MERTON’S PORTFOLIO OPTIMIZATION PROBLEM IN A BLACK AND
SCHOLES MARKET WITH NON-GAUSSIAN STOCHASTIC VOLATILITY
OF ORNSTEIN-UHLENBECK TYPE

FRED ESPEN BENTH

Department of Mathematics, University of Oslo, Norway and MaPhySto—Centre
for Mathematical Physics and Stochastics, University of Aarhus, Denmark

KENNETH HVISTENDAHL KARLSEN

Department of Mathematics, University of Bergen, Norway

KRISTIN REIKVAM

Department of Mathematics, University of Oslo, Norway

We study Merton’s classical portfolio optimization problem for an investor who can
trade in a risk-free bond and a stock. The goal of the investor is to allocate money
so that her expected utility from terminal wealth is maximized. The special feature
of the problem studied in this paper is the inclusion of stochastic volatility in the
dynamics of the risky asset. The model we use is driven by a superposition of non-
Gaussian Ornstein-Uhlenbeck processes and it was recently proposed and intensively
investigated for real market data by Barndorff-Nielsen and Shephard (2001). Using
the dynamic programming method, explicit trading strategies and expressions for the
value function via Feynman-Kac formulas are derived and verified for power utilities.
Some numerical examples are also presented.

KEY WORDS: portfolio optimization, stochastic volatility, verification theorem, Feynman-Kac
formula, non-Gaussian Ornstein-Uhlenbeck process

1. INTRODUCTION

Recently, Barndorff-Nielsen and Shephard (2001) suggested modeling the volatility in
asset price dynamics as a weighted sum of non-Gaussian Ornstein-Uhlenbeck processes.
This volatility model possesses many of the observed features of financial log-return data,
such as heavy tails and long-range dependency. Barndorff-Nielsen and Shephard intro-
duced subordinators (i.e., nondecreasing pure-jump Lévy process) as the noise driving
the Ornstein-Uhlenbeck processes. That is, the volatility level is allowed to have sudden
shifts in the upward direction, while decreasing exponentially between such shifts. Their empirical investigations on exchange rates demonstrate that such volatility models fit the empirical autocorrelation and the leptokurtic behavior of log-return data remarkably well. In the present paper we investigate Merton’s classical portfolio optimization problem under this volatility model.

Merton (1969, 1971) explicitly solved the question of optimal portfolio allocation in a market with a risk-free bond and a stock as investment alternatives. The price of the risky asset (e.g., a stock) follows a geometric Brownian motion (or a geometric Brownian motion with Poisson jumps). The investor wants to maximize her terminal wealth under a power utility function. Using a stochastic volatility model as in Barndorff-Nielsen and Shephard (2001), we prove that it is still possible to explicitly solve Merton’s problem, however now via a Feynman-Kac representation. Not surprisingly, the investor should follow Merton’s original strategy of constant allocation (i.e., an allocation independent of current level of wealth), but she should rebalance the portfolio according to changes in the volatility.

Our approach to solve the stochastic optimization problem goes via the dynamic programming method and the associated Hamilton-Jacobi-Bellman (HJB) integro-differential equation. Using a verification theorem, we are able to identify the expected value derived from utility as the solution of a second-order integro-differential equation. The solution to this HJB equation can be written in terms of the power utility function and a Feynman-Kac representation. In addition, it is possible to explicitly write the optimal allocation strategy. All results are derived under exponential integrability assumptions on the Lévy measures coming from the volatility process. We suppose that there are borrowing and short-selling constraints in the market. In a companion paper (Benth, Karlsen, and Reikvam 2003) we study the portfolio problem with a general utility function and an extension of the stock price dynamics. In that paper we use viscosity solution methods to relate the value function to the HJB equation.

A major drawback with the dynamic programming approach is that the allocation strategies must depend on the volatility (explicitly or through information generated by it). Of course, volatility is not directly observable in the market, unlike the stock price, and it is therefore in practice impossible to follow portfolio rules where one must take the level of volatility explicitly into account. An alternative to our approach is stochastic control under partial observation. This would solve the problem of dependence of the volatility on the controls. Pham and Quenez (2001) used the approach of partial observation to solve a portfolio problem similar to ours, but with a stochastic volatility process driven by a Brownian motion correlated to the dynamics of the risky asset. Choosing their point of view in our context would lead to nonlinear filtering problems for jump processes. For more information on stochastic control under partial observation, see the monograph by Bensoussan (1992).

Despite the (possible) practical shortcomings, we still believe our analysis contributes to further insight into portfolio optimization problem in markets with stochastic volatility. Investors do have a feeling for the current level of volatility, and it therefore may be desirable to include this (time-varying) information in their portfolio management strategies. Our results confirm that it is optimal to invest according to Merton’s strategy, but one must update the investment according to the level of volatility. In practice this is what an investor does. She rebalances the portfolio when level of volatility changes. Our results confirm the optimality in such a strategy.

Our results are also relevant for option pricing in incomplete markets. The asset price model suggested by Barndorff-Nielsen and Shephard (2001) leads to a market that is incomplete. Nicolato and Venardos (2001) derived an interval of arbitrage-free prices for a
plain vanilla European option. Alternatively, following Hodges and Neuberger (1989) and Davis, Panas, and Zariphopoulou (1993), one may use portfolio optimization techniques to determine the price of derivatives. In this context the natural choice of utility function is the exponential one, contrary to the power utility treated in the present paper. However, as will be shown in detail in the forthcoming paper by Benth and Karlsen (2003), the lines of argumentation and analysis for exponential utilities are more or less identical to the power utility case. Using a utility optimization approach to pricing in incomplete markets, one obtain prices dependent on the risk aversion of the investor. Utility optimization techniques for pricing derivatives in incomplete markets have recently drawn a lot of attention (see, e.g., Rouge and El Karoui 2000 and Delbaen et al. 2002).

Fouque, Papanicolaou, and Sircar (2000) studied Merton’s problem with stochastic volatility being a mean-reverting process. The volatility process is driven by a Brownian motion correlated to the risky asset. Using the same line of argumentation as we do, they derived solutions for both the optimal investment strategy and the investor’s value function. The authors investigated their solutions by means of asymptotic expansions around the (inverse of the) speed of mean-reversion, and thereby obtained results of practical interest. Since our volatility process will not be of a pure diffusion nature, we face the problem of solving an integro-differential equation to find the value function. This equation cannot be solved explicitly but, in general, only by means of a Feynman-Kac formula. In addition, integrability conditions must be imposed in order to ensure the finiteness of the expectations that we encounter.

For implications of stochastic volatility models on derivatives pricing and hedging, the interested reader is advised to look into the monograph of Fouque et al. (2000) and the references therein. See also the discussions and reference list in Barndorff-Nielsen and Shephard (2001). Recently, many authors have studied portfolio optimization problems with asset dynamics going beyond the classical geometric Brownian motion (or the Samuelson model). We would like to mention Bank and Riedel (1999), Benth, Karlsen, and Reikvam (2001a, 2001b, 2001c, 2002), Goll and Kallsen (2000), and Kallsen (2000), which model the risky asset as an exponential of a Lévy process, and Framstad, Øksendal, and Sulem (1999, 2001) who used a geometric stochastic differential equation driven by a Lévy process as stock price model. All of these references are based on asset dynamics, which do not take long range dependency of log returns into account, since the noise is driven by a Lévy process.

This paper is organized as follows: In Section 2 we give a rigorous formulation of the portfolio optimization problem together with some basic assumption. Section 3 shows some basic results on moments of the value process and the stochastic volatility model which are useful later. To prove the explicit solution derived in Section 5, we prove a verification theorem under some (natural) integrability conditions in Section 4. In Section 6, we discuss our results (and the conditions put forth) and relate them to the empirically based models of Barndorff-Nielsen and Shephard (2001). Finally, in Section 7 optimal investment strategies where varying volatility is taken into account are compared numerically with the classical solution of Merton (1969, 1971). We demonstrate that stochastic volatility may lead to significantly different market behavior.

2. FORMULATION OF THE OPTIMIZATION PROBLEM

2.1. The Financial Market Model

For $0 \leq t \leq T < \infty$, let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a filtration $\{\mathcal{F}_s\}_{0 \leq s \leq T}$ satisfying the usual conditions. Introduce $m$ independent subordinators
$Z_j(s)$ (i.e., pure-jump Lévy processes with no drift and positive increments) and denote the corresponding Lévy measures by $\ell_j(dz)$, $j = 1, \ldots, m$. The Lévy measure of a subordinator satisfies the integrability condition
\[ \int_{0+}^{\infty} \min(1, z) \ell(dz) < \infty. \]

We choose the (unique) c.admissible (i.e., RCLL) version of $Z_j(s)$. Let $B(s)$ be a Wiener process independent of all the subordinators.

Let us introduce the stochastic volatility model proposed by Barndorff-Nielsen and Shephard (2001): Denote by $Y_j(s)$, $j = 1, \ldots, m$, the Ornstein-Uhlenbeck stochastic processes with dynamics
\[ dY_j(s) = -\lambda_j Y_j(s) \, ds + dZ_j(\lambda_j s), \tag{2.1} \]
where $\lambda_j$ is a positive number. Observe that
\[ Y_j(s) = y_j e^{-\lambda_j (s-t)} + \int_t^s e^{-\lambda_j (s-u)} \, dZ_j(\lambda_j u), \]
and thus
\[ Y(s) = (Y_1(s), \ldots, Y_m(s)) \geq 0 \text{ a.s. for all } t \leq s \leq T, \]
as soon as $y_j = Y_j(t) \geq 0$, $j = 1, \ldots, m$. Define the (stochastic) volatility of the stock to have the dynamics
\[ \sigma^{i,y}(s) = \sum_{j=1}^{m} \omega_j Y_j(s), \quad s \in [t, T]. \]

where $\omega_j > 0$ are weights summing to one and $\sigma^{i,y}(t) = \sum_{j=1}^{m} \omega_j y_j =: \sigma$ is the initial volatility at time $t$. We shall frequently write $\sigma^{y}(s)$ for $\sigma^{0,y}(s)$. This will also be the case for other processes when they start at time zero. Inserting $Y_j(s)$ into the volatility process leads to
\[ \sigma^{t,y}(s) = \sum_{j=1}^{m} \omega_j y_j e^{-\lambda_j (s-t)} + \int_t^s \sum_{j=1}^{m} \omega_j e^{-\lambda_j (s-u)} \, dZ_j(\lambda_j u). \]

To follow the modeling perspective of Barndorff-Nielsen and Shephard (2001), we introduce the stock price dynamics as the stochastic process ($s \geq t$)
\[ S(s) = x \exp \left( \int_t^s \mu + \beta \sigma(u) \, du + \int_t^s \sqrt{\sigma(u)} \, dB(u) \right). \tag{2.2} \]
Here, $\mu$, $\beta$ are constants and $S(t) = x$ is the initial stock price at time $t$. The stock price dynamics follows the stochastic differential equation
\[ dS(s) = (\mu + \left( \frac{1}{2} + \beta \right) \sigma(s)) S(s) \, ds + \sqrt{\sigma(s)} S(s) \, dB(s). \]

Note that $\sigma(s)$ in fact is the square of what we usually call volatility. In order to avoid more cumbersome notation later, when dynamic programming equations are considered, we have chosen to model the volatility and stock price dynamics as in (2.2) (or, equivalently, (2.1)). We assume the usual risk-free asset dynamics
\[ dR(s) = r R(s) \, ds, \]
with interest-rate $r > 0$. 

2.2. Discussion of the Asset Price Model

Frequently, log returns of financial data over small time scales are heavy tailed and possess (quasi-)long-range dependence. These stylized facts cannot be modeled using geometric Brownian motion (the Black and Scholes or Samuelson model). By modeling the volatility as a weighted sum of non-Gaussian Ornstein-Uhlenbeck processes, Barndorff-Nielsen and Shephard (2001) obtain a flexible family of asset dynamics that capture typical autocorrelation and heavy-tailed behavior observed in data (see, e.g., their analysis on deutsche mark–dollar exchange rates).

We now discuss the asset price model (2.2) in more detail. When $\beta = 0$, it is seen directly from (2.2) that the logarithmic return over the interval $t$ to $t + \Delta$ is normally distributed when conditioned on the aggregated volatility $\int_t^{t+\Delta} \sqrt{\sigma(s)} \, ds$. Skewness in the conditional distribution of the log returns is obtained by choosing $\beta \neq 0$.

The time scaling $Z_j(\lambda_j s)$ in (2.1) makes the invariant distribution of $Y_j$ independent of $\lambda_j$. The idea of Barndorff-Nielsen and Shephard (2001) is to separate the modeling of long-range dependency and non-Gaussianity in log return data. Using the Ornstein-Uhlenbeck dynamics to build up the volatility dynamics leads to an autocorrelation function being a weighted sum of exponentials decaying at rate $\lambda_j$. The correlation between the squared of the log returns becomes an explicit function of this autocorrelation function yielding models of log returns with long-range or quasi-long-range dependence.

As mentioned, (2.2) gives normally distributed log returns conditional on the aggregated volatility. By suitably choosing the subordinator $Z$, we get a normal mixture as the resulting log-return distribution. The philosophy of Barndorff-Nielsen and Shephard (2001) is to choose the invariant distribution of $Y_j$ instead, so that the log-return distribution may fit observations. This leads to a problem of choosing the right Lévy process $Z$ yielding the desired invariant distribution. They call these Lévy processes the background driving Lévy processes (BDLP, for short), and provide a characterization of them.

We may, for example, choose the invariant distribution of $Y$ (assuming $m = 1$) to be inverse Gaussian, implying (approximately) normal inverse Gaussian distributed logreturns. The Lévy measure for the BDLP is known in this case. The normal inverse Gaussian distribution has proved its relevance in finance as a flexible, yet good, model for log-return data (see, e.g, Barndorff-Nielsen 1998, Eberlein and Keller 1995, Rydberg 1997, and Bølviken and Benth 2000). The framework of Barndorff-Nielsen and Shephard (2001) also includes the variance-gamma models frequently used in finance (see, Madan, Carr, and Chang 1990, 1998). The log returns become (approximately) variance-gamma distributed by letting the invariant distribution of $Y$ be a gamma law. Also in this case the subordinator $Z$ is explicitly known through the Lévy measure.

We refer the interested reader to Barndorff-Nielsen and Shephard (2001) for a more detailed account on modeling issues and estimation procedures for these non-Gaussian Ornstein-Uhlenbeck stochastic volatility models.

2.3. The Control Problem and Basic Assumptions

Let $\pi(s)$ be the fraction of wealth invested in the stock at time $s$, and assume that there are no transaction costs in the market. The wealth process $W(s)$ is the sum of the position in the stock and the risk-free asset:

$$W(s) = \frac{\pi(s) S(s)}{R(s)} + \frac{(1 - \pi(s)) R(s)}{R(s)}.$$
The self-financing hypothesis yields the wealth dynamics for \( t < s \leq T \):
\[
\begin{align*}
dW(s) &= \pi(s) \left( (\mu + (\frac{1}{2} + \beta) \sigma(s)) - r \right) W(s) \, ds \\
&\quad + r W(s) \, ds + \pi(s) \sqrt{\sigma(s)} W(s) \, dB(s),
\end{align*}
\]
with initial wealth \( W(0) = w \).

The set of admissible controls \( \pi \) is defined as follows.

**Definition 2.1.** An investment strategy (control) \( \pi = \{ \pi(s); t \leq s \leq T \} \) is said to be admissible, and we write \( \pi \in \mathcal{A} \), if \( \pi \) is progressively measurable, \( \pi(s) \in [0, 1] \) almost surely for all \( t \leq s \leq T \), and a unique solution \( W^\pi \) of (2.3) exists.

Note that the control set \( \mathcal{A} \) only depends on the current time and not on the level of wealth nor volatility. We emphasize that the restriction \( \pi \in [0, 1] \) is just for mathematical convenience. Instead, we could have assumed \( \pi \in [\underline{\pi}, \overline{\pi}] \) for some constants \( -\infty < \underline{\pi} < \overline{\pi} < \infty \). All arguments below go through with minor modifications in this case; however, the analysis is more transparent when \( \underline{\pi} = 0 \) and \( \overline{\pi} = 1 \). Since the processes \( Y_j(s) \) are right-continuous it follows that \( \sigma(s) \) is right-continuous. Thus \( \int_0^T \sigma(s) \, ds < \infty \) almost surely, which together with \( \pi \in [0, 1] \) yields that \( \int_0^T \pi(s) \sqrt{\sigma(s)} \, dB(s) \) is a well-defined local martingale. In this paper we need to impose exponential integrability conditions on the Lévy measures \( \ell_j(dz) \). These conditions imply the martingale property of the Itô integral.

**Condition 2.1.** For a constant \( c_j > 0 \),
\[
\int_{0^+}^\infty (e^{c_jz} - 1) \ell_j(\,dz\,) < \infty.
\]

Later we shall specify in detail the choice of the constant \( c_j \) in Condition 2.1 for each \( j = 1, \ldots, m \). At this stage we only assume for every \( j = 1, \ldots, m \) the existence of a \( c_j \) for which Condition 2.1 holds. Note that Condition 2.1 is a condition on the integrability of the tails of the Lévy measures, since \( e^{c_jz} - 1 \sim z \) when \( z \) is in the neighborhood of zero.

Under exponential integrability, we have
\[
\mathbb{E}[e^{aZ_j(k,t)}] = \exp \left( \lambda_j \int_{0^+}^\infty (e^{az} - 1) \ell_j(\,dz\,) t \right)
\]
as long as Condition 2.1 holds with \( c_j = a \).

Let \( y := (y_1, \ldots, y_m) \in \mathbb{R}^m \). We will use the symbol \( \mathbb{R}_+ \) for the set \((0, \infty)\). The spatial domain of our stochastic control problem is
\[
D := \{(w, y) \in \mathbb{R} \times \mathbb{R}^m : w > 0, y_1, \ldots, y_m \geq 0 \}.
\]

The functional to be optimized takes the form
\[
J(t, w, y; \pi) = \mathbb{E}^{t,w,y}[U(W^\pi(T))],
\]
where the notation \( \mathbb{E}^{t,w,y} \) means that we take the expectation conditioned on \( W(t) = w \) and \( Y_j(t) = y_j, j = 1, \ldots, m \). \( U \) is the investor's utility function, being concave, non-decreasing, and bounded from below. In addition, we assume that \( U \) is of sublinear growth—that is, there exist positive constants \( k \) and \( \gamma \in (0, 1) \) so that \( U(w) \leq k(1 + w^\gamma) \)
for all \( w \geq 0 \). Our optimal stochastic control problem consists of determining the value function

\[
V(t, w, y) = \sup_{\pi \in A_t} J(t, w, y; \pi), \quad (t, x) \in [0, T] \times \mathbb{D},
\]

along with an optimal investment strategy \( \pi^* \in A_t \) such that

\[
V(t, w, y) = J(t, w, y; \pi^*).
\]

Observe that

\[
V(T, w, y) = U(w) \quad \forall (w, y) \in \mathbb{D}, \quad V(t, 0, y) \equiv U(0) \quad \forall (t, y) \in [0, T] \times \mathbb{R}^m_+.
\]

The HJB equation associated to our stochastic control problem is

\[
v_t + \max_{\pi \in [0, 1]} \left\{ \pi \left( \mu + \left( \frac{1}{2} + \beta \right) \sigma - r \right) w v_w + \frac{1}{2} \pi^2 \sigma w^2 v_{ww} \right\} + rwv_w
\]

\[
- \sum_{j=1}^m \lambda_j y_j v_{y_j} + \sum_{j=1}^m \lambda_j \int_0^\infty \left( v(t, w, y + z \cdot e_j) - v(t, w, y) \right) \ell_j(dz) = 0,
\]

for \((t, w, y) \in [0, T) \times D\). In view of (2.6), we augment (2.7) with the terminal condition

\[
v(T, w, y) = U(w) \quad \forall (w, y) \in \mathbb{D}
\]

and the boundary condition

\[
v(t, 0, y) = U(0) \quad \forall (t, y) \in [0, T] \times \mathbb{R}^m_+.
\]

We have used the following notational convention: \( v_t = \partial v / \partial t \), \( v_w = \partial v / \partial w \), \( v_{ww} = \partial^2 v / \partial w^2 \), and \( v_{y_j} = \partial v / \partial y_j \), \( j = 1, \ldots, m \).

### 3. PRELIMINARY ESTIMATES

The following lemmas are useful in relating the existence of exponential moments of \( Y \) to exponential integrability conditions on the Lévy measures (see also Benth et al. 2003).

**Lemma 3.1.** For \( \xi_j > 0 \), assume Condition 2.1 holds with \( c_j = \xi_j / \lambda_j \). Then

\[
\mathbb{E} \left[ \exp \left( \xi_j \int_t^s Y_j(u) \, du \right) \right] \leq \exp \left( \xi_j \lambda_j y_j + \lambda_j \int_0^\infty \{ \exp(\xi_j z / \lambda_j) - 1 \} \ell_j(dz)(s - t) \right).
\]

**Proof.** From the dynamics of \( Y_j \) we find

\[
\lambda_j \int_t^s Y_j(u) \, du = y_j + Z_j(\lambda_j s) - Z_j(\lambda_j t) - Y_j(s)
\]

\[
\leq y_j + Z_j(\lambda_j s) - Z_j(\lambda_j t)
\]

\[
\overset{d}{=} y_j + Z_j(\lambda_j(s - t)),
\]

since \( Y_j(s) \geq 0 \). The last equality is in the sense of equality in distribution. Hence, from (2.4),
Lemma 3.1 with $\xi_j$ where $H^n$ by Hölder’s inequality and using that $K$ for $E^s_j$ then

\[ \mathbb{E}\left[ \exp\left( \xi_j \int_t^s Y_j(u) \, du \right) \right] \leq \exp\left( \frac{\xi_j}{\lambda_j} y_j \right) \mathbb{E}\left[ \exp\left( \frac{\xi_j}{\lambda_j} Z(\lambda_j(s - t)) \right) \right] \]

\[ = \exp\left( \frac{\xi_j}{\lambda_j} y_j + \lambda_j \int_{0+}^\infty \{\exp(\xi_j z/\lambda_j) - 1\} \ell_j(dz)(s - t) \right), \]

which proves the lemma. \hfill \square

In a completely analogous way, we can show the following lemma.

**Lemma 3.2.** Assume Condition 2.1 holds for some positive constant $c_j$. Then

\[ \mathbb{E}[\exp(c_j Y_j(s))] \leq \exp \left( c_j y_j + \lambda_j \int_{0+}^\infty \{\exp(c_j z) - 1\} \ell_j(dz)(s - t) \right). \]

The following result provides us with a useful moment estimate on the wealth process.

**Lemma 3.3.** For some $\theta > 0$, assume Condition 2.1 holds with $c_j = 2\theta(|\frac{1}{2} + \beta| + \theta) \frac{\omega_j}{\lambda_j}$ for $j = 1, \ldots, m$. Then

\[ \sup_{\pi \in \mathcal{A}} \mathbb{E}^\pi \left[ (W^\pi(s))^\theta \right] \leq w^\theta \exp \left( K(\theta) \sum_{j=1}^m \frac{\omega_j}{\lambda_j} y_j + C(\theta)(s - t) \right), \]

where $K(\theta) = \theta \left( |\frac{1}{2} + \beta| + \theta \right)$ and

\[ C(\theta) = \theta(|\mu - r| + r) + \frac{1}{2} \sum_{j=1}^m \lambda_j \int_{0+}^\infty \left( e^{\frac{\omega_j}{\lambda_j} s} - 1 \right) \ell_j(dz). \]

**Proof.** First we have

\[ W^\pi(s) = w \exp \left( \int_t^s \alpha(u, \sigma(u)) \, du + \int_t^s \pi(u) \sqrt{\sigma(u)} \, dB(u) \right), \]

where

\[ \alpha(u, \sigma) = \pi(u) (\mu + (\frac{1}{2} + \beta) \sigma - r) + r - \frac{1}{2} (\pi(u))^2 \sigma. \]

Define

\[ X(s) = \exp \left( \int_t^s 2\theta \pi(u) \sqrt{\sigma(u)} \, dB(u) - \frac{1}{2} \int_t^s (2\theta)^2 (\pi(u))^2 \sigma(u) \, du \right). \]

Then $\mathbb{E}[X(s)] = 1$ since $X(s)$ is a martingale by Novikov’s condition. We have, by Lemma 3.1 with $\xi_j = 2\theta^2 \omega_j$ for $j = 1, \ldots, m$,

\[ \mathbb{E} \left[ e^{\xi_j \int_t^s (2\theta)^2 (\sigma(u))^2 \sigma(u) du} \right] \leq \mathbb{E} \left[ e^{\theta^2 \int_t^s (2\theta)^2 (\sigma(u))^2 \sigma(u) du} \right] = \prod_{j=1}^m \mathbb{E} \left[ e^{\theta^2 \omega_j \xi_j Y_j(u) du} \right] < \infty. \]

Hence, by Hölder’s inequality and using that $\pi \in [0, 1]$,

\[ \mathbb{E}[(W^\pi(s))^\theta] = w^\theta \mathbb{E} \left[ \exp \left( \theta \int_t^s \alpha(u, \sigma(u)) \, du + \theta \int_t^s \pi(u) \sqrt{\sigma(u)} \, dB(u) \right) \right] \]

\[ = w^\theta \mathbb{E} \left[ \exp \left( \theta \int_t^s \alpha(u, \sigma(u)) \, du + \theta^2 \int_t^s (\pi(u))^2 \sigma(u) \, du \right) X(s)^{1/2} \right] \]
\[ \leq w^\theta \mathbb{E} \left[ \exp \left( \int_t^s \left( 2\theta \alpha(u, \sigma(u)) + 2\theta^2 (\pi(u))^2 \sigma(u) du \right) \right) \right]^{1/2} \mathbb{E} [X(s)]^{1/2} \]
\[ \leq w^\theta e^{(s-t)(\mu-r)+r} \mathbb{E} \left[ \exp \left( (2\theta \left( |\frac{1}{2} + \beta| + 2\theta^2 \right) \int_t^s \sigma(u) du \right) \right]^{1/2} \]
\[ = w^\theta e^{(s-t)(\mu-r)+r} \prod_{j=1}^m \mathbb{E} \left[ \exp \left( 2\theta \left( |\frac{1}{2} + \beta| + \theta \right) \omega_j \int_t^s Y_j(u) du \right) \right]^{1/2}. \]

Choosing \( \xi_j = 2\theta \left( |\frac{1}{2} + \beta| + \theta \right) \omega_j \), \( j = 1, \ldots, m \), in Lemma 3.1 we obtain the desired estimate.

The next proposition shows that the value function of our control problem is well-defined.

**Proposition 3.4.** Assume Condition 2.1 holds with \( c_j = 2\gamma \left( |\frac{1}{2} + \beta| + \gamma \right) \omega_j \lambda_j \) for \( j = 1, \ldots, m \). Then

\[ U(0) \leq V(t, w, y) \leq k \left( 1 + w^\gamma e^{K(\gamma) \sum_{j=1}^m \omega_j \lambda_j y_j + C(\gamma)(T-t)} \right), \]

where \( K(\gamma) \) and \( C(\gamma) \) are defined in Lemma 3.3 and \( k \) is a positive constant.

**Proof.** Since \( U \) is nondecreasing, we have \( U(w) \geq U(0) \). Hence \( \mathbb{E}[U(W^\pi(T))] \geq U(0) \) for any control \( \pi \), which implies \( V(t, w, y) \geq U(0) \). The upper bound follows from the sublinear growth of \( U \) and Lemma 3.3:

\[ V(t, w, y) = \sup_{\pi \in A} \mathbb{E} [U(W^\pi(T))] \leq k \left( 1 + \sup_{\pi \in A} \mathbb{E} [(W^\pi(T))^\gamma] \right), \]

which proves the proposition.

**Remark 3.1.** Under additional exponential integrability conditions on the Lévy measures \( \ell_j (dz) \), local Hölder continuity of the value function in all variables is proved in Benth et al. (2003). To establish such a result, one needs of course that the utility function \( U \) is Hölder continuous.

From now on we suppose that Condition 3.1—which ensures that the value function is well-defined—holds.

**Condition 3.1.** For all \( j = 1, \ldots, m \), Condition 2.1 holds with \( c_j = \gamma \left( |\frac{1}{2} + \beta| + \gamma \right) \omega_j \lambda_j \).

**4. A VERIFICATION THEOREM**

We state and prove the following verification theorem for our stochastic control problem.

**Theorem 4.1.** Assume that \( v(t, w, y) \in C^{1,2,1}([0, T] \times (0, \infty) \times [0, \infty)^m) \cap C([0, T] \times \overline{D}) \) is a solution of the HJB equation (2.7) with terminal and boundary conditions
(2.8) and (2.9). For \( j = 1, \ldots, m \), assume
\[
\sup_{\pi \in \mathcal{A}_t} \int_0^T \int_{0+} \mathbb{E} \left[ \left| v(s, W^\pi(s), Y(s) + z \cdot e_j) - v(s, W^\pi(s), Y(s)) \right| \right] \ell_j (dz) \, ds < \infty
\]
and
\[
\sup_{\pi \in \mathcal{A}_t} \int_0^T \mathbb{E} \left[ (\pi(s))^2 \sigma(s) (W^\pi(s))^2 (v_w(s, W^\pi(s), Y(s)))^2 \right] \, ds < \infty.
\]
Then
\[
v(t, w, y) \geq V(t, w, y), \quad \forall (t, w, y) \in [0, T] \times \overline{D}.
\]
If, in addition, there exists a measurable function \( \pi^*(t, w, y) \in [0, 1] \) being a maximizer for the max-operator in (2.7), then \( \pi^* \) defines an optimal investment strategy in feedback form if (2.3) admits a unique solution \( W^{\pi^*} \) and
\[
V(t, w, y) = v(t, w, y) = \mathbb{E}^{t, w, y} \left[ U(W^{\pi^*}(T)) \right], \quad \forall (t, w, y) \in [0, T] \times \overline{D}.
\]

**Remark 4.1.** The notation \( C^{1,2,1}([0, T] \times (0, \infty) \times [0, \infty)^m) \) means twice continuously differentiable in \( w \) on \( (0, \infty) \) and once continuously differentiable in \( t, y \) on \( (0, T) \times (0, \infty)^m \) with continuous extensions of the derivatives to \( t = 0 \) and \( y_j = 0, j = 1, \ldots, m \).

**Proof.** Let \( (t, w, y) \in [0, T] \times \overline{D} \) and \( \pi \in \mathcal{A}_t \), and introduce the operator
\[
\mathcal{M}^\pi v = \pi \left( \mu + \left( \frac{1}{2} + \beta \right) \sigma - r \right) w v_w - \frac{1}{2} \sigma w^2 \pi^2 v_{ww} + r w v_w - \sum_{j=1}^m \lambda_j y_j v_{y_j}.
\]
Itô’s Formula (see, e.g., Ikeda and Watanabe 1989) yields (with \( t \leq s \leq T \))
\[
v(s, W^\pi(s), Y(s)) = v(t, w, y) + \int_t^s \left\{ v_t(u, W^\pi(u), Y(u)) + \mathcal{M}^\pi v(u, W^\pi(u), Y(u)) \right\} \, du
\]
\[
+ \int_t^s \pi(u) \sqrt{\sigma(u)} W^\pi(u) v_w(u, W^\pi(u), Y(u)) \, dB(u)
\]
\[
+ \sum_{j=1}^m \int_t^s \int_{0+} \left( v(u, W^\pi(u), Y(u -) + z \cdot e_j) - v(u, W^\pi(u), Y(u -)) \right) N_j(\lambda_j \, du, dz),
\]
where \( N_j \) is the (homogeneous) Poisson random measure coming from the Lévy-Kinchine representation of the subordinator \( Z_j \). From the assumptions we know that the Itô integral is a martingale and that the integrals with respect to \( N_j \) are semimartingales (not only local semimartingales). Hence, taking expectations on both sides implies
\[
\mathbb{E} [v(s, W^\pi(s), Y(s))] = v(t, w, y) + \mathbb{E} \left[ \int_t^s \left( v_t + \mathcal{L} v \right)(u, W^\pi(u), Y(u)) \, du \right]
\]
\[
\leq v(t, w, y) + \mathbb{E} \left[ \int_t^s \left( v_t + \max_{\pi \in [0, 1]} \mathcal{L} v \right)(u, W^\pi(u), Y(u)) \, du \right]
\]
\[
= v(t, w, y),
\]
where
\[ \mathcal{L}^\pi v = \mathcal{M}^\pi v + \sum_{j=1}^m \lambda_j \int_0^\infty (v(t, w, y + z \cdot e_j) - v(t, w, y)) \ell_j (dz). \]

Putting \( s = T \) and invoking the terminal condition for \( v \), we find
\[ v(t, w, y) \geq \mathbb{E} [U(W^{\pi^*}(T))], \]
for all \( \pi \in \mathcal{A}_t \). Therefore the first conclusion in the theorem holds for \((t, w, y)\in[0, T) \times D\).

To prove the second part, observe that since \( \pi^*(t, w, y) \) is assumed to be a measurable function, we have that \( \pi^*(s, W(s), Y(s)) \) is \( \mathcal{F}_s \)-measurable for \( t \leq s \leq T \). This, together with the assumptions that \( \pi^* \in [0, 1] \) and the existence of a unique solution \( W^{\pi^*} \) of (2.3), implies that \( \pi^*(s, W(s), Y(s)) \) is an admissible (feedback) control (i.e., an element of \( \mathcal{A}_t \)). Moreover, since \( \pi^* \) is a maximizer, \( \max_{\pi \in [0, 1]} \mathcal{L}^\pi v = \mathcal{L}^{\pi^*} v \). The above calculations using Itô’s Formula go through with equality by letting \( \pi = \pi^* \). Hence,
\[ v(t, w, y) = \mathbb{E}[U(W^{\pi^*}(T))] \leq V(t, w, y). \]
Together with the first part of the theorem, this yields
\[ v(t, w, y) = V(t, w, y) = \mathbb{E}[U(W^{\pi^*}(T))], \]
for \((t, w, y)\in[0, T) \times D\).

Observe that from the terminal and boundary conditions (2.8) and (2.9), the two conclusions of the theorem obviously hold when \( t = T \) and \( w = 0 \). Hence the theorem is proved.

**Remark 4.2.** In Section 5 we construct an explicit solution of the HJB equation (2.7) when \( U \) is a power utility. Theorem 4.1 is used to prove that this solution coincides with the value function.

## 5. Explicit Solution

In this section we shall construct and verify an explicit solution to the control problem (2.5) together with an explicit optimal control \( \pi^* \) when the utility function is of power type; that is,
\[ U(w) = \gamma^{-1} w^\gamma, \]
where \( 1 - \gamma \) is known as the relative risk aversion of the investor and \( \gamma \in (0, 1) \). These power functions are also known as HARA utility functions.

### 5.1. Reduction of the HJB equation

Define
\[ v(t, w, y) = \gamma^{-1} w^\gamma h(t, y), \quad (t, w, y) \in [0, T) \times D, \]
for some function \( h(t, y) \). Observe that \( v(t, 0, y) = U(0) \). Inserted into the HJB equation (2.7) we get a first-order integro-differential equation for \( h \):
\begin{align*}
(5.1) \quad h_t(t, y) + \gamma \Pi(\sigma) h(t, y) - \sum_{j=1}^{m} \lambda_j y_j h_j(t, y) \\
+ \sum_{j=1}^{m} \lambda_j \int_{0^+}^{\infty} (h(t, y + z \cdot e_j) - h(t, y)) \ell_j (dz) = 0,
\end{align*}

where \((t, y) \in [0, T] \times [0, \infty)^m\). The terminal condition is \(h(T, y) = 1\) for all \(y \in [0, \infty)^m\), since \(v(T, w, y) = U(w) = \gamma^{-1}w^\gamma\). Recall here that \(\sigma = \sum_{j=1}^{m} \omega_j y_j\). The function \(\Pi : [0, \infty) \to \mathbb{R}\) is defined as

\begin{align*}
(5.2) \quad \Pi(\sigma) = \max_{\pi \in [0, 1]} \left\{ \pi \left( \mu + \left( \frac{1}{2} + \beta \right) \sigma - r \right) - \frac{1}{2} \pi^2 \sigma (1 - \gamma) \right\} + r.
\end{align*}

We calculate an explicit representation of \(\Pi\). A first-order condition for an interior optimum is

\begin{align*}
\left( \mu + \left( \frac{1}{2} + \beta \right) \sigma - r \right) - \pi \sigma (1 - \gamma) = 0.
\end{align*}

If we denote the interior optimum by \(\bar{\pi} = \bar{\pi}(\sigma)\), then this gives

\begin{align*}
\bar{\pi}(\sigma) = \frac{1}{1 - \gamma} \left( \frac{\mu - r}{\sigma} + \frac{1}{2} + \beta \right).
\end{align*}

Note that \(\bar{\pi}(\sigma)\) is a function in \(\sigma\) only, and not in its different components \(y_j\) explicitly. We can thus treat this interior optimum as a function on \((0, \infty)\). Note from the constraints that \(\bar{\pi}(\sigma)\) is an optimum if \(\bar{\pi}(\sigma) \in (0, 1]\). If this is not the case, the optimum is reached either in 0 or in 1, depending on the parameters of the problem. We now investigate this more closely.

In the rest of this section we assume \(\mu > r\) (the analysis for \(\mu < r\) is analogous) and aim at finding \(\pi^*\), the value of \(\pi\) for which the maximum is reached in the expression of \(\Pi\). Observe that \(\bar{\pi}(\sigma)\) is nonincreasing, \(\lim_{\sigma \to 0} \bar{\pi}(\sigma) = +\infty\), and

\begin{align*}
\lim_{\sigma \to \infty} \bar{\pi}(\sigma) = \frac{\frac{1}{2} + \beta}{1 - \gamma}.
\end{align*}

We separate the further discussion into three cases:

\textit{Case 5.1:} \(\frac{1}{2} + \beta \geq 1\). Under this assumption we see that \(\bar{\pi} \geq 1\) for all \(\sigma\), and hence

\begin{align*}
\pi^*(\sigma) = 1, \quad \sigma \in [0, \infty).
\end{align*}

Inserting this into the expression for \(\Pi\) we get

\begin{align*}
\Pi(\sigma) = \mu + \left( \frac{\gamma}{2} + \beta \right) \sigma, \quad \sigma \in [0, \infty).
\end{align*}

Define the constant \(b_1 := \beta + \frac{1}{2} \gamma\), and observe that \(b_1 > 0\).

\textit{Case 5.2:} \(\frac{1}{2} + \beta \in (0, 1)\). Under this assumption we see that there exists a \(\hat{\sigma}_1\) such that \(\bar{\pi}(\hat{\sigma}_1) = 1\) and \(\bar{\pi}(\sigma) \in (0, 1)\) for all \(\sigma > \hat{\sigma}_1\). A straightforward calculation gives

\begin{align*}
\hat{\sigma}_1 = \frac{\mu - r}{(1 - \gamma) - \left( \frac{1}{2} + \beta \right)}.
\end{align*}
Hence, the optimal $\pi$ is given as
\begin{equation}
\pi^*(\sigma) = \begin{cases} 1, & \sigma \in [0, \hat{\sigma}_1), \\ \bar{\pi}(\sigma), & \sigma \in [\hat{\sigma}_1, \infty). \end{cases}
\end{equation}

The expression for $\Pi$ now becomes
\begin{equation}
\Pi(\sigma) = \begin{cases} \mu + \left(\frac{\gamma}{2} + \beta\right)\sigma, & \sigma \in [0, \hat{\sigma}_1), \\ \frac{(\mu - r)^2}{2(1 - \gamma)\sigma} + \frac{(\mu - r)(\frac{\gamma}{2} + \beta)}{(1 - \gamma)} + \frac{1}{2(1 - \gamma)} + r, & \sigma \in [\hat{\sigma}_1, \infty). \end{cases}
\end{equation}

Moreover, it is easily seen that
\[|\Pi(\sigma)| \leq a + b_2\sigma,\]
for $b_2 = \left(\frac{1}{2} + \beta\right)^2 / 2(1 - \gamma)$ and some constant $a$.

**Case 5.3:** $\frac{1}{2} + \beta < 0$. Observe that this assumption is equivalent to $\frac{1}{2} + \beta < 0$ since $\gamma \in (0, 1)$. In this situation there will exist a $\hat{\sigma}_1$ such that $\bar{\pi}(\hat{\sigma}_1) = 1$ and a $\hat{\sigma}_0$ such that $\bar{\pi}(\hat{\sigma}_0) = 0$. The former we calculated above; the latter is easily found to be
\[\hat{\sigma}_0 = -\frac{\mu - r}{\frac{1}{2} + \beta}.
\]
The optimal $\pi$ is
\begin{equation}
\pi^*(\sigma) = \begin{cases} 1, & \sigma \in [0, \hat{\sigma}_1), \\ \bar{\pi}(\sigma), & \sigma \in [\hat{\sigma}_1, \hat{\sigma}_0], \\ 0, & \sigma \in (\hat{\sigma}_0, \infty). \end{cases}
\end{equation}

Hence the expression for $\Pi$ becomes
\begin{equation}
\Pi(\sigma) = \begin{cases} \mu + \left(\frac{\gamma}{2} + \beta\right)\sigma, & \sigma \in [0, \hat{\sigma}_1), \\ \frac{(\mu - r)^2}{2(1 - \gamma)\sigma} + \frac{(\mu - r)(\frac{\gamma}{2} + \beta)}{(1 - \gamma)} + \frac{1}{2(1 - \gamma)} + r, & \sigma \in [\hat{\sigma}_1, \hat{\sigma}_0], \\ r, & \sigma \in (\hat{\sigma}_0, \infty). \end{cases}
\end{equation}

We now prove that $\Pi(\sigma)$ is, in fact, continuously differentiable on $[0, \infty)$ in all three cases.

**Lemma 5.1.** Assume $\mu > r$. Then the function $\Pi(\sigma)$ defined in (5.2) is continuously differentiable on $[0, \infty)$.

**Proof.** It is obvious that $\Pi$ is continuous since $\pi^*$ defined in (5.3), (5.4), and (5.5) are continuous, and
\[
\Pi(\sigma) = \pi^*(\sigma) \left(\mu + \left(\frac{1}{2} + \beta\right)\sigma - r\right) - \frac{1}{2}\pi^*(\sigma)^2\sigma(1 - \gamma) + r.
\]

In Case 5.1, $\Pi$ is trivially differentiable. Furthermore, to prove differentiability of $\Pi$ in the two subsequent cases, we must show that $\Pi$ is differentiable at $\sigma = \hat{\sigma}_1$ in Case 5.2
and at $\sigma = \hat{\sigma}_1$ and $\sigma = \hat{\sigma}_0$ in Case 5.3. But it is then sufficient to only consider Case 5.3:

$$
\Pi'(\sigma) = \begin{cases} 
\frac{\xi}{2} + \beta, & \sigma \in [0, \hat{\sigma}_1) \\
\frac{(\mu - r)^2}{2(1 - \gamma)\sigma^2} + \frac{(\frac{1}{\gamma} - \beta)^2}{2(1 - \gamma)}, & \sigma \in (\hat{\sigma}_1, \hat{\sigma}_0) \\
0, & \sigma \in (\hat{\sigma}_0, \infty).
\end{cases}
$$

Straightforward calculations show

$$
\lim_{\sigma \uparrow \hat{\sigma}_1} \Pi'(\sigma) = \frac{\gamma^2}{2} + \beta = \lim_{\sigma \downarrow \hat{\sigma}_1} \Pi'(\sigma)
$$

and

$$
\lim_{\sigma \uparrow \hat{\sigma}_0} \Pi'(\sigma) = 0 = \lim_{\sigma \downarrow \hat{\sigma}_0} \Pi'(\sigma),
$$

implying the differentiability on $[0, \infty)$ of $\Pi$. This proves the lemma. \(\square\)

### 5.2. A Feynman-Kac formula for $h(t, y)$

Define the function $g(t, y)$ by

$$(5.6) \quad g(t, y) = \mathbb{E}^\gamma \left[ \exp^{\gamma \Pi(\sigma(s)) ds} \right], \quad (t, y) \in [0, T] \times [0, \infty)^m,$$

and recall that $\sigma^y(0) = \sum_{j=1}^m \omega_j y_j =: \sigma$. Note that $g(0, y) = 1$. We first show that $g$ is of exponential growth in $\sigma$ and thus well-defined under a growth hypothesis.

**Lemma 5.2.** Assume Condition 2.1 holds with $c_j = \gamma b \omega_j \lambda_j$ for $j = 1, \ldots, m$, where $b$ is equal to $b_1$ in Case 5.1, $b_2$ in Case 5.2, and $b_3 = 0$ in Case 5.3. Then

$$
g(t, y) \leq \exp \left( kt + \gamma b \sum_{j=1}^m \omega_j \lambda_j y_j \right),
$$

for some positive constant $k$.

**Proof.** From the discussion in the previous subsection, we know that there exist positive constants $a$ and $b$ as given in the assumptions such that $|\Pi(\sigma)| \leq a + b\sigma$.

Therefore,

$$
g(t, y) = \mathbb{E}^\gamma \left[ \exp^{\gamma \Pi(\sigma(s)) ds} \right] \leq \mathbb{E}^\gamma \left[ \exp^{\gamma (at + t^2 b \lambda_j y_j) ds} \right] = \mathbb{E}^{\gamma at} \prod_{j=1}^m \mathbb{E}^{\gamma b \omega_j \lambda_j y_j}\left( e^{\int_{0}^{t} \frac{\omega_j \lambda_j y_j}{\xi_j} dt} \right).$$

By independence of the $Y_j$’s we get

$$
g(t, y) \leq \exp^{\gamma at} \prod_{j=1}^m \mathbb{E}^{\gamma b \omega_j \lambda_j y_j}\left( e^{\int_{0}^{t} \frac{\omega_j \lambda_j y_j}{\xi_j} dt} \right) \leq \exp^{\gamma at} \prod_{j=1}^m e^{\frac{\omega_j \lambda_j y_j}{\xi_j}} e^{\int_{0}^{t} \frac{\omega_j \lambda_j y_j}{\xi_j} dt} = \exp^{\gamma at} \prod_{j=1}^m e^{\int_{0}^{t} \frac{\omega_j \lambda_j y_j}{\xi_j} dt}.
$$

To derive the last inequality, we used Lemma 3.1 with $\xi_j = \gamma b \omega_j$. Hence, there exists a positive constant $k$ such that

$$
g(t, y) \leq \exp^{kt + \gamma \sum_{j=1}^m \omega_j \lambda_j y_j},
$$

and the lemma is proved. \(\square\)
We show next that \( g \) is continuously differentiable in \( y \).

**Lemma 5.3.** Assume Condition 2.1 holds with \( c_j = \gamma b_j \) for \( j = 1, \ldots, m \), where \( b \) is equal to \( b_1 \) in Case 5.1, \( b_2 \) in Case 5.2, and \( b_3 = 0 \) in Case 5.3. Then \( g \in C^0,1([0, T] \times [0, \infty)^m) \); that is, \( g(\cdot, y) \) is continuous for all \( y \in [0, \infty)^m \) and \( g(t, \cdot) \) is once continuously differentiable for all \( t \in [0, T] \).

**Proof.** To prove differentiability, we will use the dominated convergence theorem to show that we may interchange expectation and differentiation. The condition for this is contained in Theorem 2.27 of Folland (1984), which essentially says that we need to bound the derivative by an integrable function independent of \( y \).

Let \((t, y) \in [0, T] \times \mathbb{R}_+^m \) and introduce the function

\[
F(t, y) = e^{\int_0^t \gamma \Pi(\sigma^y(s)) \, ds}.
\]

For each \( j = 1, \ldots, m \), we have

\[
\frac{\partial F(t, y)}{\partial y_j} = \left( \frac{\partial}{\partial y_j} \int_0^t \gamma \Pi(\sigma^y(s)) \, ds \right) e^{\int_0^t \gamma \Pi(\sigma^y(s)) \, ds}.
\]

By Lemma 5.1, \( \Pi \) is continuously differentiable and \( \Pi' \) is bounded. Hence

\[
\gamma \Pi'(\sigma^y(s)) \frac{\partial \sigma^y(s)}{\partial y_j} = \gamma \Pi'(\sigma^y(s)) \omega_j e^{-\lambda_j s} \leq c e^{-\lambda_j s}
\]

for some strictly positive constant \( c \). Theorem 2.27(b) in Folland (1984) says that differentiation and integration now commute:

\[
\frac{\partial F(t, y)}{\partial y_j} = \left( \gamma \omega_j \int_0^t \Pi'(\sigma^y(s)) e^{-\lambda_j s} \, ds \right) e^{\int_0^t \gamma \Pi(\sigma^y(s)) \, ds}.
\]

From the discussion in Section 5.1 we know there exist constants \( a \) and \( b \) such that \(|\Pi(\sigma)| \leq a + b \sigma \) in Cases 5.1 and 5.2, and \( \Pi(\sigma) \leq a \) in Case 5.3, where \( b = \frac{1}{2} + \beta \) and \( b = \frac{1 + \beta}{2(1 - \gamma)} \) in Cases 5.1 and 5.2, respectively. Hence,

\[
\left| \frac{\partial F(t, y)}{\partial y_j} \right| \leq \left( c \int_0^t e^{-\lambda_j s} \, ds \right) e^{\int_0^t \gamma \Pi(\sigma^y(s)) \, ds}
\]

\[
\leq \frac{c}{\lambda_j} e^{\gamma a T} \left\{ \begin{array}{ll}
e^{y b \int_0^t \sigma^y(s) \, ds}, & \text{Cases 5.1 and 5.2}, \\
1, & \text{Case 5.3}.
\end{array} \right.
\]

As in Lemma 3.1,

\[
\omega_j \gamma b \int_0^t \sigma^y_j(s) \, ds \leq \omega_j \gamma b \left( \frac{y_j}{\lambda_j} + \frac{1}{\lambda_j} Z_j(\lambda_j t) \right) \leq \frac{\omega_j \gamma b}{\lambda_j} (y_j + Z_j(\lambda_j t)),
\]

from which there exists a positive constant \( k \) such that

\[
\left| \frac{\partial F(t, y)}{\partial y_j} \right| \leq k \left\{ \sum_{j=1}^m e^{y b \int_0^t \sigma^y_j(s) \, ds}, \right. \text{Cases 5.1 and 5.2},
\]
\[
1, \text{ Case 5.3}.
\]

By (2.4), we have

\[
E\left[ e^{y b \int_0^t \sigma^y_j(\lambda_j t)} \right] = \int_0^\infty \left( e^{y b \sigma^y_j - 1} \right) \ell_j(dz),
\]

which is assumed finite.
For Cases 5.1 and 5.2, choose a compact set where \( y \) is in the interior. On this compact we have that \(|\partial F(t, y)/\partial y_j|\) is uniformly bounded (in \( y \)) by the random variable
\[
\exp\left(\frac{\omega_j \gamma b}{\lambda_j} Z_j(\lambda_j t)\right),
\]
which is integrable since by assumption \( \int_0^\infty (e^{\omega_j \gamma b z/\lambda_j} - 1) \ell_j (dz) < \infty \), for \( j = 1, \ldots, m \).

Theorem 2.27(b) in Folland (1984) implies that \( g(t, y) = \mathbb{E}[F(t, y)] \) is differentiable in \( y \). Differentiability is a local notion, hence the result is independent of the choice of the compact set. We conclude that
\[
\frac{\partial g(t, y)}{\partial y_j} = \mathbb{E}\left[ \frac{\partial F(t, y)}{\partial y_j} \right], \quad \forall y \in \mathbb{R}_+^m, \quad j = 1, \ldots, m.
\]

Moreover, we have that \( y \mapsto \partial F(t, y)/\partial y_j \) is continuous since \( y \mapsto \sigma^y(s), \sigma \mapsto \Pi(\sigma) \), and \( \sigma \mapsto \Pi'(\sigma) \) all are continuous mappings. Using Theorem 2.27(a) in Folland (1984) we conclude that the mapping \( (t, y) \mapsto \partial g(t, y)/\partial y_j \) is continuous.

For Case 5.3, \(|F(t, y)/\partial y_j| \leq k \) for some positive constant \( k \). Hence, Theorem 2.27(a)–(b) in Folland (1984) immediately applies to conclude that \( (t, y) \mapsto \partial g(t, y)/\partial y_j \) is continuous and
\[
\frac{\partial g(t, y)}{\partial y_j} = \mathbb{E}\left[ \frac{F(t, y)}{\partial y_j} \right], \quad j = 1, \ldots, m.
\]

Since the limit of \( \partial F(t, y)/\partial y_j \) exists when \( y_i \downarrow 0 \) for any \( i = 1, \ldots, m \), we can argue as above to show that \( \partial g(t, y)/\partial y_j \) has a limit when \( y_i \downarrow 0 \). This concludes the proof of the lemma. \( \square \)

**Remark 5.1.** Note that for Case 5.3 in Lemmas 5.2 and 5.3 we do not impose any integrability condition on the Lévy measures \( \ell_j (dz) \).

**Lemma 5.4.** Assume Condition 2.1 holds with \( c_j = 2\gamma b \omega_j^2 \) for \( j = 1, \ldots, m \), where \( b \) is equal to \( b_1 \) in Case 5.1 and \( b_2 \) in Case 5.2. In Case 5.3, assume \( \int_0^\infty z \ell_j (dz) < \infty, j = 1, \ldots, m \). Then
\[
\sum_{j=1}^m \mathbb{E}\left[ \int_0^T \int_{0+}^{\infty} |g(u, Y(u) + z \cdot e_j) - g(u, Y(u))| \ell_j (dz) du \right] < \infty.
\]

**Proof.** By the mean value theorem and differentiability of \( g \) we have
\[
|g(u, y + z \cdot e_j) - g(u, y)| \leq \sup_{x \in [0, z]} \left| \frac{\partial g(u, y + x \cdot e_j)}{\partial y_j} \right| z
\]
\[
\leq k z \begin{cases} 
\sum_{j=1}^m \frac{b y \omega_j}{\lambda_j} (y_j + Z_j(\lambda_j u)) & \text{Cases 5.1 and 5.2}, \\
1 & \text{Case 5.3},
\end{cases}
\]
where \( k \) is a positive constant only dependent on \( T \) and the parameters of the problem. Since
\[
\frac{b y \omega_j}{\lambda_j} (Z_j(\lambda_j u) + Y_j^y(u)) \leq \frac{b y \omega_j}{\lambda_j} (Z_j(\lambda_j u) + y_j + Z_j(\lambda_j u)) = \frac{b y \omega_j}{\lambda_j} y_j + \frac{2 b y}{\lambda_j} Z_j(\lambda_j u),
\]
we have

\[ |g(u, Y(u) + z \cdot e_j) - g(u, Y(u))| \leq k z \left\{ \sum_{j=1}^{m} \left( \frac{b_{y_j} u_{y_j}}{\lambda_j} y_j + 2 b_{y_j} y_j \right) \right\}, \]

Cases 5.1 and 5.2, Case 5.3.

From the integrability assumptions on \( \ell_j (dz) \) in Cases 5.1 and 5.2, we have from (2.4)

\[ \int_0^T \mathbb{E} \left[ \int_{0+}^{\infty} k z \ell_j (dz) e^{\sum_{j=1}^{m} \frac{2 b_{y_j} u_{y_j}}{\lambda_j} y_j} \right] du 
\]

\[ = k \int_{0+}^{\infty} z \ell_j (dz) e^{\sum_{j=1}^{m} \lambda_j \int_{0+}^{\infty} \left( e^{\frac{2 b_{y_j}}{\lambda_j} u_{y_j} y_j} - 1 \right) \ell_j (dz)} < \infty. \]

The exponential integrability conditions in Cases 5.1 and 5.2 imply \( \int_{0+}^{\infty} z \ell_j (dz) < \infty \). In Case 5.3 we have \( b = b_3 = 0 \) and hence

\[ \int_0^T \mathbb{E} \left[ \int_{0+}^{\infty} k z \ell_j (dz) \right] du \leq kT \int_{0+}^{\infty} z \ell_j (dz), \]

which is finite by assumption. This proves the lemma. \( \square \)

We now prove that \( g(t, y) \) is a (classical) solution to the related forward problem of (5.1).

**Proposition 5.5.** Assume there exists \( \varepsilon > 0 \) such that Condition 2.1 is satisfied with \( c_j = 2 \gamma b_{\gamma_j} \) for \( j = 1, \ldots, m \), where \( b = b_1 \) in Case 5.1 and \( b = b_2 \) in Case 5.2. Then \( g(t, \cdot) \) belongs to the domain of the infinitesimal generator of \( Y \) and

\[
\frac{\partial g}{\partial t} (t, y) = \gamma \Pi(\sigma) g(t, y) - \sum_{j=1}^{m} \lambda_j y_j \frac{\partial g}{\partial y_j} (t, y) \\
+ \sum_{j=1}^{m} \lambda_j \int_{0+}^{\infty} (g(t, y + z \cdot e_j) - g(t, y)) \ell_j (dz),
\]

for \((t, y) \in (0, T] \times [0, \infty)^m\). Moreover, \( \partial g(t, y)/\partial t \) is continuous, so that \( g \in C^{1,1}((0, T] \times [0, \infty)^m) \).

**Proof.** First observe that the conditions in Lemmas 5.3 and 5.7 are fulfilled. The first two terms on the right-hand side of (5.8) are continuous since \( \Pi \) is continuous and \( g(t, \cdot) \in C^1 \) by Lemma 5.3 for all \( t \in [0, T] \). The integral operator also defines a continuous function in time and space. This follows from the integrability conditions on the Lévy measures \( \ell_j (dz) \) and Theorem 2.27 in Folland (1984) together with arguments along the lines of the proofs of Lemmas 5.3 and 5.7. Thus, if \( g \) solves (5.8) then \( \partial g(t, y)/\partial t \) is continuous for \((t, y) \in (0, T) \times [0, \infty)^m\), and may be continuously extended to \( t = T \). Hence, \( g \in C^{1,1}((0, T] \times [0, \infty)^m) \).

From Lemma 5.3 we have that \( y \mapsto g(t, y) \) is a continuously differentiable map. Hence, we know from Itô’s lemma (see, e.g., Ikeda and Watanabe 1989) that the mapping \( s \mapsto g(t, Y(s)) \) is a (local) semimartingale with dynamics
\[ g(t, Y(s)) = g(t, y) - \sum_{j=1}^{m} \lambda_j \int_{0}^{s} Y_j(u) \frac{\partial g}{\partial y_j}(t, Y(u)) \, du \]

\[ + \sum_{j=1}^{m} \int_{0}^{s} \int_{0+}^{\infty} (g(t, Y(u-) + z \cdot e_j) - g(t, Y(u-))) N_j(\lambda_j \, du, dz), \]

where \( N_j(\lambda_j \, du, dz) \) is the Poisson random measure in the Lévy-Kintchine representation of \( Z_j(\lambda_j u) \). From Lemma 5.7 we know that

\[ \mathbb{E} \left[ \int_{0}^{T} \int_{0+}^{\infty} |g(t, Y(u) + z \cdot e_j) - g(t, Y(u))| \ell_j (dz) \, du \right] < \infty, \]

and thus \( g(t, Y(u) + z \cdot e_j) - g(t, Y(u)) \in \mathbf{L}^{1} \) (see Ikeda and Watanabe 1989, pp. 61–62, for this notation). This implies that \( g(t, Y(s)) \) is a semimartingale (and not only local semimartingale). Taking expectations on both sides and rearranging terms give

\[
\frac{\mathbb{E}[g(t, Y(s))] - g(t, y)}{s} = -\sum_{j=1}^{m} \lambda_j \int_{0}^{s} \mathbb{E} \left[ \frac{\partial g}{\partial y_j}(t, Y(u)) \right] \, du \]

\[ + \sum_{j=1}^{m} \lambda_j \int_{0}^{s} \int_{0+}^{\infty} \mathbb{E} [g(t, Y(u) + z \cdot e_j) - g(t, Y(u))] \ell_j (dz) \, du. \]

Therefore, by letting \( s \downarrow 0 \) we get that \( g(t, \cdot) \) is in the domain of the infinitesimal generator of \( Y \), which is denoted by \( G \), and

\[ Gg(t, y) = -\sum_{j=1}^{m} \lambda_j y_j \frac{\partial g}{\partial y_j}(t, y) + \sum_{j=1}^{m} \lambda_j \int_{0+}^{\infty} (g(t, y + z \cdot e_j) - g(t, y)) \ell_j (dz). \]

Since \( g(t, Y(s)) \in L^{1}(\Omega, P) \) for all \( s > 0 \) in a neighborhood of zero, we can calculate

\[
\mathbb{E} \left[ g(t, Y(s)) \right] = \mathbb{E} \left[ \mathbb{E} \left[ e_{h_0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du \right] \left| \mathcal{F}_{t} \right] \right] \]

\[ = \mathbb{E} \left[ e_{h_0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du \right] \]

\[ = \mathbb{E} \left[ e_{0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du \cdot e^{-f_0^{\gamma} \gamma \Pi(\sigma^{\gamma}(u)) du} \right], \]

where we have used the Markov property of \( Y \) together with the law of double expectation. Hence,

\[
\frac{\mathbb{E}[g(t, Y(s))] - g(t, y)}{s} = \frac{1}{s} \left[ e_{0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du e^{-f_0^{\gamma} \gamma \Pi(\sigma^{\gamma}(u)) du} - e_{h_0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du \right]
\]

\[ = \frac{1}{s} \left[ e_{0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du e^{-f_0^{\gamma} \gamma \Pi(\sigma^{\gamma}(u)) du} - e_{h_0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du \right] \]

\[ + \frac{1}{s} \left[ \mathbb{E} \left[ e_{h_0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du \right] - \mathbb{E} \left[ e_{0}^{\ell_{\gamma} \gamma \Pi(\sigma^{\gamma}(u))} du \right] \right] \]

\[ + \frac{1}{s} \left\{ g(t + s, y) - g(t, y) \right\}. \]
By the fundamental theorem of calculus we have that
\[
e^{\int_0^t \gamma \Pi(\sigma^*(u)) \, du} \frac{1}{s} \left( e^{-\int_0^t \gamma \Pi(\sigma^*(u)) \, du} - 1 \right) \to -\gamma \Pi(\sigma) e^{\int_0^t \gamma \Pi(\sigma^*(u)) \, du} \quad \text{as } s \downarrow 0.
\]
In order to show that limit and integration commute, define the function
\[
f(s) = e^{-\int_0^t \gamma \Pi(\sigma^*(u)) \, du}.
\]
The mean value theorem gives
\[
\frac{1}{s} |(f(s) - f(0))| \leq \frac{1}{s} \sup_{s \in [0, T]} |f'(s)|s = \sup_{s \in [0, T]} |\gamma \Pi(\sigma^*(s)) e^{-\int_0^s \gamma \Pi(\sigma^*(u)) \, du}| \leq \gamma e^{\int_0^t (a + b \sigma^*(u)) \, du} (a + b \sup_{s \in [0, T]} \sigma^*(s)).
\]
In the last estimation we have used the linear growth of \( \Pi \). The constant \( b \) is \( b_1 \) in Case 5.1, \( b_2 \) in Case 5.2, and \( b_3 = 0 \) in Case 5.3. Since each \( Z_j \) is a nondecreasing process,
\[
\sup_{s \in [0, T]} \sigma^*(s) \leq \sigma + \sum_{j=1}^m \omega_j Z_j(\lambda_j T).
\]
This implies
\[
e^{\int_0^t \gamma \Pi(\sigma^*(u)) \, du} \frac{1}{s} \left( e^{-\int_0^t \gamma \Pi(\sigma^*(u)) \, du} - 1 \right) \leq k \sum_{j=1}^m \omega_j e^{2\gamma \int_0^t (a + b \sigma^*(u)) \, du} Z_j(\lambda_j T),
\]
for some positive constant \( k \). But from (2.4)
\[
\sum_{j=1}^m \omega_j \mathbb{E} \left[ e^{2\gamma \int_0^t (a + b \sigma^*(u)) \, du} Z_j(\lambda_j T) \right] \leq k \sum_{j=1}^m \omega_j \mathbb{E} \left[ e^{\left( \frac{2u_j b_j}{\gamma} \right) Z_j(\lambda_j T)} \right]
= k \sum_{j=1}^m \omega_j e^{\lambda_j \int_0^\infty \left( e^{\frac{2u_j b_j}{\gamma}} - 1 \right) \xi_j (dz)},
\]
where \( k \) is some positive constant (different than above). In our estimation, we have used that there exists a positive constant \( k_\epsilon \) such that \( \epsilon \leq k_\epsilon e^{\epsilon z} \) for all \( z \geq 0 \). The last sum is finite by our integrability assumption. Hence by dominated convergence (see Theorem 2.27(a) in Folland 1984) \( \partial g/\partial t \) exists and
\[
G g(t, y) = -\gamma \Pi(\sigma) g(t, y) + \frac{\partial g(t, y)}{\partial t}.
\]
This concludes the proof of the proposition. \( \square \)

Let \( h(t, y) = g(T - t, y) \); that is,
\[
h(t, y) = \mathbb{E}^y \left[ e^{\int_0^{T-t} \gamma \Pi(\sigma(s)) \, ds} \right].
\]
We can represent this equivalently as
\[
(5.9) \quad h(t, y) = \mathbb{E}^{t,y} \left[ e^{\int_t^{T-t} \gamma \Pi(\sigma(s)) \, ds} \right],
\]
by using the time-homogeneity of \( Y \). Hence our explicit solution candidate is
\[
(5.10) \quad v(t, w, y) = \gamma^{-1} w^y h(t, y).
\]
The candidate for the optimal feedback control $\pi^*(\sigma)$ is given in (5.3), (5.4), or (5.5), depending on the size of $\left(\frac{1}{2} + \beta\right)/(1 - \gamma)$. In the next section we prove that (5.10) coincides with the value function (2.5).

5.3. Explicit Solution of the Control Problem

We apply the verification theorem to connect our explicit solution to the value function of the control problem. But first we need the integrability results stated in the following two lemmas.

**Lemma 5.6.** Assume Condition 2.1 holds with $c_j = 8\gamma \left(\left|\frac{1}{2} + \beta\right| + 4\gamma\right)\frac{1}{\gamma_j}$ for $j = 1, \ldots, m$. Then

$$
\int_0^T \mathbb{E}\left[ (\pi(u))^2 (W^\pi(u))^2 \sigma(u) (W^\pi(u))^{2(\gamma - 1)} h(u, Y(u)) \right] du < \infty, \quad \forall \pi \in \mathcal{A}_0.
$$

**Proof.** Observe that the function $h$ has the same growth as $g$. Hence, by Lemma 5.2 and $\pi \in [0, 1]$,

$$
\int_0^T \mathbb{E}\left[ (\pi(u))^2 (W^\pi(u))^2 \sigma(u) (W^\pi(u))^{2(\gamma - 1)} h(u, Y(u)) \right] du
\leq \int_0^T \mathbb{E}\left[ (W^\pi(u))^{2\gamma} \sigma(u) e^{k_1 + \gamma b \sum_{j=1}^m \frac{1}{\gamma_j} Y_j(u)} \right] du
\leq k_2 e^{k_3 T} \int_0^T \mathbb{E}\left[ (W^\pi(u))^{2\gamma} e^{(\gamma b + \varepsilon) \sum_{j=1}^m \frac{1}{\gamma_j} Y_j(u)} \right] du,
$$

where $k_2$ is a positive constant such that $\sigma \leq k_2 e^{\sum_{j=1}^m \frac{1}{\gamma_j} Y_j}$. Hölder's inequality gives

$$
\int_0^T \mathbb{E}\left[ (\pi(u))^2 (W^\pi(u))^2 \sigma(u) (W^\pi(u))^{2(\gamma - 1)} h(u, Y(u)) \right] du
\leq k_2 \int_0^T \prod_{j=1}^m \mathbb{E}\left[ e^{2(\gamma b + \varepsilon) \frac{1}{\gamma_j} Y_j(u)} \right]^{1/2} \mathbb{E}\left[ (W^\pi(u))^{4\gamma} \right]^{1/2} du
\leq k_2 \int_0^T \prod_{j=1}^m \mathbb{E}\left[ e^{2(\gamma b + \varepsilon) + \varepsilon} Y_j(u) \right]^{1/2} \mathbb{E}\left[ (W^\pi(u))^{4\gamma} \right]^{1/2} du,
$$

where $b$ is $b_1$ in Case 5.1, $b_2$ in Case 5.2, and $b_3 = 0$ in Case 5.3. In the last estimation we have redefined $\varepsilon$ to get a more tractable integrability condition.

We now argue that the two expectations are finite. Note that we are free to choose $\varepsilon$ as long as it is positive. Let $\xi_j = 2\gamma b \omega_j / \lambda_j + \varepsilon$. In Case 5.1, we have $\xi_j = 2\gamma \left(\frac{1}{2} + \beta\right) + \varepsilon < c_j$ for a suitably chosen $\varepsilon$, where $c_j$ is defined in the lemma. In Case 5.2, $\xi_j = 2\gamma \left(\frac{1}{2} + \beta\right) + \varepsilon$, and since we have $\frac{1}{2} + \beta < 1$, $\xi_j < \gamma \left(\frac{1}{2} + \beta\right) \frac{1}{\lambda_j} + \varepsilon < c_j$ if we choose $\varepsilon$ small enough. In Case 5.3, we obviously have $\xi_j = \varepsilon < c_j$ when choosing $\varepsilon$ smaller than $c_j$. Thus, the integrability condition in Lemma 3.2 holds and the terms involving the expectation of $Y_j(u)$ above are finite. Finally, invoking Lemma 3.3 yields the desired result.  \[\square\]
Lemma 5.7. Assume Condition 2.1 holds with \( c_j = 8\gamma \left( \left| \frac{1}{2} + \beta \right| + 4\gamma \right) \frac{a_j}{\lambda_j} \) for \( j = 1, \ldots, m \). Then
\[
\int_0^T \int_{0+}^{\infty} \mathbb{E}[(W^*(s))' | h(u, Y(u) + z - h(u, Y(u))] \ell_j(dz) \, du < \infty, \quad j = 1, \ldots, m.
\]

Proof. We follow the arguments in the proof of Lemma 5.7:
\[
\int_0^T \int_{0+}^{\infty} \mathbb{E}[(W^*(u))' | h(u, Y(u) + z) - h(u, Y(u))] \ell_j(dz) \, du
\leq \int_0^T \mathbb{E} \left[(W^*(u))' \mathbb{E}^{\gamma \sum_{j=1}^m \frac{a_j}{\lambda_j} h(y + 2Z_j(\lambda_j, u)) \ell_j(dz) \right] \, du
\leq k_1 e^{k_2 \sigma} \int_{0+}^{\infty} \ell_j(dz) \int_0^T \mathbb{E} \left[(W^*(u))^\gamma e^{2\gamma \sum_{j=1}^m \frac{a_j}{\lambda_j} Z_j(\lambda_j, u)} \right] \, du,
\]
where \( k_1, k_2 \) are positive constants. Using Hölder’s inequality with \( p = 4 \) and \( q = 4/3 \) gives
\[
\int_0^T \int_{0+}^{\infty} \mathbb{E}[(W^*(u))' | h(u, Y(u) + z) - h(u, Y(u))] \ell_j(dz) \, du
\leq k_1 e^{k_2 \sigma} \int_{0+}^{\infty} \ell_j(dz) \int_0^T \mathbb{E}[(W^*(u))^{\gamma/4}]^{1/4} \left[ e^{\gamma \sum_{j=1}^m \frac{a_j}{\lambda_j} Z_j(\lambda_j, u)} \right]^{3/4} \, du
\leq k_1 e^{k_2 \sigma} \int_{0+}^{\infty} \ell_j(dz) \prod_{j=1}^m \mathbb{E} \left[e^{\gamma \sum_{j=1}^m \frac{a_j}{\lambda_j} Z_j(\lambda_j, T)} \right]^{1/4} \int_0^T \mathbb{E}[(W^*(u))^{\gamma}]^{1/4} \, du
\leq k_1 e^{k_2 \sigma} \int_{0+}^{\infty} \ell_j(dz) e^{\gamma \sum_{j=1}^m \frac{a_j}{\lambda_j}} (\gamma \sum_{j=1}^m \frac{a_j}{\lambda_j} - 1) \ell_j(dz) \int_0^T \mathbb{E}[(W^*(u))^{\gamma}]^{1/4} \, du.
\]
Let \( \xi_j = 8\gamma b_j \frac{a_j}{\lambda_j} \). In Case 5.1, \( \xi_j = 8\gamma (\frac{1}{2} + \beta) \frac{a_j}{\lambda_j} \), which obviously is less than the \( c_j \) given in the lemma. In Case 5.2, we have \( \xi_j < 8\gamma (\frac{1}{2} + \beta) \frac{a_j}{\lambda_j} < c_j \). Finally, in Case 5.3, \( \xi_j = 0 < c_j \). Hence, the integrability condition in the lemma implies that \( \int_{0+}^{\infty} (\gamma \sum_{j=1}^m \frac{a_j}{\lambda_j} - 1) \ell_j(dz) < \infty \) for \( j = 1, \ldots, m \). From Lemma 3.3 the desired result follows. \( \square \)

We sum up our results in this section in the following theorem.

Theorem 5.8. Assume Condition 2.1 holds with \( c_j = 8\gamma \left( \left| \frac{1}{2} + \beta \right| + 4\gamma \right) \frac{a_j}{\lambda_j} \) for \( j = 1, \ldots, m \). Then the value function of the control problem is
\[
V(t, w, y) = \gamma^{-1} w^y h(t, y),
\]
where \( h \) is defined in (5.9). Furthermore, the optimal investment strategy is \( \pi^*(\sigma) = 1 \) in Case 5.1,
\[
\pi^*(\sigma) = \begin{cases} 1, & \sigma \in [0, \hat{\sigma}_1), \\ \pi(\sigma), & \sigma \in [\hat{\sigma}_1, \infty) \end{cases}
\]
in Case 5.2, and

\[ \pi^*(\sigma) = \begin{cases} 
1, & \sigma \in [0, \hat{\sigma}_1), \\
\bar{\pi}(\sigma), & \sigma \in [\hat{\sigma}_1, \hat{\sigma}_0], \\
0, & \sigma \in (\hat{\sigma}_0, \infty). 
\end{cases} \]

in Case 5.3. The function \( \bar{\pi}(\sigma) \) is defined as

\[ \bar{\pi}(\sigma) = \frac{1}{1 - \gamma} \left( \frac{\mu - r}{\sigma} + \frac{1}{2} + \beta \right) \]

and

\[ \hat{\sigma}_1 = \frac{\mu - r}{(1 - \gamma) - \left( \frac{1}{2} + \beta \right)}, \quad \hat{\sigma}_0 = -\frac{\mu - r}{\frac{1}{2} + \beta}. \]

**Proof.** Condition 3.1 holds under our assumption. Moreover, the integrability condition implies by Lemma 5.2 that \( h \) is well-defined. Observe that in all three cases \( \pi^* \) depends only on \( \sigma \) and not on the wealth \( w \). Together with the fact that \( \pi^* \in [0, 1] \), there obviously exists a unique solution \( W^{\pi^*} \) to (2.3), being a geometric Brownian motion with stochastic coefficients given by \( \sigma(s) \). Hence, \( \pi^* \in \mathcal{A}_t \).

Let \( v(t, w, y) = \gamma^{-1} w^y h(t, y) \), and observe that the condition in Lemmas 5.6 and 5.7 hold. Moreover, we claim that Proposition 5.5 holds true. Let \( \xi_j = 2\gamma b \omega \lambda_j + \epsilon \), where \( \epsilon \) is any positive number which we are free to choose. Going through all the three cases for the constant \( b \), we see that \( \xi_j < c_j \), where \( c_j \) is given in the theorem for \( \epsilon \) appropriately chosen. Hence, the integrability condition assumed in the theorem is stronger than the required integrability in Proposition 5.5. We also see that the integrability conditions in Lemma 5.3 hold, which implies \( v \in C([0, T] \times \overline{D}) \) since \( v \) obviously is continuous in \( w \) on \([0, \infty)\). Therefore \( v \) is a classical solution of the HJB equation (2.7) and we can apply the verification theorem (Theorem 5.5) to conclude the proof. \( \square \)

**Remark 5.2.** The reader should note that the above arguments are valid also for \( \mu < r \). However, the Cases 5.1, 5.2, and 5.3 will be slightly different and the optimal solution must be changed accordingly. Also observe that by letting \( \lambda = 0, \beta = 0 \) and \( m = 1 \) we get back the classical Merton solution

\[ \pi^* = \frac{\mu + \frac{1}{2} \sigma^2 - r}{(1 - \gamma)\sigma^2} \]

with \( \sigma \equiv \sqrt{\sigma(t)} \) as the constant volatility (\( t \) is the starting time).

**Remark 5.3.** Note that \( \pi^* \) is exactly like in traditional Merton, except that we now react on changes in \( \sigma \)—that is, \( \pi^* = \pi^*(\sigma(t)) \), where \( \sigma(t) \) is the underlying volatility. It is still optimal to choose a fraction inversely proportional to volatility, however now it varies with the changing volatility rather than being fixed, as is the assumption in the classical Merton case. In some sense this is how investors following the Merton optimal strategy behave. At every instant of rebalancing of the portfolio, they will calculate the current volatility and invest according to the Merton optimal strategy. But then they have
in effect invested according to a strategy with changing volatility, and not according to a strategy where the level of volatility is fixed from the start.

In Case 5.1, it is possible to calculate the value function explicitly. This is the content of the following proposition.

**Proposition 5.9.** Assume Condition 2.1 holds with \( c_j = 8 \gamma \left( \frac{1}{2} + \beta \right) + 4 \gamma \frac{m_j}{y} \) for \( j = 1, \ldots, m \). Then the solution \( h \in C^{1,1} ([0, T] \times [0, \infty)^m) \) of the first-order integro-differential equation

\[
 h_t(t, y) + \gamma h(t, y) - \sum_{j=1}^{m} \lambda_j y_j h_{y_j}(t, y) + \sum_{j=1}^{m} \lambda_j \int_0^\infty (h(t, y + z \cdot e_j) - h(t, y)) \ell_j (dz) = 0,
\]

with terminal condition \( h(T, y) = 1 \forall y \in [0, \infty)^m \), is

\[
 h(t, y) = \exp \left( \gamma \mu (T - t) + \gamma \left( \frac{y}{2} + \beta \right) \sum_{j=1}^{m} \frac{\omega_j}{\lambda_j} y_j (1 - e^{-\lambda_j (T - t)}) \right)
 + \lambda_j \int_t^T \int_{0^+}^\infty \left( e^{(\bar{z} + \beta) \frac{\omega_j}{\lambda_j} (1 - e^{-\lambda_j s})} - 1 \right) \ell_j (dz) ds.
\]

Furthermore, the value function (2.5) is explicitly given by

\[
 (5.12) \quad V(t, w, y) = \gamma^{-1} \omega \gamma h(t, y),
\]

where \( h \) is given in (5.11).

**Proof.** Let us calculate the function \( g(t, y) \) defined in (5.6):

\[
 g(t, y) = E \left[ e^{\int_0^t \gamma \Pi(\sigma(s)) ds} \right] = e^{\nu \mu t} E \left[ e^{\Psi(\tau + \beta) \int_0^t \omega(\sigma(s)) ds} \right]
 = e^{\nu \mu t + \gamma \Psi(\tau + \beta) \sum_{j=1}^{m} \omega_j y_j \int_0^t e^{-\lambda_j s} ds} E \left[ e^{\Psi(\tau + \beta) \sum_{j=1}^{m} \int_0^t e^{-\lambda_j s} \omega_j e^{\omega_j u} dZ_j(u) ds} \right]
 = e^{\nu \mu t + \gamma \Psi(\tau + \beta) \sum_{j=1}^{m} \omega_j y_j (1 - e^{-\lambda_j t})} \prod_{j=1}^{m} \left[ e^{\Psi(\tau + \beta) \omega_j y_j \int_0^t (1 - e^{-\lambda_j u}) dZ_j(u) ds} \right].
\]

The Fubini theorem yields

\[
 \int_0^t \int_0^s e^{-\lambda_j (t - u)} dZ_j(\lambda_j u) ds = \frac{1}{\lambda_j} \int_0^t \left( 1 - e^{-\lambda_j (t - s)} \right) dZ_j(\lambda_j s),
\]

so that

\[
 g(t, y) = e^{\nu \mu t + \gamma \Psi(\tau + \beta) \sum_{j=1}^{m} \omega_j y_j (1 - e^{-\lambda_j t})} \prod_{j=1}^{m} \left[ e^{\Psi(\tau + \beta) \omega_j y_j (1 - e^{-\lambda_j t})} \right]
 = \exp \left( \gamma \mu t + \gamma \left( \frac{y}{2} + \beta \right) \sum_{j=1}^{m} \frac{\omega_j}{\lambda_j} y_j (1 - e^{-\lambda_j t}) \right)
 + \lambda_j \int_t^T \int_{0^+}^\infty \left( e^{\Psi(\tau + \beta) \omega_j y_j (1 - e^{-\lambda_j s})} - 1 \right) \ell_j (dz) ds.
\]
\[
= \exp \left( \gamma \mu t + \gamma \left( \frac{\nu}{2} + \beta \right) \sum_{j=1}^{m} \frac{\omega_j}{\lambda_j} y_j (1 - e^{-\lambda_j t}) \right) \\
+ \lambda_j \int_0^t \int_{0+} \left( e^{\gamma (\xi + \beta) / \nu} \frac{1}{\nu} (1 - e^{-\lambda_j s}) z - 1 \right) \ell_j (dz) ds .
\]

Hence, by recalling that \( h(t, y) = g(T - t, y) \) and after a change of variables in the \( ds \)-integration, we recover (5.11). □

**Remark 5.4.** In Case 5.1, we could have first calculated the function (5.12), and then afterward verified that this coincides with the value function of our control problem. This approach would have given us weaker conditions on exponential integrability of the Lévy measures. In Cases 5.2 and 5.3, explicit results seem to be impossible to obtain due to the much more complicated structure of \( \Pi(\sigma) \). Indeed, the dependency on the level and inverse of \( \sigma(s) \) makes it a difficult task to calculate the expectation.

### 6. Discussion

Almost all of our results are based on sufficient exponential integrability of the Lévy measures. The question arises, how are these conditions related to the models that we would like to use for stochastic volatility dynamics? To be able to discuss this question, we need to give a brief description of the modeling approaches suggested by Barndorff-Nielsen and Shephard (2001). Two ways of finding a stochastic volatility model are suggested. In their first (and main) approach, they start out with the stationary distribution (on the positive axis) of the volatility and from this derive the Lévy processes \( Z_j \), where \( Z_j \) is coined the background driving Lévy process. To achieve this, the stationary distribution must be chosen from among the so-called self-decomposable distributions. Some examples of such distributions are given in Barndorff-Nielsen and Shephard. In the class of generalized inverse Gaussian distributions, the authors calculate the upper tail integral of the Lévy densities (i.e., the density of the Lévy measure) for the background driving Lévy process in several cases, thereby in effect giving the tail behavior of the Lévy measure. Formulas for the inverse Gaussian, positive hyperbolic, reciprocal gamma, and gamma distributions are given, all showing an exponential damping in the tails. However, the rate of damping is given by (some combination of) the parameters in the respective distributions which lead to restrictions on the choices of risk aversion \( 1 - \gamma \) and skewness \( \beta \) when applied in our control problem. The choice of stationary distribution for the volatility is based on empirical decisions. For example, Barndorff-Nielsen and Shephard note that if we choose \( \sigma_j(t) \) to have an inverse Gaussian distribution, then the log returns will be approximately normal inverse Gaussian distributed, a distribution that models semi-heavy tails observed in market data very well (see Barndorff-Nielsen [1998] for more on this class of distributions and, e.g., Eberlein and Keller [1995], Rydberg [1997], and Bølviken and Benth [2000] for applications to empirical finance). Their statistical studies of deutsche mark–dollar exchange rate (data on 5 minutes periods over 10 years) suggest using inverse Gaussian marginals for the volatility. Additionally, one has to use a superposition of Ornstein-Uhlenbeck processes to correctly model the dependency in the log returns. The analysis shows a good fit for the autocorrelation structure when \( m = 4 \), a superposition of four non-Gaussian Ornstein-Uhlenbeck processes.
A second way of introducing stochastic volatility dynamics through non-Gaussian Ornstein-Uhlenbeck models is to directly model the Lévy process, as also suggested by the authors. Even more, one may model the Lévy process by directly specifying the Lévy measure. This approach may give more room for the parameters in the control problem. In both of the suggested approaches, we see there will be a competition between the parameters of the control problem and the parameters in the model of the risky asset. But these constraints need not be too binding, as can be seen from the examples below. We would finally like to remark that our conditions on the exponential integrability of the Lévy measures may be weakened by going through the estimates more carefully. This task will not be pursued in any further detail here.

7. NUMERICAL EXAMPLES

In this section, we present some numerical examples. We focus on the variability of the volatility and how an investor (with complete knowledge of the present level of volatility) should optimally diversify her portfolio. Our purpose is to demonstrate some effects incurred by a varying volatility. In order to simplify matters, we choose to model the stochastic volatility by one non-Gaussian Ornstein-Uhlenbeck process; thus, we set \( m = 1 \) and have

\[
\sigma(s) = e^{-\lambda s} \sigma(0) + e^{-\lambda s} \int_0^s e^{\lambda u} dZ(u),
\]

when we start the process at \( t = 0 \) (which we indeed shall do in the examples below). Furthermore, because we want to simulate numerically the volatility it is convenient to choose the stationary probability distribution for \( \sigma(t) \) to be in the class of Gamma distributions, \( \sigma(t) \sim \Gamma(\nu, \alpha) \). The Gamma distribution is a member of the family of generalized inverse Gaussian distributions and has a density

\[
\frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} \exp(-\alpha x), \quad x > 0.
\]

As is calculated in Barndorff-Nielsen and Shephard (2001), the background driving Lévy process \( Z \) will then have a Lévy density

\[
\ell(dz) = \nu \alpha \exp(-\alpha z) dz.
\]

As will be explained below, this density makes it particularly easy to simulate \( \sigma(t) \). In the examples below we choose \( \alpha = 10, \lambda = 0.01, \) and \( \nu = 3 \).

We describe the procedure we use to simulate paths of \( \sigma(s) \). The suggested algorithm was introduced by Marcus (1987) and Rosinski (1991), and explained in our context of stochastic volatility in Barndorff-Nielsen and Shephard (2001). We adopt here the notation in Barndorff-Nielsen and Shephard. Assume we discretize the time line \([0, T]\) by homogeneous time intervals of length \( \Delta > 0 \). Then a straightforward calculation shows

\[
\sigma(s + \Delta) = e^{-\lambda \Delta} \sigma(s) + e^{-\lambda \Delta} z(s),
\]

where

\[
z(s) = e^{-\lambda s} \int_s^{s+\Delta} e^{\lambda u} dZ(u).
\]
and $s$ is a time point in our discretization of $[0, T]$. Note that $z(s)$ is independent of $z(t)$ when $t \neq s$. By a change of variables, we find that

$$z(s) = \int_0^\lambda \Delta e^u dZ(u),$$

where equality is in distribution. The integral $z_s$ can be represented as an infinite series suitable for simulation: Let $\{r_i\}$ be independent samples from a uniform probability distribution on $[0, 1]$ and $a_1 < a_2 < \cdots < a_i < \cdots$ be the arrival times of a Poisson process with intensity 1. Then (in distribution)

$$z(s) = \sum_{i=1}^\infty W^{-1}(a_i) e^{\lambda r_i \Delta}.$$

In the above expansion, the function $W^{-1}(x)$ appears, which is the inverse of $W^+(x)$, where $W^+$ is the upper tail integral of the Lévy density of $Z$. For the Gamma distribution this is explicitly invertible:

$$W^{-1}(x) = \max \left(0, -\frac{1}{\alpha} \ln \left(\frac{x}{\nu}\right)\right).$$

Introducing this function in the series expansion, Barndorff-Nielsen and Shephard (2001) suggest to simulate $z(s)$ from the representation (in law)

$$z(s) = \frac{1}{\alpha} \sum_{i=1}^{N(\nu)} \ln \left(\frac{\nu}{a_i}\right) e^{\lambda \Delta r_i},$$

where $N(\nu)$ is the number of arrivals $a_i$ before time $\nu$. The inversion of the upper tail integral $W^+$ is in general not analytically possible, thus leading to more complicated simulation algorithms. The Gamma distribution is the only case where we can invert $W^+$ among the examples of distributions suggested by Barndorff-Nielsen and Shephard, which is the reason we use this distribution for our numerical investigations. In all the simulations below, $\Delta = 1$.

We present two examples: one where the log returns are symmetric and one with skew log returns. The purpose of the numerical examples is to show how optimal allocation taking stochastic volatility into account may dramatically deviate from the classical Merton investment strategy. We choose as $\sigma(0)$ the historical volatility—that is, the volatility a Merton investor would choose to pin down her strategy at time 0 and follow until time $T$. We choose $\sigma(0) = 0.25$. The investment horizon is assumed to be $T = 5000$.

The optimal strategy in the stochastic volatility case is found by simulating one path of the volatility using the method described above, and then calculating the optimal strategy using the rules in Theorem 5.8. Note that the parameters we choose in the examples below satisfy the integrability condition in Theorem 5.8. The numerical algorithms were implemented in MATLAB, and simulation of one path of 5000 points took about 4–5 sec.

**Example 7.1.** Let $\beta = 0$, the risk aversion $1 - \gamma = \frac{7}{8}$ (i.e., the investor is very risk averse) and $\mu - r = \frac{5}{32}$. An investor following the classical Merton strategy would choose $\pi_M = 1$ and invest her whole wealth in the risky asset. Since $\frac{1}{4} + \beta / 1 - \gamma = \frac{4}{7} \in (0, 1)$, 

---

1 Note that we follow the notation of Barndorff-Nielsen and Shephard (2001) here. The function $W^{-1}$ is not to be confused with the (inverse of the) wealth dynamics in the present paper.
we are in Case 5.2 where $\pi^*$ is not necessarily equal to one. Figure 7.1 shows one possible scenario, where we observe that in fact the optimal strategy may be to invest below 100% in the risky asset. Indeed, we see that in periods of times we should go down to about 80%, significantly more conservative than putting all the money in the risky asset. However, this happens in very short periods, compared to the long periods where the investor is advised to place 100% of her total wealth in the risky asset.

**EXAMPLE 7.2.** Again we choose parameters such that the classical Merton investor puts all of her money in the risky asset. Let $\beta = -\frac{2}{3}$ and $1 - \gamma = \frac{1}{2}$; the log returns are skewed to the left and the investor is moderately risk averse. Furthermore, we assume $\mu - r = \frac{1}{10}$. The Merton investor choose $\pi_M = 1$. Since $\frac{1}{2} + \beta = -\frac{1}{6} < 0$, we are in Case 5.3, where the optimal strategy may be to choose $\pi^* = 0$ in periods of very high volatility. And indeed this may happen, as is seen in Figure 7.2. Note also that periods where the investor should allocate less than 100% in the risky assets are dominating.

We conclude from these two examples that varying volatility both with and without skew log returns may lead to significantly different optimal investment behavior. Note that we have used different seeds in the simulation of the stochastic volatility in the two examples, thus the paths are different. The average of $\sigma(s)$ over the paths is about 0.25 in each example.

**REMARK 7.1.** Note that the Gamma models we chose for these two numerical examples is not necessarily market relevant. The parameters are not estimated from empirically
observed prices, but admittedly were chosen to highlight some deviations from the classical Merton case. We do believe, however, that similar observations can be made when operating with stochastic volatility models that are statistically fitted to observed price data. Since the procedure of estimating the parameters in the stochastic volatility model is rather involved, we leave such considerations to future research.

REFERENCES


