Université de Montréal

Numerical Methods for Discrete-Time Quadratic Hedging

par

Fabrizzio Sabelli

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en mathématiques

Orientation mathématiques appliquées

January 6, 2025

Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

Numerical Methods for Discrete-Time Quadratic Hedging

présenté par

Fabrizzio Sabelli

a été évalué par un jury composé des personnes suivantes :

Florian Maire
(président-rapporteur)

Maciej Augustyniak
(directeur de recherche)

Lucas Benigni
(membre du jury)

Résumé

La couverture quadratique en temps discret et le calcul des ratios de couverture dynamique pour les options ont été peu explorés malgré leur importance dans la gestion des risques financiers. Ce mémoire de maîtrise présente deux avancées significatives dans ce domaine. Tout d'abord, une nouvelle représentation du ratio de couverture quadratique est introduite, permettant le calcul simultané des prix des options et des ratios de couverture en utilisant des méthodes numériques efficaces sans contraintes sur l'espacement des prix d'exercice. Ensuite, des expressions récursives en forme fermée pour les cumulants des modèles de la classe affine à facteurs multiples sont dérivées, permettant une mise en œuvre plus précise et plus efficace sur le plan computationnel de la stratégie de couverture. Des expériences numériques approfondies, utilisant des modèles estimés sur un ensemble de données d'options sur l'indice S&P 500 de 1996 à 2018, démontrent l'efficacité de ces méthodes à travers diverses maturités et prix d'exercice. Les résultats montrent des améliorations substantielles en termes de vitesse de calcul et de précision, offrant un cadre solide pour la gestion des risques dans les marchés financiers complexes.

Mots-clés: Évaluation du prix d'options; Couverture de risque quadratique; Modèles de volatilité multifactoriels; GARCH; Transformation de Fourier rapide (FFT); Expansion Fourier-cosine (Méthode COS); Transformation de Fourier fractionnaire (FRFT)

Abstract

Discrete-time quadratic hedging and the computation of dynamic hedging ratios for options have been underexplored despite their importance in financial risk management. This master's thesis presents two significant advancements in this area. First, a novel representation of the quadratic hedging ratio is introduced, allowing for the simultaneous computation of option prices and hedging ratios using efficient numerical methods without constraints on strike price spacing. Second, closed form recursive expressions for the cumulants of affine multi-factor models are derived, enabling a more accurate and computationally efficient implementation of the hedging strategy. Extensive numerical experiments, using models fitted on a comprehensive dataset of S&P 500 options from 1996 to 2018, demonstrate the effectiveness of these methods across various option maturities and strike prices. The results show substantial improvements in both computational speed and accuracy, offering a robust framework for risk management in complex financial markets.

Keywords: Option pricing, Quadratic hedging, Multi-factor volatility models, GARCH, Fast Fourier transform (FFT), Fourier-cosine expansion (COS Method), Fractional Fourier transform (FRFT)

Contents

List of tables

Acronyms

CAR: compound autoregressive class of models of Darolles et al. (2006)
CGF: cumulant generating function
COS method: Fourier-cosine expansion method of Fang and Oosterlee (2009)
DFRFT: discrete fractional Fourier transform
DFT: discrete Fourier transform
FFT: Fast-Fourier transform
FRFT: Fractional Fast-Fourier transform
GARCH: generalized autoregressive conditional heteroskedasticity
GARCH(C): Christoffersen et al. (2008) Gaussian GARCH component
HN: Heston and Nandi (2000) GARCH
HNSV: Heston et al. (2023) Stochastic Volatility
IDFT: inverse discrete Fourier transform

Remerciements

Je tiens à exprimer ma profonde gratitude envers ma famille pour leur incroyable et constant soutien tout au long de cette aventure. À ma mère Maryse, mon père Remo et ma sœur Renata, vos encouragements et votre amour ont été essentiels pour mener à bien cette maîtrise. Je souhaite également remercier chaleureusement mon superviseur, Dr. Maciej Augustyniak, pour son soutien inestimable, ses conseils éclairés, et sa patience tout au long de ce parcours académique. Vos conseils avisés ont grandement contribué à l'aboutissement de ce travail. Enfin, un merci tout particulier à mon fidèle compagnon, mon chien T-Bone, qui a su, à sa manière, me soutenir et m'accompagner durant les moments les plus intenses de ce périple.

Introduction

Option pricing is a central problem in mathematical finance, concerned with determining the fair value of derivative securities—contracts that derive their value from an underlying asset. These derivatives are vital in financial markets, enabling investors to hedge risks or speculate on future asset prices. The complexity of these instruments often means that closed-form solutions for option prices are unavailable, particularly when certain assumptions about the underlying asset's dynamics are relaxed. This has led to the development of various numerical methods to approximate option prices, which are crucial when analytical solutions are infeasible.

Over the past three decades, numerical methods for option pricing have been extensively studied. A notable contribution came from Carr et al. (2001), who formulated option prices as an inverse Fourier transform, incorporating a damping parameter to ensure the integrability of the call option's payoff at maturity. This approach enabled the use of the Fast-Fourier transform (FFT) to price multiple options on the same underlying asset but at different strike prices simultaneously. This innovation paved the way for several Fourier-based methods, such as the Fourier-cosine expansion method of Fang and Oosterlee (2009) (COS method) and the Fractional Fast-Fourier transform (FRFT), which have since become central to the field. Among these, the COS method stands out for its flexibility and computational efficiency, offering faster convergence than other numerical approaches and avoiding the need to impose restrictions on strike prices when pricing multiple options simultaneously.

Despite extensive research and the availability of option pricing methods, discrete-time quadratic hedging and the computation of dynamic hedging ratios for options have not received the same level of attention. Quadratic hedging, a technique aimed at minimizing the variance of the hedging error, is crucial for effective financial risk management. However, existing methods for computing hedging ratios are often developed for the single-strike setting (Augustyniak et al., 2023), with little comparative analysis of their practical effectiveness, especially in the multi-strike setting.

This master's thesis addresses this gap by making two important contributions to the literature. *First*, we propose a new representation of the quadratic hedging ratio, expressed in terms of the option's price under two different measures: the risk-neutral measure and a measure derived from the risk-neutral measure. This representation allows us to leverage the most efficient numerical methods for option pricing to simultaneously compute both the option's price and its quadratic hedging ratio for multiple strikes simultaneously. Moreover, it overcomes previous limitations that restricted the use of methods like the COS method for option pricing only. Importantly, this approach introduces a novel capability: the ability to simultaneously compute the hedge ratio for multiple options at different strikes without imposing any constraints on the spacing between strikes.

Second, in addition to this novel representation, we derive closed form recursive expressions to compute the first four cumulants of all affine multi-factor models fitting the modeling assumptions provided in Augustyniak et al. (2023). These models, which fall within the compound autoregressive class of models of Darolles et al. (2006) (CAR), encompass most of the option pricing models proposed in discrete-time literature and cover a wide array of Gaussian and non-Gaussian time series models. The closed form expressions of the cumulants are crucial for the efficiency of our proposed method, as they allow for a more accurate and computationally efficient implementation of the quadratic hedging strategy using the COS method.

To conduct our numerical experiments, we focus on two generalized autoregressive conditional heteroskedasticity (GARCH) models that incorporate either one or two volatility components with Gaussian innovations. Indeed, we use Christoffersen et al. (2008) Gaussian GARCH component (GARCH(C)) model and we use Augustyniak et al. (2024)'s formulation of the Heston et al. (2023) Stochastic Volatility (HNSV) model. We use model parameters from Augustyniak et al. (2023) who estimated the model parameters from S&P 500 returns and options data over the period of 1996 to 2018. In our analysis, we rigorously compare the efficiency of various numerical methods by evaluating their error convergence and computational speed across different maturities and numbers of strikes. In particular, we emphasize a thorough comparison of these methods' performance in computing both single-strike and multi-strike scenarios, providing insights into their effectiveness under varying market conditions.

The remainder of this master's thesis is organized as follows. Chapter 1 provides an introduction to discrete-time option pricing using Fourier-based methods. Chapter 2 introduces the general class of discrete-time affine multi-factor models of Augustyniak et al. (2023) and derives a semi-closed form expression for the quadratic hedging ratio of an option. Chapter 3 describes specific GARCH models included in the discrete-time affine framework. Chapter 4 presents the data and evaluation procedure for the numerical experiments, and presents the results from such numerical experiments. Finally, Chapter 4 is followed by the Conclusion.

Chapter 1

Numerical Methods for Option Pricing

1.1 Market Setting

In finance, there are two different scenarios to consider when formulating the filtered probability space in which the financial assets and instruments will evolve. We will focus on formulating the necessary assumptions and definitions for a discrete-time market. In other words, the index set with total order modelling time will be discrete, so over a subset of the natural numbers $\mathbb{N}$. However, we will still include remarks regarding the equivalent formulations for a continuous-time market.

We consider a discrete-time market defined on the time set $\mathcal{I} = \{0,...,T\}$ with one risky asset $S = \{S_t\}_{t\in\mathcal{I}}$ and one risk-free bond $B = \{B_t\}_{t\in\mathcal{I}}$. The market uncertainty is modelled on an augmented filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where $\mathbb{P}$ corresponds to the physical probability measure and $\mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathcal{I}}$ represents the information available to a market participant progressively over time. For any $t \in \mathcal{I}$, we assume the risk-free bond evolves correspondingly to a constant one-period risk-free rate of return r and thus follows the deterministic equation $B_t = e^{rt}$. We denote $Y = \{Y_t\}_{t\in\mathcal{I}}$ as the log-price process of the risky asset such that $Y_t = \log(S_t)$. The log-return process denoted as $y = \{y_t\}_{t\in\mathcal{I}\setminus\{0\}}$ is defined by $y_t = Y_t - Y_{t-1}$. An option is a financial instrument whose price is based on the value of an underlying asset or group of underlying assets. The option's contract allows the owner to buy or sell the underlying asset at a predetermined price depending on the type of option contract they hold. The expiration date T is called the option's maturity, the predetermined price K is called the strike price and buying or selling the underlying at the given strike price is called exercising the option. The option owner can exercise the option up to or at the maturity date depending on the type of option contract they hold. Finally, at the option's maturity, the owner of the option contract is not required to buy or sell the underlying asset. An option giving its owner the opportunity to buy the underlying asset is called a call and an option giving its owner the opportunity to sell the underlying asset is called a put. If an option can be exercised any time before its maturity date then it is an American option and if an option can only be exercised at maturity then it is called a European option. We will specifically focus on European options. Let H represent the payoff of a European option at maturity T then $H : \mathbb{R} \to \mathbb{R}^+$ is a function of the underlying asset's log-price Y_T at maturity T. Indeed, European call and put option payoffs are respectively defined as:

$$H(Y_T) = \max\left[e^{Y_T} - K, 0\right], \qquad H(Y_T) = \max\left[K - e^{Y_T}, 0\right]. \tag{1.1.1}$$

Now, in our simplified market setting, we will define an option's fair value as its no-arbitrage price. We will now introduce some necessary mathematical finance notions to understand no-arbitrage pricing, starting with the first fundamental theorem of asset pricing in discrete-time.

Theorem 1.1.1. *A discrete-time market on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is arbitrage-free (i.e. free of arbitrage opportunities) if and only if there exists at least one probability measure equivalent to the physical measure $\mathbb{P}$ under which the discounted risky asset's price process is a martingale. These equivalent probability measures are usually called equivalent martingale measures or risk-neutral measures and are denoted as $\mathbb{Q}$.*

Now, the price of an option in an arbitrage-free discrete-time market corresponds to the price at which a trader could not profit without taking on risk given the creation of a trading strategy replicating the option's cashflows. A consequence of Theorem 1.1.1 is that the no-arbitrage price of an option is given by the following equation:

$$\begin{aligned} \forall\, t \in \mathcal{T} : V_t &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[H(Y_T) \mid \mathcal{F}_t\right] \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} H(x) p^{\mathbb{Q}}_{Y_T}(x \mid \mathcal{F}_t) dx, \end{aligned} \tag{1.1.2}$$

where $p^{\mathbb{Q}}_{Y_T}(x \mid \mathcal{F}_t)$ corresponds to the $\mathcal{F}_t$-conditional probability density function of the log-price at maturity T.[1]

1.2 Option Pricing and Fourier Based Methods

Now, in order to price an option under $\mathbb{Q}$ (i.e. numerically compute Equation 1.1.2), depending on the modelling assumptions made, various methods have been proposed. Indeed, the most computationally costly method, especially in a high-dimensional setting with respect to the number of underlying assets, is Monte-Carlo simulation. This method is especially used when the option's payoff is heavily path-dependent. If the option's price is not path dependent, it is very common to use quadrature methods and/or Fourier based methods based

[1] We assume the random variables Y_t for $t \in \mathcal{T}$ are absolutely continuous.

on the availability of the underlying assets' characteristic functions as they are very computationally efficient and accurate. Indeed, the Fourier transform of a function corresponds to the following integral transform:

Definition 1.2.1. If $f : \mathbb{R} \to \mathbb{R}$ and $f \in \mathcal{L}^1$ then the Fourier transform $F : \mathbb{R} \to \mathbb{C}$ of f is defined as

$$F(w) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi wx} \, dx. \tag{1.2.1}$$

Now, the Fourier transform is a unitary operator in the $\mathcal{L}^2$ space equipped with the Hilbert inner product. Consequently, the Fourier transform preserves the space $\mathcal{L}^2$ meaning that the Fourier transform of an $\mathcal{L}^2$ function remains in $\mathcal{L}^2$. Thus, we obtain the following definition for $\mathcal{L}^2$-functions:

Definition 1.2.2. If $f : \mathbb{R} \to \mathbb{R}$ and $f \in \mathcal{L}^2$ then we obtain the following pair of Fourier transforms which denote the Fourier transform $F : \mathbb{R} \to \mathbb{C}$ of f and the inverse fourier transform of F:

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-iux} \, dx, \tag{1.2.2}$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(u) e^{iux} \, du, \tag{1.2.3}$$

where we applied the change of variable $u = 2\pi w$ in Equation (1.2.1).

It now follows that the characteristic function is a transformation similar to the Fourier transform. Indeed, the characteristic function of a random variable X is defined as follows:

$$\phi_X(t) = \mathbb{E}\left[e^{itX}\right], \qquad t \in \mathbb{R}. \tag{1.2.4}$$

If we let $f(x)$ in Equation (1.2.2) denote the probability density function of a random variable X then with sign reversal (i.e. $u \to -u$) in Equation (1.2.2), we recover the definition of the characteristic function. Consequently, given the characteristic function of a real random variable always exists and given the conditions of Definition 1.2.2 are respected, it is possible to recover the density $f(x)$ using the integral representation:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixw} \phi_X(-w) \, dw. \tag{1.2.5}$$

This is a crucial detail as often times in finance, it is often not tractable to derive a closed-form expression for the $\mathcal{F}_t$-conditional probability density function of the log-price at maturity T which is necessary to compute the expectation in Equation (1.1.2). However, given certain modelling assumptions, it is possible to obtain a closed-form or semi-closed form expression for the $\mathcal{F}_t$-conditional characteristic function of the log-price at maturity T. Consequently, given the option's payoff function satisfies certain assumptions, it is possible to rephrase Equation (1.1.2) in terms of the $\mathcal{F}_t$-conditional characteristic function of the log-price at maturity T and obtain a semi-closed form expression for an option's price. We introduce the following assumption on the option's payoff function with this goal in mind:

Assumption 1. We assume that the function $H : \mathbb{R} \to \mathbb{R}$ used to define the payoff at time T of the European derivative $H(Y_T)$ where $Y_T = \log(S_T)$ admits the following integral representation:

$$H(Y_T) = \frac{1}{2\pi i}\int_{R-i\infty}^{R+i\infty} e^{zY_T}\check{f}(z)\,dz = \mathrm{Re}\left[\frac{1}{\pi i}\int_{R+i0}^{R+i\infty} e^{zY_T}\check{f}(z)\,dz\right], \tag{1.2.6}$$

where $\mathrm{Re}\,[g]$ denotes the real part of a function g, for $\check{f} : \mathbb{C} \to \mathbb{C}$ and $R \in \mathbb{R}$ such that $z \mapsto \check{f}(R+iz)$ is integrable and $\mathbb{E}^{\mathbb{Q}}[S_T^{2R}] < \infty$.[2]

The second equality is true, since the payoff function at time T of a European derivative is real meaning the integrand is odd in its imaginary part and even in its reals part under complex conjugation. Now, this assumption simply states that we can express the payoff of a European derivative as an inverse Laplace transform. It is common in option pricing and satisfied for most common derivatives. For example, the European call option payoff with a strike price of $K = e^k$ where $k = \log(K)$ admits such an integral representation:

$$H(Y_T) = \max\left[e^{Y_T} - e^k, 0\right] = \mathrm{Re}\left[\frac{1}{\pi i}\int_{R+i0}^{R+i\infty} e^{zY_T}\frac{e^{(1-z)k}}{z(z-1)}\,dz\right], \qquad \text{for any } R > 1, \tag{1.2.7}$$

so that $\check{f}(z) = \frac{e^{(1-z)k}}{z(z-1)}$. The European put option payoff can be expressed similarly by setting $R < 0$. We can link the inverse Fourier transform to the inverse Laplace transform by the following change of variables $u = R + iz$ and obtain a modified version of Assumption 1:

$$H = \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[e^{(R+iu)Y_T}\check{f}\,(R+iu)\right] du. \tag{1.2.8}$$

From now on, we will proceed with the inverse Fourier transform representation. With Assumption 1, we are able to obtain an integral representation for the price of an option in terms of the $\mathcal{F}_t$-conditional cumulant generating function (CGF) of Y_T[3] and the Fourier transform of the option's payoff given the $\mathcal{F}_t$-conditional CGF of the log-price process is available in

[2]Laplace transform representations or equivalently Fourier transform representations of a variety of payoff functions are given in Hubalek et al. (2006) and Raible (2000).

[3]The definition of the $\mathcal{F}_t$-conditional CGF of Y_T is:

$$\mathcal{C}^{\mathbb{Q}}_{Y_T}(u \mid \mathcal{F}_t) = \log(\phi(-iu \mid \mathcal{F}_t)),$$

where $\phi(\,\cdot \mid \mathcal{F}_t)$ represents the $\mathcal{F}_t$-conditional characteristic function of Y_T.

closed-form and a specified risk-neutral measure $\mathbb{Q}$ defined analogously to Theorem 1.1.1:

$$V_t = \frac{e^{-r(T-t)}}{\pi} \int_0^\infty \operatorname{Re}\left[\exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(R+iu \mid \mathcal{F}_t\right)\right) \check{f}(R+iu)\right] du, \tag{1.2.9}$$

where $\mathcal{C}^{\mathbb{Q}}_{Y_T}$ corresponds to the $\mathcal{F}_t$-conditional CGF of Y_T. A proof of the derivation can be found in Appendix D.4.[4] We will focus on the pricing of European call options as the put option price can be obtained via the well known Put-Call Parity:

$$C_t - P_t = e^{Y_t} - e^{k-r(T-t)}, \tag{1.2.10}$$

where C_t and P_t respectively correspond to the call and put prices at time t on an underlying asset paying no dividends with strike e^k and maturity T. Now, in the case of a call option, the R parameter can be viewed as a damping parameter similarly to the damping parameter introduced by Carr et al. (2001) to ensure that the call option's payoff is square integrable.[5] Indeed, the authors derived a similar semi-closed form expression for the price of an option based on the Fourier transform of the modified call price $c_t = e^{\alpha \log(K)} C_t$. If we set $R = \alpha + 1$ for $\alpha > 0$ we obtain their representation of the payoff of a call option:

$$\begin{aligned} C_t &= \frac{e^{-r(T-t)}}{\pi} \int_0^\infty \operatorname{Re}\left[\exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(\alpha+1+iu \mid \mathcal{F}_t\right)\right) \check{f}(\alpha+1+iu)\right] du \\ &= \frac{e^{-r(T-t)-\alpha k}}{\pi} \int_0^\infty \operatorname{Re}\left[\frac{e^{-iuk}}{(\alpha+1+iu)(\alpha+iu)} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(\alpha+1+iu \mid \mathcal{F}_t\right)\right)\right] du. \end{aligned} \tag{1.2.11}$$

Similarly, we obtain the price of a European put option payoff by setting $\alpha < 0$. To numerically compute the integral representation in Equation (1.2.9), any sufficiently accurate quadrature rule such as the Newton–Cotes quadrature rules (i.e. Midpoint, Trapedzoidal or Simpson's rule) can be applied to discretize the integral. Indeed, using Simpson's rule after

[4]Lewis (2002) derived a similar formulation for the price of an option for Levy processes with a known characteristic function. He introduced transform formulas based on the Laplace transform of an option price with respect to the spot price S_t. Lee (2004) then derived general formulas for both discrete and continuous-time markets in terms of the bilateral Laplace transform of an option price with respect to its log-strike price.

[5]Carr et al. (2001) introduced the damping parameter α, since the call option's payoff function is not square integrable and square integrability is required for the inverse Fourier transform to be unique. Indeed, the Parseval relation only stands for $\mathcal{L}^2(\mathbb{R})$ integrable functions and is necessary for the unitary property of the inverse Fourier transform. A sufficient condition for a real valued function f to be square integrable is $\lim_{|x|\to\infty} f(x) = 0$. Clearly, $\lim_{k\to-\infty} e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\max\left(e^{Y_T} - e^k, 0\right) \mid \mathcal{F}_t\right] = e^{Y_t} > 0$.

truncating the integral at a sufficiently large value a,[6] we obtain

$$V_t \approx \frac{e^{-r(T-t)}}{\pi} \sum_{j=0}^{N} \mathrm{Re}\left[\exp\left(\mathcal{C}_{Y_T}^{\mathbb{Q}}\left(R + iu_j \mid \mathcal{F}_t\right)\right) \check{f}(R + iu_j) w_j\right], \tag{1.2.12}$$

where $u_j = j\Delta_u$, $\Delta_u = \frac{a}{N}$ and for $0 \leqslant j \leqslant N$:

$$w_j = \begin{cases} \frac{1}{3}\Delta_u & \text{if } j = 0, N, \\ \frac{2}{3}\Delta_u & \text{if } j \text{ is even}, \\ \frac{4}{3}\Delta_u & \text{if } j \text{ is odd}. \end{cases} \tag{1.2.13}$$

1.3 Pricing Multiple Options Simultaneously

1.3.1 The Fast-Fourier Transform

Carr et al. (2001) also proposed the use of the FFT in order to simultaneously price multiple options of the same underlying and maturity at different strike prices. The FFT is an algorithm that reduces the computational complexity of computing the discrete Fourier transform (DFT) of a sequence or its inverse, the inverse discrete Fourier transform (IDFT), from $\mathcal{O}(n^2)$ to $\mathcal{O}(n\log(n))$.

Definition 1.3.1. Let $\boldsymbol{x} = \{x_j\}_{0\leqslant j\leqslant N-1}$ represent a sequence of N complex numbers then the DFT transforms of the sequence is another sequence of N complex numbers $\boldsymbol{X} = \{X_j\}_{0\leqslant j\leqslant N-1}$ defined by the following equation:

$$X_k = \sum_{n=0}^{N-1} x_n W_N^{-nk}, \quad \text{for} \quad 0 \leqslant k \leqslant N-1, \tag{1.3.1}$$

where $W_N = e^{i\frac{2\pi}{N}}$ is the N^{th} complex root of unity (i.e. $W_N^N = 1$). The sequence $\boldsymbol{x}$ can also be recovered from the sequence $\boldsymbol{X}$ through the IDFT:

$$x_n = \sum_{k=0}^{N-1} X_k W_N^{nk}, \quad \text{for} \quad 0 \leqslant n \leqslant N-1. \tag{1.3.2}$$

Various FFT algorithms have been proposed to compute the discrete Fourier transform of a sequence. The most popular algorithm by Cooley and Tukey (1965) recursively re-expresses a DFT of a composite size $N = N_1N_2$ into smaller DFTs of size N_1 and N_2. Indeed, for $0 \leqslant k_1{,}n_1 \leqslant N_1-1$ and $0 \leqslant k_2{,}n_2 \leqslant N_2-1$, let $k = N_2k_1+k_2$ and $n = N_1n_2+n_1$. By substituting

[6] Lee (2004) linked the two approaches of Lewis (2002) and Carr et al. (2001) under one unified framework. He provides strategies on how to choose the value of α or equivalently R and subsequently choose quadrature parameters Δ_u and N in order to minimize sampling error and truncation error introduced by the discretization of the Fourier integral in Equation (1.2.12).

these definitions of k and n into Equation (1.3.1) and since $W_N^{-N_1 n_2 N_2 k_1} = (W_N^N)^{-n_2 k_1} = 1$, we obtain:

$$X_{N_2 k_1 + k_2} = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x_{N_1 n_2 + n_1} W_N^{-(N_1 n_2 + n_1)(N_2 k_1 + k_2)} \tag{1.3.3}$$

$$= \sum_{n_1=0}^{N_1-1} \left(W_N^{-n_1 k_2} \sum_{n_2=0}^{N_2-1} x_{N_1 n_2 + n_1} W_{N_2}^{-n_2 k_2} \right) W_{N_1}^{-n_1 k_1}, \tag{1.3.4}$$

where the terms $W_N^{-n_1 k_2}$ are called the twiddle factors. Now, the inner-sum represents a DFT of size N_2 and the outer-sum represents a DFT of size N_1. Equation (1.3.3) represents the general Cooley and Tukey (1965) factorization of the DFT. This factorization allows the use of any FFT type algorithm as it simply divides a larger DFT into smaller DFTs. Initially, Cooley and Tukey (1965) demonstrated it was advantageous to use powers of 2 as values of N (i.e. $N = 2^m$ for $m \in \mathbb{N}$) to obtain high computational speedups. Depending on the size of the input data, it even became common practice to zero-pad the input sequence until its length became a power of 2.[7] This special case of the algorithm is known as the radix-2 algorithm.[8] Nevertheless, Cooley and Tukey (1965) also showed that an arbitrary radix $r \in \mathbb{N}$ is also a valid option and modern implementations typically use larger radices such as $r = 32$ in order to take advantage of modern hardware and processors (Frigo & Johnson, 2005).

Now, in order to simultaneously price multiple options of the same underlying and maturity at different strike prices, Equation (1.2.12) can be reformulated as an IDFT. For $0 \leqslant l \leqslant N-1$, let e^{k_l} represent N distinct strike prices corresponding to N options of the same underlying and maturity. Let m_{max} represent largest value of log-moneyness and define the log-moneyness of each option as

$$m_l = Y_t - k_l = -m_{max} + l\Delta_k \quad \text{for} \quad 0 \leqslant l \leqslant N-1, \tag{1.3.5}$$

such that the log-moneyness values are equally distributed from $-m_{max}$ to $m_{max} - \Delta_k$ with a spacing of Δ_k. We also set $m_{max} = \frac{N\Delta_k}{2}$ as this allows us to obtain N different values of log-moneyness centered around the at-the-money strike price. This is the usual formulation as we are interested in options whose strike price is near the current price of the underlying. Given this setting and let $V_t^{(l)}$ represent the price of the option associated with the $(l+1)^{th}$ strike price within the N distinct strikes prices then we can reformulate Equation (1.2.12)

[7] Let $\boldsymbol{x} = \{x_j\}_{0 \leqslant j \leqslant N-1}$ represent a sequence of arbitrary length N where $N \neq 2^m$ for $m \in \mathbb{N}$ then zero-padding the sequence until its length is a power of 2 represents appending zeros to the end of the sequence until its length equals the smallest power of 2 larger than N.

[8] Please refer to the following resources for more information about the Fast-Fourier transform and Cooley-Tukey algorithm

into an IDFT:

$$V_t^{(l)} \approx \mathrm{Re}\left[\sum_{j=0}^{N-1} W_N^{jl} v_t^{(j)}\right], \tag{1.3.6}$$

where for $0 \leqslant j \leqslant N-1$:

$$v_t^{(j)} = \frac{e^{-r(T-t)+(R-1)m_l+Y_t}}{\pi} \frac{e^{-ij\Delta_u m_{max}} w_j}{(R+iu_j)(R+iu_j-1)} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T-Y_t}(R+iu_j \mid \mathcal{F}_t)\right), \tag{1.3.7}$$

represents the IDFT multiplied by a factor of N of $v_t^{(j)}$ and where the log-moneyness spacing Δ_k and the integral spacing Δ_u satisfy the Nyquist relation:[9]

$$\Delta_u \Delta_k = \frac{2\pi}{N}. \tag{1.3.9}$$

The justification of Equation (1.3.6) can be found in Appendix D.1. Equation (1.3.6) can be computed using the inverse FFT and multiplying by a factor N.

Remark 1.3.2. It is possible to obtain an equivalent formulation of Equation (1.3.6) in terms of the DFT by setting $m_l = m_{max} - l\Delta_k$. Indeed, the values of moneyness obtained are now between $-m_{max} + \Delta$ and m_{max} and, using similar steps as in Appendix D.1, we obtain:

$$V_t^{(l)} \approx \mathrm{Re}\left[\sum_{j=0}^{N-1} W_N^{-jl} v_t^{(j)}\right], \tag{1.3.10}$$

where for $0 \leqslant j \leqslant N-1$:

$$v_t^{(j)} = \frac{e^{-r(T-t)+(R-1)m_l+Y_t}}{\pi} \frac{e^{ij\Delta_u m_{max}} w_j}{(R+iu_j)(R+iu_j-1)} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T-Y_t}(R+iu_j \mid \mathcal{F}_t)\right).$$

Now, as defined in Equation (1.3.9), it is only possible to assign a predetermined value to one of Δ_u and Δ_k to ensure the Nyquist-Shannon sampling theorem is respected and prevent inaccurate results. Furthermore, Δ_u and Δ_k are inversely proportional to each other

[9] The Nyquist-Shannon theorem states that to perfectly reconstruct a signal whose highest frequency component is ν, it is necessary to have a sample rate f_s more than 2ν in order to avoid aliasing. Aliasing happens when the highest frequencies cannot be recovered in signal processing and thus are portrayed as lower frequencies. Now let Δt represent the spacing between the values in the time domain and Δf represent the spacing between the values obtained after the signal is transformed in the frequency domain (i.e. Fourier domain). If N represents the number of values sampled then we have the following two relations:

$$\Delta t = \frac{1}{f_s}, \qquad \Delta f = \frac{1}{N\Delta t}.$$

Nevertheless, the FFT assumes that the signal is periodic, so using the relation between angular frequency ω in radians per second and fundamental frequency f in Hertz, we obtain:

$$\omega = 2\pi f \implies \Delta\omega\Delta t = \frac{2\pi}{N}. \tag{1.3.8}$$

meaning we need the approximation of Equation (1.2.12) to remain accurate for larger values of Δ_u if we want the moneyness spacing to be realistic. As in Carr et al. (2001), the use of Simpson's rule is recommended as it is more precise for larger value of Δ_u than the Midpoint and Trapezoidal rule.

1.3.2 The Fractional Fast-Fourier Transform

As mentioned in the previous section, in order to use to take advantage of the computational speedup of the radix-2 FFT algorithm, the input sequence was often zero-padded until its length became a power of 2. This led to parts of the output vector being discarded and unnecessary computations. For example, using the FFT, Carr et al. (2001) computed a 4096-point FFT and obtained 4096 option prices. However, in practice, as Chourdakis (2005) pointed out, only 67 of those options fell within ±20% of the at-the-money option price. Bailey and Swarztrauber (1991) and Bailey and Swarztrauber (1994) proposed an alternative way to compute the discrete Fourier transform of a sequence without the need to satisfy Equation (1.3.9) called the discrete fractional Fourier transform (DFRFT). Indeed, for an arbitrary rational or complex value α, let

$$G_l(\boldsymbol{x}, \alpha) = \sum_{j=0}^{N-1} e^{-i2\pi jl\alpha} x_j, \tag{1.3.11}$$

denote the DFRFT of an arbitrary sequence of N complex numbers $\boldsymbol{x} = \{x_j\}_{0 \leqslant j \leqslant N-1}$. By setting $\alpha = \frac{1}{m}$, it is clear that the discrete Fourier transform is a special case of the DFRFT. Now, inspired by the work of Bluestein (1970), Bailey and Swarztrauber (1991) developed a fast algorithm for the DFRFT which rephrases Equation (1.3.11) as a circular convolution which can be computed by a combination of three FFT. Indeed, noting that $2jl = j^2 + l^2 - (l-j)^2$, we obtain the modified version of Equation (1.3.11):

$$\begin{aligned} G_l(\boldsymbol{x}, \alpha) &= \sum_{j=0}^{N-1} e^{-i\pi\left[j^2+l^2-(l-j)^2\right]\alpha} x_j \\ &= e^{-\pi i l^2\alpha} \sum_{j=0}^{N-1} x_j e^{-i\pi j^2\alpha} e^{i\pi(l-j)^2\alpha} \\ &= e^{-\pi i l^2\alpha} \sum_{j=0}^{N-1} y_j z_{l-j}, \end{aligned} \tag{1.3.12}$$

where for $0 \leqslant j \leqslant N-1$:

$$y_j = x_j e^{-i\pi j^2\alpha}, \qquad z_j = e^{i\pi j^2\alpha}. \tag{1.3.13}$$

Now, Equation (1.3.12) corresponds to a discrete convolution of two functions. Fast algorithms to compute discrete convolutions of two functions usually require the function convolution to be circular. Since Equation (1.3.12) does not satisfy this condition, the authors pad the sequence up to a length $2p$ with $p \geqslant N-1$ as follows:

$$\begin{aligned}
y_j &= 0 \qquad N \leqslant j < 2p, \\
z_j &= 0 \qquad N \leqslant j < 2p-N, \\
z_j &= e^{i\pi(j-2p)^2\alpha} \qquad 2p-N \leqslant j < 2p.
\end{aligned} \tag{1.3.14}$$

It is now clear that

$$G_l\left(\boldsymbol{x}, \alpha\right) = e^{-\pi i l^2 \alpha} \sum_{j=0}^{2p-1} y_j z_{k-j}, \tag{1.3.15}$$

represents a circular convolution, since $z_{k-j} = z_{j-k}$. Thus, as mentioned, the circular convolution is computed by the means of three FFT:

$$G_l\left(\boldsymbol{x}, \alpha\right) = e^{-\pi i l^2 \alpha} F_k^{-1}\left(F_k(\boldsymbol{y}),\, F_k(\boldsymbol{z})\right), \tag{1.3.16}$$

where $\boldsymbol{y} = \{y_j\}_{0 \leqslant j \leqslant 2p-1}$, $\boldsymbol{z} = \{z_j\}_{0 \leqslant j \leqslant 2p-1}$ and

$$F_k(\boldsymbol{x}) = \sum_{j=0}^{N-1} W_N^{-jl} x_j, \qquad F_k^{-1}(\boldsymbol{x}) = \frac{1}{N} \sum_{j=0}^{N-1} W_N^{jl} x_j, \tag{1.3.17}$$

correspond to the discrete Fourier transform and inverse discrete Fourier transform respectively. Choudakis (2005) then proposed setting $\alpha = \frac{\Delta_u \Delta_k}{2\pi}$ to obtain a modified version of Equation (1.3.10):

$$V_t^{(l)}\left(\alpha\right) \approx \operatorname{Re}\left[\sum_{j=0}^{N-1} e^{-i2\pi jl\alpha} v_j^l\right] \tag{1.3.18}$$

which can be computed using the FRFT algorithm described above. This formulation is much more flexible as it allows us to choose the moneyness spacing and integral spacing independently. However, Černỳ (2004) points out that a rule of thumb to follow: *if the initial spacing $\Delta_k = \frac{2\pi}{N\Delta_u}$ obtained after deciding on a value of Δ_u is six times coarser than the desired spacing for the moneyness, one should use the FRFT, otherwise it will be faster to pad the input sequence up to a length equal to a power of 2 and use the FFT.*

1.3.3 The Fourier-Cosine Series Expansion

Fang and Oosterlee (2009) proposed an alternative method to price an option via its Fourier-cosine series expansion. Indeed, as mentioned in Equation (1.2.5), the probability density function of a random variable can be derived from its characteristic function using a transformation analogous to the inverse Fourier transform. The authors proposed instead to reconstruct the integral in Equation (1.2.5) from its Fourier-cosine series. The Fourier-cosine expansion of a function f supported on a bounded subset $[a,b] \subset \mathbb{R}$ is:

$$f(x) = {\sum_{k=0}^{\infty}}' A_k(x) \cos\left(k\pi \frac{x-a}{b-a}\right), \tag{1.3.19}$$

where

$$A_k(x) = \frac{2}{b-a} \int_a^b f(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx, \tag{1.3.20}$$

and ${\sum_{k=0}^{\infty}}'$ denotes that the first term of the summation is weighted by one-half.[10] Now, let $[a,b] \subset \mathbb{R}$ correspond to a large enough subset such that

$$\hat{\phi}(\omega) = \int_a^b e^{i\omega x} f(x) dx \approx \int_{-\infty}^{\infty} e^{i\omega x} f(x) dx = \phi(\omega). \tag{1.3.21}$$

Furthermore, if we rephrase Equation (1.3.20) in terms of $\hat{\phi}$ as follows:

$$A_k(x) = \frac{2}{b-a} \operatorname{Re}\left[\hat{\phi}\left(\frac{k\pi}{b-a}\right) \exp\left(-i\frac{k\pi a}{b-a}\right)\right],$$

and replace the coefficients A_k in Equation (1.3.19) by $F_k \approx A_k$ where

$$F_k(x) = \frac{2}{b-a} \operatorname{Re}\left[\phi\left(\frac{k\pi}{b-a}\right) \exp\left(-i\frac{k\pi a}{b-a}\right)\right]. \tag{1.3.22}$$

then finally, after truncating the summation, we obtain:

$$f(x) \approx {\sum_{k=0}^{N-1}}' F_k(x) \cos\left(k\pi \frac{x-a}{b-a}\right). \tag{1.3.23}$$

Now, for $t \in \mathcal{T}$ let $Z_t = Y_t - \log(K)$. If we reformulate the payoff of a European option in terms of Z_t then we obtain the following reformulation of Equation (1.1.2) in terms of Z_T:[11]

$$\begin{aligned} \forall\, t \in \mathcal{T} : V_t &= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}} \left[H(Z_T) \mid \mathcal{F}_t\right] \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} H(x) p_{Z_T}^{\mathbb{Q}}(x \mid \mathcal{F}_t) dx. \end{aligned} \tag{1.3.24}$$

[10]Fang and Oosterlee (2009) mention that Fourier-cosine expansions usually give optimal approximations of functions with a finite support and, as usual, are also superior when the functions are periodic. Now, let f be the probability density function of the log-price process Y_t then f rapidly decays to zero when $Y_t \to \pm\infty$. Given assumptions for the uniqueness of the inverse Fourier transform, it is assumed that $\lim_{|x|\to\infty} f(x) = 0$ so it is possible to truncate the integration range in Equation (1.2.5) without any loss of accuracy.

[11]For example, the European call option pay reformulated in terms of Z_T instead of S_T is

$$H = \max(S_T - K,0) = K \cdot \max(e^{Z_T} - 1,0).$$

Furthermore, if we expand the $\mathcal{F}_t$-conditional probability density function of Z_T in Equation (1.3.24) in terms of its Fourier-cosine series, we realize that ϕ in Equation (1.3.21) corresponds to the $\mathcal{F}_t$-conditional characteristic function of Z_T.[12] Thus, we obtain

$$V_t \approx e^{-r(T-t)} \int_a^b H(x) \sum_{k=0}^{N-1}{}' F_k(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx, \tag{1.3.25}$$

where H represents the payoff function of a European option in terms of Z_T. Now, after switching the order of the summation and the integral, setting

$$M_k = \frac{2}{b-a} \int_a^b H(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx, \tag{1.3.26}$$

and introducing the $\mathcal{F}_t$-conditional CGF instead of the $\mathcal{F}_t$-conditional characteristic function of Z_T, we obtain the general COS method formula for European options:

$$V_t \approx e^{-r(T-t)} \sum_{k=0}^{N-1}{}' \operatorname{Re}\left[\exp\left(\mathcal{C}_{Z_T}^{\mathbb{Q}}\left(\frac{ik\pi}{b-a} \mid \mathcal{F}_t\right)\right) \exp\left(-i\frac{k\pi a}{b-a}\right)\right] M_k, \tag{1.3.27}$$

where the coefficients M_k can be obtained in closed-form for a variety of options. Indeed, by substituting the payoff function of a European call option $H = K \cdot \max\left(e^{Z_T} - 1, 0\right)$ into Equation (1.3.26) and integrating, Fang and Oosterlee (2009) obtain:

$$M_k^{Call} = \frac{2}{b-a} K \left(\chi_k(0,b) - \psi_k(0,b)\right). \tag{1.3.28}$$

Similarly, by substituting the payoff function of a European put option $H = K \cdot \max\left(1 - e^{Z_T}, 0\right)$ into Equation (1.3.26) and integrating, Fang and Oosterlee (2009) also obtain:

$$M_k^{Put} = \frac{2}{b-a} K \left(-\chi_k(a,0) + \psi_k(a,0)\right), \tag{1.3.29}$$

where χ_k and ψ_k in Equation (1.3.28) and Equation (1.3.29) are defined as follows:

$$\begin{aligned}\chi_k(c,d) = {} & \frac{1}{1+\left(\frac{k\pi}{b-a}\right)^2}\left[\cos\left(k\pi\frac{d-a}{b-a}\right)e^d - \cos\left(k\pi\frac{c-a}{b-a}\right)e^c\right] \\ & + \frac{k\pi}{(b-a)\left[1+\left(\frac{k\pi}{b-a}\right)^2\right]}\left[\sin\left(k\pi\frac{d-a}{b-a}\right)e^d - \sin\left(k\pi\frac{c-a}{b-a}\right)e^c\right],\end{aligned} \tag{1.3.30}$$

and

$$\psi_k(c,d) = \begin{cases} \left[\sin\left(k\pi\frac{d-a}{b-a}\right) - \sin\left(k\pi\frac{c-a}{b-a}\right)\right]\frac{b-a}{k\pi} & k \neq 0, \\ d-c & k = 0. \end{cases} \tag{1.3.31}$$

[12] Since the $\mathcal{F}_t$-conditional probability density function of Z_T rapidly decays as $Y_T \to \pm\infty$, ϕ in Equation (1.3.21) represents a sufficiently accurate approximation of $\hat{\phi}$ given a and b are chosen sufficiently large.

The authors also provide closed-form formulas for the coefficients of digital and gap options and propose the following method to determine the truncation range of Equation (1.3.21):

$$[a,b] := \left[\kappa_1 - L\sqrt{\kappa_2 + \sqrt{\kappa_4}}, \qquad \kappa_1 + L\sqrt{\kappa_2 + \sqrt{\kappa_4}}\right], \tag{1.3.32}$$

where κ_n represents the n^{th} $\mathcal{F}_t$-conditional cumulant of Z_T. The authors determined that Equation (1.3.32) combined with a value of $L = 10$ was accurate for maturities going from 0.1 year to 10 years.[13] Furthermore, the authors suggest using the Put-Call Parity as presented in Equation (1.2.10) to price call options. Indeed, the payoff of a call option grows exponentially with the log-price of the risky asset which introduces a significant cancellation error for large values of L. The authors mention to stick with values of $L \in [7.5,\ 10]$ when pricing call options without using the Put-Call Parity. The authors also determined experimentally that the computational complexity of the method is linear in the number of terms N chosen in the Fourier-cosine series expansion.[14]

Now, in the cases of models that are considered *affine* as in the continuous-time frameworks of Duffie et al. (2000) and Duffie et al. (2003) or the affine discrete-time framework of Augustyniak et al. (2023),[15] it is possible to reformulate the $\mathcal{F}_t$-conditional cumulant generating function of Z_T in terms of the $\mathcal{F}_t$-conditional cumulant generating function of $Y_T - Y_t$:[16]

$$\mathcal{C}^{\mathbb{Q}}_{Z_T}\left(u \mid \mathcal{F}_t\right) = uZ_t + \mathcal{C}^{\mathbb{Q}}_{Z_T - Z_t}\left(u \mid \mathcal{F}_t\right)(u) = uZ_t + \mathcal{C}^{\mathbb{Q}}_{Y_T - Y_t}\left(u \mid \mathcal{F}_t\right)(u). \tag{1.3.33}$$

Let $\boldsymbol{Z}_t = Y_t\mathbf{1}_m - \log\left(\boldsymbol{K}\right)$ where $\boldsymbol{K} \in \mathbb{R}^m$ represents a vector of m different arbitrary strike prices[17] and where $\mathbf{1}_m$ is a vector of 1s of size m Furthermore, let $\boldsymbol{M}_k = N_k\boldsymbol{K}$ represent the vectorized Fourier-Cosine coefficients obtained by multiplying N_k element-wise with the

[13] Fang and Oosterlee (2009) give a truncation rule for very short maturities such as $T = 0.001$ which includes the sixth cumulant:

$$[a,b] := \left[\kappa_1 - L\sqrt{\kappa_2 + \sqrt{\kappa_4 + \sqrt{\kappa_6}}}, \qquad \kappa_1 + L\sqrt{\kappa_2 + \sqrt{\kappa_4 + \sqrt{\kappa_6}}}\right].$$

However, in practice, deriving the sixth cumulant is not very tractable.

[14] Generalizations and extensions of the COS method for a multidimensional risky asset price process, for non-parametric approaches and for Bermudan and discrete-barrier options were derived by Ruijter and Oosterlee (2012), Vladimirov (2024), Fang and Oosterlee (2009), and Fang and Oosterlee (2011) respectively.

[15] Please refer to Section 2 and Definition 2.1 of Duffie et al. (2003) for a general explanation of affine processes.

[16] This reformulation is possible for any stochastic process that has a scale invariant distribution. However, in the mentioned frameworks, the $\mathcal{F}_t$-conditional cumulant generating function of the $Y_T - Y_t$ is available in closed-form formulas.

[17] The logarithm function is applied element-wise on the vector $\boldsymbol{K}$.

vector $\boldsymbol{K}$ where

$$N_k = \begin{cases} \frac{2}{b-a}\left(\chi_k(0,b) - \psi_k(0,b)\right) & \text{for a call option,} \\ \frac{2}{b-a}\left(-\chi_k(a,0) + \psi_k(a,0)\right) & \text{for a put option,} \end{cases} \tag{1.3.34}$$

then we obtain a vectorized version of the general COS method formula for European options:

$$\boldsymbol{V}_t \approx \boldsymbol{K} \odot e^{-r(T-t)} {\sum_{k=0}^{N-1}}' \operatorname{Re}\left[\exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T-Y_t}\left(\frac{ik\pi}{b-a} \mid \mathcal{F}_t\right)\right) \exp\left(ik\pi\frac{\boldsymbol{Z}_t - a}{b-a}\right)\right] N_k, \tag{1.3.35}$$

where the summation can be rewritten as a matrix-vector product. Furthermore, using Equation (1.3.33) and given that Equation (1.3.35) contains the $\mathcal{F}_t$-conditional CGF of $Y_T - Y_t$ then the cumulants used to compute the truncation range of Equation (1.3.32) can be taken in terms of $Y_T - Y_t$ for all strikes without a significant loss in accuracy. This results in a significant gain in computational speed provided that the vector of strikes $\boldsymbol{K}$ does not include very deep in-the-money or out-of-money strike prices. Now, using this vectorized version of the COS method, it is now possible to simultaneously price multiple options with different strikes without any restrictions on the strike prices. Furthermore, the vector $\boldsymbol{K}$ of strikes does not depend on the underlying asset's log-price Y_t and does not require a specific spacing between the strikes as in the FFT or FRFT. Consequently, the COS method is a much more flexible option pricing method than the FFT and the FRFT.

1.4 Other Formulations

Before Carr et al. (2001) derived their semi-closed form expression for the price of an option based on the Fourier transform, Heston (1993) derived a general formula for the price of a European call option. Indeed, the authors made use of the Gil-Pelaez (1951)'s inversion formula[18] to obtain:

$$C_t = S_t P_1 - K e^{-r(T-t)} P_2, \tag{1.4.1}$$

where

$$P_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \operatorname{Re}\left[\frac{K e^{-iu} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(iu + 1 \mid \mathcal{F}_t\right)\right)}{iu \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(1 \mid \mathcal{F}_t\right)\right)}\right] du, \tag{1.4.2}$$

[18]To derive this more general result, the authors made use of the Gil-Pelaez (1951)'s inversion formula:

$$\begin{aligned} \mathbb{P}(X \leqslant x) &= \frac{1}{2} + \frac{1}{2\pi}\int_0^\infty \frac{e^{iux}\phi(-u) - e^{-iux}\phi(u)}{iu} du \\ &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \operatorname{Re}\left[\frac{e^{-iux}\phi(u)}{iu}\right] du, \end{aligned}$$

where $\operatorname{Re}[f]$ denotes the real part of a function f and ϕ denotes the characteristic function of the random variable X.

$$P_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{K e^{-iu} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(iu \mid \mathcal{F}_t \right) \right)}{iu} \right] du, \tag{1.4.3}$$

and where $\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(u \mid \mathcal{F}_t \right)$ denotes the $\mathcal{F}_t$-conditional CGF of the log-price Y_T. Equation (1.4.1) can also be formulated as a single integral:

$$\begin{aligned} C_t &= e^{-r(T-t)} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(1 \mid \mathcal{F}_t \right) \right) \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{K^{-iu} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(iu + 1 \mid \mathcal{F}_t \right) \right)}{iu \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(1 \mid \mathcal{F}_t \right) \right)} \right] du \right) \\ &\quad - e^{-r(T-t)} K \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{K^{-iu} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(iu \mid \mathcal{F}_t \right) \right)}{iu} \right] du \right) \\ &= e^{-r(T-t)} \left[\frac{S_t - K}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{K^{-iu} \left(\exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(iu + 1 \mid \mathcal{F}_t \right) \right) - \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(iu \mid \mathcal{F}_t \right) \right) \right)}{iu} \right] du \right]. \end{aligned}$$

However, this integral formulation of a European call option's price contains a singularity at 0. Consequently, it cannot directly be numerically evaluated using traditional quadrature rules such as Newton-Cotes rules. Nevertheless, it is possible to compute the integrals using various other quadrature rules such as a Gaussian quadrature.

The celebrated Black-Scholes-Merton formula[19] of Black and Scholes (1973) for the price of a European call option can be obtained as a special case of Equation (1.4.1). Indeed, the Black-Scholes-Merton formula for the price of a European call option is:

$$V_t = S_t \Phi\left(d_1 \right) - K e^{-r(T-t)} \Phi\left(d_2 \right). \tag{1.4.4}$$

where $Y_T = \log(S_T)$ is assumed to have a normal distribution, where Φ denotes the cumulative probability function of the standard normal distribution with mean zero and unit variance and

$$d_1 = \frac{1}{\sigma\sqrt{T}} \left(\log\left(\frac{S_t}{K} \right) + \left(r + \frac{1}{2}\sigma^2 \right) T \right), \qquad d_2 = d_1 - \sigma\sqrt{T}, \qquad \sigma^2 = \frac{\mathrm{Var}(Y_T \mid \mathcal{F}_t)}{T - t}.$$

[19] The Black-Scholes-Merton formula was presented under simplified market conditions in continuous-time in which the volatility σ of the returns is constant.

Chapter 2

Numerical Methods for Discrete-Time Quadratic Hedging

Before defining the problem of quadratic hedging, we will introduce necessary mathematical finance terminology as in Augustyniak et al. (2017).

2.1 Terminology

Definition 2.1.1. (Trading strategy) A trading strategy $\boldsymbol{\theta} = \left(\theta^B, \theta^S\right)$ is a pair of stochastic processes where $\theta^B = \left\{\theta_t^B\right\}_{t\in\mathcal{T}}$ is $\mathbb{F}$-adapted and $\theta^S = \left\{\theta_{t+1}^S\right\}_{t\in\mathcal{T}}$ is $\mathbb{F}$-predictable. Now θ_t^B and θ_{t+1}^S respectively represent the number of risk-free bond shares and the number of shares of the risky asset held over the time period $[t,t+1)$.

Definition 2.1.2. (Value process of a hedging portfolio) Let $V^{\boldsymbol{\theta}} = \left\{V_t^{\boldsymbol{\theta}}\right\}_{t\in\mathcal{T}}$ denote the value (after rebalancing) of a hedging portfolio at time $t \in \mathcal{T}$ based on the trading strategy $\boldsymbol{\theta}$ then for $t \in \mathcal{T}$ we define:

$$V_t^{\boldsymbol{\theta}} = \theta_t^B B_t + \theta_{t+1}^S S_t. \tag{2.1.1}$$

We assume that the stock position isn't adjusted at time T which entails that $\theta_{T+1}^S = \theta_T^S$. Initially, $\theta_0^S = 0$ and thus $V_0^{\boldsymbol{\theta}} = \theta_0^B B_0$.

Definition 2.1.3. (Gains process) Let $\tilde{G}^{\boldsymbol{\theta}} = \left\{\tilde{G}_t^{\boldsymbol{\theta}}\right\}_{t\in\mathcal{T}}$ denote the discounted gains process associated with a hedging portfolio then $\tilde{G}_t^{\boldsymbol{\theta}}$ represents the discounted hedging gain for time 0 to t for a strategy θ. We set $\tilde{G}_0^{\boldsymbol{\theta}} = 0$ and for $t \in \mathcal{T} \setminus \{0\}$:

$$\tilde{G}_t^{\boldsymbol{\theta}} = \sum_{k=1}^{t} \theta_t^S \left(\tilde{S}_k - \tilde{S}_{k-1}\right), \tag{2.1.2}$$

where $\forall t \in \mathcal{T} : \tilde{S}_t = B_t^{-1} S_t$.

Definition 2.1.4. (Cost process) Let $\tilde{C}^{\boldsymbol{\theta}} = \{\tilde{C}_t^{\boldsymbol{\theta}}\}_{t\in\mathcal{T}}$ denote the discounted hedging cost process associated with a strategy $\boldsymbol{\theta}$ then the discounted hedging cost from time 0 to t is defined as:

$$\tilde{C}_t^{\boldsymbol{\theta}} = \tilde{V}_t^{\boldsymbol{\theta}} - \tilde{G}_t^{\boldsymbol{\theta}}, \tag{2.1.3}$$

where $\tilde{C}_0^{\boldsymbol{\theta}} = V_0^{\boldsymbol{\theta}}$.

Definition 2.1.5. (Self-financing hedging strategy) A trading strategy is called self-financing if its cost process is constant. In mathematical terms, $\forall\, t \in \mathcal{T}:\ \tilde{C}_t^{\boldsymbol{\theta}} = V_0^{\boldsymbol{\theta}}$. Note that, for a self-financing strategy, since $\tilde{C}_t^{\boldsymbol{\theta}} = \tilde{C}_{t-1}^{\boldsymbol{\theta}}$, we have for $t \in \mathcal{T} \setminus \{0\}$: $\theta_t^B = \theta_{t-1}^B - \left(\theta_{t+1}^S - \theta_t^S\right)\tilde{S}_t$. Note that for a self-financing portfolio,(2.1.1) can be reformulated in terms of the gains process $\tilde{G}_t^{\boldsymbol{\theta}}$:

$$\tilde{V}_t^{\boldsymbol{\theta}} = V_0^{\boldsymbol{\theta}} + \tilde{G}_t^{\boldsymbol{\theta}}. \tag{2.1.4}$$

2.2 Quadratic Hedging

Quadratic hedging has objective to minimize a given quadratic measure of risk or quadratic hedging criterion by solving an optimization problem. See Schweizer (2001) for a survey of quadratic hedging approaches. We will focus on *risk minimization* which aims to minimize the terminal variance of $\Lambda_T = H(Y_T) - V_T^{\boldsymbol{\theta}}$ under a specified risk-neutral measure $\mathbb{Q}$. Indeed, given an initial amount of capital $V_0^{\boldsymbol{\theta}}$, the goal is to find a self-financing strategy $\boldsymbol{\theta}$ that solves the following optimization problem:

$$\arg\min_{V_0^{\boldsymbol{\theta}},\boldsymbol{\theta}} \mathbb{E}^{\mathbb{Q}}\left[\Lambda_T^2\right]. \tag{2.2.1}$$

The solution of Equation (2.2.1) thus replicates the payoff of the option with minimal error. Schweizer (1995) derived the general solution for this risk-minimization strategy. Indeed, for $t \in \mathcal{T} \setminus \{T\}$ we have:

$$\theta_{t+1}^S = e^{-r(T-t)} \frac{\mathbb{E}^{\mathbb{Q}}\left[H(Y_T)\left(e^{-r}S_{t+1} - S_t\right) \mid \mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{Q}}\left[\left(e^{-r}S_{t+1} - S_t\right)^2 \mid \mathcal{F}_t\right]}. \tag{2.2.2}$$

2.3 Modelling Framework: Assumptions and Affine Volatility Factor Models

We will now introduce certain assumptions about the driving factors of the asset return dynamics. We will specifically introduce the general discrete-time affine framework of Augustyniak et al. (2023).

Assumption 2. The asset return dynamics are driven by an $\mathbb{F}$-adapted multi-factor process $\boldsymbol{l} = \{\boldsymbol{l}_t\}_{t\in\mathcal{T}}$ where

$$\boldsymbol{l}_t = \begin{bmatrix} l_{1,t} & \dots & l_{N,t} \end{bmatrix}^\top, \tag{2.3.1}$$

denotes a N-dimensional vector of factors.

These factors can be observable or latent. Now, for discrete-time models, in order to have access to semi-closed form formulas for θ^S_{t+1} based on Equation (2.2.2), it is necessary to have access to semi-closed form expressions of the one-step $\mathcal{F}_t$-conditional joint CGF of the return and factor processes. We introduce the following assumption to ensure this.

Assumption 3. For any real u and $\mathbf{v} = \begin{bmatrix} v_1 & \dots & v_N \end{bmatrix}^\top$, the joint CGF of y_t and $\boldsymbol{l}_t$ conditional on $\mathcal{F}_{t-1}$ has the following affine representation under the physical measure $\mathbb{P}$:

$$\begin{aligned}\mathcal{C}^{\mathbb{P}}_{y_t,\boldsymbol{l}_t}(u,\boldsymbol{v} \mid \mathcal{F}_{t-1}) &= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp(uy_t + \boldsymbol{v}^\top \boldsymbol{l}_t) \mid \mathcal{F}_{t-1}\right]\right)\\ &= \mathcal{A}(u,\boldsymbol{v};t-1,t) + \mathcal{B}_0(u,\boldsymbol{v};t-1,t)y_{t-1} + \mathcal{B}(u,\boldsymbol{v};t-1,t)^\top \boldsymbol{l}_{t-1},\end{aligned}$$

where $\mathcal{F}_{t-1}$ denotes the σ-algebra generated by the noise processes that drive the dynamics of $\{y_s\}_{s=0}^{t-1}$ and $\{\boldsymbol{l}_s\}_{s=0}^{t-1}$. Here, the one-step coefficients $\mathcal{A}(u,\boldsymbol{v};t-1,t)$, $\mathcal{B}_0(u,\boldsymbol{v};t-1,t)$, and

$$\boldsymbol{\mathcal{B}}(u,\boldsymbol{v};t-1,t) = [\mathcal{B}_1(u,\boldsymbol{v};t-1,t) \cdots \mathcal{B}_N(u,\boldsymbol{v};t-1,t)]^\top \tag{2.3.2}$$

are three real-valued functions that satisfy $\mathcal{A}(0,\mathbf{0};t-1,t) = 0$, $\mathcal{B}_0(0,\mathbf{0};t-1,t) = 0$, and $\mathcal{B}(0,\mathbf{0};t-1,t) = \mathbf{0}$ for any $t \in \mathcal{T} \setminus \{0\}$.

Assumption 3 states that the one-step conditional joint CGF of the return and factor processes has an *affine* representation. It also makes it possible to obtain an affine representation of the multi-step conditional joint CGF (i.e. $C^{\mathbb{P}}_{y_T,\boldsymbol{l}_T}(u,\boldsymbol{v} \mid \mathcal{F}_t)$) of the return and factor processes under the physical measure $\mathbb{P}$. Assumption 3 also entails that the joint process $(y,\boldsymbol{l})$ is included in the general class of discrete-time affine processes of Darolles et al. (2006), better known as the CAR class of models. As stated in Augustyniak et al. (2023), CAR processes cover a large array of Gaussian and non-Gaussian time series models. Now, the expectations in Equation (2.2.2) are computed under the risk-neutral measure $\mathbb{Q}$. Consequently, this requires to have access to the risk-neutral multi-step $\mathcal{F}_t$-conditional joint CGF of the log-price and factor processes. Augustyniak et al. (2023) introduce the following general pricing kernel which preserves the affine representation of the multi-step $\mathcal{F}_t$-conditional joint CGF after the change of measure from the physical measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$:

Assumption 4. The risk-neutral probability measure denoted by $\mathbb{Q}$ is defined via the following Radon-Nikodym derivative:

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \prod_{t=1}^{T} \exp\left(\theta_y y_t + \boldsymbol{\theta}_l^\top \boldsymbol{l}_t - \mathcal{C}^{\mathbb{P}}_{y_t,\boldsymbol{l}_t}\left(\theta_y, \boldsymbol{\theta}_l \mid \mathcal{F}_{t-1}\right)\right), \tag{2.3.3}$$

where θ_y and $\boldsymbol{\theta}_l = \begin{bmatrix}\theta_{l,1} & \dots & \theta_{l,K}\end{bmatrix}^\top$, respectively, represent the equity and model factor risk preference parameters (i.e., the market prices of risk), which are assumed to satisfy the following system of equations for $t \in \mathcal{T} \setminus \{0\}$:

$$\mathcal{A}\left(1+\theta_y, \boldsymbol{\theta}_l; t-1, t\right) = \mathcal{A}\left(\theta_y, \boldsymbol{\theta}_l; t-1, t\right) + r, \tag{2.3.4}$$

$$\mathcal{B}_0\left(1+\theta_y, \boldsymbol{\theta}_l; t-1, t\right) = \mathcal{B}_0\left(\theta_y, \boldsymbol{\theta}_l; t-1, t\right), \tag{2.3.5}$$

$$\boldsymbol{\mathcal{B}}\left(1+\theta_y, \boldsymbol{\theta}_l; t-1, t\right) = \boldsymbol{\mathcal{B}}\left(\theta_y, \boldsymbol{\theta}_l; t-1, t\right). \tag{2.3.6}$$

Equations (2.3.4) to (2.3.6) were originally derived under a similar CAR setting by Bertholon et al. (2008) and denoted as the internal consistency conditions. They are often referred to as the no-arbitrage constraints in literature. Indeed, these constraints are equivalent to the condition $C^{\mathbb{Q}}_{y_t,\boldsymbol{l}_t}(1, \mathbf{0} \mid \mathcal{F}_{t-1}) = r$ which ensures the discounted risky asset price process is a martingale under the risk-neutral measure $\mathbb{Q}$. Finally, it is possible to obtain a *linear monotonic pricing kernel* by setting $\boldsymbol{\theta}_l = \mathbf{0}$ in Equation (2.3.3). Now, let $\boldsymbol{l}^* = \{\boldsymbol{l}_t^*\}_{t\in\mathcal{T}}$ represent the factor process under risk-neutral measure $\mathbb{Q}$ and $\odot$ and $\oslash$ represent the Hadamard multiplication and division operators respectively. We define $\boldsymbol{l}_t^* = \boldsymbol{l}_t \odot \boldsymbol{\delta}$ such that $\boldsymbol{\delta} = \begin{bmatrix}\delta & \dots & \delta\end{bmatrix}^\top$ is a vector of model-specific real-valued constants where $\boldsymbol{\delta}$ represents the scaling relationship between the physical and risk-neutral $\mathcal{F}_t$-conditional factors. Using Assumption 4, we can derive an affine representation for the multi-step $\mathcal{F}_t$-conditional joint CGF of the log-price of the risky asset and factor processes under the risk-neutral measure $\mathbb{Q}$ for any given scaling vector $\boldsymbol{\delta}$.

Proposition 2.3.1. *Assume that the process* $(y,\boldsymbol{l})$ *satisfies Assumption 3, define the scaled vector of factors* $\boldsymbol{l}_t^* = \boldsymbol{l}_t \odot \boldsymbol{\delta}$ *and assume that Assumption 4 holds. Then, for any* $t \in \mathcal{T}$*, the* $\mathcal{F}_t$*-conditional risk-neutral joint CGF of* Y_T *and* $\boldsymbol{l}_t^*$ *is given by*

$$\mathcal{C}^{\mathbb{Q}}_{Y_T,\boldsymbol{l}_t^*}(u,\boldsymbol{v} \mid \mathcal{F}_t) = uY_t + \mathcal{A}^*(u,\boldsymbol{v};t,T) + \mathcal{B}_0^*(u,\boldsymbol{v};t,T)Y_t + \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,T)^\top \boldsymbol{l}_t^* \tag{2.3.7}$$

where the multi-step coefficients $\mathcal{A}^*(u,\boldsymbol{v};t,T), \mathcal{B}_0^*(u,\boldsymbol{v};t,T)$ *and* $\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,T)$ *satisfy the following recursions:*

$$\begin{aligned}\mathcal{A}^*(u,\boldsymbol{v};t,T) = \mathcal{A}^*(u,\boldsymbol{v};t+1,T) + \mathcal{A}^*(u + \mathcal{B}_0^*(u,\boldsymbol{v};t+1,T)& \\ + \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T);t,t+1)&,\\ \mathcal{B}_0^*(u,\boldsymbol{v};t,T) = \mathcal{B}_0^*(u + \mathcal{B}_0^*(u,\boldsymbol{v};t+1,T), \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T);t,t+1)&,\end{aligned}$$

$$\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,T) = \mathcal{B}^*(u+\mathcal{B}_0^*(u,\boldsymbol{v};t+1,T),\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T);t,t+1),$$

with the terminal conditions $\mathcal{A}^*(u,\boldsymbol{v};T,T)=0, \mathcal{B}_0^*(u,\boldsymbol{v};T,T)=0$ *and* $\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};T,T)=\boldsymbol{v}$*. The one-step risk-neutral coefficients* $\mathcal{A}^*(u,\boldsymbol{v};t-1,t), \mathcal{B}_0^*(u,\boldsymbol{v};t-1,t)$ *and* $\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t-1,t)$ *are related to their physical counterparts given in Assumption 3 via the following relationships:*

$$\mathcal{A}^*(u,\boldsymbol{v};t-1,t) = \mathcal{A}(u+\theta_y,\boldsymbol{v}\odot\boldsymbol{\delta}+\boldsymbol{\theta}_l;t-1,t) - \mathcal{A}(\theta_y,\boldsymbol{\theta_l};t-1,t), \tag{2.3.8}$$

$$\mathcal{B}_0^*(u,\boldsymbol{v};t-1,t) = \mathcal{B}_0(u+\theta_y,\boldsymbol{v}\odot\boldsymbol{\delta}+\boldsymbol{\theta_l};t-1,t) - \mathcal{B}_0(\theta_y,\boldsymbol{\theta_l};t-1,t), \tag{2.3.9}$$

$$\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t-1,t) = (\boldsymbol{\mathcal{B}}(u+\theta_y,\boldsymbol{v}\odot\boldsymbol{\delta}+\boldsymbol{\theta}_l;t-1,t) - \mathcal{B}(\theta_y,\boldsymbol{\theta_l};t-1,t))\oslash\boldsymbol{\delta}, \tag{2.3.10}$$

where θ_y *and* $\boldsymbol{\theta}_l$ *are assumed to satisfy the no-arbitrage constraints of Equations* (2.3.4) *to* (2.3.6)*.*

Proof. See Appendix A.1 of Augustyniak et al. (2023). ☐

Remark 2.3.2. Note that if we set $\boldsymbol{v}=\mathbf{0}$ in Equation (2.3.7) then we obtain the risk-neutral CGF of Y_T:

$$\mathcal{C}^{\mathbb{Q}}_{Y_T}(u\mid\mathcal{F}_T) = \mathcal{C}^{\mathbb{Q}}_{Y_T,\boldsymbol{l}^*_T}(u,\mathbf{0}\mid\mathcal{F}_t) = uY_t + \mathcal{A}^*(u,\mathbf{0};t,T) + \mathcal{B}_0^*(u,\mathbf{0};t,T)y_t + \boldsymbol{\mathcal{B}}^*(u,\mathbf{0};t,T)^\top\boldsymbol{l}^*_t\ .$$

Now, given Proposition 2.3.7 and Assumptions 3 and 4, it is possible to simplify the general solution for the risk-minimization strategy presented in Equation (2.2.2). Indeed, Augustyniak et al. (2023) simplified it and obtained a representation under one single integral. This integral representation can be computed using similar numerical methods as the ones introduced in Chapter 1. However, the COS method is not applicable to this integral formulation. We introduce in the following proposition a new simplified representation of the quadratic hedging ratio (i.e. Equation (2.2.2)) in terms of the price of an option under two different measures: the risk-neutral measure $\mathbb{Q}$ and the measure $\widetilde{\mathbb{Q}}_t$, a measure derived from the risk-neutral measure. This new representation can be computed using all the numerical methods presented in Chapter 1, including the COS method. Furthermore, since it is formulated in terms of the options prices under two measures, this representation allows us to leverage the most efficient numerical methods for option pricing in order to simultaneously compute the option's price and quadratic hedging ratio.

Proposition 2.3.3. *Assume that the process* $(y,\boldsymbol{l}_t)_{t\geqslant 0}$ *satisfies Assumption 3 and define the scaled vector of factors* $\boldsymbol{l}^*_t=\boldsymbol{l}_t\odot\delta$*. Then, for any European derivative whose payoff function* $H(Y_T)$ *at maturity* T *satisfies Assumption 1, the risk-minimizing hedging strategy given in Equation* (2.2.2) *admits the following integral representation in terms of the risk-neutral* *measure $\mathbb{Q}$ defined in Assumption 4 and in terms of a measure $\widetilde{\mathbb{Q}}_t$ derived from $\mathbb{Q}$:*

$$\theta_{t+1}^{S} = \frac{e^{-r(T-t)} S_t}{e^{-2r} \exp\left(\mathcal{C}_{Y_{t+1},\boldsymbol{l}_{t+1}^*}^{\mathbb{Q}} (2, \mathbf{0} \mid \mathcal{F}_t)\right) - S_t^2} \left(\mathbb{E}^{\widetilde{\mathbb{Q}}_t}[H \mid \mathcal{F}_t] - \mathbb{E}^{\mathbb{Q}}[H \mid \mathcal{F}_t]\right) \tag{2.3.11}$$

where the Radon-Nikodym derivative of $\widetilde{\mathbb{Q}}_t$ with respect to $\mathbb{Q}$ is given by:

$$\left.\frac{d\widetilde{\mathbb{Q}}_t}{d\mathbb{Q}}\right|_{\mathcal{F}_T} = e^{y_{t+1}-r}, \tag{2.3.12}$$

and the expectation under measure $\widetilde{\mathbb{Q}}_t$ is computed using the following $\mathcal{F}_t$-conditional joint CGFs of Y_T and $\boldsymbol{l}_t^*$:

$$\mathcal{C}_{Y_T}^{\widetilde{\mathbb{Q}}_t}(u \mid \mathcal{F}_t) = uY_t - r + \mathcal{A}^{\widetilde{\mathbb{Q}}_t}(u, \mathbf{0}; t, T) + \mathcal{B}_0^{\widetilde{\mathbb{Q}}_t}(u, \mathbf{0}; t, T)\, y_t + \boldsymbol{\mathcal{B}}^{\widetilde{\mathbb{Q}}_t}(u, \mathbf{0}; t, T)^\top \boldsymbol{l}_t^*, \tag{2.3.13}$$

where the multi-step coefficients $\mathcal{A}^{\widetilde{\mathbb{Q}}_t}(u, \boldsymbol{v}; t, T)$, $\mathcal{B}_0^{\widetilde{\mathbb{Q}}_t}(u, \boldsymbol{v}; t, T)$ and $\boldsymbol{\mathcal{B}}^{\widetilde{\mathbb{Q}}_t}(u, \boldsymbol{v}; t, T)$ are given by:

$$\mathcal{A}^{\widetilde{\mathbb{Q}}_t}(u, \boldsymbol{v}; t, T) = \mathcal{A}^*(u, \boldsymbol{v}; t+1, T) + \mathcal{A}^*\left(1 + u + \mathcal{B}_0^*(u, \boldsymbol{v}; t+1, T), \boldsymbol{\mathcal{B}}^*(u, \boldsymbol{v}; t+1, T); t, t+1\right),$$

$$\mathcal{B}_0^{\widetilde{\mathbb{Q}}_t}(u, \boldsymbol{v}; t, T) = \mathcal{B}_0^*\left(1 + u + \mathcal{B}_0^*(u, \boldsymbol{v}; t+1, T), \boldsymbol{\mathcal{B}}^*(u, \boldsymbol{v}; t+1, T); t, t+1\right),$$

$$\boldsymbol{\mathcal{B}}^{\widetilde{\mathbb{Q}}_t}(u, \boldsymbol{v}; t, T) = \boldsymbol{\mathcal{B}}^*\left(1 + u + \mathcal{B}_0^*(u, \boldsymbol{v}; t+1, T), \boldsymbol{\mathcal{B}}^*(u, \boldsymbol{v}; t+1, T); t, t+1\right),$$

with terminal conditions $\mathcal{A}^*(u, \boldsymbol{v}; T, T) = 0$, $\mathcal{B}_0^*(u, \boldsymbol{v}; T, T) = 0$ and $\boldsymbol{\mathcal{B}}^*(u, \boldsymbol{v}; T, T) = \boldsymbol{v}$.

Proof. See Appendix D.5. $\square$

Proposition 2.3.3 allows us to simultaneously compute the price of an option and its corresponding hedge ratio. The expectations in Equation (2.3.11) can be computed numerically using every method introduced in Chapter 1. Consequently, it is now possible to use the COS method to simultaneously compute the hedge ratios of options with different strikes without needing to impose any restrictions on the strikes of each option. Furthermore, Appendix C.1 contains closed form expressions for the first four cumulants of the risky asset log-price process Y_T given the return dynamics are driven by an N-dimensional $\mathbb{F}$-adapted multi-factor process. These closed form expressions are available given that the derivatives up to degree four of the one-step conditional joint CGF of the return and factor processes under risk-neutral measure $\mathbb{Q}$ are derivable in closed-form or through numerical differentiation algorithms. The derivation of the N-dimensional cumulants can be found in Appendix D.3. The necessary notation to understand the cumulant expressions can be found at the start of Appendix D.3 as well. Appendix C.2 also contains closed-form expressions for the first two cumulants given the return dynamics are driven by an 1-dimensional $\mathbb{F}$-adapted multi-factor process and given the return dynamics are not autoregressive (i.e. $\forall\, t \in \mathcal{T} : \mathcal{B}_0^*(u, \boldsymbol{v}; t, T) = 0$). The third and fourth cumulants can be recursively computed as a special case of the closed form expressions of the N-dimensional cumulants presented in Appendix C.1. It remains possible to derive closed-form expressions for the 1-dimensional case, but the analytical expressions are involved and not very tractable.

Chapter 3

Relevant Cases of GARCH models

3.1 Example of a Gaussian Framework

In this chapter, we will investigate the computational performance of different numerical methods in order to compute quadratic hedging ratios of one single option and multiple options with different strikes simultaneously. We will limit ourselves to a Gaussian setting and will only be focusing on the error convergence of the different numerical methods and their computational speed. Refer to Augustyniak et al. (2023) for a complete study of the hedging effectiveness of different models ranging from Gaussian dynamics to inverse Gaussian dynamics. They also present an empirical study of the extent to which multi-component volatility factors, fat tails, and a non-monotonic pricing kernel can improve hedging performance. In our Gaussian framework, the one-period risky assets return dynamics under the physical measure $\mathbb{P}$ are given by the following equations:

$$y_{t+1} = r + \lambda h_t + \sqrt{h_t}\epsilon_{t+1}, \tag{3.1.1}$$

$$h_{t+1} = \mathbf{1}^\top \boldsymbol{l}_{t+1}, \tag{3.1.2}$$

$$l_{i,t+1} = \phi_i l_{i,t} + \alpha_i \left(\psi_{i,t+1}^2 - 2\gamma_i \sqrt{h_t}\psi_{i,t+1} - 1\right), \qquad 1 \leqslant i < N, \tag{3.1.3}$$

$$l_{N,t+1} = (1-\phi_N)\sigma^2 + \phi_N l_{N,t} + \alpha_N \left(\psi_{N,t+1}^2 - 2\gamma_N \sqrt{h_t}\psi_{N,t+1} - 1\right), \tag{3.1.4}$$

where $(\epsilon_t, \psi_{1,t}, \dots, \psi_{N,t})$ follows a $\mathcal{F}_t$-conditional multivariate Gaussian distribution with mean zero and covariance matrix Σ with unit variance on the diagonal. We set $\mathcal{F}_t = \sigma\left(\epsilon_s, \psi_{1,s}, \dots, \psi_{N,s} \mid s \leqslant t\right)$ such that $\mathbb{F}$ is the natural filtration with respect to the innovation process $\{\epsilon_t, \psi_{1,t}, \dots, \psi_{N,t}\}_{t\in\mathcal{I}}$. The parameter λ corresponds to the equity risk premium. Now the conditional variance h_t is decomposed into N components whose dynamics are all driven by their respective innovations.[1] These dynamics also rely on model parameters:

[1] If we set all pairwise correlations between the innovations processes to 1, we obtain the Gaussian case of the general affine framework presented in Augustyniak et al. (2024).

$\boldsymbol{\phi} = \begin{bmatrix}\phi_1 & \dots & \phi_N\end{bmatrix}^\top$ is a vector of persistence parameters for each of the conditional variance components which satisfy the following relation $0 < \phi_i < \phi_N < 1$ for all $1 \leqslant i < N$, $\boldsymbol{\alpha} = \begin{bmatrix}\alpha_1 & \dots & \alpha_N\end{bmatrix}^\top$ is a vector of non-negative autoregressive conditional heteroskedasticity parameters, $\boldsymbol{\gamma} = \begin{bmatrix}\gamma_1 & \dots & \gamma_N\end{bmatrix}^\top$ is the vector of leverage parameters and σ^2 corresponds to the unconditional variance of y_t (i.e. $\sigma^2 = \mathbb{E}^{\mathbb{P}}[h_t] = \mathbb{E}^{\mathbb{P}}[l_{N,t}]$). Furthermore, we assume the model parameters satisfy parametric constraints that ensure the conditional variance's positivity (i.e. $\forall\, t : h_t > 0$).

3.2 Heston-Nandi Stochastic Volatility Model

A generalization of the Heston and Nandi (2000) GARCH (HN) model was recently proposed by Heston et al. (2023). We will refer to this generalization as the HNSV model. We will use a different parametrization introduced by Augustyniak et al. (2024) that fits the general discrete-time framework presented in Augustyniak et al. (2023). The discretized HNSV dynamics under the physical measure $\mathbb{P}$ are given by:

$$y_t = r + \lambda h_{t-1} + \sqrt{h_{t-1}}\epsilon_t, \tag{3.2.1}$$

$$h_t = \sigma^2 + \phi(h_{t-1} - \sigma^2) + \alpha\left(\psi_t^2 - 2\gamma\sqrt{h_{t-1}}\psi_t - 1\right). \tag{3.2.2}$$

where (ϵ_t, ψ_t) follows an $\mathcal{F}_{t-1}$-conditional bivariate standard Gaussian distribution with correlation coefficient ρ. Apart from the parameter ρ, this model shares a similar parameterization as the one presented in Augustyniak and Badescu (2021). We recover the HN model as a special case of HNSV with $\rho = 1$. Indeed, the mapping becomes clear if we rephrase the model dynamics using the relations:

$$\omega = (1-\phi)\sigma^2 - \alpha, \qquad \phi = \beta + \alpha\gamma^2,$$

to obtain

$$y_t = r + \lambda h_{t-1} + \sqrt{h_{t-1}}\epsilon_t, \tag{3.2.3}$$

$$h_t = \omega + \alpha\left(\psi_t - \gamma\sqrt{h_{t-1}}\right)^2 + \beta h_{t-1}. \tag{3.2.4}$$

We show in Appendix D.6 that HNSV satisfies the affine assumptions presented in Assumption 3. Indeed, we obtain the following affine representation for the one-step $\mathcal{F}_{t-1}$-conditional joint CGF of y_t and h_t under the physical measure $\mathbb{P}$:

$$\mathcal{C}^{\mathbb{P}}_{y_t,h_t}(u, v \mid \mathcal{F}_{t-1}) = \mathcal{A}(u, v; t-1, t) + \mathcal{B}_0(u, v; t-1, t)y_{t-1} + \mathcal{B}(u, v; t-1, t)h_{t-1}, \tag{3.2.5}$$

where the coefficients $\mathcal{A}(u,v; t-1,t)$, $\mathcal{B}_0(u,v; t-1,t)$ and $\mathcal{B}(u,v; t-1,t)$ are given by

$$\mathcal{A}(u, v; t-1, t) = ur + v\left((1-\phi)\sigma^2 - \alpha\right) - \frac{1}{2}\log(1 - 2v\alpha), \tag{3.2.6}$$

$$\mathcal{B}_0(u,v;t-1,t) = 0, \tag{3.2.7}$$

$$\mathcal{B}(u,v;t-1,t) = u\lambda + v\phi + \frac{u^2(1-\rho^2)}{2} + \frac{(u\rho - 2v\alpha\gamma)^2}{2(1-2v\alpha)}. \tag{3.2.8}$$

Now, since this model respects the affine assumption, we use the Radon-Nikodym derivative defined in Assumption 4 to obtain the one-step $\mathcal{F}_{t-1}$-conditional joint CGF of y_t and h_t under the risk-neutral measure $\mathbb{Q}$:

$$\mathcal{C}^{\mathbb{Q}}_{y_t,h^*_t}(u,v \mid \mathcal{F}_{t-1}) = \mathcal{A}^*(u,v;t-1,t) + \mathcal{B}^*_0(u,v;t-1,t)y_{t-1} + \mathcal{B}^*(u,v;t-1,t)h^*_{t-1}, \tag{3.2.9}$$

where we have the following relation between risk-neutral and physical conditional variance:

$$\delta = \frac{1-2\theta_l\alpha\left(1-\rho^2\right)}{1-2\theta_l\alpha},$$
$$h^*_t = \delta h_t,$$

and where the coefficients $\mathcal{A}^*$ $(u,v;t-1,t)$, $\mathcal{B}^*_0(u,v;t-1,t)$ and $\mathcal{B}^*$ $(u,v;t-1,t)$ are given by

$$\mathcal{A}^*(u,v;t-1,t) = ur + v\omega^* - \frac{1}{2}\log(1-2v\alpha^*), \tag{3.2.10}$$

$$\mathcal{B}^*_0(u,v;t-1,t) = 0, \tag{3.2.11}$$

$$\mathcal{B}^*(u,v;t-1,t) = -\frac{u}{2} + \left(\beta + \alpha^* q^{*2}\right)v + \frac{u^2\left(1-\rho^{*2}\right)}{2} + \frac{\left(u\rho^* - 2v\alpha^* q^*\right)^2}{2\left(1-2v\alpha^*\right)}, \tag{3.2.12}$$

where

$$\omega^* = \omega\delta, \qquad \alpha^* = \frac{\alpha\delta}{1-2\theta_l\alpha}, \qquad \pi = \frac{1}{1-2\theta_l\alpha\left(1-\rho^2\right)},$$

$$q^* = (\gamma - \theta_y\rho)\sqrt{\pi}, \qquad \rho^* = \rho\sqrt{\pi}.$$

Using the explicit Equations (3.2.10), (3.2.11), and (3.2.12), we derive closed-form expressions for the derivatives of the one-step coefficients under the HNSV modelling assumptions, since they are essential to the use of the COS method for hedging. The one-step derivatives can be found in Appendix B.1. Finally, in order to verify the values of the cumulants using Monte-Carlo simulations, we derive the HNSV dynamics under the risk-neutral measure $\mathbb{Q}$ in Appendix D.6 and obtain

$$y_t = r - \frac{h^*_{t-1}}{2} + \sqrt{h^*_{t-1}}\epsilon^*_t, \tag{3.2.13}$$

$$h^*_t = \omega^* + \alpha^*\left(\psi^*_t - q^*\sqrt{h^*_{t-1}}\right)^2 + \beta h^*_{t-1}, \tag{3.2.14}$$

where (ϵ^*_t, ψ^*_t) follows a standard bivariate Gaussian distribution with correlation coefficient ρ^*.

3.3 Gaussian GARCH Component Model

GARCH(C) model represents an affine adaptation of the two-factor GARCH model of Engle and Lee (1999). The asset return dynamics of the GARCH(C) model under $\mathbb{P}$ are given by:

$$y_t = r + \lambda h_{t-1} + \sqrt{h_{t-1}} z_t, \tag{3.3.1}$$

$$h_t = s_t + q_t, \tag{3.3.2}$$

$$s_t = \rho_s s_{t-1} + \xi_t^{(s)}, \tag{3.3.3}$$

$$q_t = \sigma^2 + \rho_q (q_{t-1} - \sigma^2) + \xi_t^{(q)}, \tag{3.3.4}$$

where z_t has an $\mathcal{F}_{t-1}$-conditional Gaussian distribution with zero mean and unit variance. The driving processes for the two volatility components, $\xi_t^{(s)}$ and $\xi_t^{(q)}$, are defined by the following equations:

$$\xi_t^{(s)} = a_s \left(z_t^2 - 2\gamma_s \sqrt{h_{t-1}} z_t - 1 \right), \qquad \xi_t^{(q)} = a_q \left(z_t^2 - 2\gamma_q \sqrt{h_{t-1}} z_t - 1 \right). \tag{3.3.5}$$

Note again that $\mathbb{E}^{\mathbb{P}}[\xi_t^{(s)} | \mathcal{F}_{t-1}] = \mathbb{E}^{\mathbb{P}}[\xi_t^{(q)} | \mathcal{F}_{t-1}] = 0$ and $\mathbb{E}^{\mathbb{P}}[h_t] = \mathbb{E}^{\mathbb{P}}[q_t] = \sigma^2$. Augustyniak et al. (2023) show that GARCH(C) satisfies the affine assumption presented in Assumption 3. Indeed, recall the one-step joint $\mathcal{F}_{t-1}$-conditional joint CGF of y_t and $\boldsymbol{l}_t$:

$$\mathcal{C}^{\mathbb{P}}_{y_t, \boldsymbol{l}_t}(u, \boldsymbol{v} | \mathcal{F}_{t-1}) = \mathcal{A}(u, \boldsymbol{v}; t-1, t) + \mathcal{B}_0(u, \boldsymbol{v}; t-1, t) y_{t-1} + \boldsymbol{\mathcal{B}}(u, \boldsymbol{v}; t-1, t)^\top \boldsymbol{l}_{t-1}, \tag{3.3.6}$$

with

$$\mathcal{A}(u, \boldsymbol{v}; t, t+1) = ur + \boldsymbol{v}^\top \boldsymbol{K}_1 - \frac{1}{2} \log(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2), \tag{3.3.7}$$

$$\mathcal{B}_0(u, \boldsymbol{v}; t, t+1) = 0, \tag{3.3.8}$$

$$\boldsymbol{\mathcal{B}}(u, \boldsymbol{v}; t, t+1) = \begin{bmatrix} \mathcal{B}_s(u, \boldsymbol{v}; t, t+1) & \mathcal{B}_q(u, \boldsymbol{v}; t, t+1) \end{bmatrix}^\top, \tag{3.3.9}$$

$$\mathcal{B}_s(u, \boldsymbol{v}; t, t+1) = u\lambda + \boldsymbol{v}^\top \boldsymbol{K}_4 + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3)^2}{2(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2)}, \tag{3.3.10}$$

$$\mathcal{B}_q(u, \boldsymbol{v}; t, t+1) = u\lambda + \boldsymbol{v}^\top \boldsymbol{K}_5 + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3)^2}{2(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2)}, \tag{3.3.11}$$

where the two-dimensional vectors $\boldsymbol{K}_1, \ldots, \boldsymbol{K}_5$ are given by:

$$\boldsymbol{K}_1 = \begin{bmatrix} 0 \\ \sigma^2(1-\rho_q) \end{bmatrix} - \boldsymbol{K}_2, \quad \boldsymbol{K}_2 = \begin{bmatrix} a_s \\ a_q \end{bmatrix}, \quad \boldsymbol{K}_3 = \begin{bmatrix} a_s\gamma_s \\ a_q\gamma_q \end{bmatrix}, \quad \boldsymbol{K}_4 = \begin{bmatrix} \rho_s \\ 0 \end{bmatrix}, \quad \boldsymbol{K}_5 = \begin{bmatrix} 0 \\ \rho_q \end{bmatrix}.$$

Now, under the pricing kernel defined in Assumption 4, Augustyniak et al. (2023) derived the relationship between the market prices of risk θ_y and $\boldsymbol{\theta_l}$:

$$\theta_y = -\frac{1}{2} - \lambda + 2\boldsymbol{\theta_l}^\top (\boldsymbol{K}_3 + \lambda \boldsymbol{K}_2). \tag{3.3.12}$$

We define the scaled factor process $\boldsymbol{l}_t^* = \begin{bmatrix} s_t^* & q_t^* \end{bmatrix}^\top$ by setting the scaling vector $\boldsymbol{\delta} = \begin{bmatrix} \delta & \delta \end{bmatrix}^\top$ where $\delta = \frac{1}{1-2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2}$ such that:

$$\boldsymbol{l}_t^* = \boldsymbol{l}_t \odot \boldsymbol{\delta} = \begin{bmatrix} s_t \frac{1}{1-2\boldsymbol{\theta_l}^\top \boldsymbol{K}_2} & q_t \frac{1}{1-2\boldsymbol{\theta_l}^\top \boldsymbol{K}_2} \end{bmatrix}^\top . \tag{3.3.13}$$

This scaling term δ represents the relation between the conditional risk-neutral variance under risk-neutral measure $\mathbb{Q}$ and the physical conditional variance under physical measure $\mathbb{P}$. Furthermore, under the risk-neutral measure $\mathbb{Q}$, the asset returns follow a conditional Gaussian distribution with dynamics:

$$y_t = r - \frac{1}{2} h_{t-1}^* + \sqrt{h_{t-1}^*} z_t^*, \tag{3.3.14}$$

where z_t^* is a Gaussian random variable with zero mean and unit variance. A derivation of the $\mathcal{F}_t$ conditional risk-neutral variance h_t^* can be found in Appendix D.8. Now, if we rewrite Equation (3.3.12) as

$$\theta_y - 2\boldsymbol{\theta_l}^\top \boldsymbol{K}_3 = -\frac{1}{2} - \lambda^*, \tag{3.3.15}$$

where $\lambda^* = \frac{\lambda}{\delta}$ then using Equations (2.3.8) to (2.3.10), we obtain the one-step risk-neutral joint CGF of Y_t and $\boldsymbol{l}_t^*$ under the risk-neutral measure $\mathbb{Q}$:

$$\mathcal{C}^{\mathbb{Q}}_{y_t, \boldsymbol{l}_t^*}(u, \boldsymbol{v} | \mathcal{F}_t) = \mathcal{A}^*(u, \boldsymbol{v}; t-1, t) + \mathcal{B}_0^*(u, \boldsymbol{v}; t-1, t) y_{t-1} + \boldsymbol{\mathcal{B}}^*(u, \boldsymbol{v}; t-1, t)^\top \boldsymbol{l}_{t-1}^*, \tag{3.3.16}$$

with

$$\mathcal{A}^*(u, \boldsymbol{v}; t, t+1) = ur + \boldsymbol{v}^\top \boldsymbol{K}_1^* - \frac{1}{2} \log(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*), \tag{3.3.17}$$

$$\mathcal{B}_0^*(u, \boldsymbol{v}; t, t+1) = 0, \tag{3.3.18}$$

$$\boldsymbol{\mathcal{B}}^*(u, \boldsymbol{v}; t, t+1) = \begin{bmatrix} \mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1) & \mathcal{B}_q^*(u, \boldsymbol{v}; t, t+1) \end{bmatrix}^\top, \tag{3.3.19}$$

$$\mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1) = u\lambda^* + \boldsymbol{v}^\top \boldsymbol{K}_4^* - \frac{1}{2}\left(\lambda^* + \frac{1}{2}\right)^2 + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2})^2}{2(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)}, \tag{3.3.20}$$

$$\mathcal{B}_q^*(u, \boldsymbol{v}; t, t+1) = u\lambda^* + \boldsymbol{v}^\top \boldsymbol{K}_5^* - \frac{1}{2}\left(\lambda^* + \frac{1}{2}\right)^2 + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2})^2}{2(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)}, \tag{3.3.21}$$

where the two-dimensional vectors $\boldsymbol{K}_1^*, \ldots, \boldsymbol{K}_5^*$ are given by:

$$\boldsymbol{K}_1^* = \boldsymbol{K}_1 \odot \boldsymbol{\delta}, \qquad \boldsymbol{K}_2^* = \boldsymbol{K}_2 \odot \boldsymbol{\delta}^{\circ 2}, \qquad \boldsymbol{K}_3^* = \boldsymbol{K}_3 \odot \boldsymbol{\delta}, \qquad \boldsymbol{K}_4^* = \boldsymbol{K}_4, \qquad \boldsymbol{K}_5^* = \boldsymbol{K}_5. \tag{3.3.22}$$

and $\boldsymbol{\delta}^{\circ k}$ represents the Hadamard power raised to the k^{th} power. Using Equations (3.3.17) through (3.3.21), we derive closed-form expressions for the derivatives of the one-step coefficients under the GARCH(C) modelling assumptions, since they are essential to the use of the COS method for hedging. The one-step derivatives can be found in Appendix B.2. Now, as mentioned in Augustyniak et al. (2023), we can obtain the single component GARCH model of Heston and Nandi (2000) as a special case of the GARCH(C) model. Indeed, if we set

$$\alpha = a_s, \qquad \gamma = \gamma_s, \qquad \omega = (1-\rho_s)\sigma^2 - a_s, \qquad \beta = \rho_s - \alpha_s\gamma_s^2,$$

in the HN dynamics under the physical measure $\mathbb{P}$ from Augustyniak and Badescu (2021), we obtain an equivalent representation as setting $a_q = \rho_q = 0$ in Equations (3.3.1) through (3.3.4). We consequently obtain the following two-dimensional vectors $\boldsymbol{K}_1, \ldots, \boldsymbol{K}_5$:

$$\boldsymbol{K}_1 = \begin{bmatrix} 0 \\ \frac{\omega+\alpha}{1-\beta-\alpha\gamma^2} \end{bmatrix} - \boldsymbol{K}_2, \qquad \boldsymbol{K}_2 = \begin{bmatrix} \alpha \\ 0 \end{bmatrix}, \qquad \boldsymbol{K}_3 = \begin{bmatrix} \alpha\gamma \\ 0 \end{bmatrix}, \qquad \boldsymbol{K}_4 = \begin{bmatrix} \beta + \alpha\gamma^2 \\ 0 \end{bmatrix}, \qquad \boldsymbol{K}_5 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Finally, the one-step derivatives under the HN modelling assumptions can obtained as a special case of the one-step derivatives derived under HNSV assumptions with $\rho = 1$.

Chapter 4

Empirical Study of Hedging Error and Speed

This section conducts an empirical study of hedging error and computational speed of a variety of different numerical methods to determine the most efficient method to compute quadratic hedging ratios in discrete-time. Our study considers the hedging of SPX options with different maturities. Our computer codes were written in Python 3.10.13 and they were ran on a MacBook Air with M2 chip with 16 GB of RAM.

To assess the incremental impact of the dimensionality of the factor process $\boldsymbol{l}_t$ on the numerical methods efficiency, we implement the risk-minimization hedging framework presented in Chapter 2 with the two GARCH models discussed in Chapter 3: the HNSV model and the GARCH(C) model. The HNSV parameters are taken from Table 2 in Augustyniak et al. (2024). Please refer to the paper for details on the estimation of the model. Here are the HNSV parameters:

$$\lambda = 2.1257, \qquad \sigma^2 = 1.2007 \times 10^{-4}, \qquad \phi = 0.967, \tag{4.0.1}$$

$$\alpha = 4.4440 \times 10^{-6} \qquad \gamma = 1.8927 \times 10^{2}, \qquad \theta_l = 1.6252 \times 10^{-4}. \tag{4.0.2}$$

The GARCH(C) parameters are taken from Table 1 Augustyniak et al. (2024) or Table 1 from the supplementary material of Augustyniak et al. (2023). The model parameters were estimated from S&P 500 returns and options data over the period of 1996 to 2018. Please refer to Augustyniak et al. (2023) for more details on the estimation of the model. Here are the GARCH(C) parameters:

$$\lambda = 2.1190, \qquad \sigma^2 = 1.2150 \times 10^{-4}, \qquad \rho_s = 0.87662, \tag{4.0.3}$$

$$\alpha_s = 2.5842 \times 10^{-6} \qquad \gamma_s = 3.6089 \times 10^{2}, \qquad \rho_q = 0.98939, \tag{4.0.4}$$

$$\alpha_q = 1.8801 \times 10^{-6}, \qquad \gamma_q = 1.3387 \times 10^{2}, \qquad \boldsymbol{\theta_l} = \begin{bmatrix} 2.2798 \times 10^{3} & 3.7886 \times 10^{4} \end{bmatrix}^{\top}. \tag{4.0.5}$$

To evaluate the effectiveness of the numerical methods presented in Chapter 1 to hedge one option at a time, we compare the performance of the Fourier quadrature method with the COS method. We do not consider the formulation in Equation (1.2.11) of Carr et al. (2001) as it is equivalent to the Fourier quadrature formulation. Furthermore, we do not present results for the formulation in Equation (1.4.1) of Heston and Nandi (2000) as our numerical experiments established it converged much slower using a Gaussian quadrature than the Fourier quadrature using Simpson's rule or the COS method.

We will compare the error convergence and computational speed of the different methods by computing the quadratic hedging ratios of 21 options at different strikes written on the same underlying asset. We specifically compare the convergence of the sum of the absolute errors and the maximum absolute error between the reference quadratic hedging ratios and those computed by the specific methods. We also compare the computational speeds of the different methods in milliseconds. The maximum absolute error (max. abs. error) corresponds to the largest absolute difference between the 21 computed quadratic hedging ratios and their reference values. The total absolute error (tot. abs. error) corresponds to the sum of the absolute differences between the 21 computed quadratic hedging ratios and their reference values. Finally, the CPU time in milliseconds (msec.) represents the total time needed to compute these 21 hedging ratios sequentially for each method.

The strike prices are spaced around the at-the-money strike price of $K = 100$ (i.e. $S_t = 100$) with log-moneyness increments of $\Delta_k = \log(0.005)$. We compute quadratic hedging ratios across four different maturities: 3 months, 6 months, 1 year and 3 years. We compute the reference values using the Fourier quadrature with a value of $N = 2^{17}$ and truncation value $a = 1000$ in the case of HNSV and $a = 1200$ in the case of GARCH(C).

The method specific parameters for the Fourier quadrature method to price a single option for all values of N and under all modelling assumptions are:

$$R = 4.5, \tag{4.0.6}$$

and the upper truncation a depends on the maturity only as follows:

$$a = \begin{cases} 150 & \text{if } T = 3 \text{ months}, \\ 100 & \text{if } T > 3 \text{ months}. \end{cases} \tag{4.0.7}$$

In the case of the COS method, the parameter $L = 10$ for all numerical experiments (i.e. all maturities and values of N). In the case of the integral spacing used in the FFT, it was determined via the Nyquist relation with the log-moneyness spacing set to $\Delta_k = \log(0.005)$ for all maturities and values of N. It follows that the moneyness spacing was set to the same value for the FRFT for all numerical experiments. The integral spacing however for FRFT was set to the following values for all maturities depending on the value of the parameter N respectively:

N	32	64	128	256	512
Δ_u	1.5	1	0.75	0.5	0.25.

he values of the integral spacing were tuned for each value of N to demonstrate the magnitude of the error convergence.

Tables 4.1 and 4.2 present the results under HNSV and GARCH(C) modelling assumptions respectively. In both modelling assumptions, the computational speed is basically equivalent. In the case of HNSV, the COS method method's errors seems to converge slightly faster than the Fourier quadrature in terms of the parameter N, but, in practice, this speed of convergence of the errors is negligible. Indeed, the magnitude of the difference between the errors of the Fourier quadrature and the COS method is very small compared to the hedge ratios values residing in between 0.25 and 0.80. Consequently, when pricing a single strike under HNSV assumptions, we would consider both methods equivalent in computational efficiency. However, it is also important to note that the COS method can be vectorized and price the 21 strikes simultaneously.

In the case of GARCH(C), we can see in Table 4.2 that the COS method's errors converge faster. Indeed, for the shorter maturities of 3 months and 6 months, the COS method's errors converge approximately 1.5 times faster than the Fourier quadrature's errors. For the longer maturities, this factor is approximately 1.3 times faster. However, once again, this faster error convergence can be considered negligible if we realize that the computational speed is in milliseconds and we are computing hedge ratios for maturities much larger in magnitude than milliseconds.

Now, in the case of simultaneously computing these quadratic hedge ratios, Tables 4.3 and 4.4 present the results under HNSV and GARCH(C) modelling assumptions respectively. We only consider the maximum absolute error in this case as it can be considered an upper bound on the absolute error of these methods for computing a single strike. Under both modelling assumptions, it is clear that FFT is outclassed by the FRFT and the COS method in terms of the ratio of its error convergence to its speed. When comparing the results of the FRFT and the COS method, we establish that for a given execution time, the COS method is simply more accurate than the FRFT. This is even more apparent if we take a look at the results under GARCH(C) assumptions in Table 4.4. However, it is once again important to consider that the execution time is in milliseconds, so, for the same arguments related to the magnitude of the hedge ratios and the hedging time mentioned previously, the FRFT and the COS method can be considered equivalent in terms of error convergence. Nevertheless, it is important to note that the COS method can compute quadratic hedge ratios for strikes without imposing any restrictions on the spacing of these strikes. Consequently, due to this flexibility, the COS method can be seen as superior to the FRFT and, of course, the FFT when computing multiple quadratic hedge ratios simultaneously.

Table 4.1. Error convergence and cpu time comparing the Fourier Quadrature and COS method for computing quadratic hedging ratios of 21 European call options successively under HNSV modelling assumptions at different maturities.

Panel A: Call Option with maturity 3 months

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	1.306e-01	83.699	9.215e-03	5.115e-01	82.648	2.467e-02
40	1.181e-02	89.057	9.762e-04	3.142e-01	92.667	1.496e-02
80	4.337e-05	99.993	3.958e-06	1.970e-02	94.677	9.383e-04
160	3.893e-10	111.156	3.990e-11	5.642e-05	106.783	2.687e-06
320	1.929e-10	132.923	1.422e-11	4.672e-10	148.678	2.452e-11

Panel B: Call Option with maturity 6 months

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	7.271e-01	165.446	4.504e-02	5.170e-01	164.406	2.474e-02
40	2.117e-02	173.181	1.548e-03	8.292e-02	177.964	3.949e-03
80	8.326e-05	188.707	6.959e-06	1.058e-03	188.141	5.040e-05
160	1.661e-09	211.848	1.939e-10	1.602e-07	211.001	7.628e-09
320	3.562e-10	256.628	5.625e-11	7.492e-12	253.396	7.293e-13

Panel C: Call Option with maturity 1 year

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	7.487e+00	326.697	3.845e-01	6.346e-01	325.285	3.811e-02
40	1.557e-01	336.610	7.824e-03	8.278e-02	337.748	3.944e-03
80	3.504e-04	374.047	2.213e-05	1.058e-03	372.011	5.040e-05
160	5.214e-09	421.982	4.385e-10	1.602e-07	419.188	7.628e-09
320	4.420e-09	508.822	3.980e-10	1.287e-11	500.699	1.182e-12

Panel D: Call Option with maturity 3 years

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	1.366e+02	980.000	6.983e+00	2.902e+00	975.281	2.106e-01
40	6.892e-01	1004.036	4.672e-02	3.138e+00	1006.787	2.175e-01
80	5.709e-04	1120.127	4.851e-05	1.059e-03	1119.372	5.051e-05
160	8.232e-08	1287.953	9.390e-09	1.602e-07	1263.210	7.630e-09
320	8.232e-08	1527.188	9.395e-09	5.020e-11	1510.433	3.532e-12

Table 4.2. Error convergence and cpu time comparing the Fourier Quadrature and COS method for computing quadratic hedging ratios of 21 European call options successively under GARCH(C) modelling assumptions at different maturities.

Panel A: Call Option with maturity 3 months

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	1.432e-02	96.096	1.014e-03	2.795e+00	94.550	1.624e-01
40	1.279e-05	99.362	9.024e-07	4.811e-01	97.932	2.997e-02
80	3.519e-11	110.729	3.504e-12	1.970e-02	108.470	9.383e-04
160	3.513e-11	129.803	3.491e-12	5.642e-05	123.520	2.687e-06
320	3.513e-11	152.998	3.491e-12	4.543e-10	150.556	2.168e-11

Panel B: Call Option with maturity 6 months

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	8.193e-02	187.908	4.156e-03	4.947e+00	220.130	2.613e-01
40	2.872e-05	194.922	2.861e-06	5.947e-01	193.146	4.329e-02
80	1.250e-09	217.380	8.197e-11	1.058e-03	217.153	5.040e-05
160	1.250e-09	250.445	8.192e-11	1.602e-07	245.581	7.628e-09
320	1.250e-09	301.576	8.192e-11	1.321e-12	303.029	1.223e-13

Panel C: Call Option with maturity 1 year

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	2.493e+00	373.162	1.279e-01	3.416e+01	372.016	2.056e+00
40	1.415e-04	388.050	1.554e-05	1.769e+01	385.364	1.028e+00
80	9.475e-09	441.252	6.252e-10	9.262e-02	430.677	7.455e-03
160	9.475e-09	497.351	6.253e-10	1.602e-07	496.028	7.627e-09
320	9.475e-09	601.785	6.253e-10	6.783e-12	595.522	4.118e-13

Panel D: Call Option with maturity 3 years

	COS			Quadrature		
N	tot. abs. error	cpu time (msec.)	max. abs. error	tot. abs. error	cpu time (msec.)	max. abs. error
20	1.185e+03	1112.865	5.958e+01	8.288e+03	1130.688	4.675e+02
40	1.536e-02	1156.521	1.109e-03	2.912e+03	1154.878	1.547e+02
80	1.356e-06	1296.423	9.667e-08	7.143e+02	1299.353	3.663e+01
160	1.356e-06	1514.486	9.667e-08	2.624e+02	1500.150	1.687e+01
320	1.356e-06	1778.954	9.667e-08	1.063e-09	1804.575	6.079e-11

Table 4.3. Error convergence and cpu time comparing the FFT, FRFT and COS method for computing quadratic hedging ratios of 21 European call options simultaneously under HNSV modelling assumptions at different maturities.

Panel A: Call Option with maturity 3 months								
FFT			FRFT			COS		
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	4.936	2.404e-03	32	4.055	3.484e-04	16	4.010	2.167e-02
256	5.729	2.450e-02	64	4.455	1.392e-05	32	4.098	2.590e-03
512	7.426	3.689e-03	128	4.962	1.756e-07	64	4.517	3.246e-05
1024	10.839	4.376e-05	256	5.776	2.004e-10	128	5.083	3.760e-09
2028	16.867	6.848e-09	512	7.639	1.088e-10	160	5.330	4.274e-11

Panel B: Call Option with maturity 6 months								
FFT			FRFT			COS		
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	9.810	2.723e-02	32	8.020	2.194e-04	16	7.870	1.105e-01
256	11.460	2.429e-02	64	8.785	5.608e-06	32	8.088	2.142e-03
512	15.229	3.689e-03	128	9.839	1.432e-07	64	8.821	6.827e-05
1024	22.201	4.376e-05	256	11.397	9.379e-11	128	9.883	1.417e-08
2028	33.680	6.847e-09	512	15.022	5.864e-13	160	10.244	2.105e-10

Panel C: Call Option with maturity 1 year								
FFT			FRFT			COS		
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	19.559	1.331e-01	32	16.042	2.180e-04	16	15.720	8.759e-01
256	23.023	3.586e-02	64	17.471	5.592e-06	32	16.193	1.269e-02
512	30.824	3.685e-03	128	19.405	1.431e-07	64	17.610	2.373e-04
1024	43.942	4.376e-05	256	22.613	9.419e-11	128	19.508	2.671e-08
2028	67.399	6.848e-09	512	29.844	5.931e-13	160	20.525	5.778e-10

Panel D: Call Option with maturity 3 years								
FFT			FRFT			COS		
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	58.380	1.529e+00	32	47.468	3.101e-04	16	47.144	1.887e+01
256	68.018	2.391e-01	64	51.761	5.592e-06	32	47.700	1.750e-01
512	90.395	1.983e-01	128	58.112	1.431e-07	64	52.394	8.021e-04
1024	132.083	4.381e-05	256	67.367	9.437e-11	128	58.671	1.340e-08
2028	200.926	6.848e-09	512	88.183	1.970e-12	160	59.895	1.260e-08

Table 4.4. Error convergence and cpu time comparing the FFT, FRFT and COS method for computing quadratic hedging ratios of 21 European call options simultaneously under GARCH(C) modelling assumptions at different maturities.

Panel A: Call Option with maturity 3 months								
	FFT			FRFT			COS	
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	5.700	4.117e-02	32	4.664	2.180e-04	16	4.548	4.040e-03
256	6.731	9.867e-02	64	5.124	5.592e-06	32	4.676	2.257e-05
512	8.997	3.699e-03	128	5.781	1.431e-07	64	5.197	3.010e-11
1024	13.265	4.376e-05	256	6.769	9.382e-11	128	5.837	5.952e-12
2028	20.529	6.847e-09	512	9.212	8.871e-14	160	6.167	5.952e-12
Panel B: Call Option with maturity 6 months								
	FFT			FRFT			COS	
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	11.370	6.431e-01	32	9.103	2.185e-04	16	8.865	1.263e-02
256	13.275	2.729e-01	64	10.130	5.592e-06	32	9.151	5.353e-05
512	17.833	3.795e-02	128	11.383	1.431e-07	64	10.104	8.662e-11
1024	26.308	4.376e-05	256	13.437	9.388e-11	128	11.300	9.223e-11
2028	40.749	6.847e-09	512	18.426	8.049e-14	160	12.052	9.223e-11
Panel C: Call Option with maturity 1 year								
	FFT			FRFT			COS	
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	22.451	4.056e+00	32	18.350	1.132e-01	16	17.631	5.281e-01
256	26.592	2.025e+00	64	20.161	1.231e-05	32	18.283	7.577e-04
512	36.217	1.019e+00	128	22.549	1.431e-07	64	20.031	7.837e-10
1024	53.490	5.340e-03	256	26.687	9.397e-11	128	22.613	6.697e-10
2028	81.392	6.848e-09	512	36.262	1.460e-13	160	23.727	6.697e-10
Panel D: Call Option with maturity 3 years								
	FFT			FRFT			COS	
N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error	N	cpu time (msec.)	max. abs. error
128	67.127	9.196e+02	32	54.162	2.239e+02	16	52.691	2.826e+02
256	80.163	4.602e+02	64	60.079	1.322e+02	32	54.289	9.978e-02
512	107.945	1.595e+02	128	67.377	8.950e+01	64	59.843	1.006e-07
1024	160.247	4.577e+01	256	79.863	7.579e-02	128	67.401	1.006e-07
2028	243.660	1.545e+01	512	108.114	1.646e-11	160	70.430	1.006e-07

Conclusion

The analysis presented in this thesis contributes significantly to the field of option pricing and hedging, particularly by addressing the limitations in existing methodologies for discrete-time quadratic hedging. Through the development of a new, simplified representation of the quadratic hedging ratio and the derivation of closed form recursive expressions for the $\mathcal{F}_t$-conditional cumulants of the log-price at time maturity of affine multi-factor models, this work extends the applicability and efficiency of Fourier-based numerical methods in hedging contexts.

Our proposed approach not only enhances the computational efficiency of hedging strategies but also introduces the novel capability to simultaneously compute hedge ratios for multiple options across different strike prices without the need to impose constraints on the spacing of the strike prices. This advancement overcomes the constraints typically associated with traditional methods, thereby broadening the practical utility of techniques like the COS method. The results from our extensive numerical experiments demonstrate the robustness and efficiency of the proposed methods across a wide range of scenarios, particularly in the context of multi-strike option hedging under different market conditions.

The findings underscore the importance of integrating advanced numerical methods with a deeper understanding of model dynamics, especially in the context of affine multi-factor models. By leveraging the efficiency of closed form cumulant expressions, our work facilitates more accurate and faster computations, which are critical for both theoretical advancements and practical implementations in financial markets.

Future research could expand on this work by deriving general closed form recursive expressions for the $\mathcal{F}_t$-conditional joint cumulants of the log-price and multi-factor processes. These advanced expressions would enable the pricing of variance swaps and higher-order moment swaps under discrete-time modeling assumptions, offering a deeper insight into the risk management and hedging strategies of these complex derivatives. Additionally, further exploration could involve extending these formulas to accommodate path-dependent options, such as American options, thereby broadening their applicability and enhancing their utility in more complex financial instruments. Such advancements would not only refine the current methodologies but also open new avenues for practical applications in derivative pricing and hedging under various market conditions.

References

Augustyniak, M., & Badescu, A. (2021). On the computation of hedging strategies in affine garch models. *Journal of Futures Markets*, *41*(5), 710–735. https://doi.org/https://doi.org/10.1002/fut.22187

Augustyniak, M., Badescu, A., & Bégin, J.-F. (2023). A discrete-time hedging framework with multiple factors and fat tails: On what matters. *Journal of Econometrics*, *232*(2), 416–444. https://doi.org/https://doi.org/10.1016/j.jeconom.2021.08.002

Augustyniak, M., Badescu, A., Bégin, J.-F., & Jayaraman, S. K. (2024). On the relation between discrete and continuous-time affine option pricing models. *Submitted*. https://doi.org/10.2139/ssrn.4819165

Augustyniak, M., Godin, F., & Simard, C. (2017). Assessing the effectiveness of local and global quadratic hedging under garch models. *Quantitative Finance*, *17*(9), 1305–1318. https://doi.org/10.1080/14697688.2017.1279342

Badescu, A., Cui, Z., & Ortega, J.-P. (2019). Closed-form variance swap prices under general affine garch models and their continuous-time limits. *Annals of Operations Research*, *282*. https://doi.org/10.1007/s10479-018-2941-9

Bailey, D. H., & Swarztrauber, P. N. (1991). The fractional fourier transform and applications. *SIAM Review*, *33*(3), 389–404. https://doi.org/10.1137/1033097

Bailey, D. H., & Swarztrauber, P. N. (1994). A fast method for the numerical evaluation of continuous fourier and laplace transforms. *SIAM Journal on Scientific Computing*, *15*(5), 1105–1110. https://doi.org/10.1137/0915067

Bekele, A. (2016). *Cooley-tukey fft algorithms* (tech. rep.). Carleton University. Retrieved August 8, 2024, from https://people.scs.carleton.ca/~maheshwa/courses/5703COMP/16Fall/FFT_Report.pdf

Bertholon, H., Monfort, A., & Pegoraro, F. (2008). Econometric Asset Pricing Modelling. *Journal of Financial Econometrics*, *6*(4), 407–458. https://doi.org/10.1093/jjfinec/nbn011

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654. Retrieved August 5, 2024, from http://www.jstor.org/stable/1831029

Bluestein, L. (1970). A linear filtering approach to the computation of discrete fourier transform. *IEEE Transactions on Audio and Electroacoustics*, *18*(4), 451–455. https://doi.org/10.1109/TAU.1970.1162132

Carr, P., Stanley, M., & Madan, D. (2001). Option valuation using the fast fourier transform. *J. Comput. Finance*, *2*. https://doi.org/10.21314/JCF.1999.043

Černỳ, A. (2004). Introduction to fast fourier transform in finance. *Journal of Derivatives*, *12*(1), 73–88.

Chourdakis, K. (2005). Option pricing using the fractional fft. *Journal of Computational Finance*, *8*. https://doi.org/10.21314/JCF.2005.137

Christoffersen, P., Jacobs, K., Ornthanalai, C., & Wang, Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, *90*(3), 272–297. https://doi.org/https://doi.org/10.1016/j.jfineco.2007.12.003

Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, *19*(90), 297–301. Retrieved August 8, 2024, from http://www.jstor.org/stable/2003354

Darolles, S., Gourieroux, C., & Jasiak, J. (2006). Structural laplace transform and compound autoregressive models. *Journal of Time Series Analysis*, *27*(4), 477–503. https://doi.org/https://doi.org/10.1111/j.1467-9892.2006.00479.x

Duffie, D., Filipović, D., & Schachermayer, W. (2003). Affine processes and applications in finance. *The Annals of Applied Probability*, *13*(3), 984–1053. https://doi.org/10.1214/aoap/1060202833

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, *68*(6), 1343–1376. https://doi.org/https://doi.org/10.1111/1468-0262.00164

Engle, R. F., & Lee, G. g. J. (1999). A Long-Run and Short-Run Component Model of Stock Return Volatility. https://doi.org/10.1093/oso/9780198296836.003.0020

Fang, F., & Oosterlee, C. W. (2009). A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, *31*(2), 826–848. https://doi.org/10.1137/080718061

Fang, F., & Oosterlee, C. (2009). Pricing early-exercise and discrete barrier options by fourier-cosine series expansions. *Numerische Mathematik*, *114*, 27–62. https://doi.org/10.1007/s00211-009-0252-4

Fang, F., & Oosterlee, C. (2011). A fourier-based valuation method for bermudan and barrier options under heston's model. *SIAM J. Financial Math.*, *2*, 439–463. https://doi.org/10.1137/100794158

Frigo, M., & Johnson, S. (2005). The design and implementation of fftw3. *Proceedings of the IEEE*, *93*(2), 216–231. https://doi.org/10.1109/JPROC.2004.840301

Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika*, *38*(3/4), 481–482. Retrieved August 5, 2024, from http://www.jstor.org/stable/2332598

Heckbert, P. (1998). *Fourier transforms and the fast fourier transform (fft) algorithm* (tech. rep.). Carnegie Mellon University. Retrieved August 8, 2024, from https://www.cs.cmu.edu/afs/andrew/scs/cs/15-463/2001/pub/www/notes/fourier/fourier.pdf

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, *6*(2), 327–343. Retrieved August 5, 2024, from http://www.jstor.org/stable/2962057

Heston, S. L., Jacobs, K., & Kim, H. J. (2023). A new closed-form discrete-time option pricing model with stochastic volatility. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4607397

Heston, S. L., & Nandi, S. (2000). A closed-form garch option valuation model. *The Review of Financial Studies*, *13*(3), 585–625. Retrieved August 5, 2024, from http://www.jstor.org/stable/2645997

Hubalek, F., Kallsen, J., & Krawczyk, L. (2006). Variance-optimal hedging for processes with stationary independent increments. *The Annals of Applied Probability*, *16*(2), 853–885. https://doi.org/10.1214/105051606000000178

Lee, R. (2004). Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance*, *7*, 51–86. https://api.semanticscholar.org/CorpusID:16935634

Lewis, A. (2002). A simple option formula for general jump-diffusion and other exponential levy processes. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.282110

Raible, S. (2000). *Lévy processes in finance: Theory, numerics, and empirical facts* [Doctoral dissertation, Albert Ludwig University of Freiburg, Germany].

Ruijter, M. J., & Oosterlee, C. W. (2012). Two-dimensional fourier cosine series expansion method for pricing financial options. *SIAM Journal on Scientific Computing*, *34*(5), B642–B671. https://doi.org/10.1137/120862053

Schweizer, M. (2001). A guided tour through quadratic hedging approaches. In E. Jouini, J. Cvitanic, & M. Musiela (Eds.), *Handbooks in mathematical finance: Option pricing, interest rates and risk management* (pp. 538–574). Cambridge University Press.

Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, *20*(1), 1–32. Retrieved August 28, 2024, from http://www.jstor.org/stable/3690105

Seber, G. A. F. (2007). Differentiation. In *A matrix handbook for statisticians* (pp. 351–382). John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1002/9780470226797.ch17

Vladimirov, E. (2024). Icos: Option-implied cos method. https://arxiv.org/abs/2309.00943

Wikipedia contributors. (2024a). Kronecker product — Wikipedia, the free encyclopedia [[Online; accessed 29-August-2024]].

Wikipedia contributors. (2024b). Tensor reshaping — Wikipedia, the free encyclopedia [[Online; accessed 29-August-2024]].

Appendix A

Preliminary Notions

A.1 Probability Notions

Definition A.1.1. Let $\mathcal{C}$ be a collection of subsets of a set Ω then $\mathcal{C}$ is a σ-algebra on Ω if

(1) $\Omega \in \mathcal{C}$

(2) if $E \in \mathcal{C}$ then $E^C \in \mathcal{C}$

(3) let $E_1, E_2, ...$ be a series of subsets in $\mathcal{C}$ then $\cup_{n=1}^{\infty} E_n \in \mathcal{C}$.

Definition A.1.2. A measure $\mu : \Omega \to [0 + \infty]$ on $\Omega, \mathcal{F}$ is a function on a σ-algebra $\mathcal{F}$ of Ω satisfying the following conditions:

(1) $\mu(E)\ \forall\ E \in \mathcal{F}$

(2) let $E_1, E_2, ...$ be a series of mutually disjoint subsets in $\mathcal{F}$ then $\mu(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$

If $\mu(\Omega) = 1$, μ represents a probability measure usually denoted by $\mathbb{P}$. We denote the following pair $(\Omega, \mathcal{F})$ as a measurable space and the following triplet $(\Omega, \mathcal{F}, \mathbb{P})$ as a measured space.

Definition A.1.3. A probability space denotes the triplet $(\Omega, \mathcal{F}, \mathbb{P})$ where

(1) Ω : a sample space which corresponds to the set of all possible outcomes of a random experiment.

(2) $\mathcal{F}$: an event space which corresponds to a σ-algebra composed of all events to be considered. An event is a set composed of outcomes from the sample space.

(3) $\mathbb{P}$: a probability measure assigning a each event of the sample space a number between 0 and 1 (i.e. a probability).

Definition A.1.4. Let $(\mathcal{X}, \mathcal{A})$ and $(\mathcal{Y}, \mathcal{B})$ be two measurable spaces. A function $f : \mathcal{X} \to \mathcal{Y}$ is said to be $\mathcal{A} - \mathcal{B}$ measurable if $\forall B \in \mathcal{B} : f^{-1}(\mathcal{B}) \in \mathcal{A}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{X}, \mathcal{A})$ be a measurable space then $X : \Omega \to \mathcal{X}$ is called a random variable if it is $\mathcal{F} - \mathcal{A}$ measurable.

Now a random variable can either have a countable or discrete set of outcomes (e.g. the outcomes of rolling a dice) or an uncountable or continuous set of outcomes (e.g. a subset of $\mathbb{R}$). If the random variable is discrete then it has a discrete probability distribution and otherwise it has an absolutely continuous probability distribution.[1] Discrete probability distributions are characterized by the probability mass function

$$p(x) = \mathbb{P}(X = x). \tag{A.1.1}$$

This allows us to calculate the probability of the results of a discrete random variable X

$$\mathbb{P}\left(X \in E\right) = \sum_{\omega \in \Omega \cap E} \mathbb{P}\left(X = \omega\right),$$

where E is a subset of Ω. This gives rise to the discrete cumulative distribution function

$$F(x) = \mathbb{P}(X \leqslant x) = \sum_{\omega \leqslant x} p(\omega), \tag{A.1.2}$$

since a discrete random variable can also be defined as a random variable whose cumulative distribution function is a step function that only increases at its discontinuities where the discontinuities form countable sets. Now an absolutely continuous random variable has an absolutely continuous distribution which is uniquely identified by its probability density function $f : \mathbb{R} \to [0, +\infty]$. Indeed, the probability that the outcome of a random variable X belongs to an arbitrary interval $[a, b] \subset \mathbb{R}$ is given by

$$\mathbb{P}(a \leqslant X \leqslant b) = \int_b^a f(x)\ dx, \tag{A.1.3}$$

It now follows that the probability $\mathbb{P}(X = a) = 0$ for a continuous random variable.[2] We also obtain a continuous equivalent to the discrete cumulative distribution function using the probability density function of a continuous random variable X

$$F(x) = \mathbb{P}(X \leqslant x) = \sum_{-\infty}^{x} f(t) dt. \tag{A.1.4}$$

We will focus on real continuous random variables for the remainder of this section as the price of financial instruments follows a real continuous distribution.

[1] We will omit the case of continuous random variables as they are not present in this scope.

[2] If we replace the interval $[a,b]$ by the set E, we obtain the measure-theoretic form

$$\mathbb{P}(X \in E) = \int_E f(x) dx.$$

Definition A.1.5. For $p \geqslant 1$, The p^{th} moment of a real continuous random variable X following an arbitrary probability distribution is defined as

$$\mathbb{E}[X^p] = \int_{-\infty}^{\infty} x^p f(x) dx$$

where $f(x)$ denotes the probability density function of the distribution. If $\mathbb{E}[|X^p|] < \infty$ then we write $X \in \mathcal{L}^p$. If $X \in \mathcal{L}^1$, we say X is an integrable random variable.

Now the probability density function isn't the only unique identifier of a probability distribution. The characteristic function

$$\phi_X(t) = \mathbb{E}\left[e^{itX}\right], \qquad i \in \mathbb{C} \tag{A.1.5}$$

is often denoted as the principal identifier of a probability distribution as it always exists which is not the case of the probability density function. Indeed, since $|e^{itX}| = |\cos(tX) + i\sin(tX)| = 1$ where $|\cdot|$ denotes the norm of a complex number, $\left|\mathbb{E}[e^{itX}]|\right| < \mathbb{E}\left[\left|e^{itX}\right|\right] = 1 < \infty$ by Jensen's inequality.[3]

Definition A.1.6. If $f : \mathbb{R} \to \mathbb{R}$ and $f \in \mathcal{L}^1$ then the Fourier transform $F : \mathbb{R} \to \mathbb{C}$ of f is defined as

$$F(w) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi wx}\, dx. \tag{A.1.6}$$

Now, the Fourier transform is a unitary operator in the $\mathcal{L}^2$ space equipped with the Hilbert inner product. Consequently, the Fourier transform preserves the space $\mathcal{L}^2$ meaning that the Fourier transform of an $\mathcal{L}^2$ function remains in $\mathcal{L}^2$. Thus, we obtain the following definition.

Definition A.1.7. If $f : \mathbb{R} \to \mathbb{R}$ and $f \in \mathcal{L}^2$ then we obtain the following pair of Fourier transforms which denote the Fourier transform $F : \mathbb{R} \to \mathbb{C}$ of f and the inverse fourier transform of F:

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-iux}\, dx, \tag{A.1.7}$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(u) e^{iux}\, du, \tag{A.1.8}$$

where we applied the change of variable $u = 2\pi w$.

[3]Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, X be an integrable real-valued random variable and φ are convex function then

$$\varphi\left(\mathbb{E}\left[X\right]\right) \leqslant \mathbb{E}\left[\varphi\left(X\right)\right]$$

Now we can also obtain the CGF of a random variable X in terms of its characteristic function

$$\mathcal{C}_X(t) = \log\left(\phi\left(-it\right)\right). \tag{A.1.9}$$

The CGF allows us to easily recover the cumulants of a random variable such as the mean, variance. Indeed, the n^{th} cumulant can be recovered by taking the n^{th} degree derivative of the CGF evaluated at 0

$$\kappa_n = \frac{\partial^n}{\partial t^n}\mathcal{C}_X(t)\bigg|_{t=0}.$$

Definition A.1.8. Let μ and ν be two finite measures on the same measurable space $(\mathcal{X}, \mathcal{F})$. We say ν is absolutely continuous with respect to the measure μ if for every measurable set $A \in \mathcal{F}$, $\mu(A) = 0$ implies $\nu(A) = 0$.

Theorem A.1.9 (Radon-Nikodym)**.** *Let $(\mathcal{X}, \mathcal{F})$ be a measurable space on which two finite measures μ and ν are defined. If $\nu \ll \mu$ then there exists a $\mathcal{F}$-measurable function $f : \mathcal{X} \to [0, +\infty)$ such that for every set $A \in \mathcal{F}$ we have*

$$\nu(A) = \int_A f \, d\mu.$$

We call f the Radon-Nikodym derivative and is usually denoted as $\frac{d\nu}{d\mu}$.

Definition A.1.10. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{F}_0 \subset \mathcal{F}$ be a sub-σ-algebra and $X \in \mathcal{L}^1$ be a random variable. We say $\mathbb{E}\left[X \mid \mathcal{F}_0\right]$ is the conditional expectation of X given $\mathcal{F}_0$ if

(1) $\mathbb{E}\left[X \mid \mathcal{F}_0\right]$ is $\mathcal{F}_0$-measurable,

(2) $\forall\ E \in \mathcal{F}_0$ we have $\mathbb{E}\left[X \mathbb{1}_E\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_0\right] \mathbb{1}_E\right]$.

where $\mathbb{1}$ corresponds to the indicator function.

Theorem A.1.11 (Conditional expectation under change of measure)**.** *Let $(\mathcal{X}, \mathcal{F})$ be a measurable space on which two finite measures μ and ν are defined. Let $\nu << \mu$ and $\frac{d\nu}{d\mu}$ be the Radon-Nikodym derivative of ν with respect to μ and let $\mathcal{F}_0 \subset \mathcal{F}$ be a sub-σ-algebra then for any random variable X we have*

$$\mathbb{E}^{\nu}\left[X \mid \mathcal{F}_0\right] = \frac{\mathbb{E}^{\mu}\left[X \frac{d\nu}{d\mu} \mid \mathcal{F}_0\right]}{\mathbb{E}^{\mu}\left[\frac{d\nu}{d\mu} \mid \mathcal{F}_0\right]}.$$

Definition A.1.12. Let $\mathcal{T}$ be a index set. We say an index set has a total order if any two elements part of the index set are comparable. In other words, for any elements x, y, z in $\mathcal{T}$, the following conditions are satisfied:

(1) $x \leqslant x$,

(2) $x \leqslant y$ and $y \leqslant z \implies x \leqslant z$,
(3) $x \leqslant y$ and $y \leqslant x \implies x = y$,
(4) either $x \leqslant y$ or $y \leqslant x$.

If the index set is countable, we say it is discrete and if the index set is uncountable then we say it is continuous.

Definition A.1.13. A stochastic process denoted as $\{X_t\}_{t\in\mathcal{T}}$ is a family of random variables X_t indexed by an index set $\mathcal{T}$ with a total order. If the index set is countable then we say this stochastic process is in discrete time and if the index set is uncountable then we say this stochastic process is in continuous time.

Definition A.1.14. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2$,... be a series of sub-$\sigma$-algebras of $\mathcal{F}$. Let $\mathcal{I}$ be a discrete index set with a total order then we say $\mathbb{F} = \{\mathcal{F}_i\}_{i\in I}$ is a filtration if $\forall n \geqslant 1$

$$\mathcal{F}_{n-1} \subset \mathcal{F}_n.$$

We denote $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ as a filtered probability space.

Remark A.1.15. If $\mathcal{I}$ was a continuous index set with a total order then we say $\mathbb{F} = \{\mathcal{F}_i\}_{i\in I}$ is a filtration if $\forall\ s \in \mathbb{R}^+$ and $i \in \mathcal{I}$

$$\mathcal{F}_i \subset \mathcal{F}_{i+s}$$

Definition A.1.16. Let $\mathbb{F} = \{\mathcal{F}_i\}_{i\in I}$ be a filtration then we say $\mathbb{F}$ is right continuous if

$$\mathbb{F} = \mathbb{F}^+,$$

where

$$\mathbb{F}^+ = \left\{\mathcal{F}_i^+\right\}_{i\in I}, \qquad \mathcal{F}_i^+ = \cap_{i<s}\mathcal{F}_s.$$

Definition A.1.17. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let

$$\mathcal{N}_{\mathbb{P}} = \{E \subset \Omega \mid \exists G \in \mathcal{F} \text{ such that } E \subset G \text{ and } \mathbb{P}(G) = 0\},$$

be the set containing all sets of measure zero in $\mathcal{F}$. We call this set the null set of $\mathbb{P}$. A filtration is called complete if $\mathcal{F}_0$ contains $\mathcal{N}_{\mathbb{P}}$.

Remark A.1.18. A filtration $\mathbb{F}$ is called augmented if it is complete and right continuous. Instead of mentioning that a filtration is augmented, it is usually mentioned that the filtration satisfies the usual hypotheses.

Definition A.1.19. Let $\{X_t\}_{t\in\mathcal{T}}$ be a stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\sigma(X_n \mid n \leqslant t)$ denote the σ-algebra then let

$$\mathcal{F}_t = \sigma(X_n \mid n \leqslant t),$$

be this σ-algebra and $\mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathcal{T}}$ be a filtration. We call this filtration $\mathbb{F}$ the natural filtration of $\mathcal{F}$ with respect to $\{X_t\}_{t\in\mathcal{T}}$.

Definition A.1.20. Let $\{X_t\}_{t\in\mathcal{T}}$ be a stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We say that the stochastic processes is adapted if $\forall\, t \in \mathcal{T}\ X_t$ is $\mathcal{F}_t$-measurable.

Definition A.1.21. Let $\{X_t\}_{t\in\mathcal{T}}$ be a discrete stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We say that the stochastic processes is predictable if $\forall\ t \in \mathcal{T}\ X_t$ is $\mathcal{F}_{t-1}$-measurable.[4]

Definition A.1.22. Let $\{X_t\}_{t\in\mathcal{T}}$ be a discrete stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We say that $\{X_t\}_{t\in\mathbb{N}}$ is a martingale if

(1) $\{X_t\}_{t\in\mathbb{N}}$ is adapted to $\mathbb{F}$,

(2) $\forall\, t \in \mathcal{T}$: $X_t \in \mathcal{L}^1$,

(3) $\forall\, t \in \mathcal{T}$: $\mathbb{E}\left[X_t \mid X_{t-1}\right] = X_{t-1}$ almost surely.

Remark A.1.23. In continuous time, we simply replace the index set $\mathcal{I}$ by a continuous time equivalent (e.g. $\mathbb{N}$ with $\mathbb{R}^+$) and obtain a modified version of the third condition

$$\forall\ s > 0 : \mathbb{E}\left[X_{t+s} \mid X_t\right] = X_t \text{ almost surely.}$$

A.2 Linear Algebra Notions

This section includes useful results and theorems for linear algebra.

Theorem A.2.1. *Let $A \in \mathbb{R}^{n\times n}$ with spectral radius $p(A) = \max(|\lambda_1|,\ldots,|\lambda_n|)$ then if $p(A) < 1$ then the following identities hold*

$$\sum_{i=0}^{\infty} A^i = (I - A)^{-1}, \tag{A.2.1}$$

$$\sum_{i=0}^{N-1} A^i = (I - A)^{-1}(I - A^N). \tag{A.2.2}$$

Theorem A.2.2. *(See Wikipedia contributors, 2024a, for more details) Let $\boldsymbol{A}$, $\boldsymbol{B} \in \mathbb{R}^{m\times n}$ be real matrices and $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ be column vectors then we have the following results for $m,n \in \mathbb{R}$:*

(1) $\operatorname{vec}(\boldsymbol{A}) = \left[a_{1,1}, \ldots a_{n,1},\ldots,a_{1,2},\ldots,a_{n,2},\ldots,a_{1,m},\ldots,a_{n,m}\right]^\top$,

(2) $\boldsymbol{ab}^\top = \boldsymbol{a} \otimes \boldsymbol{b}^\top$,

(3) $\operatorname{vec}(\boldsymbol{ab}^\top) = \boldsymbol{b} \otimes \boldsymbol{a}$,

[4] Defining a predictable stochastic process in continuous time is not as simple as in discrete time. Since we will focus on discrete time, the reader can refer to this online resource for more information.

where vec *corresponds to the vectorization or flattening operation of a matrix obtained by stacking the columns of a matrix on top of each other and* $\otimes$ *corresponds to the Kronecker product. Now let* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times q}$ *and* $\boldsymbol{C} \in \mathbb{R}^{q\times r}$ *where* $m,n,q,r \in \mathbb{R}$ *be real matrices then we have the following identity:*

$$\operatorname{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}^\top \otimes \boldsymbol{A}) \operatorname{vec}(\boldsymbol{B}). \tag{A.2.3}$$

Theorem A.2.3. *(See Result 17.25 in Seber, 2007, for more details) Suppose* $y = \boldsymbol{w}^\top \boldsymbol{A}\boldsymbol{z} \in \mathbb{R}$, *where* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *is a real matrix,* $\boldsymbol{w} \in \mathbb{R}^m$ *and* $\boldsymbol{z} \in \mathbb{R}^n$ *are real column vectors, and* $\boldsymbol{A}, \boldsymbol{w}, \boldsymbol{z}$ *are all functions of* $\boldsymbol{x} \in \mathbb{R}^k$ *where* $m,n,k \in \mathbb{R}$. *We wish to find the column vector* $\frac{\partial y}{\partial \boldsymbol{x}}$. *We first note using the identities in Theorem A.2.2 that*

$$y = \operatorname{vec}(y) = (\boldsymbol{z}^\top \otimes \boldsymbol{w}^\top) \operatorname{vec}(\boldsymbol{A}) = (\boldsymbol{z} \otimes \boldsymbol{w})^\top \operatorname{vec}(\boldsymbol{A}) = [vec(\boldsymbol{w}\boldsymbol{z}^\top)]^\top \operatorname{vec}(\boldsymbol{A}),$$

then using $w^\top A z = z^\top A^\top w$, *we get*

$$\frac{\partial y}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{w}^\top}{\partial \boldsymbol{x}} \boldsymbol{A}\boldsymbol{z} + \frac{\partial \operatorname{vec}(\boldsymbol{A})^\top}{\partial \boldsymbol{x}} \operatorname{vec}(\boldsymbol{w}\boldsymbol{z}^\top) + \frac{\partial \boldsymbol{z}^\top}{\partial \boldsymbol{x}} \boldsymbol{A}^\top \boldsymbol{w}.$$

Appendix B

One-step Derivatives

B.1 One-step derivatives under HNSV

Using the explicit equations presented in Equations (3.2.10) through (3.2.11) , we can derive the one-step derivatives under $\mathbb{Q}$ and $\widetilde{\mathbb{Q}}_t$ under HNSV modelling assumptions given $g(u) = u$ under $\mathbb{Q}$ and $g(u) = 1 + u$ under $\widetilde{\mathbb{Q}}_t$. Thus, they are respectively evaluated at $u = 0, v = 0$ under $\mathbb{Q}$ and at $u = 1, v = 0$ under $\widetilde{\mathbb{Q}}_t$. The necessary notation to understand the use of the function g and the one-step derivative expressions in this section can be found at the start of Appendix D.3.

B.1.1 First Order Partial Derivatives

We have

$$\begin{aligned}
\mathcal{O}^*_{10}(\boldsymbol{h}_{t+1}(0)) &= \left[-\frac{1}{2} + \frac{2u\left(1-\rho^{*2}\right)}{2} + \frac{u\rho^{*2} - 2v\alpha^* q^* \rho^*}{1-2\alpha^* v}\right]_{u=g(0),v=0} = -\frac{1}{2} + g(0), \\
\mathcal{O}^*_{01}(\boldsymbol{h}_{t+1}(0)) &= \left[\beta + \alpha^* q^{*2} - \frac{2\alpha^* q^*\left(u\rho^* - 2v\alpha^* q^*\right)}{1-2v\alpha^*} + \frac{\alpha^*\left(u\rho^* - 2v\alpha^* q^*\right)^2}{\left(1-2v\alpha^*\right)^2}\right]_{u=g(0),v=0} \\
&= \beta + \alpha^* q^{*2} + \alpha^* \rho^* g(0)\left(1-2q^*\right),
\end{aligned}$$

and

$$\mathcal{M}^*_{10}(\boldsymbol{h}_{t+1}(0)) = r, \qquad \mathcal{M}^*_{01}(\boldsymbol{h}_{t+1}(0)) = \left[\omega^* + \frac{\alpha^*}{1-2v\alpha^*}\right]_{u=g(0),v=0} = \omega^* + \alpha^*.$$

B.1.2 Second Order Partial Derivatives

We have

$$\mathcal{O}^*_{20}(\boldsymbol{h}_{t+1}(0)) = \left[1 - \rho^{*2} + \frac{\rho^{*2}}{1-2\alpha^* v}\right]_{u=g(0),v=0} = 1,$$

$$\mathcal{O}^*_{11}(\boldsymbol{h}_{t+1}(0)) = \left[-\frac{2\alpha^* q^* \rho^*}{1-2\alpha^* v} + \frac{2\alpha^*\left(u\rho^{*2} - 2v\alpha^* q^* \rho^*\right)}{\left(1-2\alpha^* v\right)^2}\right]_{u=g(0),v=0} = 2\alpha^* g(0)\rho^{*2} - 2\alpha^* q^* \rho^*,$$

$$\begin{aligned}\mathcal{O}^*_{02}(\boldsymbol{h}_{t+1}(0)) &= \left[\frac{4\alpha^{*2} q^{*2}}{1-2v\alpha^*} - \frac{8\alpha^{*2} q^*\left(u\rho^* - 2v\alpha^* q^*\right)}{(1-2v\alpha^*)^2} + \frac{4\alpha^{*2}\left(u\rho^* - 2v\alpha^* q^*\right)^2}{\left(1-2v\alpha^*\right)^3}\right]_{u=g(0),v=0} \\ &= 4\alpha^{*2} q^{*2} + 4\alpha^{*2}\rho^* g(0)\left(1-2q^*\right),\end{aligned}$$

and

$$\begin{aligned}&\mathcal{M}^*_{20}(\boldsymbol{h}_{t+1}(0)) = 0, \qquad \mathcal{M}^*_{11}(\boldsymbol{h}_{t+1}(0)) = 0, \\ &\mathcal{M}^*_{02}(\boldsymbol{h}_{t+1}(0)) = \left[\frac{2\alpha^{*2}}{(1-2v\alpha^*)^2}\right]_{u=g(0),v=0} = 2\alpha^{*2},\end{aligned}$$

B.1.3 Third Order Partial Derivatives

We have

$$\begin{aligned}\mathcal{O}^*_{30}(\boldsymbol{h}_{t+1}(0)) &= 0, \qquad \mathcal{O}^*_{21}(\boldsymbol{h}_{t+1}(0)) = 0 \\ \mathcal{O}^*_{12}(\boldsymbol{h}_{t+1}(0)) &= \left[-\frac{8\alpha^{*2} q^* \rho^*}{\left(1-2\alpha^* v\right)^2} + \frac{8\alpha^{*2}\left(u\rho^{*2} - 2v\alpha^* q^* \rho^*\right)}{\left(1-2\alpha^* v\right)^3}\right]_{u=g(0),v=0} \\ &= 8\alpha^{*2} g(0)\rho^{*2} - 8\alpha^{*2} q^* \rho^*, \\ \mathcal{O}^*_{03}(\boldsymbol{h}_{t+1}(0)) &= \left[\frac{24\alpha^{*3} q^{*2}}{\left(1-2v\alpha^*\right)^2} - \frac{48\alpha^{*3} q^*\left(u\rho^* - 2v\alpha^* q^*\right)}{(1-2v\alpha^*)^3} + \frac{24\alpha^{*3}\left(u\rho^* - 2v\alpha^* q^*\right)^2}{\left(1-2v\alpha^*\right)^4}\right]_{u=g(0),v=0} \\ &= 24\alpha^{*3} q^{*2} + 24\alpha^{*3}\rho^* g(0)\left(1-2q^*\right),\end{aligned}$$

and

$$\begin{aligned}&\mathcal{M}^*_{30}(\boldsymbol{h}_{t+1}(0)) = 0, \qquad \mathcal{M}^*_{21}(\boldsymbol{h}_{t+1}(0)) = 0, \qquad \mathcal{M}^*_{12}(\boldsymbol{h}_{t+1}(0)) = 0, \\ &\mathcal{M}^*_{03}(\boldsymbol{h}_{t+1}(0)) = \left[\frac{8\alpha^{*3}}{(1-2v\alpha^*)^3}\right]_{u=g(0),v=0} = 8\alpha^{*3}\end{aligned}$$

B.1.4 Fourth Order Partial Derivatives

We have

$$\begin{aligned}\mathcal{O}^*_{40}(\boldsymbol{h}_{t+1}(0)) &= 0, \qquad \mathcal{O}^*_{31}(\boldsymbol{h}_{t+1}(0)) = 0, \\ \mathcal{O}^*_{22}(\boldsymbol{h}_{t+1}(0)) &= \left[\frac{8\alpha^{*2}\rho^{*2}}{\left(1-2\alpha^* v\right)^3}\right]_{u=g(0),v=0} = 8\alpha^{*2}\rho^{*2}, \\ \mathcal{O}^*_{13}(\boldsymbol{h}_{t+1}(0)) &= \left[-\frac{48\alpha^{*3} q^* \rho^*}{\left(1-2\alpha^* v\right)^3} + \frac{48\alpha^{*3}\left(u\rho^{*2} - 2v\alpha^* q^* \rho^*\right)}{\left(1-2\alpha^* v\right)^4}\right]_{u=g(0),v=0}\end{aligned}$$

$$= 4\alpha^{*3} g(0)\rho^{*2} - 48\alpha^{*3} q^* \rho^*,$$

$$\mathcal{O}^*_{04}(\boldsymbol{h}_{t+1}(0)) = \left[\frac{192\alpha^{*4}q^{*2}}{(1-2v\alpha^*)^3} - \frac{384\alpha^{*4}q^*\left(u\rho^* - 2v\alpha^* q^*\right)}{(1-2v\alpha^*)^4} + \frac{192\alpha^{*4}\left(u\rho^* - 2v\alpha^* q^*\right)^2}{(1-2v\alpha^*)^5}\right]_{u=g(0),v=0}$$

$$= 192\alpha^{*4}q^{*2} + 192\alpha^{*4}\rho^* g(0)\left(1-2q^*\right),$$

and

$$\mathcal{M}^*_{40}(\boldsymbol{h}_{t+1}(0)) = 0, \mathcal{M}^*_{31}(\boldsymbol{h}_{t+1}(0)) = 0, \qquad \mathcal{M}^*_{22}(\boldsymbol{h}_{t+1}(0)) = 0, \qquad \mathcal{M}^*_{13}(\boldsymbol{h}_{t+1}(0)) = 0,$$

$$\mathcal{M}^*_{04}(\boldsymbol{h}_{t+1}(0)) = \left[\frac{48\alpha^{*4}}{(1-2v\alpha^*)^4}\right]_{u=g(0),v=0} = 48\alpha^{*4}$$

B.2 One-step derivatives under GARCH(C)

Using the explicit equations presented in Equations (3.3.17) through (3.3.20), we can derive the one-step derivatives under $\mathbb{Q}$ and $\widetilde{\mathbb{Q}}_t$ under the GARCH(C) model assumptions given $g(u) = u$ under $\mathbb{Q}$ and $g(u) = 1 + u$ under $\widetilde{\mathbb{Q}}_t$. Thus, they are respectively evaluated at $u = 0, v = 0$ under $\mathbb{Q}$ and at $u = 1, v = 0$ under $\widetilde{\mathbb{Q}}_t$. The necessary notation to understand the use of the function g and the one-step derivative expressions in this section can be found at the start of Appendix D.3. We also introduce the following notation to simplify the expressions of the one-step derivatives:

$$\phi = u - 2\boldsymbol{v}^\top \boldsymbol{K}^*_3 - \lambda^* - \frac{1}{2}, \qquad \tau = \frac{1}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2}, \qquad \psi = \lambda^* + \frac{1}{2} - g(0),$$

where we have

$$\frac{\partial}{\partial u}\left[\phi^k\right]_{u=g(0),\boldsymbol{v}} = \left[k\phi^{k-1}\right]_{u=g(0),\boldsymbol{v}} = k\left(-1\right)^{k-1}\psi^{k-1}, \qquad \frac{\partial}{\partial u}\tau^k = 0,$$

$$\frac{\partial}{\partial \boldsymbol{v}}\left[\phi^k\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = \left[-2k\phi^{k-1}\boldsymbol{K}^*_3\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \qquad \frac{\partial}{\partial \boldsymbol{v}}\left[\tau^k\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = 2k\tau^{k+1}\boldsymbol{K}^*_2.$$

$$= -2k\left(-1\right)^{k-1}\psi^{k-1}\boldsymbol{K}^*_3,$$

We will use these results throughout the derivations without referring to them. Furthermore, for the high-order partial derivatives with respect to $\boldsymbol{v}$, we will represent them as Kronecker product instead of tensors. This is common practice and simplifies notation extensively.

B.2.1 First Order Partial Derivatives

We have

$$\mathcal{O}^*_{1,10}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial}{\partial u}\,\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}}$$

$$
\begin{aligned}
&= \frac{\partial}{\partial u}\left[u\lambda^* + \boldsymbol{v}^\top \boldsymbol{K}_4^* - \frac{1}{2}\left(\lambda^* + \frac{1}{2}\right)^2 + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2})^2}{2(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)}\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \left[\lambda^* + \frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2}}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= g(0) - \frac{1}{2},
\end{aligned}
$$

and

$$
\begin{aligned}
\boldsymbol{\mathcal{O}}_{1,01}^*(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial}{\partial \boldsymbol{v}}\, \mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1)\big|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[u\lambda^* + \boldsymbol{v}^\top \boldsymbol{K}_4^* - \frac{1}{2}\left(\lambda^* + \frac{1}{2}\right)^2 + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2})^2}{2(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)}\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \left[\boldsymbol{K}_4^* - 2\frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2}}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\boldsymbol{K}_3^* + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2})^2}{(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)^2}\boldsymbol{K}_2^*\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \boldsymbol{K}_4^* + 2\left(\lambda^* + \frac{1}{2} - g(0)\right)\boldsymbol{K}_3^* + \left(\lambda^* + \frac{1}{2} - g(0)\right)^2 \boldsymbol{K}_2^* \\
&= \boldsymbol{K}_4^* + 2\psi \boldsymbol{K}_3^* + \psi^2 \boldsymbol{K}_2^*.
\end{aligned}
$$

Now similarly we obtain

$$
\mathcal{O}_{2,10}^*(\boldsymbol{h}_{t+1}(0)) = \left.\frac{\partial}{\partial u}\mathcal{B}_q^*(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} = g(0) - \frac{1}{2},
$$

and

$$
\boldsymbol{\mathcal{O}}_{2,01}^*(\boldsymbol{h}_{t+1}(0)) = \left.\frac{\partial}{\partial \boldsymbol{v}}\mathcal{B}_q^*(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} = \boldsymbol{K}_5^* + 2\psi \boldsymbol{K}_3^* + \psi^2 \boldsymbol{K}_2^*.
$$

Furthermore we have

$$
\begin{aligned}
\mathcal{M}_{10}^*(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial}{\partial u}\mathcal{A}^*(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial u}\left[ur + \boldsymbol{v}^\top \boldsymbol{K}_1^* - \frac{1}{2}\log(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= r,
\end{aligned}
$$

and

$$
\begin{aligned}
\boldsymbol{\mathcal{M}}_{01}^*(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial}{\partial \boldsymbol{v}}\mathcal{A}^*(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[ur + \boldsymbol{v}^\top \boldsymbol{K}_1^* - \frac{1}{2}\log(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \left[\boldsymbol{K}_1^* + \frac{\boldsymbol{K}_2^*}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}}
\end{aligned}
$$

$$= \boldsymbol{K}_1^* + \boldsymbol{K}_2^*.$$

B.2.2 Second Order Partial Derivatives

We have

$$\begin{aligned}
\mathcal{O}_{1,20}^*(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^2}{\partial u^2}\mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial u}\left[\frac{\partial}{\partial u}\mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1)\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial u}\left[\lambda^* + \frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2}}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \left[\frac{1}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= 1,
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{\mathcal{O}}_{1,11}^*(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^2}{\partial u \partial \boldsymbol{v}}\ \mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1)|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial}{\partial u}\mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1)\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\lambda^* + \frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2}}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \left[2\frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2}}{(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)^2}\boldsymbol{K}_2^* - \frac{2\boldsymbol{K}_3^*}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= -2\left(\lambda^* + \frac{1}{2} - g(0)\right)\boldsymbol{K}_2^* - 2\boldsymbol{K}_3^* \\
&= -2\psi \boldsymbol{K}_2^* - 2\boldsymbol{K}_3^*,
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{\mathcal{O}}_{1,02}^*(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^2}{\partial \boldsymbol{v}^2}\mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial}{\partial \boldsymbol{v}}\mathcal{B}_s^*(u, \boldsymbol{v}; t, t+1)\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\boldsymbol{K}_4^* - 2\frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2}}{1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*}\boldsymbol{K}_3^* + \frac{(u - 2\boldsymbol{v}^\top \boldsymbol{K}_3^* - \lambda^* - \frac{1}{2})^2}{(1 - 2\boldsymbol{v}^\top \boldsymbol{K}_2^*)^2}\boldsymbol{K}_2^*\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \left[4\phi^2\tau^3 \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} - 4\phi\tau^2 \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} + 4\tau \boldsymbol{K}_3^* \boldsymbol{K}_3^{*\top} - 4\phi\tau^2 \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top}\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= 4\psi^2 \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} + 4\psi \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} + 4\boldsymbol{K}_3^* \boldsymbol{K}_3^{*\top} + 4\psi \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top}.
\end{aligned}$$

Now similarly we obtain

$$\mathcal{O}^*_{2,20}(\boldsymbol{h}_{t+1}(0)) = \left.\frac{\partial^2}{\partial u^2}\mathcal{B}^*_q(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} = 1,$$

and

$$\boldsymbol{\mathcal{O}}^*_{2,11}(\boldsymbol{h}_{t+1}(0)) = \left.\frac{\partial^2}{\partial u \partial \boldsymbol{v}}\mathcal{B}^*_q(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} = -2\psi \boldsymbol{K}^*_2 - 2\boldsymbol{K}^*_3,$$

and

$$\begin{aligned}\boldsymbol{\mathcal{O}}^*_{2,02}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^2}{\partial \boldsymbol{v}^2}\mathcal{B}^*_q(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= 4\psi^2\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 4\psi\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3 + 4\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_3 + 4\psi\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2.\end{aligned}$$

Furthermore we have

$$\mathcal{M}^*_{20}(\boldsymbol{h}_{t+1}(0)) = \left.\frac{\partial^2}{\partial u^2}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} = 0$$

$$\boldsymbol{\mathcal{M}}^*_{11}(\boldsymbol{h}_{t+1}(0)) = \left.\frac{\partial^2}{\partial u \partial \boldsymbol{v}}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} = \mathbf{0}$$

and

$$\begin{aligned}\boldsymbol{\mathcal{M}}^*_{02}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^2}{\partial \boldsymbol{v}^2}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial}{\partial \boldsymbol{v}}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= \frac{\partial}{\partial \boldsymbol{v}}\left[\boldsymbol{K}^*_1 + \frac{\boldsymbol{K}^*_2}{1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2}\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= \left[\frac{2}{(1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^2}\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= 2\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\end{aligned}$$

B.2.3 Third Order Partial Derivatives

We have

$$\begin{aligned}\mathcal{O}^*_{1,30}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^3}{\partial u^3}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= \frac{\partial}{\partial u}\left[\frac{\partial^2}{\partial u^2}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= \frac{\partial}{\partial u}\left[\frac{1}{1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2}\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\ &= 0\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\mathcal{O}}^*_{1,21}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^3}{\partial u^2 \partial \boldsymbol{v}} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial u}\left[\frac{\partial^2}{\partial u \partial \boldsymbol{v}} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1)\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial u}\left[2\frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}^*_3 - \lambda^* - \frac{1}{2}}{(1-2\boldsymbol{v}^\top \boldsymbol{K}^*_2)^2}\boldsymbol{K}^*_2 - \frac{2\boldsymbol{K}^*_3}{1-2\boldsymbol{v}^\top \boldsymbol{K}^*_2}\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= \left[\frac{2}{(1-2\boldsymbol{v}^\top \boldsymbol{K}^*_2)^2}\boldsymbol{K}^*_2\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= 2\boldsymbol{K}^*_2
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\mathcal{O}}^*_{1,12}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^3}{\partial u \partial \boldsymbol{v}^2} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial^2}{\partial u \partial \boldsymbol{v}} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1)\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[2\frac{u - 2\boldsymbol{v}^\top \boldsymbol{K}^*_3 - \lambda^* - \frac{1}{2}}{(1-2\boldsymbol{v}^\top \boldsymbol{K}^*_2)^2}\boldsymbol{K}^*_2 - \frac{2\boldsymbol{K}^*_3}{1-2\boldsymbol{v}^\top \boldsymbol{K}^*_2}\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= \left[-4\tau^2 \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_2 + 8\phi\tau^3 \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_2 - 4\tau^2 \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_3\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&= -4\boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_2 - 8\psi \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_2 - 4\boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_3
\end{aligned}$$

and

$$\begin{aligned}
&\boldsymbol{\mathcal{O}}^*_{1,03}(\boldsymbol{h}_{t+1}(0)) \\
&\quad= \left.\frac{\partial^3}{\partial \boldsymbol{v}^3} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&\quad= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial^2}{\partial \boldsymbol{v}^2} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1)\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&\quad= \frac{\partial}{\partial \boldsymbol{v}}\left[4\phi^2\tau^3 \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_2 - 4\phi\tau^2 \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_3 + 4\tau \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_3 - 4\phi\tau^2 \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_2\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&\quad= -\left[16\phi\tau^3 \boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_2 - 24\phi^2\tau^4 \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_2 - 8\tau^2 \boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_3\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&\qquad - \left[16\phi\tau^3 \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_3 - 8\tau^2 \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_3 - 8\tau^2 \boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_2\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&\qquad - \left[16\phi\tau^3 \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_2\right]_{u=g(0), \boldsymbol{v}=\mathbf{0}} \\
&\quad= 16\psi \boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_2 + 24\psi^2 \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_2 + 8\boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_3 + 16\psi \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2 \boldsymbol{K}^{*\top}_3 \\
&\quad+ 8\boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_3 + 8\boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_2 + 16\psi \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_3 \boldsymbol{K}^{*\top}_2
\end{aligned}$$

Now similarly we obtain

$$\mathcal{O}^*_{2,30}(\boldsymbol{h}_{t+1}(0)) = \left.\frac{\partial^3}{\partial u^3} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\right|_{u=g(0), \boldsymbol{v}=\mathbf{0}} = 0$$

$$\boldsymbol{\mathcal{O}}^*_{2,21}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^3}{\partial u^2 \partial \boldsymbol{v}} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = 2\boldsymbol{K}^*_2$$

$$\begin{aligned}\boldsymbol{\mathcal{O}}^*_{2,12}(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^3}{\partial u \partial \boldsymbol{v}^2} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\ &= -4\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2 - 8\psi\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 - 4\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{\mathcal{O}}^*_{2,03}(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^3}{\partial \boldsymbol{v}^3} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\ &= 16\psi\boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 24\psi^2\boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 8\boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3 \\ &\quad + 16\psi\boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3 + 8\boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_3 + 8\boldsymbol{K}^*_3 \otimes \boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2 \\ &\quad + 16\psi\boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2\end{aligned}$$

Furthermore we have

$$\mathcal{M}^*_{30}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^3}{\partial u^3} \mathcal{A}^*(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = 0$$

$$\boldsymbol{\mathcal{M}}^*_{21}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^3}{\partial u^2 \partial \boldsymbol{v}} \mathcal{A}^*(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = \boldsymbol{0}$$

$$\boldsymbol{\mathcal{M}}^*_{12}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^3}{\partial u \partial \boldsymbol{v}^2} \mathcal{A}^*(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = \boldsymbol{0}$$

and

$$\begin{aligned}\boldsymbol{\mathcal{M}}^*_{03}(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^3}{\partial \boldsymbol{v}^3} \mathcal{A}^*(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\ &= \frac{\partial}{\partial \boldsymbol{v}} \left[\frac{\partial^2}{\partial \boldsymbol{v}^2} \mathcal{A}^*(u, \boldsymbol{v}; t, t+1)\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\ &= \frac{\partial}{\partial \boldsymbol{v}} \left[\frac{2}{(1 - 2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^2} \boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\ &= \left[\frac{8}{(1 - 2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^3} \boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\ &= 8\boldsymbol{K}^*_2 \otimes \boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\end{aligned}$$

B.2.4 Fourth Order Partial Derivatives

We have

$$\begin{aligned}
\mathcal{O}^*_{1,40}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^4}{\partial u^4}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial u}\left[\frac{\partial^3}{\partial u^3}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{\mathcal{O}}^*_{1,31}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^4}{\partial u^3\partial \boldsymbol{v}}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial u}\left[\frac{\partial^2}{\partial u^2\partial \boldsymbol{v}}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial u}\left[\frac{2}{(1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^2}\boldsymbol{K}^*_2\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \mathbf{0},
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\mathcal{O}}^*_{1,22}(\boldsymbol{h}_{t+1}(0)) &= \left.\frac{\partial^4}{\partial u^2\partial \boldsymbol{v}^2}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial^3}{\partial u^2\partial \boldsymbol{v}}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{2}{(1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^2}\boldsymbol{K}^*_2\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= \left[\frac{8}{(1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^3}\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&= 8\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2,
\end{aligned}$$

$$\begin{aligned}
&\boldsymbol{\mathcal{O}}^*_{1,13}(\boldsymbol{h}_{t+1}(0)) \\
&\quad= \left.\frac{\partial^4}{\partial u\partial \boldsymbol{v}^3}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right|_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&\quad= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial^3}{\partial u\partial \boldsymbol{v}^2}\mathcal{B}^*_s(u,\boldsymbol{v};t,t+1)\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&\quad= \frac{\partial}{\partial \boldsymbol{v}}\left[-4\tau^2\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2+8\phi\tau^3\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2-4\tau^2\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&\quad= \left[-16\tau^3\boldsymbol{K}_2\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2-16\tau^3\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2+48\phi\tau^4\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}} \\
&\qquad-\left[16\tau^3\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3\right]_{u=g(0),\boldsymbol{v}=\mathbf{0}}
\end{aligned}$$

$$= -16\boldsymbol{K}_2 \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} - 16\boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} - 48\psi \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} - 16\boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top},$$

$$\begin{aligned}
&\boldsymbol{\mathcal{O}}^*_{1,04}(\boldsymbol{h}_{t+1}(0)) \\
&\quad = \frac{\partial^4}{\partial \boldsymbol{v}^4} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} = \frac{\partial}{\partial \boldsymbol{v}} \left[\frac{\partial^3}{\partial \boldsymbol{v}^3} \mathcal{B}^*_s(u, \boldsymbol{v}; t, t+1) \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\quad = \frac{\partial}{\partial \boldsymbol{v}} \left[-16\phi\tau^3 \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} + 24\phi^2\tau^4 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} + 8\tau^2 \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\qquad - \left[16\phi\tau^3 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} - 8\tau^2 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_3^{*\top} - 8\tau^2 \boldsymbol{K}_3^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\qquad - \left[16\phi\tau^3 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\quad = \left[32\tau^3 \boldsymbol{K}_3^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} - 96\phi\tau^4 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\qquad - \left[96\phi\tau^4 \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} - 192\phi^2\tau^5 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\qquad + \left[32\tau^3 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} + 32\tau^3 \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\qquad - \left[96\phi\tau^4 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} - 32\tau^3 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_3^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\qquad + \left[32\tau^3 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} + 32\tau^3 \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\qquad - \left[96\phi\tau^4 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} \right]_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\quad = 32\boldsymbol{K}_3^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} + 96\psi \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} + 96\psi \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} \\
&\qquad + 192\psi^2 \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} + 32\boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} + 32\boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} \\
&\qquad + 96\psi \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top} + 32\boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_3^{*\top} + 32\boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} \\
&\qquad + 32\boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} + 96\psi \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top}.
\end{aligned}$$

Now similarly we obtain

$$\mathcal{O}^*_{2,40}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^4}{\partial u^4} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} = 0,$$

$$\boldsymbol{\mathcal{O}}^*_{2,31}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^4}{\partial u^3 \partial \boldsymbol{v}} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} = \boldsymbol{0},$$

$$\boldsymbol{\mathcal{O}}^*_{2,22}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^4}{\partial u^2 \partial \boldsymbol{v}^2} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} = 8\boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top},$$

$$\begin{aligned}
&\boldsymbol{\mathcal{O}}^*_{2,13}(\boldsymbol{h}_{t+1}(0)) \\
&\quad = \frac{\partial^4}{\partial u \partial \boldsymbol{v}^3} \mathcal{B}^*_q(u, \boldsymbol{v}; t, t+1)\Big|_{u=g(0), \boldsymbol{v}=\boldsymbol{0}} \\
&\quad = -16\boldsymbol{K}_2 \otimes \boldsymbol{K}_3^* \boldsymbol{K}_2^{*\top} - 16\boldsymbol{K}_3^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} - 48\psi \boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_2^{*\top} - 16\boldsymbol{K}_2^* \otimes \boldsymbol{K}_2^* \boldsymbol{K}_3^{*\top},
\end{aligned}$$

and

$$
\begin{aligned}
&\boldsymbol{\mathcal{O}}^*_{2,04}(\boldsymbol{h}_{t+1}(0)) \\
&\quad= \frac{\partial^4}{\partial \boldsymbol{v}^4}\mathcal{B}^*_q(u,\boldsymbol{v};t,t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&\quad= \frac{\partial}{\partial \boldsymbol{v}}\left[-16\phi\tau^3\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 24\phi^2\tau^4\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 8\tau^2\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&\qquad - \left[16\phi\tau^3\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3 - 8\tau^2\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_3 - 8\tau^2\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&\qquad - \left[16\phi\tau^3\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&\quad= 32\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 96\psi\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 96\psi\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 \\
&\qquad + 192\psi^2\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2 + 32\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3 + 32\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3 \\
&\qquad + 96\psi\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_3 + 32\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_3 + 32\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2 \\
&\qquad + 32\boldsymbol{K}^*_3\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2 + 96\psi\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_3\boldsymbol{K}^{*\top}_2.
\end{aligned}
$$

Furthermore we have

$$
\mathcal{M}^*_{40}(\boldsymbol{h}_{t+1}(0)) = \frac{\partial^4}{\partial u^4}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = 0,
$$

$$
\begin{aligned}
\boldsymbol{\mathcal{M}}^*_{31}(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^4}{\partial u^3\partial \boldsymbol{v}}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
\boldsymbol{\mathcal{M}}^*_{22}(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^4}{\partial u^2\partial \boldsymbol{v}^2}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
\boldsymbol{\mathcal{M}}^*_{13}(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^4}{\partial u\partial \boldsymbol{v}^3}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} = \boldsymbol{0},
\end{aligned}
$$

and

$$
\begin{aligned}
\boldsymbol{\mathcal{M}}^*_{04}(\boldsymbol{h}_{t+1}(0)) &= \frac{\partial^4}{\partial \boldsymbol{v}^4}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\Big|_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{\partial^3}{\partial \boldsymbol{v}^3}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \frac{\partial}{\partial \boldsymbol{v}}\left[\frac{8}{(1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^3}\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= \left[\frac{48}{(1-2\boldsymbol{v}^\top\boldsymbol{K}^*_2)^4}\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2\right]_{u=g(0),\boldsymbol{v}=\boldsymbol{0}} \\
&= 48\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\otimes\boldsymbol{K}^*_2\boldsymbol{K}^{*\top}_2.
\end{aligned}
$$

Appendix C

Cumulants of the log-price process

C.1 N-Dimensional cumulants

In this section, we will use the notation introduced in Appendix D.3 to simplify the proofs of the N-dimensional cumulants.

C.1.1 Cumulants under risk-neutral measure $\mathbb{Q}$

C.1.1.1 Cumulant 1

We have

$$\kappa_1^{\mathbb{Q}} = Y_t + \mathcal{A}_1^*(0;t,T) + \mathcal{B}_{0,1}^*(0;t,T)y_t + \boldsymbol{\mathcal{B}}_1^*(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.1}$$

where $\mathcal{A}_1^*(0;t,T)$, $\mathcal{B}_{0,1}^*(0;t,T)$ and $\boldsymbol{\mathcal{B}}_1^*(0;t,T)$ satisfy the recursions presented in Equations (D.3.10), (D.3.11), (D.3.12) with terminal constraints $\mathcal{A}_1^*(0;T,T) = 0$, $\mathcal{B}_{0,1}^*(0;T,T) = 0$ and $\boldsymbol{\mathcal{B}}_1^*(0;T,T) = \mathbf{0}$ and $g(u) = u$.

C.1.1.2 Cumulant 2

We have

$$\kappa_2^{\mathbb{Q}} = \mathcal{A}_2^*(0;t,T) + \mathcal{B}_{0,2}^*(0;t,T)y_t + \boldsymbol{\mathcal{B}}_2^*(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.2}$$

where $\mathcal{A}_2^*(0;t,T)$, $\mathcal{B}_{0,2}^*(0;t,T)$ and $\boldsymbol{\mathcal{B}}_2^*(0;t,T)$ satisfy the recursions presented in Equations (D.3.15), (D.3.17), (D.3.19) with terminal constraints $\mathcal{A}_2^*(0;T,T) = 0$, $\mathcal{B}_{0,2}^*(0;T,T) = 0$ and $\boldsymbol{\mathcal{B}}_2^*(0;T,T) = \mathbf{0}$ and $g(u) = u$.

C.1.1.3 Cumulant 3

We have

$$\kappa_3^{\mathbb{Q}} = \mathcal{A}_3^*(0;t,T) + \mathcal{B}_{0,3}^*(0;t,T)y_t + \boldsymbol{\mathcal{B}}_3^*(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.3}$$

where $\mathcal{A}_3^*(0;t,T)$, $\mathcal{B}_{0,3}^*(0;t,T)$ and $\boldsymbol{\mathcal{B}}_3^*(0;t,T)$ satisfy the recursions presented in Equations (D.3.21), (D.3.22), (D.3.24) with terminal constraints $\mathcal{A}_3^*(0;T,T)=0$, $\mathcal{B}_{0,3}^*(0;T,T)=0$ and $\boldsymbol{\mathcal{B}}_3^*(0;T,T)=\mathbf{0}$ and $g(u)=u$.

C.1.1.4 Cumulant 4

We have

$$\kappa_4^{\mathbb{Q}} = \mathcal{A}_4^*(0;t,T) + \mathcal{B}_{0,4}^*(0;t,T)y_t + \boldsymbol{\mathcal{B}}_4^*(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.4}$$

where $\mathcal{A}_4^*(0;t,T)$, $\mathcal{B}_{0,4}^*(0;t,T)$ and $\boldsymbol{\mathcal{B}}_4^*(0;t,T)$ satisfy the recursions presented in Equations (D.3.27), (D.3.29), (D.3.33) with terminal constraints $\mathcal{A}_4^*(0;T,T)=0$, $\mathcal{B}_{0,4}^*(0;T,T)=0$ and $\boldsymbol{\mathcal{B}}_4^*(0;T,T)=\mathbf{0}$ and $g(u)=u$.

C.1.2 Cumulants under measure $\widetilde{\mathbb{Q}}_t$

As shown in Appendix D.5, we recall that the $\mathcal{F}_t$-conditional CGF of Y_T under $\widetilde{\mathbb{Q}}_t$ is

$$\mathcal{C}_{Y_T}^{\widetilde{\mathbb{Q}}_t}(u \mid \mathcal{F}_t) = uYt - r + \mathcal{A}^{\widetilde{\mathbb{Q}}_t}(u,\mathbf{0};t,T) + \mathcal{B}_0^{\widetilde{\mathbb{Q}}_t}(u,\mathbf{0};t,T)\, y_t + \boldsymbol{\mathcal{B}}^{\widetilde{\mathbb{Q}}_t}(u,\mathbf{0};t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.5}$$

where the multi-step coefficients $\mathcal{A}^{\widetilde{\mathbb{Q}}_t}(u,\boldsymbol{v};t,T)$, $\mathcal{B}_0^{\widetilde{\mathbb{Q}}_t}(u,\boldsymbol{v};t,T)$ and $\boldsymbol{\mathcal{B}}^{\widetilde{\mathbb{Q}}_t}(u,\boldsymbol{v};t,T)$ are given by:

$$\mathcal{A}^{\widetilde{\mathbb{Q}}_t}(u,\boldsymbol{v};t,T) = \mathcal{A}^*(u,\boldsymbol{v};t+1,T) + \mathcal{A}^*\left(1+u+\mathcal{B}_0^*(u,\boldsymbol{v};t+1,T), \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T); t,t+1\right),$$

$$\mathcal{B}_0^{\widetilde{\mathbb{Q}}_t}(u,\boldsymbol{v};t,T) = \mathcal{B}_0^*\left(1+u+\mathcal{B}_0^*(u,\boldsymbol{v};t+1,T), \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T); t,t+1\right),$$

$$\boldsymbol{\mathcal{B}}^{\widetilde{\mathbb{Q}}_t}(u,\boldsymbol{v};t,T) = \boldsymbol{\mathcal{B}}^*\left(1+u+\mathcal{B}_0^*(u,\boldsymbol{v};t+1,T), \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T); t,t+1\right),$$

with terminal conditions $\mathcal{A}^*(u,\boldsymbol{v};T,T)=0$, $\mathcal{B}_0^*(u,\boldsymbol{v};T,T)=0$ and $\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};T,T)=\boldsymbol{v}$. We can now obtain closed-form formulas for the cumulants under measure $\widetilde{\mathbb{Q}}$ using vector calculus and the chain rule.

C.1.2.1 Cumulant 1

We have

$$\kappa_1^{\widetilde{\mathbb{Q}}_t} = Y_t + \mathcal{A}_1^{\widetilde{\mathbb{Q}}_t}(0;t,T) + \mathcal{B}_{0,1}^{\widetilde{\mathbb{Q}}_t}(0;t,T)\, y_t + \boldsymbol{\mathcal{B}}_1^{\widetilde{\mathbb{Q}}_t}(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.6}$$

where the multi-step coefficients $\mathcal{A}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T)$, $\mathcal{B}_{0,1}^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ and $\boldsymbol{\mathcal{B}}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ are given by:

$$\mathcal{A}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \mathcal{A}_1^*(u;t+1,T) + \left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right)\mathcal{M}_{10}^*(u) + \boldsymbol{\mathcal{M}}_{01}^*(u)\boldsymbol{\mathcal{B}}_1^*(u;t+1,T),$$

$$\mathcal{B}_{0,1}^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right)\mathcal{N}_{10}^*(u) + \boldsymbol{\mathcal{N}}_{01}^*(u)\boldsymbol{\mathcal{B}}_1^*(u;t+1,T),$$

$$\boldsymbol{\mathcal{B}}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right)\boldsymbol{\mathcal{O}}_{10}^*(u) + \boldsymbol{\mathcal{O}}_{01}^*(u)\boldsymbol{\mathcal{B}}_1^*(u;t+1,T),$$

and where $\mathcal{A}_1^*(u;t,T)$, $\mathcal{B}_{u,1}^*(u;t,T)$ and $\boldsymbol{\mathcal{B}}_1^*(0;t,T)$ satisfy the recursions presented in Equations (D.3.10), (D.3.11), (D.3.12) with terminal constraints $\mathcal{A}_1^*(u;T,T)=0$, $\mathcal{B}_{0,1}^*(u;T,T)=0$ and $\boldsymbol{\mathcal{B}}_1^*(u;T,T)=\mathbf{0}$ and $g(u)=1+u$.

C.1.2.2 Cumulant 2

We have

$$\kappa_2^{\widetilde{\mathbb{Q}}_t} = \mathcal{A}_2^{\widetilde{\mathbb{Q}}_t}(0;t,T) + \mathcal{B}_{0,2}^{\widetilde{\mathbb{Q}}_t}(0;t,T)\,y_t + \boldsymbol{\mathcal{B}}_2^{\widetilde{\mathbb{Q}}_t}(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.7}$$

where the multi-step coefficients $\mathcal{A}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T)$, $\mathcal{B}_{0,2}^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ and $\boldsymbol{\mathcal{B}}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ are given by:

$$\begin{aligned}
\mathcal{A}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T) = {}& \mathcal{A}_2^*(u;t+1,T) + \mathcal{M}_{20}^*(u)\left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right)^2 \\
& + 2\boldsymbol{\mathcal{M}}_{11}^*(u)^\top \boldsymbol{\mathcal{B}}_1^*(u;t+1,T)\left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right) \\
& + \boldsymbol{\mathcal{B}}_1^*(u;t+1,T)^\top \boldsymbol{\mathcal{M}}_{02}^*(u)\boldsymbol{\mathcal{B}}_1^*(u;t+1,T) \\
& + \mathcal{M}_{10}^*(u)\mathcal{B}_{0,2}^*(u;t+1,T) + \boldsymbol{\mathcal{M}}_{01}^*(u)\boldsymbol{\mathcal{B}}_2^*(u;t+1,T),
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{B}_{0,2}^{\widetilde{\mathbb{Q}}_t}(u;t,T) = {}& \mathcal{N}_{20}^*(u)\left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right)^2 + 2\boldsymbol{\mathcal{N}}_{11}^*(u)^\top \boldsymbol{\mathcal{B}}_1^*(u;t+1,T)\left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right) \\
& + \boldsymbol{\mathcal{B}}_1^*(u;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(u)\boldsymbol{\mathcal{B}}_1^*(u;t+1,T) \\
& + \mathcal{N}_{10}^*(u)\mathcal{B}_{0,2}^*(u;t+1,T) + \boldsymbol{\mathcal{N}}_{01}^*(u)\boldsymbol{\mathcal{B}}_2^*(u;t+1,T),
\end{aligned}$$

and

$$\boldsymbol{\mathcal{B}}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \begin{bmatrix} H_{\mathcal{B}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T)} & \cdots & H_{\mathcal{B}_N^{\widetilde{\mathbb{Q}}_t}(u;t,T)} \end{bmatrix}^\top,$$

and where $\mathcal{A}_2^*(u;t,T)$, $\mathcal{B}_{0,2}^*(u;t,T)$ and $\boldsymbol{\mathcal{B}}_2^*(u;t,T)$ satisfy the recursions presented in Equations (D.3.15), (D.3.17), (D.3.19) with terminal constraints $\mathcal{A}_2^*(u;T,T)=0$, $\mathcal{B}_{0,2}^*(u;T,T)=0$ and $\boldsymbol{\mathcal{B}}_2^*(u;T,T)=\mathbf{0}$. Finally, $g(u)=1+u$ and for $k=1,\dots,N$ we have:

$$\begin{aligned}
H_{\mathcal{B}_k^{\widetilde{\mathbb{Q}}_t}(u;t,T)} = {}& \mathcal{O}_{k,20}^*(u)\left(2+\mathcal{B}_{0,1}^*(u;t+1,T)\right)^2 + 2\boldsymbol{\mathcal{O}}_{k,11}^*(u)^\top \boldsymbol{\mathcal{B}}_1^*(u;t+1,T)\left(1+\mathcal{B}_{0,1}^*(u;t+1,T)\right) \\
& + \boldsymbol{\mathcal{B}}_1^*(u;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,02}^*(u)\boldsymbol{\mathcal{B}}_1^*(u;t+1,T) + \mathcal{O}_{k,10}^*(u)\mathcal{B}_{0,2}^*(x;t+1,T) \\
& + \boldsymbol{\mathcal{O}}_{k,01}(u)^*\boldsymbol{\mathcal{B}}_2^*(u;t+1,T).
\end{aligned}$$

C.1.2.3 Cumulant 3

We have

$$\kappa_3^{\widetilde{\mathbb{Q}}_t} = \mathcal{A}_3^{\widetilde{\mathbb{Q}}_t}(0;t,T) + \mathcal{B}_{0,3}^{\widetilde{\mathbb{Q}}_t}(0;t,T)\,y_t + \boldsymbol{\mathcal{B}}_3^{\widetilde{\mathbb{Q}}_t}(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.1.8}$$

where the multi-step coefficients $\mathcal{A}_3^{\widetilde{\mathbb{Q}}_t}(u;t,T)$, $\mathcal{B}_{0,3}^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ and $\boldsymbol{\mathcal{B}}_3^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ are given by:

$$\begin{aligned}
\mathcal{A}_3^{\widetilde{\mathbb{Q}}_t}(u;t,T) = {}& \mathcal{A}_3^*(x;t+1,T) + 3\mathcal{B}_{0,2}^*(x;t+1,T)\mathcal{M}_{20}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \\
& + 3\mathcal{B}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{M}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
& + 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
& + 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{11}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)
\end{aligned}$$

$$
\begin{aligned}
&+ \mathcal{M}_{30}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^3 \\
&+ 3\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^2 \boldsymbol{\mathcal{M}}_{21}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 3\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x) \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ \mathrm{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \,\mathrm{vec}\left(\boldsymbol{\mathcal{M}}_{03}^*(x)\right) \\
&+ \mathcal{M}_{10}^*(x)\mathcal{B}_{0,3}^*(x;t+1,T) + \boldsymbol{\mathcal{M}}_{01}^*(x)\boldsymbol{\mathcal{B}}_3^*(x;t+1,T),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{B}_{0,3}^{\widetilde{\mathbb{Q}}_t}(u;t,T) = {}& 3\mathcal{B}_{0,2}^*(x;t+1,T)\mathcal{N}_{20}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \\
&+ 3\left(\frac{\partial^2 g}{\partial u^2}(x) + \mathcal{B}_{0,2}^*(x;t+1,T)\right)\boldsymbol{\mathcal{N}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{11}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \\
&+ \mathcal{N}_{30}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^3 \\
&+ 3\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^2 \boldsymbol{\mathcal{N}}_{21}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 3\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{12}^*(x) \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ \mathrm{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \,\mathrm{vec}\left(\boldsymbol{\mathcal{N}}_{03}^*(x)\right) \\
&+ \mathcal{N}_{10}^*(x)\mathcal{B}_{0,3}^*(x;t+1,T) + \boldsymbol{\mathcal{N}}_{01}^*(x)\boldsymbol{\mathcal{B}}_3^*(x;t+1,T),
\end{aligned}
$$

and

$$
\boldsymbol{\mathcal{B}}_3^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \left[\Upsilon_{\mathcal{B}_1^{\widetilde{\mathbb{Q}}_t}(x;t,T)} \quad \cdots \quad \Upsilon_{\mathcal{B}_N^{\widetilde{\mathbb{Q}}_t}(x;t,T)}\right]^\top,
$$

and where $\mathcal{A}_3^*(u;t,T)$, $\mathcal{B}_{0,3}^*(u;t,T)$ and $\boldsymbol{\mathcal{B}}_3^*(u;t,T)$ satisfy the recursions presented in Equations (D.3.21), (D.3.22), (D.3.24) with terminal constraints $\mathcal{A}_3^*(u;T,T)=0$, $\mathcal{B}_{0,3}^*(u;T,T)=0$ and $\boldsymbol{\mathcal{B}}_3^*(u;T,T)=\mathbf{0}$. Finally, $g(u)=1+u$ and for $K=1,...,N$ we have:

$$
\begin{aligned}
\Upsilon_{\mathcal{B}_k^{\widetilde{\mathbb{Q}}_t}(x;t,T)} = {}& 3\mathcal{B}_{0,2}^*(x;t+1,T)\mathcal{O}_{k,20}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \\
&+ 3\mathcal{B}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{O}}_{k,11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,11}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \\
&+ \mathcal{O}_{k,30}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^3 \\
&+ 3\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^2 \boldsymbol{\mathcal{O}}_{k,21}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 3\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,12}^*(x) \boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ \mathrm{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \,\mathrm{vec}\left(\boldsymbol{\mathcal{O}}_{k,03}^*(x)\right)
\end{aligned}
$$

$$+ \mathcal{O}^*_{k,10}(x) + \mathcal{B}^*_{0,3}(x;t+1,T) + \boldsymbol{\mathcal{O}}^*_{k,01}(x)\boldsymbol{\mathcal{B}}^*_3(x;t+1,T).$$

C.1.2.4 Cumulant 4

We have

$$\kappa_4^{\widetilde{\mathbb{Q}}_t} = \mathcal{A}_4^{\widetilde{\mathbb{Q}}_t}(0;t,T) + \mathcal{B}_{0,4}^{\widetilde{\mathbb{Q}}_t}(0;t,T)\,y_t + \boldsymbol{\mathcal{B}}_4^{\widetilde{\mathbb{Q}}_t}(0;t,T)^\top \boldsymbol{l}^*_t,$$

where the multi-step coefficients $\mathcal{A}_4^{\widetilde{\mathbb{Q}}_t}(u;t,T)$, $\mathcal{B}_{0,4}^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ and $\boldsymbol{\mathcal{B}}_4^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ are given by:

$$\begin{aligned}
&\mathcal{A}_4^{\widetilde{\mathbb{Q}}_t}(u;t,T)\\
&= \mathcal{A}^*_4(x;t+1,T) + 4\mathcal{B}^*_{0,3}(x;t+1,T)\mathcal{M}^*_{20}(x)\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\\
&\quad + 4\mathcal{B}^*_{0,3}(x;t+1,T)\boldsymbol{\mathcal{M}}^*_{11}(x)^\top\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}^*_3(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{02}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}^*_3(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{11}(x)\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\\
&\quad + 3\mathcal{B}^*_{0,2}(x;t+1,T)^2\mathcal{M}^*_{20}(x) + 6\mathcal{B}^*_{0,2}(x;t+1,T)\boldsymbol{\mathcal{M}}^*_{11}(x)^\top\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)\\
&\quad + 3\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{02}(x)\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)\\
&\quad + 6\mathcal{M}^*_{30}(x)\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^2 + \mathcal{B}^*_{0,2}(x;t+1,T)\\
&\quad + 6\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^2\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{21}(x)\\
&\quad + 6 + \mathcal{B}^*_{0,2}(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{12}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\quad + 12\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\mathcal{B}^*_{0,2}(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{21}(x)\\
&\quad + 12\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{12}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\quad + 6\,\mathrm{vec}(\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top)^\top\,\mathrm{vec}\left(\boldsymbol{\mathcal{M}}^*_{03}(x)\right)\\
&\quad + \left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^4\mathcal{M}^*_{40}(x)\\
&\quad + 4\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^3\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{31}(x)\\
&\quad + 6\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^2\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{22}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\quad + 4\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\mathrm{vec}\left(\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\right)^\top\mathrm{vec}\left(\boldsymbol{\mathcal{M}}^*_{13}(x)\right)\\
&\quad + \mathrm{vec}\left(\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\right)^\top\mathrm{vec}\left(\boldsymbol{\mathcal{M}}^*_{04}(x)\right)\\
&\quad + \mathcal{M}^*_{10}(x)\mathcal{B}^*_{0,4}(x;t+1,T) + \boldsymbol{\mathcal{M}}^*_{01}(x)\boldsymbol{\mathcal{B}}^*_4(x;t+1,T),
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{B}_{0,4}^{\widetilde{\mathbb{Q}}_t}(u;t,T)\\
&= 4\mathcal{B}^*_{0,3}(x;t+1,T)\mathcal{N}^*_{20}(x)\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\\
&\quad + 4\mathcal{B}^*_{0,3}(x;t+1,T)\boldsymbol{\mathcal{N}}^*_{11}(x)^\top\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)
\end{aligned}$$

$$+ 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$
$$+ 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{11}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)$$
$$+ 3 + \mathcal{B}_{0,2}^*(x;t+1,T)^2\mathcal{N}_{20}^*(x) + 6 + \mathcal{B}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{N}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)$$
$$+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)$$
$$+ 6\mathcal{N}_{30}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^2 + \mathcal{B}_{0,2}^*(x;t+1,T)$$
$$+ 6\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^2 \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{21}^*(x)$$
$$+ 6 + \mathcal{B}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$
$$+ 12\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right) + \mathcal{B}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{21}^*(x)$$
$$+ 12\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$
$$+ 6\,\mathrm{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top\,\mathrm{vec}\left(\boldsymbol{\mathcal{N}}_{03}^*(x)\right)$$
$$+ \left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^4 \mathcal{N}_{40}^*(x)$$
$$+ 4\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^3 \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{31}^*(x)$$
$$+ 6\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^2 \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{22}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$
$$+ 4\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)\mathrm{vec}\left(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\right)^\top \mathrm{vec}\left(\boldsymbol{\mathcal{N}}_{13}^*(x)\right)$$
$$+ \mathrm{vec}\left(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\right)^\top \mathrm{vec}\left(\boldsymbol{\mathcal{N}}_{04}^*(x)\right)$$
$$+ \mathcal{N}_{10}^*(x) + \mathcal{B}_{0,4}^*(x;t+1,T) + \boldsymbol{\mathcal{N}}_{01}^*(x)\boldsymbol{\mathcal{B}}_4^*(x;t+1,T),$$

and

$$\boldsymbol{\mathcal{B}}_4^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \left[\Phi_{\mathcal{B}_1^{\widetilde{\mathbb{Q}}_t}(x;t,T)} \quad \cdots \quad \Phi_{\mathcal{B}_N^{\widetilde{\mathbb{Q}}_t}(x;t,T)},\right]^\top$$

where $\mathcal{A}_4^*(u;t,T)$, $\mathcal{B}_{0,4}^*(u;t,T)$ and $\boldsymbol{\mathcal{B}}_4^*(u;t,T)$ satisfy the recursions presented in Equations (D.3.27), (D.3.29), (D.3.33) with terminal constraints $\mathcal{A}_4^*(u;T,T)=0$, $\mathcal{B}_{0,4}^*(u;T,T)=0$ and $\boldsymbol{\mathcal{B}}_4^*(u;T,T)=\mathbf{0}$. Finally, $g(u)=1+u$ and for $k=1,\ldots,N$ we have:

$$\Phi_{\mathcal{B}_k^{\widetilde{\mathbb{Q}}_t}(x;t,T)}$$
$$= 4\mathcal{B}_{0,3}^*(x;t+1,T)\mathcal{O}_{k,20}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)$$
$$+ 4\mathcal{B}_{0,3}^*(x;t+1,T)\boldsymbol{\mathcal{O}}_{k,11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$
$$+ 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$
$$+ 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,11}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)$$
$$+ 3 + \mathcal{B}_{0,2}^*(x;t+1,T)^2\mathcal{O}_{k,20}^*(x) + 6\mathcal{B}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{O}}_{k,11}^*(x)^\top \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)$$
$$+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,02}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)$$
$$+ 6\mathcal{O}_{k,30}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)^2 \mathcal{B}_{0,2}^*(x;t+1,T))$$

$$+ 6\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^2 \boldsymbol{\mathcal{B}}^*_2(x;t+1,T)^\top \boldsymbol{\mathcal{O}}^*_{k,21}(x)$$
$$+ 6\mathcal{B}^*_{0,2}(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top \boldsymbol{\mathcal{O}}^*_{k,12}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)$$
$$+ 12\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\mathcal{B}^*_{0,2}(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top \boldsymbol{\mathcal{O}}^*_{k,21}(x)$$
$$+ 12\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)^\top \boldsymbol{\mathcal{O}}^*_{k,12}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)$$
$$+ 6\,\mathrm{vec}(\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}^*_2(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top)^\top\,\mathrm{vec}\left(\boldsymbol{\mathcal{O}}^*_{k,03}(x)\right)$$
$$+ \left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^4 \mathcal{O}^*_{k,40}(x)$$
$$+ 4\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^3 \boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top \boldsymbol{\mathcal{O}}^*_{k,31}(x)$$
$$+ 6\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)^2 \boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top \boldsymbol{\mathcal{O}}^*_{k,22}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)$$
$$+ 4\left(1+\mathcal{B}^*_{0,1}(x;t+1,T)\right)\mathrm{vec}\left(\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\right)^\top \mathrm{vec}\left(\boldsymbol{\mathcal{O}}^*_{k,13}(x)\right)$$
$$+ \mathrm{vec}\left(\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\right)^\top \mathrm{vec}\left(\boldsymbol{\mathcal{O}}^*_{k,04}(x)\right)$$
$$+ \mathcal{O}^*_{k,10}(x)\mathcal{B}^*_{0,4}(x;t+1,T) + \boldsymbol{\mathcal{O}}^*_{k,01}(x)\boldsymbol{\mathcal{B}}^*_4(x;t+1,T).$$

C.2 One-Dimensional Cumulants

C.2.1 Cumulants under risk-neutral measure $\mathbb{Q}$

C.2.1.1 Cumulant 1

We have

$$\kappa_1 = Y_t + \mathcal{A}^*_{10}(t,T) + \mathcal{B}^*_{10}(t,T)^\top h^*_t, \tag{C.2.1}$$

where $\mathcal{A}^*_{10}(t,T)$ and $\mathcal{B}^*_{10}(t,T)$ satisfy the following recursions:

$$\mathcal{A}^*_{10}(t,T) = \mathcal{A}^*_{10}(t+1,T) + \mathcal{M}^*_{10} + \mathcal{B}^*_{10}(t+1,T)\mathcal{M}^*_{01}, \tag{C.2.2}$$

and

$$\mathcal{B}^*_{10}(t,T) = \mathcal{O}^*_{10} + \mathcal{B}^*_{10}(t+1,T)\mathcal{O}^*_{01}, \tag{C.2.3}$$

with terminals constraints $\mathcal{A}^*_{10}(T,T)=0$ and $\mathcal{B}^*_{10}(T,T)=0$. Solving Equation (C.2.3) using the constraint $\mathcal{B}^*_{10}(T,T)=0$, we obtain

$$\mathcal{B}^*_{10}(t,T) = \mathcal{O}^*_{10}\frac{1-\mathcal{O}^{*T-t}_{01}}{1-\mathcal{O}^*_{01}}. \tag{C.2.4}$$

Similarly, solving Equation (C.2.2) using the constraint $\mathcal{A}^*_{10}(T,T)=0$, we obtain

$$\begin{aligned}\mathcal{A}^*_{10}(t,T) &= (T-t)\mathcal{M}^*_{10} + \frac{\mathcal{M}^*_{01}\mathcal{O}^*_{10}}{1-\mathcal{O}^*_{01}}\left(T-t-\frac{1-\mathcal{O}^{*T-t}_{01}}{1-\mathcal{O}^*_{01}}\right)\\ &= (T-t)\mathcal{M}^*_{10} + \frac{\mathcal{M}^*_{01}}{1-\mathcal{O}^*_{01}}\left(\mathcal{O}^*_{10}(T-t)-\mathcal{B}^*_{10}(t,T)\right).\end{aligned} \tag{C.2.5}$$

C.2.1.2 Cumulant 2

We have

$$\kappa_2 = \mathcal{A}^*_{20}(t,T) + \boldsymbol{\mathcal{B}}^*_{20}(t,T)^\top \boldsymbol{l}^*_t, \tag{C.2.6}$$

where $\mathcal{A}^*_{20}(t,T)$ and $\boldsymbol{\mathcal{B}}^*_{20}(t,T)$ satisfy the following recursions:

$$\mathcal{A}^*_{20}(t,T) = \mathcal{A}^*_{20}(t+1,T) + \mathcal{M}^*_{20} + 2\mathcal{M}^*_{11}\boldsymbol{\mathcal{B}}^*_{10}(t+1,T) + \mathcal{M}^*_{01}\boldsymbol{\mathcal{B}}^*_{20}(t+1,T) + \mathcal{M}^*_{02}\boldsymbol{\mathcal{B}}^{*2}_{10}(t+1,T), \tag{C.2.7}$$

and

$$\boldsymbol{\mathcal{B}}^*_{20}(t,T) = \boldsymbol{\mathcal{O}}^*_{20} + 2\boldsymbol{\mathcal{O}}^*_{11}\boldsymbol{\mathcal{B}}^*_{10}(t+1,T) + \boldsymbol{\mathcal{O}}^*_{01}\boldsymbol{\mathcal{B}}^*_{20}(t+1,T) + \boldsymbol{\mathcal{O}}^*_{02}\boldsymbol{\mathcal{B}}^{*2}_{10}(t+1,T), \tag{C.2.8}$$

with terminals constraints $\mathcal{A}^*_{20}(T,T) = 0$ and $\boldsymbol{\mathcal{B}}^*_{20}(T,T) = 0$. Solving Equation (C.2.8) using the constraint $\boldsymbol{\mathcal{B}}^*_{20}(T,T) = 0$, we obtain

$$\begin{aligned}\boldsymbol{\mathcal{B}}^*_{20}(t,T) = {}& \boldsymbol{\mathcal{O}}^*_{20}\frac{1-\boldsymbol{\mathcal{O}}^{*T-t}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}} + 2\frac{\boldsymbol{\mathcal{O}}^*_{11}\boldsymbol{\mathcal{O}}^*_{10}}{1-\boldsymbol{\mathcal{O}}^*_{01}}\left(\frac{1-\boldsymbol{\mathcal{O}}^{*T-t-1}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}} - (T-t-1)\boldsymbol{\mathcal{O}}^{*T-t-1}_{01}\right)\\ &+ \frac{\boldsymbol{\mathcal{O}}^*_{02}\boldsymbol{\mathcal{O}}^{*2}_{10}}{(1-\boldsymbol{\mathcal{O}}^*_{01})^2}\left((1+\boldsymbol{\mathcal{O}}^{*T-t}_{01})\frac{1-\boldsymbol{\mathcal{O}}^{*T-t-1}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}} - 2(T-t-1)\boldsymbol{\mathcal{O}}^{*T-t-1}_{01}\right)\end{aligned} \tag{C.2.9}$$

Similarly, solving Equation (C.2.7) using the constraint $\mathcal{A}^*_{20}(T,T) = 0$, we obtain

$$\begin{aligned}&\mathcal{A}^*_{20}(t,T)\\ &= (T-t)\mathcal{M}^*_{20} + \frac{2\mathcal{M}^*_{11}\boldsymbol{\mathcal{O}}^*_{10} + \mathcal{M}^*_{01}\boldsymbol{\mathcal{O}}^*_{20}}{1-\boldsymbol{\mathcal{O}}^*_{01}}\left(T-t-\frac{1-\boldsymbol{\mathcal{O}}^{*T-t}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}}\right)\\ &\quad + \frac{\mathcal{M}^*_{02}\boldsymbol{\mathcal{O}}^{*2}_{10}}{(1-\boldsymbol{\mathcal{O}}^*_{01})^2}\left(T-t-2\frac{1-\boldsymbol{\mathcal{O}}^{*T-t}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}} + \frac{1-\boldsymbol{\mathcal{O}}^{*2(T-t)}_{01}}{1-\boldsymbol{\mathcal{O}}^{*2}_{01}}\right)\\ &\quad + 2\frac{\mathcal{M}^*_{01}\boldsymbol{\mathcal{O}}^*_{11}\boldsymbol{\mathcal{O}}^*_{10}}{(1-\boldsymbol{\mathcal{O}}^*_{01})^2}\left(T-t-1-\frac{1-\boldsymbol{\mathcal{O}}^{*T-t-1}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}}\right)\\ &\quad - 2\left(\frac{\mathcal{M}^*_{01}\boldsymbol{\mathcal{O}}^*_{11}\boldsymbol{\mathcal{O}}^*_{10}}{(1-\boldsymbol{\mathcal{O}}^*_{01})^3} + \frac{\mathcal{M}^*_{01}\boldsymbol{\mathcal{O}}^*_{02}\boldsymbol{\mathcal{O}}^{*2}_{10}}{(1-\boldsymbol{\mathcal{O}}_{01})^4}\right)\left((T-t-2)\boldsymbol{\mathcal{O}}^{T-t}_{01} - (T-t-1)\boldsymbol{\mathcal{O}}^{T-t-1}_{01} + \boldsymbol{\mathcal{O}}_{01}\right)\\ &\quad + \frac{\mathcal{M}^*_{01}\boldsymbol{\mathcal{O}}^*_{02}\boldsymbol{\mathcal{O}}^{*2}_{10}}{(1-\boldsymbol{\mathcal{O}}_{01})^3}\left(T-t-2+\frac{1-\boldsymbol{\mathcal{O}}^{*T-t}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}} - \frac{1-\boldsymbol{\mathcal{O}}^{*T-t-1}_{01}}{1-\boldsymbol{\mathcal{O}}^*_{01}} - \frac{1}{\boldsymbol{\mathcal{O}}^*_{01}}\left(\frac{1-\boldsymbol{\mathcal{O}}^{*2(T-t)}_{01}}{1-\boldsymbol{\mathcal{O}}^{*2}_{01}} - 1\right)\right)\end{aligned} \tag{C.2.10}$$

C.2.2 Cumulants under measure $\widetilde{\mathbb{Q}}_t$

C.2.2.1 Cumulant 1

We have

$$\kappa^{\widetilde{\mathbb{Q}}_t}_1 = Y_t + \mathcal{A}^{\widetilde{\mathbb{Q}}_t}_1(0;t,T) + \boldsymbol{\mathcal{B}}^{\widetilde{\mathbb{Q}}_t}_1(0;t,T)^\top \boldsymbol{l}^*_t, \tag{C.2.11}$$

where the multi-step coefficients $\mathcal{A}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ and $\boldsymbol{\mathcal{B}}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ are given by:

$$\mathcal{A}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T) = (T-t)\mathcal{M}_{10}^* + \frac{\mathcal{M}_{01}^*}{1-\mathcal{O}_{01}^*}\left(\mathcal{O}_{10}^*(T-t) - \mathcal{B}_{10}^*(t,T)\right),$$

$$\boldsymbol{\mathcal{B}}_1^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \mathcal{O}_{10}^*\frac{1-\mathcal{O}_{01}^{*T-t}}{1-\mathcal{O}_{01}^*},$$

with $g(u) = 1+u$.

C.2.2.2 Cumulant 2

We have

$$\kappa_2^{\widetilde{\mathbb{Q}}_t} = \mathcal{A}_2^{\widetilde{\mathbb{Q}}_t}(0;t,T) + \boldsymbol{\mathcal{B}}_2^{\widetilde{\mathbb{Q}}_t}(0;t,T)^\top \boldsymbol{l}_t^*, \tag{C.2.12}$$

where the multi-step coefficients $\mathcal{A}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ and $\boldsymbol{\mathcal{B}}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T)$ are given by:

$$\begin{aligned}
&\mathcal{A}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T) \\
&= (T-t)\mathcal{M}_{20}^* + \frac{2\mathcal{M}_{11}^*\boldsymbol{\mathcal{O}}_{10}^* + \mathcal{M}_{01}^*\boldsymbol{\mathcal{O}}_{20}^*}{1-\boldsymbol{\mathcal{O}}_{01}^*}\left(T-t-\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t}}{1-\boldsymbol{\mathcal{O}}_{01}^*}\right) \\
&\quad + \frac{\mathcal{M}_{02}^*\boldsymbol{\mathcal{O}}_{10}^{*2}}{(1-\boldsymbol{\mathcal{O}}_{01}^*)^2}\left(T-t-2\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t}}{1-\boldsymbol{\mathcal{O}}_{01}^*} + \frac{1-\boldsymbol{\mathcal{O}}_{01}^{*2(T-t)}}{1-\boldsymbol{\mathcal{O}}_{01}^{*2}}\right) \\
&\quad + 2\frac{\mathcal{M}_{01}^*\boldsymbol{\mathcal{O}}_{11}^*\boldsymbol{\mathcal{O}}_{10}^*}{(1-\boldsymbol{\mathcal{O}}_{01}^*)^2}\left(T-t-1-\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t-1}}{1-\boldsymbol{\mathcal{O}}_{01}^*}\right) \\
&\quad - 2\left(\frac{\mathcal{M}_{01}^*\boldsymbol{\mathcal{O}}_{11}^*\boldsymbol{\mathcal{O}}_{10}^*}{(1-\boldsymbol{\mathcal{O}}_{01}^*)^3} + \frac{\mathcal{M}_{01}^*\boldsymbol{\mathcal{O}}_{02}^*\boldsymbol{\mathcal{O}}_{10}^{*2}}{(1-\boldsymbol{\mathcal{O}}_{01})^4}\right)\left((T-t-2)\boldsymbol{\mathcal{O}}_{01}^{T-t} - (T-t-1)\boldsymbol{\mathcal{O}}_{01}^{T-t-1} + \boldsymbol{\mathcal{O}}_{01}\right) \\
&\quad + \frac{\mathcal{M}_{01}^*\boldsymbol{\mathcal{O}}_{02}^*\boldsymbol{\mathcal{O}}_{10}^{*2}}{(1-\boldsymbol{\mathcal{O}}_{01})^3}\left(T-t-2+\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t}}{1-\boldsymbol{\mathcal{O}}_{01}^*} - \frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t-1}}{1-\boldsymbol{\mathcal{O}}_{01}^*} - \frac{1}{\boldsymbol{\mathcal{O}}_{01}^*}\left(\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*2(T-t)}}{1-\boldsymbol{\mathcal{O}}_{01}^{*2}} - 1\right)\right),
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_2^{\widetilde{\mathbb{Q}}_t}(u;t,T) = \boldsymbol{\mathcal{O}}_{20}^*\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t}}{1-\boldsymbol{\mathcal{O}}_{01}^*} &+ 2\frac{\boldsymbol{\mathcal{O}}_{11}^*\boldsymbol{\mathcal{O}}_{10}^*}{1-\boldsymbol{\mathcal{O}}_{01}^*}\left(\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t-1}}{1-\boldsymbol{\mathcal{O}}_{01}^*} - (T-t-1)\boldsymbol{\mathcal{O}}_{01}^{*T-t-1}\right) \\
&+ \frac{\boldsymbol{\mathcal{O}}_{02}^*\boldsymbol{\mathcal{O}}_{10}^{*2}}{(1-\boldsymbol{\mathcal{O}}_{01}^*)^2}\left((1+\boldsymbol{\mathcal{O}}_{01}^{*T-t})\frac{1-\boldsymbol{\mathcal{O}}_{01}^{*T-t-1}}{1-\boldsymbol{\mathcal{O}}_{01}^*} - 2(T-t-1)\boldsymbol{\mathcal{O}}_{01}^{*T-t-1}\right),
\end{aligned}$$

with $g(u) = 1+u$.

Appendix D

Proofs

D.1 Derivation of the inverse Discrete Fourier Transform

After truncating the integral at a sufficiently large value a, we obtain

$$V_t \approx \frac{e^{-r(T-t)}}{\pi} \sum_{j=0}^{N} \mathrm{Re}\left[\exp\left(\mathcal{C}_{Y_T}^{\mathbb{Q}}(R+iu_j \mid \mathcal{F}_t)\right) \check{f}(R+iu_j) w_j\right], \tag{D.1.1}$$

where $\Delta_u = \frac{a}{N}$, $u_j = j\Delta_u$ and w_j is defined according to a quadrature rule such as Newton-Cotes for $1 \leqslant j \leqslant N$. Let $m_l = Y_t - k_l = -m_{max} + l\Delta_k$ represent the log-moneyness of an option with strike e^{k_l} where m_{max} represents the largest value of log-moneyness of an option, Δ_k represents the spacing between the different values of moneyness ranging from $-m_{max}$ to $m_{max} - \Delta_k$ and let $m_{max} = \frac{N\Delta_k}{2}$ then we can reformulate equation (D.1.1) for a call option payoff:

$$\begin{aligned}
V_t^{(l)} &\approx \frac{e^{-r(T-t)}}{\pi} \sum_{j=0}^{N-1} \mathrm{Re}\left[\exp\left(\mathcal{C}_{Y_T}^{\mathbb{Q}}(R+iu_j \mid \mathcal{F}_t)\right) \frac{e^{(1-R-iu_j)k_l}}{(R+iu_j)(R+iu_j-1)} w_j\right] \\
&= \frac{e^{-r(T-t)-(1-R)m_l+Y_t}}{\pi} \sum_{j=0}^{N-1} \mathrm{Re}\left[\frac{e^{iu_j m_l} w_j}{(R+iu_j)(R+iu_j-1)} \exp\left(\mathcal{C}_{Y_T-Y_t}^{\mathbb{Q}}(R+iu_j \mid \mathcal{F}_t)\right)\right] \\
&= \frac{e^{-r(T-t)+(R-1)m_l+Y_t}}{\pi} \sum_{j=0}^{N-1} \mathrm{Re}\left[\frac{e^{ij\Delta_u(-m_{max}+l\Delta_k)} w_j}{(R+iu_j)(R+iu_j-1)} \exp\left(\mathcal{C}_{Y_T-Y_t}^{\mathbb{Q}}(R+iu_j \mid \mathcal{F}_t)\right)\right] \\
&= \frac{e^{-r(T-t)+(R-1)m_l+Y_t}}{\pi} \sum_{j=0}^{N-1} \mathrm{Re}\left[e^{i\Delta_u\Delta_k jl} \frac{e^{-ij\Delta_u m_{max}} w_j}{(R+iu_j)(R+iu_j-1)} \exp\left(\mathcal{C}_{Y_T-Y_t}^{\mathbb{Q}}(R+iu_j \mid \mathcal{F}_t)\right)\right] \\
&= \mathrm{Re}\left[\sum_{j=0}^{N-1} e^{i\frac{2\pi}{N}jl} v_t^{(j)}\right],
\end{aligned}$$

since by the Nyquist relation we have

$$\frac{2\pi}{N} = \Delta_u \Delta_k,$$

and where for $0 \leqslant j \leqslant N-1$:

$$v_t^{(j)} = \frac{e^{-r(T-t)+(R-1)m_l+Y_t}}{\pi} \frac{e^{-ij\Delta_u m_{max}} w_j}{(R+iu_j)(R+iu_j-1)} \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T-Y_t}(R+iu_j \mid \mathcal{F}_t)\right).$$

which completes the proof.

D.2 Proof of Theorem D.3.1

We will use the following notation to simplify the proof. We have

$$\begin{aligned} J_{\boldsymbol{h}}(x) &= \begin{bmatrix} \frac{\partial h_0}{\partial x}(x) & \frac{\partial h_1}{\partial x}(x) & \frac{\partial h_N}{\partial x}(x) \\ & = & \end{bmatrix}^\top \\ &= \frac{\partial h_0}{\partial x}(x) +\!\!+ \begin{bmatrix} \frac{\partial h_1}{\partial x}(x) & \dots & \frac{\partial h_N}{\partial x}(x) \end{bmatrix}^\top \\ &= h_{0,1}(x) +\!\!+ \boldsymbol{g}_1(x)^\top, \end{aligned}$$

where $h_{0,1}$ designates the first partial derivative of h_0 with respect to x and $\boldsymbol{g}_1^\top$ designates the vector of partial derivatives of h_1 to h_n with respect to x. We will use a similar notation analogous for the second derivative as well:

$$\begin{aligned} \frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x) &= \begin{bmatrix} \frac{\partial^2 h_0}{\partial x^2}(x) & \frac{\partial^2 h_1}{\partial x^2}(x) & \frac{\partial^2 h_N}{\partial x^2}(x) = \end{bmatrix}^\top \\ &= \frac{\partial^2 h_0}{\partial x^2}(x) +\!\!+ \begin{bmatrix} \frac{\partial^2 h_1}{\partial x^2}(x) & \dots & \frac{\partial^2 h_N}{\partial x^2}(x) \end{bmatrix}^\top \\ &= h_{0,2}(x) +\!\!+ \boldsymbol{g}_2(x)^\top. \end{aligned}$$

Now let J_f designate the $1 \times N+1$ Jacobian vector of f, H_f designate the $N+1 \times N+1$ Hessian matrix of f. Similar to the derivatives of $\boldsymbol{h}$, to simplify of the tensor representation of the 3rd order partial derivatives of f with respect to $\boldsymbol{x}$, we will introduce a notation to help visualize. We introduce the following notation for the Jacobian vector of f:

$$J_f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_0}(\boldsymbol{x}) & \frac{\partial f}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f}{\partial x_N}(\boldsymbol{x}) = \end{bmatrix} = \frac{\partial f}{\partial x_0}(\boldsymbol{x}) +\!\!+ \begin{bmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{x}) & \dots & \frac{\partial f}{\partial x_N}(\boldsymbol{x}) \end{bmatrix} = f_{10}(\boldsymbol{x}) +\!\!+ \boldsymbol{f}_{01}(\boldsymbol{x}).$$

Now the definition of the Hessian matrix states:

$$H_f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_0^2}(\boldsymbol{x}) & \dots & \frac{\partial^2 f}{\partial x_0 \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_0 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^2 f}{\partial x_N^2}(\boldsymbol{x}). \end{bmatrix}$$

We introduce a notation partitioning the Hessian matrix into three different sequences:

$$f_{20}(\boldsymbol{x}) = \frac{\partial^2 f}{\partial x_0^2}(\boldsymbol{x}),$$

$$\boldsymbol{f}_{11}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_0 \partial x_1}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_0 \partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial^2 f}{\partial x_0 \partial x_N}(\boldsymbol{x}) \end{bmatrix}^\top,$$

$$\boldsymbol{f}_{02}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^2 f}{\partial x_N^2}(\boldsymbol{x}) \end{bmatrix}.$$

Finally we introduce a similar notation to partition the 3-d tensor of the 3rd order partial-derivatives of f with respect to $\boldsymbol{x}$. First off, we will visualize the tensor by giving you a view of the frontal slices of the 3d-tensor. Let $\Upsilon_f(\boldsymbol{x})_i$ denote the frontal slice i of the 3d-tensor with $i = 1,\dots,N+1$ then we have:

$$\Upsilon_f(\boldsymbol{x})_1 = \begin{bmatrix} \frac{\partial^3 f}{\partial x_0^3}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0^2 \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_0^2 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0 \partial x_N^2}(\boldsymbol{x}) \end{bmatrix},$$

$$\Upsilon_f(\boldsymbol{x})_2 = \begin{bmatrix} \frac{\partial^3 f}{\partial x_0^2 \partial x_1}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0 \partial x_1 \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_0 \partial x_1 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1 \partial x_N^2}(\boldsymbol{x}) \end{bmatrix},$$

$$\vdots$$

$$\Upsilon_f(\boldsymbol{x})_N = \begin{bmatrix} \frac{\partial^3 f}{\partial x_0^2 \partial x_{N-1}}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0 \partial x_{N-1} \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_0 \partial x_{N-1} \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_{N-1} \partial x_N^2}(\boldsymbol{x}) \end{bmatrix},$$

$$\Upsilon_f(\boldsymbol{x})_{N+1} = \begin{bmatrix} \frac{\partial^3 f}{\partial x_0^2 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0 \partial x_N^2}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_0 \partial x_N^2}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_N^3}(\boldsymbol{x}) \end{bmatrix}.$$

Similarly to the partitioning of the Hessian Matrix of f, we partition $\Upsilon_f(\boldsymbol{x})$ into four different sequences of its third-order partial derivatives:

$$f_{30}(\boldsymbol{x}) = \frac{\partial^3 f}{\partial x_0^3}(\boldsymbol{x}),$$

$$\boldsymbol{f}_{21}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^3 f}{\partial x_0^2 \partial x_1}(\boldsymbol{x}) & \frac{\partial^3 f}{\partial x_0^2 \partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0^2 \partial x_N}(\boldsymbol{x}) \end{bmatrix}^\top,$$

$$
\boldsymbol{f}_{12}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^3 f}{\partial x_0 \partial x_1^2}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0 \partial x_1 \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_0 \partial x_1 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_0 \partial x_N^2}(\boldsymbol{x}) \end{bmatrix},
$$

and $\boldsymbol{f}_{03}$ is a 3-d tensor composed of all 3rd order partial derivatives of f with respect to x_1,...,x_N. Here is a visualization the tensor by giving you a view of its frontal slices of the 3d-tensor. We have

$$
\boldsymbol{f}_{03}(\boldsymbol{x})_1 = \begin{bmatrix} \frac{\partial^3 f}{\partial x_1^3}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1^2 \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_1^2 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1 \partial x_N^2}(\boldsymbol{x}) \end{bmatrix},
$$

$$
\boldsymbol{f}_{03}(\boldsymbol{x})_2 = \begin{bmatrix} \frac{\partial^3 f}{\partial x_1^2 \partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_2 \partial x_N^2}(\boldsymbol{x}) \end{bmatrix},
$$

$$
\vdots
$$

$$
\boldsymbol{f}_{03}(\boldsymbol{x})_{N-1} = \begin{bmatrix} \frac{\partial^3 f}{\partial x_1^2 \partial x_{N-1}}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1 \partial x_{N-1} \partial x_N}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_1 \partial x_{N-1} \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_{N-1} \partial x_N^2}(\boldsymbol{x}) \end{bmatrix},
$$

$$
\boldsymbol{f}_{03}(\boldsymbol{x})_N = \begin{bmatrix} \frac{\partial^3 f}{\partial x_1^2 \partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1 \partial x_N^2}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^3 f}{\partial x_1 \partial x_N^2}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_N^3}(\boldsymbol{x}) \end{bmatrix}.
$$

One final clarification is that $\operatorname{vec}(\Upsilon_f(\boldsymbol{x}))$ is a column-wise vectorization of the tensor. In other words, the mode-1 fibers or columns are stacked over each other to create a vector iterating from the first frontal slice to the last one. We are finally ready to prove the theorem. Now by definition of the Kronecker product, we have

$$
\begin{aligned}
&\operatorname{vec}(J_{\boldsymbol{h}}(x) \otimes \frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x) J_{\boldsymbol{h}}(x)^\top)^\top \operatorname{vec}(\Upsilon_f(\boldsymbol{h}(x))) \\
&\quad = \sum_{i=0}^{N} \sum_{j=0}^{N} \sum_{k=0}^{N} \frac{\partial \boldsymbol{h}_i}{\partial x}(x) \frac{\partial^2 \boldsymbol{h}_k}{\partial x^2}(x) \frac{\partial \boldsymbol{h}_j}{\partial x}(x) \left[\operatorname{vec}(\Upsilon_f(\boldsymbol{h}(x)))\right]_{k+(N+1)\times j+(N+1)^2\times i}.
\end{aligned}
$$

Now if $i = j = k = 0$ we have

$$
\left[\operatorname{vec}(\Upsilon_f(\boldsymbol{h}(x)))\right]_0 = f_{30}(\boldsymbol{h}(x)).
$$

If $i = j = 0$ and $k = 1,\ldots,N$ we have

$$[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_k = [\boldsymbol{f}_{21}(\boldsymbol{h}(x))]_{k-1}\,.$$

If $i = 0, k = 0$ and $j = 1,\ldots,N$, we have

$$[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{(N+1)\times j} = [\boldsymbol{f}_{21}(\boldsymbol{h}(x))]_{j-1}\,.$$

If $i = 0$ and $1 \leqslant j,k \leqslant N$ we have

$$[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{k+(N+1)\times j} = [\boldsymbol{f}_{12}(\boldsymbol{h}(x))]_{k-1,j-1}\,.$$

Using a similar argument and by symmetry, if j=k=0, for $i = 1,\ldots N$ we have

$$[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{(N+1)^2\times i} = [\boldsymbol{f}_{21}(\boldsymbol{h}(x))]_{i-1}\,.$$

By symmetry of the tensor of 3rd-order partial derivatives, If $k = 0$ and $1 \leqslant i,j \leqslant N$ we have

$$[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{(N+1)\times j+(N+1)^2\times i} = [\boldsymbol{f}_{12}(\boldsymbol{h}(x))]_{j-1,i-1}\,.$$

and if $j = 0$ and $1 \leqslant i,k \leqslant N$ we have

$$[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{k+(N+1)^2\times i} = [\boldsymbol{f}_{12}(\boldsymbol{h}(x))]_{k-1,i-1}\,.$$

Finally if $1 \leqslant i,j,k \leqslant N$ then we have

$$[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{k+(N+1)j+(N+1)^2\times i} = [\boldsymbol{f}_{03}(\boldsymbol{h}(x))]_{k-1,j-1,i-1}\,.$$

Consequently, if we partition the summation analogously, we obtain

$$\begin{aligned}
&\mathrm{vec}(J_{\boldsymbol{h}}(x) \otimes \frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x) J_{\boldsymbol{h}}(x)^\top)^\top\,\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))\\
&\quad = \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x)\,[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_0\\
&\qquad + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \sum_{k=1}^{N} \frac{\partial^2 h_k}{\partial x^2}(x)\,[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_k\\
&\qquad + \frac{\partial h_0}{\partial x}(x)\frac{\partial^2 h_0}{\partial x^2}(x) \sum_{j=1}^{N} \frac{\partial h_j}{\partial x}(x)\,[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{(N+1)\times j}\\
&\qquad + \frac{\partial h_0}{\partial x}(x) \sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial^2 h_k}{\partial x^2}(x)\frac{\partial h_j}{\partial x}(x)\,[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{k+(N+1)\times j}\\
&\qquad + \frac{\partial^2 h_0}{\partial x^2}(x)\frac{\partial h_0}{\partial x}(x) \sum_{i=1}^{N} \frac{\partial h_i}{\partial x}(x)\,[\mathrm{vec}\,(\Upsilon_f(\boldsymbol{h}(x)))]_{(N+1)^2\times i}
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{\partial h_0}{\partial x}(x) \sum_{i=1}^{N}\sum_{k=1}^{N} \frac{\partial h_i}{\partial x}(x) \frac{\partial^2 h_k}{\partial x^2}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)^2\times i} \\
&+ \frac{\partial^2 h_0}{\partial x^2}(x) \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial h_i}{\partial x}(x) \frac{\partial h_j}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{(N+1)\times j+(N+1)^2\times i} \\
&+ \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial h_i}{\partial x}(x) \frac{\partial^2 h_k}{\partial x^2}(x) \frac{\partial h_j}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)\times j+(N+1)^2\times i},
\end{aligned}
$$

and we obtain

$$
\begin{aligned}
&\mathrm{vec}(J_{\boldsymbol{h}}(x) \otimes \frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x) J_{\boldsymbol{h}}(x)^\top)^\top \mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right) \\
&\quad = \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x) f_{30}(\boldsymbol{h}(x)) + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \boldsymbol{g}_2(x)^\top \boldsymbol{f}_{21}(\boldsymbol{h}(x)) \\
&\qquad + 2\frac{\partial h_0}{\partial x}(x) \frac{\partial^2 h_0}{\partial x^2}(x) \boldsymbol{g}_1(x)^\top \boldsymbol{f}_{21}(\boldsymbol{h}(x)) + 2\frac{\partial h_0}{\partial x}(x) \boldsymbol{g}_2(x)^\top \boldsymbol{f}_{12}(\boldsymbol{h}(x)) \boldsymbol{g}_1(x) \\
&\qquad + \frac{\partial^2 h_0}{\partial x^2}(x) \boldsymbol{g}_1(x)^\top \boldsymbol{f}_{12}(\boldsymbol{h}(x)) \boldsymbol{g}_1(x) + \mathrm{vec}\left(\boldsymbol{g}_1 \otimes \boldsymbol{g}_2 \boldsymbol{g}_1^\top\right)^\top \mathrm{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right).
\end{aligned}
$$

Now using a similar argument, we can also partition the following sum

$$
\begin{aligned}
&\mathrm{vec}(J_{\boldsymbol{h}}(x) \otimes J_{\boldsymbol{h}}(x) \frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x)^\top)^\top \mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right) \\
&\quad = \sum_{i=0}^{N}\sum_{j=0}^{N}\sum_{k=0}^{N} \frac{\partial h_j}{\partial x}(x) \frac{\partial^2 h_i}{\partial x^2}(x) \frac{\partial h_k}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)\times j+(N+1)^2\times i} \\
&\quad = \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_0 \\
&\qquad + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \sum_{i=1}^{N} \frac{\partial^2 h_i}{\partial x^2}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{(N+1)^2\times i} \\
&\quad + \frac{\partial h_0}{\partial x}(x) \frac{\partial^2 h_0}{\partial x^2}(x) \sum_{j=1}^{N} \frac{\partial h_j}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{(N+1)\times j} \\
&\qquad + \frac{\partial^2 h_0}{\partial x^2}(x) \frac{\partial h_0}{\partial x}(x) \sum_{k=1}^{N} \frac{\partial h_k}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k} \\
&\quad + \frac{\partial^2 h_0}{\partial x^2}(x) \sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial h_k}{\partial x}(x) \frac{\partial h_j}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)\times j} \\
&\qquad + \frac{\partial h_0}{\partial x}(x) \sum_{i=1}^{N}\sum_{k=1}^{N} \frac{\partial h_k}{\partial x}(x) \frac{\partial^2 h_i}{\partial x^2}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)^2\times i}
\end{aligned}
$$

$$
+ \frac{\partial h_0}{\partial x}(x) \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial^2 h_i}{\partial x^2}(x) \frac{\partial h_j}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{(N+1)\times j+(N+1)^2\times i}
$$

$$
+ \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\partial h_j}{\partial x}(x) \frac{\partial^2 h_i}{\partial x^2}(x) \frac{\partial h_k}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)\times j+(N+1)^2\times i},
$$

and obtain

$$
\begin{aligned}
&\mathrm{vec}(J_{\boldsymbol{h}}(x) \otimes J_{\boldsymbol{h}}(x) \frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x)^\top)^\top \mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right) \\
&\quad = \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x) f_{30}(\boldsymbol{h}(x)) + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \boldsymbol{g}_2(x)^\top \boldsymbol{f}_{21}(\boldsymbol{h}(x)) \\
&\qquad + 2\frac{\partial h_0}{\partial x}(x) \frac{\partial^2 h_0}{\partial x^2}(x) \boldsymbol{g}_1(x)^\top \boldsymbol{f}_{21}(\boldsymbol{h}(x)) + 2\frac{\partial h_0}{\partial x}(x) \boldsymbol{g}_2(x)^\top \boldsymbol{f}_{12}(\boldsymbol{h}(x)) \boldsymbol{g}_1(x) \\
&\qquad + \frac{\partial^2 h_0}{\partial x^2}(x) \boldsymbol{g}_1(x)^\top \boldsymbol{f}_{12}(\boldsymbol{h}(x)) \boldsymbol{g}_1(x) + \mathrm{vec}\left(\boldsymbol{g}_1 \otimes \boldsymbol{g}_1 \boldsymbol{g}_2^\top\right)^\top \mathrm{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right).
\end{aligned}
$$

We apply the same logic once again to the following sum

$$
\begin{aligned}
&\mathrm{vec}(\frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x) \otimes J_{\boldsymbol{h}}(x) J_{\boldsymbol{h}}(x)^\top)^\top \mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right) \\
&\quad = \sum_{i=0}^{N} \sum_{j=0}^{N} \sum_{k=0}^{N} \frac{\partial h_i}{\partial x}(x) \frac{\partial^2 h_j}{\partial x^2}(x) \frac{\partial h_k}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)\times j+(N+1)^2\times i} \\
&\quad = \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_0 \\
&\qquad + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \sum_{j=1}^{N} \frac{\partial^2 h_j}{\partial x^2}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{(N+1)\times j} \\
&\qquad + \frac{\partial h_0}{\partial x}(x) \frac{\partial^2 h_0}{\partial x^2}(x) \sum_{i=1}^{N} \frac{\partial h_i}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{(N+1)^2\times i} \\
&\qquad + \frac{\partial^2 h_0}{\partial x^2}(x) \frac{\partial h_0}{\partial x}(x) \sum_{k=1}^{N} \frac{\partial h_k}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_k \\
&\qquad + \frac{\partial^2 h_0}{\partial x^2}(x) \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{\partial h_k}{\partial x}(x) \frac{\partial h_i}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)^2\times i} \\
&\qquad + \frac{\partial h_0}{\partial x}(x) \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\partial h_k}{\partial x}(x) \frac{\partial^2 h_j}{\partial x^2}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)\times j} \\
&\qquad + \frac{\partial h_0}{\partial x}(x) \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial^2 h_j}{\partial x^2}(x) \frac{\partial h_i}{\partial x}(x) \left[\mathrm{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{(N+1)\times j+(N+1)^2\times i}
\end{aligned}
$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial h_i}{\partial x}(x)\frac{\partial^2 h_j}{\partial x^2}(x)\frac{\partial h_k}{\partial x}(x)\left[\operatorname{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right)\right]_{k+(N+1)\times j+(N+1)^2\times i},$$

and obtain

$$\begin{aligned}
&\operatorname{vec}\Big(\frac{\partial^2\boldsymbol{h}}{\partial x^2}(x)\otimes J_{\boldsymbol{h}}(x)J_{\boldsymbol{h}}(x)^\top\Big)^\top\operatorname{vec}\left(\Upsilon_f(h(x))\right)\\
&=\left(\frac{\partial h_0}{\partial x}(x)\right)^2\frac{\partial^2 h_0}{\partial x^2}(x)f_{30}(\boldsymbol{h}(x))+\left(\frac{\partial h_0}{\partial x}(x)\right)^2\boldsymbol{g}_2(x)^\top\boldsymbol{f}_{21}(\boldsymbol{h}(x))\\
&\quad+2\frac{\partial h_0}{\partial x}(x)\frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^\top\boldsymbol{f}_{21}(\boldsymbol{h}(x))+2\frac{\partial h_0}{\partial x}(x)\boldsymbol{g}_2(x)^\top\boldsymbol{f}_{12}(\boldsymbol{h}(x))\boldsymbol{g}_1(x)\\
&\quad+\frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^\top\boldsymbol{f}_{12}(\boldsymbol{h}(x))\boldsymbol{g}_1(x)+\operatorname{vec}\left(\boldsymbol{g}_2\otimes\boldsymbol{g}_1\boldsymbol{g}_1^\top\right)^\top\operatorname{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right).
\end{aligned}$$

It now remains to show that

$$\begin{aligned}
\operatorname{vec}\left(\boldsymbol{g}_1\otimes\boldsymbol{g}_2\boldsymbol{g}_1^\top\right)^\top\operatorname{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right)&=\operatorname{vec}\left(\boldsymbol{g}_1\otimes\boldsymbol{g}_1\boldsymbol{g}_2^\top\right)^\top\operatorname{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right)\\
&=\operatorname{vec}\left(\boldsymbol{g}_2\otimes\boldsymbol{g}_1\boldsymbol{g}_1^\top\right)^\top\operatorname{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right).
\end{aligned}\tag{D.2.1}$$

We will prove it by induction. For the base case of $N=1$, it is clear that the equalities stand. Indeed, all the expressions simplify to the following scalar result:

$$g_1^2 g_2\frac{\partial f}{\partial x_1^3}(\boldsymbol{h}(x)).$$

Let equation (D.2.1) be valid for an arbitrary value of $N-1$, we will show it stands for N. Now, we know $\boldsymbol{f}_{03}$ is a 3-d tensor composed of all 3rd order partial derivatives of f with respect to x_1,...,x_N. Thus, we can decompose the tensor the same way we decomposed Υ_f. Let

$$\begin{aligned}
\phi_{30}(\boldsymbol{x})&=\frac{\partial^3 f}{\partial x_1^3}(\boldsymbol{x}),\\
\boldsymbol{\phi}_{21}(\boldsymbol{x})&=\begin{bmatrix}\frac{\partial^3 f}{\partial x_1^2\partial x_2}(\boldsymbol{x}) & \frac{\partial^3 f}{\partial x_1^2\partial x_3}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1^2\partial x_N}(\boldsymbol{x})\end{bmatrix},\\
\boldsymbol{\phi}_{12}(\boldsymbol{x})&=\begin{bmatrix}\frac{\partial^3 f}{\partial x_1\partial x_2^2}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1\partial x_2\partial x_N}(\boldsymbol{x})\\ \vdots & \ddots & \vdots\\ \frac{\partial^3 f}{\partial x_1\partial x_2\partial x_N}(\boldsymbol{x}) & \dots & \frac{\partial^3 f}{\partial x_1\partial x_N^2}(\boldsymbol{x})\end{bmatrix},
\end{aligned}$$

and $\boldsymbol{\phi}_{03}$ is a 3-d tensor composed of all 3rd order partial derivatives of f with respect to x_2,...,x_N. Now if $i=j=k=0$ we have

$$\left[\operatorname{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right)\right]_0=\frac{\partial^3 f}{\partial x_1^3}(\boldsymbol{h}(x))=\phi_{30}(\boldsymbol{h}(x)).$$

If $i=j=0$ and $k=1$,...,N we have

$$\left[\operatorname{vec}\left(\boldsymbol{f}_{03}(\boldsymbol{h}(x))\right)\right]_k=\left[\boldsymbol{\phi}_{21}(\boldsymbol{h}(x))\right]_{k-1}.$$

If $i=0, k=0$ and $j=1,...,N$, we have

$$[\operatorname{vec}(\boldsymbol{f}_{03}(\boldsymbol{h}(x)))]_{(N+1)\times j} = [\boldsymbol{\phi}_{21}(\boldsymbol{h}(x))]_{j-1}.$$

If $i=0$ and $1 \leqslant j,k \leqslant N$ we have

$$[\operatorname{vec}(\boldsymbol{f}_{03}(\boldsymbol{h}(x)))]_{k+(N+1)\times j} = [\boldsymbol{\phi}_{12}(\boldsymbol{h}(x))]_{k-1,j-1}.$$

Using a similar argument and by symmetry, if j=k=0, for $i=1,...N$ we have

$$[\operatorname{vec}(\boldsymbol{f}_{03}(\boldsymbol{h}(x)))]_{(N+1)^2\times i} = [\boldsymbol{\phi}_{21}(\boldsymbol{h}(x))]_{i-1}.$$

By symmetry of the tensor of 3rd-order partial derivatives, If $k=0$ and $1 \leqslant i,j \leqslant N$ we have

$$[\operatorname{vec}(\boldsymbol{f}_{03}(\boldsymbol{h}(x)))]_{(N+1)\times j+(N+1)^2\times i} = [\boldsymbol{\phi}_{12}(\boldsymbol{h}(x))]_{j-1,i-1},$$

and if $j=0$ and $1 \leqslant i,k \leqslant N$ we have

$$[\operatorname{vec}(\boldsymbol{f}_{03}(\boldsymbol{h}(x)))]_{k+(N+1)^2\times i} = [\boldsymbol{\phi}_{12}(\boldsymbol{h}(x))]_{k-1,i-1}.$$

Finally if $1 \leqslant i,j,k \leqslant N$ then we have

$$[\operatorname{vec}(\boldsymbol{f}_{03}(\boldsymbol{h}(x)))]_{k+(N+1)j+(N+1)^2\times i} = [\boldsymbol{\phi}_{03}(\boldsymbol{h}(x))]_{k-1,j-1,i-1}.$$

Consequently, we obtain the three following recursive equations:

$$\begin{aligned}
&\operatorname{vec}\left(\boldsymbol{g}_2 \otimes \boldsymbol{g}_1\boldsymbol{g}_1^\top\right)^\top \operatorname{vec}(\boldsymbol{f}_{03}(h(x))) \\
&\quad= \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x)\phi_{30}(h(x)) + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \boldsymbol{g}_2(x)^\top \boldsymbol{\phi}_{21}(h(x)) \\
&\qquad+ 2\frac{\partial h_0}{\partial x}(x)\frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^\top\boldsymbol{\phi}_{21}(h(x)) + 2\frac{\partial h_0}{\partial x}(x)\boldsymbol{g}_2(x)^\top\boldsymbol{\phi}_{12}(h(x))\boldsymbol{g}_1(x) \\
&\qquad+ \frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^\top\boldsymbol{\phi}_{12}(h(x))\boldsymbol{g}_1(x) + \operatorname{vec}\left(\boldsymbol{g}_2 \otimes \boldsymbol{g}_1\boldsymbol{g}_1^\top\right)^\top \operatorname{vec}(\boldsymbol{\phi}_{03}(h(x)))
\end{aligned}$$

and

$$\begin{aligned}
&\operatorname{vec}\left(\boldsymbol{g}_1 \otimes \boldsymbol{g}_2\boldsymbol{g}_1^\top\right)^\top \operatorname{vec}(\boldsymbol{f}_{03}(h(x))) \\
&\quad= \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x)\phi_{30}(h(x)) + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \boldsymbol{g}_2(x)^\top \boldsymbol{\phi}_{21}(h(x)) \\
&\qquad+ 2\frac{\partial h_0}{\partial x}(x)\frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^\top\boldsymbol{\phi}_{21}(h(x)) + 2\frac{\partial h_0}{\partial x}(x)\boldsymbol{g}_2(x)^\top\boldsymbol{\phi}_{12}(h(x))\boldsymbol{g}_1(x) \\
&\qquad+ \frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^\top\boldsymbol{\phi}_{12}(h(x))\boldsymbol{g}_1(x) + \operatorname{vec}\left(\boldsymbol{g}_1 \otimes \boldsymbol{g}_2\boldsymbol{g}_1^\top\right)^\top \operatorname{vec}(\boldsymbol{\phi}_{03}(h(x)))
\end{aligned}$$

and

$$
\begin{aligned}
&\operatorname{vec}\left(\boldsymbol{g}_1 \otimes \boldsymbol{g}_1\boldsymbol{g}_2^{\top}\right)^{\top} \operatorname{vec}\left(\boldsymbol{f}_{03}(h(x))\right) \\
&\quad= \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \frac{\partial^2 h_0}{\partial x^2}(x)\phi_{30}(h(x)) + \left(\frac{\partial h_0}{\partial x}(x)\right)^2 \boldsymbol{g}_2(x)^{\top}\boldsymbol{\phi}_{21}(h(x)) \\
&\qquad + 2\frac{\partial h_0}{\partial x}(x)\frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^{\top}\boldsymbol{\phi}_{21}(h(x)) + 2\frac{\partial h_0}{\partial x}(x)\boldsymbol{g}_2(x)^{\top}\boldsymbol{\phi}_{12}(h(x))\boldsymbol{g}_1(x) \\
&\qquad + \frac{\partial^2 h_0}{\partial x^2}(x)\boldsymbol{g}_1(x)^{\top}\boldsymbol{\phi}_{12}(h(x))\boldsymbol{g}_1(x) + \operatorname{vec}\left(\boldsymbol{g}_1 \otimes \boldsymbol{g}_1\boldsymbol{g}_2^{\top}\right)^{\top} \operatorname{vec}\left(\boldsymbol{\phi}_{03}(h(x))\right)
\end{aligned}
$$

Now, since $\boldsymbol{\phi}_{03}$ is a 3-d tensor composed of all 3rd order partial derivatives of f with respect to x_2,...,x_N, we can once again decompose it as we did with $\boldsymbol{f}_{03}$. Indeed, for $k = N-1$ to 1, if we recursively decompose each $\mathbb{R}^{k\times k\times k}$ tensor respectively composed of the 3rd order partial derivatives of f with respect to $x_{N-k+1}, x_{N-k+2}, x_{N-k+3}, \ldots, x_{N-1}, x_N$, we realize that the last tensor is a 1-dimensional tensor element of $\mathbb{R}$ equal to $\frac{\partial^3 f}{\partial x^3}(\boldsymbol{h}(x))$. Consequently, if we recursively combine the the decompositions, it is clear that Equation (D.2.1) holds which completes the proof.

D.3 Proof of N-Dimensional Derivatives

As mentioned in Proposition 2.3.7, we recall that the multi-step coefficients $\mathcal{A}^*(u,\boldsymbol{v};t,T)$, $\mathcal{B}_0^*(u,\boldsymbol{v};t,T)$ and $\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,T)$ satisfy the following recursions derived in Augustyniak et al. (2023):

$$
\begin{aligned}
\mathcal{A}^*(u,\boldsymbol{v};t,T) &= \mathcal{A}^*(u,\boldsymbol{v};t+1,T) + \mathcal{A}^*(u+\mathcal{B}_0^*(u,\boldsymbol{v};t+1,T),\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T);t,t+1), \\
\mathcal{B}_0^*(u,v;t,T) &= \mathcal{B}_0^*(u+\mathcal{B}_0^*(u,\boldsymbol{v};t+1,T),\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T);t,t+1), \\
\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,T) &= \boldsymbol{\mathcal{B}}^*(u+\mathcal{B}_0^*(u,\boldsymbol{v};t+1,T),\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t+1,T);t,t+1),
\end{aligned}
\tag{D.3.1}
$$

with the terminal conditions $\mathcal{A}^*(u,\boldsymbol{v};T,T) = 0$, $\mathcal{B}_0^*(u,\boldsymbol{v};T,T) = 0$ and $\boldsymbol{\mathcal{B}}^*(u,v;T,T) = \boldsymbol{v}$. We recall that the $\mathcal{F}_t$-conditional risk-neutral CGF of Y_T is

$$
\begin{aligned}
\mathcal{C}_{Y_T}^{\mathbb{Q}}(u, \mid \mathcal{F}_t) &= \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp\left(uY_T\right) \mid \mathcal{F}_t\right]\right) \\
&= uY_t + \mathcal{A}^*\left(u,\mathbf{0};t,T\right) + \mathcal{B}_0^*\left(u,\mathbf{0};t,T\right)y_t + \boldsymbol{\mathcal{B}}^*\left(u,\mathbf{0};t,T\right)^{\top}\boldsymbol{l}_t^*.
\end{aligned}
\tag{D.3.2}
$$

The n^{th} cumulant of $Y_T = \log(S_T)$ under the risk-neutral measure $\mathbb{Q}$ is obtained by taking the the nth derivative of Equation (D.3.2). Thus the nth cumulant is defined by

$$
\kappa_n = \left.\frac{\partial^n \mathcal{C}_{Y_T}^{\mathbb{Q}}\left(u \mid \mathcal{F}_t\right)}{\partial u^n}\right|_{u=g(0)}.
$$

Now the nth cumulant of Y_t involves the n^{th} derivatives of $A^*\left(u,\boldsymbol{v};t,T\right)$, $B_0^*\left(u,\boldsymbol{v};t,T\right)$ and $\boldsymbol{B}^*\left(u,\boldsymbol{v};t,T\right)$ with respect to u. We will now derive the first four derivatives of those terms with respect to u. In this section, we will use a similar notation as Badescu et al. (2019) in order to avoid cumbersome expressions in the cumulant derivations. For any $i,j = 0,1,2,\ldots$

$$\mathcal{A}^*_i(x;t,T) = \left.\frac{\partial^i \mathcal{A}^*(u;t,T)}{\partial u^i}\right|_{u=x},$$
$$\mathcal{B}^*_{0,i}(x;t,T) = \left.\frac{\partial^i \mathcal{B}^*_0(u;t,T)}{\partial u^i}\right|_{u=x},$$
$$\boldsymbol{\mathcal{B}}^*_i(x;t,T) = \left.\frac{\partial^i \boldsymbol{\mathcal{B}}^*(u;t,T)}{\partial u^i}\right|_{u=x}.$$

In the following derivations, we assume an arbitrary N number of volatility processes. Since $\boldsymbol{v} = \boldsymbol{0}$ in Equation (D.3.2), we can further simplify notation by considering the multi-step coefficients as functions of only u. Thus for $k = 1,...,N$:

$$\begin{aligned}
\mathcal{M}^* &:= \mathcal{A}^*(u,\boldsymbol{v};t,t+1) = \mathcal{A}^*(u,v_1,...,v_N;t,t+1),\\
\mathcal{N}^* &:= \mathcal{B}^*_0(u,\boldsymbol{v};t,t+1) = \mathcal{B}^*_0(u,v_1,...,v_N;t,t+1),\\
\mathcal{O}^*_k &:= \mathcal{B}^*_k(u,\boldsymbol{v};t,t+1) = \mathcal{B}^*_k(u,v_1,...,v_N;t,t+1),\\
\boldsymbol{\mathcal{O}}^* &:= \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,t+1) = \begin{bmatrix}\mathcal{O}^*_1 & \ldots & \mathcal{O}^*_N\end{bmatrix}^\top,\\
\boldsymbol{h}_{t+1}(u) &= \begin{bmatrix}g(u)+\mathcal{B}^*_0(u,\boldsymbol{0};t+1,T) & \mathcal{B}^*_1(u,\boldsymbol{0};t+1,T) & \ldots & \mathcal{B}^*_N(u,\boldsymbol{0};t+1,T)\end{bmatrix}^\top,
\end{aligned} \tag{D.3.3}$$

where $g:\mathbb{R}\to\mathbb{R}$. We have

$$\begin{aligned}
\mathcal{A}^*\left(g(u)+\mathcal{B}^*_0\left(u,\boldsymbol{0};t+1,T\right),\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{0};t,T);t,t+1\right) &= \left(\mathcal{M}^*\circ \boldsymbol{h}_{t+1}\right)(u),\\
\mathcal{B}^*_0\left(g(u)+\mathcal{B}^*_0\left(u,\boldsymbol{0};t+1,T\right),\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{0};t,T);t,t+1\right) &= \left(\mathcal{N}^*\circ \boldsymbol{h}_{t+1}\right)(u),\\
\boldsymbol{\mathcal{B}}^*\left(g(u)+\mathcal{B}^*_0\left(u,\boldsymbol{0};t+1,T\right),\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{0};t,T);t,t+1\right) &= \left(\boldsymbol{\mathcal{O}}^*\circ \boldsymbol{h}_{t+1}\right)(u).
\end{aligned} \tag{D.3.4}$$

Finally, we introduce the subsequent notation to compliment the previous one:

$$\begin{aligned}
\mathcal{M}^*_{ij}(x) &= \left.\frac{\partial^{i+j}\mathcal{A}^*(u,\boldsymbol{v};t,t+1)}{\partial u^i\partial \boldsymbol{v}^j}\right|_{u=g(x)+\mathcal{B}^*_0(x,\boldsymbol{0};t+1,T),\boldsymbol{v}=\boldsymbol{\mathcal{B}}^*(x,\boldsymbol{0};t,T)},\\
\mathcal{N}^*_{ij}(x) &= \left.\frac{\partial^{i+j}\mathcal{B}^*_0(u,\boldsymbol{v};t,t+1)}{\partial u^i\partial \boldsymbol{v}^j}\right|_{u=g(x)+\mathcal{B}^*_0(x,\boldsymbol{0};t+1,T),\boldsymbol{v}=\boldsymbol{\mathcal{B}}^*(x,\boldsymbol{0};t,T)},\\
\mathcal{O}^*_{k,ij}(x) &= \left.\frac{\partial^{i+j}\mathcal{B}^*_k(u,\boldsymbol{v};t,t+1)}{\partial u^i\partial \boldsymbol{v}^j}\right|_{u=g(x)+\mathcal{B}^*_0(x,\boldsymbol{0};t+1,T),\boldsymbol{v}=\boldsymbol{\mathcal{B}}^*(x,\boldsymbol{0};t,T)},\\
\boldsymbol{\mathcal{O}}^*_{ij}(x) &= \left.\frac{\partial^{i+j}\boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,t+1)}{\partial u^i\partial \boldsymbol{v}^j}\right|_{u=g(x)+\mathcal{B}^*_0(x,\boldsymbol{0};t+1,T),\boldsymbol{v}=\boldsymbol{\mathcal{B}}^*(x,\boldsymbol{0};t,T)}.
\end{aligned} \tag{D.3.5}$$

We will omit mentioning the following fact in the derivations as it is used extensively $u = g(0), \boldsymbol{v} = \boldsymbol{0}$ for all $t \leqslant T$: $\mathcal{B}^*_0(0,\boldsymbol{0};t,T) = 0$ and $\boldsymbol{\mathcal{B}}^*(0,\boldsymbol{0};t,T) = \boldsymbol{0}$. Finally, let use define the following four abbreviations to alleviate the length of the derived expressions:

$$\mathcal{G}^*_{0,1}(x;t+1,T) = \frac{\partial g}{\partial u}(x) + \mathcal{B}^*_{0,1}(x;t+1,T), \tag{D.3.6}$$

$$\mathcal{G}^*_{0,2}(x;t+1,T) = \frac{\partial^2 g}{\partial u^2}(x) + \mathcal{B}^*_{0,2}(x;t+1,T), \tag{D.3.7}$$

$$\mathcal{G}^*_{0,3}(x;t+1,T) = \frac{\partial^3 g}{\partial u^3}(x) + \mathcal{B}^*_{0,3}(x;t+1,T), \tag{D.3.8}$$

$$\mathcal{G}^*_{0,4}(x;t+1,T) = \frac{\partial^4 g}{\partial u^4}(x) + \mathcal{B}^*_{0,4}(x;t+1,T). \tag{D.3.9}$$

D.3.1 First derivative

We have

$$\begin{aligned}\mathcal{A}^*_1(x;t,T) &= \mathcal{A}^*_1(x;t+1,T) + J_{\mathcal{M}^*}(x)J_{\boldsymbol{h}_{t+1}}(x) \\ &= \mathcal{A}^*_1(x;t+1,T) + \mathcal{M}^*_{10}(x)\mathcal{G}^*_{0,1}(x;t+1,T) \\ &\quad + \boldsymbol{\mathcal{M}}^*_{01}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T),\end{aligned} \tag{D.3.10}$$

$$\begin{aligned}\mathcal{B}^*_{0,1}(x;t,T) &= J_{\mathcal{N}^*}(x)J_{\boldsymbol{h}_{t+1}}(x) \\ &= \mathcal{N}^*_{10}(x)\mathcal{G}^*_{0,1}(x;t+1,T) + \boldsymbol{\mathcal{N}}^*_{01}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T),\end{aligned} \tag{D.3.11}$$

$$\begin{aligned}\boldsymbol{\mathcal{B}}^*_1(x;t,T) &= J_{\boldsymbol{\mathcal{O}}^*}(x)J_{\boldsymbol{h}_{t+1}}(x) \\ &= \boldsymbol{\mathcal{O}}^*_{10}(x)\mathcal{G}^*_{0,1}(x;t+1,T) + \boldsymbol{\mathcal{O}}^*_{01}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T),\end{aligned} \tag{D.3.12}$$

with terminals constraints $\mathcal{A}^*_1(x;T,T) = 0$, $\mathcal{B}^*_{0,1}(x;T,T) = 0$ and $\boldsymbol{\mathcal{B}}^*_1(x;T,T) = \mathbf{0}$. Now if we assume $\mathcal{B}^*_0(u,v;t,T) = 0$ for all $t = 0,1,...,T$ and $g(u) = u$, we can obtain closed form equations for the recursions. Indeed, solving Equation (C.2.3) using the constraint $\boldsymbol{\mathcal{B}}^*_1(x;T,T) = \mathbf{0}$ and under the assumption that $\boldsymbol{\mathcal{O}}^*_{01}(x)$ satisfies Theorem A.2.1, we obtain

$$\boldsymbol{\mathcal{B}}^*_1(x;t,T) = (I_N - \boldsymbol{\mathcal{O}}^*_{01}(x))^{-1}\,(I_N - (\boldsymbol{\mathcal{O}}^*_{01}(x))^{T-t})\boldsymbol{\mathcal{O}}^*_{10}(x), \tag{D.3.13}$$

where $\boldsymbol{I}_N$ corresponds to the $N \times N$ identity matrix. Similarly, solving Equation (C.2.2) using the constraint $\mathcal{A}^*_1(T,T) = 0$, we obtain

$$\begin{aligned}&\mathcal{A}^*_1(x;t,T) \\ &\quad = \boldsymbol{\mathcal{M}}^*_{01}(x)\,(\boldsymbol{I}_N - \boldsymbol{\mathcal{O}}^*_{01}(x))^{-1}\left((T-t)\boldsymbol{I}_N - (\boldsymbol{I}_N - \boldsymbol{\mathcal{O}}^*_{01}(x))^{-1}\left(\boldsymbol{I}_N - (\boldsymbol{\mathcal{O}}^*_{01}(x))^{T-t}\right)\right)\boldsymbol{\mathcal{O}}^*_{10}(x) \\ &\qquad + (T-t)\mathcal{M}^*_{10}(x) \\ &\quad = (T-t)\mathcal{M}^*_{10}(x) + \boldsymbol{\mathcal{M}}^*_{01}(x)\,(\boldsymbol{I}_N - \boldsymbol{\mathcal{O}}^*_{01}(x))^{-1}\,((T-t)\boldsymbol{\mathcal{O}}^*_{10}(x) - \boldsymbol{\mathcal{B}}^*_1(x;t,T))\,.\end{aligned} \tag{D.3.14}$$

D.3.2 Second Derivative

We have

$$\mathcal{A}_2^*(x;t,T) = \mathcal{A}_2^*(x;t+1,T) + J_{\boldsymbol{h}_{t+1}}(x)^\top H_{\mathcal{M}^*}(x) J_{\boldsymbol{h}_{t+1}}(x) + \sum_{i=0}^{N} \left[J_{\mathcal{M}^*}(x)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(x)$$

$$= \mathcal{A}_2^*(x;t+1,T) + \mathcal{M}_{20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2 \tag{D.3.15}$$

$$+ 2\boldsymbol{\mathcal{M}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\mathcal{G}_{0,1}^*(x;t+1,T) \tag{D.3.16}$$

$$+ \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$

$$+ \mathcal{M}_{10}^*(x)\mathcal{G}_{0,2}^*(x;t+1,T) + \boldsymbol{\mathcal{M}}_{01}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T),$$

where $H_{\boldsymbol{h}_{t+1}^{(i)}}$ denotes the second derivative of the i^{th} component of the vector-valued function h_{t+1}.

$$\mathcal{B}_{0,2}^*(x;t,T) = J_{\boldsymbol{h}_{t+1}}(x)^\top H_{\mathcal{N}^*}(x) J_{\boldsymbol{h}_{t+1}}(x) + \sum_{i=0}^{N} \left[J_{\mathcal{N}^*}(x)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(x)$$

$$= \mathcal{N}_{20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2 \tag{D.3.17}$$

$$+ 2\boldsymbol{\mathcal{N}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\mathcal{G}_{0,1}^*(x;t+1,T)$$

$$+ \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$

$$+ \mathcal{N}_{10}^*(x)\mathcal{G}_{0,2}^*(x;t+1,T)$$

$$+ \boldsymbol{\mathcal{N}}_{01}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T),$$

and

$$\boldsymbol{\mathcal{B}}_2^*(x;t,T) = \begin{bmatrix} H_{\mathcal{B}_1^*(x;t,T)} & \dots & H_{\mathcal{B}_N^*(x;t,T)} \end{bmatrix}^\top, \tag{D.3.18}$$

with terminals constraints $\mathcal{A}_2^*(x;T,T) = 0$, $\mathcal{B}_{0,2}^*(x;T,T) = 0$ and $\boldsymbol{\mathcal{B}}_2^*(x;T,T) = \mathbf{0}$ and where for $k = 1,\dots,N$ we have

$$H_{\mathcal{B}_k^*(x;t,T)} = J_{\boldsymbol{h}_{t+1}}(x)^\top H_{\mathcal{O}_k^*}(x) J_{\boldsymbol{h}_{t+1}}(x) + \sum_{i=0}^{N} \left[J_{\mathcal{O}_k^*}(x)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(x)$$

$$= \mathcal{O}_{k,20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2 \tag{D.3.19}$$

$$+ 2\boldsymbol{\mathcal{O}}_{k,11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\mathcal{G}_{0,1}^*(x;t+1,T)$$

$$+ \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)$$

$$+ \mathcal{O}_{k,10}^*(x)\mathcal{G}_{0,2}^*(x;t+1,T)$$

$$+ \boldsymbol{\mathcal{O}}_{k,01}(x)^*\boldsymbol{\mathcal{B}}_2^*(x;t+1,T).$$

D.3.3 Third Derivative

Before proceding with the proofs of the third and fourth derivatives of the multi-step coefficients of the cumulants of the log-price process Y_T, we derive a useful theorem which simplifies notation and the proofs:

Theorem D.3.1. *Let $f : \mathbb{R}^{N+1} \to \mathbb{R}$ and $\boldsymbol{h} : \mathbb{R} \to \mathbb{R}^{N+1}$ be a vector valued function such that*

$$\boldsymbol{h}(x) = \begin{bmatrix} h_0(x) & h_1(x) & \dots & h_N(x) \end{bmatrix}^\top = h_0 +\!\!+ \begin{bmatrix} h_1 & \dots & h_N \end{bmatrix}^\top = h_0 +\!\!+ \boldsymbol{g}^\top,$$

where $+\!\!+$ stands for concatenation of two sequences and $\boldsymbol{g}^\top = \begin{bmatrix} h_1 & \dots & h_N \end{bmatrix}^\top$ then the following expressions are equivalent

$$\begin{aligned}\operatorname{vec}(\frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x) \otimes J_{\boldsymbol{h}}(x) J_{\boldsymbol{h}}(x)\top)^\top \operatorname{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right) &= \operatorname{vec}(J_{\boldsymbol{h}}(x) \otimes \frac{\partial^2 h}{\partial x^2}(x) J_{\boldsymbol{h}}(x)^\top)^\top \operatorname{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right) \\ &= \operatorname{vec}(J_{\boldsymbol{h}}(x) \otimes J_{\boldsymbol{h}}(x) \frac{\partial^2 h}{\partial x^2}(x)^\top)^\top \operatorname{vec}\left(\Upsilon_f(\boldsymbol{h}(x))\right),\end{aligned}$$

where $\otimes$ stands for the Kronecker product, $J_{\boldsymbol{h}}(x)$ stands for the Jacobian of $\boldsymbol{h}$, $\frac{\partial^2 \boldsymbol{h}}{\partial x^2}(x)$ stands for the vector of partial-second derivatives of $\boldsymbol{h}$, vec *stands for the column-wise vectorization operator of a matrix and $\Upsilon_f(\boldsymbol{h}(x))$ represents the 3-d tensor formed by all 3rd degree partial derivatives of f evaluated at $\boldsymbol{h}(x)$.*[1]

Proof. See Appendix D.2. □

We have

$$\begin{aligned}\mathcal{A}_3^*(x;t,T) = \mathcal{A}_3^*(x;t+1,T) &+ \frac{\partial}{\partial u}\left[J_{\boldsymbol{h}_{t+1}}(u)^\top H_{\mathcal{M}^*}(u) J_{\boldsymbol{h}_{t+1}}(u)\right]_{u=x} \\ &+ \sum_{i=0}^{N} \frac{\partial}{\partial u}\left[\left[J_{\mathcal{M}^*}(u)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x},\end{aligned} \tag{D.3.20}$$

and using Theorem A.2.3 and the symmetricity of the Hessian we obtain

$$\begin{aligned}&\frac{\partial}{\partial u}\left[J_{\boldsymbol{h}_{t+1}}(u)^\top H_{\mathcal{M}^*}(u) J_{\boldsymbol{h}_{t+1}}(u)\right]_{u=x} \\ &= \left[2\frac{\partial J_{\boldsymbol{h}_{t+1}}(u)^\top}{\partial u} H_{\mathcal{M}^*}(u) J_{\boldsymbol{h}_{t+1}}(u) + \operatorname{vec}\left(J_{\boldsymbol{h}_{t+1}}(u) J_u^\top(u)\right)^\top \frac{\partial \operatorname{vec}\left(H_{\mathcal{M}^*}(u)\right)}{\partial u}\right]_{u=x} \\ &= 2\sum_{i=0}^{N} H_{\boldsymbol{h}_{t+1}^{(i+1)}}(x)\left[H_{\mathcal{M}^*}(x) J_{\boldsymbol{h}_{t+1}}(x)\right]_i + \operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(x) \otimes J_{\boldsymbol{h}_{t+1}}(x) J_{\boldsymbol{h}_{t+1}}(x)^\top)^\top \operatorname{vec}\left(\Upsilon_{M^*}(x)\right) \\ &= 2\frac{\partial^2 h_{t+1}}{\partial u^2}(x) H_{\mathcal{M}^*}(x) J_{\boldsymbol{h}_{t+1}}(x) + \operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(x) \otimes J_{\boldsymbol{h}_{t+1}}(x) J_{\boldsymbol{h}_{t+1}}(x)^\top)^\top \operatorname{vec}\left(\Upsilon_{\mathcal{M}^*}(x)\right)\end{aligned}$$

[1] The tensor is vectorized analogously to how matrices are vectorized using the vec operator defined in Theorem A.2.2. Please refer to the following Wikipedia page on Tensor Reshaping for more information.

$$
\begin{aligned}
&= 2\mathcal{G}^*_{0,2}(x;t+1,T)\mathcal{M}^*_{20}(x)\mathcal{G}^*_{0,1}(x;t+1,T)\\
&\qquad + 2\mathcal{G}^*_{0,2}(x;t+1,T)\boldsymbol{\mathcal{M}}^*_{11}(x)^\top\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\qquad + 2\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{02}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\qquad + 2\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{11}(x)\mathcal{G}^*_{0,1}(x;t+1,T)\\
&\qquad + \mathcal{M}^*_{30}(x)\mathcal{G}^*_{0,1}(x;t+1,T)^3\\
&\qquad + 3\mathcal{G}^*_{0,1}(x;t+1,T)^2\boldsymbol{\mathcal{M}}^*_{21}(x)^\top\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\qquad + 3\mathcal{G}^*_{0,1}(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{12}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\qquad + \mathrm{vec}(\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\otimes\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top)^\top\,\mathrm{vec}\left(\boldsymbol{\mathcal{M}}^*_{03}(x)\right),
\end{aligned}
$$

where $\Upsilon_{M^*}(x)$ represents the 3-d tensor formed by all 3rd degree partial derivatives of $\mathcal{M}^*$ and $\mathrm{vec}\left(\boldsymbol{\mathcal{M}}^*_{03}\right)$ represents a vectorized version of the 3-d tensor formed by all 3rd degree partial derivatives of $\mathcal{M}^*$ with respect to $\boldsymbol{v}$.[2] .We will use Υ to refer to the matrix of third partial derivatives of a certain function from now on. We also obtain

$$
\begin{aligned}
&\sum_{i=0}^{N}\frac{\partial}{\partial u}\left[\left[J_{\mathcal{M}^*}(u)\right]_i H_{\boldsymbol{h}^{(i+1)}_{t+1}}(u)\right]_{u=x}\\
&\qquad = \sum_{i=0}^{N}\left[J_{\boldsymbol{h}_{t+1}}(x)^\top H_{\mathcal{M}^*}(x)\right]_i H_{\boldsymbol{h}^{(i+1)}_{t+1}}(x) + \sum_{i=0}^{N}\left[J_{\mathcal{M}^*}(x)\right]_i \Upsilon_{\boldsymbol{h}^{(i+1)}_{t+1}}(x),
\end{aligned}
$$

where

$$
\begin{aligned}
&\sum_{i=0}^{N}\left[J_{\boldsymbol{h}_{t+1}}(x)^\top H_{\mathcal{M}^*}(x)\right]_i H_{\boldsymbol{h}^{(i+1)}_{t+1}}(x)\\
&\qquad = J_{\boldsymbol{h}_{t+1}}(x)^\top H_{\mathcal{M}^*}(x)\frac{\partial^2 h_{t+1}}{\partial u^2}(x)\\
&\qquad = \mathcal{G}^*_{0,1}(x;t+1,T)\mathcal{M}^*_{20}(x)\mathcal{G}^*_{0,2}(x;t+1,T)\\
&\qquad\qquad + \boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{11}(x)\mathcal{G}^*_{0,2}(x;t+1,T)\\
&\qquad\qquad + \mathcal{G}^*_{0,1}(x;t+1,T)\boldsymbol{\mathcal{M}}^*_{11}(x)^\top\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)\\
&\qquad\qquad + \boldsymbol{\mathcal{B}}^*_1(x;t+1,T)^\top\boldsymbol{\mathcal{M}}^*_{02}(x)\boldsymbol{\mathcal{B}}^*_2(x;t+1,T),
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{i=0}^{N}\left[J_{\mathcal{M}^*}(x)\right]_i \Upsilon_{\boldsymbol{h}^{(i+1)}_{t+1}}(x) &= J_{\mathcal{M}^*}(x)\frac{\partial^3 h_{t+1}}{\partial u^3}(x)\\
&= \mathcal{M}^*_{10}(x)\mathcal{G}^*_{0,3}(x;t+1,T) + \boldsymbol{\mathcal{M}}^*_{01}(x)\boldsymbol{\mathcal{B}}^*_3(x;t+1,T),
\end{aligned}
$$

[2]These tensors are vectorized analogously to how matrices are vectorized. Please refer to Tensor Reshaping.

which allows us to obtain the following recursion:

$$
\begin{aligned}
&\mathcal{A}_3^*(x;t,T)\\
&= \mathcal{A}_3^*(x;t+1,T) + 3\mathcal{G}_{0,2}^*(x;t+1,T)\mathcal{M}_{20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad + 3\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{M}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{11}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad + \mathcal{M}_{30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^3\\
&\quad + 3\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{M}}_{21}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 3\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + \operatorname{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \operatorname{vec}\left(\boldsymbol{\mathcal{M}}_{03}^*(x)\right)\\
&\quad + \mathcal{M}_{10}^*(x)\mathcal{G}_{0,3}^*(x;t+1,T) + \boldsymbol{\mathcal{M}}_{01}^*(x)\boldsymbol{\mathcal{B}}_3^*(x;t+1,T).
\end{aligned}
\tag{D.3.21}
$$

Using similar steps, we obtain the following recursion

$$
\begin{aligned}
\mathcal{B}_{0,3}^*(x;t,T) &= \frac{\partial}{\partial u}\left[J_{\boldsymbol{h}_{t+1}}(u)^\top H_{\mathcal{N}^*}(u)J_{\boldsymbol{h}_{t+1}}(u)\right]_{u=x} + \sum_{i=0}^{N}\frac{\partial}{\partial u}\left[\left[J_{\mathcal{N}^*}(u)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x}\\
&= 3\mathcal{G}_{0,2}^*(x;t+1,T)\mathcal{N}_{20}^*(x)\left(1+\mathcal{B}_{0,1}^*(x;t+1,T)\right)\\
&\quad + 3\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{N}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{11}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad + \mathcal{N}_{30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^3\\
&\quad + 3\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{N}}_{21}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 3\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + \operatorname{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \operatorname{vec}\left(\boldsymbol{\mathcal{N}}_{03}^*(x)\right)\\
&\quad + \mathcal{N}_{10}^*(x)\mathcal{G}_{0,3}^*(x;t+1,T) + \boldsymbol{\mathcal{N}}_{01}^*(x)\boldsymbol{\mathcal{B}}_3^*(x;t+1,T).
\end{aligned}
\tag{D.3.22}
$$

Now

$$
\boldsymbol{\mathcal{B}}_3^*(x;t,T) = \left[\Upsilon_{\mathcal{B}_1^*(x;t,T)} \quad \ldots \quad \Upsilon_{\mathcal{B}_N^*(x;t,T)}\right]^\top,
\tag{D.3.23}
$$

where for $k = 1,\ldots,N$ we have

$$
\begin{aligned}
\Upsilon_{\mathcal{B}_k^*(x;t,T)} &= \frac{\partial}{\partial u}\left[J_{\boldsymbol{h}_{t+1}}(u)^\top H_{\mathcal{O}_k^*}(u)J_{\boldsymbol{h}_{t+1}}(u)\right]_{u=x} + \sum_{i=0}^{N}\frac{\partial}{\partial u}\left[\left[J_{\mathcal{O}_k^*}(u)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x}\\
&= 3\mathcal{G}_{0,2}^*(x;t+1,T)\mathcal{O}_{k,20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)
\end{aligned}
$$

$$
\begin{aligned}
&3\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{O}}_{k,11}^*(x)^\top\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top\boldsymbol{\mathcal{O}}_{k,02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top\boldsymbol{\mathcal{O}}_{k,11}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&+\mathcal{O}_{k,30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^3\\
&+3\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{O}}_{k,21}^*(x)^\top\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+3\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\boldsymbol{\mathcal{O}}_{k,12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+\operatorname{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top\operatorname{vec}\left(\boldsymbol{\mathcal{O}}_{k,03}^*(x)\right)\\
&+\mathcal{O}_{k,10}^*(x)\mathcal{G}_{0,3}^*(x;t+1,T)+\boldsymbol{\mathcal{O}}_{k,01}^*(x)\boldsymbol{\mathcal{B}}_3^*(x;t+1,T).
\end{aligned}
\tag{D.3.24}
$$

D.3.4 Fourth Derivative

We have

$$
\begin{aligned}
\mathcal{A}_4^*(x;t,T)=\mathcal{A}_4^*(x;t+1,T)&+2\sum_{i=0}^N\frac{\partial}{\partial u}\left[H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\left[H_{\mathcal{M}^*}(u)J_{\boldsymbol{h}_{t+1}}(u)\right]_i\right]_{u=x}\\
&+\frac{\partial}{\partial u}\left[\operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(u)\otimes J_{\boldsymbol{h}_{t+1}}(u)J_{\boldsymbol{h}_{t+1}}(u)^\top)^\top\operatorname{vec}\left(\Upsilon_{\mathcal{M}^*}(u)\right)\right]_{u=x}\\
&+\sum_{i=0}^N\frac{\partial}{\partial u}\left[\left[J_{\boldsymbol{h}_{t+1}}(u)^\top H_{\mathcal{M}^*}(u)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x}\\
&+\sum_{i=0}^N\frac{\partial}{\partial u}\left[\left[J_{\mathcal{M}^*}(u)\right]_i\Upsilon_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x},
\end{aligned}
\tag{D.3.25}
$$

and using Theorem A.2.3 we obtain

$$
\begin{aligned}
&2\sum_{i=0}^N\frac{\partial}{\partial u}\left[H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\left[H_{\mathcal{M}^*}(u)J_{\boldsymbol{h}_{t+1}}(u)\right]_i\right]_{u=x}\\
&=2\sum_{i=0}^N\Upsilon_{\boldsymbol{h}_{t+1}^{(i+1)}}(x)\left[H_{\mathcal{M}^*}(x)J_{\boldsymbol{h}_{t+1}}(x)\right]_i+2\sum_{i=0}^N\sum_{j=0}^N H_{\boldsymbol{h}_{t+1}^{(i+1)}}(x)\left[H_{\mathcal{M}^*}(x)\right]_{ij}H_{\boldsymbol{h}_{t+1}^{(j+1)}}(x)\\
&\quad+2\sum_{i=0}^N\sum_{j=0}^N\sum_{k=0}^N\left[J_{\boldsymbol{h}_{t+1}}(x)\right]_i H_{\boldsymbol{h}_{t+1}^{(k+1)}}(x)\left[J_{\boldsymbol{h}_{t+1}}(x)\right]_j\left[\operatorname{vec}\left(\Upsilon_{\mathcal{M}^*}(x)\right)\right]_{k+(N+1)\times j+(N+1)^2\times i}\\
&=2\frac{\partial^3 h_{t+1}}{\partial u^3}(x)^\top H_{\mathcal{M}^*}(x)J_{\boldsymbol{h}_{t+1}}(x)+2\frac{\partial^2 h_{t+1}}{\partial u^2}(x)^\top H_{\mathcal{M}^*}(x)\frac{\partial^2 h_{t+1}}{\partial u^2}(x)\\
&\quad+2\operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(x)\otimes\frac{\partial^2 h_{t+1}}{\partial u^2}(x)J_{\boldsymbol{h}_{t+1}}(x)^\top)^\top\operatorname{vec}\left(\Upsilon_{\mathcal{M}^*}(x)\right)\\
&=2\mathcal{B}_{0,3}^*(x;t+1,T)\mathcal{M}_{20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad+2\mathcal{B}_{0,3}^*(x;t+1,T)\boldsymbol{\mathcal{M}}_{11}^*(x)^\top\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad+2\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top\boldsymbol{\mathcal{M}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)
\end{aligned}
$$

$$
\begin{aligned}
&+ 2\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{11}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T) \\
&+ 2\mathcal{G}_{0,2}^*(x;t+1,T)^2\mathcal{M}_{20}^*(x) \qquad\qquad (\text{D.3.26}) \\
&+ 2\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{M}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_2^*(x;t+1,T) \\
&+ 2\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{02}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T) \\
&+ 2\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{11}^*(x)\mathcal{G}_{0,2}^*(x;t+1,T) \\
&+ 2\mathcal{M}_{30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2\mathcal{G}_{0,2}^*(x;t+1,T) \\
&+ 2\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{21}^*(x) \\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{21}^*(x) \\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 2\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T) \\
&+ 2\operatorname{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \operatorname{vec}\left(\boldsymbol{\mathcal{M}}_{03}^*(x)\right).
\end{aligned}
$$

We also obtain using Theorem D.3.1

$$
\begin{aligned}
&\frac{\partial}{\partial u}\left[\operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(u)\otimes J_{\boldsymbol{h}_{t+1}}(u)J_{\boldsymbol{h}_{t+1}}(u)^\top)^\top \operatorname{vec}\left(\Upsilon_{\mathcal{M}^*}(u)\right)\right]_{u=g(0)} \\
&= 3\operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(x)\otimes \frac{\partial^2 h_{t+1}}{\partial u^2}(x)J_{\boldsymbol{h}_{t+1}}(x)^\top)^\top \operatorname{vec}\left(\Upsilon_{M^*}(x)\right) \\
&\qquad + \operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(x)\otimes J_{\boldsymbol{h}_{t+1}}(x)\otimes J_{\boldsymbol{h}_{t+1}}(x)J_{\boldsymbol{h}_{t+1}}(x)^\top)^\top \operatorname{vec}\left(\Phi_{\mathcal{M}^*}(x)\right) \\
&= 3\mathcal{M}_{30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2\mathcal{G}_{0,2}^*(x;t+1,T) \\
&\qquad + 6\mathcal{G}_{0,1}^*(x;t+1,T)\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(t+1,T)^\top \boldsymbol{\mathcal{M}}_{21}^*(x) \\
&\qquad + 3\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_2^*(t+1,T)^\top \boldsymbol{\mathcal{M}}_{21}^*(x) \\
&\qquad + 6\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_2^*(t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(t+1,T) \\
&\qquad + 3\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(t+1,T) \\
&\qquad + 3\operatorname{vec}(\boldsymbol{\mathcal{B}}_2^*(t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(t+1,T)\boldsymbol{\mathcal{B}}_1^*(t+1,T)^\top)^\top \operatorname{vec}\left(\boldsymbol{\mathcal{M}}_{03}^*(x)\right) \\
&\qquad + \mathcal{G}_{0,1}^*(x;t+1,T)^4\mathcal{M}_{40}^* + 4\mathcal{G}_{0,1}^*(x;t+1,T)^3\boldsymbol{\mathcal{B}}_1^*(t+1,T)^\top \boldsymbol{\mathcal{M}}_{31}^* \\
&\qquad + 6\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_1^*(t+1,T)^\top \boldsymbol{\mathcal{M}}_{22}^*\boldsymbol{\mathcal{B}}_1^*(t+1,T) \\
&\qquad + 4\mathcal{G}_{0,1}^*(x;t+1,T)\operatorname{vec}\left(\boldsymbol{\mathcal{B}}_1^*(t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(t+1,T)\boldsymbol{\mathcal{B}}_1^*(t+1,T)^\top\right)^\top \operatorname{vec}\left(\boldsymbol{\mathcal{M}}_{13}^*\right) \\
&\qquad + \operatorname{vec}\left(\boldsymbol{\mathcal{B}}_1^*(t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(t+1,T)\boldsymbol{\mathcal{B}}_1^*(t+1,T)^\top\right)^\top \operatorname{vec}\left(\boldsymbol{\mathcal{M}}_{04}^*\right),
\end{aligned}
$$

where $\Phi_{M^*}(x)$ represents the 4-d tensor formed by all 4th degree partial derivatives of $\mathcal{M}^*$, $\operatorname{vec}(\boldsymbol{\mathcal{M}}_{04}^*)$ represents a vectorized version of the 4-d tensor formed by all 4th degree partial derivatives of $\mathcal{M}^*$ with respect to $\boldsymbol{v}$ and $\mathcal{M}^*$, $\operatorname{vec}(\boldsymbol{\mathcal{M}}_{13}^*)$ represents a vectorized version of the 3-d tensor formed by all 4th degree partial derivatives of $\mathcal{M}^*$ of the form $\frac{\partial^4}{\partial u \partial v_i v_j v_k}$ where

$1 \leqslant i,j,k, \leqslant N$. Now, using Theorem A.2.3, the symmetry of the Hessian matrix and by Theorem D.3.1 we obtain

$$\sum_{i=0}^{N} \frac{\partial}{\partial u}\left[\left[J_{\boldsymbol{h}_{t+1}}(u)^{\top} H_{\mathcal{M}^*}(u)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x} = \frac{\partial}{\partial u}\left[J_{\boldsymbol{h}_{t+1}}(u)^{\top} H_{\mathcal{M}^*}(u) \frac{\partial^2 h_{t+1}}{\partial u^2}(u)\right]_{u=x}$$

$$= \frac{\partial^2 h_{t+1}}{\partial u^2}(x)^{\top} H_{\mathcal{M}^*}(x) \frac{\partial^2 h_{t+1}}{\partial u^2}(x) + J_{\boldsymbol{h}_{t+1}}(x)^{\top} H_{\mathcal{M}^*}(x) \frac{\partial^3 h_{t+1}}{\partial u^3}(x)$$
$$+ \left[\operatorname{vec}\left(J_{\boldsymbol{h}_{t+1}}(u) \frac{\partial^2 h_{t+1}}{\partial u^2}(u)^{\top}\right)^{\top} \frac{\partial}{\partial u} \operatorname{vec}\left(H_{\mathcal{M}^*}(u)\right)\right]_{u=x}$$

$$= \frac{\partial^2 h_{t+1}}{\partial u^2}(x)^{\top} H_{\mathcal{M}^*}(x) \frac{\partial^2 h_{t+1}}{\partial u^2}(x) + J_{\boldsymbol{h}_{t+1}}(x)^{\top} H_{\mathcal{M}^*}(x) \frac{\partial^3 h_{t+1}}{\partial u^3}(x)$$
$$+ \operatorname{vec}\left(J_{\boldsymbol{h}_{t+1}}(x) \otimes J_{\boldsymbol{h}_{t+1}}(x) \frac{\partial^2 h_{t+1}}{\partial u^2}(x)^{\top}\right)^{\top} \operatorname{vec}\left(\Upsilon_{\mathcal{M}^*}(x)\right),$$

which is equivalent to Equation (D.3.26) without the multiplier 2. Due to the symmetry of the Hessian matrix once again, we obtain

$$\sum_{i=0}^{N} \frac{\partial}{\partial u}\left[\left[J_{\mathcal{M}^*}(u)\right]_i \Upsilon_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x}$$
$$= \frac{\partial}{\partial u}\left[J_{\mathcal{M}^*}(u) \frac{\partial^3 h_{t+1}}{\partial u^3}(u)\right]_{u=x} = J_{\boldsymbol{h}_{t+1}}(x)^{\top} H_{\mathcal{M}^*}(x) \frac{\partial^3 h_{t+1}}{\partial u^3}(x) + J_{\mathcal{M}^*}(x) \frac{\partial^4 h_{t+1}}{\partial u^4}(x)$$
$$= \mathcal{G}^*_{0,3}(x;t+1,T)\mathcal{M}^*_{20}(x)\mathcal{G}^*_{0,1}(x;t+1,T)$$
$$+ \mathcal{G}^*_{0,3}(x;t+1,T)\boldsymbol{\mathcal{M}}^*_{11}(x)^{\top}\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)$$
$$+ \boldsymbol{\mathcal{B}}^*_3(x;t+1,T)^{\top}\boldsymbol{\mathcal{M}}^*_{02}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)$$
$$+ \boldsymbol{\mathcal{B}}^*_3(x;t+1,T)^{\top}\boldsymbol{\mathcal{M}}^*_{11}(x)\mathcal{G}^*_{0,1}(x;t+1,T)$$
$$+ \mathcal{M}^*_{10}(x)\mathcal{G}^*_{0,4}(x;t+1,T) + \boldsymbol{\mathcal{M}}^*_{01}(x)\boldsymbol{\mathcal{B}}^*_4(x;t+1,T),$$

which allows us to obtain the following recursion

$$\begin{aligned}
&\mathcal{A}^*_4(t,T)\\
&= \mathcal{A}^*_4(x;t+1,T) + 4\mathcal{G}^*_{0,3}(x;t+1,T)\mathcal{M}^*_{20}(x)\mathcal{G}^*_{0,1}(x;t+1,T)\\
&\quad + 4\mathcal{G}^*_{0,3}(x;t+1,T)\boldsymbol{\mathcal{M}}^*_{11}(x)^{\top}\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}^*_3(x;t+1,T)^{\top}\boldsymbol{\mathcal{M}}^*_{02}(x)\boldsymbol{\mathcal{B}}^*_1(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}^*_3(x;t+1,T)^{\top}\boldsymbol{\mathcal{M}}^*_{11}(x)\mathcal{G}^*_{0,1}(x;t+1,T)\\
&\quad + 3\mathcal{G}^*_{0,2}(x;t+1,T)^2\mathcal{M}^*_{20}(x) \qquad\qquad \text{(D.3.27)}\\
&\quad + 6\mathcal{G}^*_{0,2}(x;t+1,T)\boldsymbol{\mathcal{M}}^*_{11}(x)^{\top}\boldsymbol{\mathcal{B}}^*_2(x;t+1,T)
\end{aligned}$$

$$
\begin{aligned}
&+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{02}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T) && \text{(D.3.28)}\\
&+ 6\mathcal{M}_{30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2\mathcal{G}_{0,2}^*(x;t+1,T)\\
&+ 6\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{21}^*(x)\\
&+ 6\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 12\mathcal{G}_{0,1}^*(x;t+1,T)\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{21}^*(x)\\
&+ 12\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 6\,\mathrm{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \mathrm{vec}\left(\boldsymbol{\mathcal{M}}_{03}^*(x)\right)\\
&+ \mathcal{G}_{0,1}^*(x;t+1,T)^4\mathcal{M}_{40}^*(x)\\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)^3\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{31}^*(x)\\
&+ 6\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{M}}_{22}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)\,\mathrm{vec}\left(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\right)^\top \mathrm{vec}\left(\boldsymbol{\mathcal{M}}_{13}^*(x)\right)\\
&+ \mathrm{vec}\left(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\right)^\top \mathrm{vec}\left(\boldsymbol{\mathcal{M}}_{04}^*(x)\right)\\
&+ \mathcal{M}_{10}^*(x)\mathcal{G}_{0,4}^*(x;t+1,T) + \boldsymbol{\mathcal{M}}_{01}^*(x)\boldsymbol{\mathcal{B}}_4^*(x;t+1,T).
\end{aligned}
$$

Using similar steps, we have

$$
\begin{aligned}
\mathcal{B}_{0,4}^*(t,T) = 2\sum_{i=0}^{N}\frac{\partial}{\partial u}&\left[H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\left[H_{\mathcal{N}^*}(u)J_{\boldsymbol{h}_{t+1}}(u)\right]_i\right]_{u=x}\\
&+ \frac{\partial}{\partial u}\left[\mathrm{vec}(J_{\boldsymbol{h}_{t+1}}(u)\otimes J_{\boldsymbol{h}_{t+1}}(u)J_{\boldsymbol{h}_{t+1}}(u)^\top)^\top \mathrm{vec}\left(\Upsilon_{\mathcal{N}^*}(u)\right)\right]_{u=x}\\
&+ \sum_{i=0}^{N}\frac{\partial}{\partial u}\left[\left[J_{\boldsymbol{h}_{t+1}}(u)^\top H_{\mathcal{N}^*}(u)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x}\\
&+ \sum_{i=0}^{N}\frac{\partial}{\partial u}\left[\left[J_{\mathcal{N}^*}(u)\right]_i \Upsilon_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x},
\end{aligned}
$$

and obtain the following recursion

$$
\begin{aligned}
&\mathcal{B}_{0,4}^*(x;t,T) && \text{(D.3.29)}\\
&= 4\mathcal{G}_{0,3}^*(x;t+1,T)\mathcal{N}_{20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad + 4\mathcal{G}_{0,3}^*(x;t+1,T)\boldsymbol{\mathcal{N}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{11}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad + 3\mathcal{G}_{0,2}^*(x;t+1,T)^2\mathcal{N}_{20}^*(x)\\
&\quad + 6\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{N}}_{11}^*(x)^\top \boldsymbol{\mathcal{B}}_2^*(x;t+1,T) && \text{(D.3.30)}
\end{aligned}
$$

$$
\begin{aligned}
&+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{02}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)\\
&+ 6\mathcal{N}_{30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2\mathcal{G}_{0,2}^*(x;t+1,T)\\
&+ 6\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{21}^*(x)\\
&+ 6\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 12\mathcal{G}_{0,1}^*(x;t+1,T)\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{21}^*(x)\\
&+ 12\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 6\operatorname{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \operatorname{vec}(\boldsymbol{\mathcal{N}}_{03}^*(x))\\
&+ \mathcal{G}_{0,1}^*(x;t+1,T)^4\mathcal{N}_{40}^*(x)\\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)^3\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{31}^*(x)\\
&+ 6\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{N}}_{22}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)\operatorname{vec}\left(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\right)^\top \operatorname{vec}(\boldsymbol{\mathcal{N}}_{13}^*(x))\\
&+ \operatorname{vec}\left(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\right)^\top \operatorname{vec}(\boldsymbol{\mathcal{N}}_{04}^*(x))\\
&+ \mathcal{N}_{10}^*(x)\mathcal{G}_{0,4}^*(x;t+1,T) + \boldsymbol{\mathcal{N}}_{01}^*(x)\boldsymbol{\mathcal{B}}_4^*(x;t+1,T). \qquad \text{(D.3.31)}
\end{aligned}
$$

Now

$$
\boldsymbol{\mathcal{B}}_4^*(x;t,T) = \begin{bmatrix}\Phi_{\mathcal{B}_1^*(x;t,T)} & \ldots & \Phi_{\mathcal{B}_N^*(x;t,T)},\end{bmatrix}^\top \qquad \text{(D.3.32)}
$$

where for $k = 1,...,N$ we have

$$
\begin{aligned}
\Phi_{\mathcal{B}_k^*(x;t,T)} = {}& 2\sum_{i=0}^{N}\frac{\partial}{\partial u}\left[H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\left[H_{\mathcal{O}_k^*}(u)J_{\boldsymbol{h}_{t+1}}(u)\right]_i\right]_{u=x}\\
&+ \frac{\partial}{\partial u}\left[\operatorname{vec}(J_{\boldsymbol{h}_{t+1}}(u)\otimes J_{\boldsymbol{h}_{t+1}}(u)J_{\boldsymbol{h}_{t+1}}(u)^\top)^\top \operatorname{vec}(\Upsilon_{M^*}(u))\right]_{u=x}\\
&+ \sum_{i=0}^{N}\frac{\partial}{\partial u}\left[\left[J_{\boldsymbol{h}_{t+1}}(u)^\top H_{\mathcal{O}_k^*}(u)\right]_i H_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x} + \sum_{i=0}^{N}\frac{\partial}{\partial u}\left[\left[J_{\mathcal{O}_k^*}(u)\right]_i \Upsilon_{\boldsymbol{h}_{t+1}^{(i+1)}}(u)\right]_{u=x},
\end{aligned}
$$

and obtain the following recursions for for $k = 1,...,N$:

$$
\begin{aligned}
&\Phi_{\mathcal{B}_k^*(x;t,T)}\\
&= 4\mathcal{G}_{0,3}^*(x;t+1,T)\mathcal{O}_{k,20}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad + 4\mathcal{G}_{0,3}^*(x;t+1,T)\boldsymbol{\mathcal{O}}_{k,11}^*(x)^\top \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,02}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&\quad + 4\boldsymbol{\mathcal{B}}_3^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,11}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)\\
&\quad + 3\mathcal{G}_{0,2}^*(x;t+1,T)^2\mathcal{O}_{k,20}^*(x)\\
&\quad + 6\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{O}}_{k,11}^*(x)^\top \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)
\end{aligned}
$$

$$
\begin{aligned}
&+ 3\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,02}^*(x)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)\\
&+ 6\mathcal{O}_{k,30}^*(x)\mathcal{G}_{0,1}^*(x;t+1,T)^2\mathcal{G}_{0,2}^*(x;t+1,T)\\
&+ 6\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,21}^*(x) \qquad \text{(D.3.33)}\\
&+ 6\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 12\mathcal{G}_{0,1}^*(x;t+1,T)\mathcal{G}_{0,2}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,21}^*(x)\\
&+ 12\mathcal{G}_{0,1}^*(x;t+1,T)\boldsymbol{\mathcal{B}}_2^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,12}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 6\operatorname{vec}(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_2^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top)^\top \operatorname{vec}\big(\boldsymbol{\mathcal{O}}_{k,03}^*(x)\big)\\
&+ \mathcal{G}_{0,1}^*(x;t+1,T)^4\mathcal{O}_{k,40}^*(x)\\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)^3\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,31}^*(x)\\
&+ 6\mathcal{G}_{0,1}^*(x;t+1,T)^2\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top \boldsymbol{\mathcal{O}}_{k,22}^*(x)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\\
&+ 4\mathcal{G}_{0,1}^*(x;t+1,T)\operatorname{vec}\big(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\big)^\top \operatorname{vec}\big(\boldsymbol{\mathcal{O}}_{k,13}^*(x)\big)\\
&+ \operatorname{vec}\big(\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\otimes \boldsymbol{\mathcal{B}}_1^*(x;t+1,T)\boldsymbol{\mathcal{B}}_1^*(x;t+1,T)^\top\big)^\top \operatorname{vec}\big(\boldsymbol{\mathcal{O}}_{k,04}^*(x)\big)\\
&+ \mathcal{O}_{k,10}^*(x)\mathcal{G}_{0,4}^*(x;t+1,T) + \boldsymbol{\mathcal{O}}_{k,01}^*(x)\boldsymbol{\mathcal{B}}_4^*(x;t+1,T).
\end{aligned}
$$

D.4 Proof of Equation (1.2.9)

Using the integral representation of the payoff function of an option provided by Assumption 1, we obtain:

$$
\begin{aligned}
V_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[f(Y_T) \mid \mathcal{F}_t\right] &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{\pi}\int_0^\infty \operatorname{Re}\left[e^{(R+iu)Y_T}\check{f}(R+iu)\right]\, du \mid \mathcal{F}_t\right]\\
&= \frac{e^{-r(T-t)}}{\pi}\int_0^\infty \operatorname{Re}\left[\mathbb{E}^{\mathbb{Q}}\left[e^{(R+iu)Y_T} \mid \mathcal{F}_t\right]\check{f}(R+iu)\right]\, du\\
&= \frac{e^{-r(T-t)}}{\pi}\int_0^\infty \operatorname{Re}\left[\mathcal{C}_{Y_T}^{\mathbb{Q}}\left(R+iu \mid \mathcal{F}_t\right)\check{f}(R+iu)\right]\, du.
\end{aligned}
$$

D.5 Proof of Proposition 2.3.3

Given that the risk-neutral measure $\mathbb{Q}$ is defined as in Assumption 4, the denominator in Equation (2.2.2) can be simplified to obtain:

$$
\begin{aligned}
\operatorname{Var}^{\mathbb{Q}}[e^{-r}S_{t+1}|\mathcal{F}_t] &= e^{-2r}\mathbb{E}^{\mathbb{Q}}[S_{t+1}^2|\mathcal{F}_t] - \mathbb{E}^{\mathbb{Q}}[e^{-r}S_{t+1}|\mathcal{F}_t]^2\\
&= e^{-2r}\mathbb{E}^{\mathbb{Q}}[e^{2Y_{t+1}}|\mathcal{F}_t] - S_t^2.
\end{aligned}
$$

The numerator can also be reformulated in terms of two separate expectations:

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[H(e^{-r}S_{t+1}-S_t)|\mathcal{F}_t] &= \mathbb{E}^{\mathbb{Q}}[HS_{t+1}e^{-r}|\mathcal{F}_t] - S_t\mathbb{E}^{\mathbb{Q}}[H|\mathcal{F}_t]\\
&= \mathbb{E}^{\mathbb{Q}}[HS_te^{y_{t+1}}e^{-r}|\mathcal{F}_t] - S_t\mathbb{E}^{\mathbb{Q}}[H|\mathcal{F}_t]
\end{aligned}
$$

$$= S_t\mathbb{E}^{\mathbb{Q}}[He^{y_{t+1}-r}|\mathcal{F}_t] - S_t\mathbb{E}^{\mathbb{Q}}[H|\mathcal{F}_t]$$
$$= S_t\left(\mathbb{E}^{\mathbb{Q}}[He^{y_{t+1}-r}|\mathcal{F}_t] - \mathbb{E}^{\mathbb{Q}}[H|\mathcal{F}_t]\right).$$

We thus obtain the following equation for the hedge ratio in the affine setting of Augustyniak et al. (2023):

$$\theta^S_{t+1} = e^{-r(T-t)}S_t\frac{\mathbb{E}^{\mathbb{Q}}[He^{y_{t+1}-r}|\mathcal{F}_t] - \mathbb{E}^{\mathbb{Q}}[H|\mathcal{F}_t]}{e^{-2r}\mathbb{E}^{\mathbb{Q}}[e^{2Y_{t+1}}|\mathcal{F}_t] - S_t^2}. \tag{D.5.1}$$

The ingredients needed to derive a semi-explicit expression for the risk-minimizing hedging ratio in an affine setting are:

(1) The $\mathcal{F}_t$-conditional risk-neutral joint CGF of Y_{t+1} and $\boldsymbol{l}^*_{t+1}$:

$$\begin{aligned}\mathcal{C}^{\mathbb{Q}}_{Y_{t+1},\boldsymbol{l}^*_{t+1}}(u,\boldsymbol{v}\mid\mathcal{F}_t) &= \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp\left(uY_{t+1}+\boldsymbol{v}^\top\boldsymbol{l}^*_{t+1}\right)\mid\mathcal{F}_t\right]\right)\\ &= uY_t + \mathcal{A}^*(u,\boldsymbol{v};t,t+1) + \mathcal{B}^*_0(u,\boldsymbol{v};t,t+1)\,y_t + \boldsymbol{\mathcal{B}}^*(u,\boldsymbol{v};t,t+1)^\top\boldsymbol{l}^*_t.\end{aligned}$$

(2) The $\mathcal{F}_t$-conditional risk-neutral CGF of Y_T:

$$\begin{aligned}\mathcal{C}^{\mathbb{Q}}_{Y_T}(u,\mid\mathcal{F}_t) &= \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp\left(uY_T\right)\mid\mathcal{F}_t\right]\right)\\ &= uY_t + \mathcal{A}^*(u;t,T) + \mathcal{B}^*_0(u;t,T)\,y_t + \boldsymbol{\mathcal{B}}^*(u;t,T)^\top\boldsymbol{l}^*_t.\end{aligned}$$

The denominator in Equation (D.5.1) is given by:

$$e^{-2r}\,\mathbb{E}^{\mathbb{Q}}\left[e^{2Y_{t+1}}\mid\mathcal{F}_t\right] - S_t^2 = e^{-2r}\,\exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_{t+1},\boldsymbol{l}^*_{t+1}}(2,\mathbf{0}\mid\mathcal{F}_t)\right) - S_t^2.$$

To compute the second term in the numerator of Equation (D.5.1), $\mathbb{E}^{\mathbb{Q}}[H\mid\mathcal{F}_t]$, where $H = f(Y_T)$, we need to know the $\mathcal{F}_t$-conditional risk-neutral CGF of Y_T (which is available by assumption). To compute the first term in the numerator of Equation (D.5.1), $\mathbb{E}^{\mathbb{Q}}[He^{y_{t+1}-r}\mid\mathcal{F}_t]$, we define a new probability measure on $(\Omega,\mathcal{F}_T)$, denoted by $\widetilde{\mathbb{Q}}_t$, as follows:

$$\text{For any } A\in\mathcal{F}_T: \qquad \widetilde{\mathbb{Q}}_t(A) := \mathbb{E}^{\mathbb{Q}}\left[e^{y_{t+1}-r}\mathbf{1}_A\right].$$

Therefore, the Radon-Nikodym derivative of $\widetilde{\mathbb{Q}}_t$ with respect to $\mathbb{Q}$ is given by:

$$\left.\frac{d\widetilde{\mathbb{Q}}_t}{d\mathbb{Q}}\right|_{\mathcal{F}_T} = e^{y_{t+1}-r}$$

We can then write:

$$\mathbb{E}^{\widetilde{\mathbb{Q}}_t}[H\mid\mathcal{F}_t] = \frac{\mathbb{E}^{\mathbb{Q}}\left[H\frac{d\widetilde{\mathbb{Q}}_t}{d\mathbb{Q}}\,\middle|\,\mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{Q}}\left[\frac{d\widetilde{\mathbb{Q}}_t}{d\mathbb{Q}}\,\middle|\,\mathcal{F}_t\right]} = \mathbb{E}^{\mathbb{Q}}\left[He^{y_{t+1}-r}\mid\mathcal{F}_t\right].$$

To compute the above expression, we need the CGF of Y_T under $\widetilde{\mathbb{Q}}_t$:

$\mathcal{C}^{\widetilde{\mathbb{Q}}_t}_{Y_T}(u\mid\mathcal{F}_t)$

$$
\begin{aligned}
&= \log\left(\mathbb{E}^{\widetilde{\mathbb{Q}}_t}[\exp(uY_T) \mid \mathcal{F}_t]\right) = \log\left(\frac{\mathbb{E}^{\mathbb{Q}}\left[\exp(uY_T)\frac{d\widetilde{\mathbb{Q}}_t}{d\mathbb{Q}} \,\middle|\, \mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{Q}}\left[\frac{d\widetilde{\mathbb{Q}}_t}{d\mathbb{Q}} \,\middle|\, \mathcal{F}_t\right]}\right) \\
&= \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp(uY_T + y_{t+1}) \mid \mathcal{F}_t\right]\right) - r = \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp(y_{t+1})\mathbb{E}^{\mathbb{Q}}\left[\exp(uY_T) \mid \mathcal{F}_{t+1}\right] \mid \mathcal{F}_t\right]\right) - r \\
&= \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp(Y_{t+1} - Y_t)\exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_T}\left(u \mid \mathcal{F}_{t+1}\right)\right) \mid \mathcal{F}_t\right]\right) - r \\
&= \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp\left((1 + u + \mathcal{B}^*_0\left(u; t+1, T\right))Y_{t+1} + \mathcal{A}^*\left(u; t+1, T\right) + \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right)^\top \boldsymbol{l}^*_{t+1}\right) \,\middle|\, \mathcal{F}_t\right]\right) \\
&\quad - Y_t\left(1 + \mathcal{B}^*_0\left(u; t+1, T\right)\right) - r \\
&= \log\left(\mathbb{E}^{\mathbb{Q}}\left[\exp\left(\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right)\right)Y_{t+1} + \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right)^\top \boldsymbol{l}^*_{t+1}\right) \,\middle|\, \mathcal{F}_t\right]\right) \\
&\quad + \mathcal{A}^*\left(u; t+1, T\right) - Y_t\left(1 + \mathcal{B}^*_0\left(u; t+1, T\right)\right) - r \\
&= \mathcal{A}^*\left(u; t+1, T\right) - Y_t\left(1 + \mathcal{B}^*_0\left(u; t+1, T\right)\right) - r \\
&\quad + \mathcal{C}^{\mathbb{Q}}_{Y_{t+1}, \boldsymbol{l}^*_{t+1}}\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right), \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right) \mid \mathcal{F}_t\right) \\
&= \mathcal{A}^*\left(u; t+1, T\right) - Y_t\left(1 + \mathcal{B}^*_0\left(u; t+1, T\right)\right) - r + Y_t(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right)) \\
&\quad + \mathcal{A}^*\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right), \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right); t, t+1\right) \\
&\quad + \mathcal{B}^*_0\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right), \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right); t, t+1\right) y_t \\
&\quad + \boldsymbol{\mathcal{B}}^*\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right), \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right); t, t+1\right)^\top \boldsymbol{l}^*_t \\
&= \mathcal{A}^*\left(u; t+1, T\right) - r + uYt + \mathcal{A}^*\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right), \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right); t, t+1\right) \\
&\quad + \mathcal{B}^*_0\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right), \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right); t, t+1\right) y_t \\
&\quad + \boldsymbol{\mathcal{B}}^*\left(1 + u + \mathcal{B}^*_0\left(u; t+1, T\right), \boldsymbol{\mathcal{B}}^*\left(u; t+1, T\right); t, t+1\right)^\top \boldsymbol{l}^*_t.
\end{aligned}
$$

In summary, the risk-minimizing hedging ratio can be expressed in the following way:

$$
\theta^S_{t+1} = \frac{e^{-r(T-t)} S_t}{e^{-2r}\, \exp\left(\mathcal{C}^{\mathbb{Q}}_{Y_{t+1}, \boldsymbol{l}^*_{t+1}}\left(2, \mathbf{0} \mid \mathcal{F}_t\right)\right) - S_t^2} \left(\mathbb{E}^{\widetilde{\mathbb{Q}}_t}[H \mid \mathcal{F}_t] - \mathbb{E}^{\mathbb{Q}}\left[H \mid \mathcal{F}_t\right]\right).
$$

D.6 Proof of Affine Property of HNSV and HNSV Risk-Neutral Dynamics

For any $u, v \in \mathbb{R}$, the $\mathcal{F}_{t-1}$-conditional CGF of y_t and h_t is computed as

$$
\begin{aligned}
\mathcal{C}^{\mathbb{P}}_{y_t, h_t}(u, v \mid \mathcal{F}_{t-1}) &= \log\left(\mathbb{E}^{\mathbb{P}}[\exp\left(uy_t + vh_t\right) \mid \mathcal{F}_{t-1}]\right) \\
&= \log\left(\mathbb{E}^{\mathbb{P}}[\exp\left(u\sqrt{h_{t-1}}\epsilon_t + v\alpha\left(\psi_t^2 - 2\gamma\sqrt{h_{t-1}}\psi_t\right)\right) \mid \mathcal{F}_{t-1}]\right) \\
&\qquad + ur + v\left(\sigma^2(1 - \phi) - \alpha\right) + \left(u\lambda + v\phi\right) h_{t-1} \qquad \text{(D.6.1)} \\
&= ur + v\left(\sigma^2(1 - \phi) - \alpha\right) + \left(u\lambda + v\phi\right) h_{t-1}
\end{aligned}
$$

$$+\mathcal{C}^{\mathbb{P}}_{\epsilon_t,\psi_t,\psi_t^2}\left(u\sqrt{h_{t-1}},-2v\alpha\gamma\sqrt{h_{t-1}},v\alpha \mid \mathcal{F}_{t-1}\right).$$

Let U_t be a standard Gaussian random variable independent of ϵ_t conditionally on $\mathcal{F}_{t-1}$ that satisfies the following representation:

$$\psi_t = \rho\epsilon_t + \sqrt{1-\rho^2}U_t.$$

Now the joint CGF of $X_1, X_1^2, X_2, X_2^2, X_1X_2$ where X_1 and X_2 are two independent standard Gaussian random variables is given by

$$\mathcal{C}(u_1,u_2,u_3,u_4,u_5) = \frac{u_1^2(1-2u_4)+u_3^2(1-2u_2)+2u_1u_3u_5}{2\left[(1-2u_2)(1-2u_4)-u_5^2\right]} - \frac{1}{2}\log\left[(1-2u_2)(1-2u_4)-u_5^2\right].$$

Thus we can reformulate Equation (D.6.1) the following way:

$$\begin{aligned}&\mathcal{C}^{\mathbb{P}}_{y_t,h_t}(u,v \mid \mathcal{F}_{t-1})\\ &= \mathcal{C}^{\mathbb{P}}_{\epsilon_t,\epsilon_t^2,U_t,U_t^2,\epsilon_tU_t}\left(\eta_1,\eta_2,\eta_3,\eta_4,\eta_5 \mid \mathcal{F}_{t-1}\right) + ur + v\left(\sigma^2(1-\phi)-\alpha\right) + \left(u\lambda+v\phi\right)h_{t-1}.\end{aligned}$$

where $\eta_1, ..., \eta_5$ are given by

$$\begin{aligned}&\eta_1 = (u-2v\rho\alpha\gamma)\sqrt{h_{t-1}}, \qquad \eta_2 = v\rho^2\alpha, \qquad \eta_3 = -2v\sqrt{1-\rho^2}\alpha\gamma\sqrt{h_{t-1}},\\ &\eta_4 = v\left(1-\rho^2\right)\alpha, \qquad \eta_5 = 2v\rho\sqrt{1-\rho^2}\alpha.\end{aligned}$$

With some simple algebraic manipulations and by adding and subtracting $\frac{u^2\rho^2}{2(1-2v\alpha)}$, we obtain

$$\begin{aligned}&\mathcal{C}^{\mathbb{P}}_{\epsilon_t,\epsilon_t^2,U_t,U_t^2,\epsilon_tU_t}\left(\eta_1,\eta_2,\eta_3,\eta_4,\eta_5 \mid \mathcal{F}_{t-1}\right)\\ &= \left[\frac{u^2\left(1-\rho^2\right)}{2} + \frac{\left(u\rho-2v\alpha\gamma\right)^2}{2\left(1-2v\alpha\right)}\right]h_{t-1} - \frac{1}{2}\log\left(1-2v\alpha\right).\end{aligned}$$

Finally, after some reordering, we obtain the affine representation of the joint-conditional CGF of y_t and h_t given $\mathcal{F}_{t-1}$ under $\mathbb{P}$:

$$\mathcal{C}^{\mathbb{P}}_{y_t,h_t}(u,v \mid \mathcal{F}_{t-1}) = \mathcal{A}\left(u,v;t-1,t\right) + \mathcal{B}\left(u,v;t-1,t\right)h_{t-1}, \tag{D.6.2}$$

where the coefficients $\mathcal{A}\left(u,v;t-1,t\right)$ and $\mathcal{B}\left(u,v;t-1,t\right)$ are given by

$$\begin{aligned}\mathcal{A}\left(u,v;t-1,t\right) &= ur + v\omega - \frac{1}{2}\log\left(1-2v\alpha\right),\\ \mathcal{B}\left(u,v;t-1,t\right) &= u\lambda + v\phi + \frac{u^2\left(1-\rho^2\right)}{2} + \frac{\left(u\rho-2v\alpha\gamma\right)^2}{2\left(1-2v\alpha\right)}.\end{aligned}$$

Now, let $\mathbb{Q}$ be the risk-neutral measure defined in Assumption 4 then using Equations (2.3.8) to (2.3.10), we obtain the one-step joint-conditional CGF of y_t and h_t given $\mathcal{F}_{t-1}$ under the risk-neutral measure $\mathbb{Q}$:

$$\mathcal{C}^{\mathbb{Q}}_{y_t,h_t}(u,v|\mathcal{F}_{t-1}) = \mathcal{A}^*(u,v;t-1,t) + \mathcal{B}_0^*(u,v;t-1,t)y_{t-1} + \mathcal{B}^*(u,v;t-1,t)h_{t-1}^*, \tag{D.6.3}$$

where we have the following relationship between risk-neutral and physical conditional variance:

$$\begin{aligned}\delta &= \frac{1-2\theta_l\alpha\left(1-\rho^2\right)}{1-2\theta_l\alpha},\\ h_t^* &= \delta h_t,\end{aligned}$$

and where the single-step coefficients $\mathcal{A}^*\left(u,v;t-1,t\right)$, $\mathcal{B}_0^*(u,v;t-1,t)$ and $\mathcal{B}^*\left(u,v;t-1,t\right)$ under risk-neutral measure $\mathbb{Q}$ are given by

$$\begin{aligned}\mathcal{A}^*(u,v;t-1,t) &= ur + v\omega\delta - \frac{1}{2}\log(1-2v\alpha\tau),\\ \mathcal{B}_0^*(u,v;t-1,t) &= 0,\\ \mathcal{B}^*(u,v;t-1,t) &= u\frac{\lambda}{\delta} + v\phi + \frac{(u^2+2u\theta_y)(1-\rho^2)\left(1-2\theta_l\alpha\right)}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)}\\ &\quad + \frac{\left(u\rho-2v\alpha\gamma\delta+\theta_y\rho-2\theta_l\alpha\gamma\right)^2}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)\left(1-2v\alpha\tau\right)} - \frac{\left(\theta_y\rho-2\theta_l\alpha\gamma\right)^2}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)},\end{aligned}$$

where $\tau = \frac{1-2\theta_l\alpha\left(1-\rho^2\right)}{\left(1-2\theta_l\alpha\right)^2}$. Now, by imposing the martingale constraint according to which the discounted asset price is a martingale after the change of measure defined in Assumption 4, we can determine the relationship between θ_l and θ_y:

$$\begin{aligned}&\log\left(\mathbb{E}^{\mathbb{Q}}\left[e^{y_t}\mid\mathcal{F}_{t-1}\right]\right)\\ &\quad = \mathcal{C}^{\mathbb{Q}}_{y_t,h_t}(1,0|\mathcal{F}_{t-1}) = \mathcal{A}^*(1,0;t-1,t) + \mathcal{B}^*(1,0;t-1,t)h^*_{t-1}\\ &\quad = r + \frac{\lambda}{\delta} + \frac{\left(1+2\theta_y\right)\left(1-\rho^2\right)\left(1-2\theta_l\alpha\right)}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)} + \frac{\left(\rho+\theta_y\rho-2\theta_l\alpha\gamma\right)^2}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)} - \frac{\left(\theta_y\rho-2\theta_l\alpha\gamma\right)^2}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)}\\ &\quad = r + \frac{\lambda}{\delta} + \frac{\left(1+2\theta_y\right)\left(1-\rho^2-2\theta_l\alpha\left(1-\rho^2\right)\right)}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)} + \frac{\rho^2+2\theta_y\rho^2-4\theta_l\rho\alpha\gamma}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)}\\ &\quad = r + \frac{\lambda}{\delta} + \frac{1}{2} + \theta_y - \frac{\rho^2\left(1+2\theta_y\right)}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)} + \frac{\rho^2+2\theta_y\rho^2-4\theta_l\rho\alpha\gamma}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)}\\ &\quad = r + \frac{\lambda}{\delta} + \frac{1}{2} + \theta_y - \frac{2\theta_l\rho\alpha\gamma}{1-2\theta_l\alpha\left(1-\rho^2\right)}.\end{aligned}$$

and, since the martingale constraint imposes that $\log\left(\mathbb{E}^{\mathbb{Q}}\left[e^{y_t}\mid\mathcal{F}_{t-1}\right]\right) = r$, we obtain

$$\theta_y = -\frac{1}{2} + \frac{2\theta_l\rho\alpha\gamma - \lambda(1-2\theta_l\alpha)}{1-2\theta_l\alpha(1-\rho^2)}, \tag{D.6.4}$$

which, after some simple rearranging, we rewrite as:

$$\frac{\theta_y\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)-2\theta_l\rho\alpha\gamma}{1-2\theta_l\alpha\left(1-\rho^2\right)} = -\frac{1}{2} - \lambda^*, \tag{D.6.5}$$

where $\lambda^* = \frac{1-2\theta_l\alpha}{1-2\theta_l\alpha(1-\rho^2)}\lambda$. Now lets revisit $\mathcal{B}^*(u,v;t-1,t)$ in order to obtain a similar representation under the risk-neutral measure $\mathbb{Q}$ as the one under the physical measure $\mathbb{P}$:

$$\begin{aligned}
&\mathcal{B}^*(u,v;t-1,t)\\
&\quad= u\lambda^* + v\phi + \frac{u^2(1-2\theta_l\alpha(1-\rho^2)) - u^2\rho^2}{2(1-2\theta_l\alpha(1-\rho^2))} + \frac{2u\theta_y(1-2\theta_l\alpha(1-\rho^2)) - 2u\theta_y\rho^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\qquad - \frac{2v\alpha\tau(2u\theta_y(1-2\theta_l\alpha(1-\rho^2)) - 2u\theta_y\rho^2)}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)} + \frac{(u\rho - 2v\alpha\gamma\tau + 4v\theta_l\alpha^2\gamma\tau)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\qquad + \frac{2u\theta_y\rho^2 - 4v\theta_y\rho\alpha\gamma\tau + 8v\theta_y\theta_l\rho\alpha^2\gamma\tau - 4u\theta_l\rho\alpha\gamma + 8v\theta_l\alpha^2\gamma^2\tau - 16v\theta_l^2\alpha^3\gamma^2\tau}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\qquad + \frac{2v\alpha\tau(\theta_y\rho - 2\theta_l\alpha\gamma)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}.
\end{aligned}$$

We add and subtract $\frac{8uv\theta_l\rho\alpha^2\gamma\tau}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}$ and use Equation (D.6.5) to obtain

$$\begin{aligned}
&\mathcal{B}^*(u,v;t-1,t)\\
&\quad= u\lambda^* + v\phi + \frac{u^2}{2} - \frac{u^2\rho^2}{2(1-2\theta_l\alpha(1-\rho^2))} + \frac{u(\theta_y(1-2\theta_l\alpha(1-\rho^2)) - 2\theta_l\rho\alpha\gamma)}{(1-2\theta_l\alpha(1-\rho^2))}\\
&\qquad + \frac{(u\rho - 2v\alpha\gamma\tau + 4v\theta_l\alpha^2\gamma\tau)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)} + \frac{2v\alpha\tau(\theta_y\rho - 2\theta_l\alpha\gamma)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\qquad + \frac{4uv\theta_y\rho^2\alpha\tau - 4v\theta_y\rho\alpha\gamma\tau + 8v\theta_y\theta_l\rho\alpha^2\gamma\tau - 8uv\theta_l\rho\alpha^2\gamma\tau + 8v\theta_l\alpha^2\gamma^2\tau - 16v\theta_l^2\alpha^3\gamma^2\tau}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\quad= u\lambda^* + v\phi + \frac{u^2}{2} - \frac{u^2\rho^2}{2(1-2\theta_l\alpha(1-\rho^2))} - u\left(\frac{1}{2} + \lambda^*\right)\\
&\qquad + \frac{(u\rho - 2v\alpha\gamma\tau + 4v\theta_l\alpha^2\gamma\tau)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\qquad + 2\frac{(\theta_y\rho - 2\theta_l\alpha\gamma)(2uv\rho\alpha\tau - 2v\alpha\gamma\tau + 4v\theta_l\alpha^2\gamma\tau)}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)} + \frac{2v\alpha\tau(\theta_y\rho - 2\theta_l\alpha\gamma)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\quad= -\frac{u}{2} + v\phi + \frac{u^2}{2} - \frac{u^2\rho^2}{2(1-2\theta_l\alpha(1-\rho^2))} + \frac{(u\rho - 2v\alpha\tau(\gamma(1-2\theta_l\alpha)))^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}\\
&\qquad + 4v\alpha\tau\frac{(\theta_y\rho - 2\theta_l\alpha\gamma)(u\rho - \gamma(1-2\theta_l\alpha))}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)} + \frac{2v\alpha\tau(\theta_y\rho - 2\theta_l\alpha\gamma)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}.
\end{aligned}$$

We add and subtract $\frac{4v^2\alpha^{*2}(\theta_y\rho-2\theta_l\alpha\gamma)^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}$ and $2v\alpha\tau\frac{4v\alpha\tau(\theta_y\rho-2\theta_l\alpha\gamma)\gamma(1-2\theta_l\alpha)}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}$ and use the relation $\phi = \beta - \alpha\gamma^2$ to obtain

$$\begin{aligned}
&\mathcal{B}^*(u,v;t-1,t)\\
&\quad= -\frac{u}{2} + v\beta + \frac{u^2}{2} - \frac{u^2\rho^2}{2(1-2\theta_l\alpha(1-\rho^2))} + \frac{(u\rho - 2v\alpha\tau(2\theta_l\alpha\gamma - \theta_y\rho + \gamma(1-2\theta_l\alpha)))^2}{2(1-2\theta_l\alpha(1-\rho^2))(1-2v\alpha\tau)}
\end{aligned}$$

$$
\begin{aligned}
&+ v\alpha\tau \left(\frac{(\theta_y\rho - 2\theta_l\alpha\gamma)^2}{1-2\theta_l\alpha\left(1-\rho^2\right)} - 2\frac{(\theta_y\rho - 2\theta_l\alpha\gamma)\,\gamma\,(1-2\theta_l\alpha)}{1-2\theta_l\alpha\left(1-\rho^2\right)} + \gamma^2 \frac{(1-2\theta_l\alpha)^2}{1-2\theta_l\alpha\left(1-\rho^2\right)} \right) \\
=& -\frac{u}{2} + v\beta + + \frac{u^2}{2} - \frac{u^2\rho^2}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)} + \frac{\left(u\rho - 2v\alpha\tau\left(\gamma-\theta_y\rho\right)\right)^2}{2\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)\left(1-2v\alpha\tau\right)} \\
&+ v\alpha\tau \left(\frac{\theta_y\rho - \gamma}{\sqrt{1-2\theta_l\alpha\left(1-\rho^2\right)}} \right)^2 .
\end{aligned}
$$

Now if we set

$$
\omega^* = \omega\delta, \qquad \alpha^* = \alpha\tau, \qquad \pi = \frac{1}{1-2\theta_l\alpha\left(1-\rho^2\right)}, \qquad q^* = \left(\gamma-\theta_y\rho\right)\sqrt{\pi}, \qquad \rho^* = \rho\sqrt{\pi},
$$

then we can simplify the risk-neutral one-step coefficients to

$$
\begin{aligned}
\mathcal{A}^*(u,v;t-1,t) &= ur + v\omega^* - \frac{1}{2}\log(1-2v\alpha^*), \\
\mathcal{B}^*(u,v;t-1,t) &= -\frac{u}{2} + \left(\beta + \alpha^* q^{*2}\right)v + \frac{u^2\left(1-\rho^{*2}\right)}{2} + \frac{\left(u\rho^* - 2v\alpha^* q^*\right)^2}{2\left(1-2v\alpha^*\right)}.
\end{aligned}
$$

which completes the derivation.

D.7 Proof of Joint Dynamics of HNSV Noise Process

We will now derive the joint dynamics of the joint-noise process (ϵ_t, ψ_t) under the risk-neutral measure $\mathbb{Q}$. Let X_t, Y_t be two stochastic processes following a $\mathcal{F}_{t-1}$-conditional bivariate standard Gaussian distribution with means μ_1, μ_2, correlation coefficient ρ and covariance matrix

$$
\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},
$$

then their $\mathcal{F}_{t-1}$-conditional joint CGF under physical measure $\mathbb{P}$ is given by:

$$
\mathcal{C}^{\mathbb{P}}_{X_t,Y_t}(w,v \mid \mathcal{F}_{t-1}) = w\mu_1 + v\mu_2 + \frac{1}{2}\left(w^2\sigma_1^2 + 2wv\rho\sigma_1\sigma_2 + v^2\sigma_2^2\right). \tag{D.7.1}
$$

For any $u,v \in \mathbb{R}$, the $\mathcal{F}_{t-1}$-conditional CGF of ϵ_t and ψ_t is computed as

$$
\begin{aligned}
&\mathcal{C}^{\mathbb{Q}}_{\epsilon_t,\psi_t}(u,v \mid \mathcal{F}_{t-1}) \\
&\quad= \log\left(\mathbb{E}^{\mathbb{Q}}[\exp(u\epsilon_t + v\psi_t) \mid \mathcal{F}_{t-1}]\right) = \log\left(\frac{\mathbb{E}^{\mathbb{P}}\left[\exp(u\epsilon_t + v\psi_t)\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_{t-1}\right]}{\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_{t-1}\right]}\right) \\
&\quad= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(u\epsilon_t + v\psi_t + \theta_y y_t + \theta_l h_t - \mathcal{C}^{\mathbb{P}}_{y_t,h_t}(\theta_y,\theta_l \mid \mathcal{F}_{t-1})\right) \mid \mathcal{F}_{t-1}\right]\right) \\
&\quad= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(u\epsilon_t + v\psi_t + \theta_y y_t + \theta_l h_t\right) \mid \mathcal{F}_{t-1}\right]\right) - \mathcal{C}^{\mathbb{P}}_{y_t,h_t}(\theta_y,\theta_l \mid \mathcal{F}_{t-1}) \\
&\quad= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left((u+\theta_y\sqrt{h_{t-1}})\epsilon_t + (v - 2\alpha\gamma\theta_l\sqrt{h_{t-1}})\psi_t + \alpha\theta_l\psi_t^2\right) \mid \mathcal{F}_{t-1}\right]\right) \qquad \text{(D.7.2)} \\
&\qquad + \theta_y(r+\lambda h_{t-1}) + \theta_l((1-\phi)(\sigma^2 - h_{t-1}) - \alpha) - \mathcal{C}^{\mathbb{P}}_{y_t,h_t}(\theta_y,\theta_l \mid \mathcal{F}_{t-1})
\end{aligned}
$$

$$
\begin{aligned}
&= \mathcal{C}^{\mathbb{P}}_{\epsilon_t,\psi_t,\psi_t^2}\left(u+\theta_y\sqrt{h_{t-1}}, v-2\alpha\gamma\theta_l\sqrt{h_{t-1}},\alpha\theta_l \mid \mathcal{F}_{t-1}\right) + \theta_y(r+\lambda h_{t-1}) \\
&\qquad + \theta_l((1-\phi)(\sigma^2-h_{t-1})-\alpha) - \mathcal{C}^{\mathbb{P}}_{y_t,h_t}(\theta_y,\theta_l \mid \mathcal{F}_{t-1}).
\end{aligned}
$$

Let U_t be a standard Gaussian random variable independent of ϵ_t conditionally on $\mathcal{F}_{t-1}$ that satisfies the following representation:

$$
\psi_t = \rho\epsilon_t + \sqrt{1-\rho^2}U_t.
$$

Now the joint CGF of $X_1, X_1^2, X_2, X_2^2, X_1X_2$ where X_1 and X_2 are two independent standard Gaussian random variables is given by

$$
\mathcal{C}(u_1,u_2,u_3,u_4,u_5) = \frac{u_1^2(1-2u_4)+u_3^2(1-2u_2)+2u_1u_3u_5}{2\left[(1-2u_2)(1-2u_4)-u_5^2\right]} - \frac{1}{2}\log\left[(1-2u_2)(1-2u_4)-u_5^2\right].
$$

Thus we can reformulate (D.7.2) the following way:

$$
\begin{aligned}
&\mathcal{C}^{\mathbb{Q}}_{\epsilon_t,\psi_t}(u,v \mid \mathcal{F}_{t-1}) \\
&= \mathcal{C}^{\mathbb{P}}_{\epsilon_t,\epsilon_t^2,U_t,U_t^2,\epsilon_tU_t}(\eta_1,\eta_2,\eta_3,\eta_4,\eta_5 \mid \mathcal{F}_{t-1}) + \theta_y(r+\lambda h_{t-1}) + \theta_l((1-\phi)(\sigma^2-h_{t-1})-\alpha) \\
&\qquad - \mathcal{C}^{\mathbb{P}}_{y_t,h_t}(\theta_y,\theta_l \mid \mathcal{F}_{t-1}),
\end{aligned}
$$

where $\eta_1, \ldots, \eta_5$ are given by

$$
\eta_1 = (\theta_y - 2\theta_l\rho\alpha\gamma)\sqrt{h_{t-1}} + u + v\rho, \qquad \eta_2 = \theta_l\rho^2\alpha, \qquad \eta_3 = v\sqrt{1-\rho^2} - 2\theta_l\sqrt{1-\rho^2}\alpha\gamma\sqrt{h_{t-1}},
$$
$$
\eta_4 = \theta_l\left(1-\rho^2\right)\alpha, \qquad \eta_5 = 2\theta_l\rho\sqrt{1-\rho^2}\alpha.
$$

Now we can also rewrite Equation (D.6.4) as

$$
\frac{\theta_y\left(1-2\theta_l\alpha\left(1-\rho^2\right)\right)-2\theta_l\rho\alpha\gamma}{1-2\theta_l\alpha} = -\frac{1-2\theta_l\alpha(1-\rho^2)}{2\left(1-2\theta_l\alpha\right)} - \lambda. \tag{D.7.3}
$$

With some simple algebraic manipulations and using Equation (D.7.3), we obtain

$$
\begin{aligned}
&\mathcal{C}^{\mathbb{P}}_{\epsilon_t,\epsilon_t^2,U_t,U_t^2,\epsilon_tU_t}(\eta_1,\eta_2,\eta_3,\eta_4,\eta_5 \mid \mathcal{F}_{t-1}) \\
&\qquad = u\left[-\frac{1-2\theta_l\alpha(1-\rho^2)}{2\left(1-2\theta_l\alpha\right)} - \lambda\right]\sqrt{h_{t-1}} + v\frac{\theta_y\rho-2\theta_l\alpha\gamma}{1-2\theta_l\alpha}\sqrt{h_{t-1}} \\
&\qquad\quad + \frac{1}{2}\left[u^2\left[\frac{1-2\theta_l\alpha\left(1-\rho^2\right)}{1-2\theta_l\alpha}\right] + 2\frac{uv\rho}{1-2\theta_l\alpha} + \frac{v^2}{1-2\theta_l\alpha}\right] \\
&\qquad\quad + \left[\frac{\theta_y^2(1-\rho^2)}{2} + \frac{\left(\theta_y\rho-2\theta_l\alpha\gamma\right)^2}{2\left(1-2\theta_l\alpha\right)}\right]h_{t-1} - \frac{1}{2}\log\left(1-2\alpha\theta_l\right),
\end{aligned}
$$

and finally

$$
\begin{aligned}
\mathcal{C}^{\mathbb{Q}}_{\epsilon_t,\psi_t}(u,v \mid \mathcal{F}_{t-1}) &= u\left[-\frac{1-2\theta_l\alpha(1-\rho^2)}{2\left(1-2\theta_l\alpha\right)} - \lambda\right]\sqrt{h_{t-1}} + v\frac{\theta_y\rho-2\theta_l\alpha\gamma}{1-2\theta_l\alpha}\sqrt{h_{t-1}} \\
&\quad + \frac{1}{2}\left[u^2\left[\frac{1-2\theta_l\alpha\left(1-\rho^2\right)}{1-2\theta_l\alpha}\right] + 2\frac{uv\rho}{1-2\theta_l\alpha} + \frac{v^2}{1-2\theta_l\alpha}\right].
\end{aligned}
$$

Using Equation (D.7.1), we establish that the joint noise process (ϵ_t, ψ_t) under $\mathbb{Q}$ follows a bivariate Gaussian distribution with correlation coefficient $\rho^* = \frac{\rho}{\sqrt{1-2\alpha\theta_l(1-\rho^2)}}$ where

$$\epsilon_t = \mu_1^* + \sigma_1^* \epsilon_t^*, \qquad \psi_t = \mu_2^* + \sigma_2^* \psi_t^*,$$

where

$$\mu_1^* = \left[-\frac{1-2\theta_l\alpha(1-\rho^2)}{2\,(1-2\theta_l\alpha)} - \lambda \right] \sqrt{h_{t-1}}, \qquad \sigma_1^{*2} = \frac{1-2\theta_l\alpha\,(1-\rho^2)}{1-2\theta_l\alpha},$$
$$\mu_2^* = \frac{\theta_y\rho - 2\theta_l\alpha\gamma}{1-2\theta_l\alpha}\sqrt{h_{t-1}}, \qquad \sigma_2^{*2} = \frac{1}{1-2\theta_l\alpha},$$

and

$$\epsilon_t^* \sim \mathcal{N}(0,1), \qquad \psi_t^* \sim \mathcal{N}(0,1).$$

Indeed, Equation (D.7.1) gives us the relationship betweenn the physical correlation coefficient and the risk-neutral correlation coefficient:

$$2uv\sigma_{\epsilon_t}^{\mathbb{Q}}\rho^*\sigma_{\psi_t}^{\mathbb{Q}} = 2\frac{uv\rho}{1-2\theta_l\alpha} \iff \rho^* = \frac{\rho}{\sqrt{1-2\alpha\theta_l\,(1-\rho^2)}}.$$

We plug these dynamics into Equation (3.2.3) and Equation (3.2.4), multiply h_t by δ and obtain the risk-neutral HNSV dynamics:

$$y_t = r - \frac{h_{t-1}^*}{2} + \sqrt{h_{t-1}^*}\epsilon_t^*,$$
$$h_t^* = \omega^* + \alpha^* \left(\psi_t^* - q^* \sqrt{h_{t-1}} \right)^2 + \beta h_{t-1}^*.$$

where

$$\omega^* = \delta\omega, \qquad \alpha^* = \frac{\delta\alpha}{1-2\alpha\theta_l}, \qquad q^* = \frac{\gamma - \rho\theta_y}{\sqrt{1-2\alpha\theta_l\,(1-\rho^2)}}.$$

D.8 Proof of the Dynamics of GARCH(C) Noise Process

In order to have a direct relationship between the physical parameters and risk-neutral parameters under the GARCH(C) model parametrization, we first reformulate the GARCH(C) dynamics presented in Equations (3.3.1) through (3.3.4) the following way:

$$\begin{aligned} y_t &= r + \lambda h_{t-1} + \sqrt{h_{t-1}} z_t \\ h_t &= s_t + q_t \\ s_t &= \omega_s + \rho_s s_{t-1} + \xi_t^{(s)} \\ q_t &= \omega_q + \rho_q q_{t-1} + \xi_t^{(q)} \end{aligned} \tag{D.8.1}$$

where

$$\omega_s = -\alpha_s, \qquad \omega_q = \sigma^2(1-\rho_q) - \alpha_q, \qquad \beta_s = \rho_s - \alpha_s\gamma_s^2, \qquad \beta_q = \rho_q - \alpha_q\gamma_q^2,$$

and where we redefine the driving processes for the two volatility components, $\xi_t^{(s)}$ and $\xi_t^{(q)}$, as the following equations:

$$\xi_t^{(s)} = a_s\left(z_t - \gamma_s\sqrt{h_{t-1}}\right)^2 + (\beta_s - \rho_s)\,h_{t-1}, \qquad \xi_t^{(q)} = a_q\left(z_t - \gamma_q\sqrt{h_{t-1}}\right)^2 + (\beta_q - \rho_q)\,h_{t-1}.$$

Let $\boldsymbol{\theta}_l = \begin{bmatrix}\theta_s & \theta_q\end{bmatrix}^\top$. Using the notation from Equation (3.3.1) through (3.3.4), for any $u \in \mathbb{R}$, the $\mathcal{F}_{t-1}$-conditional CGF of z_t is computed as

$$\begin{aligned}
&\mathcal{C}^{\mathbb{Q}}_{z_t}(u \mid \mathcal{F}_{t-1}) \\
&= \log\left(\mathbb{E}^{\mathbb{Q}}[\exp(uz_t) \mid \mathcal{F}_{t-1}]\right) = \log\left(\frac{\mathbb{E}^{\mathbb{P}}\left[\exp(uz_t)\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_{t-1}\right]}{\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_{t-1}\right]}\right) \\
&= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(uz_t + \theta_y y_t + \boldsymbol{\theta}_l^\top \boldsymbol{l}_t - \mathcal{C}^{\mathbb{P}}_{y_t,\boldsymbol{l}_t}(\theta_y, \boldsymbol{\theta}_l \mid \mathcal{F}_{t-1})\right) \mid \mathcal{F}_{t-1}\right]\right) \\
&= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(uz_t + \theta_y y_t + \theta_s s_t + \theta_q q_t\right) \mid \mathcal{F}_{t-1}\right]\right) - \mathcal{C}^{\mathbb{P}}_{y_t,\boldsymbol{l}_t}(\theta_y, \boldsymbol{\theta}_l \mid \mathcal{F}_{t-1}) \\
&= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(uz_t + \theta_y\sqrt{h_{t-1}}z_t + \theta_s\xi_t^{(s)} + \theta_q\xi_t^{(q)}\right) \mid \mathcal{F}_{t-1}\right]\right) + \theta_y\left(r + \lambda h_{t-1}\right) \qquad \text{(D.8.2)} \\
&\quad + \theta_s\rho_s s_{t-1} + \theta_q\left(\sigma^2 + \rho_q(q_{t-1} - \sigma^2)\right) - \mathcal{C}^{\mathbb{P}}_{y_t,\boldsymbol{l}_t}(\theta_y, \boldsymbol{\theta}_l \mid \mathcal{F}_{t-1}) \\
&= \log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(\left(u + \theta_y\sqrt{h_{t-1}} - 2\theta_s a_s\gamma_s\sqrt{h_{t-1}} - 2\theta_q a_q\gamma_q\sqrt{h_{t-1}}\right)z_t + \left(\theta_s a_s + \theta_q a_q\right)z_t^2\right) \mid \mathcal{F}_{t-1}\right]\right) \\
&\quad + \theta_y\left(r + \lambda h_{t-1}\right) + \theta_s\left(\rho_s s_{t-1} - a_s\right) + \theta_q\left(\sigma^2 + \rho_q(q_{t-1} - \sigma^2) - a_q\right) - \mathcal{C}^{\mathbb{P}}_{y_t,\boldsymbol{l}_t}(\theta_y, \boldsymbol{\theta}_l \mid \mathcal{F}_{t-1}) \\
&= \mathcal{C}^{\mathbb{P}}_{z_t,z_t^2}\left(u + \left(\theta_y - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_3\right)\sqrt{h_{t-1}}, \boldsymbol{\theta}_l^\top \boldsymbol{K}_2 \mid \mathcal{F}_{t-1}\right) \\
&\quad + \theta_y\left(r + \lambda h_{t-1}\right) + \boldsymbol{\theta}_l^\top \boldsymbol{K}_4 s_{t-1} + \boldsymbol{\theta}_l^\top \boldsymbol{K}_5 q_{t-1} + \theta_q\sigma^2(1-\rho_q) - \boldsymbol{\theta}_l^\top \boldsymbol{K}_2 - \mathcal{C}^{\mathbb{P}}_{y_t,\boldsymbol{l}_t}(\theta_y, \boldsymbol{\theta}_l \mid \mathcal{F}_{t-1}) \\
&= \mathcal{C}^{\mathbb{P}}_{z_t,z_t^2}\left(u + \left(\theta_y - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_3\right)\sqrt{h_{t-1}}, \boldsymbol{\theta}_l^\top \boldsymbol{K}_2 \mid \mathcal{F}_{t-1}\right) + \frac{1}{2}\log(1 - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2) - \frac{(\theta_y - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_3)^2}{2(1 - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2)}h_{t-1}.
\end{aligned}$$

Now the joint CGF of $X_1, X_1^2, X_2, X_2^2, X_1X_2$ where X_1 and X_2 are two independent standard Gaussian random variables is given by

$$\mathcal{C}(u_1,u_2,u_3,u_4,u_5) = \frac{u_1^2(1-2u_4) + u_3^2(1-2u_2) + 2u_1u_3u_5}{2\left[(1-2u_2)(1-2u_4) - u_5^2\right]} - \frac{1}{2}\log\left[(1-2u_2)(1-2u_4) - u_5^2\right]$$

and thus the joint CGF of X_1, X_1^2 is

$$\mathcal{C}(u_1,u_2) = \frac{u_1^2}{2(1-2u_2)} - \frac{1}{2}\log(1-2u_2)$$

Thus we obtain

$$\mathcal{C}^{\mathbb{Q}}_{z_t}(u \mid \mathcal{F}_{t-1}) = u\frac{\theta_y - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_3}{1 - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2}\sqrt{h_{t-1}} + u^2\frac{1}{2(1 - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2)}$$

Now we can also rewrite (3.3.12) as

$$\frac{\theta_y - 2\boldsymbol{\theta_l}^\top \boldsymbol{K}_3}{1 - 2\boldsymbol{\theta_l}^\top \boldsymbol{K}_2} = -\frac{1}{2\left(1 - 2\boldsymbol{\theta_l}^\top \boldsymbol{K}_2\right)} - \lambda \tag{D.8.3}$$

so we obtain

$$\mathcal{C}_{z_t}^{\mathbb{Q}}(u \mid \mathcal{F}_{t-1}) = u\left(-\frac{1}{2\left(1 - 2\boldsymbol{\theta_l}^\top \boldsymbol{K}_2\right)}\sqrt{h_{t-1}} - \lambda\right) + \frac{1}{2}u^2 \frac{1}{1 - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2} \tag{D.8.4}$$

The expression derived for $\mathcal{C}_{z_t}^{\mathbb{Q}}(u \mid \mathcal{F}_{t-1})$ in Equation (D.8.4) implies that the noise process z_t follow a Gaussian distribution under $\mathbb{Q}$. Specifically, we can represent z_t as follows:

$$z_t = \mu_1^* + \sigma_1^* z_t^*$$

where

$$\mu^* = -\left(\frac{1}{2\left(1 - 2\boldsymbol{\theta_l}^\top \boldsymbol{K}_2\right)} + \lambda\right)\sqrt{h_{t-1}}, \qquad \sigma^{*2} = \frac{1}{1 - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2},$$

and z_t^* follows a standard Gaussian distribution with zero mean and unit variance. Letting $h_t^* = \delta h_t$, $s_t^* = \delta s_t$ and $q_t^* = \delta q_t$ where

$$\delta = \frac{1}{1 - 2\boldsymbol{\theta}_l^\top \boldsymbol{K}_2},$$

and plugging the above representation of z_t into Equations (D.8.1), we obtain the risk-neutral GARCH(C) dynamics:

$$\begin{aligned} y_t &= r - \frac{1}{2}h_{t-1}^* + \sqrt{h_{t-1}^*}z_t^* \\ h_t^* &= s_t^* + q_t^* \\ s_t^* &= \omega_s^* + \rho_s s_{t-1}^* + \xi_t^{*(s)} \\ q_t^* &= \omega_q^* + \rho_q q_{t-1}^* + \xi_t^{*(q)} \end{aligned} \tag{D.8.5}$$

where

$$\omega_s^* = \omega_s \delta, \qquad \omega_q^* = \omega_q \delta,$$

and where driving processes for the two risk-neutral volatility components, $\xi_t^{*(s)}$ and $\xi_t^{*(q)}$, are defined by the following equations:

$$\xi_t^{*(s)} = a_s^*\left(z_t^* - \gamma_s^*\sqrt{h_{t-1}^*}\right)^2 + \left(\beta_s - \rho_s\right)h_{t-1}^*, \qquad \xi_t^{*(q)} = a_q^*\left(z_t^* - \gamma_q^*\sqrt{h_{t-1}^*}\right)^2 + \left(\beta_q - \rho_q\right)h_{t-1}^*. \tag{D.8.6}$$

with

$$\alpha_s^* = \delta^2 \alpha_s, \qquad \alpha_q^* = \delta^2 \alpha_q, \qquad \gamma_s^* = \frac{1}{2} + \frac{(\lambda + \gamma_s)}{\delta}, \qquad \gamma_q^* = \frac{1}{2} + \frac{(\lambda + \gamma_q)}{\delta}.$$