\partial_x^j \Phi_{n,m}(x)| \leq C\big(1+2(2m+1)^{(\rho+1)(m+2)}n^{-1}\big)^n \leq C\exp\big(2(2m+1)^{(\rho+1)(m+2)}\big).$$

Thus, for any $0 \leq p \leq n$, we have $\|\partial_x^j \Phi_{p,m}\|_\infty \leq C\exp\big(2(2m+1)^{(\rho+1)(m+2)}\big)$ and note that $C\exp\big(2(2m+1)^{(\rho+1)(m+2)}\big)$ is independent of n. This concludes the proof. $\square$

Remark A.3.3. *Formulas* (2.3.12)-(2.3.13) *show that the expansion solution, when* $f \equiv 0$, *denoted by* $u^{\mathtt{tc}}_{0,m,n}(t,x)$, *satisfies the following recursion*

$$u^{\mathtt{tc}}_{0,m,n}(t_i,x) = \sum_{j=0}^{m}\sum_{|\gamma|=j}\frac{\partial_x^\gamma u^{\mathtt{tc}}_{0,m,n}(t_{i+1},x)}{\gamma!}v^{(\gamma)}(\hat{t}_i,x)n^{-\left[\frac{j+1}{2}\right]},$$

where $\hat{t}_i \in [t_i, t_{i+1})$ *and* $v^{(\gamma)}$ *is the multi-variate Gaussian moment corresponding to the multi-index* γ, *which is equal to* $\partial_v^\gamma G(v)$, *where* $v = (v_1, v_2, \cdots, v_d)$ *and* G *is the moment generating function of the multi-variate Gaussian process with transition density* Γ_0^x.

In what follows, C will always represent a constant that depends only on T, m and l, unless stated otherwise. It will be helpful to indicate the dependence of u on the terminal data ψ and the driver f. To this end, we write $u = u^{(\psi,f)} = u^{(\psi,0)} + u^{(0,f)}$ with $u^{(\psi,0)}$ and $u^{(0,f)}$ given by

$$\begin{aligned}
(\partial_t + \mathcal{A})u^{(\psi,0)} &= 0, & u^{(\psi,0)}(T,\cdot) &= \psi, \\
(\partial_t + \mathcal{A})u^{(0,f)} + f &= 0, & u^{(0,f)}(T,\cdot) &= 0.
\end{aligned}$$

By the linearity of Cauchy problem (2.3.3) we have $\partial^\beta u = \partial^\beta u^{(\psi,0)} + \partial^\beta u^{(0,f)}$. Moreover, by Duhamel's principle, the functions $\partial^\beta u^{(\psi,0)}$ and $\partial^\beta u^{(0,f)}$ are given by

$$\partial_x^\beta u^{(\psi,0)}(t,x;T) = \int\limits_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma(t,x;T,y)\psi(y), \quad \partial_x^\beta u^{(0,f)}(t,x) = \int\limits_t^T \mathrm{d}\tau \partial_x^\beta u^{(f,0)}(t,x;\tau),$$

where Γ is the fundamental solution corresponding to $(\partial_t + \mathcal{A})$, that is, Γ is the transition density of the Markov process whose generator is $\mathcal{A}$. Note, we have explicitly indicated the dependence of the function $u^{(\psi,0)}$ on the terminal time T by writing $u^{(\psi,0)}(t,x;T)$. We will drop the T-dependence when doing so causes no confusion.

Let us define $v_{l,n}^{(\psi,f),\bar{x}} \equiv v_{l,n}^{\bar{x}}$ as the solution of the following sequence of PDEs

$$\begin{cases}
\big(\partial_t + \mathcal{A}_0^{\bar{x}}\big) v_{0,n}^{\bar{x}} + f = 0, \quad t \in [t_{n-1}, T), \\
v_{0,n}^{\bar{x}}(T,\cdot) = \psi, \\
\big(\partial_t + \mathcal{A}_0^{\bar{x}}\big) v_{l,n}^{\bar{x}} + \displaystyle\sum_{i=1}^{l} \mathcal{A}_i^{\bar{x}} v_{l-i,n}^{\bar{x}} = 0, \\
v_{l,n}^{\bar{x}}(T,\cdot) = 0, \quad l \geq 1,
\end{cases} \tag{A.3.5}$$

and, for $j = 1, 2, 3, \dots, n-1$,

$$\begin{cases} \big(\partial_t + \mathcal{A}_0^{\bar{x}}\big) v_{0,n}^{\bar{x}} + f = 0, & t \in [t_{n-j-1}, t_{n-j}), \\ v_{0,n}^{\bar{x}}(t_{n-j}, \cdot) = v_{0,n}(t_{n-j}, \cdot), & \\ \big(\partial_t + \mathcal{A}_0^{\bar{x}}\big) v_{l,n}^{\bar{x}} + \sum_{i=1}^{l} \mathcal{A}_i^{\bar{x}} v_{l-i,n}^{\bar{x}} = 0, & \\ v_{l,n}^{\bar{x}}(t_{n-j}, \cdot) = v_{l,n}(t_{n-j}, \cdot), & l \geq 1, \end{cases} \tag{A.3.6}$$

where we have defined

$$v_{l,n}(t,x) := v_{l,n}^{\bar{x}}(t,x)\Big|_{\bar{x}=x}, \qquad \partial_x^\beta v_{l,n}(t,x) := \partial_x^\beta v_{l,n}^{\bar{x}}(t,x)\Big|_{\bar{x}=x}.$$

Comparing (2.3.9) and (2.3.10) with (A.3.5) and (A.3.6), we see that the only difference is that in (A.3.5) and (A.3.6) we have not applied the Taylor expansion operator $\mathbf{T}_m^{\bar{x}}$ to the terminal condition ψ, driver f or intermediate terminal solutions $v_{l,n}(t_{n-j}, \cdot)$.

By the triangle inequality, we have

$$\begin{aligned} |\partial^\beta u - \partial^\beta \bar{u}_{l,m,n}| &\leq |\partial^\beta u^{(\psi,0)} - \partial^\beta \bar{v}_{l,n}^{(\psi,0)}| + |\partial^\beta \bar{v}_{l,n}^{(\psi,0)} - \partial^\beta \bar{u}_{l,m,n}^{(\psi,0)}| \\ &\quad + |\partial^\beta u^{(0,f)} - \partial^\beta \bar{v}_{l,n}^{(0,f)}| + |\partial^\beta \bar{v}_{l,n}^{(0,f)} - \partial^\beta \bar{u}_{l,m,n}^{(0,f)}|. \end{aligned} \tag{A.3.7}$$

We will bound the supremum of each term on the right-hand side of (A.3.7) separately, which will imply a bound for the supremum over the left-hand side of (A.3.7). This is done in the following lemmas whose proofs are given at the end of this section.

Lemma A.3.4. *Under Assumptions 2.4.5 and 2.4.6, the first term on the right-hand side of* (A.3.7) *satisfies*

$$\sup_{x \in \mathbb{R}^d} \left|\partial_x^\beta u^{(\psi,0)}(t', x) - \partial_x^\beta \bar{v}_{l,n}^{(\psi,0)}(t', x)\right| \leq C \left(\frac{T-t}{n}\right)^{(l+3-|\beta|)/2}, \quad \forall t' \in [t, T], \quad \forall |\beta| \leq 1. \tag{A.3.8}$$

Lemma A.3.5. *Under Assumptions 2.4.5 and 2.4.6, the second term in the right-hand side of* (A.3.7) *satisfies*

$$\sup_{x\in\mathbb{R}^d}\left|\partial_x^\beta \bar{v}_{l,n}^{(\psi,0)}(t',x)-\partial_x^\beta \bar{u}_{l,m,n}^{(\psi,0)}(t',x)\right| \\ \le Cn^{l+1}\Big(\frac{T-t}{n}\Big)^{(m+1-|\beta|-2l)/2},\quad \forall t'\in[t,T],\quad \forall|\beta|\le m+1-2l. \quad \text{(A.3.9)}$$

Lemma A.3.6. *Given Assumptions 2.4.5 and 2.4.7, we have*

$$\partial_x^\beta v_{l,n}^{(0,f),\bar{x}}(t_0,x)=\sum_{i=0}^{n-1}\int_{t_{n-i-1}}^{t_{n-i}} \mathrm{d}t'\partial_x^\beta v_{l,n}^{(f,0),\bar{x}}(t_0,x;t'), \quad \text{(A.3.10)}$$

$$\partial_x^\beta u_{l,m,n}^{(0,f),\bar{x}}(t_0,x)=\sum_{i=0}^{n-1}\int_{t_{n-i-1}}^{t_{n-i}} \mathrm{d}t'\partial_x^\beta u_{l,m,n}^{(f,0),\bar{x}}(t_0,x;t'). \quad \text{(A.3.11)}$$

Proof of Lemma A.3.6. We only prove (A.3.11), as (A.3.10) is established in similar fashion. We discuss two cases, $l=0$ and $l=1$, and note that the case $l>1$ is analogous with only more tedious computations.

Case: $l=0$**.** The case of $l=0$ follows from the formula

$$\begin{aligned}&\partial_x^\beta u_{0,m,n}^{(0,f),\bar{x}}(t_i,x)\\ &=\partial_x^\beta u_{0,m,n}^{(0,f),\bar{x}}(t_{i+1},x)+\int_{t_i}^{t_{i+1}}\mathrm{d}t'\int_{\mathbb{R}^d}\mathrm{d}y_1\\ &\quad\partial_x^\beta\Gamma_0^{\bar{x}}(t_0,x;t_1,y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d}\mathrm{d}y_2\Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\ldots\\ &\quad\times\mathbf{T}_m^{y_{i-1}}\Big[\int_{\mathbb{R}^d}\mathrm{d}y_{i+1}\Gamma_0^{y_i}(t_i,y_i;t',y_{i+1})\mathbf{T}_m^{y_i}f(t',y_{i+1})\Big]\Big].\end{aligned} \quad \text{(A.3.12)}$$

Here $u_{0,m,n}^{(0,f),\bar{x}}(t_n,x)=0$. Iterating (A.3.12) gives

$$\partial_x^\beta u_{0,m,n}^{(0,f),\bar{x}}(t_0,x)=\int_{t_{n-1}}^{T}\mathrm{d}t'\int_{\mathbb{R}^d}\mathrm{d}y_1\partial_x^\beta\Gamma_0^{\bar{x}}(t_0,x;t_1,y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d}\mathrm{d}y_2\Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\ldots$$

$$
\begin{aligned}
&\times \mathbf{T}_m^{y_{n-2}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_n \Gamma_0^{y_{n-1}}(t_{n-1}, y_{n-1}; t', y_n)\mathbf{T}_m^{y_{n-1}} f(t', y_n)\Big]\Big] \\
&+ \int_{t_{n-2}}^{t_{n-1}} \mathrm{d}t' \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0, x; t_1, y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1, y_1; t_2, y_2)\ldots \\
&\times \mathbf{T}_m^{y_{n-3}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_{n-1} \Gamma_0^{y_{n-2}}(t_{n-2}, y_{n-2}; t', y_{n-1})\mathbf{T}_m^{y_{n-2}} f(t', y_{n-1})\Big]\Big] \\
&+ \ldots + \int_{t_0}^{t_1} \mathrm{d}t' \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0, x; t', y_1)\mathbf{T}_m^{\bar{x}} f(t', y_1) \\
= &\sum_{i=0}^{n-1} \int_{t_{n-i-1}}^{t_{n-i}} \mathrm{d}t' \partial_x^\beta u_{0,m,n}^{(f,0),\bar{x}}(t_0, x; t'). \qquad \text{(A.3.13)}
\end{aligned}
$$

Formula (A.3.13) can also be obtained by iterating (2.3.12) and (2.3.13), writing the intermediate solutions $u_{0,m,n}^{\mathtt{tc}}(t_{n-j}, x)$ explicitly and working backwards in time. Here, in (A.3.13), we interchange the Taylor expansion operator $\mathbf{T}$ and the integration operator. Also, we know from (A.3.13) that

$$
\partial_x^\beta u_{0,m,n}^{(0,f),\bar{x}}(t_j, x) = \sum_{i=0}^{n-j-1} \int_{t_{n-i-1}}^{t_{n-i}} \mathrm{d}t' \partial_x^\beta u_{0,m,n}^{(f,0),\bar{x}}(t_j, x; t'). \qquad \text{(A.3.14)}
$$

Case: $l = 1$**.** Because (A.3.14) holds, iterating (2.3.12) and (2.3.13) again by writing the intermediate solutions $u_{1,m,n}^{\mathtt{tc}}(t_{n-j}, x)$ explicitly and working backwards in time yield the following recursive relationship

$$
\begin{aligned}
&\partial_x^\beta u_{1,m,n}^{(0,f),\bar{x}}(t_i, x) \qquad \text{(A.3.15)} \\
&= \partial_x^\beta u_{1,m,n}^{(0,f),\bar{x}}(t_{i+1}, x) + \int_{t_i}^{t_{i+1}} \mathrm{d}t' \int_{\mathbb{R}^d} \mathrm{d}y_1 \\
&\quad \partial_x^\beta \Gamma_0^{\bar{x}}(t_0, x; t_1, y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1, y_1; t_2, y_2)\ldots \\
&\quad \times \mathbf{T}_m^{y_{i-1}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_{i+1} \Gamma_0^{x'}(t_i, y_i; t', y_{i+1})\mathcal{A}_1^{x'} u_{0,m,n}^{(0,f),x'}(t', y_{i+1})\Big]_{x'=y_i}\Big].
\end{aligned}
$$

Here $u^{(0,f),\bar{x}}_{1,m,n}(t_n,x)=0$. Iterate (A.3.15) to obtain

$$
\begin{aligned}
\partial_x^\beta u^{(0,f),\bar{x}}_{1,m,n}(t_0,x) = & \int_{t_{n-1}}^{T} \mathrm{d}t' \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;t_1,y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\dots \\
& \times \mathbf{T}_m^{y_{n-2}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_n \Gamma_0^{x'}(t_{n-1},y_{n-1};t',y_n)\mathcal{A}_1^{x'} u^{(0,f),x'}_{0,m,n}(t',y_n)\Big]_{x'=y_{n-1}}\Big] \\
& + \int_{t_{n-2}}^{t_{n-1}} \mathrm{d}t' \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;t_1,y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\dots \\
& \times \mathbf{T}_m^{y_{n-3}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_{n-1} \Gamma_0^{x'}(t_{n-2},y_{n-2};t',y_{n-1}) \\
& \mathcal{A}_1^{x'} u^{(0,f),x'}_{0,m,n}(t',y_{n-1})\Big]_{x'=y_{n-2}}\Big] \\
& + \ldots + \int_{t_0}^{t_1} \mathrm{d}t' \int_{\mathbb{R}^d} \mathrm{d}y_1 \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;t',y_1)\mathcal{A}_1^{\bar{x}} u^{(0,f),\bar{x}}_{0,m,n}(t',y_1)\Big] \\
= & \int_{t_{n-1}}^{T} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;t_1,y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\dots \\
& \times \mathbf{T}_m^{y_{n-2}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_n \Gamma_0^{x'}(t_{n-1},y_{n-1};\tau,y_n) \\
& \mathcal{A}_1^{x'} \int_{\tau}^{T} \mathrm{d}t' u^{(f,0),x'}_{0,m,n}(\tau,y_n;t')\Big]_{x'=y_{n-1}}\Big] \\
& + \int_{t_{n-2}}^{t_{n-1}} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;t_1,y_1)\mathbf{T}_m^{\bar{x}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\dots \\
& \times \mathbf{T}_m^{y_{n-3}}\Big[\int_{\mathbb{R}^d} \mathrm{d}y_{n-1} \Gamma_0^{x'}(t_{n-2},y_{n-2};\tau,y_{n-1}) \\
& \times \mathcal{A}_1^{x'} \sum_{i=0}^{1} \int_{t_{n-i-1}\vee\tau}^{t_{n-i}} \mathrm{d}t' u^{(f,0),x'}_{0,m,n}(\tau,y_{n-1};t')\Big]_{x'=y_{n-2}}
\end{aligned}
$$

$$
\begin{aligned}
&+\ldots+\int_{t_0}^{t_1} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y_1 \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;\tau,y_1) \\
&\mathcal{A}_1^{\bar{x}} \sum_{i=0}^{n-1} \int_{t_{n-i-1}\vee\tau}^{t_{n-i}} \mathrm{d}t' u_{0,m,n}^{(f,0),\bar{x}}(\tau,y_1;t')\Big] \\
=&\int_{t_{n-1}}^{T} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;t_1,y_1) \mathbf{T}_m^{\bar{x}} \Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\ldots \\
&\times \mathbf{T}_m^{y_{n-2}} \Big[\int_{\mathbb{R}^d} \mathrm{d}y_n \Gamma_0^{x'}(t_{n-1},y_{n-1};\tau,y_n) \\
&\mathcal{A}_1^{x'} \int_{\tau}^{T} \mathrm{d}t' u_{0,m,n}^{(f,0),x'}(\tau,y_n;t')\Big]_{x'=y_{n-1}}\Big] \\
&+\int_{t_{n-2}}^{t_{n-1}} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y_1 \partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;t_1,y_1) \mathbf{T}_m^{\bar{x}} \Big[\int_{\mathbb{R}^d} \mathrm{d}y_2 \Gamma_0^{y_1}(t_1,y_1;t_2,y_2)\ldots \\
&\times \mathbf{T}_m^{y_{n-3}} \Big[\int_{\mathbb{R}^d} \mathrm{d}y_{n-1} \Gamma_0^{x'}(t_{n-2},y_{n-2};\tau,y_{n-1}) \\
&\mathcal{A}_1^{x'} \int_{t_{n-1}}^{T} \mathrm{d}t' u_{0,m,n}^{(f,0),x'}(\tau,y_{n-1};t')\Big]_{x'=y_{n-2}}\Big] \\
&+\ldots+\int_{t_0}^{t_1} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y_1 \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;\tau,y_1) \mathcal{A}_1^{\bar{x}} \int_{t_{n-1}}^{T} \mathrm{d}t' u_{0,m,n}^{(f,0),\bar{x}}(\tau,y_1;t')\Big] \\
&+\ldots \\
&+\int_{t_0}^{t_1} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y_1 \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t_0,x;\tau,y_1) \mathcal{A}_1^{\bar{x}} \int_{\tau}^{t_1} \mathrm{d}t' u_{0,m,n}^{(f,0),\bar{x}}(\tau,y_1;t')\Big] \\
=&\sum_{i=0}^{n-1} \int_{t_{n-i-1}}^{t_{n-i}} \mathrm{d}t' \partial_x^\beta u_{1,m,n}^{(f,0),\bar{x}}(t_0,x;t').
\end{aligned}
$$

The last equality is obtained by rearranging terms and interchanging the summation and integration operators. The proof for $l > 1$ is analogous with only more tedious computations. □

As we show in the next lemma, equation (A.3.8) together with Lemmas A.3.5 and A.3.6, yields a bound on the last two terms on the right-hand side of (A.3.7).

Lemma A.3.7. *Under Assumptions 2.4.5 and 2.4.7, the third term on the right-hand side of* (A.3.7) *satisfies*

$$\sup_{x\in\mathbb{R}^d}\left|\partial_x^\beta u^{(0,f)}(t',x)-\partial_x^\beta \bar{v}_{l,n}^{(0,f)}(t',x)\right| \leq C\left(\frac{T-t}{n}\right)^{(l+3-|\beta|)/2}, \qquad \forall t'\in[t,T], \quad \forall|\beta|\leq 1, \tag{A.3.16}$$

and the fourth term satisfies

$$\begin{aligned}&\sup_{x\in\mathbb{R}^d}\left|\partial_x^\beta \bar{v}_{l,n}^{(0,f)}(t',x)-\partial_x^\beta \bar{u}_{l,m,n}^{(0,f)}(t',x)\right|\\ &\leq Cn^{l+1}\left(\frac{T-t}{n}\right)^{(m+1-|\beta|-2l)/2}, \quad \forall t'\in[t,T], \quad \forall|\beta|\leq m+1-2l.\end{aligned} \tag{A.3.17}$$

The error bound (2.4.4) follows from (A.3.7) and Lemmas A.3.4, A.3.5 and A.3.7. Hence, what remains is to give proofs of Lemmas A.3.4, A.3.5 and A.3.7.

Proof of Lemma A.3.4. Equation (A.3.8) is established in (Lorig et al., 2013, Theorem 3.12) for $|\beta|=0$, and the case of $|\beta|=1$ is a straightforward extension. □

Proof of Lemma A.3.5. To ease notation, throughout this proof, we drop the superscript $(\psi,0)$ and simply write $v^{(\psi,0)}=v$ and likewise for u. To establish a bound on $|\partial^\beta \bar{v}_{l,n}-\partial^\beta \bar{u}_{l,m,n}|$, we examine two cases: $l=0$ and $l\geq 1$.

Case: $l=0$**.** For the case of $l=0$, we are going to prove the result for any $t'\in[t,T]$ by an induction argument on the index i of time nodes $\{t_i\}_{i=0}^n$. First, we note from Taylor's theorem and Assumption 2.4.6 on ψ that there exists a $z_{\bar{x},y,\alpha}\in\mathbb{R}^d$, which depends on $\bar{x}$, y and α, such that

$$\psi(y)-\mathbf{T}_m^{\bar{x}}\psi(y)=\sum_{|\alpha|=m+1}\frac{\partial_z^\alpha\psi(z_{\bar{x},y,\alpha})}{\alpha!}(y-\bar{x})^\alpha.$$

Given the multi-index operations defined in (2.3.4), we see that

$$|(y-x)+(x-\bar{x})|^{\theta+\kappa}=\prod_{i=1}^d|(y_i-x_i)+(x_i-\bar{x}_i)|^{\theta_i+\kappa_i}$$

$$
\begin{aligned}
&= \prod_{i=1}^{d} \sum_{j=0}^{\theta_i+\kappa_i} \binom{\theta_i+\kappa_i}{j} |y_i - x_i|^j |x_i - \bar{x}_i|^{\theta_i+\kappa_i-j} \\
&\le C \sum_{|\eta+\xi|=|\theta+\kappa|} |y-x|^{\eta} |x-\bar{x}|^{\xi}.
\end{aligned}
$$

Using these results and the structure of $v^{\bar{x}}_{0,n}$, $u^{\bar{x}}_{0,m,n}$ in (2.3.12)-(2.3.13) for the case $(\psi, f) = (\psi, 0)$, we have

$$
\begin{aligned}
&\left|\partial^\beta_x v^{\bar{x}}_{0,n}(t',x) - \partial^\beta_x u^{\bar{x}}_{0,m,n}(t',x)\right| \\
&= \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial^\beta_x \Gamma^{\bar{x}}_0(t',x;T,y) \left(\psi(y) - \mathbf{T}^{\bar{x}}_m \psi(y)\right) \Big| \\
&= \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial^\beta_x \Gamma^{\bar{x}}_0(t',x;T,y) \sum_{|\alpha|=m+1} \frac{\partial^\alpha_z \psi(z_{\bar{x},y,\alpha})}{\alpha!} (y-\bar{x})^\alpha \Big| \\
&\le \sum_{|\alpha|=m+1} \int_{\mathbb{R}^d} \mathrm{d}y \Big| \partial^\beta_x \Gamma^{\bar{x}}_0(t',x;T,y) \frac{\partial^\alpha_z \psi(z_{\bar{x},y,\alpha})}{\alpha!} (y-\bar{x})^\alpha \Big| \\
&= \sum_{|\alpha|=m+1} \int_{\mathbb{R}^d} \mathrm{d}y \Big| \partial^\beta_x \Gamma^{\bar{x}}_0(t',x;T,y) \frac{\partial^\alpha_z \psi(z_{\bar{x},y,\alpha})}{\alpha!} [(y-x)+(x-\bar{x})]^\alpha \Big| \\
&\le \sum_{|\alpha|=m+1} \int_{\mathbb{R}^d} \mathrm{d}y \Big| \partial^\beta_x \Gamma^{\bar{x}}_0(t',x;T,y) \frac{\partial^\alpha_z \psi(z_{\bar{x},y,\alpha})}{\alpha!} C \sum_{\gamma+\theta=\alpha} |y-x|^\gamma |x-\bar{x}|^\theta \Big| \\
&\le C \sum_{|\gamma+\theta|=m+1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|)/2} |x-\bar{x}|^\theta, \\
&\quad \forall |\beta| \le m+1, \quad \forall t' \in [t_{n-1}, T],
\end{aligned} \tag{A.3.18}
$$

where in the last inequality, we have used Taylor's theorem and Lemma A.3.1. The constant C is independent of n. Observe that (A.3.18) holds only for $t' \in [t_{n-1}, T]$ and we will give the bound for $|\partial^\beta_x v^{\bar{x}}_{0,n}(t',x) - \partial^\beta_x u^{\bar{x}}_{0,m,n}(t',x)|$ for all $t' \in [t, T]$. Now we begin the induction argument on the index i of time nodes $\{t_i\}_{i=0}^n$. Let us assume that the following holds for some $1 \le i \le n$ with constant C independent of n

$$
\begin{aligned}
&\left|\partial^\beta_x v^{\bar{x}}_{0,n}(t',x) - \partial^\beta_x u^{\bar{x}}_{0,m,n}(t',x)\right| \\
&\le iC \sum_{|\gamma+\theta|=m+1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|)/2} |x-\bar{x}|^\theta, \\
&\quad \forall |\beta| \le m+1, \quad \forall t' \in [t_{n-i}, t_{n-i+1}).
\end{aligned} \tag{A.3.19}
$$

The case of $i = 1$ was proved in (A.3.18). If the bound (A.3.19) holds for an arbitrary i, we will show that a similar bound holds for $i + 1$. We have

$$
\begin{aligned}
&\left|\partial_x^\beta v_{0,n}^{\bar{x}}(t',x) - \partial_x^\beta u_{0,m,n}^{\bar{x}}(t',x)\right| \\
&= \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \left(v_{0,n}(t_{n-i},y) - \mathbf{T}_m^{\bar{x}} u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y)\right) \Big| \\
&= \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big[(v_{0,n}(t_{n-i},y) - v_{0,n}^{x'}(t_{n-i},y)) \\
&\quad + (v_{0,n}^{x'}(t_{n-i},y) - u_{0,m,n}^{x'}(t_{n-i},y)) \\
&\quad + (u_{0,m,n}^{x'}(t_{n-i},y) - u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y)) + (u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y) - \mathbf{T}_m^{\bar{x}} u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y)) \Big]_{x'=y} \Big| \\
&\le \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big[v_{0,n}(t_{n-i},y) - v_{0,n}^{x'}(t_{n-i},y) \Big]_{x'=y} \Big| \\
&\quad + \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big[v_{0,n}^{x'}(t_{n-i},y) - u_{0,m,n}^{x'}(t_{n-i},y) \Big]_{x'=y} \Big| \\
&\quad + \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big[u_{0,m,n}^{x'}(t_{n-i},y) - u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y) \Big]_{x'=y} \Big| \\
&\quad + \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big[u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y) - \mathbf{T}_m^{\bar{x}} u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y) \Big] \Big| \\
&\le \underbrace{\int_{\mathbb{R}^d} \mathrm{d}y \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big| v_{0,n}(t_{n-i},y) - v_{0,n}^{x'}(t_{n-i},y) \Big| \Big]_{x'=y}}_{=0} \\
&\quad + \int_{\mathbb{R}^d} \mathrm{d}y \Big[\Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big| \partial_y^\beta v_{0,n}^{x'}(t_{n-i},y) - \partial_y^\beta u_{0,m,n}^{x'}(t_{n-i},y) \Big| \Big]_{x'=y} \\
&\quad \underbrace{+ \int_{\mathbb{R}^d} \mathrm{d}y \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big| u_{0,m,n}^{x'}(t_{n-i},y) - u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y) \Big| \Big]_{x'=y}}_{=0} \\
&\quad + \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y) \Big| u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y) - \mathbf{T}_m^{\bar{x}} u_{0,m,n}^{\mathtt{tc}}(t_{n-i},y) \Big| \\
&\le (i+1) C \sum_{|\gamma+\theta|=m+1} \left(\frac{T-t}{n}\right)^{(|\gamma|-|\beta|)/2} |x-\bar{x}|^\theta, \quad \forall |\beta| \le m+1, \quad \forall t' \in [t_{n-i-1}, t_{n-i}),
\end{aligned}
\tag{A.3.20}
$$

where x' is an arbitrary point in $\mathbb{R}^d$ and we have used (2.3.13), the following symmetry property of the Gaussian kernel

$$\partial_x^\beta \Gamma_0^{\bar{x}}(t,x;T,y) = (-1)^{|\beta|}\partial_y^\beta \Gamma_0^{\bar{x}}(t,x;T,y),$$

equation (A.3.19), Lemma A.3.1 and the integration by parts formula. This establishes (A.3.19) for all $i = 1, 2, \ldots, n$, where the constant C is independent of n by Lemmas A.3.1 and A.3.2, because the intermediate solutions $u^{\mathtt{tc}}_{0,m,n}(t_{n-i}, y)$ have the same recursive structure as in (A.3.3); see Remark A.3.3. Now, setting $\bar{x} = x$ in (A.3.19), we have

$$\begin{aligned} &\left|\partial_x^\beta v_{0,n}(t',x) - \partial_x^\beta u_{0,m,n}(t',x)\right| \\ &\le iC\Big(\frac{T-t}{n}\Big)^{(m+1-|\beta|)/2}, \quad \forall|\beta| \le m+1, \quad \forall t' \in [t_{n-i}, t_{n-i+1}). \end{aligned} \tag{A.3.21}$$

As the right-hand side of (A.3.21) is independent of x and (A.3.21) holds for all $i = 0, 1, \ldots, n-1$ and $|\beta| \le m+1$, we have

$$\begin{aligned} &\sup_{x\in\mathbb{R}^d}\left|\partial_x^\beta v_{0,n}(t',x) - \partial_x^\beta u_{0,m,n}(t',x)\right| \\ &\le Cn\Big(\frac{T-t}{n}\Big)^{(m+1-|\beta|)/2}, \quad \forall\,|\beta| \le m+1, \quad \forall\, t' \in [t,T]. \end{aligned} \tag{A.3.22}$$

This establishes (A.3.9) for $l = 0$.

Case: $l \ge 1$. We are going to prove an analogue of (A.3.22) in the case of $l \ge 1$, which is

$$\begin{aligned} \sup_{x\in\mathbb{R}^d}\left|\partial_x^\beta v_{l,n}(t',x) - \partial_x^\beta u_{l,m,n}(t',x)\right| &\le Cn^{l+1}\Big(\frac{T-t}{n}\Big)^{(m+1-|\beta|-2l)/2}, \\ &\forall\,|\beta| \le m+1-2l, \quad \forall\, t' \in [t,T]. \end{aligned}$$

We will first perform a induction argument on l and inside the induction on l, we will perform a nested induction on time index i. For $l = 1$, arbitrary $\bar{x} \in \mathbb{R}^d$, $t' \in [t_{n-1}, T]$ and $|\beta| \le m-1$, similar to the argument in (A.3.18), we have

$$\Big|\int_{\mathbb{R}^d} \mathrm{d}y\Gamma_0^{\bar{x}}(t',x;\tau,y)\big(\partial_y^\beta \mathcal{A}_1^{\bar{x}} v_{0,n}^{\bar{x}}(\tau,y) - \partial_y^\beta \mathcal{A}_1^{\bar{x}} u_{0,m,n}^{\bar{x}}(\tau,y)\big)\Big|$$

$$
\begin{aligned}
&= \Big| \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t', x; \tau, y) \\
&\quad \times \Big(\partial_y^\beta \Big(\sum_{1 \le |\alpha| \le 2} \sum_{|\kappa|=1} \frac{1}{\kappa!} \partial_x^\kappa a_\alpha(t, \bar{x})(y - \bar{x})^\kappa \partial_y^\alpha \big(v_{0,n}^{\bar{x}}(\tau, y) - u_{0,m,n}^{\bar{x}}(\tau, y) \big) \Big) \Big) \Big| \\
&\le \Big| \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t', x; \tau, y) \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} C \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} |y - \bar{x}|^{\theta+\kappa} \Big| \\
&= \Big| \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t', x; \tau, y) \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} C \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} |(y - x) + (x - \bar{x})|^{\theta+\kappa} \Big| \\
&\le C \Big| \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t', x; \tau, y) \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} C \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} \\
&\quad \times \sum_{\eta+\xi=\theta} |y - x|^\eta |x - \bar{x}|^\xi \sum_{\eta+\xi=\kappa} |y - x|^\eta |x - \bar{x}|^\xi \Big| \\
&\le C \Big| \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t', x; \tau, y) \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} C \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} \\
&\quad \times \sum_{\eta+\xi=\theta} |y - x|^\eta |x - \bar{x}|^\xi \big(\sum_{\eta=\kappa} |y - x|^\eta + \sum_{\xi=\kappa} |y - \bar{x}|^\xi \big) \\
&\le C \sum_{|\gamma+\theta|=m+1} \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-1)/2} |x - \bar{x}|^\theta \\
&\quad + C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} |x - \bar{x}|^{\theta+\kappa} \\
&\le C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} \Big(|x - \bar{x}|^\theta + |x - \bar{x}|^{\theta+\kappa} \Big),
\end{aligned}
\tag{A.3.23}
$$

where constant C is independent of n. Thus, we get the following general bound for arbitrary $\bar{x}$ and for all $t' \in [t_{n-1}, T]$ using again the structure of $v_{1,n}^{\bar{x}}$, $u_{1,m,n}^{\bar{x}}$, (2.3.12) and (2.3.13) for the case $(\psi, f) = (\psi, 0)$, we have

$$
\begin{aligned}
&\big| \partial_x^\beta v_{1,n}^{\bar{x}}(t', x) - \partial_x^\beta u_{1,m,n}^{\bar{x}}(t', x) \big| \\
&= \Big| \int_{t'}^{T} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t', x; \tau, y) \big(\mathcal{A}_1^{\bar{x}} v_{0,n}^{\bar{x}}(\tau, y) - \mathcal{A}_1^{\bar{x}} u_{0,m,n}^{\bar{x}}(\tau, y) \big) \Big| \\
&= \Big| \int_{t'}^{T} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t', x; \tau, y) \big(\partial_y^\beta \mathcal{A}_1^{\bar{x}} v_{0,n}^{\bar{x}}(\tau, y) - \partial_y^\beta \mathcal{A}_1^{\bar{x}} u_{0,m,n}^{\bar{x}}(\tau, y) \big) \Big|
\end{aligned}
$$

$$
\begin{aligned}
&\leq C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2)/2} (|x-\bar{x}|^{\theta} + |x-\bar{x}|^{\theta+\kappa}),\\
&\forall |\beta| \leq m-1, \quad \forall t' \in [t_{n-1}, T].
\end{aligned}
$$

Here we have used the definition of $\mathcal{A}_1^{\bar{x}}$ in (2.3.5), inequalities (A.3.20) and (A.3.23) and Lemma A.3.1. Now we have proved the claim for $l = 1$ and $i = 0$. The next step is to prove it for $l = 1$ and any $0 \leq i \leq n-1$. This is done by induction on index i. To carry out the induction on i, we will need the following estimate: for any $i = 0, 1, 2, \ldots, n-1$,

$$
\begin{aligned}
&\Big| \int_{t'}^{t_{n-i}} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t', x; \tau, y) \big(\partial_y^{\beta} \mathcal{A}_1^{\bar{x}} v_{0,n}^{\bar{x}}(\tau, y) - \partial_y^{\beta} \mathcal{A}_1^{\bar{x}} u_{0,m,n}^{\bar{x}}(\tau, y)\big) \Big|\\
&\leq (i+1) C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2)/2} \Big(|x-\bar{x}|^{\theta+\kappa} + |x-\bar{x}|^{\theta}\Big),\\
&\qquad \forall |\beta| \leq m-1, \quad \forall t' \in [t_{n-i-1}, t_{n-i}), \quad i = 0, 1, 2, \ldots, n-1, \qquad \text{(A.3.24)}
\end{aligned}
$$

which follows from (2.3.5), (A.3.19) and Lemma A.3.1. Now, suppose that the following holds for an arbitrary i

$$
\begin{aligned}
&\big| \partial_x^{\beta} v_{1,n}^{\bar{x}}(t', x) - \partial_x^{\beta} u_{1,m,n}^{\bar{x}}(t', x) \big|\\
&\leq w(i, n, 1) C\\
&\quad \times \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2)/2} \Big(|x-\bar{x}|^{\theta+\kappa} + |x-\bar{x}|^{\theta}\Big), \qquad \text{(A.3.25)}\\
&\quad \forall |\beta| \leq m-1, \quad \forall t' \in [t_{n-i}, t_{n-i+1}),
\end{aligned}
$$

where $w(i, n, l) = (i-1)(n^l+1)+1, i = 1, 2, \ldots, n$. Note that w satisfies the following recursive relationship

$$
w(i+1, n, l) = w(i, n, l) + (n^l + 1), \quad w(1, n, l) = 1.
$$

Simple computation reveals that $w(i, n, l) = in^l + i - n^l - 1 \leq n^{l+1}$ for any $1 \leq i \leq n$. We will show that (A.3.25) holds for $i+1$. Using (2.3.12) and (2.3.13) for the case $(\psi, f) = (\psi, 0)$, we obtain

$$
\big| \partial_x^{\beta} v_{1,n}^{\bar{x}}(t', x) - \partial_x^{\beta} u_{1,m,n}^{\bar{x}}(t', x) \big|
$$

$$
\begin{aligned}
&= \Big| \int_{t'}^{t_{n-i}} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t', x; \tau, y) \big(\mathcal{A}_1^{\bar{x}} v_{0,n}^{\bar{x}}(\tau, y) - \mathcal{A}_1^{\bar{x}} u_{0,m,n}^{\bar{x}}(\tau, y) \big) \Big| \\
&\quad + \underbrace{\Big| \int_{\mathbb{R}^d} \mathrm{d}y \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t', x; t_{n-i}, y) \big(v_{1,n}(t_{n-i}, y) - v_{1,n}^{x'}(t_{n-i}, y) \big) \Big]_{x'=y} \Big|}_{=0} \\
&\quad + \Big| \int_{\mathbb{R}^d} \mathrm{d}y \Big[\Gamma_0^{\bar{x}}(t', x; t_{n-i}, y) \big(\partial_y^\beta v_{1,n}^{x'}(t_{n-i}, y) - \partial_y^\beta u_{1,m,n}^{x'}(t_{n-i}, y) \big) \Big]_{x'=y} \Big| \\
&\quad + \underbrace{\Big| \int_{\mathbb{R}^d} \mathrm{d}y \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t', x; t_{n-i}, y) \big(u_{1,m,n}^{x'}(t_{n-i}, y) - u_{1,m,n}(t_{n-i}, y) \big) \Big]_{x'=y} \Big|}_{=0} \\
&\quad + \underbrace{\Big| \int_{\mathbb{R}^d} \mathrm{d}y \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t', x; t_{n-i}, y) \big(u_{1,m,n}^{x'}(t_{n-i}, y) - u_{1,m,n}^{\mathtt{tc}}(t_{n-i}, y) \big) \Big]_{x'=y} \Big|}_{=0} \\
&\quad + \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t', x; t_{n-i}, y) \big(u_{1,m,n}^{\mathtt{tc}}(t_{n-i}, y) - \mathbf{T}_{m-2}^{\bar{x}} u_{1,m,n}^{\mathtt{tc}}(t_{n-i}, y) \big) \Big| \\
&\leq ((i+1) + w(i,n,1) + 1) C \\
&\quad \times \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} \Big(|x-\bar{x}|^{\theta+\kappa} + |x-\bar{x}|^\theta \Big) \\
&\leq (n + w(i,n,1) + 1) C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} \Big(|x-\bar{x}|^{\theta+\kappa} + |x-\bar{x}|^\theta \Big) \\
&= w(i+1,n,1) C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} \Big(|x-\bar{x}|^{\theta+\kappa} + |x-\bar{x}|^\theta \Big) \\
&\leq n^2 C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa|=1} \Big(\frac{T-t}{n} \Big)^{(|\gamma|-|\beta|-2)/2} \Big(|x-\bar{x}|^{\theta+\kappa} + |x-\bar{x}|^\theta \Big), \quad \forall |\beta| \leq m-1, \\
&\quad \forall t' \in [t_{n-i-1}, t_{n-i}).
\end{aligned}
$$

Here we have used inequalities (A.3.24), (A.3.25), Taylor's theorem, the triangle inequality and Lemma A.3.1. The constant C is independent of n by Lemmas A.3.1 and A.3.2.

Now that we have proved the claim for $l = 1$ and any $0 \leq i \leq n-1$, we will proceed by continuing the induction argument on l. Suppose that the following holds for a given $l \geq 1$ and $i = 1, 2, \ldots, n$,

$$\begin{aligned}
&\left|\partial_x^\beta v_{l,n}^{\bar{x}}(t', x) - \partial_x^\beta u_{l,m,n}^{\bar{x}}(t', x)\right| \\
&\leq i n^l C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa_1|=1} \cdots \sum_{|\kappa_l|=l} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2l)/2} \\
&\quad \times \Big(|x-\bar{x}|^{\theta+\kappa_1} + \ldots + |x-\bar{x}|^{\theta+\kappa_l} + |x-\bar{x}|^{\theta}\Big), \\
&\quad \forall |\beta| \leq m+1-2l, \quad \forall t' \in [t_{n-i}, t_{n-i+1}), \quad l \geq 1.
\end{aligned} \tag{A.3.26}$$

We will show that (A.3.26) holds for $l+1$ and any i satisfying $1 \leq i \leq n$. This requires several steps. First, assuming (A.3.26) holds, we have for all $i = 1, 2, \ldots, n$,

$$\begin{aligned}
&\Big| \int_{t'}^{t_{n-i+1}} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y \Big[\Gamma_0^{\bar{x}}(t', x; \tau, y) \\
&\quad \times \big(\sum_{j=1}^{l+1} \partial_y^\beta \mathcal{A}_j^{\bar{x}} v_{l+1-j,n}^{\bar{x}}(\tau, y) - \sum_{j=1}^{l+1} \partial_y^\beta \mathcal{A}_j^{\bar{x}} u_{l+1-j,m,n}^{\bar{x}}(\tau, y)\big)\Big]\Big| \\
&\leq i n^l C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa_1|=1} \cdots \sum_{|\kappa_{l+1}|=l+1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2(l+1))/2} \\
&\quad \times \Big(|x-\bar{x}|^{\theta+\kappa_1} + \ldots + |x-\bar{x}|^{\theta+\kappa_{l+1}} + |x-\bar{x}|^{\theta}\Big), \\
&\quad \forall |\beta| \leq m+1-2(l+1), \quad t' \in [t_{n-i}, t_{n-i+1}).
\end{aligned} \tag{A.3.27}$$

Inequality (A.3.27) is the result of applying the definition of $\mathcal{A}_j^{\bar{x}}$ in (2.3.5), inequality (A.3.26), the triangle inequality and Lemma A.3.1. Having obtained (A.3.27), we can estimate $|\partial_x^\beta v_{l+1,n}^{\bar{x}}(t', x) - \partial_x^\beta u_{l+1,m,n}^{\bar{x}}(t', x)|$. First, for $l+1$ and $i = 1$, we have

$$\begin{aligned}
&\left|\partial_x^\beta v_{l+1,n}^{\bar{x}}(t', x) - \partial_x^\beta u_{l+1,m,n}^{\bar{x}}(t', x)\right| \\
&= \Big| \int_{t'}^{T} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y \Big[\Gamma_0^{\bar{x}}(t', x; \tau, y) \big(\sum_{j=1}^{l+1} \partial_y^\beta \mathcal{A}_j^{\bar{x}} v_{l+1-j,n}^{\bar{x}}(\tau, y) - \sum_{j=1}^{l+1} \partial_y^\beta \mathcal{A}_j^{\bar{x}} u_{l+1-j,m,n}^{\bar{x}}(\tau, y)\big)\Big]\Big| \\
&\leq n^l C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa_1|=1} \cdots \sum_{|\kappa_{l+1}|=l+1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2(l+1))/2} \\
&\quad \times \Big(|x-\bar{x}|^{\theta+\kappa_1} + \ldots + |x-\bar{x}|^{\theta+\kappa_{l+1}} + |x-\bar{x}|^{\theta}\Big),
\end{aligned}$$

$$\forall |\beta| \le m+1-2(l+1), \quad t' \in [t_{n-1}, T).$$

Suppose that for $l+1$, an arbitrary i and for all $t' \in [t_{n-i}, t_{n-i+1})$, we have

$$\begin{aligned}
&\left|\partial_x^\beta v_{l+1,n}^{\bar{x}}(t',x) - \partial_x^\beta u_{l+1,m,n}^{\bar{x}}(t',x)\right| \\
&\le w(i,n,l+1)Cn^l \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa_1|=1} \cdots \sum_{|\kappa_{l+1}|=l+1} \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2(l+1))/2} \\
&\quad \times \Big(|x-\bar{x}|^{\theta+\kappa_1} + \ldots + |x-\bar{x}|^{\theta+\kappa_{l+1}} + |x-\bar{x}|^{\theta}\Big), \\
&\quad \forall |\beta| \le m+1-2(l+1), \quad \forall t' \in [t_{n-i}, t_{n-i+1}).
\end{aligned} \tag{A.3.28}$$

Next, applying (A.3.27) and (A.3.28) and proceeding as in the case $l=1$, we obtain for $l+1$ and $i+1$

$$\begin{aligned}
&\left|\partial_x^\beta v_{l+1,n}^{\bar{x}}(t',x) - \partial_x^\beta u_{l+1,m,n}^{\bar{x}}(t',x)\right| \\
&\le \Big| \int_{t'}^{t_{n-i}} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}y \Gamma_0^{\bar{x}}(t',x;\tau,y) \big(\sum_{j=1}^{l+1} \partial_y^\beta \mathcal{A}_j^{\bar{x}} v_{l+1-j,n}^{\bar{x}}(\tau,y) - \sum_{j=1}^{l+1} \partial_y^\beta \mathcal{A}_j^{\bar{x}} u_{l+1-j,m,n}^{\bar{x}}(\tau,y)\big)\Big| \\
&\quad + \underbrace{\Big| \int_{\mathbb{R}^d} \mathrm{d}y \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y)\big(v_{l+1,n}(t_{n-i},y) - v_{l+1,n}^{x'}(t_{n-i},y)\big)\Big]_{x'=y}\Big|}_{=0} \\
&\quad + \Big| \int_{\mathbb{R}^d} \mathrm{d}y \Big[\Gamma_0^{\bar{x}}(t',x;t_{n-i},y)\big(\partial_y^\beta v_{l+1,n}^{x'}(t_{n-i},y) - \partial_y^\beta u_{l+1,m,n}^{x'}(t_{n-i},y)\big)\Big]_{x'=y}\Big| \\
&\quad + \underbrace{\Big| \int_{\mathbb{R}^d} \mathrm{d}y \Big[\partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y)\big(u_{l+1,m,n}^{x'}(t_{n-i},y) - u_{l+1,m,n}(t_{n-i},y)\big)\Big]_{x'=y}\Big|}_{=0} \\
&\quad + \Big| \int_{\mathbb{R}^d} \mathrm{d}y \partial_x^\beta \Gamma_0^{\bar{x}}(t',x;t_{n-i},y)\big(u_{l+1,m,n}^{\mathtt{tc}}(t_{n-i},y) - \mathbf{T}_{m-2(l+1)}^{\bar{x}} u_{l+1,m,n}^{\mathtt{tc}}(t_{n-i},y)\big)\Big| \\
&\le (in^l + w(i,n,l+1)+1)C \\
&\quad \times \Big(\sum_{|\gamma+\theta|=m+1} \sum_{|\kappa_1|=1} \cdots \sum_{|\kappa_{l+1}|=l+1} \\
&\quad \times \Big(\frac{T-t}{n}\Big)^{(|\gamma|-|\beta|-2(l+1))/2} \Big(|x-\bar{x}|^{\theta+\kappa_1} + \ldots + |x-\bar{x}|^{\theta+\kappa_{l+1}} + |x-\bar{x}|^{\theta}\Big)\Big) \\
&\le (n^{l+1} + w(i,n,l+1)+1)C
\end{aligned}$$

$$
\begin{aligned}
&\times \Bigg(\sum_{|\gamma+\theta|=m+1} \sum_{|\kappa_1|=1} \cdots \sum_{|\kappa_{l+1}|=l+1} \\
&\times \left(\frac{T-t}{n}\right)^{(|\gamma|-|\beta|-2(l+1))/2} \Big(|x-\bar{x}|^{\theta+\kappa_1} + \ldots + |x-\bar{x}|^{\theta+\kappa_{l+1}} + |x-\bar{x}|^{\theta} \Big) \Bigg) \\
= {} & w(i+1,n,l+1) C \sum_{|\gamma+\theta|=m+1} \sum_{|\kappa_1|=1} \cdots \sum_{|\kappa_{l+1}|=l+1} \left(\frac{T-t}{n}\right)^{(|\gamma|-|\beta|-2(l+1))/2} \\
&\times \Big(|x-\bar{x}|^{\theta+\kappa_1} + \ldots + |x-\bar{x}|^{\theta+\kappa_{l+1}} + |x-\bar{x}|^{\theta} \Big), \\
& t' \in [t_{n-i-1}, t_{n-i}), \quad |\beta| \leq m - 2(l+1) + 1.
\end{aligned}
$$

Thus, we have established that (A.3.26) holds for any $i = 1, 2, 3, \ldots, n$. Setting $\bar{x} = x$ in (A.3.26), we have

$$
\begin{aligned}
&\left| \partial_x^\beta v_{l+1,n}(t',x) - \partial_x^\beta u_{l+1,m,n}(t',x) \right| \leq n^{l+2} C \left(\frac{T-t}{n}\right)^{(m+1-|\beta|-2(l+1))/2}, \\
&\quad \forall t' \in [t,T], \quad |\beta| \leq m - 2l - 1.
\end{aligned}
$$

As a result, (A.3.26) holds for any $l \geq 0$ and $1 \leq i \leq n$. Next, setting $\bar{x} = x$, the right-hand side of (A.3.26) becomes independent of x and summing (A.3.26) with respect to index l, where this index ranges from 0 to l, yields an upper bound for $|\partial_x^\beta \bar{v}_{l,n}(t',x) - \partial_x^\beta \bar{u}_{l,m,n}(t',x)|$. Then, we obtain

$$
\begin{aligned}
\sup_{x \in \mathbb{R}^d} \left| \partial_x^\beta \bar{v}_{l,n}(t',x) - \partial_x^\beta \bar{u}_{l,m,n}(t',x) \right| &\leq C n^{l+1} \left(\frac{T-t}{n}\right)^{(m+1-|\beta|-2l)/2}, \\
&\forall t' \in [t,T], \quad |\beta| \leq m + 1 - 2l.
\end{aligned}
$$

This establishes (A.3.9). □

Proof of Lemma A.3.7. From (2.3.9) and (2.3.10), respectively (A.3.5) and (A.3.6), it follows that

$$
\partial_x^\beta u_{l,m,n}^{(0,f),\bar{x}}(t_0,x;T) = \sum_{i=0}^{n-1} \int_{t_{n-i-1}}^{t_{n-i}} \mathrm{d}t' \partial_x^\beta u_{l,m,n}^{(f,0),\bar{x}}(t_0,x;t'), \tag{A.3.29}
$$

$$
\partial_x^\beta v_{l,n}^{(0,f),\bar{x}}(t_0,x;T) = \sum_{i=0}^{n-1} \int_{t_{n-i-1}}^{t_{n-i}} \mathrm{d}t' \partial_x^\beta v_{l,n}^{(f,0),\bar{x}}(t_0,x;t'). \tag{A.3.30}
$$

Equations (A.3.29) and (A.3.30) are proved in Lemma A.3.6. Using (A.3.29) and

(A.3.30), we compute

$$
\begin{aligned}
&\left|\partial_x^\beta u^{(0,f)}(t',x) - \partial_x^\beta \bar{v}_{l,n}^{(0,f)}(t',x;\tau)\right| \\
&= \Big| \sum_{i=0}^{n-1} \int\limits_{t_{n-i-1}\vee t'}^{t_{n-i}\vee t'} \mathrm{d}\tau (\partial_x^\beta u^{(f,0)}(t',x,\tau) - \partial_x^\beta \bar{v}_{l,n}^{(f,0)}(t',x;\tau))\Big| \\
&\leq C\left(\frac{T-t}{n}\right)^{(l+3-|\beta|)/2}, \quad \forall t' \in [t,T], \quad \forall |\beta| \leq m+1-2l, \qquad \text{(A.3.31)}
\end{aligned}
$$

and

$$
\begin{aligned}
&\left|\partial_x^\beta \bar{v}_{l,n}^{(0,f)}(t',x;\tau) - \partial_x^\beta \bar{u}_{l,m,n}^{(0,f)}(t',x)\right| \\
&= \Big| \sum_{i=0}^{n-1} \int\limits_{t_{n-i-1}\vee t'}^{t_{n-i}\vee t'} \mathrm{d}\tau (\partial_x^\beta \bar{v}_{l,n}^{(f,0)}(t',x;\tau) - \partial_x^\beta \bar{u}_{l,m,n}^{(f,0)}(t',x;\tau))\Big| \\
&\leq Cn^{l+1}\Big(\frac{T-t}{n}\Big)^{(m+1-|\beta|-2l)/2}, \quad \forall t' \in [t,T], \quad \forall |\beta| \leq m+1-2l. \qquad \text{(A.3.32)}
\end{aligned}
$$

Taking the supremum over x in (A.3.31) and (A.3.32), we obtain (A.3.16) and (A.3.17). □

A.4 Background Results

This section reports some useful results from Fujii and Takahashi (2016a).

Lemma A.4.1. *Assume that the driver f of Equation (3.2.1) satisfies Assumption 3.4.2 and $\exp(\lambda \exp(\beta T)|\phi(X_T)|)$ is in $\mathbb{L}^1(\Omega, \mathcal{F}_T, \mathbb{P})$. Then there exists a solution to the QEFBSDEJ (3.2.1), which satisfies*

$$
|Y_t| \leq \frac{1}{\lambda} \ln \mathbb{E}\Big[\exp\left(\lambda \exp(\beta(T-t))|\phi(X_T)| + \lambda \int\limits_t^T \exp(\beta(v-t)) l \,\mathrm{d}v\right)\Big|\mathcal{F}_t\Big].
$$

In particular, when $\|\phi(X_T)\|_\infty < \infty$, Y is essentially bounded with

$$
\|Y\|_{\mathbb{S}^\infty} \leq \exp(\beta T)(\|\phi(X_T)\|_\infty + Tl).
$$

Lemma A.4.2. *Assume that $\|\phi(X_T)\|_\infty < \infty$ and let Assumption 3.4.2 hold. If* *there exists a solution* (Y, Z, U) *to the QEFBSDEJ* (3.2.1)*, then* $Z \in \mathbb{H}^2_{\mathbf{BMO}}$ *and* $U \in \mathbb{J}^2_{\mathbf{BMO}}$ *(and hence* $U \in \mathbb{J}^\infty$*) and* $\|Z\|_{\mathbb{H}^2_{\mathbf{BMO}}} + \|U\|_{\mathbb{J}^2_{\mathbf{BMO}}}$ *is bounded above by a constant depending only on* $\big(\lambda, \beta, T, \|\phi(X_T)\|_\infty, l\big)$*.*

Now, consider two QEFBSDEJs, with $i \in \{1, 2\}$, satisfying Assumptions 3.4.2 and 3.4.3

$$\begin{aligned} \mathrm{d}Y^i_t &= -f^i(t, X^i_t, Y^i_t, Z^i_t, V^i_t)\,\mathrm{d}t \\ &\quad + Z^i_t\,\mathrm{d}W_t + \int_E U^i_t(e)\tilde{\mu}(\,\mathrm{d}t,\,\mathrm{d}e) \qquad\qquad Y^i_T = \phi^i(X^i_T). \end{aligned} \tag{A.4.1}$$

In addition, let $\delta\phi(X_T) := \phi^1(X^1_T) - \phi^2(X^2_T)$, $\delta X := X^1 - X^2$, $\delta Y := Y^1 - Y^2$, $\delta Z := Z^1 - Z^2$, $\delta V := V^1 - V^2$, $\delta f(t) := (f^1 - f^2)(t, X^1_t, Y^1_t, Z^1_t, V^1_t)$ and $\Theta_t := (X_t, Y_t, Z_t, V_t)$. Then we have the following lemma that can be proved with minor modifications on Fujii and Takahashi (2016a)

Lemma A.4.3 (A Priori Estimates)**.** *Suppose Assumptions 3.4.2-3.4.3 hold for the QEFBSDEJ* (A.4.1) *with* $i = 1, 2$*. If there exists a solution* (X^i, Y^i, Z^i, U^i)*, it satisfies the following inequalities*

$$\begin{aligned} &\|\delta Z\|_{\mathbb{H}^2_{\mathbf{BMO}}} + \|\delta U\|_{\mathbb{J}^2_{\mathbf{BMO}}} \\ &\quad \le C\Bigg(\|\delta Y\|_{\mathbb{S}^\infty} + \|\delta\phi(X_T)\|_\infty + \|\rho(\delta X)\|_{\mathbb{S}^\infty} + \sup_{\tau \in \mathcal{T}_0^T}\Bigg\|\mathbb{E}\Bigg[\int_\tau^T |\delta f(v)|\,\mathrm{d}v\Bigg|\mathcal{F}_\tau\Bigg]\Bigg\|_\infty\Bigg) \\ &\|(\delta Y, \delta Z, \delta U)\|^p_{\mathcal{K}^p[0,T]} \\ &\quad \le C'\Bigg(\mathbb{E}\Bigg[|\delta\phi(X_T)|^{pq^2} + \Bigg|\int_0^T \rho(\delta X_v)\,\mathrm{d}v\Bigg|^{pq^2} + \Bigg(\int_0^T |\delta f(v)|\,\mathrm{d}v\Bigg)^{pq^2}\Bigg]\Bigg)^{\frac{1}{q^2}} \\ &\quad \forall p \ge 2, \quad \forall q \ge q_*. \end{aligned}$$

C *and* $q_* \ge 1$ *are positive and rely on* $(K_M, \lambda, \beta, T, \|\phi(X_T)\|_\infty, l)$ *and the constant* M *is chosen such that* $\|Y^i\|_{\mathbb{S}^\infty} \le M$ *and* $\|U^i\|_{\mathbb{J}^\infty} \le M$ *for* $i = 1, 2$*.* C' *is a positive constant depending only on* $\big(p, q, K_M, \lambda, \beta, T, \|\phi(X_T)\|_\infty, l\big)$*.*

The result below follows from Lemma A.4.1, A.4.2 and A.4.3

Theorem A.4.4. *Suppose that the QEFBSDEJ* (3.2.1) *satisfies Assumptions 3.4.2 and 3.4.3. Then, there exists a solution* (X, Y, Z, U) *which is unique in the space* $\mathbb{S}^2_r[0,T] \times \mathbb{S}^\infty \times \mathbb{H}^2_{\mathbf{BMO}} \times \mathbb{J}^2_{\mathbf{BMO}}$.

A.5 Proofs of the Second Method

Proof of Lemma A.4.3. The proof of the first inequality proceeds as follows. Due to the universal bounds given in Lemmas A.4.1 and A.4.2, we can choose a constant M such that $\|Y^i\|_{\mathbb{S}_\infty[0,T]} \leq M$ and $\|U^i\|_{\mathbb{J}_\infty} \leq M$ for both $i = 1, 2$. Set a sequence of $\mathbb{F}$-stopping times as

$$\tau_n := \inf\left\{ t \geq 0 : \int_0^t |\delta Z_s|^2 \,\mathrm{d}s + \int_0^t \int_E |\delta U_s(e)|^2 N(\,\mathrm{d}s,\,\mathrm{d}e) \geq n \right\} \wedge T.$$

Then, for any $\tau \in \mathcal{T}_0^T$, we have

$$\begin{aligned}
&|\delta Y_\tau|^2 + \mathbb{E}\left[\int_\tau^T |\delta Z_s|^2 \,\mathrm{d}s + \int_\tau^T \int_E |\delta U_s(e)|^2 N(\,\mathrm{d}s,\,\mathrm{d}e) \middle| \mathcal{F}_\tau \right] \\
&= \mathbb{E}\left[|\delta\phi(X_T)|^2 + \int_\tau^T 2\delta Y_s \left(\delta f(s) + f^2(t, \Theta^1_s) - f^2(t, \Theta^2_s) \right) \middle| \mathcal{F}_\tau \right].
\end{aligned}$$

Taking $\sup_{\tau \in \mathcal{T}_0^T}$ for each term on the left-hand side gives

$$\begin{aligned}
&\|\delta Z\|^2_{\mathbb{H}^2_{\mathbf{BMO}}} + \|\delta U\|^2_{\mathbb{J}^2_{\mathbf{BMO}}} \\
&\leq 2\|\delta\phi(X_T)\|^2_\infty + 4\|\delta Y\|_{\mathbb{S}^\infty[0,T]} \\
&\times \sup_{\tau \in \mathcal{T}_0^T} \left\| \mathbb{E}\left[\int_\tau^T \left(|\delta f(s)| + \rho(\delta X_s) + K_M\big(|\delta Y_s| + \|\delta V_s\|_{\mathbb{L}^2(\nu)} + H_s|\delta Z_s| \big) \right) \mathrm{d}s \middle| \mathcal{F}_\tau \right] \right\|_\infty
\end{aligned}$$

where the process H is defined by $H_s := 1 + \sum_{i=1}^2 \left(|Z^i_s| + \|\delta V^i_s\|_{\mathbb{L}^2(\nu)} \right)$. We know that $H \in \mathbb{H}^2_{\mathbf{BMO}}$ with norm dominated by the universal bound given in Lemma A.4.1. We can see that

$$\sup_{\tau \in \mathcal{T}_0^T} \left\| \mathbb{E}\left[\int_\tau^T H_s|\delta Z_s| \,\mathrm{d}s \middle| \mathcal{F}_\tau \right] \right\|_\infty$$

$$\leq \sup_{\tau \in \mathcal{T}_0^T} \left\| \mathbb{E}\left[\int_\tau^T |H_s|^2 \, \mathrm{d}s \middle| \mathcal{F}_\tau \right]^{\frac{1}{2}} \right\|_\infty \left\| \mathbb{E}\left[\int_\tau^T |\delta Z_s|^2 \, \mathrm{d}s \middle| \mathcal{F}_\tau \right]^{\frac{1}{2}} \right\|_\infty$$
$$\leq \|H\|_{\mathbb{H}^2_{\mathbf{BMO}}} \|\delta Z\|_{\mathbb{H}^2_{\mathbf{BMO}}}.$$

Then, with arbitrary positive constant $\epsilon > 0$

$$\begin{aligned}
&\|\delta Z\|^2_{\mathbb{H}^2_{\mathbf{BMO}}} + \|\delta U\|^2_{\mathbb{J}^2_{\mathbf{BMO}}} \\
&\leq 2\|\delta\phi(X_T)\|^2_\infty + 2T\|\rho(\delta X)\|_{\mathbb{S}^\infty} + 2 \sup_{\tau \in \mathcal{T}_0^T} \left\| \mathbb{E}\left[\int_\tau^T |\delta f(s)| \, \mathrm{d}s \middle| \mathcal{F}_\tau \right] \right\|^2_\infty \\
&\quad + \|\delta Y\|^2_{\mathbb{S}^\infty[0,T]} \left(2 + 4K_M T + \frac{4K_M^2}{\epsilon} + \frac{4K_M^2}{\epsilon} \|H\|^2_{\mathbb{H}^2_{\mathbf{BMO}}} \right) \\
&\quad + \epsilon \left(\|\delta Z\|^2_{\mathbb{H}^2_{\mathbf{BMO}}} + \|\delta U\|^2_{\mathbb{J}^2_\infty} \right).
\end{aligned}$$

Given that $\|\delta U\|_{\mathbb{J}^2_\infty} \leq \|\delta U\|_{\mathbb{J}^2_{\mathbf{BMO}}}$ as shown in Fujii and Takahashi (2016a), choosing $\epsilon < 1$ yields the desired result. For the second inequality, define a d-dimensional $\mathbb{F}$-progressively measurable process $(b_s, s \in [0,T])$ by

$$b_s := \frac{f^2(s, X_s^1, Y_s^1, Z_s^1, V_s^1) - f^2(s, X_s^1, Y_s^1, Z_s^2, V_s^1)}{|\delta Z_s|^2} \mathbf{1}_{\delta Z_s \neq 0} \delta Z_s$$

and also the map $\tilde{f} : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{L}^2(E, \nu; \mathbb{R}^q) \to \mathbb{R}$ by

$$\tilde{f}(s, \tilde{x}, \tilde{y}, \tilde{V}) := \delta f(s) - f^2(s, \Theta_s^2) + f^2(s, \tilde{x} + X_s^2, \tilde{y} + Y_s^2, Z_s^2, \tilde{V} + V_s^2).$$

Then, $(\delta Y, \delta Z, \delta U)$ can be interpreted as the solution to the following BSDE

$$\begin{aligned}
\delta Y_t = \delta\phi(X_T) + \int_t^T \big(\tilde{f}(s, \delta X_s, \delta Y_s, \delta V_s) + b_s \cdot \delta Z_s \big) \, \mathrm{d}s \\
- \int_t^T \delta Z_s \, \mathrm{d}W_s - \int_t^T \int_E \delta U_s(e) \tilde{N}(\, \mathrm{d}s, \, \mathrm{d}e).
\end{aligned}$$

Because $b_s \leq K_M\big(1 + |Z_s^1| + |Z_s^2| + 2\|V_s^1\|_{\mathbb{L}^2(\nu)}\big)$, b process belongs to $\mathbb{H}^2_{\mathbf{BMO}}$. Furthermore, $\tilde{f}$ satisfies the linear growth property $|\tilde{f}(s, \tilde{x}, \tilde{y}, \tilde{V})| \leq |\delta f(s)| + \rho(\tilde{x}) + K_M\big(|\tilde{y}| + \|\tilde{V}\|_{\mathbb{L}^2(\nu)}\big)$. The rest of the proof follows from that of (Fujii and Takahashi, 2016a, Lemma 3.3). $\square$

Proof of Lemma 3.4.5. If a function g is globally Lipschitz continuous, a candidate for the approximating sequence $\{g_h\}_{h=1}^{+\infty}$ is $g_h(x) = h^q \int_{\mathbb{R}^q} \mathrm{d}y\, g(x-y)\phi(hy)$, where ϕ is a non-negative C^∞-function defined in $\mathbb{R}^q$ with support in the unit ball such that $\int_{\mathbb{R}^q} \mathrm{d}y\, \phi(y) = 1$. The function $g_h(x)$ is a.e. C^∞, globally Lipschitz continuous with the same Lipschitz constant as g and satisfies the relation

$$\big|g_h(x) - g(x)\big| \leq \frac{C}{h}.$$

We refer the readers to N'ZI et al. (2006) for more details. $\square$

Proof of Theorem 3.4.6. The proof follows from the a priori estimates in Lemma A.4.3 and (Bouchard and Elie, 2008, Lemma A.1). First, because of Equation (3.4.1) and (Bouchard and Elie, 2008, Lemma A.1), we know that $\lim_{h\to\infty} \|X^{(h)} - X\|_{\mathbb{S}^2_r[0,T]} = 0$. The fact that $\big\|\big(Y^{(h)} - Y, Z^{(h)} - Z, U^{(h)} - U\big)\big\|_{\mathcal{K}^2[0,T]} \to 0$ as $h \to \infty$ follows from the a priori estimates in Lemma A.4.2. $\square$

Proof of Theorem 3.4.7. Letting $\delta X_t^{(h,s)} := X_t^{(h,s)} - X_t$, we have for $w = 2$

$$\begin{aligned}
\|\delta X_t^{(h,s)}\|^w_{\mathbb{S}^w_r[0,T]} \leq\; & \mathbb{E}\Bigg[\bigg(\int_t^T \big|\mu_{h,s}\big(v, X_v\big) - \mu\big(v, X_v\big)\big|^2 \,\mathrm{d}v\bigg)^{\frac{w}{2}}\Bigg]^{\frac{1}{w}} \\
& + \mathbb{E}\Bigg[\bigg(\int_t^T \big|\sigma_{h,s}\big(v, X_v\big) - \sigma\big(v, X_v\big)\big|^2 \,\mathrm{d}v\bigg)^{\frac{w}{2}}\Bigg]^{\frac{1}{w}} \\
& + \mathbb{E}\Bigg[\int_t^T \int_E \big|\gamma_{h,s}\big(v, X_v\big) - \gamma\big(v, X_v\big)\big|^w \nu(\,\mathrm{d}e)\,\mathrm{d}v\Bigg]^{\frac{1}{w}}.
\end{aligned}$$

A bound for the second term of $\big|\sigma_{h,s}\big(v, X_v\big) - \sigma\big(v, X_v\big)\big|$ is obvious. Now consider the bounds of the first and third terms. By analogy, we only need to discuss $\big|\mu_{h,s}\big(v, X_v\big) - \mu\big(v, X_v\big)\big|$. We have

$$\begin{aligned}
&\big|\mu_{h,s}\big(v, X_v\big) - \mu\big(v, X_v\big)\big| \\
&\leq \big|\mu_{h,s}\big(v, X_v\big) - \mu_h\big(v, X_v\big)\big| + \big|\mu_h\big(v, X_v\big) - \mu\big(v, X_v\big)\big| \\
&\leq \frac{C}{h} + C(s+1)\mathbf{1}_{s \leq X_v \leq s+1} + C\big|X_v\big|\mathbf{1}_{X_v > s+1}.
\end{aligned}$$

This establishes $\lim_{(h,s)\to\infty} \left\|X^{(h,s)} - X\right\|_{\mathbb{S}^2_r[0,T]} = 0$ as $(h,s) \to \infty$. The fact that $\left\|\left(Y^{(h,s)} - Y, Z^{(h,s)} - Z, U^{(h,s)} - U\right)\right\|_{\mathcal{K}^2[0,T]} \to 0$ as $(h,s) \to \infty$ follows from the a priori estimates in Lemma A.4.2. □

Proof of Theorem 3.4.9. The proof follows from the a priori estimates in Lemma A.4.3 and (Bouchard and Elie, 2008, Lemma A.1). In the non-degeneracy transformation step, only the diffusion matrix is changed. Because of the smoothness and boundedness of the new diffusion matrix, we can deduce that $\|\sigma_{h,s,i}(t,\cdot) - \sigma_{h,s}(t,\cdot)\|_\infty \leq C(i) \to 0$ as $i \to \infty$. Therefore, we have shown $\lim_{i\to\infty} \left\|X^{(h,s,i)} - X^{(h,s)}\right\|_{\mathbb{S}^2_r[0,T]} = 0$ as $i \to \infty$. The fact that $\left\|\left(Y^{(h,s,i)} - Y, Z^{(h,s,i)} - Z, U^{(h,s,i)} - U\right)\right\|_{\mathcal{K}^2[0,T]} \to 0$ as $i \to \infty$ follows from the a priori estimates in Lemma A.4.2. □

Proof of Theorem 3.4.12. The proof of this theorem follows from (Menaldi and Garroni, 1992, Theorem 3.1, Chapter II). □

Proof of Theorem 3.4.13. The proof follows from the a priori estimates given in Halle (2010) and is similar to the arguments in El Karoui et al. (1997b) showing the convergence of Picard iteration. □

Proof of Theorem 3.4.14. The proof follows from the Lipschitz continuity property of the driver f with respect to variables (y, z, ψ), together with Theorem 3.4.15. □

Proof of Theorem 3.4.15. The proof follows from the definition of $v^x_{k,n}$ in Equation (3.3.8), the discussions (Liu and Li, 2000, Theorem 3.3), (Jum, 2015, Theorem 3.2) and (Jum, 2015, Theorem 3.3). We should notice that Equation (3.3.8) applies the law of iterated expectations. It is the same as the expectation of the Euler discretized process Y_K in (Liu and Li, 2000, Theorem 3.3). Therefore we have shown the case with zero-th order derivative. The error bounds for higher order derivatives are established by differentiating both sides of (Liu and Li, 2000, Equation 4.7) with respect to x_0 (it is straightforward to verify the differentiability with respect to x_0 given smooth coefficients and terminal condition). □

Before proving Theorem 3.4.16, we need the following lemma

Lemma A.5.1. *Let $u^x_{k,m,n}(t_j, x)$ be the intermediate solution at the j-th time discretization point. Then*

$$u^{\bar{x}}_{k,m,n}(t_j, x) = \sum_{|\xi|=0}^{m} G_{\xi,j}(t_j, \bar{x}) \frac{(x - x_0)^\xi}{\xi!}$$

and

$$u^{x}_{k,m,n}(t_j,x) - \mathbf{T}^{x_0}_{m} u^{x}_{k,m,n}(t_j,x) = \sum_{|\theta|=0}^{m} \sum_{|\xi+\theta|=m} \frac{(x-x_0)^{\xi}}{\xi!}\Big(G_{\xi,j}(t_j,x) - \mathbf{T}^{x_0}_{|\theta|} G_{\xi,j}(t_j,x)\Big).$$

$\{G_{\xi,j}\}_{\xi,j}$ *are some smooth functions, bounded and with bounded derivatives of all orders. Also, we have* $|\partial^{\beta}_{x} G_{\xi,j}(t,x)| \leq M$ *for* M *independent of* j *and* n *and* $|\beta| \leq m$.

Proof. The first claim of the lemma is obtained by expanding the intermediate solutions, collecting terms with the same powers $(x-x_0)^{|\xi|}$, $|\xi| = 0, 1, \cdots, m+1$ and using induction arguments. The second claim follows similarly from (Detemple et al., 2015, Lemma C.2) by replacing the estimates of the higher order derivatives of the coefficients with a constant M. $\square$

Proof of Theorem 3.4.16. We will only show the ϕ-part in the probabilistic representations (3.3.7) and (3.3.8) becauss of the analogy to the f-part. If we set an additional localization compact set of x_n as $\mathbb{B}(x_{n-1},\epsilon)$ for $\gamma(t,x_n,e)$, where $\mathbb{B}(x_{n-1},\epsilon)$ is the closed ball centered at x_{n-1} with radius ϵ and x_{n-1} is the backward variable in the transition density $\Gamma^{\bar{x}}_{0}(t',x_{n-1};T,x_n)$, then we have

$$\begin{aligned}
&\Big|v^{\bar{x}}_{k,n}(t',x_{n-1}) - u^{\bar{x}}_{k,m,n}(t',x_{n-1})\Big| \qquad\qquad \text{(A.5.1)}\\
&= \Big|\int_{\mathbb{R}^r} \mathrm{d}x_n \Gamma^{\bar{x}}_{0}(t',x_{n-1};T,x_n)\left(\phi(x_n) - \mathbf{T}^{x_0}_{m}\phi(x_n)\right)\Big|\\
&\leq \Big|\int_{\mathbb{R}^r} \mathrm{d}x_n \Gamma^{\bar{x}}_{0}(t',x_{n-1};T,x_n) \sum_{|\alpha|=m+1} \frac{\partial^{\alpha}_{z}\phi(z_{x_0,x_n,\alpha})}{\alpha!}(x_n-x_0)^{\alpha}\Big|\\
&\leq \sum_{|\alpha|=m+1}\int_{\mathbb{R}^r} \mathrm{d}x_n \Big|\Gamma^{\bar{x}}_{0}(t',x_{n-1};T,x_n)\frac{\partial^{\alpha}_{z}\phi(z_{x_0,x_n,\alpha})}{\alpha!}(x_n-x_0)^{\alpha}\Big|\\
&= \underbrace{\sum_{|\alpha|=m+1}\int_{\mathbb{B}(x_{n-1},\epsilon)} \mathrm{d}x_n \Big|\Gamma^{\bar{x}}_{0}(t',x_{n-1};T,x_n)\frac{\partial^{\alpha}_{z}\phi(z_{x_0,x_n,\alpha})}{\alpha!}(x_n-x_0)^{\alpha}\Big|}_{\text{In the closed-ball (1)}}\\
&\quad + \underbrace{\sum_{|\alpha|=m+1}\int_{\mathbb{R}^r-\mathbb{B}(x_{n-1},\epsilon)} \mathrm{d}x_n \Big|\Gamma^{\bar{x}}_{0}(t',x_{n-1};T,x_n)\frac{\partial^{\alpha}_{z}\phi(z_{x_0,x_n,\alpha})}{\alpha!}(x_n-x_0)^{\alpha}\Big|}_{\text{The Gaussian part (2)}}
\end{aligned}$$

$$\leq \underbrace{C \sum_{|\theta+\xi|=m+1} \epsilon^{|\theta|} |x_{n-1} - x_0|^{\xi}}_{(1)} + \underbrace{C \sum_{|\theta+\xi|=m+1} \left(\frac{T-t}{n}\right)^{\frac{|\theta|}{2}} |x_{n-1} - x_0|^{\xi}}_{(2)}.$$

Therefore

$$|\partial_x^{\beta} v_{k,n}^{x_0}(t', x_0) - \partial_x^{\beta} u_{k,m,n}^{x_0}(t', x_0)| \leq C\epsilon^{m+1-|\beta|} + C\left(\frac{T-t}{n}\right)^{\frac{m+1-|\beta|}{2}} \tag{A.5.2}$$

for $|\beta| \leq m$ according to Equation (A.5.1) and the definition of the partial derivatives of the intermediate solutions. Next, for $t' \in [t_{n-2}, t_{n-1})$, we have by setting $\bar{x} = x_{n-2}$

$$\begin{aligned}
&\left| v_{k,n}^{x_{n-2}}(t', x_{n-2}) - u_{k,m,n}^{x_{n-2}}(t', x_{n-2}) \right| \\
&\leq \int_{\mathbb{R}^r} \mathrm{d}x_{n-1} \Gamma_0^{x_{n-2}}(t', x_{n-2}; t_{n-1}, x_{n-1}) \left| v_{k,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) - u_{k,m,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) \right| \\
&\quad + \int_{\mathbb{R}^r} \mathrm{d}x_{n-1} \Gamma_0^{x_{n-2}}(t', x_{n-2}; t_{n-1}, x_{n-1}) \left| u_{k,m,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) - \mathbf{T}_m^{x_0} u_{k,m,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) \right| \\
&\leq \underbrace{\int_{\mathbb{R}^r} \mathrm{d}x_{n-1} \Gamma_0^{x_{n-2}}(t', x_{n-2}; t_{n-1}, x_{n-1}) C \sum_{|\theta+\xi|=m+1} \epsilon^{|\theta|} |x_{n-1} - x_0|^{\xi}}_{\text{(1)-From previous error bounds}} \\
&\quad + \underbrace{\int_{\mathbb{R}^r} \mathrm{d}x_{n-1} \Gamma_0^{x_{n-2}}(t', x_{n-2}; t_{n-1}, x_{n-1}) C \sum_{|\theta+\xi|=m+1} \left(\frac{T-t}{n}\right)^{\frac{|\theta|}{2}} |x_{n-1} - x_0|^{\xi}}_{\text{(2)-From previous error bounds}} \\
&\quad + \int_{\mathbb{B}(x_{n-2},\epsilon)} \mathrm{d}x_{n-1} \Gamma_0^{x_{n-2}}(t', x_{n-2}; t_{n-1}, x_{n-1}) \\
&\quad \times \underbrace{\left| u_{k,m,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) - \mathbf{T}_m^{x_0} u_{k,m,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) \right|}_{(3)} \\
&\quad + \int_{\mathbb{R}^r - \mathbb{B}(x_{n-2},\epsilon)} \mathrm{d}x_{n-1} \Gamma_0^{x_{n-2}}(t', x_{n-2}; t_{n-1}, x_{n-1}) \\
&\quad \times \underbrace{\left| u_{k,m,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) - \mathbf{T}_m^{x_0} u_{k,m,n}^{x_{n-1}}(t_{n-1}, x_{n-1}) \right|}_{(3)}
\end{aligned}$$

$$\leq \underbrace{C \sum_{|\theta+\xi|=m+1} \epsilon^{|\theta|} |x_{n-2} - x_0|^{\xi}}_{(1)} + \underbrace{C \sum_{|\theta+\xi|=m+1} \left(\frac{T-t}{n}\right)^{\frac{|\theta|}{2}} |x_{n-2} - x_0|^{\xi}}_{(2)}$$

$$+ \underbrace{C \sum_{|\theta+\xi|=m+1} \epsilon^{|\theta|} |x_{n-2} - x_0|^{\xi} + C \sum_{|\theta+\xi|=m+1} \left(\frac{T-t}{n}\right)^{\frac{|\theta|}{2}} |x_{n-2} - x_0|^{\xi}}_{(3)}$$

$$= 2C \sum_{|\theta+\xi|=m+1} \epsilon^{|\theta|} |x_{n-2} - x_0|^{\xi} + 2C \sum_{|\theta+\xi|=m+1} \left(\frac{T-t}{n}\right)^{\frac{|\theta|}{2}} |x_{n-2} - x_0|^{\xi}.$$

Suppose that for $t' \in [t_{n-i-1}, t_{n-i})$, we have

$$\begin{aligned}
&\left| v_{k,n}^{x_{n-i-1}}(t', x_{n-i-1}) - u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) \right| \\
&\leq (i+1)C \sum_{|\theta+\xi|=m+1} \epsilon^{|\theta|} |x_{n-i-1} - x_0|^{\xi} + (i+1)C \sum_{|\theta+\xi|=m+1} \left(\frac{T-t}{n}\right)^{\frac{|\theta|}{2}} |x_{n-i-1} - x_0|^{\xi}
\end{aligned}$$

then when n is sufficiently large, we have for $t' \in [t_{n-i-2}, t_{n-i-1})$

$$\begin{aligned}
&\left| v_{k,n}^{x_{n-i-2}}(t', x_{n-i-2}) - u_{k,m,n}^{x_{n-i-2}}(t', x_{n-i-2}) \right| \\
&\leq \int_{\mathbb{R}^r} \mathrm{d}x_{n-i-1} \, \Gamma_0^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2}; t_{n-i-1}, x_{n-i-1}) \\
&\quad \left| v_{k,n}^{x_{n-i-1}}(t', x_{n-i-1}) - u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) \right| \\
&\quad + \int_{\mathbb{R}^r} \mathrm{d}x_{n-i-1} \, \Gamma_0^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2}; t_{n-i-1}, x_{n-i-1}) \\
&\quad \left| u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) - \mathbf{T}_m^{x_0} u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) \right| \\
&\leq \int_{\mathbb{R}^r} \mathrm{d}x_{n-i-1} \, \Gamma_0^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2}; t_{n-i-1}, x_{n-i-1}) \\
&\quad \underbrace{\left| v_{k,n}^{x_{n-i-1}}(t', x_{n-i-1}) - u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) \right|}_{\text{error bound from the previous time step}} \\
&\quad + \int_{\mathbb{B}(x_{n-i-2}, \epsilon)} \mathrm{d}x_{n-i-1} \, \Gamma_0^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2}; t_{n-i-1}, x_{n-i-1})
\end{aligned}$$

$$
\times \underbrace{\left| u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) - \mathbf{T}_m^{x_0} u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) \right|}_{\text{jump error bound}}
$$

$$
+ \int\limits_{\mathbb{R}^r - \mathbb{B}(x_{n-i-2},\epsilon)} \mathrm{d}x_{n-i-1}\, \Gamma_0^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2}; t_{n-i-1}, x_{n-i-1})
$$

$$
\times \underbrace{\left| u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) - \mathbf{T}_m^{x_0} u_{k,m,n}^{x_{n-i-1}}(t', x_{n-i-1}) \right|}_{\text{Gaussian error bound}}
$$

$$
\leq (i+2)C \sum_{|\theta+\xi|=m+1} \epsilon^{|\theta|} |x_{n-i-2} - x_0|^{\xi}
$$

$$
+ (i+2)C \sum_{|\theta+\xi|=m+1} \left(\frac{T-t}{n} \right)^{\frac{|\theta|}{2}} |x_{n-i-2} - x_0|^{\xi}
$$

with the intermediate coefficients canceled out by the additional ϵ and $\frac{T-t}{n}$ when ϵ is small enough and n is large enough. Therefore, by induction and setting $x = x_0$ in the error bound, we have for $t' \in [t_0, t_1)$

$$
\left| \partial_x^{\beta} v_{k,n}^{x_0}(t', x_0) - \partial_x^{\beta} u_{k,m,n}^{x_0}(t', x_0) \right| \leq nC(\beta)\epsilon^{m+1-|\beta|} + nC(\beta) \left(\frac{T-t}{n} \right)^{\frac{m+1-|\beta|}{2}}.
$$

The error bounds of the higher order derivatives follow from Equation (A.5.2). The last step is to set $\epsilon = n^{\frac{1}{4}} \frac{C}{\sqrt{n}}$. The error bound announced follows. The interpretation of ϵ is the number $(n^{\frac{1}{4}})$ of standard deviations $(\frac{C}{\sqrt{n}})$ of the distribution given by $\Gamma_0^{x_i}(t_i, x_i; t_{i+1}, x_{i+1})$, for $0 \leq i \leq n$. Moreover, let us analyze the additional error introduced by the localization of $\mathbb{B}(x_i, \epsilon)$ compared to the solution before this localization. Denote by $\Gamma_0^{\bar{x}}$ as the approximate transition density before localization and $\Gamma_{0,\epsilon}^{\bar{x}}$ after. Then, we can decompose the errors

$$
\begin{aligned}
&\left| v_{k,n}^{x_{n-i-2}}(t', x_{n-i-2}) - u_{k,m,n}^{x_{n-i-2}}(t', x_{n-i-2}) \right| \\
&\leq \left| v_{k,n}^{x_{n-i-2}}(t', x_{n-i-2}) - v_{k,n,\epsilon}^{x_{n-i-2}}(t', x_{n-i-2}) \right| \\
&\quad + \left| v_{k,n,\epsilon}^{x_{n-i-2}}(t', x_{n-i-2}) - u_{k,m,n,\epsilon}^{x_{n-i-2}}(t', x_{n-i-2}) \right| \\
&\quad + \left| u_{k,m,n,\epsilon}^{x_{n-i-2}}(t', x_{n-i-2}) - u_{k,m,n}^{x_{n-i-2}}(t', x_{n-i-2}) \right|.
\end{aligned} \tag{A.5.3}
$$

Here subscript ϵ is the solution obtained after localizing the coefficients in $\mathbb{B}(x_{n-i-3}, \epsilon)$. The bound for the second term on the right-hand side of Equation (A.5.3) is obtained above, while the first term on the right-hand side is bounded above by a constant $(i+1)C_\epsilon$ while $C_\epsilon \to 0$ as $\epsilon = \frac{C}{n^{\frac{1}{4}}} \to 0$. $\lim_{n\to\infty} nC_\epsilon = 0$ can be seen from the fact that, for both Lévy density $\Gamma_0^{\bar{x}}$ or Gaussian density $\Gamma_0^{\bar{x}}$, the integral can be represented by

$$\begin{aligned}
&\Bigg|\int\limits_{\mathbb{R}^r} \mathrm{d}x_{n-i-2}\Gamma_0^{x_{n-i-3}}(t_{n-i-3}, x_{n-i-3}; t_{n-i-2}, x_{n-i-2})v_{k,n}^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2}) \\
&\quad - \Gamma_{0,\epsilon}^{x_{n-i-3}}(t_{n-i-3}, x_{n-i-3}; t_{n-i-2}, x_{n-i-2})v_{k,n,\epsilon}^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2})\Bigg| \\
&\leq \int\limits_{\mathbb{R}^r} \mathrm{d}x_{n-i-2}|\Gamma_0^{x_{n-i-3}}(t_{n-i-3}, x_{n-i-3}; t_{n-i-2}, x_{n-i-2}) \\
&\quad \underbrace{-\Gamma_{0,\epsilon}^{x_{n-i-3}}(t_{n-i-3}, x_{n-i-3}; t_{n-i-2}, x_{n-i-2})\big|\big|v_{k,n}^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2})\big|}_{(1)} \\
&\quad + \int\limits_{\mathbb{R}^r} \mathrm{d}x_{n-i-2}|\Gamma_{0,\epsilon}^{x_{n-i-3}}(t_{n-i-3}, x_{n-i-3}; t_{n-i-2}, x_{n-i-2}) \\
&\quad \underbrace{\times|v_{k,n}^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2}) - v_{k,n,\epsilon}^{x_{n-i-2}}(t_{n-i-2}, x_{n-i-2})|}_{(2)} \\
&\leq \underbrace{C_\epsilon^{i+2}}_{(1)} + \underbrace{M_\epsilon^{i+1}}_{(2)}
\end{aligned}$$

where

$$\begin{aligned}
M_\epsilon^{i+1} &:= \sum_{k=1}^{i+1} C_\epsilon^k \\
C_\epsilon^{i+2} &:= M \int\limits_{\mathbb{R}^r - \mathbb{B}(x_{n-i-3},\epsilon)} \mathrm{d}x_{n-i-2}|\Gamma_0^{x_{n-i-3}}(t_{n-i-3}, x_{n-i-3}; t_{n-i-2}, x_{n-i-2}) \\
&\quad - \Gamma_{0,\epsilon}^{x_{n-i-3}}(t_{n-i-3}, x_{n-i-3}; t_{n-i-2}, x_{n-i-2})|.
\end{aligned}$$

By Assumption 3.4.3 and the tail property of the Gaussian and Lévy transition densities, one can show that C_ϵ^{i+2} is exponentially decaying in n as $n \to \infty$ (for Lévy density we can see this from the transition density expansion formula in Filipović et al. (2013) and Assumption 3.4.3). The third term in the error decomposition (A.5.3) is analogous and can be seen from

$$
\begin{aligned}
&\left|u^x_{k,m,n}(t,x) - u^x_{k,m,n,\epsilon}(t,x)\right| \qquad \text{(A.5.4)}\\
&= \underbrace{\int\limits_{\mathbb{R}^r} \mathrm{d}y \left(\Gamma^x_0(t,x;\tau,y)\mathbf{T}^{x_0}_m u^y_{k,m,n}(\tau,y) - \Gamma^x_0(t,x;\tau,y)u^y_{k,m,n}(\tau,y)\right)}_{C\left(\frac{T-t}{n}\right)\text{ multiplied by a polynomial of }(x-x_0).}\\
&\quad + \underbrace{\int\limits_{\mathbb{B}(x,\epsilon)} \mathrm{d}y \left(\Gamma^x_0(t,x;\tau,y)u^y_{k,m,n}(\tau,y) - \Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n}(\tau,y)\right)}_{=0}\\
&\quad + \underbrace{\int\limits_{\mathbb{R}^r-\mathbb{B}(x,\epsilon)} \mathrm{d}y \left(\Gamma^x_0(t,x;\tau,y)u^y_{k,m,n}(\tau,y) - \Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n}(\tau,y)\right)}_{\text{exponential decay}}\\
&\quad + \underbrace{\int\limits_{\mathbb{B}(x,\epsilon)} \mathrm{d}y \left(\Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n}(\tau,y) - \Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n,\epsilon}(\tau,y)\right)}_{\text{error from the previous time step}}\\
&\quad + \underbrace{\int\limits_{\mathbb{R}^r-\mathbb{B}(x,\epsilon)} \mathrm{d}y \left(\Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n}(\tau,y) - \Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n,\epsilon}(\tau,y)\right)}_{\text{error from the previous time step multiplied by an exponentially decaying function of }n}\\
&\quad + \underbrace{\int\limits_{\mathbb{B}(x,\epsilon)} \mathrm{d}y \left(\Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n,\epsilon}(\tau,y) - \Gamma^x_{0,\epsilon}(t,x;\tau,y)\mathbf{T}^{x_0}_m u^y_{k,m,n,\epsilon}(\tau,y)\right)}_{C\epsilon^{m+1}\text{ multiplied by a polynomial of }(x-x_0).}\\
&\quad + \underbrace{\int\limits_{\mathbb{R}^r-\mathbb{B}(x,\epsilon)} \mathrm{d}y \left(\Gamma^x_{0,\epsilon}(t,x;\tau,y)u^y_{k,m,n,\epsilon}(\tau,y) - \Gamma^x_{0,\epsilon}(t,x;\tau,y)\mathbf{T}^{x_0}_m u^y_{k,m,n,\epsilon}(\tau,y)\right)}_{C\left(\frac{\tau-t}{n}\right)^{\frac{m+1}{2}}\text{ multiplied by a polynomial of }(x-x_0).}
\end{aligned}
$$

and

$$\begin{aligned}
&\int_{\mathbb{R}^r} \mathrm{d}y \big|\Gamma_0^x(t,x;\tau,y) - \Gamma_{0,\epsilon}^x(t,x;\tau,y)\big| |y-x|^\beta \\
&= 2 \int_{\mathbb{R}^r - \mathbb{B}(x,\epsilon)} \mathrm{d}y \Gamma_0^x(t,x;\tau,y)|y-x|^\beta \\
&\le 2C\mathrm{e}^{-\epsilon^{-\alpha}}
\end{aligned}$$

for some positive α. In Equation (A.5.4), the second to the third terms on the right-hand side are either exponential decaying in n, e.g., $\exp(-n^\alpha)$ with $\alpha > 0$ or of higher polynomial order, e.g., $n^{-\theta}$ with $\theta \ge 2$. Therefore, the convergence and the error bound follow. $\square$

Remark A.5.2. *The striking feature of the localization of X is that, for convergence, we do not ask $\epsilon \to \infty$ due to the time discretization. At every time step, the variance of the distribution of X is of order $O\left(\frac{1}{n}\right)$. This permits a localization such that when $\epsilon \to 0$ and $n \to \infty$, ϵ will cover more and more standard deviations of the distribution and stretch to the tails.*

Proof of Theorem 3.4.17. We will only show the bound of C_ζ. Suppose that at zeroth iteration, we obtain $u_{0,m,n}^x(t,x)$, which we plug into $f(t,x,y,z,\psi)$. Localize f on $x \in \mathbb{B}(x_0,\zeta)$. The resulting error is

$$\int_{\mathbb{R}^r} \mathrm{d}y \ \Gamma(t,x;T,y)|f(t,y) - f^\zeta(t,y)| = C_\zeta \to 0 \qquad \zeta \to \infty$$

because of the boundedness of f. Because at every Picard iteration we localize the function f, the error C_ζ accumulates k times. $\square$

A.6 Proofs of the Econometrics of SDE with Jumps

Proof of Theorem 4.5.3. First, observe that

$$\begin{aligned}
&|\Gamma_{h,s,i,J,m,n}(t,x;T,y) - \Gamma_{h,s,i}(t,x;T,y)| \\
&\le |\Gamma_{h,s,i}(t,x;T,y) - \Gamma_{h,s,i,J}(t,x;T,y)| \\
&\quad + |\Gamma_{h,s,i,J}(t,x;T,y) - \Gamma_{h,s,i,J,m,n}(t,x;T,y)| \\
&\le C_{h,s,i,J} + C_{h,s,i,J,m,n}.
\end{aligned}$$

Both $C_{h,s,i,J}$ and $C_{h,s,i,J,m,n}$ are independent of (x, y). The first constant $C_{h,s,i,J}$ is due to the fact that both $\Gamma_{h,s,i}(t, x; T, y)$ and $\Gamma_{h,s,i,J}(t, x; T, y)$ are uniformly bounded in (x, y) and the second constant $C_{h,s,i,J,m,n}$ can be obtained directly from Theorem 4.5.2 and the definition of $\Gamma_{h,s,i,J}(t, x; T, y)$ and $\Gamma_{h,s,i,J,m,n}(t, x; T, y)$. □

Proof of Theorem 4.5.5. The proof is analogous to that of (Yu, 2007, Theorem 3, Theorem 4) with minor modifications. It is obvious the true MLE sets the score to 0. Therefore, by denoting θ_0 as the true values of the parameters, we have

$$\partial_\theta l_N(\theta_0) = -\partial_\theta^2 l_N(\tilde{\theta})(\widehat{\theta}_{0,N} - \theta_0)$$

for some $\tilde{\theta}$ in between θ_0 and $\widehat{\theta}_{0,N}$. In addition, we have

$$\begin{aligned}
&I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{0,N} - \theta_0) \\
&= I_N(\theta_0)^{\frac{1}{2}}(-\partial_\theta^2 l_N(\tilde{\theta}))^{-1}\partial_\theta l_N(\theta_0) \\
&= -(I_N(\theta_0)^{\frac{1}{2}}\partial_\theta^2 l_N(\tilde{\theta})I_N(\theta_0)^{\frac{1}{2}})^{-1}I_N(\theta_0)^{\frac{1}{2}}\partial_\theta l_N(\theta_0) \\
&= -(I_N(\theta_0)^{\frac{1}{2}}\partial_\theta^2 l_N(\theta_0)I_N(\theta_0)^{\frac{1}{2}})^{-1}(I_N(\theta_0)^{\frac{1}{2}}\partial_\theta l_N(\theta_0)) + \mathbf{O}_r(1) \\
&= G_N^{-1}(\theta_0)S_N(\theta_0) + \mathbf{O}_r(1)
\end{aligned} \tag{A.6.1}$$

where $G_N^{-1}(\theta) := -I_N(\theta)^{-\frac{1}{2}}\partial_\theta^2 l_N(\theta)I_N(\theta)^{-\frac{1}{2}}$ and $S_N(\theta) := I_N(\theta)^{-\frac{1}{2}}\partial_\theta l_N(\theta)$. The Equation (A.6.1) holds uniformly for $\theta_0 \in \Theta$ and $h < \bar{h}$, where h is the time gap between each observation and $\bar{h}$ is specified in the Assumption 6 of Yu (2007). Because of Assumption 6 in (Yu, 2007, Appendix A), we know that $\widehat{\theta}_{0,N}$ is in the $(I_N(\theta_0)^{-\frac{1}{2}}$-neighborhood of θ_0, by repeating the previous argument, we have

$$\begin{aligned}
&I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{0,N} - \theta_0) \\
&= I_N(\theta_0)^{\frac{1}{2}}(-\partial_\theta^2 l_N(\tilde{\theta}))^{-1}\partial_\theta l_N(\theta_0) \\
&= -(I_N(\theta_0)^{\frac{1}{2}}\partial_\theta^2 l_N(\tilde{\theta})I_N(\theta_0)^{\frac{1}{2}})^{-1}I_N(\theta_0)^{\frac{1}{2}}\partial_\theta l_N(\theta_0) \\
&= -(I_N(\theta_0)^{\frac{1}{2}}\partial_\theta^2 l_N(\theta_0)I_N(\theta_0)^{\frac{1}{2}})^{-1}(I_N(\theta_0)^{\frac{1}{2}}\partial_\theta l_N(\theta_0)) + \mathbf{o}_r(1) \\
&= G_N^{-1}(\theta_0)S_N(\theta_0) + \mathbf{o}_r(1)
\end{aligned}$$

The asymptotic distribution of $\widehat{\theta}_{0,N}$ now follows from the limiting distributions of G_N and S_N. In the stationary case, $G_N(\theta_0)$ is non-random and it follows that

$$(N \times i_N(\theta_0))^{\frac{1}{2}}(\widehat{\theta}_{0,N} - \theta_0) = \mathcal{N}(\mathbf{0}, \mathbf{I}_{r\times r}) + \mathbf{o}_r(1).$$

Next, we investigate the stochastic difference between $\widehat{\theta}_{h,s,i,J,m,n,N}$ and $\widehat{\theta}_{0,N}$. We have

$$\begin{aligned}
&\partial_\theta l_N(\widehat{\theta}_{h,s,i,J,m,n,N}) - \partial_\theta l_{N,h,s,i,J,m,n}(\widehat{\theta}_{h,s,i,J,m,n,N}) \\
&= \partial_\theta l_N(\widehat{\theta}_{h,s,i,J,m,n,N}) \\
&= \partial_\theta l_N(\widehat{\theta}_{0,N}) + \partial_\theta^2 l_N(\bar{\theta})(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_{0,N}) \\
&= \partial_\theta^2 l_N(\bar{\theta})(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_{0,N})
\end{aligned}$$

where $\bar{\theta}$ lies in between $\widehat{\theta}_{h,s,i,J,m,n,N}$ and $\widehat{\theta}_{0,N}$. Therefore

$$\begin{aligned}
&I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_{0,N}) \\
&= -(I_N(\theta_0)^{\frac{1}{2}}\partial_\theta^2 l_N(\bar{\theta}) I_N(\theta_0)^{\frac{1}{2}})^{-1} I_N(\theta_0)^{-\frac{1}{2}} \\
&\quad \times (\partial_\theta l_N(\widehat{\theta}_{h,s,i,J,m,n,N}) - \partial_\theta l_{N,h,s,i,J,m,n}(\widehat{\theta}_{h,s,i,J,m,n,N})) \\
&= (G_N(\theta_0)^{-1} + \mathbf{O}_r(1)) I_N(\theta_0)^{-\frac{1}{2}}(\partial_\theta l_N(\widehat{\theta}_{h,s,i,J,m,n,N}) \\
&\quad - \partial_\theta l_{N,h,s,i,J,m,n}(\widehat{\theta}_{h,s,i,J,m,n,N})).
\end{aligned}$$

We will show later that there exists a sequence of $(h, s, i, J, n) \to \infty$, so that for any $h < \bar{h}$, we have $\partial_\theta l_{N,h,s,i,J,m,n}(\cdot) = (1 + \mathbf{o}_r(1))\partial_\theta l_N(\cdot)$ uniformly for any $\theta_0 \in \Theta$. Assume that this holds now, we have, for some $\tilde{\theta}$ between θ_0 and $\widehat{\theta}_{h,s,i,J,m,n,N}$

$$\begin{aligned}
&I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_{0,N}) \\
&= \mathbf{o}_r(1) I_N(\theta_0)^{-\frac{1}{2}} \partial_\theta l_N(\widehat{\theta}_{h,s,i,J,m,n,N}) \\
&= \mathbf{o}_r(1) I_N(\theta_0)^{-\frac{1}{2}} (\partial_\theta l_N(\theta_0) + \partial_\theta^2 l_N(\tilde{\theta})(\widehat{\theta}_{h,s,i,J,m,n,N} - \theta_0)) \\
&= \mathbf{o}_r(1)(I_N(\theta_0)^{-\frac{1}{2}} \partial_\theta l_N(\theta_0) + \mathbf{O}_r(1) I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{h,s,i,J,m,n,N} - \theta_0) \\
&= \mathbf{o}_r(1)(S + \mathbf{O}_r(1) I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_{0,N} + \widehat{\theta}_{0,N} - \theta_0))
\end{aligned}$$

where S is the probability limit of $I_N(\theta_0)^{-\frac{1}{2}}\partial_\theta l_N(\theta_0)$ under $\mathbb{P}_{\theta_0}$ as $N \to \infty$. Therefore, the theorem is proved. The last step is to show that as $(h, s, i, J, n) \to \infty$, for any $h < \bar{h}$, we have $\partial_\theta l_{N,h,s,i,J,m,n}(\cdot) = (1 + \mathbf{o}_r(1))\partial_\theta l_N(\cdot)$ uniformly for any $\theta_0 \in \Theta$. First, we have uniformly for all $\theta \in \Theta$, $\Gamma_{h,s,i,J,m,n}(t, x; t + h, y|\theta)$ is an approximate of the true transition density $\Gamma(t, x; t + h, y|\theta)$ with the relative error denoted by $\epsilon^{(1)}_{h,s,i,J,m,n}(t, x; t + h, y|\theta)$. That is to say, we have that $\Gamma_{h,s,i,J,m,n}(t, x; t + h, y|\theta) := \Gamma(t, x; t+h, y|\theta)(1+\epsilon^{(1)}_{h,s,i,J,m,n}(t, x; t+h, y|\theta))$. Similarly, $\partial_\theta \Gamma_{h,s,i,J,m,n}(t, x; t+h, y|\theta) := \partial_\theta \Gamma(t, x; t+h, y|\theta)(1+\epsilon^{(2)}_{h,s,i,J,m,n}(t, x; t+h, y|\theta))$. Now we bound the functions $\epsilon^{(1)}$ and $\epsilon^{(2)}$. Let $q_{h,s,i,J,m,n,t}$ and $\xi_{h,s,i,J,m,n,t}$ be two sequences of positive numbers converging to 0. Because of the Assumption 7 of (Yu, 2007, Appendix A), we can ask that when $X_t \in U_{h,s,i,J,m,n}$, where $U_{h,s,i,J,m,n}$ is a compact subset of $\mathbb{R}^r$, we have $\mathbb{P}(X_t \in U_{h,s,i,J,m,n}) = 1 - q_{h,s,i,J,m,n,t}$ for all $t \leq T$. Then, we know that we can ask $|\epsilon^{(1)}|$ and $|\epsilon^{(2)}|$, for observations inside this compact set, to be bounded by $\xi_{h,s,i,J,m,n,t}$ uniformly for all $\theta \in \Theta$. This argument shows that $\partial_\theta l_{N,h,s,i,J,m,n}(\cdot) = (1 + \mathbf{o}_r(1))\partial_\theta l_N(\cdot)$. The distribution relationship $\sqrt{N i_N(\theta_0)}(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_0) = \mathcal{N}(\mathbf{0}, \mathbf{I}_{r\times r}) + \mathbf{o}_r(1)$ can be seen from the equation $I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{h,s,i,J,m,n,N} - \theta_0) = I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{h,s,i,J,m,n,N} - \widehat{\theta}_{0,N}) + I_N(\theta_0)^{\frac{1}{2}}(\widehat{\theta}_{0,N} - \theta_0)$. $\square$

References

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed form approximation approach. *Econometrica*, 70:223–262.

Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *Annals of Statistics*, 36:906–937.

Antonelli, F. (1993). Backward-forward stochastic differential equations. *Annals of Applied Probability*, 3:777–793.

Bichuch, M., Capponi, A., and Sturm, S. (2015a). Arbitrage-free pricing of XVA part I: Framework and explicit examples. *Preprint*.

Bichuch, M., Capponi, A., and Sturm, S. (2015b). Arbitrage-free pricing of XVA part II: PDE representation and numerical analysis. *Preprint*.

Bick, B., Kraft, H., and Munk, C. (2013). Solving constrained consumption investment problems by simulation of artificial market strategies. *Management Science*, 59(2):485–503.

Bismut, J. (1973). Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44:384–404.

Bouchard, B. and Elie, R. (2008). Discrete-time approximation of decoupled forward-backward sde with jumps. *Stochastic Processes and Their Applications*, 118:53–75.

Briand, P. and Hu, Y. (2006). BSDE with quadratic growth and unbounded terminal value. *Probability Theory and Related Fields*, 136:604–618.

Briand, P. and Hu, Y. (2008). Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probability Theory and Related Fields*, 141:543–567.

Briand, P. and Labart, C. (2012). Simulation of BSDEs by wiener chaos expansion. *The Annals of Applied Probability*, 24(3):1129–1171.

Carmona, R. (2009). *Indifference Pricing*. Princeton University Press.

Chakraborty, A. K. and Chatterjee, M. (2013). On multivariate folded normal distribution. *The Indian Journal of Statistics*, 75B(1):1–15.

Choi, S. (2013). Closed-form likelihood expansions for inhomogeneous diffusions. *Journal of Econometrics*, 174:45–65.

Choi, S. (2015). Explicit form of approximate transition probability density functions of diffusion processes. *Journal of Econometrics*, 187:57–73.

Cvitanic, J. and Ma, J. (1996). Hedging options for a large investor and forward-backward SDEs. *Annals of Applied Probability*, 6(2):370–398.

Delong, L. (2013). *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*. Springer.

Detemple, J., Lorig, M., Rindisbacher, M., Sturm, S., and Zhang, L. (2015). Analytical expansion to forward-backward stochastic differential equations: The original scheme. *Working Paper*.

Detemple, J. and Zapatero, F. (1991). Asset prices in an exchange economy with habit formation. *Econometrica*, 59(6):1633–1657.

Detemple, J. and Zapatero, F. (1992). Optimal consumption-portfolio policies with habit formation. *Mathematical Finance*, 2(4):251–274.

Duffie, D. and Epstein, L. (1992). Asset pricing with stochastic differential utility. *Review of Financial Studies*, 5(3):411–436.

El Karoui, N., Hamadène, S., and Matoussi, A. (2008). Backward stochastic differential equations and applications. *Paris-Princeton Lecture Notes on Mathematical Finance*.

El Karoui, N., Kapoudjian, C., Pardoux, E., and Quenez, M. (1997a). Reflected solutions of backward SDEs and related obstacle problems for PDEs. *Annals of Probability*, 25(2):702–737.

El Karoui, N., Peng, S., and Quenez, M. C. (1997b). Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71.

El Karoui, N. and Quenez, M. (1997). Non-linear pricing theory and BSDEs. *Financial Mathematics*, pages 191–246.

Filipović, D., Mayerhofer, E., and Schneider, P. (2013). Density approximations for multivariate affine jump-diffusion processes. *Journal of Econometrics*, 176(2):93–111.

Fouque, J. P., Papanicolaou, G., Sircar, R., and Solna, K. (2011). *Multiscale Stochastic Volatility for Equity, Interest Rate and Credit Derivatives*. Cambridge University Press.

Fujii, M. (2016). A polynomial scheme of asymptotic expansion for backward SDEs and option pricing. *Quantitative Finance*, 16(3):427–445.

Fujii, M. and Takahashi, A. (2012a). Analytical approximation for non-linear FBSDEs with perturbation scheme. *International Journal of Theoretical and Applied Finance*, 15(5).

Fujii, M. and Takahashi, A. (2012b). Perturbative expansion of FBSDE in an incomplete market with stochastic volatility. *The Quarterly Journal of Finance*, 2(3).

Fujii, M. and Takahashi, A. (2016a). Quadratic-exponential growth BSDEs with jumps and their Malliavin differentiability. *Working Paper,* *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705670*.

Fujii, M. and Takahashi, A. (2016b). Solving backward stochastic differential equations by connecting the short-term expansions. *Working Paper,* *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2795490*.

Gobet, E. and Pagliarani, S. (2014). Analytical approximations of BSDEs with non-smooth driver. *Available at SSRN 2448691.*

Halle, J. O. (2010). Backward stochastic differential equations with jumps. *Master Thesis, University of Oslo.*

He, H. and Pearson, N. (1991). Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. *Journal of Economic Theory*, 54(2):259–304.

Hu, Y., Imkeller, P., and Müller, M. (2005). Utility maximization in incomplete markets. *Annals of Applied Probability*, 15:1691–1712.

Jum, E. (2015). Numerical approximation of stochastic differential equations driven by lévy motion with infinitely many jumps. *Ph.D. Thesis, University of Tennessee, Knoxville.*

Kawai, K. and Maekawa, K. (2004). A note on Yu-Phillips' estimation for a continuous time model of a diffusion process. *Lecture Notes,* *https://www.hue.ac.jp/prfssr/rcfe/w_papers/note_yu_phillips.pdf*.

Kobylanski, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability*, 28(2):558–602.

Ladyzenskaja, O. A., Solonnikov, V. A., and Ural'ceva, N. N. (1986). *Linear and Quasi-linear Equations of Parabolic Type.* American Mathematical Society.

Lejay, A., Mordecki, E., and Torres, S. (2014). Numerical approximation of backward stochastic differential equations with jumps. *Working Paper.*

Li, C. and Chen, D. (2016). Estimating jump-diffusions using closed-form likelihood expansions. *Journal of Econometrics*, 195:51–70.

Liu, J. (2007). Portfolio selection in stochastic environments. *Review of Financial Studies*, 20(1):1–39.

Liu, X. and Li, C. (2000). Weak approximations and extrapolations of stochastic differential equations with jumps. *SIAM Journal on Numerical Analysis*, 37(6):1747–1767.

Lorig, M., Pagliarani, S., and Pascucci, A. (2013). A Taylor series approach to pricing and implied vol for LSV models. *Working Paper, http://arxiv.org/abs/1308.5019v1.*

Lorig, M., Pagliarani, S., and Pascucci, A. (2014). Asymptotics for d-dimensional Lévy-type processes. *Working Paper, http://arxiv.org/abs/1404.3153.*

Lorig, M., Pagliarani, S., and Pascucci, A. (2015a). Analytical expansions for parabolic equations. *SIAM Journal on Applied Mathematics*, 75(2):468–491.

Lorig, M., Pagliarani, S., and Pascucci, A. (2015b). A family of density expansions for Lévy-type processes. *Annals of Applied Probability*, 25(1):235–267.

Lorig, M., Pagliarani, S., and Pascucci, A. (2015c). Pricing approximations and error estimates for local Lévy-type models with default. *Computers and Mathematics with Applications*, 69(10):1189–1219.

Ma, J. and Yong, J. (2000). *Forward-Backward Stochastic Differential Equations and their Applications.* Springer.

Ma, J. and Zhang, J. (2002). Representation theorems for backward stochastic differential equations. *The Annals of Applied Probability*, 12(4):1390–1418.

Menaldi, J. L. and Garroni, M. G. (1992). *Green Functions for Second Order Parabolic Integro-Differential Problems.* Research Notes in Mathematics 275, Longman Scientific and Technical, Essex.

Morlais, M. A. (2009). Quadratic BSDEs driven by a continuous martingales and applications to the utility maximization problem. *Finance and Stochastics*, 13:121–150.

N'ZI, M., Ouknine, Y., and Sulem, A. (2006). Regularity and representation of viscosity solutions of partial differential equations via backward stochastic differential equations. *Stochastic Processes and Their Applications*, 116:1319–1339.

Pagliarani, S. and Pascucci, A. (2012). Analytical approximation of the transition density in a local volatility model. *Central European Journal of Mathematics*, 10(1):250–270.

Pardoux, E. and Peng, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems and Control Letters*, 14(1):55–61.

Pardoux, E. and Peng, S. (1992). *Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations*. Springer.

Pardoux, E. and Tang, S. (1999). Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probability Theory and Related Fields*, 116(9):123–150.

Peng, S. (2014). A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation. *Stochastics*, 38:119–134.

Takahashi, A. and Yamada, T. (2014). An asymptotic expansion for forward-backward SDEs a malliavin calculus approach. *Preprint*.

Tang, S. and Li, X. (1994). Maximum principle for optimal control of distributed parameter stochastic systems with random jumps. *Differential equations, dynamical systems and control science*, 152:867–890.

Tevzadze, R. (2008). Solvability of backward stochastic differential equations with quadratic growth. *Stochastic Processes and Their Applications*, 118:503–515.

Thomée, V. and Zhang, N. (1989). Error estimates for semidiscrete finite element methods for parabolic integro-differential equations. *Mathematics of Computation*, 53:121–139.

Yu, J. (2007). Closed-form likelihood approximation and estimation of jump-diffusions with an application to the realignment risk of the chinese yuan. *Journal of Econometrics*, 141:1245–1280.

Zhan, S. (2005). On the determinantal inequalities. *Journal of Inequalities in Pure and Applied Mathematics*, 6(4):1–7.

Zhang, J. (2005). Representation of solutions to BSDEs associated with a degenerate FSDE. *The Annals of Probability*, 15(3):1798–1831.

Zhang, J. (2006a). The wellposedness of FBSDEs. *Discrete and Continuous Dynamic Systems*, 6(4):927–940.

Zhang, J. (2006b). The wellposedness of FBSDEs (ii). *Preprint*.

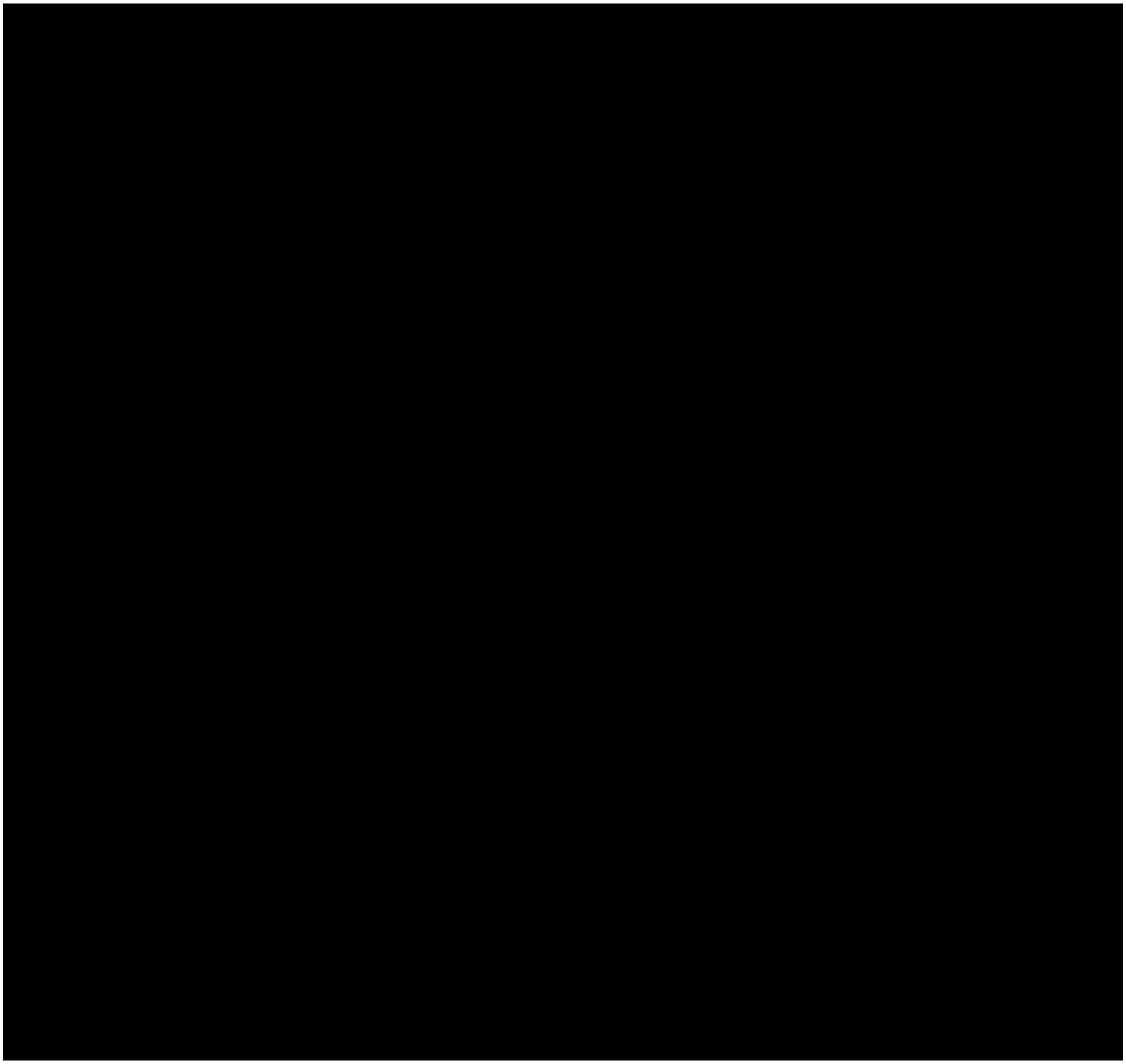
CURRICULUM VITAE

[redacted]

[redacted]

[redacted]

[redacted]



[redacted]

[redacted]