Les Cahiers du GERAD ISSN: 0711–2440

Optimal Hedging in Discrete and Continuous Time

B. Rémillard
S. Rubenthaler

G–2009–77

November 2009

Optimal Hedging in Discrete and Continuous Time

Bruno Rémillard

GERAD & Service de l'enseignement des méthodes quantitatives de gestion
HEC Montréal
3000, chemin de la Côte Sainte-Catherine
Montréal (Québec) Canada, H3T 2A7
bruno.remillard@hec.ca

Sylvain Rubenthaler

Laboratoire J.-A. Dieudonné
Université de Nice-Sophia Antipolis
Parc Valrose
06108 Nice Cedex 2, France
rubentha@unice.fr

November 2009

Les Cahiers du GERAD

G–2009–77

Abstract

In this article we find the optimal solution of the hedging problem in discrete time by minimizing the mean square hedging error, when the underlying assets are multidimensional, extending the results of Schweizer (1995). We also find explicit expressions for the optimal hedging problem in continuous time when the underlying assets are modeled by a regime-switching geometric Lévy process. It is also shown that the continuous time solution can be approximated by discrete time Hidden Markov models processes. In addition, in the case of the regime-switching geometric Brownian motion, the optimal prices are the same as the prices under an equivalent martingale measure, making that measure a natural choice. However, the optimal hedging strategy is not the usual delta hedging but it can be easily computed by Monte Carlo methods.

Key Words: Hedging, option pricing, regime-switching, Lévy processes.

Résumé

Dans cet article, nous trouvons la solution optimale du problème du portefeuille de réplication en temps discret lorsque l'on désire minimiser l'erreur quadratique moyenne entre la valeur finale du portefeuille et la valeur intrinsèque d'une option, lorsqu'il y a plusieurs actifs sous-jacents. Nous généralisons ainsi les résultats obtenus par Schweizer (1995) qui a traité le cas d'un seul actif, sous des conditions plus fortes. Nous trouvons aussi la solution en temps continu de ce problème de réplication lorsque les actifs sous-jacents sont des processus de Lévy avec changement de régimes. On montre aussi que la solution optimale en temps continu peut être approchée par celle en temps discret lorsque les actifs sont modélisés par des chaînes de Markov avec états cachés. De plus, dans le cas de mouvements browniens avec changement de régimes, on montre que les prix obtenus et la stratégie de réplication obtenus par la solution optimale sont les mêmes que ceux obtenus sous une certaine mesure martingale et que contrairement avec la plupart des modèles connus, la stratégie de réplication n'est pas donnée par le gradient des prix par rapport aux sous-jacents; par contre la stratégie optimale peut être obtenue facilement par une méthode Monte Carlo.

Acknowledgments: Partial funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada, by the Fonds Québécois de Recherche sur la Nature et les Technologies, the Institut de Finance Mathématique de Montréal and by the PPF Complexité-Modélisation-Finance de l'Universtité Nice-Sophia Antipolis.

1 Introduction

In many applications one is interested in finding a portfolio which will be traded dynamically at discrete time period so that its value at maturity is as close as possible as a target function of the underlying assets. Of course, it can be interpreted as option pricing and hedging, but sometimes the target function is not a payoff. For example, that kind of problem arises when one tries to replicate hedge funds or create synthetic funds with prescribed law and dependence with a given portfolio. See, for example Papageorgiou et al. (2008).

The hedging problem for one risky asset was first solved by Schweizer (1995), when the error measure is the average quadratic hedging error. He showed that the initial value of the portfolio, which can be interpreted as the "value" of the option, is the average, under the "real probability measure", of the discounted payoff, multiplied by a martingale. However the "price" can be negative since the martingale is not necessarily positive. In the latter case, which is more the norm that the exception, the martingale cannot be used as the density of an equivalent martingale measure. However, the discounted asset price process, multiplied by that martingale is itself a martingale.

Even if that hedging problem as been solved quite generally by Schweizer (1995) in the one-dimensional case, it seems to have been ignored or forgotten, e.g., (Bouchaud and Potters, 2002) or Cornalba et al. (2002). More troubling, delta hedging, based on the Black-Scholes-Merton model, is sometimes used in practice even if it is known that the geometric Brownian motion model is inadequate for the underlying assets (Kat and Palaro, 2005). Furthermore, even when the geometric Brownian motion model is adequate, the hedging error in discrete time is not zero. It converges to zero as the number of hedging periods tends to infinity. That problem is well documented. See, e.g., Boyle and Emanuel (1980), Wilmott (2006)[Chapters 46-47] and references therein.

Motivated by replication applications, Papageorgiou et al. (2008) proposed a locally optimal solution minimizing the average quadratic hedging error at each period, in the general multidimensional asset case. They erroneously claimed that it was globally optimal, which is true only if the discounted underlying assets are martingales. A first motivation for the present paper is to correct that mistake and generalize the results of Schweizer (1995) to the multidimensional case. This in done in the next section. Another motivation is to give a partial answer to the question: What happens when the number of hedging periods tends to infinity, specially when one uses Gaussian Hidden Markov models (HMM). To answer that, we first solve the optimal hedging problem in continuous time for possible limits of general HMM.

Minimizing the average quadratic hedging error in continuous time has received much attention. Unfortunately, most of the time it is assumed that the discounted prices are martingales (Cont and Tankov, 2004) or that the discounted portfolio is a martingale under a class of equivalent martingale measures (Föllmer and Sondermann, 1986). See also Pham (2000). As mentioned in Cont and Tankov (2004), minimizing the hedging error under an equivalent martingale measure is not realistic.

When the market is not complete but there is no arbitrage, there are infinitely many martingale measures. One has then to choose the "best" martingale measure with respect to some utility criterion. There is a huge literature on that subject. One interesting paper is Duan (1995) where the author proposes a choice of the martingale measure when the log-returns are distributed as a GARCH-M process. Unfortunately he also proposed a hedging strategy which has been shown to be wrong by Garcia and Renault (1998). Hence the need to find optimal hedging strategies.

The optimal solution of the discrete time hedging problem is stated in Section 2, extending the results of Schweizer (1995). It is interesting to note that when the price process is Markovian, or a component of a Markov process, then the optimal solution can be implemented using approximation techniques. Such examples include the GARCH-type models and some HMM.

In Section 3 we find the solution of optimal hedging problem in continuous time when the log-returns of the price process follow a regime-switching Lévy process. It is shown that for all but the so-called geometric Brownian motion, the optimal strategy ϕ_t is not given by the so-called Delta and in fact depends on the whole trajectory up to time t. When the discounted price process is a martingale, one recovers the formula established by Cont and Tankov (2004) for one-dimensional Lévy processes. A very interesting case is the regime-switching geometric Brownian motion. It is proven that in that case, the martingale appearing in the pricing formula is indeed positive, thus permitting an equivalent change of measure under which the discounted prices are martingales. Surprisingly, under the change of measure, the Markov chain associated with the regime changes in non homogeneous. That is totally different from the equivalent martingale measure proposed by Guo (2001). However Monte Carlo simulations can be used to find the associated price and compute the optimal hedging solution.

In Section 4, we show that under weak assumptions, a regime-switching geometric random walk converges in law to a regime-switching geometric Lévy process. Moreover, under additional conditions, the associated discrete time optimal strategy converges to the optimal strategy of the limiting continuous time process. In particular, the optimal hedging solution in the regime-switching geometric Brownian motion case can be approximated by a regime-switching geometric Gaussian random walk. All results are proved in a series of appendices. Finally, an example of application involving the regime-switching geometric Brownian motion is given in Section 5.

2 Optimal hedging strategy in discrete time

Denote the price process by S, i.e., S_k is the value of the d underlying assets at period k and let $\mathbb{F} = \{\mathcal{F}_k, k = 0, \ldots, n\}$ be a filtration under which S is adapted. Assume that S is square integrable. Set $\Delta_k = \beta_k S_k - \beta_{k-1} S_{k-1}$, where the discounting factors β_k are predictable, i.e. β_k is $\mathcal{F}_{k-1}$-measurable for $k = 1, \ldots, n$.

The aim of this section is to find an initial investment amount V_0 and a predictable investment strategy $\overrightarrow{\phi} = (\phi_k)_{k=1}^n$ that minimize the expected quadratic hedging error $E\left[\left\{G\left(V_0, \overrightarrow{\phi}\right)\right\}^2\right]$, where

$$G = G\left(V_0, \overrightarrow{\phi}\right) = \beta_n C - V_n,$$

and

$$V_k = V_0 + \sum_{j=1}^{k} \phi_j^\top \Delta_j, \quad k = 0, \ldots, n.$$

Set $P_{n+1} = 1$, and for $k = n, \ldots, 1$, define

$$\begin{aligned} A_k &= E\left(\Delta_k \Delta_k^\top P_{k+1} | \mathcal{F}_{k-1}\right), \\ b_k &= A_k^{-1} E\left(\Delta_k P_{k+1} | \mathcal{F}_{k-1}\right), \\ \alpha_k &= A_k^{-1} E\left(\beta_n C \Delta_k P_{k+1} | \mathcal{F}_{k-1}\right), \\ P_k &= \prod_{j=k}^{n} \left(1 - b_j^\top \Delta_j\right). \end{aligned}$$

Theorem 2.0.1 *Suppose that $E(P_k|\mathcal{F}_{k-1}) \neq 0$ P-a.s., for $k = 1, \ldots, n$. Then the solution $\left(V_0, \overrightarrow{\phi}\right)$ of the minimization problem is $V_0 = E(\beta_n C P_1)/E(P_1)$, and*

$$\phi_k = \alpha_k - V_{k-1} b_k, \quad k = 1, \ldots, n.$$

The proof is given in Appendix A.

Let C_k be the optimal investment at period k so that the value of the portfolio at period n is as close as possible to C, in terms of mean square error. It follows from Theorem 2.0.1 that C_k is given by

$$\beta_k C_k = \frac{E(\beta_n C P_{k+1} | \mathcal{F}_k)}{E(P_{k+1} | \mathcal{F}_k)}, \qquad k = 0, \ldots, n. \tag{2.0.1}$$

Using (2.0.1), it is easy to check that an alternative expression for C_k is given by

$$\begin{aligned}
\beta_{k-1}C_{k-1} &= E(\beta_k C_k \mathcal{U}_k|\mathcal{F}_{k-1}) \\
&= \frac{1}{\gamma_k} E\left\{\beta_k C_k \left(1 - b_k^\top \Delta_k\right)\gamma_{k+1}|\mathcal{F}_{k-1}\right) \\
&= E\left(\beta_n C \mathcal{U}_k \cdots \mathcal{U}_n|\mathcal{F}_{k-1}\right),
\end{aligned} \tag{2.0.2}$$

where $\mathcal{U}_k = \frac{E(P_k|\mathcal{F}_k)}{E(P_k|\mathcal{F}_{k-1})}$, $\gamma_k = E(P_k|\mathcal{F}_{k-1})$, $k = 1, \ldots, n+1$, while an alternative expression for α_k is

$$\alpha_k = A_k^{-1} E\left(\beta_k C_k \Delta_k \gamma_{k+1}|\mathcal{F}_{k-1}\right). \tag{2.0.3}$$

Remark 2.0.2 *Setting $\mathcal{Z}_0 = 1$ and $\mathcal{Z}_k = \prod_{j=1}^k \mathcal{U}_j$, $k = 1, \ldots, n$, one obtains that $(\mathcal{Z}_k, \beta_k C_k \mathcal{Z}_k, \beta_k S_k \mathcal{Z}_k)_{k=0}^n$ are martingales, since $E(\mathcal{U}_k|\mathcal{F}_{k-1}) = 1$ and $E(\Delta_k \mathcal{U}_k|\mathcal{F}_{k-1}) = 0$, because $E\left\{\Delta_k(1 - b_k^\top \Delta_k)P_{k+1}|\mathcal{F}_{k-1}\right\} = 0$. However, in most applications, $\mathcal{Z}$ does not define a change of measure in general since it can take negative values.*

Set $G_k = \beta_k C_k - V_k$, $k = 0, \ldots, n$. The following properties of the hedging error process G will be important in the next section.

Proposition 2.0.3 *$\gamma_{k+1}G_k$ is a martingale and so is $\beta_k \gamma_{k+1} S_k G_k$.*

Proof: First, $V_k = \alpha_k^\top \Delta_k + V_{k-1}\left(1 - b_k^\top \Delta_k\right)$, so

$$\begin{aligned}
E\left\{V_k E(P_{k+1}|\mathcal{F}_k)|\mathcal{F}_{k-1}\right\} &= E\left\{\alpha_k^\top \Delta_k E(P_{k+1}|\mathcal{F}_k)|\mathcal{F}_{k-1}\right\} \\
&\quad + E\left\{V_{k-1}\left(1 - b_k^\top \Delta_k\right) E(P_{k+1}|\mathcal{F}_k)|\mathcal{F}_{k-1}\right\} \\
&= \alpha_k^\top A_k b_k + V_{k-1} E(P_k|\mathcal{F}_{k-1}) \\
&= b_k^\top E(\beta_n C \Delta_k P_{k+1}|\mathcal{F}_{k-1}) + V_{k-1} E(P_k|\mathcal{F}_{k-1}) \\
&= E(\beta_n C P_{k+1}|\mathcal{F}_{k-1}) - E(\beta_n C P_k|\mathcal{F}_{k-1}) \\
&\qquad + V_{k-1} E(P_k|\mathcal{F}_{k-1}) \\
&= E\left\{\beta_k C_k E(P_{k+1}|\mathcal{F}_k)\mathcal{F}_{k-1}\right\} - G_{k-1} E(P_k|\mathcal{F}_{k-1}).
\end{aligned}$$

Hence, $\gamma_{k+1}G_k$ is a martingale. Finally, the last claim follows from the fact that $E(G_n \Delta_k|\mathcal{F}_{k-1}) = 0$, combined with the martingale property of $\gamma_{k+1}G_k$. ■

2.1 Markovian models

If the price process is Markov and $C_n = C_n(S_n)$, then $C_k = C_k(S_k)$, $\alpha_k = \alpha_k(S_{k-1})$, and $b_k = b_k(S_{k-1})$. It follows that all these functions can be approximated using the methodology developed in Papageorgiou et al. (2008).

Another interesting case encountered in practice is when S_k is not a Markov process but (S_k, h_k) is Markov, even if h_k is not observable, as in GARCH models or Hidden Markov models (HMM for short).

If $C_n = C_n(S_n)$, then $C_k = C_k(S_k, h_k)$, $\alpha_k = \alpha_k(S_{k-1}, h_{k-1})$, and $b_k = b_k(S_{k-1}, h_{k-1})$.

More precisely, setting $\gamma_k(S_{k-1}, h_{k-1}) = E(P_k|\mathcal{F}_{k-1})$, for $k = 1, \ldots, n+1$, then, assuming for simplicity that $\beta_k = \beta_1 \beta_{k-1}$, one gets, for $k = n, \ldots, 1$,

$$\begin{aligned}
A_k(s,h) &= \beta_{k-1}^2 E_{s,h}\left\{(\beta_1 S_1 - s)(\beta_1 S_1 - s)^\top \gamma_{k+1}(S_1, h_1)\right\}, \\
b_k(s,h) &= \beta_{k-1} A_k^{-1}(s,h) E_{s,h}\left\{(\beta_1 S_1 - s)\gamma_{k+1}(S_1, h_1)\right\}, \\
\gamma_k(s,h) &= E_{s,h}\{\gamma_{k+1}(S_1, h_1)\} - b_k^\top(s,h) A_k(s,h) b_k(s,h), \\
C_{k-1}(s,h) &= \frac{\beta_1}{\gamma_k(s,h)} E_{s,h}\left[C_k(S_1, h_1)\gamma_{k+1}(S_1, h_1)\left\{1 - \beta_{k-1} b_k(s,h)^\top (\beta_1 S_1 - s)\right\}\right], \\
\alpha_k(s,h) &= \beta_k \beta_{k-1} A_k^{-1}(s,h) E_{s,h}\left\{C_k(S_1, h_1)\gamma_{k+1}(S_1, h_1)(\beta_1 S_1 - s)\right\}.
\end{aligned}$$

Again, all these functions can be approximated using the methodology developed in Rémillard et al. (2009). Implementation of the hedging strategy then requires prediction of h_t given $S_0, \ldots, S_t$, which is a filtering problem. See Rémillard et al. (2009) for an implementation in the HMM case, where $P(\tau_t = i | S_0, \ldots, S_t)$ occurs naturally when estimating the parameters of the HMM model using the EM algorithm.

Remark 2.1.1 *One could suggest to use the smallest filtration to get rid of the unobservable process h but in that case, all conditional expectations based on $\mathcal{F}_k$ would depend on all past values $S_0, \ldots, S_k$, making it impossible to implement in practice.*

2.1.1 GARCH type models

For that model, one assumes that

$$\Delta_k = \beta_k S_k - \beta_{k-} S_{k-1} = \beta_{k-1} S_{k-1} \xi_k,$$

with

$$\begin{aligned} \xi_k &= \pi_1(h_{k-1}, \epsilon_k) \\ h_k &= \pi_2(h_{k-1}, \epsilon_k), \end{aligned}$$

where the innovations ϵ_k are independent and identically distributed with probability law ν. It is immediate that (S_k, h_k) is a Markov process. Furthermore, almost all known GARCH(1,1) models can be written in that way. Further assume that $C = C_n(S_n)$.

It is easy to check that for all $k = n, \ldots, 1$, $\gamma_k = \gamma_k(h_{k-1})$ and

$$\begin{aligned} A_k(s,h) &= \beta_{k-1}^2 s^2 B_k(h), \\ b_k(s,h) &= \frac{\mu_k(h)}{s\beta_{k-1} B_k(h)}, \\ b_k(S_{k-1}, h_{k-1})^\top \Delta_k &= \frac{\xi_k \mu_k(h_{k-1})}{B_k(h_{k-1})}, \\ C_{k-1}(s,h) &= \frac{\beta_1}{\gamma_k(h)} \int C_k \left[\frac{s}{\beta_1}\{1+\pi_1(h,y)\}, \pi_2(h,y) \right] \gamma_{k+1}\{\pi_2(h,y)\} \\ &\qquad \times \left\{ 1 - \frac{\mu_k(h)}{B_k(h)} \pi_1(h,y) \right\} \nu(dy), \\ \alpha_k(s,h) &= \frac{\beta_1}{s B_k(h)} \int C_k \left[\frac{s}{\beta_1}\{1+\pi_1(h,y)\}, \pi_2(h,y) \right] \gamma_{k+1}\{\pi_2(h,y)\} \\ &\qquad \times \pi_1(h,y) \nu(dy), \end{aligned}$$

where

$$\begin{aligned} B_k(h) &= \int \pi_1^2(h,y) \gamma_{k+1}\{\pi_2(h,y)\} \nu(dy), \\ \mu_k(h) &= \int \pi_1(h,y) \gamma_{k+1}\{\pi_2(h,y)\} \nu(dy), \\ \gamma_k(h) &= \int \left\{ 1 - \frac{\mu_k(h)}{B_k(h)} \pi_1(h,y) \right\} \gamma_{k+1}\{\pi_2(h,y)\} \nu(dy). \end{aligned}$$

Example 2.1.2 (Binomial tree) *Suppose that $d = 1$ and $S_k = S_{k-1}\zeta_k$, $\beta_k = (1+R)^{-k}$, where $P(\zeta_k = U) = p$ and $P(\zeta_k = D) = 1 - p$, where $D < 1 + R < U$. Then $\xi_k = \frac{\zeta_k}{1+R} - 1$, $\mu = p\frac{U-D}{1+R} - \frac{1+R-D}{1+R}$ and $B = p(1-p)\frac{(U-D)^2}{(1+R)^2} + \mu^2$. Furthermore, setting $q = \frac{1+R-D}{U-D}$, it follows that $\mathcal{U}_k = \frac{q}{p}$ with probability p and $\mathcal{U}_k = \frac{1-q}{1-p}$ with probability $1-p$.*

It is easy to check that one recovers the usual formulas from (Cox et al., 1979) for C_k and ϕ_k. In addition, $G \equiv 0$, i.e., there is no hedging error.

2.1.2 Regime switching geometric random walks

An important alternative model to the usual geometric random walk is to consider a regime-switching geometric random walk. That model displays serial dependence in the log-returns and accounts for much variability of the asset behavior. For implementation issues, including estimation, prediction and goodness-of-fit tests, one may consult Rémillard et al. (2009).

To define the process, suppose that τ is a finite homogeneous Markov chain with transition matrix Q with values in $\{1,\ldots,l\}$. Further assume that given $\tau_1 = i_1,\ldots,\tau_n = i_n$, $\xi_1,\ldots,\xi_n$ are independent with $\xi_j \sim \mathbb{P}_{i_j}$, $j = 1,\ldots,n$, have mean $\mathbb{E}_i(\xi_j) = E(\xi_j|\tau_j = i) = \mu(i)$ and $\mathbb{E}_i\left(\xi_j\xi_j^\top\right) = B(i)$. Setting $X_k = \beta_k S_k$, suppose that $\Delta_k = X_k - X_{k-1} = D(X_{k-1})\xi_k$, $k = 1,\ldots,n$, where $D(s)$ be the diagonal matrix with $(D(s))_{ii} = s_i$, for all $i = 1,\ldots,d$.

It then follows that given the regimes, the log-returns associated with S are independent, hence the name regime-switching geometric random walk.

Note that S is not a Markov process in general but (S,τ) is a Markov process.

Next, set $\gamma_k(\tau_{k-1}) = E(P_k|\mathcal{F}_{k-1})$, $k = 1,\ldots,n$, and $\gamma_{n+1} \equiv 1$. For $k = 1,\ldots,n+1$, and $i = 1,\ldots,l$, further set

$$\rho_k(i) = \left\{\sum_{j=1}^{l} Q_{ij}\gamma_k(j)B(j)\right\}^{-1}\left\{\sum_{j=1}^{l} Q_{ij}\gamma_k(j)\mu(j)\right\}. \tag{2.1.1}$$

Then, it is easy to check that on $\{S_{k-1} = s \text{ and } \tau_{k-1} = i\}$,

$$\begin{aligned} b_k &= b_k(s,i) = e^{r(k-1)}D^{-1}(s)\rho_{k+1}(i), \\ b_k^\top\Delta_k &= \rho_{k+1}(i)^\top\xi_k, \\ \gamma_k(i) &= \sum_{j=1}^{l} Q_{ij}\gamma_{k+1}(j)\left\{1-\rho_{k+1}(i)^\top\mu(j)\right\}, \end{aligned}$$

for all $k = 1,\ldots,n$.

The following proposition, proved in Appendix E.1 is important in the sequel.

Proposition 2.1.3 *For any $k = 1,\ldots,n+1$ and $i = 1,\ldots,l$, $\gamma_k(i) \in (0,1]$.*

For simplicity, set

$$(Q_k)_{ij} = \frac{Q_{ij}\gamma_{k+1}(j)}{\gamma_k(i)}\left\{1-\rho_{k+1}(i)^\top\mu(j)\right\}, \quad 1 \le i,j \le l.$$

It follows from Proposition 2.1.3 and the definition of γ_k that Q_k is the transition matrix of a non homogeneous Markov chain.

If in addition $C = \Phi(S_n)$, then $C_k = C_k(S_k,\tau_k)$ and $\alpha_k = \alpha_k(S_{k-1},\tau_{k-1})$, where

$$\begin{aligned} C_{k-1}(s,i) &= \frac{\beta_k}{\beta_{k-1}}\sum_{j=1}^{l} Q_{ij}\frac{\gamma_{k+1}(j)}{\gamma_k(i)} \\ &\quad \times\int C_k\left\{\frac{\beta_{k-1}}{\beta_k}D(s)(\mathbf{1}+y),j\right\}\left\{1-\rho_{k+1}(i)^\top y\right\}\mathbb{P}_j(dy) \\ &= \frac{\beta_k}{\beta_{k-1}}\sum_{j=1}^{l}(Q_k)_{ij}\int C_k\left\{\frac{\beta_{k-1}}{\beta_k}D(s)(\mathbf{1}+y),j\right\}\left\{\frac{1-\rho_{k+1}(i)^\top y}{1-\rho_{k+1}(i)^\top\mu(j)}\right\}\mathbb{P}_j(dy), \end{aligned}$$

and

$$\begin{aligned}\alpha_k(s,i) &= \frac{\beta_k}{\beta_{k-1}} D^{-1}(s) \left\{ \sum_{j=1}^{l} Q_{ij}\gamma_{k+1}(j)B(j) \right\}^{-1} \sum_{j=1}^{l} Q_{ij}\gamma_{k+1}(j) \\ &\quad \times \int C_k \left\{ \frac{\beta_{k-1}}{\beta_k} D(s)(\mathbf{1}+y), j \right\} y\mathbb{P}_j(dy).\end{aligned}$$

3 Optimal hedging strategy for regime-switching geometric Levy processes

We first define the models, state some important properties and then prove the optimality of the proposed solution.

3.1 Regime switching Lévy processes

First we recall the definition of a Lévy process. In the following, we consider Lévy processes with exponential moments, i.e., L is a Lévy process with parameters (υ, a, ν) if it is a càdlàg process with independent increments, such that for all $\theta \in B_d(0, 2+\epsilon)$,

$$E\left(e^{\theta^\top L_t}\right) = e^{t\Psi_{\upsilon,a,\nu}(\theta)},$$

where

$$\Psi(\theta) = \theta^\top \upsilon + \frac{1}{2}\theta^\top a\theta + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{y^\top\theta} - 1 - \theta^\top y\right)\nu(dy). \tag{3.1.1}$$

Here $\upsilon \in \mathbb{R}^d$, a is a non-negative definite $d \times d$ matrix, and ν is a Lévy measure, that is a non-negative measure such that $\int_{\mathbb{R}^d\setminus\{0\}} \min\left(1, |y|^2\right)\nu(dy) < \infty$. In particular, $E(L_t) = t\upsilon$, $\mathrm{Cov}(L_t, L_t) = t(a + a_\nu)$, where $a_\nu = \int_{\mathbb{R}^d\setminus\{0\}} yy^\top \nu(dy)$, and $E\left\{e^{(L_t)_j}\right\} = e^{t\psi_j}$, where

$$\psi_j = \upsilon_j + \frac{a_{jj}}{2} + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{y_j} - 1 - y_j\right)\nu(dy), \quad j = 1, \ldots, d.$$

Remark 3.1.1 *Note Ψ is usually written as follows:*

$$\Psi(\theta) = \theta^\top \upsilon' + \frac{1}{2}\theta^\top a\theta + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{y^\top\theta} - 1 - \theta^\top y 1_{\{|y|\le 1\}}\right)\nu(dy).$$

However, since it is assumed that the moment generating function exists in a ball of radius at least 2, then the Lévy process has moments of all orders, and it follows that $\int |y| 1_{\{|y|>1\}}\nu(dy)$ is finite, and since $e^{y^\top\theta} - 1 - \theta^\top y = O(|y|^2)$, one has

$$\begin{aligned}\int_{\mathbb{R}^d\setminus\{0\}} \left\{e^{y^\top\theta} - 1 - \theta^\top y\right\}\nu(dy) &= \int_{\mathbb{R}^d\setminus\{0\}} \left\{e^{y^\top\theta} - 1 - \theta^\top y 1_{\{|y|\le 1\}}\right\}\nu(dy) \\ &\quad -\theta^\top \int y 1_{\{|y|>1\}}\nu(dy).\end{aligned}$$

Therefore the two representations coincide if

$$\upsilon = \upsilon' + \int y 1_{\{|y|>1\}}\nu(dy).$$

The infinitesimal generator $\mathcal{L}_L$ of L is thus given by

$$\begin{aligned}\mathcal{L}_L f(x) &= \sum_{i=1}^{d} \upsilon_i \partial_{x_i} f(x) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} a_{ij}\partial_{x_i}\partial_{x_j} f(x) \\ &\quad + \int_{\mathbb{R}^d\setminus\{0\}} \left\{ f(x+y) - f(x) - y^\top \nabla f(x) \right\} \nu(dy),\end{aligned}$$

for all nice function f, in particular for infinitely differentiable functions with compact support and their limits. That includes for example exponentials.

The main property of infinitesimal generators $\mathcal{L}$ of a Markov process x_t that will be used throughout the paper is that for any nice function,

$$f(x_t) - \int_0^t \mathcal{L} f(x_u) du$$

is a martingale. In fact, in most interesting cases, the latter property characterizes the law of the process and it is basically the definition of the so-called martingale problem. See, e.g., Ethier and Kurtz (1986).

Next, to define a regime-switching Lévy process, let τ be a continuous time Markov chain on $\{1,\ldots,l\}$, with infinitesimal generator Λ. In particular, $P(\tau_t = j|\tau_0 = i) = P_{ij}(t)$, where the transition matrix P can be written as $P(t) = e^{t\Lambda}$, $t \ge 0$. The process L_t is a regime-switching Lévy process with parameters $(\upsilon(i), a(i), \nu_i)$, $i = 1,\ldots,l$, and Λ, if the process (L,τ) is a Markov process with infinitesimal generator

$$\mathcal{L}_L f(s,i) = \mathcal{L}_{L_i} f_i(s) + \sum_{j=1}^{l} \Lambda_{ij} f(s,j),$$

where $\mathcal{L}_{L_i}$ is the infinitesimal generator of a Lévy process L_i with parameters $(\upsilon(i), a(i), \nu_i)$, $i = 1,\ldots,l$.

Such a process is easy to construct. If T_k denotes the time of the k-th jump of τ, and if τ jumped from state i to state j, then

$$L_t = L_{T_k} + L_j(t) - L_j(T_k), \quad T_k \le t \le T_{k+1}.$$

In particular, $L_{T_{k+1}} - L_{T_k} = L_{T_{k+1},\tau_{T_k}} - L_{T_k,\tau_{T_k}}$.

In Appendix D one shows another construction than can be applied in the more general case of non-homogeneous Markov chains.

It follows that L is continuous if and only if each Lévy process $L_{i,t}$ is continuous. Therefore the only continuous regime-switching Lévy process is the so-called regime-switching Brownian motion (Hamilton, 1990) with generator

$$\mathcal{L}_L f(s,i) = \mathcal{L}_{L_i} f_i(s) + \sum_{j=1}^{l} \Lambda_{ij} f(s,j),$$

with

$$\mathcal{L}_{L_i} f(x) = \sum_{i=1}^{d} \upsilon_i \partial_{x_i} f(x) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} a_{ij}\partial_{x_i}\partial_{x_j} f(x),$$

$i = 1,\ldots,l$.

Next, since L plays the role of the log-return of the price process S, hereafter called a regime-switching geometric Lévy process, the process S is defined by

$$S_t = D(s)\mathbf{e}^{L_t}, \qquad t \ge 0,$$

i.e., for all $j = 1,\ldots,d$, the j-th component $(S_t)_j$ of S_t is $s_j e^{(L_t)_j}$.

As a result, (S,τ) is a Markov process with with infinitesimal generator $\mathcal{L}$ defined by

$$\mathcal{L}f(s,i) = \mathcal{L}_i f(s,i) + \sum_{j=1}^{l} \Lambda_{ij} f(s,j), \tag{3.1.2}$$

where for each $i = 1,\ldots,l$, $\mathcal{L}_i$ is the infinitesimal generator associated with the (geometric Lévy) process $S_{i,t} = D(s)\mathbf{e}^{L_{i,t}}$, and is given by

$$\mathcal{L}_i f(s) = \psi(i)^\top D(s)\nabla f(s) + \frac{1}{2}\sum_{k=1}^{d}\sum_{j=1}^{d} a_{kj}(i) s_k s_j \partial_{s_k}\partial_{s_j} f(s) + \mathcal{L}_{J,i} f(s),$$

with

$$\psi(i) = v(i) + \frac{1}{2}\text{diag}\{a(i)\} + \int_{\mathbb{R}^d\setminus\{0\}} (\mathbf{e}^y - \mathbf{1} - y)\,\nu_i(dy),$$

where $\text{diag}(a)$ is the diagonal matrix formed with the elements of the diagonal of a,

$$\mathcal{L}_{J,i} f(s) = \int_{\mathbb{R}^d\setminus\{0\}} \left[f\left\{D(s)(\mathbf{1}+y)\right\} - f(s) - y^\top D(s)\nabla f(s)\right]\tilde{\nu}_i(dy),$$

and

$$\int_{\mathbb{R}^d\setminus\{0\}} f(y)\tilde{\nu}_i(dy) = \int_{\mathbb{R}^d\setminus\{0\}} f\left(\mathbf{e}^y - \mathbf{1}\right)\nu_i(dy).$$

Finally, for $i = 1,\ldots,l$, set

$$\mathbb{A}(i) = a(i) + \int_{\mathbb{R}^d\setminus\{0\}} \left(\mathbf{e}^y - \mathbf{1}\right)\left(\mathbf{e}^y - \mathbf{1}\right)^\top \nu_i(dy) = a(i) + a_{\tilde{\nu}_i},$$

$m(i) = (\psi(i) - r\mathbf{1})$.

It is assumed that for each $i \in \{1,\ldots,l\}$, $\mathbb{A}(i)$ is invertible and so $\rho(i) = \{\mathbb{A}(i)\}^{-1}m(i)$ and $\ell(i) = \rho(i)^\top m(i)$ are well defined.

We are now in a position to state some properties of the Markov process (S,τ). First, note that S and τ are semimartingales. In fact, if $g(s,j) = s$, then $\mathcal{L}g(s,i) = D(s)\psi(i)$, so

$$M_t^{(g)} = S_t - S_0 - \int_0^t D(S_u)\psi(\tau_u)du \tag{3.1.3}$$

is a martingale. As a result, one obtains the following representation for the discounted value X of S:

$$X_t = e^{-rt}S_t = S_0 + \int_0^t D(X_u)m(\tau_u)du + \int_0^t e^{-ru}dM_u^{(g)}. \tag{3.1.4}$$

Next, setting $h(s,j) = j$, one obtains that

$$M_t^{(h)} = \tau_t - \tau_0 - \int_0^t \Lambda h(\tau_u)du$$

is a martingale. Moreover, it follows from Lemma C.0.4 that $\left[M^{(g)}, M^{(h)}\right]_t$ is a martingale.

Next, set

$$\gamma(t) = e^{t\{\Lambda - D(\ell)\}}\mathbf{1}. \tag{3.1.5}$$

Then $\gamma(0,i) = 1$ for all $i = 1,\ldots,l$ and

$$\dot{\gamma}(t,i) = \frac{d}{dt}\gamma(t,i) = -\ell(i)\gamma(t,i) + \sum_{j=1}^{l}\Lambda_{ij}\gamma(t,j).$$

Finally, set

$$(\Lambda_t)_{ij} = \Lambda_{ij}\gamma(t,j)/\gamma(t,i), \quad i \neq j, \qquad (\Lambda_t)_{ii} = -\sum_{j\neq i}(\Lambda_t)_{ij}. \tag{3.1.6}$$

Remark 3.1.2 *Note that Λ_t, defined by (3.1.6) is the infinitesimal generator of a time non homogeneous Markov chain $\tilde\tau$. Moreover, for any function f on $\{1,\ldots,l\}$,*

$$\Lambda_t f(i) = \frac{1}{\gamma(t,i)}\sum_{j=1}^{l}\Lambda_{ij}\gamma(t,j)\{f(j)-f(i)\}.$$

In that case, given $\tilde\tau_t = i$, one has

$$\begin{aligned} P(\tilde\tau_{t+u} = i \text{ for all } u \leq s|\tilde\tau_t = i) &= e^{\int_t^{t+s}(\Lambda_u)_{ii}du} \\ &= \frac{\gamma(t,i)}{\gamma(t+s,i)}e^{\{\Lambda_{ii}-\ell(i)\}s}, \end{aligned} \tag{3.1.7}$$

which is the distribution function the time in state i after time t, if $\tilde\tau_t = i$. When it jumps after time t, $\tilde\tau$ chooses state $j \neq i$ with probability

$$\int_t^\infty (\Lambda_s)_{ij}e^{\int_t^s(\Lambda_u)_{ii}du}ds.$$

Finally, for every $t \leq s$, and $i,j \in \{1,\ldots,l\}$,

$$P_(\tilde\tau_s = j|\tau_t = i) = \left(e^{\int_t^{t+h}\Lambda_u du}\right)_{ij}.$$

The following lemma is fundamental for the analysis of the optimal solution. Its proof is given in Appendix B.1. Before stating it, set, for any $i = 1,\ldots,l$,

$$\mathcal{K}_i f(s) = \int_{\mathbb{R}^d\setminus\{0\}} y\left[f\{D(s)(\mathbf{1}+y)\} - f(s) - y^\top D(s)\nabla f(s)\right]\tilde\nu_i(dy).$$

Lemma 3.1.3 *If $X_t = e^{-rt}S_t$, $M_t = \int_0^t \rho^\top(\tau_{u-})D^{-1}(X_{u-})dX_u$ and $Z = \mathcal{E}\{-M\}$, then Z is a multiplicative functional, and if*

$$v_t(s,i) = v(t,s,i) = E\{f(S_t,\tau_t)Z_t|S_0 = s, \tau_0 = i\}/\gamma(t,i),$$

then $v(0,s,i) = f(s,i)$ and

$$\partial_t v_t = \mathcal{H}_t v_t, \tag{3.1.8}$$

where

$$\begin{aligned} \mathcal{H}_t f(s,i) &= \mathcal{L}_i f(s,i) - m^\top(i)D(s)\nabla f(s) - \rho^\top(i)\mathcal{K}_i f(s,i) + \Lambda_t f(s,i) \\ &= r\sum_{k=1}^d s_k\partial_{s_k}f(s,i) + \frac{1}{2}\sum_{j=1}^d\sum_{k=1}^d a_{jk}(i)s_js_k\partial_{s_j}\partial_{s_k}f(s,i) \\ &\quad + \int_{\mathbb{R}^d\setminus\{0\}}\left\{1-\rho^\top(i)y\right\}\left[f\{D(s)(\mathbf{1}+y)\} - f(s) - y^\top D(s)\nabla f(s)\right]\tilde\nu_i(dy) \\ &\quad + \sum_{j=1}^l(\Lambda_t)_{ij}f(s,j). \end{aligned} \tag{3.1.9}$$

In particular,

$$\gamma(t,i) = E(Z_t|S_0 = s, \tau_0 = i), \qquad i = 1,\ldots,l.$$

and

$$E(Z_T|\mathcal{F}_t) = Z_t\gamma(T-t,\tau_t), \quad 0 \leq t \leq T. \tag{3.1.10}$$

Furthermore, one can write

$$v_t = e^{\int_0^t\mathcal{H}_u du}f.$$

Remark 3.1.4 *Set $Y_t = \gamma(T-t,\tau_t)S_t Z_t$. Using Lemma B.0.8, it is easy to check that*

$$Y_t - Y_0 - r\int_0^t Y_u du$$

is a martingale, proving that $e^{-rt}Y_t = X_t\gamma(T-t,\tau_t)Z_t$ is a martingale. As a result, using the last result and Lemma 3.1.3, one obtains that

$$\frac{E\{X_T Z_T|\mathcal{F}_t\}}{E(Z_T|\mathcal{F}_t)} = \frac{E\left\{e^{-rT}Y_T|\mathcal{F}_t\right\}}{E(Z_T|\mathcal{F}_t)} = \frac{e^{-rt}Y_t}{\gamma(T-t,\tau_t)} = X_t Z_t.$$

If it happens that Z is positive, then $\frac{d\tilde{P}_i}{dP_i} = Z_T/\gamma(T,i)$ defines a change of measure under which X is a martingale. For example, for the regime-switching geometric Brownian motion, S is continuous so Z is positive, being an exponenttial.

3.2 Optimal solution

Let C be the unique solution of

$$\partial_t C_t(s,i) + \mathcal{H}_{T-t}C_t(s,i) = rC_t(s,i), \quad C_T(s,i) = \Phi(s). \tag{3.2.1}$$

Using Lemma 3.1.3, one can write

$$C_t(S_t,\tau_t) = E\left\{\Phi(S_T)Z_T|\mathcal{F}_t\right\}/E(Z_T|\mathcal{F}_t) = E\left\{\Phi(S_T)Z_T\mathcal{F}_t\right\}/\gamma_{T-t}(\tau_t), \tag{3.2.2}$$

where $M_t = \int_0^t \rho^\top(\tau_{u-})D^{-1}(X_{u-})dX_u$ and $Z = \mathcal{E}\left\{-M\right\}$.

Set

$$\begin{aligned}
\alpha(t,s,i) &= D^{-1}(s)\mathbb{A}(i)^{-1}D^{-1}(s)\left\{\mathcal{L}_i(C_t g) - g\mathcal{L}_i(C_t) - rgC_t\right\}(s,i)\\
&= D^{-1}(s)\mathbb{A}(i)^{-1}\left\{m(i)C_t(s,i) + \mathbb{A}(i)D(s)\nabla C_t(s,i) + \mathcal{K}_i C_t(s,i)\right\}\\
&= \nabla C_t(s,i) + D^{-1}(s)\mathbb{A}(i)^{-1}\left\{C_t(s,i)m(i) + \mathcal{K}_i C_t(s,i)\right\}.
\end{aligned} \tag{3.2.3}$$

Suppose that V satisfies the following stochastic differential equation:

$$V_t = C(0,s,i) + \int_0^t \alpha(u-,S_{u-},\tau_{u-})^\top dX_u - \int_0^t V_{u-}dM_u. \tag{3.2.4}$$

It follows from Protter (2004)[Theorem V.7] that V is uniquely determined by M and S, since the solution of (3.2.4) is unique.

Next, set

$$\phi_t = \alpha(t,S_{t-},\tau_{t-}) - V_{t-}D^{-1}(X_{t-})\rho(\tau_{t-}) \tag{3.2.5}$$

$$\begin{aligned}
&= \nabla C_t(S_{t-},\tau_{t-}) + G_{t-}D^{-1}(\beta_t S_{t-})\rho(\tau_{t-})\\
&\quad + D^{-1}(S_{t-})\mathbb{A}^{-1}(\tau_{t-})\mathcal{K}_{\tau_{t-}}C_t(S_{t-},\tau_{t-}),
\end{aligned} \tag{3.2.6}$$

with $G_t = e^{-rt}C(t,S_t,\tau_t) - V_t$. Note that ϕ is predictable.

As a result, one can also write

$$V_t = C(0,s,i) + \int_0^t \phi_u^\top dX_u, \tag{3.2.7}$$

i,.e., V_t can be seen as the actualized value at time t of a portfolio with strategy ϕ, while G_t is the corresponding hedging error at that period. One can now find an expression for the hedging error associated with strategy ϕ.

Lemma 3.2.1 *Let $M^{(C)}$ and $M^{(g)}$ be the martingales respectively defined by*

$$\begin{aligned} M_t^{(C)} &= C(t,S_t,\tau_t) - C(0,s,i) - \int_0^t \mathcal{L}C_u(S_u,\tau_u)du, \\ M_t^{(g)} &= S_t - s - \int_0^t D(S_u)\psi(\tau_u)du. \end{aligned}$$

Then, for $0 \le t \le T$,

$$\begin{aligned} G_t &= \int_0^t e^{-ru} dM_u^{(C)} - \int_0^t e^{-ru}\phi_u^\top dM_u^{(g)} \\ &\quad - \int_0^t \ell(\tau_u) G_u du + \int_0^t e^{-ru}\left\{(\Lambda - \Lambda_{T-u})C_{u,S_u}\right\}(\tau_u)du. \end{aligned}$$

The proof is given in Appendix B.2.

Finally, here are some interesting properties of the hedging error which are essential in proving the optimality of the strategy based on ϕ.

Lemma 3.2.2 *For all $0 \le t \le T$, $\gamma(T-t,\tau_t)G_t$ and $\beta_t\gamma(T-t,\tau_t)S_tG_t$ are martingales. In particular $E\{G_T\} = 0$ and for any $0 \le u \le t \le T$,*

$$E\left(G_T X_t | \mathcal{F}_u\right) = E\left(G_T X_u | \mathcal{F}_u\right). \tag{3.2.8}$$

The proof is given in Appendix B.3.

Finally, using (3.2.8), one can state the main theorem of the section, whose proof is in Appendix B.4.

Theorem 3.2.3 *The optimal solution of the hedging problem for a regime-switching geometric Lévy process is given by ϕ, as defined by equation (3.2.5) and the actualized value of the associated portfolio satisfies (3.2.4).*

We now give some examples of calculations, the most interesting being the regime-switching geometric Brownian motion which should be used in practice instead the geometric Brownian motion.

3.2.1 Geometric Brownian motion

In that case, one can write $M_t^{(g)} = \int_0^t D(S_u)\sigma dW_u$, where $\sigma\sigma^\top = a$ and W is a Brownian motion. Also $M_t^{(C)} = \int_0^t \nabla C_u(S_u)^\top D(S_u)\sigma dW_u$. It follows from Lemma 3.2.1 that

$$\begin{aligned} e^{\ell t}G_t &= \int_0^t e^{(\ell-r)u}\nabla C_u(S_u)^\top D(S_u)\sigma dW_u - \int_0^t e^{(\ell-r)u}\phi_u^\top D(S_u)\sigma dW_u \\ &= -\rho^\top\sigma\int_0^t e^{\ell u} G_u dW_u. \end{aligned}$$

Since the solution of the last stochastic differential equation is unique, it follows that $G \equiv 0$, proving the perfect hedging, as it is well known for the Black-Scholes model. One also obtains the usual expression for ϕ, that is $\phi_t = \nabla C_t(S_t)$.

3.2.2 Risk neutral measure

To recover known results from the literature, suppose that $X_t = e^{-rt}S_t$ is a martingale. It then follows from (3.1.4) that $m \equiv 0$, so $\psi = r\mathbf{1}$, $\mathcal{H}_t = \mathcal{L}$, $\rho \equiv 0$, $\ell \equiv 0$, so $c \equiv 1$, $M \equiv 0$ and $Z \equiv 1$.

Next, we get from (3.2.2) – (3.2.5) that

$$\begin{aligned}
C_t(s,i) &= e^{-r(T-t)}E\left\{\Phi(S_{T-t})|S_0=s,\tau_0=i\right\},\\
\alpha(t,s,i) &= \nabla C_t(s,i)+D^{-1}(s)\mathbb{A}(i)^{-1}\mathcal{K}_iC_t(s,i)\\
&= \left\{I-D^{-1}(s)\mathbb{A}^{-1}(i)D(s)a_{\tilde\nu_i}\right\}\nabla C_t(s,i)\\
&\qquad +D^{-1}(s)\mathbb{A}^{-1}(i)\int_{\mathbb{R}^d\setminus\{0\}} y\left[C_t\left\{D(s)(\mathbf{1}+y),i\right\}-C_s(s,i)\right]\tilde\nu_i(dy),\\
V_t &= C(0,s,i)+\int_0^t \alpha(u-,S_{u-},\tau_{u-})^\top dX_u,\\
\phi_t &= \alpha(t,S_{t-},\tau_{t-}).
\end{aligned}$$

In particular, if there is no regime-switching and $d=1$, one obtains formula (10.35) of Cont and Tankov (2004).

3.2.3 Regime switching geometric Brownian motion

Suppose that $\nu_i\equiv 0$. First note that $\mathbb{A}=a$, S is continuous, and its infinitesimal generator is given by

$$\begin{aligned}
\mathcal{L}f(s,i) &= \psi(i)^\top\nabla f_i(s)+f(s,i)+\frac12\sum_{j=1}^d\sum_{k=1}^d a_{jk}(i)s_js_k\partial_{s_j}\partial_{s_k}f(s,i)\\
&\qquad +\sum_{j=1}^l\Lambda_{ij}f(s,j).
\end{aligned}\tag{3.2.9}$$

Next, it follows that (3.1.9) reduces to

$$\begin{aligned}
\mathcal{H}_tf(s,i) &= r\sum_{k=1}^d s_k\partial_{s_k}f(s,i)+\frac12\sum_{j=1}^d\sum_{k=1}^d a_{jk}(i)s_js_k\partial_{s_j}\partial_{s_k}f(s,i)\\
&\qquad +\sum_{j=1}^l(\Lambda_t)_{ij}f(s,j).
\end{aligned}\tag{3.2.10}$$

As a result, $\mathcal{H}_t$ is the infinitesimal generator of a time non homogeneous Markov process $(\tilde S,\tilde\tau)$, where the Markov chain $\tilde\tau$ has infinitesimal generator (Λ_t), so

$$C_t(s,i)=e^{-r(T-t)}E\left\{\Phi(\tilde S_T)|\tilde S_t=s,\tilde\tau_t=i\right\}.\tag{3.2.11}$$

That law corresponds to the change of measure described in Remark 3.1.4. Note that using the algorithm described in Appendix D, it is easy to use a Monte-Carlo method to estimated C_t.

Next, according to (3.2.3),

$$\alpha(t,s,i)=\nabla C_t(s,i)+C_t(s,i)D^{-1}(s)\rho(i),\quad i=1,\ldots,l.$$

Again, using an obvious extension to the multivariate case of the "pathwise method" in Broadie and Glasserman (1996), one can use simulations to obtain an unbiased estimate of α_t. More precisely, if Φ is differentiable almost everywhere, then

$$\nabla C_t(s,i)=e^{-r(T-t)}D^{-1}(s)E\left\{D(\tilde S_T)\nabla\Phi(\tilde S_T)|\tilde S_t=s,\tilde\tau_t=i\right\},\tag{3.2.12}$$

so α_t can be written as an expectation of a function of $\tilde{S}_T$.

Finally, from (3.2.5), one gets

$$\phi_t = \nabla C_t(S_t, \tau_{t-}) + C_t(S_t, \tau_{t-}) D^{-1}(S_t) \rho(\tau_{t-}) - V_{t-} D^{-1}(X_t) \rho(\tau_{t-}).$$

In particular, $\phi_0 = \nabla C_0(S_0, \tau_0)$. It follows from (3.2.11) and (3.2.12) that ϕ_t can be estimated by Monte-Carlo methods.

Remark 3.2.4 *Since the martingale M is continuous, Theorem* **??** *can be applied to yield the following representation for V:*

$$V_t = Z_t \left\{ H_0 + \int_{0+}^{t} Z_u^{-1} d(H_u + [H, M]_u) \right\},$$

where

$$\begin{aligned} H_t &= C_0(s,i) + \int_0^t \alpha(u, S_u, \tau_u)^\top dX_u \\ &= C_0(s,i) + \int_0^t \left\{ \nabla C_u(S_u, \tau_u) + C_u(S_u, \tau_u) D^{-1}(S_u) \rho(\tau_u) \right\}^\top D(X_u) m(\tau_u) du \\ &\quad + \int_0^t e^{-ru} \left\{ \nabla C_u(S_u, \tau_u) + C_u(S_u, \tau_u) D^{-1}(S_u) \rho(\tau_u) \right\}^\top dM_u^{(g)} \end{aligned}$$

and

$$M_t = \int_0^t \ell(\tau_u) du + \int_0^t \rho(\tau_u)^\top D^{-1}(S_u) dM_u^{(g)},$$

with

$$[H, M]_t = \int_0^t e^{-ru} \alpha(u, S_u, \tau_u)^\top D(S_u) m(\tau_u) du + martingale,$$

by Lemma C.0.4.

4 Continuous time approximation

In what follows, we state some conditions under which the HMM model described in Section 2.1.2 can be approximated by a regime-switching geometric Lévy process. We then show that under slightly the same conditions, the "option prices" and the optimal strategy under a HMM model converge in some sense to the optimal strategy of a regime-switching geometric Lévy process.

4.1 Continuous time limit of the HMM price process

Suppose now that for each n, one has a HMM model $\left(S_k^{(n)}, \tau_k^{(n)}\right)$, where $\beta_k^{(n)} = e^{-rTk/n}$. Define $S^{(n)}(t) = S_{[nt/T]}^{(n)}$. From now on, when talking of convergence in law, denoted by $\rightsquigarrow$, we mean convergence in law in the space $D([0,T])$ with the Skorohod topology.

We will now state conditions under which $S^{(n)} \rightsquigarrow S$, where S is a regime-switching geometric Lévy process. For simplicity, let $\mathbb{E}_i$ denote expectation under the law of $\xi_1^{(n)}$ given $\tau_1^{(n)} = i$ and recall the following notations from Section 2.1.2: $\mathbb{E}_i\left(\xi_1^{(n)}\right) = \mu^{(n)}(i)$ and $\mathbb{E}_i\left(\xi_1^{(n)} \xi_1^{(n)\top}\right) = B^{(n)}(i)$, $i = m \ldots, l$.

Further let $C_2(\mathbb{R}^d)$ be the set of continuous functions f on $\mathbb{R}^d$ so that $f(y) = O(|y|^2)$ and $f(y)/|y|^2 \to 0$ as $y \to 0$.

Theorem 4.1.1 *Suppose that* $\lim_{n\to\infty} n\left(Q^{(n)} - I\right) \to \Lambda T$*. Assume also that for any* $i = 1,\ldots,l$*, the following conditions are satisfied, as* $n\to\infty$*:*

(i) $n\mu^{(n)}(i) \to Tm(i)$*,*

(ii) $nB^{(n)}(i) \to T\mathbb{A}(i)$*,*

(iii) for all $f \in C_2(\mathbb{R}^d)$*,*

$$n\mathbb{E}_i\left\{f\left(\xi_1^{(n)}\right)\right\} \to T\int_{\mathbb{R}^d\setminus\{0\}} f(y)\tilde{\nu}_i(dy).$$

Then $\left(S^{(n)},\tau^{(n)}\right) \rightsquigarrow (S,\tau)$ *with infinitesimal generator* $\mathcal{L}$ *defined by* (3.1.2)*.*

The proof of the theorem is given in Section B.5.

Example 4.1.2 *Consider a regime-switching geometric Gaussian random walk with*

$$\xi_k^{(n)} = e^{R_k^{(n)} - rT/n} - \mathbf{1},$$

where under $\mathbb{P}_i$*,* $R_k^{(n)}$ *is Gaussian with mean* $\upsilon(i)T/n$ *and covariance matrix* $a(i)T/n$*, where* $\upsilon(i) = \psi(i) - \frac{1}{2}\mathrm{diag}\{a(i)\}$*. It is easy to check that the conditions of Theorem 4.1.1 are met with* $\psi(i)$*,* $\mathbb{A}(i) = a(i)$ *and* $\nu_i \equiv 0$*. In other words, the limiting process is a regime-switching geometric Brownian motion with infinitesimal generator* (3.2.9)*, as one might have guessed.*

4.2 Continuous time limit of the optimal hedging strategy

For the rest of the section, suppose that the assumptions of Theorem 4.1.1 are met. We also use the definitions of Section 2.1.2, by adding the subscript n to denote dependence on n.

The first lemma deals about the behavior of $\gamma_k^{(n)}$ and $\rho_k^{(n)}$, as n tends to infinity. Its proof is given in Appendix B.6.

Lemma 4.2.1 *Set* $\gamma^{(n)}(t,i) = \gamma^{(n)}_{n+1-[nt/T]}(i)$ *and* $\rho^{(n)}(t,i) = \rho^{(n)}_{[nt/T]+1}(i)$*. Then* $\gamma_n(t) \to \gamma(t)$*, where* γ *is defined by* (3.1.5)*. Moreover* $\rho^{(n)}(t,i) \to \rho(i) = \mathbb{A}^{-1}(i)m(i)$*.*

Set $\zeta_k^{(n)} = e^{rT/n}\left(\xi_k^{(n)} + \mathbf{1}\right)$, $k = 1,\ldots,n$. Note that assumptions (i–iii) are equivalent to

(i') $n\mathbb{E}_i\left(\zeta_1^{(n)} - \mathbf{1}\right) \to T\psi(i)$,

(ii') $n\mathbb{E}_i\left\{\left(\zeta_1^{(n)} - \mathbf{1}\right)\left(\zeta_1^{(n)} - \mathbf{1}\right)^\top\right\} \to T\mathbb{A}(i)$,

(iii') for all $f \in C_2(\mathbb{R}^d)$,

$$n\mathbb{E}_i\left\{f\left(\zeta_1^{(n)} - \mathbf{1}\right)\right\} \to T\int_{\mathbb{R}^d\setminus\{0\}} f(y)\tilde{\nu}_i(dy).$$

We are now in a position to study the behavior of $C^{(n)}$.

Theorem 4.2.2 *Set* $C^{(n)}(t,s) = C^{(n)}_{[nt/T]}(s,i)$*, set* $Z^{(n)}(t) = Z^{(n)}_{[nt/T]}$*, where*

$$Z_k^{(n)} = \prod_{j=1}^{k}\left[1 - \{\rho_{k+1}^{(n)}\}^\top(i)\xi_k^{(n)}\right],$$

and

$$M_t^{(n)} = \int_0^t b^{(n)}(u-)^\top dX_t^{(n)},$$

where $X^{(n)}(t) = e^{-r[nt/T]}S^{(n)}(t)$.

Then

$$\left(S^{(n)}, X^{(n)}, \tau^{(n)}, M^{(n)}, Z^{(n)}\right) \rightsquigarrow (S, X, \tau, M, Z),$$

where $X_t = e^{-rt}S_t$, $M_t = \int_0^t \rho^\top(\tau_{u-})D^{-1}(X_{u-})dX_u$ *and* $Z = \mathcal{E}\{-M\}$, *as defined in Section 3.*

Moreover, if $\Phi(s) = O(|s|^p)$ *for some integer* p, *and for any* $j = 1, \ldots, d$,

$$E\left\{\left(\zeta^{(n)}\right)_j^k\right\} = 1 + \theta_{jk}/n + o(1/n), \quad k = 1, \ldots, 2p+2, \tag{4.2.1}$$

then $C_t^{(n)}(s,i) = O(|s|^p)$ *and*

$$C^{(n)}(t,s,i) \to C_t(s,i) = \frac{e^{-r(T-t)}}{\gamma_{T-t}(i)} E_{s,i}\left\{\Phi(S_{T-t})Z_{T-t}\right\},$$

where, by Lemma 3.1.3, C satisfies (3.2.1).

If in addition Φ *is almost everywhere differentiable with derivative* $\Phi'(s) = O(|s|^{p-1})$ *and* (4.2.1) *holds, then* $\nabla C_t^{(n)}(s,i) = O(|s|^{p-1})$ *and*

$$\begin{aligned} \nabla C^{(n)}(t,s,i) &\to \nabla C_t(s,i) \\ &= \frac{e^{-r(T-t)}}{\gamma_{T-t}(i)} D^{-1}(s) E_{s,i}\left\{\Phi'(S_{T-t})S_{T-t}Z_{T-t}\right\}. \end{aligned} \tag{4.2.2}$$

The proof is given in Appendix B.7.

Remark 4.2.3 *It is easy to check that* (4.2.1) *holds for* $p = \infty$ *for the regime-switching geometric Gaussian random walk.*

Remark 4.2.4 *The result on the convergence of the gradient on* $C^{(n)}$ *is comparable to a result of Broadie and Glasserman (1996) on the unbiased estimation of Greeks by Monte Carlo methods.*

Before stating the main approximation theorem,we need to study the convergence of $\alpha^{(n)}(t,s,i) = \alpha^{(n)}_{[nt/T]}(s,i)$.

Lemma 4.2.5 *Suppose that* Φ *is almost everywhere differentiable with derivative* $\Phi'(s) = O(|s|^{p-1})$ *and assume that* (4.2.1) *holds. Further assume that (iii) holds for all* $f \in C_p(\mathbb{R}^d)$. *Then* $\alpha^{(n)}(t,s,i) \to \alpha(t,s,i)$, *where* α *is given by* (3.2.3).

The proof is given in Appendix B.8.

Finally, one can state the main approximation result, namely the convergence of the portfolio value $V^{(n)}$ and the discrete time optimal strategy $\phi^{(n)}$, whose proof is given in Appendix B.9.

To that end, set $\phi^{(n)}(t) = \phi^{(n)}_{[nt/T]}$, and $V^{(n)}(t) = V^{(n)}_{[nt/T]}$.

Theorem 4.2.6 *Suppose that* $\Phi(s) = O(|s|^p)$, Φ *is almost everywhere differentiable with derivative* $\Phi'(s) = O(|s|^{p-1})$ *and* (4.2.1) *holds. Then*

$$\left(S^{(n)}, X^{(n)}, \tau^{(n)}, C^{(n)}, \alpha^{(n)}, V^{(n)}, \phi^{(n)}, G^{(n)}\right) \rightsquigarrow (S, X, \tau, C, \alpha, V, \phi, G),$$

with $G_t = e^{-rt}C_t(S_t, \tau_t) - V_t$, *where* V *and* ϕ *are given by* (3.2.4) *and* (3.2.5).

5 Example of application

In Rémillard et al. (2009), the authors analyzed the daily log-returns of the S&P 500 from January 1st 2007 to December 31st 2008, and concluded that a regime-switching geometric Gaussian random walk with 4 regimes was the best fit for that data set. Their estimated parameters are given in Tables 1–2.

Table 1: Parameter estimations of the daily log-returns using 4 regimes.

Regime	Mean	Variance	Stat. distr.	Prob. of next regime
1	-0.00500	0.002221	0.133	0.0084
2	-0.00134	0.000191	0.517	0.9850
3	0.00131	0.000126	0.113	4.2798e-006
4	0.00119	0.000014	0.237	0.0064

Table 2: Transition matrix Q for 4 regimes.

Regime	1	2	3	4
1	0.9842	0.0158	0	0
2	0.0043	0.9744	0	0.0213
3	0	0	0	1
4	0	0.0542	0.4754	0.4704

To find the associated parameters in continuous time (measured in years), one can multiply the mean and variance by 250 and set $\Lambda = 250(Q - I)$. Our aim is to price, using a regime-switching geometric Brownian motion, at-the-money call and put options with a maturity of 0.12 years (30 days), using an annual rate of 3% and a starting price of the underlying asset of 100. The continuous time corresponding parameters are given in Tables 3–4.

Table 3: Parameters for the continuous time case.

Regime	ψ	A	ρ	ℓ
1	-0.9724	0.5553	-1.8053	1.8096
2	-0.3111	0.0478	-7.1440	2.4370
3	0.3433	0.0315	9.9444	3.1151
4	0.2993	0.0035	76.9286	20.7130

Table 4: Generator Λ.

Regime	1	2	3	4
1	-3.9500	3.9500	0	0
2	1.0750	-6.4000	0	5.3250
3	0	0	-250.0000	250.0000
4	0	13.5500	118.8500	-132.4000

One can now evaluate γ_t and Λ_t. The graph of $-\Lambda_t(3,3)$, which is much larger than the others, is given in Figure 1. To simulate the process according to the algorithm in Appendix D, it follows that one can take $\lambda = 250$, using r instead of ψ, according to (3.2.10).

Based on equations (3.2.11)–(3.2.12), the results of the simulation, including the price of at-the-money call and put options, together with the value of $\phi_0 = \frac{d}{ds}C_0(s,i)$, are given in Table 5, using 1,000,000 repetitions and antithetic variables. Using the results of Table 1, one predicts that the next regime will be regime 2, having probability .98. Hence, from Table 5, the price of an at-the-money call option is 3.5034, while the price of an at-the-money put option is 3.1435. Note that the put-call parity principle is respected since the difference between the two prices is $0.3599 \approx 100 \times \left(1 - e^{-.03\times 30/250}\right) = 0.3606$. Furthermore, for the call

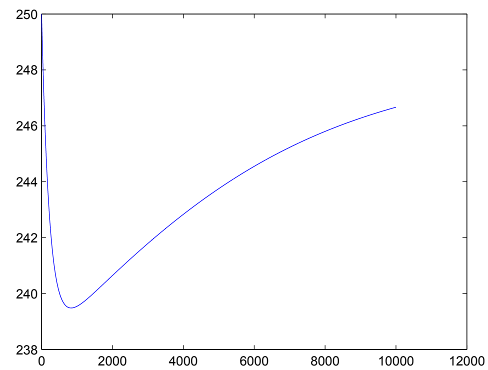

Figure 1: Graph of $-\Lambda_t(3,3)$ for $t \in [0, 0.12]$.

option, the initial number of risky asset is 0.5356, while for the put option, one should start by short-selling 0.4644 units.

Because one can evaluate C_t and ϕ_t for any t, one could do as proposed in Rémillard et al. (2009) and compare the optimal discrete hedging with the discretized version, i.e., by considering $\phi_{Tk/n}$ for $k = 1, \ldots, n$, as in the discretized version of the Black-Scholes model, using filtering to predict the regimes using information available previously.

Table 5: 95% confidence intervals for at-the-money price of calls and puts, together with initial investments, using 1,000,000 simulations.

	Call		Put	
Regime	Price	ϕ_0	Price	ϕ_0
1	9.3103 ± 0.0182	0.5524 ± 0.0004	8.9549 ± 0.0110	-0.4475 ± 0.0003
2	3.5034 ± 0.0069	0.5356 ± 0.0001	3.1435 ± 0.0055	-0.4644 ± 0.0001
3	2.6398 ± 0.0049	0.5380 ± 0.0002	2.2803 ± 0.0041	-0.4620 ± 0.0002
4	2.6469 ± 0.0049	0.5384 ± 0.0002	2.2874 ± 0.0042	-0.4616 ± 0.0002

6 Conclusion

In this paper we presented the optimal discrete hedging solution for a dynamic portfolio. If the underlying assets are Markovian or form a Markov process by adding a latent process, then the optimal hedging strategy depends on deterministic functions that can be approximated. We also find the optimal hedging strategy in the continuous case when the underlying assets are modeled by a regime-switching geometric Lévy process. For the regime-switching geometric Brownian motion, the optimal strategy can be deduced from a risk neutral measure. It is therefore natural to choose that risk neutral measure to be the one used in pricing contingent claims. Finally, it is shown that under appropriate HMM models, the optimal strategy in the discrete case converges to the one obtained in the continuous time when the number of hedging periods increases.

A Proof of Theorem 2.0.1

It is easy to check that a necessary a sufficient condition for $\left(V_0, \overrightarrow{\phi}\right)$ to minimize $E\left[\left\{G\left(V_0, \overrightarrow{\phi}\right)\right\}^2\right]$ is that $E\left\{G\left(V_0, \overrightarrow{\phi}\right)\right\} = 0$ and $E\left\{G\left(V_0, \overrightarrow{\phi}\right)\Delta_k | \mathcal{F}_{k-1}\right\} = 0$ for all $k = 1, \ldots, n$.

The necessity comes from the fact that for any event $A \in \mathcal{F}_{k-1}$, one must have

$$0 = \frac{d}{d\epsilon}\Big|_{\epsilon=0} E\left[\left\{G\left(V_0, \overrightarrow{\phi}\right) - \epsilon \mathbb{I}_A \Delta_k\right\}^2\right] = -2E\left\{G\left(V_0, \overrightarrow{\phi}\right)\Delta_k \mathbb{I}_A\right\},$$

which is equivalent to the condition $E\left\{G\left(V_0, \overrightarrow{\phi}\right)\Delta_k|\mathcal{F}_{k-1}\right\} = 0$, while the condition $E\left\{G\left(V_0, \overrightarrow{\phi}\right)\right\} = 0$ comes from the fact that for any θ, one must have

$$0 = \frac{d}{d\epsilon}\Big|_{\epsilon=0} E\left[\left\{G\left(V_0 + \epsilon\theta, \overrightarrow{\phi}\right)\right\}^2\right] = -2E\left\{G\left(V_0, \overrightarrow{\phi}\right)\right\}.$$

To see that the conditions are sufficient, it suffices to check that

$$E\left[\left\{G\left(V_0 + \theta_0, \overrightarrow{\phi + \psi}\right)\right\}^2\right] = E\left[\left\{G\left(V_0, \overrightarrow{\phi}\right)\right\}^2\right] + E\left\{\left(\theta_0 + \sum_{k=1}^{n} \psi_k^\top \Delta_k\right)^2\right\}.$$

The proof that $\overrightarrow{\phi}$ is the solution is based on the following proposition.

Proposition A.0.7 *For any $k = 1, \ldots, n$,*

$$E(V_n|\mathcal{F}_k) = V_k E(P_{k+1}|\mathcal{F}_k) + E\left\{\beta_n C(1 - P_{k+1})|\mathcal{F}_k\right\}. \tag{A.0.3}$$

Clearly, (A.0.3) holds true for $k = n$. To show that is holds for $k = n - 1$, note that $V_n = V_{n-1} + \phi_n^\top \Delta_n$, so

$$\begin{aligned} E(V_n|\mathcal{F}_{n-1}) &= V_{n-1} + E(\Delta_n|\mathcal{F}_{n-1})^\top \phi_n = V_{n-1} + b_n^\top A_n(a_n - V_{n-1}b_n) \\ &= V_{n-1}E(P_n|\mathcal{F}_{n-1}) + E\left\{\beta_n C(1 - P_n)|\mathcal{F}_{n-1}\right\}. \end{aligned}$$

Suppose now that (A.0.3) is true for $k = j$. We will prove that it is also true for $k = j - 1$. Now,

$$\begin{aligned} E(V_n|\mathcal{F}_j) &= V_j E(P_{j+1}|\mathcal{F}_j) + E\left\{\beta_n C(1 - P_{j+1})|\mathcal{F}_j\right\} \\ &= V_{j-1}E(P_{j+1}|\mathcal{F}_j) + \phi_j^\top \Delta_j E(P_{j+1}|\mathcal{F}_j) + E\left\{\beta_n C(1 - P_{j+1})|\mathcal{F}_j\right\} \\ &= V_{j-1}E(P_j|\mathcal{F}_j) + a_j^\top E(\Delta_j P_{j+1}|\mathcal{F}_j) + E\left\{\beta_n C(1 - P_{j+1})|\mathcal{F}_j\right\}, \end{aligned}$$

so

$$\begin{aligned} E(V_n|\mathcal{F}_{j-1}) &= V_{j-1}E(P_j|\mathcal{F}_{j-1}) + a_j^\top E(\Delta_j P_{j+1}|\mathcal{F}_{j-1}) + E\left\{\beta_n C(1 - P_{j+1})|\mathcal{F}_{j-1}\right\} \\ &= V_{j-1}E(P_j|\mathcal{F}_{j-1}) + a_j^\top A_j b_j + E\left\{\beta_n C(1 - P_{j+1})|\mathcal{F}_{j-1}\right\} \\ &= V_{j-1}E(P_j|\mathcal{F}_{j-1}) + E\left\{\beta_n C b_j^\top \Delta_j P_{j+1}|\mathcal{F}_{j-1}\right\} + E\left\{\beta_n C(1 - P_{j+1})|\mathcal{F}_{j-1}\right\} \\ &= V_{j-1}E(P_j|\mathcal{F}_{j-1}) + E\left\{\beta_n C(1 - P_j)|\mathcal{F}_{j-1}\right\}. \end{aligned}$$

This completes the proof of the proposition. ■

To complete the proof of theorem, note that it follows from (A.0.3) that for any $k = 0, \ldots, n$,

$$E\left\{G(\left(V_0, \overrightarrow{\phi}\right)|\mathcal{F}_k\right\} = E(\beta_n C P_{k+1}|\mathcal{F}_k) - V_k E(P_{k+1}|\mathcal{F}_k). \tag{A.0.4}$$

Now using (A.0.4), one has

$$E\left\{G(\left(V_0, \overrightarrow{\phi}\right)\Delta_k|\mathcal{F}_k\right\} = E(\beta_n C \Delta_k P_{k+1}|\mathcal{F}_k) - E(V_k \Delta_k P_{k+1}|\mathcal{F}_k),$$

so

$$\begin{aligned} E\left\{G(\left(V_0, \overrightarrow{\phi}\right)\Delta_k|\mathcal{F}_{k-1}\right\} &= E(\beta_n C\Delta_k P_{k+1}|\mathcal{F}_{k-1}) - E(V_k\Delta_k P_{k+1}|\mathcal{F}_{k-1}) \\ &= A_k(a_k - V_{k-1}b_k - \phi_k) = 0. \end{aligned}$$

Finally, it follows also from (A.0.4) that

$$E\left\{G(\left(V_0, \overrightarrow{\phi}\right)\right\} = E(\beta_n C P_1) - V_0 E(P_1) = 0.$$

B Proof of the main results

First, we need the following lemma.

Lemma B.0.8 *If $X_t = \beta_t S_t$, $M_t = \int_0^t \rho^\top(\tau_{u-}) D^{-1}(X_{u-}) dX_u$ and $Z = \mathcal{E}\{-M\}$. Further let*

$$\tilde{\mathcal{H}} f(s,i) = \mathcal{L}f(s,i) - f(s,i)\ell(i) - m^\top(i) D(s)\nabla f(s) - \rho^\top(i)\mathcal{K}_i f(s,i). \tag{B.0.5}$$

Then

$$f(S_t,\tau_t)Z_t = f(S_0,\tau_0) + \int_0^t Z_u \tilde{\mathcal{H}} f(S_u,\tau_u) du + \tilde{M}_t,$$

where $\tilde{M}$ is a martingale. In particular

$$\lim_{t\downarrow 0} \frac{1}{t}\left[E_{s,i}\left\{f(S_t,\tau_t)Z_t - f(s,i)\right\}\right] = \tilde{\mathcal{H}} f(s,i).$$

Proof: For any f in the domain of $\mathcal{L}$, there exists martingales $M^{(g)}$ and $M^{(f)}$ so that

$$S_t = s + \int_0^t D(S_u)\psi(\tau_u) du + M_t^{(g)}$$

and

$$f(S_t,\tau_t) = f(S_0,\tau_0) + \int_0^t \mathcal{L}f(S_u,\tau_u) du + M_t^{(f)}.$$

In addition, it follows from Lemma C.0.4 that there exists a martingale $M^{(f,g)}$ so that

$$\left[M^{(f)}, M^{(g)}\right]_t = M_t^{(f,g)} + \int_0^t \left\{\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\right\}(S_u,\tau_u) du. \tag{B.0.6}$$

It is easy to check that

$$Z_t = 1 - \int_0^t Z_u \ell(\tau_u) du - \int_0^t Z_{u-}\rho^\top(\tau_{u-}) D^{-1}(S_{u-}) dM_u^{(g)}.$$

It follows from Itô's formula (Theorem C.0.3) that there exists a martingale $\mathcal{M}$ so that

$$\begin{aligned} f(S_t,\tau_t)Z_t &= f(S_0,\tau_0) + \int_0^t Z_u\left\{\mathcal{L}f(S_u,\tau_u) - f(S_u,\tau_u)\ell(\tau_u)\right\} du + \mathcal{M}_t \\ &\quad - \int_0^t Z_{u-}\rho^\top(\tau_{u-}) D^{-1}(S_{u-}) d\left[M^{(f)}, M^{(g)}\right]_u. \end{aligned}$$

According to (B.0.6),

$$\left[M^{(f)}, M^{(g)}\right]_t = M_t^{(f,g)} + \int_0^t \left\{\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f)(S_u,\tau_u)\right\} du,$$

for some martingale $M^{(f,g)}$. As a result, there exists a martingale $\tilde{M}$ so that

$$\begin{aligned} f(S_t,\tau_t)Z_t &= f(S_0,\tau_0) + \int_0^t Z_u \left\{\mathcal{L}f(S_u,\tau_u) - f(S_u,\tau_u)\ell(\tau_u)\right\} du + \mathcal{M}'_t \\ &\quad - \int_0^t Z_u \rho^\top(\tau_{u-}) D^{-1}(S_u) \left\{\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\right\}(S_u,\tau_u) du \\ &= f(S_0,\tau_0) + \int_0^t Z_u \tilde{\mathcal{H}} f(S_u,\tau_u) du + \tilde{M}_t, \end{aligned}$$

where

$$\begin{aligned} \tilde{\mathcal{H}} f(s,i) &= \mathcal{L}f(s,i) - f(s,i)\ell(i) - \rho^\top(i) D^{-1}(s) \left\{\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\right\}(s,i) \\ &= \mathcal{L}f(s,i) - f(s,i)\ell(i) - \rho^\top(i) D^{-1}(s) \left\{\mathcal{L}_i(fg) - f\mathcal{L}_i g - g\mathcal{L}_i f\right\}(s,i), \end{aligned}$$

since g does not depend on τ. Finally, one can check that

$$\left\{\mathcal{L}_i(fg) - f\mathcal{L}_i g - g\mathcal{L}_i f\right\}(s,i) = D(s)\mathbb{A}(i)D(s)\nabla f(s,i) + D(s)\mathcal{K}_i f(s,i). \tag{B.0.7}$$

Hence

$$\tilde{\mathcal{H}} f(s,i) = \mathcal{L}f(s,i) - f(s,i)\ell(i) - m^\top(i)D(s)\nabla f(s) - \rho^\top(i)\mathcal{K}_i f(s,i).$$

The rest of the proof is easy. ■

B.1 Proof of Lemma 3.1.3

Proof: First,

$$Z_{t+h} = Z_h - \int_h^{t+h} Z_{u-} dM_u = Z_h - \int_0^t Z_{h+u-} d\tilde{M}_h(u)$$

where

$$\begin{aligned} M_h(t) &= M_{t+h} - M_h = -r\int_h^{t+h} \rho^\top(\tau_u)\mathbf{1} du + \int_h^{t+h} \rho^\top(\tau_{u-})D^{-1}(S_{u-})dS_u \\ &= -r\int_0^t \rho^\top(\tau_{h+u})\mathbf{1} du + \int_0^t \rho^\top(\tau_{h+u-})D^{-1}(S_{h+u-})dS_{h+u}. \end{aligned}$$

Setting

$$Z_{h,t} = 1 - \int_0^t Z_{h,u-} d\tilde{M}_h(u),$$

it follows from the uniqueness of solutions that $Z_{t+h} = Z_h Z_{h,t}$. Hence, since τ and (S,τ) are Markov processes, it follows that Z is a multiplicative functional. As a result, for $h > 0$ small enough,

$$\begin{aligned} v(t+h,s,i) &= \frac{1}{\gamma(t+h,s,i)} \left[E\left\{Z_h v(t,S_h,\tau_h)\gamma(t,\tau_h) | S_0 = s, \tau_0 = i\right\} - v(t,s,i)\gamma(t,i)\right] \\ &\quad - v(t,s,i)\left\{\frac{\gamma(t+h,s,i) - \gamma(t,i)}{\gamma(t+h,s,i)}\right\} + v(t,s,i). \end{aligned}$$

Consequently, using Lemma B.0.8, one may conclude that

$$\partial_t v_t(s,i) = \mathcal{H}_t v_t(s,i) = -\frac{\dot{\gamma}(t,i)}{\gamma(t,i)} v_t(s,i) + \frac{1}{\gamma(t,i)} \tilde{\mathcal{H}}(\gamma_t v_t)(s,i).$$

It is then easy to check that

$$\mathcal{H}_t f(s,i) = \mathcal{L}_i f(s,i) - m^\top(i)D(s)\nabla f(s) - \rho^\top(i)\mathcal{K}_i f(s,i)$$

$$+\frac{1}{\gamma(t,i)}\sum_{j=1}^{l}\Lambda_{ij}\gamma(t,j)\{f(s,j)-f(s,i)\}.$$

Because the solution to (B.0.5) is unique, $f \equiv 1$ entails that $v_t \equiv 1$, so $E(Z_t|S_0 = s, \tau_0 = i) = \gamma(t,i)$. Finally, from the multiplicative property of Z, one gets

$$E(Z_T|\mathcal{F}_t) = Z_t E_{S_t,\tau_t}(Z_{T-t}) = Z_t\gamma(T-t,\tau_t).$$

■

B.2 Proof of Lemma 3.2.1

By Lemma C.0.4 and (B.0.7), there exists a martingale $M^{(C,g)}$ so that

$$\begin{aligned}\left[M^{(C)}, M^{(g)}\right]_t &= M_t^{(C,g)} + \int_0^t \{\mathcal{L}(gC_t) - LC_t - C_t\mathcal{L}g\}(S_u,\tau_u)du \\ &= M_t^{(C,g)} + \int_0^t D(S_u)\{\mathbb{A}(\tau_u)D(S_u)\nabla C_u(S_u,\tau_u) + \mathcal{K}_{\tau_u}C_u(S_u,\tau_u)\}du.\end{aligned}$$

As a result,

$$X_t = s + \int_0^t e^{-ru}D(S_u)m(\tau_u)du + \int_0^t e^{-ru}dM_u^{(g)}$$

and

$$\begin{aligned}V_t &= C(0,s,i) + \int_0^t e^{-ru}\phi(u-,S_{u-},\tau_{u-})^\top dM_u^{(g)} \\ &\quad + \int_0^t e^{-ru}m(\tau_u)^\top D(S_u)\phi(u,S_u,\tau_u)du \\ &= C(0,s,i) + \int_0^t e^{-ru}\phi(u-,S_{u-},\tau_{u-})^\top dM_u^{(g)} \\ &\quad + \int_0^t e^{-ru}m(\tau_u)^\top D(S_u)\nabla C_u(S_u,\tau_u)du \\ &\quad + \int_0^t G_u\ell(\tau_u)du + \int_0^t e^{-ru}\rho(\tau_u)^\top\mathcal{K}_{\tau_u}C_u(S_u,\tau_u)du.\end{aligned}$$

Next,

$$\begin{aligned}e^{-rt}C(t,S_t,\tau_t) &= C(0,s,i) + \int_0^t e^{-ru}\{\partial_u C(u,S_u) - rC(u,S_u)\}du \\ &\quad + \int_0^t e^{-ru}\mathcal{L}C(u,S_u,\tau_u)du + \int_0^t e^{-ru}dM_u^{(C)}.\end{aligned}$$

Therefore, one obtains, using (3.2.1),

$$\begin{aligned}G_t &= -\int_0^t e^{-ru}\mathcal{H}_{T-u}C_u(S_u,\tau_u)du + \int_0^t e^{-ru}\mathcal{L}C_u(S_u,\tau_u)du + \int_0^t e^{-ru}dM_u^{(C)} \\ &\quad - \int_0^t e^{-ru}m(\tau_u)^\top D(S_u)\nabla C_u(S_u,\tau_u)du - \int_0^t e^{-ru}\phi(u-,S_{u-},\tau_{u-})^\top dM_u^{(g)} \\ &\quad - \int_0^t G_u\ell(\tau_u)du - \int_0^t e^{-ru}\rho(\tau_u)^\top\mathcal{K}_{\tau_u}C_u(S_u,\tau_u)du\end{aligned}$$

$$
\begin{aligned}
&= \int_0^t e^{-ru} dM_u^{(C)} - \int_0^t e^{-ru} \phi(u-, S_{u-}, \tau_{u-})^\top dM_u^{(g)} \\
&- \int_0^t \ell(\tau_u) G_u du + \int_0^t e^{-ru} \left\{ (\Lambda - \Lambda_{T-u}) C_{u,S_u} \right\} (\tau_u) du.
\end{aligned}
$$

■

B.3 Proof of Lemma 3.2.2

By Lemma 3.2.1 and Itô's formula in Theorem C.0.3, one gets

$$
\begin{aligned}
G_t f_t(S_t, \tau_t) &= \int_0^t f_u(S_{u-}, \tau_{u-}) dG_u + \int_0^t \partial_u f_u(S_u, \tau_u) G_u du \\
&\quad + \int_0^t \mathcal{L} f_u(S_u, \tau_u) G_u du + \int_0^t G_{u-} dM_u^{(f)} + \left[G, M^{(f)} \right]_t \\
&= \int_0^t f_u(S_{u-}, \tau_{u-}) dM_u^{(C)} + \int_0^t G_{u-} dM_u^{(f)} + \left[G, M^{(f)} \right]_t \\
&\quad - \int_0^t e^{-ru} f_u(S_{u-}, \tau_{u-}) \phi_u^\top dM_u^{(g)} \\
&\quad + \int_0^t e^{-ru} f_u(S_u, \tau_u) \left\{ (\Lambda - \Lambda_{T-u}) C_{u,S_u} \right\} (\tau_u) du \\
&\quad + \int_0^t \left\{ \mathcal{L} f_u(S_u, \tau_u) + \partial_u f_u(S_u, \tau_u) - f_u(S_u, \tau_u) \ell(\tau_u) \right\} G_u du,
\end{aligned}
$$

where the martingale $M^{(f)}$ is defined by

$$
M_t^{(f)} = f_t(S_t, \tau_t) - f_0(s, i) - \int_0^t \partial_u f_u(S_u, \tau_u) du - \int_0^t \mathcal{L} f_u(S_u, \tau_u) du
$$

and where, by Lemmas 3.2.1 and C.0.4, one has

$$
\begin{aligned}
\left[G, M^{(f)} \right]_t &= \int_0^t e^{-ru} d \left[M^{(C)}, M^{(f)} \right]_u \\
&\quad - \int_0^t e^{-ru} \phi_u^\top d \left[M^{(g)}, M^{(f)} \right]_u \\
&= \int_0^t e^{-ru} dM_u^{(C,f)} - \int_0^t e^{-ru} \phi_u^\top dM_u^{(f,g)} \\
&\quad + \int_0^t e^{-ru} \{ \mathcal{L}(f_u C_u) - f_u \mathcal{L}(C_u) - C_u \mathcal{L}(f_u) \} (S_u, \tau_u) du \\
&\quad - \int_0^t e^{-ru} \phi_u^\top \{ \mathcal{L}(f_u g) - f_u \mathcal{L}(g) - g \mathcal{L}(f_u) \} (S_u, \tau_u) du.
\end{aligned}
$$

As a result, if $f_t(s, i) = \gamma_{T-t}(i)$, then

$$
\mathcal{L} f_u(s, i) + \partial_u f_u(s, i) - f_u(s, i) \ell(i) \equiv 0,
$$

and

$$
\mathcal{N}(f_u, g)(s, i) = \{ \mathcal{L}(f_u g) - f_u \mathcal{L}(g) - g \mathcal{L}(f_u) \} (s, i) \equiv 0
$$

by (B.0.7). Moreover

$$
\mathcal{N}(f_u, C_u)(s, i) = \left\{ \mathcal{L}(f_u C_u) - f_u \mathcal{L}(C_u) - C_u \mathcal{L}(f_u) \right\} (s, i)
$$

$$\begin{aligned} &= -\gamma_{T-u}(i)\Lambda C_{u,s}(i) - C_{u,s}(i)\Lambda\gamma_{T-u}(i) + \Lambda(\gamma_{T-u}C_{u,s})(i) \\ &= -f_u(s,i)\left\{(\Lambda - \Lambda_{T-u})C_{u,s}\right\}(i), \end{aligned}$$

since

$$\Lambda(\gamma_t h)(i) - \gamma_t(i)\Lambda h(i) - h(i)\Lambda\gamma_t(i) = \gamma_t(i)(\Lambda_t - \Lambda)h(i), \qquad i = 1,\ldots,l.$$

Hence $G_t\gamma(T-t,\tau_t)$ is a martingale with initial value 0 and terminal value G_T, since $\gamma_0 = \mathbf{1}$. Hence $E(G_T) = 0$.

Next, take $f_t(s,i) = s_k\gamma_{T-t}(i)$, for a given $k \in \{1,\ldots,d\}$. Then

$$\{\mathcal{L}(f_u g) - f_u\mathcal{L}(g) - g\mathcal{L}(f_u)\}(s,i) = f_u(s,i)D(s)\mathbb{A}(i)e_k$$

by (B.0.7), with $(e_k)_j = I_{jk}$. Furthermore,

$$\mathcal{L}f_u(s,i) + \partial_u f_u(s,i) - f_u(s,i)\ell(i) = f_u(s,i)\psi_k(i)$$

and

$$\begin{aligned} \mathcal{N}(f_u, C_u)(s,i) &= \left\{\mathcal{L}(f_u C_u) - f_u\mathcal{L}(C_u) - C_u\mathcal{L}(f_u)\right\}(s,i) \\ &= f_u(s,i)e_k^\top D(s)\nabla C_u(s,i) + f_u(s,i)e_k^\top \mathcal{K}_i C_u(s,i) \\ &\quad - f_u(s,i)\left\{(\Lambda - \Lambda_{T-u})C_{u,s}\right\}(i). \end{aligned}$$

Therefore, setting $R_t = f_t(S_t,\tau_t)G_t$, one concludes that

$$R_t - r\int_0^t R_u du$$

is a martingale. Hence , so is $e^{-rt}R_t = X_t G_t\gamma_{T-t}(\tau_t)$.

Finally, to prove (3.2.8), note that $E(G_T|\mathcal{F}_t) = \gamma_{T-t}(\tau_t)G_t$ and $E(G_T|\mathcal{F}_u) = \gamma_{T-u}(\tau_u)G_u$. Therefore

$$E\{G_T(X_t - X_u)|\mathcal{F}_u\} = E\{G_t\gamma_{T-t}(\tau_t)X_t - G_u\gamma_{T-u}(\tau_u)X_u|\mathcal{F}_u\} = 0,$$

since we just proved that $X_t G_t\gamma_{T-t}(\tau_t)$ is a martingale. That completes the proof. ■

B.4 Proof of Theorem 3.2.3

Proof: It follows from (3.2.8) that for any $B \in \mathcal{F}_u$, and any $u \le v \le T$,

$$E\left(G_T\int_0^T \psi_t^\top dX_t\right) = 0,$$

where ψ is the predictable process given by $\psi_t = 1_B 1_{(u,v]}(t)$. Therefore, using properties of stochastic integrals, one may conclude that for any predictable process ψ so that $\int_0^T \psi_t^\top dX_t$ is square integrable, one gets

$$E\left(G_T\int_0^T \psi_t^\top dX_t\right) = 0.$$

Hence

$$E\left[\left\{e^{-rT}\Phi(S_T) - \int_0^T \psi_t^\top dX_t\right\}^2\right] = E\left(G_T^2\right) + E\left[\left\{\int_0^T (\phi_t - \psi_t)^\top dX_t\right\}^2\right],$$

because $E\left\{G_T\int_0^T(\phi_t - \psi_t)^\top dX_t\right\} = 0$. ■

B.5 Proof of Theorem 4.1.1

Without loss of generality, one may suppose that $T = 1$.

Set $T_n f(s,i) = E\left\{ f\left(S_1^{(n)}, \tau_1^{(n)}\right) \middle| S_0^{(n)} = s, \tau_0^{(n)} = i \right\}$.

According to Ethier and Kurtz (1986)[Theorem I.6.5, Theorem IV.2.6], it suffices to prove that

$$n(T_n - I)f(s,i) \to \mathcal{L}f(s,i),$$

uniformly on $[0,\infty)^d \times \{1,\ldots,l\}$ for all f so that f_i is infinitely differentiable and has compact support. Consequently, it is sufficient to show that

$$n(T_n - I)f(s,i) \to \mathcal{L}f(s,i),$$

uniformly on every compact subset of $[0,\infty)^d \times \{1,\ldots,l\}$, for all functions f that are bounded and such that f_i is twice differentiable with bounded continuous derivatives.

First, note that $S_1^{(n)} = s + \left(e^{r/n} - 1\right)s + e^{r/n} D(s)\xi_1^{(n)} = D(s)\left(\mathbf{1} + \zeta_1^{(n)}\right)$, where $\zeta_1^{(n)} = \left(e^{r/n} - 1\right)\mathbf{1} + e^{r/n}\xi_1^{(n)}$.

It follows from assumption (ii) that $\left|\xi_1^{(n)}\right| \stackrel{Pr}{\to} 0$, as $n \to \infty$.

Hence, for any $\delta > 0$, $\sup_{|s|\le\delta} \left|S_1^{(n)} - s\right| \stackrel{Pr}{\to} 0$, as $n \to \infty$.

Since

$$\begin{aligned} n(T_n - I)f(s,i) &= n\sum_{j=1}^{l} Q_{ij}^{(n)} \mathbb{E}_j\left\{ f\left(S_1^{(n)}, j\right) - f(s,i)\right\} \\ &= n\sum_{j=1}^{l} (Q^{(n)} - I)_{ij} \mathbb{E}_j\left\{ f\left(S_1^{(n)}, j\right) - f(s,i)\right\} \\ &\quad + n\mathbb{E}_i\left\{ f\left(S_1^{(n)}, i\right) - f(s,i)\right\} \end{aligned}$$

and since f is continuous and bounded, one gets that as $n \to \infty$,

$$n\sum_{j=1}^{l} (Q^{(n)} - I)_{ij} \mathbb{E}_j\left\{ f\left(S_1^{(n)}, j\right) - f(s,i)\right\} \to \sum_{j=1}^{l} \Lambda_{ij}\left\{ f(s,j) - f(s,i)\right\},$$

uniformly on $|s| \le \delta$.

Suppose now that f does not not depend on i. To simplify, just ignore i, i.e., suppose there is no regime-switching. It only remains to show that uniformly in $|s| \le \delta$, $n\mathbb{E}_i\left\{ f\left(S_1^{(n)}\right) - f(s)\right\} \to \mathcal{L}f(s)$.

Next, assumption (i-iii) imply that as $n \to \infty$, $nE\left(\zeta_1^{(n)}\right) \to r\mathbf{1} + m = \psi$, $nE\left(\zeta_1^{(n)} \zeta_1^{(n)\top}\right) \to \mathbb{A}$, and

$$nE\left\{ f\left(\zeta_1^{(n)}\right)\right\} \to \int_{\mathbb{R}^d\setminus\{0\}} f(y)\tilde{\nu}(dy).$$

Set

$$h_s(y) = f\left\{D(s)(\mathbf{1} + y)\right\} - f(s) - y^\top D(s)\nabla f(s) - \frac{1}{2} y^\top D(s) H_f(s) D(s) y,$$

where H_f is the Hessian matrix of f at s.

It follows from assumption (iii) that uniformly on $|s| \le \delta$, $h_s \in C_2(\mathbb{R}^d)$.

As a result, uniformly on $|s| \le \delta$,

$$\begin{aligned} n(T_n - I)f(s) &= nE\left[f\left\{s + D(s)\zeta_1^{(n)}\right\} - f(s)\right] \\ &= nE\left\{h_s\left(\zeta_1^{(n)}\right)\right\} + nE\left(\zeta_1^{(n)}\right)^\top D(s)\nabla f(s) \\ &\quad + \frac{n}{2}\text{Trace}\left\{E\left(\zeta_1^{(n)}\zeta_1^{(n)\top}\right)D(s)H_f(s)D(s)\right\} \\ &\to \psi^\top D(s)\nabla f(s) + +\frac{1}{2}\text{Trace}\left\{\mathbb{A}D(s)H_f(s)D(s)\right\} \\ &\quad + \int_{\mathbb{R}^d\setminus\{0\}} h_s(y)\tilde{\nu}(dy) \\ &= \mathcal{L}f(s), \end{aligned}$$

since $\mathbb{A} = a + a_{\tilde{\nu}}$. ■

B.6 Proof of Lemma 4.2.1

First, remark that

$$\gamma_k^{(n)} = F_k^{(n)} \cdots F_n^{(n)}\mathbf{1},$$

where $F_k^{(n)} = Q^{(n)} - D\left(\rho_{k+1}^{(n)}\right)Q^{(n)}D\left(\mu^{(n)}\right)$.

It follows from Proposition 2.1.3 that $\min_{1\le k\le n}\min_{1\le i\le l}\gamma_k^{(n)}(i) > 0$ and $\rho_{k+1}^{(n)} = \rho + O(1/n)$ uniformly in k, by formula (2.1.1), since

$$n\sum_{j=1}^{l} Q_{ij}^{(n)}\gamma_k^{(n)}(j)B^{(n)}(j) = \gamma_k^{(n)}(i)\mathbb{A}(i) + O(1/n)$$

and

$$n\sum_{j=1}^{l} Q_{ij}^{(n)}\gamma_k^{(n)}(j)\mu^{(n)}(j) = \gamma_k^{(n)}(i)m(i) + O(1/n).$$

As a result, $n\left(F_k^{(n)} - I\right) \to \Lambda - D(\ell)$, so

$$\gamma_n(t) = \gamma_{n+1-[nt/T]}^{(n)} \to \gamma(t) = e^{t\{\Lambda - D(\ell)\}}\mathbf{1} = \gamma(t).$$

■

B.7 Proof of Theorem 4.2.2

For simplicity we do the proof with $d = 1$, the case $d > 1$ being similar.

First, note that by assumptions of Theorem 4.1.1,

$$E\left(X_k^{(n)}\right) \le E\left(X_{k-1}^{(n)}\right)(1 + \delta/n), \qquad E\left({X_k^{(n)}}^2\right) \le E\left({X_{k-1}^{(n)}}^2\right)(1 + \delta/n),$$

for some $\delta > 0$. Hence, for any $k = 1, \ldots, n$,

$$E\left(X_k^{(n)}\right) \le s(1 + \delta/n)^n \to se^\delta, \quad E\left({X_k^{(n)}}^2\right) \le s^2(1 + \delta/n)^n \to s^2e^\delta.$$

As a result, the sequence of semimartingales $X_t^{(n)}$ is P-UT in the sense of Jacod and Shiryaev (2003), meaning that the sequence $H \cdot X^{(n)}$ is tight uniformly in $H^{(n)}$, for all $\mathcal{F}^{(n)}$-predictable process $H^{(n)}$ bounded by 1. To show that it is indeed the case, note that

$$\begin{aligned} H^{(n)} \cdot X^{(n)} &= \sum_{k=1}^{n} H_k^{(n)} \left(X_k^{(n)} - X_{k-1}^{(n)} \right) \\ &= \sum_{k=1}^{n} H_k^{(n)} X_{k-1}^{(n)} (\xi_k^{(n)} - \mu_{k-1}^{(n)}) + \sum_{k=1}^{n} H_k^{(n)} X_{k-1}^{(n)} \mu_{k-1}^{(n)}, \end{aligned}$$

where

$$\mu_{k-1}^{(n)} = E\left(\xi_k^{(n)} | \mathcal{F}_{k-1}^{(n)} \right) = \sum_{j=1}^{l} Q_{ij}^{(n)} \mu^{(n)}(j) = m(i)/n + o(1/n),$$

provided $\tau_{k-1}^{(n)} = i$. Hence there exists $\delta_1 > 0$ so that $|\mu^{(n)}| \le \delta_1/n$. Similarly, by choosing δ_1 large enough, we may suppose that

$$E\left\{ \left(\xi_k^{(n)} - \mu_{k-1}^{(n)})^2 \right) | \mathcal{F}_{k-1}^{(n)} \right\} \le \delta_1/n.$$

Hence, for any $H^{(n)}$,

$$E\left\{ \left| \sum_{k=1}^{n} H_k^{(n)} X_{k-1}^{(n)} \mu_{k-1}^{(n)} \right| \right\} \le \sum_{k=1}^{n} E(X_{k-1}^{(n)}) \delta_1/n \le s\delta_1 (1+\delta/n)^n \to s\delta_1 e^{\delta}.$$

Thus $\sum_{k=1}^{n} H_k^{(n)} X_{k-1}^{(n)} \mu_{k-1}^{(n)}$ is P-UT. Finally,
$E\left[\left\{ \sum_{k=1}^{n} H_k^{(n)} X_{k-1}^{(n)} \left(\xi_k^{(n)} - \mu_{k-1}^{(n)}) \right) \right\}^2 \right]$

$$\begin{aligned} &= E\left[\sum_{k=1}^{n} {H_k^{(n)}}^2 {X_{k-1}^{(n)}}^2 E\left\{ \left(\xi_k^{(n)} - \mu_{k-1}^{(n)} \right)^2 | \mathcal{F}_{k-1}^{(n)} \right\} \right] \\ &\le s^2 \delta_1 (1+\delta/n)^n \to s^2 \delta_1 e^{\delta}, \end{aligned}$$

completing the proof that $X^{(n)}$ is P-UT.

It is easy to check that

$$Z^{(n)}(t) = 1 - \int_0^t Z^{(n)}(u-) dM^{(n)}(u),$$

where $M^{(n)}$ can also be written as

$$M^{(n)}(t) = \int_0^t b^{(n)}(u-)^\top dX^{(n)}(u),$$

with $b^{(n)}(u) = D^{-1}\left\{ X^{(n)}(u) \right\} \rho^{(n)} \left\{ u, \tau^{(n)}(u) \right\}$.

Because the sequence of semimartingales $X^{(n)}$ is P-UT, and $\left(S^{(n)}, X^{(n)}, \tau^{(n)} \right) \rightsquigarrow (S, X, \tau)$ with $X_t = e^{-rt} S_t$, one may apply Rubenthaler (2003)[Theorem 4.3] (see also the original results in Słomiński (1989), Mémin and Słomiński (1991), Kurtz and Protter (1991a) and Kurtz and Protter (1991b)), to conclude that $\left(S^{(n)}, X^{(n)}, \tau^{(n)}, M^{(n)} \right) \rightsquigarrow (S, X, \tau, M)$, where

$$M_t = \int_0^t \rho^\top (\tau_{u-}) D^{-1}(X_{u-}) dX_u.$$

Moreover, it is easy to check that $M^{(n)}$ is P-UT, so now one can invoke Rubenthaler (2003)[Theorem 4.4] to conclude that

$$\left(S^{(n)}, X^{(n)}, \tau^{(n)}, M^{(n)}, Z^{(n)} \right) \rightsquigarrow (S, X, \tau, M, Z),$$

where $Z = \mathcal{E}\{-M\}$.

Note that if $E\left({\zeta_1^{(n)}}^k\right) = 1 + \theta_k/n + o(1/n)$, for all $k = 1, \ldots, 2p+2$, then

$$E\left({S_j^{(n)}}^{2p} {Z_j^{(n)}}^2\right) \le (1 + \delta_p/n)^j,$$

for some $\delta_p > 0$. As a result ${S_{T-t}^{(n)}}^p Z_{T-t}^{(n)}$ is uniformly integrable.

Next, if $\Phi(s) = O(|s|^p)$, the sequence $\Phi\left(S_{T-t}^{(n)}\right)$ is uniformly integrable.

Since

$$e^{-r[nt/T]T/n} C^{(n)}(t,s,i) = \frac{e^{-rT}}{\gamma^{(n)}(T-t,i)} E_{s,i}\left[\Phi\left\{S^{(n)}(T-t)\right\} Z^{(n)}(T-t)\right],$$

one can use Lemma 4.2.1 and the fact that $\left(S^{(n)}, X^{(n)}, \tau^{(n)}, M^{(n)}, Z^{(n)}\right) \rightsquigarrow (S, X, \tau, M, Z)$, to conclude that

$$C^{(n)}(t,s,i) \to C(t,s,i) = e^{-r(T-t)} E_{s,i}\left\{\Phi(S_{T-t}) Z_{T-t}\right\}/\gamma(T-t,i),$$

where, by Lemma 3.1.3, C satisfies (3.2.1).

Finally, note that $Z^{(n)}$ does not depend on $S_0^{(n)}$. As a result, if Φ is almost everywhere differentiable, with derivative $\Phi'(s) = O(|s|^{p-1})$, one gets

$$\nabla C_t^{(n)}(s,i) = \frac{e^{-rT + r[nt/T]T/n}}{\gamma^{(n)}(T-t,i)} D^{-1}(s) E_{s,i}\left[\Phi'\left\{S^{(n)}(T-t)\right\} S^{(n)}(T-t) Z^{(n)}(T-t)\right]$$

which converges to

$$\frac{e^{-r(T-t)}}{\gamma(T-t,i)} D^{-1}(s) E_{s,i}\left\{\Phi'\left(S_{T-t}\right) S_{T-t} Z_{T-t}\right\} = \nabla C_t(s,i),$$

because$\Phi'\left(S_{T-t}^{(n)}\right) S_{T-t}^{(n)} Z_{T-t}^{(n)}$ is uniformly integrable. ■

B.8 Proof of Lemma 4.2.5

First, note that

$$\begin{aligned} \alpha_k^{(n)}(s,i) &= e^{-rT/n} D^{-1}(s) \left\{\sum_{j=1}^{l} Q_{ij}^{(n)} \gamma_{k+1}^{(n)}(j) B^{(n)}(j)\right\}^{-1} \sum_{j=1}^{l} Q_{ij}^{(n)} \gamma_{k+1}^{(n)}(j) \\ &\quad \times \mathbb{E}_j\left[C_k^{(n)}\left\{e^{rT/n} D(s)\left(\mathbf{1} + \xi_1^{(n)}\right), j\right\} \xi_1^{(n)}\right]. \end{aligned}$$

Because of assumptions (i–iii) and Lemma 4.2.1, it suffices to show that

$$n\mathbb{E}_i\left[C_k^{(n)}\left\{(D(s)\left(\mathbf{1} + \xi_1^{(n)}, i\right)\right\} \xi_1^{(n)}\right] \to \mathbb{A}(i) D(s) \alpha_t(s,i).$$

By Theorem 4.2.2, one only needs to prove that

$$n\mathbb{E}_i\left[\left[C_k^{(n)}\left\{(D(s)\left(\mathbf{1} + \xi_1^{(n)}, i\right)\right\} - C_t^{(n)}(s,i)\right] \xi_1^{(n)}\right]$$

converges to

$$\mathbb{A}(i) D(s) \nabla C_t(s,i) + \mathcal{K}_i C_t(s,i),$$

which in turn is true if one can show that

$$n\mathbb{E}_i\left\{f^{(n)}\left(\xi_n\right)\right\} \to \mathcal{K}_i C_t(s,i),$$

using Theorem 4.2.2 for the convergence of the gradient, where

$$f^{(n)}(y) = y\left\{C_t^{(n)}\{D(s)(\mathbf{1}+y), i\} - C_t^{(n)}(s,i) - y^\top D(s)\nabla C_t^{(n)}(s,i)\right\}.$$

Now to $f^{(n)} \in C_2(\mathbb{R}^d)$ and is uniformly $O(|y|^2)$ in n, and $f^{(n)} \to f \in C_2(\mathbb{R}^d)$ defined by

$$f(y) = C_t\{D(s)(\mathbf{1}+y), i\} - C_t(s,i) - y^\top D(s)\nabla C_t(s,i)$$

by Theorem 4.2.2. Consequently, by dominated convergence and assumption (iii), one may conclude that

$$n\mathbb{E}_i\left\{f^{(n)}\left(\xi_n\right)\right\} \to \int_{\mathbb{R}^d\setminus\{0\}} f(y)\tilde{\nu}_i(dy) = \mathcal{K}_i C_t(s,i).$$

■

B.9 Proof of Theorem 4.2.6

First, note that

$$\begin{aligned} V^{(n)}(t) &= V^{(n)}(0) + \int_0^t \alpha^{(n)}\{u-, S^{(n)}(u-), \tau^{(n)}(u-)\}^\top dX^{(n)}(u) \\ &\quad - \int_0^t V^{(n)}(u-)dM^{(n)}(u) \\ &= V^{(n)}(0) + \int_0^t \phi_u^{(n)\top} dX^{(n)}(u). \end{aligned}$$

Because $X^{(n)}$ is P-UT, it follows from Theorem 4.2.2, Lemma 4.2.5 and Rubenthaler (2003)[Theorem 4.4] that

$$\left(S^{(n)}, X^{(n)}, \tau^{(n)}, C^{(n)}, \alpha^{(n)}, V^{(n)}\right) \rightsquigarrow (S, X, \tau, C, \alpha, V),$$

with V satisfying (3.2.4). Hence

$$\left(S^{(n)}, X^{(n)}, \tau^{(n)}, C^{(n)}, \alpha^{(n)}, V^{(n)}, \phi^{(n)}, G^{(n)}\right) \rightsquigarrow (S, X, \tau, C, \alpha, V, \phi, G),$$

where ϕ satisfies (3.2.5) and $G_t = e^{-rt}C_t(S_t, \tau_t) - V_t$.

■

C Stochastic calculus for semimartingales

Definition of quadratic covariation. Let X and Y be semimartingales. The quadratic variation of X, denoted $[X, X]$, is defined by

$$[X,X]_t = X_t^2 - 2\int_0^t X_{u-}dX_u,$$

and the quadratic covariation of X, Y, denoted $[X, Y]$, is defined by

$$[X,Y]_t = X_tY_t - \int_0^t X_{u-}dY_u - \int_0^t Y_{u-}dX_u.$$

Note that the operation $(X, Y) \mapsto [X, Y]$ is bilinear and symmetric and

$$2[X,Y] = [X+Y, X+Y] - [X,X] - [Y,Y].$$

The following result is proved in (Protter, 2004, Theorem 28, page 75).

Theorem C.0.1 *Let X and Y be two semimartingales, and let $H, K \in \mathbb{L}$. Then*

$$[H \cdot X, K \cdot Y]_t = \int_0^t H_u K_u d[X,Y]_u.$$

The following result is proved in (Protter, 2004, Theorem 37, page 84).

Theorem C.0.2 *Let X be a semimartingale with $X_0 = 0$. Then there exists a (unique) semimartingale Z that satisfies the equation $Z_t = 1 + \int_0^t Z_{u-} dX_u$. Z is given by*

$$Z_t = \mathcal{E}_t(X) = \exp\left\{X_t - \frac{1}{2}[X,X]_t\right\} \prod_{0<u\le t} (1 + \Delta X_u) \exp\left\{-\Delta X_u + \frac{1}{2}(\Delta X_u)^2\right\}$$

where the infinite product converges.

Theorem C.0.3 (Itô's formula)

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \partial_t f(u, X_u) du + \int_0^t \nabla_x f(u, X_{u-})^\top dX_u \\ &\quad + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \partial_{x_i} \partial_{x_j} f(u, X_{u-}) d[X^i, X^j]_u^c \\ &\quad + \sum_{u \le t} \{f(u, X_u) - f(u, X_{u-}) - \nabla_x f(u, X_{u-})^\top \Delta X_u\} \end{aligned}$$

Lemma C.0.4 *Suppose that $M_t^{(f)} = f(x_t) - f(x_0) - \int_0^t \mathcal{L}f(x_u)du$ and $M_t^{(g)} = g(x_t) - g(x_0) - \int_0^t \mathcal{L}g(x_u)du$ are martingales. Then*

$$M_t^{(f,g)} = \left[M^{(f)}, M^{(g)}\right]_t - \int_0^t \{\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\}(x_u)du$$

is a martingale. Moreover

$$M_t^{(f,g)} = M_t^{(fg)} - \int_0^t \left\{g(x_u) - M_u^{(g)}\right\} dM_u^{(f)} - \int_0^t \left\{f(x_u) - M_u^{(f)}\right\} dM_u^{(g)}.$$

Proof: By definition of the quadratic covariation,

$$M_t^{(f)} M_t^{(g)} = \int_0^t M_u^{(f)} dM_u^{(g)} + \int_0^t M_u^{(g)} dM_u^{(f)} + \left[M^{(f)}, M^{(g)}\right]_t,$$

and

$$M_t^{(fg)} = f(x_t)g(x_t) - f(x_0)g(x_0) - \int_0^t \mathcal{L}(fg)(x_u)du,$$

is a martingale. Setting $Vf_t = \int_0^t \mathcal{L}f(x_u)du$ and $Vg_t = \int_0^t \mathcal{L}g(x_u)du$, it follows that

$$M_t^{(f)} Vg_t = \int_0^t \mathcal{L}g(x_u) M_u^{(f)} du + \int_0^t Vg_u dM_u^{(f)},$$

and

$$M_t^{(g)} Vf_t = \int_0^t \mathcal{L}f(x_u) M_u^{(g)} du + \int_0^t Vf_u dM_u^{(g)}.$$

Now, by definition,

$$M_t^{(f)} M_t^{(g)} = f(x_t)g(x_t) - g(x_0)f(x_t) - f(x_t)Vg_t - f(x_0)g(x_t) + f(x_0)g(x_0)$$

$$
\begin{aligned}
& +f(x_0)Vg_t - g(x_t)Vf_t + g(x_0)Vf_t + Vf_tVg_t \\
= \quad & M_t^{(fg)} + \int_0^t \mathcal{L}(fg)(x_u)du - M_u^{(f)}Vg_u - M_u^{(g)}Vf_u \\
& -g(x_0)M_t^{(f)} - g(x_0)Vf_t - f(x_0)Vg_t - Vf_tVg_t \\
= \quad & M_t^{(f,g)} + \int_0^t \mathcal{L}(fg)(x_u)du - g(x_0)Vf_t - f(x_0)Vg_t - Vf_tVg_t \\
& - \int_0^t \mathcal{L}f(x_u)M_u^{(g)}du - \int_0^t \mathcal{L}f(x_u)M_u^{(g)}du \\
= \quad & M^{(f,g)} + \int_0^t \{\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\}(x_u)du.
\end{aligned}
$$

Hence the result. ■

D Construction of a non-homogeneous switching Lévy process

Suppose that the Markov chain τ_t is non-homogeneous, with generator Λ_t, i.e., for any $j \neq i$,

$$
\lim_{h \downarrow 0} \frac{1}{h} P(\tau_{t+h} = j | \tau_t = i) = (\Lambda_t)_{ij}.
$$

For a given T, assume that one can find $0 < \lambda$ so that

$$
\max_{1 \le i \le l} \sup_{0 \le t \le T} -(\Lambda_t)_{ii} \le \lambda.
$$

To construct a switching Lévy process L on $[0,T]$ based on the Markov chain τ and Lévy processes L_i, do the following:

- Generate $N \sim \text{Poisson}(\lambda T)$.
- If $N = n$, generate independent uniform variates $U_1, \ldots, U_n$ and order then. Denote the resulting sample by $U_{n:1}, \ldots, U_{n:n}$, with $U_{n:1} < U_{n:2} \cdots < U_{n:n}$. That can be done by generating $n+1$ independent exponential variates $E_1, \ldots, E_{n+1}$ and by setting

$$
U_{n:i} = \frac{\sum_{j=1}^{i} E_j}{\sum_{j=1}^{n+1} E_j}, \qquad i = 1, \ldots, n.
$$

- Set $t_i = T \times U_{n:i}$, $i = 1, \ldots, n$. These values are the possible switching regime times. Further set $t_0 = 0$ and $t_{n+1} = T$.
- For $k = 1, \ldots, n$, if $\tau_{t_{k-1}} = i$, then $\tau_{t_k} = j$ with probability $P_{k,ij}$, where

$$
P_{k,ij} = (\Lambda_{t_k})_{ij}/\lambda, \quad j \neq i, P_{k,ii} = 1 + (\Lambda_{t_k})_{ii}/\lambda.
$$

- For $k = 0, \ldots, n$, and $t \in (t_k, t_{k+1}]$, set

$$
L_t = L_{t_k} + L_{t,\tau_k} - L_{t_k,\tau_k}.
$$

In particular, if $\delta_k = t_k - t_{k-1}$, $k = 1, \ldots, n+1$, then

$$
L_T = L_0 + \sum_{k=1}^{n+1} \tilde{L}_{\delta_k, \tau_{k-1}},
$$

where $\tilde{L}_{\delta_1,i}, \ldots, \tilde{L}_{\delta_{n+1},i}$ are independent and $\tilde{L}_{\delta_k,i} \stackrel{\text{Law}}{=} L_{\delta_k,i}$, $k = 1, \ldots, n+1$.

E Auxiliary results

Proposition E.0.5 *Suppose $A = \Sigma + bb^\top$ where Σ is symmetric and invertible. Then A is invertible, and*

$$A^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1} b b^\top \Sigma^{-1}}{1 + b^\top \Sigma^{-1} b}.$$

Moreover, $1 - b^\top A^{-1} b = \frac{1}{1+b^\top \Sigma^{-1} b} > 0$.

Proof: It is easy to check that $A\left(\Sigma^{-1} - \frac{\Sigma^{-1} b b^\top \Sigma^{-1}}{1+b^\top \Sigma^{-1} b}\right) = I$, so A is invertible and its inverse is $\Sigma^{-1} - \frac{\Sigma^{-1} b b^\top \Sigma^{-1}}{1+b^\top \Sigma^{-1} b}$. Finally, setting $c = b^\top \Sigma^{-1} b$, one gets

$$1 - b^\top A^{-1} b = 1 - c + \frac{c^2}{1+c} = \frac{1}{1+c} > 0.$$

■

E.1 Proof of Proposition 2.1.3

The result is obviously true for $k = n + 1$.

Suppose that it is true for $k + 1$. For i given, set $\pi_j = Q_{ij}\gamma_{k+1}(j)/D$, where $D = \sum_{j=1}^{l} Q_{ij}\gamma_{k+1}(j)$. By hypothesis, $\pi_1, \ldots, \pi_l$ are probabilities and letting X be a random vector with law $\mathcal{P}_j$ with probability π_j, one gets

$$\gamma_k(i) = D\left(1 - \mu^\top B^{-1} \mu\right),$$

where $\mu = E(X)$ and $B = E\left(XX^\top\right)$. Let Σ be the covariance matrix of X which is non singular since the covariance of X under $\mathcal{P}_j$ is assumed non singular. It then follows from Proposition E.0.5 that $1 - \mu^\top B^{-1} \mu = \frac{1}{1+\mu^\top \Sigma^{-1} \mu} > 0$, where $\Sigma = B - \mu\mu^\top$ is the associated covariance matrix. Since $D > 0$ by hypothesis, one may conclude that $\gamma_k(i) > 0$. As a by-product we get that $\gamma_k(i) \le 1$ if $\gamma_{k+1}(j) \le 1$ for all $j = 1, \ldots$. Since that is true for $\gamma_{n+1} \equiv 1$, one may conclude that for all $k = 1, \ldots, n$, $\gamma_k(i) \le 1$. ■

References

Bouchaud, J.-P. and Potters, M. (2002). Back to basics: historical option pricing revisited. *Philosophical Transactions: Mathematical, Physical & Engineering Sciences*, 357(1758):2019–2028.

Boyle, P. P. and Emanuel, D. (1980). Discretely adjusted option hedges. *Journal of Financial Economics*, 8:259–282.

Broadie, M. and Glasserman, P. (1996). Estimating security price derivatives using simulation. *Management Science*, 42:260–285.

Cont, R. and Tankov, P. (2004). *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL.

Cornalba, L., Bouchaud, J.-P., and Potters, M. (2002). Option pricing and hedging with temporal correlations. *Int. J. Theor. Appl. Finance*, 5(3):307–320.

Cox, J., Ross, S., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263.

Duan, J.-C. (1995). The GARCH option pricing model. *Math. Finance*, 5(1):13–32.

Ethier, S. and Kurtz, T. (1986). *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York. Characterization and convergence.

Föllmer, H. and Sondermann, D. (1986). Hedging of nonredundant contingent claims. In *Contributions to mathematical economics*, pages 205–223. North-Holland, Amsterdam.

Garcia, R. and Renault, É. (1998). A note on hedging in ARCH and stochastic volatility option pricing models. *Math. Finance*, 8(2):153–161.

Guo, X. (2001). Information and option pricings. *Quantitative Finance*, 1:38–44.

Hamilton, J. (1990). Analysis of time series subject to changes in regime. *J. Econometrics*, 45(1-2):39–70.

Jacod, J. and Shiryaev, A. N. (2003). *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition.

Kat, H. and Palaro, H. (2005). Who needs hedge funds? A copula-based approach to hedge fund return replication. Technical report, Cass Business School, City University.

Kurtz, T. G. and Protter, P. (1991a). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070.

Kurtz, T. G. and Protter, P. (1991b). Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic analysis*, pages 331–346. Academic Press, Boston, MA.

Mémin, J. and Słomiński, L. (1991). Condition UT et stabilité en loi des solutions d'équations différentielles stochastiques. In *Séminaire de Probabilités, XXV*, volume 1485 of *Lecture Notes in Math.*, pages 162–177. Springer, Berlin.

Papageorgiou, N., Rémillard, B., and Hocquard, A. (2008). Replicating the properties of hedge fund returns. *Journal of Alternative Invesments*, 11:8–38.

Pham, H. (2000). Hedging and optimization problems in continuous financial models. In Yong, J. and Cont, R., editors, *Mathematical Finance, Theory and Practice*, Series in Contemporary Applied Mathematics.

Protter, P. E. (2004). *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition. Stochastic Modelling and Applied Probability.

Rémillard, B., Papageorgiou, N., and Hocquard, A. (2009). Option pricing and dynamic hedging for regime-switching geometric random walks. Technical report, DGIA-HEC Research on Alternative Investments.

Rubenthaler, S. (2003). Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. *Stochastic Process. Appl.*, 103(2):311–349.

Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Math. Oper. Res.*, 20(1):1–32.

Słomiński, L. (1989). Stability of strong solutions of stochastic differential equations. *Stochastic Process. Appl.*, 31(2):173–202.

Wilmott, P. (2006). *Paul Wilmott on Quantitative Finance*, volume 3. John Wiley & Sons.