Density quantization method in the optimal portfolio choice with partial observation of stochastic volatility

Grzegorz Halaj

Bank Pekao SA, ALM Office, Warsaw
Postdoc fellow at the Fields Institute, Toronto

Abstract
Computational aspects of the optimal consumption and investment with the partially observed stochastic volatility of the asset prices are considered. The new quantization approach to filtering – density quantization – is introduced which reduces the original infinite dimensional state space of the problem to the finite quantization set. The density quantization is embedded into the numerical algorithm to solve the dynamic programming equation related to the portfolio optimization.

Keywords: Optimal portfolio, partial observation, filtering, density quantization

1. Introduction

The optimal investment portfolio and consumption path choice have been studied extensively for about 5 decades but still poses computational difficulties. The standard Merton-like investment problem evolved into many research fields and one of them are the partial observation models studied by Lakner [15], Pham and Quenez [19], Sass and Haussmann [21], Stettner [24], Baghery and Øksendal [1], Halaj [13], Brendle [6]. Those models are imbedded into the control theory framework with partial information [4]. In the case of portfolio choice in the partial observation, the investor does not
observe all the factors defining the dynamics of the market. Hence, the opti-
mal strategies are not only functions of the state variables like the wealth
of the investor, but depend on investor’s beliefs about the true value of as-
set prices and economic factors [24]. Those beliefs are represented by the
probability distribution of values of the factors given available information
on the market. This makes the value function of the investor to be defined on
the infinite dimensional state space. Following the quantization approach we
introduced the numerical method to solve the related dynamic programming
equation.

The partial observation shows up in the portfolio models in two forms.
The first one is related to trend parameters of assets prices, the second one to
volatility. Models with an unobserved trend are usually easy to reformulate
in a full observation setting (eg. by the change of measure approach the drift
parameter of diffusion price processes vanish). The portfolio problems with
unknown volatility are, however, more realistic and we assumed in our model
that the volatility of stocks is unobserved, also driven by some unobserved
economic factors. This is a standard assumption in advanced models of
volatility [2, 12].

We consider the numerical approximation of the solution to the follow-
ing consumption and investment problem. For clarity of the presentation
of the main idea we postpone the rigorous assumptions to the section [1, 2].
Let us denote by $\mathbb{T}$ the set $\{0, 1, \ldots, T\} \subset \mathbb{N} \cap \{0\}$. We consider the prob-
ability space $(\Omega, \mathcal{F}, P)$ and the discrete filtration $\mathcal{F}^t = \{\mathcal{F}_t\}_{t \in \mathbb{T}}$. On $\mathcal{F}$ we
define $N$ sequences of the IID random variables $\{\epsilon_i(t)\}_{t \in \mathbb{T}}, i \in \{1, \ldots, N\}$,
each of which has a positive density $\phi$. By $\{r(t)\}_{t \in \mathbb{T}}$ we denote a process of the interest rates in the economy, constant within the periods $[t, t+1]$ and
by $\{S_0(t)\}_{t \in \mathbb{T}}$ we denote the bank account process given by the recurrence
$S_0(0) = 1$ and $\forall n > 0 S_0(t+1) = (1 + r(t))S_0(t)$.

We assume that the process of the economic factors measuring the eco-
nomic situation is denoted by $\{Y(t)\}_{t \in \mathbb{T}}$ and given by the following equation:

\[ Y(0) = y_0 \quad Y(t+1) = Y(t) + \alpha (\bar{y} - Y(t)) + \sigma_Y \xi(t+1), \tag{1} \]

where $(\xi(t))_{t \in \mathbb{T}}$ is the IID sequence of $K$-dimensional random variables, each
with positive density $\psi$. $\sigma_Y$ is the full-rank $K \times K$ matrix. The tradable
assets on the market have the following dynamics for $i \in \{1, \ldots, N\}$:

\[ S_i(t+1) = S_i(t) \exp(\mu(Y(t+1)) + \sigma(Y(t+1)) \epsilon_i(t+1)). \tag{2} \]
There are two types of shocks on the market: the idiosyncratic $\epsilon$ which are asset (eg. company) specific and the systemic $\xi$ which influence all the risky assets on the market.

The investor (and consumer at the same time) have preferences characterized by utility functions $u: \mathbb{R}^K \times \mathbb{R}_+ \to \mathbb{R}$ and $u_T: \mathbb{R}^K \times \mathbb{R}_+ \to \mathbb{R}$.

The return from investment in the asset $S_i$ in $t$ for $[t, t + 1]$ is denoted by $R_i(t + 1) = \frac{S_i(t + 1)}{S_i(t)}$. The logarithmic return is denoted $R^l_i(t) = \ln(R_i(t))$. The wealth of the investor starting with capital $x_0$ and investing $\pi: = \{\pi(t) | t \in \mathbb{T}\}$ and consuming $c: = \{c(t) | t \in \mathbb{T}\}$ is defined by:

$$X(t + 1) = \sum_{i=1}^{N} \pi_i(t) \frac{S_i(t + 1)}{S_i(t)} + (X(t) - \sum_{i=1}^{N} \pi_i(t) - c(t))(1 + r(t)), \quad X(0) = x_0.$$  

The investor observes only the prices $S_i$ of the securities so his information is described by filtration $F^S = \{F^S_t\}_{t \in \mathbb{T}}$ where $F^S_t = \sigma(\{S_i(s) | i \in \{1, \ldots, N\}, s \leq t\})$. It is equivalent to observing $F^{lR} = \sigma(\{R^l_i(s) | i \in \{1, \ldots, N\}, s \leq t\}$ generated by log-returns from investment in $S$ which is more convenient for our purpose. The investor wants to optimize his utility from consumption and his terminal wealth. She maximizes the functional $J: \mathcal{A}(x) \to \mathbb{R}$ given by:

$$J(c, \pi) = \mathbb{E} \left\{ \sum_{t=1}^{T-1} \delta^t u(c(t)) + \delta^T u_T(Y(T), X^{c, \pi}(T)) \right\},$$

where $T$ is the time horizon and $\mathcal{A}(x)$ is the set of the admissible controls $(c, \pi)$

$$\mathcal{A}(x) = \left\{ ((c(0), \ldots, c(T - 1)), (\pi(0), \ldots, \pi(T - 1))) | (c(t), \pi(t)) - F^{lR}_t \text{-measurable}; \quad 0 \leq c(t) \leq X^{c, \pi}(t); \quad \sum_{i=1}^{N} \pi_i(t) \leq X^{c, \pi}(t) - c(t) \right\}.$$

The solution maximizing $J$ is given by the special form of dynamic programming.

$$V(t, x, \varrho) = \sup_{\mathcal{A}(x)} \left\{ u(\bar{c}) + \delta \mathbb{E} V(t + 1, x(\hat{R}^l(t + 1)), \varrho_{t+1}(t + 1, \hat{R}^l(t + 1)) \right\}$$
The arguments $t$ and $x$ refer to usual time and current wealth state space variables (like in Bertsekas [5]). The variable $\rho$ measures the distribution of a random variable which is precisely defined in section 2 and corresponds to the process $Y$. Unlike in the full observation stochastic control in which case the value function at time $t$ depends on the realization of $Y(t)$, the partial observation allows the investor to assess only the most likely value of $Y(t)$ given the history of the asset prices. That is the intuition behind the dependence of $V$ on the whole distribution of $Y$ given $\mathcal{F}_t$.

The lack of closed form solutions to consumption/investment problems gives rise to research on numerical simulation methods to be applied to models with partial observation. The numerical procedure has to be designed in a special way to account for the natural high complexity of the state space of stochastic control. The literature does not give many examples of approximate solutions to the stochastic control with filtering. Runggaldier and Stettner [20] discretize transition probabilities and Desai et al. [10] proposed a particle filtering method to discretize a dynamic programming equation. There are numerous studies of numerical methods of filtering equations themselves. The Zakai equations which describe densities of unobserved parameters evolving in time as new information arrives to the market can be solved by means of the so-called particle filter methods [9]. Filter equations in discrete time can be estimated with the so-called quantization [16, 17, 18].

In the related paper Corsi et al. [8] also proposed the scheme to solve the dynamic programming equation of the partially observed control problem applying markovian quantization and approximated the given Markov Chain by the transition probability matrix. In this way, infinite dimensional random variables can be approximated by a finite set of point in an optimal way. We modified the standard quantization techniques to develop a different approach – density quantization – and we applied it to our control problem. Basically, the main idea is to find the optimal quantization set directly in the set of the unnormalized densities of the filter. In other words, the density quantization set is the set of the appropriately selected unnormalized density functions. It proved to be successful in recursively solving the dynamic programming equation.

The paper is organized as follows: First, we present the mathematical model of the market and decision making. Second, we show the influence partial information can exert on consumption of an investor in a simplistic theoretical example. Third, we rewrite the partially observed problem in a full observation setting, applying the change of measure approach and
we solve it using the dynamic programming. Fourth, we define the density quantization and we show how the investment/consumption problem can be solved numerically. Finally, we illustrate the application of the developed numerical method in the example of the market model.

1.1. Notation.

- $\mathcal{M}^{n \times m}$ — the set of matrices $M$ with $n$ rows and $m$ columns.
- $B^n(F)$ — the set of Borel functions $f : \Omega \times \mathbb{R}^K \to \mathbb{R}$, $f(\omega, \cdot) \in L^n(\mathbb{R}^K)$.
- $\phi_{\mu, \sigma}(x)$ — Gaussian density i.e.
  \[
  \phi_{\mu, \sigma}(x) = \frac{1}{(2\pi)^{N/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right),
  \]
  where $\mu$ is mean and $\Sigma = \sigma\sigma^T$ is covariance matrix.
- $\text{vol}(A)$ — volume of a set $A \in \mathbb{R}^n$.
- $\text{diam}(A) = \sup_{(x,y) \in A \times A} (\sum_{1 \leq i \leq n} |x_i - y_i|^2)^{1/2}$.
- $A^c$ — completion of set $A$.

1.2. Main technical assumptions. To show convergence of solutions of approximate problem to the original one we need 3 assumptions for $\phi$ and $\psi$.

- [A0] $\sigma(\cdot)$ is bounded from 0, i.e. $\exists \sigma_{\text{min}} > 0$ such that $\forall y \in \mathbb{R}^K \sigma(y) > \sigma_{\text{min}}$.
- [A1] $u$ and $u_T$ satisfy $\int_{\mathbb{R}^N \times \mathbb{R}^K} (\phi(x) + \psi(y)) (u(\exp(x)) + u_T(\exp(x), y)) dx dy < \infty$.
- [A2] For every $y \in \mathbb{R}^N$ and $a < 1$, $\int_{\mathbb{R}^K} \phi\left(\frac{z-y}{a}\right) \frac{1}{\phi(y)} dz < L_\Phi$.

[A0] and [A2] are necessary to show convergence of the numerical scheme. [A1] guarantees the integrability in the dynamic programming equations.

2. Filtering

The most fundamental task to be done before applying the standard dynamic programming toolkit to our consumption/investment selection is to transform the problem into the complete observation one. We apply changing of reference measure techniques similarly to Krishnamurthy and Le Gland
Elliot and Miao [12], Stettner [24]. After that, in the section 3 we transform the classical dynamic programming equation to make it suitable for our portfolio problem.

For each $t \in T$ we define a random variable

$$
\lambda_t = \frac{\phi \left( \sigma^{-1}(Y(t))[R^0(t) - \mu(Y(t))] \right) \psi \left( \sigma_Y^{-1}[Y(t) - Y(t-1) - \alpha(\bar{y} - Y(t-1))] \right)}{\det(\sigma(Y(t))) \phi(R^0(t)) \det(\sigma_Y) \psi(Y(t))}.
$$

Let us construct a process $\Lambda$ in the following way: $\Lambda_t = \prod_{s=1}^{t} \lambda_s$.

The process $\Lambda$ defines a new measure – the reference measure $\bar{P}$. Let us define the filtration $\mathcal{F}_{t}^{IR,Y}$:

$$
\mathcal{F}_{t}^{IR,Y} = \sigma(\{R^t(s), Y(s)| i \in \{1, \ldots, N\}, s \leq t\}).
$$

For each $t \in T$, its restriction to $\mathcal{F}_{t}^{IR,Y}$ is defined:

$$
\frac{dP}{d\bar{P}} \bigg|_{\mathcal{F}_{t}^{IR,Y}} = \Lambda_t.
$$

To study properties of stochastic processes under $\bar{P}$ we extensively use the generalized Bayes formula: for $(\mathcal{F}, P)$-measurable $Z$

$$
\mathbb{E}[Z|\mathcal{F}_t^Y] = \frac{\mathbb{E}[\Lambda_t^{-1}Z|\mathcal{F}_t^Y]}{\mathbb{E}[\Lambda_t^{-1}|\mathcal{F}_t^Y]} \quad \text{and} \quad \mathbb{E}[Z|\mathcal{F}_t^Y] = \frac{\mathbb{E}[\Lambda_t Z|\mathcal{F}_t^Y]}{\mathbb{E}[\Lambda_t|\mathcal{F}_t^Y]].
$$

We assume that there exists a function-valued process $\varrho: \Omega \times T \times \mathbb{R}^K \rightarrow L^2(\mathbb{R})$ satisfying for a $L^2$ Borel function $f$ we have

$$
\mathbb{E}[\Lambda_t f(Y)|\mathcal{F}_t^{IR}] = \int_{\mathbb{R}^K} f(z) \varrho_t(z) dz, \ t \in T.
$$

For a given $t$ the object $\varrho_t$ is the conditional unnormalized density of $Y$. We look for the recursive equation for $\varrho$ which – combining with the Bayes formula – would give the tractable expression for filters used in the optimization. It makes the numerical simulation of the optimal consumption/investment strategies easier.

The main result of the section, relying on independence of $Y$ and $R^t$ under $\bar{P}$ is given in the following theorem.

**Theorem 2.1.** Let us denote

$$
\Phi(z, R^t(t)) = \frac{\phi \left( \sigma^{-1}(z)[R^0(t) - \mu(z)] \right)}{\det(\sigma(z))},
$$

$$
\Psi(z, Y(t-1)) = \frac{\psi \left( \sigma_Y^{-1}[z - Y(t-1) - \alpha(\bar{y} - Y(t-1))] \right)}{\det(\sigma_Y)}.
$$
The process $\varrho$ satisfies the recursion

$$
\varrho_t(z) = \frac{\Phi(z, R^l(t))}{\phi(R^l(t))} \int_{\mathbb{R}^K} \Psi(z, y) \varrho_{t-1}(y) dy. \quad (7)
$$

**Proof:** (see appendix Appendix A)

Note that

$$
\mathbb{E} \left[ f(Y(t)) \mid \mathcal{F}^t_R \right] = \frac{\int_{\mathbb{R}^K} f(z) \varrho_t(z) dz}{\int_{\mathbb{R}^K} \varrho_t(z) dz}.
$$

This representation allows us to rewrite the original problem to the full observation control. Since $\Lambda$ is a martingale then – similarly to Runggaldier and Stettner [20] – we obtained

$$
J(c, \pi) = \mathbb{E} \left\{ \sum_{t=1}^{T} \delta^t u(c(t)) + \delta^T u_T(Y(T), X^{c,\pi}(T)) \right\}
$$

$$
= \sum_{t=1}^{T} \mathbb{E} \left[ \Lambda_t \delta^t u(c(t)) \right] + \mathbb{E} \left[ \Lambda_T \delta^T u_T(Y(T), X^{c,\pi}(T)) \right]
$$

$$
= \sum_{t=1}^{T} \mathbb{E} \left[ \mathbb{E} \left[ \Lambda_t \delta^t u(c(t)) \mid \mathcal{F}^t_R \right] \right] + \mathbb{E} \left[ \mathbb{E} \left[ \Lambda_T \delta^T u_T(Y(T), X^{c,\pi}(T)) \mid \mathcal{F}^T_R \right] \right]
$$

$$
= \mathbb{E} \left\{ \sum_{t=1}^{T} \delta^t u(c(t)) + \delta^T \int_{\mathbb{R}^K} u_T(z, X^{c,\pi}(T)) \varrho_T(z) dz \right\}. \quad (8)
$$

In this way all the processes used in the equation are observed, i.e. adapted to the filtration generated by the observed process $R^l$.

**Remark 2.1.** There is an analogical way of solving the filters in the continuous time. The density of the filter is described by the so-called Zakai equation. For instance, Beneš et al. [3] showed the existence of the solution to the Zakai equation in the theoretical case and Carmona and Ludkovsky [7] used this equation in the financial setting.

### 3. Dynamic programming

After transcribing the original problem to a Markov one we can formulate a dynamic programming equation for the optimal wealth of an investor. Unfortunately, this is an infinite-dimensional problem since the value functional of the optimization is a function on the space of the random densities.
Let $\mathcal{A}_t(x)$ be the set of the admissible controls of the portfolio starting at time $t$ from wealth $x$ (initial at $t$) i.e.

$$
\mathcal{A}_t(x) := \{((c(t), c(t+1), \ldots, c(T)), (\pi(t), \pi(t+1), \ldots, \pi(T))) | (c, \pi) \in \mathcal{A}(x)\}.
$$

The optimal value $V$ at time $t \in \mathbb{T}$ with the initial wealth $x$ and the updated density $\varrho$ of the unobserved process at $t$ is defined

$$
V(t, x, \varrho): = \sup_{(c, \pi) \in \mathcal{A}_t(x)} \mathbb{E}\left[ \sum_{s=t}^{T-1} \delta^{s-t} u(c(s)) + \delta^{T-t} \int_{\mathbb{R}^K} u_T(z, X^{c,\pi}(T)) \varrho_T(z) dz \bigg| \mathcal{F}_t^R \right],
$$

for $X^{c,\pi}(t) \equiv x$ and $\varrho_t \equiv \varrho$. The previous section gives

$$
V(T, x, \varrho) = \mathbb{E}\left[ \int_{\mathbb{R}^K} u_T(z, X^{c,\pi}(T)) \varrho_T(z) dz \bigg| \mathcal{F}_T^R \right] = \int_{\mathbb{R}^K} u_T(z, x) \varrho(z) dz.
$$

**Lemma 3.1.** The following backward recursion holds for $V$ and $t < T$:

$$
V(t, x, \varrho) = \sup_{\{\tilde{c}, \tilde{\pi} \in \mathbb{R} \times \mathbb{R}^N | 0 \leq \tilde{c} \leq x, \sum_{1 \leq i \leq N} \tilde{\pi}_i \leq x\}} \left\{ u(\bar{c}) + \delta \int_{\mathbb{R}^N} V\left(t + 1, \sum_{i=1}^{N} \tilde{\pi}_i \tilde{R}_i \right) + (x - \sum_{i=1}^{N} \tilde{\pi}_i - \bar{c})(1 + r(t)), \varrho_{t+1} \right\} \varrho \left( \prod_{i=1}^{N} \phi(\tilde{R}_i) d\tilde{R}_1 \cdot \ldots \cdot d\tilde{R}_N \right)
$$

and

$$
\sup_{(c, \pi) \in \mathcal{A}} J(c, \pi) = V(0, x_0, \psi). \tag{10}
$$

**Proof:** It is straightforward that the identity [10] hold.
By definition, $\mathcal{F}_t^{IR}$-measurability of $(c(t), \pi(t))$ and $\mathcal{F}_t^{IR} \subset \mathcal{F}_{t+1}^{IR}$

$$V(t, x, \varrho) = \sup_{(c, \pi) \in \mathcal{A}_t(x)} \left\{ \mathbb{E} \left[ \sum_{s=t}^T \delta^{s-t} u(c(s)) + \delta^{T-t} \int_{\mathbb{R}^K} u_T(z, X^{c,\pi}) \varrho_T(z) dz \bigg| \mathcal{F}_t^{IR} \right] \right\}$$

$$= \sup_{(c, \pi) \in \mathcal{A}_t(x)} \left\{ u(c(t)) + \delta \mathbb{E} \left[ \sum_{s=t+1}^T \delta^{s-(t+1)} u(c(s)) + \right. \right.$$  
$$\left. \left. \delta^{T-(t+1)} \int_{\mathbb{R}^K} u_T(z, X^{c,\pi}) \varrho_T(z) dz \bigg| \mathcal{F}_{t+1}^{IR} \mathcal{F}_t^{IR} \right] \right\}$$

$$= \sup_{(c, \pi) \in \mathcal{A}_t(x)} \left\{ u(c(t)) + \delta \mathbb{E} \left[ V(t+1, \kappa(x, c(t), \pi(t), R^t(t+1)), \varrho_{t+1}) \bigg| \mathcal{F}_t^{IR} \right] \right\}$$

Using definition of $\Lambda_t$, it is easy to check that $R^t_\mathcal{I}(t+1)$ has the density $\phi$ under the measure $\bar{P}$. Let us define

$$\kappa(t, \bar{x}, \hat{c}, \pi, \hat{R}) = \sum_{i=1}^N \hat{\pi}_i \hat{R}_i + (\bar{x} - \sum_{i=1}^N \hat{\pi}_i - \hat{c})(1 + r(t)).$$

$$= \sup_{(c, \pi) \in \mathcal{A}_t(x)} \left\{ u(c(t)) + \delta \mathbb{E} \left[ V \left( t+1, \kappa(x, c(t), \pi(t), R^t(t+1)), \right. \right. \right.$$  
$$\left. \left. \frac{\Phi(\cdot, R^t(t+1))}{\phi(R^t(t+1))} \int_{\mathbb{R}^K} \Psi(\cdot, y) \varrho_{t-1}(y) dy \bigg| \mathcal{F}_t^{IR} \right] \right\}$$

$$\sup_{(\bar{c}, \hat{R}) \in \mathcal{A}(x)} \left\{ u(\bar{c}) + \delta \int_{\mathbb{R}^N} V \left( t+1, \kappa(x, \bar{c}, \pi, \hat{R}), \right. \right.$$  
$$\left. \left. \frac{\Phi(\cdot, \hat{R})}{\phi(\hat{R})} \int_{\mathbb{R}^K} \Psi(\cdot, y) \varrho_{t-1}(y) dy \right] \prod_{i=1}^N \phi(\hat{R}_i) d\hat{R}_1 \cdot \ldots \cdot d\hat{R}_N \right\}.$$
3.1. Quantization. The set $B^2(\mathcal{F})$ is infinite dimensional. Hence, $V$ in the form of the equation 9 cannot be computed numerically. It has to be approximated by a sequence of the functionals $\hat{V}$ on the finite state space. It is a very complex numerical problem. To this end, Runggaldier and Stettner [20] proposed the scheme for the numerical filtering based on the Hidden Markov Model approach. They discretized the infinite-dimensional equation of the transition probabilities on the finite state space and then could solve the dynamic programming equation on the set of vectors in $\mathbb{R}^M$. The grid was chosen arbitrarily. The partition in this grid method is straightforward but easily gets untractable as the number of nodes in the partition of the codomain increases. With only 5 grids the number of the approximate value functions, which has to be computed amounts to $5^5 = 3125$. 20 grids give $10485760000000000000000000$ variants!

A much more applicable extension of this approach is the density quantization introduced to limit the number of the unnormalized densities in the set of arguments of the value function and to optimize the structure of the grid.

To reduce dimensionality of our problem from the infinite to finite space we follow the idea proposed by Stettner [24]. We discretize the codomain of the densities $\varrho$. We take a partition of each coordinate of $\mathbb{R}^K$, i.e. for the coordinate $j$ we take the set of $m_j^n$ increasing real numbers $\{z_j^n(1), \ldots, z_j^n(m_j^n)\}$. We approximate random densities $\varrho$ by the functional on the grid

$$Z_n : = \prod_{j=1}^K \{z_j^n(1), \ldots, z_j^n(m_j^n)\} \subset \mathbb{R}^K.$$ 

The cubes $B_n(k_1, \ldots, k_K) : = \prod_{j=1}^K [z_j^n(k_j), z_j^n(k_j + 1)]$ give the partition of the support of $\rho_t$. The sum of the sets $B_n$ is denoted

$$\mathfrak{B}_n = \bigcup_{\{(k_1, \ldots, k_K) | 1 \leq k_i < m_i^n\}} B_n(k_1, \ldots, k_K).$$

We assume that

$$\lim_{n \to +\infty} \vol(\mathfrak{B}_n) = +\infty,$$

$$\forall 1 \leq j \leq K \quad \lim_{n \to +\infty} \max_{i \in \{1, \ldots, m_j^n - 1\}} |z_j^n(i + 1) - z_j^n(i)| = 0.$$
Similarly, we define the partition of the codomain of $\rho: G_n := \{g_n(1), \ldots, g_n(m_n)\}$ such that $g_n(i + 1) > g_n(i)$, $\lim_{n \to \infty} g_n = \infty$ and $\max_i |g_n(i + 1) - g_n(i)| \to 0$ as $n$ tends to $+\infty$.

The idea of the quantization is to focus attention on the set of density functions and to project all random densities on this set. Unlike in the grid method where the projections are made inside every subset $B_n$ on the grid $G_n$, the number of outcomes can be quite small. Intuitively, under mild conditions the densities form a “tight” family of functions and many of them are very close to each other (in sup norm). There is no point in distinguishing between them since they lead to similar value functions. The idea of quantization gave fruitful results in finding transition probabilities for nonlinear filters considered by Pagès et al. \[17\]. It happens to be useful in the dynamic programming as well.

The fundamental idea of the approach is to project random densities on a set of functions that have only finitely many values and vanish at all point sufficiently distant from 0. The set of functions will be called density quantization set (or shortly—quantization set).

**Definition 3.1 (Density quantization set).** Any set of functions $q: \mathbb{R}^K \to G_n$, which are constant on the sets $B_n(k_1, \ldots, k_K)$, $k_i \in \{1, \ldots, m_n(i)\}$ and are equal to 0 on $\mathbb{R}^n$ is called density quantization set and is denoted $Q_n$. The number of elements of the set $Q_n$ is $N_n$ and tends to $+\infty$ with $n \to +\infty$.

The projection $\text{Proj}_{Q_n}^Z[\varrho]$ of a random density $\varrho$ on $Q_n$ is defined

$$\text{Proj}_{Q_n}^Z[\varrho] := \arg\min_{q \in Q_n} \max_{\{z(1), \ldots, z(K)\}\mid 1 \leq k_i \leq m_n(i)} |\varrho(\omega, z(1)(k_1), \ldots, z(K)(k_K)) - q(z(1)(k_1), \ldots, z(K)(k_K))|.$$  

The proper behavior of density quantization is guaranteed by the so-called Zador Theorem \[17\].

**Theorem 3.1.** Let $X$ be $n$-dimensional random variable such that $X \in L^{2+\delta}$, $\delta > 0$, with density $f_n$. Let $Q_K \subset \mathbb{R}^n$ and the projection $\text{Proj}_{Q_K}: \mathbb{R} \to Q_K$ be defined as $\text{Proj}_{Q_K}[x] = \arg\min_{q \in Q_K} \|q - x\|$. Then

$$D_n(Q_K) := \min_{\{Q_K | \#Q_K = K\}} \mathbb{E}\|X - \text{Proj}_{Q_K}(X)\|^2 \leq K^{-\frac{2}{n}} J_n \cdot \left(\int_{\mathbb{R}^n} f_n(x)^{\frac{n}{n+2}} \, dx\right)^{\frac{n+2}{n}},$$

where $J_n \simeq \frac{n}{2\pi c}$.  

11
The quantity $D_n(Q_n)$ is called the distortion of the quantization.

The straightforward application of the Zador Theorem gives very slow convergence. The structure of the stock market model allows to show that, in fact, the numerical scheme have much better properties if we assume Lipschitz continuity of $\Phi(z, \cdot)/\phi(\cdot)$.

Let us discretize the space of values of $R$ and consider the set $Q_n^{(R)} := \{\bar{r}_1, \ldots, \bar{r}_n | \bar{r}_i \in \mathbb{R}^N\}$. Let the function $\bar{\varrho}$ be defined as

$$\bar{\varrho}(z, \bar{R}) := \Phi(z, \bar{R})/\phi(\bar{R}) \int_{\mathbb{R}^K} \Psi(z, y) \rho(y) dy.$$  

From now on, we can think about $G_n$ as of the set generated implicitly by $Q_n^{(R)}$, i.e.

$$G_n \equiv \{\bar{\varrho}(z^{(j)}(l), \bar{r}_i) | j \in \{1, \ldots, n\}, l \in \{1, \ldots, m_n^{(j)}\}, i \in \{1, \ldots, n\}\}.$$

The quantization set from definition [3.1] corresponding to $Q_n^{(R)}$ is denoted $Q_{Q_n^{(R)}}$. The set of all the sets $Q_n$ is denoted $S_Q$. A separate discretization is defined for the portfolio process $X$, but in the standard way followed in the numerical solving of the dynamic programming equations (in the full observation setting or even in the deterministic setting). The projection set is $Q_n^X = \{x_1, \ldots, x_n\}$ and the set of the projection sets is denoted $S_X$.

Let us define the sequence of functionals $\hat{V}$ on $T \times \mathbb{R} \times B^1(F)$ as

$$\hat{V}(T, x, \varrho) = \int_{\mathbb{R}^N} u_T(z, x) \varrho(z) dz$$

$$\hat{V}(t, x, \varrho) = \sup_{c, \pi \in A(t)} \{u(c) + \delta \mathbb{E} \hat{V}(t + 1, \text{Proj}_{Q_n^X}[x^{c,\pi}(R)], \text{Proj}_{Q}[\varrho(R)])\}.$$  

Note that

$$V(T, x, \varrho) = \hat{V}(T, x, \varrho)$$

for $x \in Q^X$ and $\varrho \in Q$. In the following part of the subsection we showed that $\hat{V}$ approximates $V$.

For brevity, if it does not lead to confusion, let us denote $\hat{x} := \text{Proj}_{Q_n^X}[x]$ and $\hat{\varrho} := \text{Proj}_{Q}[\varrho]$.

**Lemma 3.2.** If $u_T(\cdot, z)$ is $L_2$-Lipschitz continuous and bounded by $L_1$ then

$$|V(T, x_1, \varrho_1) - \hat{V}(T, x_2, \varrho_2)|^2 \leq 2L_1^2 \int_{\mathbb{R}^K} |\varrho_1(z) - \hat{\varrho}(z)| dz + 2L_2^2 |x_1 - \hat{x}_2|.$$  

12
Proof:
\[ |V(T, x_1, \varrho_1) - \tilde{V}(T, x_2, \varrho_2)|^2 = \left| \int_{\mathbb{R}^K} u(x_1, z) \varrho_1(z) dz - \int_{\mathbb{R}^K} u(x_2, z) \varrho_2(z) dz \right|^2 < \int_{\mathbb{R}^K} |u(x_1, z) \varrho_1(z) - u(x_1, z) \varrho_2(z) + u(x_1, z) \varrho_2(z) - u(x_2, z) \varrho_2(z)|^2 dz \leq 2L_1^2 \int_{\mathbb{R}^K} |\varrho_1(z) - \varrho_2(z)|^2 dz + 2L_2^2 |x_1 - \tilde{x}_2|^2. \]  

(12)

Definition 3.2. The value function \( v \) is \((\epsilon, c_0, c_1)\)-Lipschitz continuous if and only if there exist constants \( c_0 \) and \( c_1 \) such that for each pair \((x_1, x_2)\) and \((\varrho_1, \varrho_2)\)
\[ |v(x_1, \varrho_1) - \tilde{v}(x_2, \varrho_2)|^2 < \epsilon + c_0 |x_1 - x_2|^2 + c_1 \int_{\mathbb{R}^K} |\varrho_1(z) - \varrho_2(z)|^2 dz. \]

Lemma 3.3. Let \( \varrho(z, \cdot) \) be Lipschitz continuous and is \( p \)-decaying for \( p > N \) i.e.
\[ \exists a_z > 0 \quad \varrho(z, x) < a_z |x|^{-p}. \]

Let us assume that \( \text{vol}(B(\bar{K})) \leq \left( \frac{M_0}{n} \right)^K \). Let us denote the distribution of \( x_1^{(\epsilon, \pi)}(R) \) by \( f_{x_1}^{\epsilon, \pi} \in L^1 \) and \( x_2^{(\epsilon, \pi)}(R) \) by \( f_{x_2}^{\pi} \in L^1 \). Then, for \((\epsilon, L_x, L_\varrho)\)-Lipschitz continuous function \( V \)
\[ \text{err}_n: = \min_{\{Q_x \in S^X_\#Q_x = n\}} \mathbb{E} |V(t, x_1^{(\epsilon, \pi)}(R), \varrho_1(R)) - \tilde{V}(t, \text{Proj}_{Q_x} [x_2^{(\epsilon, \pi)}(R), \text{Proj}_{Z_2}^{Q} [\varrho_2(R)]])|^2 \leq \]
\[ 2L_x(1 + r(t))^2 |x_1 - x_2|^2 + 2L_\varrho L_\varrho^2 \int_{\mathbb{R}^K} |\varrho_1(z) - \varrho_2(z)|^2 dz + L_x n^{-2} J_1 \left( \int_{\mathbb{R}} f_X(x)^3 dx \right)^3 + 2L_\varrho L_\varrho^2 \sqrt{K} \left( \frac{M_0}{n} \right)^{K+1} + \\
4L_\varrho M_0^K n^{-\frac{2}{N}} J_N \left( \int_{\mathbb{R}^N} f^\text{max}_z (x)^N dx \right)^{\frac{N+2}{N}} + 2L_\varrho a_z v_N \frac{M_0^{N-p}}{p-1} + \epsilon, \]

where
\[ v_N = \begin{cases} 
\frac{(2\pi)^{\frac{N+2-p}{2}}}{(N-p)(N-2)!}, & \text{if } N = 2k \\
\frac{2 \cdot \pi^{\frac{N+2-p}{2}}}{(N-p)(N-2)!}, & \text{if } N = 2k + 1 
\end{cases} \]
and \( f_{z_k}^{max} \) is density of \( \varrho(R)(z_k) \) for \( k \) such that

\[
k = \arg \max_i \left\{ \left( \int_{\mathbb{R}^N} f_{z_k}^{max}(x) \frac{N}{N+2} dx \right)^{\frac{N+2}{N}} \right\}.
\]

**Proof:** By Lipschitz continuity of \( V \) and the inequality \( |a-b|^2 \leq 2a^2 + 2b^2 \)

\[
\Delta V := |V(t, x_1^{(e, \pi)}(R), \varrho_1(R)) - \widehat{V}(t, \text{Proj}_{Q_X}[x_2^{(e, \pi)}(R)], \text{Proj}_{Z_n}^{Q}[\varrho_2(R)])|^2 \leq 
\epsilon + 2L_x|x_1^{(e, \pi)}(R) - \text{Proj}_{Q}[x_2^{(e, \pi)}(R)]|^2 + 2L_{\varrho} \int_{\mathbb{R}^K} |\varrho(R)(z) - \text{Proj}_{Z_n}^{Q}[\varrho(R)](z)|^2 dz \leq 
4L_x|x_1^{(e, \pi)}(R) - x_2^{(e, \pi)}(R)|^2 + 4L_{\varrho} \sum_{\vec{k} \in K_n} \int_{B(\vec{k})} |\varrho_2(R)(z) - g_k|^2 dz + 
4L_{\varrho} \int_{\mathbb{R}^n} \varrho(R)^2(z) dz
\]

for a set of \( g_k \)s, \( g_k \in G_n \).

Firstly, let us make an easy observation that

\[
|x_1^{(e, \pi)}(R) - x_2^{(e, \pi)}(R)|^2 = (1 + r(t))|x_1 - x_2|^2.
\]

Secondly, we can use the quantization on \( Q^n \) to estimate the approximation of the random variable \( x_2^{(e, \pi)}(R) \) with \( \text{Proj}_{Q_X}[x_2^{(e, \pi)}(R)] \). Since \( x_2^{(e, \pi)}(R) \) is the sum of the random variables of the form \( p_i \exp(R), \widehat{R} \sim \varrho(\cdot) \), then from the Zador theorem (see theorem 3.1) we get

\[
\min_{\{Q^n \in S^n : \sharp Q^n = n\}} \mathbb{E}|x_2^{(e, \pi)}(R) - \text{Proj}_{Q_X}[x_2^{(e, \pi)}(R)]|^2 \leq 
\max_{(p_1, ..., p_N) \in \{0, x_2\}^N} n^{-2} J_1 \left( \int_{\mathbb{R}} (f_{X_2}^{(p_1, ..., p_N)}(x))^\frac{1}{3} dx \right)^3,
\]

where \( f_{X_2}^{(p_1, ..., p_N)}(\cdot) := (\varphi_{p_1}^{(e)} * ... * \varphi_{p_N}^{(e)})(\cdot) \) is the convolution of densities of the random variables \( p_i \exp(\widehat{R}) \). The maximum always exists and we skipped the justification to the remark 3.1. We denoted the corresponding \( f_{X_2}^{(p_1, ..., p_N)} \) by \( f_X \).
Thirdly, by the assumption [A2]

\[
\int_{\mathbb{R}^K} |\varrho_1(R)(z) - \varrho_2(R)(z)|^2 \, dz = \\
\int_{\mathbb{R}^K} \left[ \frac{\Phi(z, R^{(i)})}{\phi(R^{(i)})} \int_{\mathbb{R}^K} \Psi(z, y) \varrho_1(y) \, dy - \frac{\Phi(z, R^{(i)})}{\phi(R^{(i)})} \int_{\mathbb{R}^K} \Psi(z, y) \varrho_2(y) \, dy \right]^2 \, dz \leq \\
\int_{\mathbb{R}^K} \frac{\Phi^2(z, R^{(i)})}{\phi^2(R^{(i)})} \, dz L^2_{\Psi} \left( \int_{\mathbb{R}^K} |\varrho_1(y) - \varrho_2(y)| \, dy \right)^2 \leq L^2_{\Psi} L^2_{\Psi} \int_{\mathbb{R}^K} |\varrho_1(z) - \varrho_2(z)|^2 \, dz
\]

Finally, since \(\varrho_2(R)\) is \textit{globally} Lipschitz continuous (there is a common Lipschitz constant \(L_R\) for all \(R(\omega)\)),

\[
\int_{B(\vec{k})} |\varrho_2(R)(z) - g_k|^2 \, dz \leq 2 \int_{B(\vec{k})} |\varrho_2(R)(z) - \varrho_2(R)(z_n)|^2 \, dz + \\
2 \text{vol}(B(\vec{k})) |\varrho_2(R)(z_n) - g_n|^2 \leq 2 L^2_R \text{vol}(B(\vec{k})) \text{diam}(B(\vec{k}))^2 + \\
2 \text{vol}(B(\vec{k})) |\varrho_2(R)(z_n) - g_n|^2
\]

Then,

\[
\text{err}_n \leq \epsilon + 2 L_x (1 + r(t)) |x_1 - x_2| + 2 L_\phi L^2_{\Psi} \int_{\mathbb{R}^K} |\varrho_1(z) - \varrho_2(z)|^2 \, dz + \\
2 L_\phi L^2_R \text{vol}(B(\vec{k})) \text{diam}(B(\vec{k}))^2 + 2 L_x n^{-2} J_1 \left( \int_{\mathbb{R}} (f_X(x))^{\frac{1}{3}} \, dx \right)^3 + \\
4 L^2_{\Psi} E \left[ \min_{\{Q \in \mathcal{S}_Q \mid \# Q = n\}} \sum_{\vec{k} \in K_n} \text{vol}(B(\vec{k})) |\varrho(R)(z_k) - g_k|^2 \right] + 2 L_\phi \int_{\mathbb{R}_n} \varrho(R)^2(z) \, dz.
\]

The sign "\(\sum\)" can be swapped with "\(\min\)" since for each 2 elements of the sum, the minimum is calculated with respect to 2 different arguments. Zador
Theorem and assumption on $B(\vec{k})$ imply

$$\text{err}_n \leq \epsilon + 2L_x(1+r(t))^2|x_1-x_2|^2 + 2L_\varphi L_\varphi^2 \int_{\mathbb{R}^\mathcal{K}} |\varphi_1(z)-\varphi_2(z)|^2 \, dz$$

$$L_x n^{-2} J_1 \cdot \left( \int_{\mathbb{R}} f_X(x)^3 \, dx \right)^3 + 2L_\varphi L_\varphi^2 \left( \frac{M_0}{n} \right)^K \cdot \sqrt{K} \left( \frac{M_0}{n} \right) +$$

$$4L_\varphi \left( \frac{M_0}{n} \right)^K n^{-\frac{2}{3}} J_N \cdot \sum_{k \in \mathcal{K}_n} \left( \int_{\mathbb{R}^N} f_{z_k}(x)^{N+2} \, dx \right)^{\frac{N+2}{N}} + 2L_\varphi \int_{\mathbb{R}^\mathcal{K}} \varphi(R)^2 \, dz$$

$$\leq \epsilon + 2L_x(1+r(t))^2|x_1-x_2|^2 + 2L_\varphi L_\varphi^2 \int_{\mathbb{R}^\mathcal{K}} |\varphi_1(z)-\varphi_2(z)|^2 \, dz$$

$$L_x n^{-2} J_1 \cdot \int_{\mathbb{R}^\mathcal{K}} \left( f_X(x)^3 \, dx \right)^3 + 2L_\varphi L_\varphi^2 \sqrt{K} \left( \frac{M_0}{n} \right)^{K+1} +$$

$$4L_\varphi M_0^K n^{-\frac{2}{3}} J_N \left( \int_{\mathbb{R}^N} f_{z_k}^{\max}(x)^{\frac{N+2}{N+2}} \, dx \right)^{\frac{N+2}{N}} + 2L_\varphi a_i \int_{[M_0, +\infty)^N} |x|^{-p} \, dx$$

The calculation of the constant $v_N$ is left to the Appendix.

\[ \square \]

**Remark 3.1.** The minimum with respect to $p_i$s in formula $13 \int_{\mathbb{R}} (\phi^{(e)}_{p_1} * \ldots * \phi^{(e)}_{p_N})(x)^{\frac{1}{3}} \, dx$ always exists. Let us show it in the case of $N = 2$. The other cases are analogous.

We have

$$F(p_1, p_2) : = \int_{\mathbb{R}} \left( \left( (\phi^{(e)}_{p_1} * \ldots * \phi^{(e)}_{p_2}) \right)(x) \right)^{\frac{1}{3}} \, dx$$

and obviously, by the definition of $\phi_{p_i}^{(e)}$, $F(0, p_2) = F(p_1, 0) = 0$. It is straightforward that $F$ is continuous on the compact set $[\bar{p}_1, x_1] \times [\bar{p}_2, x_2]$ and $f|_{[\bar{p}_1, x_1] \times [\bar{p}_2, x_2]}$ attains the maximum.

To show that $F$ has maximum on $[0, x_1] \times [0, x_2]$ it is sufficient to prove that $F$ has negative (and continuous) partial derivatives on $(0, \bar{p}_1) \times (0, \bar{p}_2)$, for some $\bar{p}_1$ and $\bar{p}_2$. Let us concentrate on the partial derivative with respect to $p_1$.

Firstly, we majorize $\phi^{(e)}$ by an integrable function $\hat{\phi}^{(e)}$ with the monotone tails in the sense that for every $i$ $\hat{\phi}^{(e)}(x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_N)$ is increasing for every $i$ $\hat{\phi}^{(e)}(x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_N)$ is increasing.
for sufficiently large negative arguments and decreasing for sufficiently large positive ones. It is always possible for the bounded density function. For some mild conditions we can swap the derivative and integrals.

\[ \frac{\partial F(p_1, p_2)}{\partial p_1} = \int_{\mathbb{R}} \frac{1}{2} \left( \int_{\mathbb{R}} \frac{1}{p_1 p_2} \phi^{(e)} \left( \frac{x - z}{p_1} \right) \phi^{(e)} \left( \frac{z}{p_2} \right) dz \right)^{-\frac{1}{2}} \cdot \int_{\mathbb{R}} \frac{1}{p_1^2} \left[ -\frac{1}{p_1^2} \phi^{(e)} \left( \frac{x - z}{p_1} \right) - \nabla \phi^{(e)} \left( \frac{x - z}{p_1} \right) \frac{x - z}{p_1^3} \right] \phi^{(e)} \left( \frac{z}{p_2} \right) dz dx. \]

The expression in the square brackets is negative for sufficiently small \( p_1 \). The first component is always negative. If \( x_i - z_i \geq 0 \) than for \( p_1 \) large enough \( \left( \frac{x_i - z_i}{p_1} \right) \nabla \phi^{(e)} \left( \frac{x - z}{p_1} \right) \leq 0 \). On the other hand, \( x_i - z_i < 0 \) implies

\[ \left( \frac{x_i - z_i}{p_1} \right) \nabla \phi^{(e)} \left( \frac{x - z}{p_1} \right) > 0. \]

Thus, on \([0, \bar{p}_1] \times [0, \bar{p}_2] \), \( F \) is bounded by a function that has the maximum value. This completes the proof.

A direct consequence of lemma 3.2 is the following estimate of the approximation error of the algorithm.

**Theorem 3.2.** For \( u_T(\cdot, z) \) \( L_u \)-Lipschitz continuous and \( u_T(\cdot, \cdot) \) bounded by \( L_u(b) \),

\[ |V(0, x, \varrho) - \hat{V}(0, \hat{x}, \hat{\varrho})|^2 < \frac{1}{1 - \delta} C_{n,M_0} + L_T^{(1)} |x - \hat{x}|^2 + L_T^{(2)} \int_{\mathbb{R}^K} |\varrho(z) - \hat{\varrho}(z)|^2 dz, \tag{15} \]

with

\[ C_{n,M_0} := \left( L_u n^{-2} J_1 \cdot \left( \int_{\mathbb{R}} f_X(x) \frac{1}{4} dx \right)^3 + 2 L_u(b) L_R^2 \sqrt{K} \left( \frac{M_0}{n} \right)^{K+1} + 4 L_u(b) M_0^K n^{-\frac{2}{7}} J_N \left( \int_{\mathbb{R}^N} f_{z_{k_0}}^\max(x) \frac{N}{\sqrt{N+2}} dx \right)^{N+2} \right) + 2 L_u(b) a_z v_N \left( \frac{M_0^{N-p}}{p - 1} \right), \]

\[ L_T^{(1)} := 2 T + \delta T L_u^2 \prod_{t=1}^T (1 + r(t)), \]

\[ L_T^{(2)} := 2 T + \delta T \left( L_u(b) \right)^2 L_\alpha^2 L_\varphi^2. \]
Proof: Let us suppose that the projections \( \hat{\cdot} \) realizes the minimal error over the set of projection sets and \( V(t + 1, \cdot, \cdot) \) is \( (\epsilon, L_x, L_\varrho) \)-Lipschitz continuous. In this case, applying lemma 3.2

\[
|V(t, x, \varrho) - \hat{V}(t, \hat{x}, \hat{\varrho})|^2 \\
\leq \left( \sup_{(c, \pi) \in \mathcal{A}(x) \cup \mathcal{A}(\hat{x})} \left\{ u(c) + \delta \mathbb{E} V(t + 1, x(R), \varrho(R)) - \right. \\
\left. + u(c) - \delta \mathbb{E} \hat{V}(t + 1, x(R), \varrho(R)) \right\} \right)^2 \\
\leq \sup_{(c, \pi) \in \mathcal{A}(x) \cup \mathcal{A}(\hat{x})} \left\{ \delta \mathbb{E} \left| V(t + 1, x(R), \varrho(R)) - \hat{V}(t + 1, x(R), \varrho(R)) \right|^2 \right\} \\
= \delta \epsilon + \delta C_{n,M} + 2\delta L_x (1 + r(t))^2 |x - \hat{x}|^2 + 2\delta L_\varrho L_\varrho^2 L_\varrho^2 \int_{\mathbb{R}^K} |\varrho(z) - \hat{\varrho}(z)|^2 dz.
\]

Thus, \( V(t, \cdot, \cdot) \) is \( (\delta \epsilon + \delta C_{n,M}, 2\delta L_x (1 + r(t))^2, 2\delta L_\varrho L_\varrho^2 L_\varrho^2) \)-Lipschitz continuous. By induction the proof is completed.

4. Numerical illustration

We show the performance of the algorithm for an arbitrary parametrization of the model. We consider the market for only one risky asset with the stochastic volatility driven by the one dimensional process \( y \). Even in this case, the computation of the optimal consumption / investment paths is very much time consuming for the algorithm implemented on a regular computer.

We assume that the dynamics of the market specified in general terms in the subsection is given by the following system of the three equations:

\[
S_0(t + 1) = (1 + 0.03)^{t+1}, \\
S_1(t + 1) = S_1(t) \exp \left\{ 0.05 + 0.25 (1 - 0.75 \exp \{-Y(t + 1)^2\}) \epsilon(t + 1) \right\}, \\
S_1(0) = 6.00, \\
Y(t + 1) = 0.80 \cdot Y(t) + 0.20 \xi(t + 1), \\
Y(0) = 1.50.
\]

The risk-free interest rate (the bank account interest) is equal to 3%. We assume that the mean log rate of return from the risky asset is 5%. The

\footnote{1.6MHz dual core processor with 2GB RAM.}
volatility of the risky asset can vary a lot in the model. It can be as large as almost 16% for $Y(t) = 0.8$ and as low as 6% if $Y(t) = 0.0$. For the parametrization in the equation trajectories of $Y$ usually oscillates within -0.8 and 0.8 bounds.

The variability of $Y$ is visibly reflected into the volatility of $S_1$. A sample trajectory of $S_1$ is shown on figure 1.

The utility functions meeting the requirements of boundedness and Lipschitz continuity demanded in the assumptions of the theorem 3.2 (in fact the lemma 3.2) were chosen to be the following:

$$u_T(y, x) = \left( \frac{2 \exp(-y)x}{1 + 2 \exp(-y)x} \right)^{0.5} \text{ and } u(x) = 0.25 \left( \frac{x}{2 + x} \right)^{0.5}. \quad (17)$$

This are the utility function of the so-called IRRA type, i.e. the increasing relative risk aversion.

The numerical algorithm to calculate the approximate optimal consumption / investment strategies was based on the backwards dynamic programming equation 11. We compute $\hat{V}$ recursively for 10 points of time with the time horizon $T = 10$. At each point of time equation give us the estimates of the consumption $c$ and the investment $\pi$ maximizing his joint utility.
Figure 2: The initial quantization set $Q_n^0$ (top graph) and the approximation of the optimal quantization set in the stochastic gradient descent methods in 500 steps (bottom graph)

Source: own calculations

Keeping in mind the complexity of the calculation of $\hat{V}(t, \hat{x}, \hat{\varrho})$ based on $\hat{V}(t+1, \ldots)$ we have to choose a limited number of nodes for the portfolio value $x$ and the beliefs $\varrho$. To make the calculation tractable, we assume that the quantization set of the portfolio process $X$ is $Q_X = \{0.00, 0.25, 0.50, \ldots, 9.75, 10.00\}$. The investor’s beliefs described by the unnormalized densities are projected on the appropriate quantization.

Since the minimal error in the approximate algorithm described by the recursion $[11]$ is realized for the optimally chosen quantization set we apply the stochastic gradient descent algorithm to find the optimal structure (compare Pagès and Pham [18]). We start with the initial arbitrary set of the densities $Q_n^0$ evaluated at the points from $Z_n := \{-1.50, -1.45, -1.40, \ldots, 1.45, 1.50\}$. We choose the parameters – the mean and the standard deviation – to be the equidistant point on the real line, i.e. the mean took values in the set $\{-1.50, -1.25, \ldots, 1.50\}$ and the standard deviation in the set $\{0.1, 0.3, 0.5, 0.7, 0.9\}$. The densities are presented on figure 2, top panel.

The procedure requires to construct the incremental function $H$ which defines the corrections of $Q_n^0$ leading in the limit to the optimal quantization set. The incremental function is derived by the formal differentiation under $\hat{\mathbb{E}}$ of the distortion function $D_n(Q)$ for a given realization $\hat{r}$ of return $R^t$. 

20
Figure 3: The selection of the essential elements of the quantization set $Q_n^{500}$ (25 out of 65 elements)

Source: own calculations

Namely, in the special case of 1-dimensional process $Y$ any quantization set $Q$ can be represented by the matrix from $M^{n \times m_n}$ where $(k, m)$ element is defined as the value of the $k$th element of the set $Q$ evaluated at the $m$th argument from $Z_n$. Then, the function $H : M^{n \times m_n} \times \mathbb{R} \to M^{n \times m_n}$ is defined as

$$H_{k,m}(Q, \hat{r}) : = \frac{\partial}{\partial \hat{r}} \varrho(z_k, \hat{r}) \left( \varrho(z_k, \hat{r}) - g_m(k) \right).$$

Since in the example the densities $\phi$ and $\psi$ in the example are smooth then the derivative of $\varrho$ is well-defined. Then for a given sequence $\{\beta_i\}_{i \in \mathbb{N}}$ satisfying the usual regularity conditions ($\sum \beta_i = \infty$ and $\sum \beta_i^2 < \infty$) the following sequence of the quantization sets converge to the optimal one (see Shapiro and Wardi [22] for convergence results):

$$Q_{n}^{m+1} = \max \left\{ Q_n^m + \beta_{m+1} H(Q_n^m, R_l(\omega_{m+1})), 0 \right\}.$$  

Applying this recursive procedure for $Q_n^m$ with 500 iterations of $m$ we obtain the modified set $Q_n^{500}$ which is graphically presented on figure 2 bottom panel. Firstly, one crucial advantage of $Q_n^{500}$ is the reduction of the domain of the densities from the initial $[-1.5, 1.5]$ to $[-0.5, 0.5]$ where they are essentially different from 0. This decreases the dimensionality of the numerical problem since the same number of points $Z_n$ can be denser allocated on the smaller interval $[-0.5, 0.5]$. Secondly, we can drop the densities from $Q_n^{500}$ which are very similar to each other and we take into account only a subset $Q_n^{ess} \subset Q_n^{500}$ if the densities. Namely, we apply the following procedure:

1. $Q_n^{ess} : = Q_n^{500}$; eps=0.1;
2. REPEAT i
Figure 4: Evolution of beliefs with prior at time $T - 9$ equal to the normal distribution with mean $0.1$ and standard deviation $0.2$.

Note: Each box shows a density which is recursively obtained using equation 7, i.e. the density on the box $T - n$ is obtained from the formula 7 with $\varrho_{T-n-1}$ equal to the density from the box $T - n - 1$ and $\varrho_{T-9} \equiv \phi_{0.1,0.2}$.

Source: own calculations

3. Chose a pair of densities $q_1 \in \hat{Q}^{ess}$ and $q_2 \in \hat{Q}^{ess}$;
4. IF $\frac{\sum_{z \in \mathbb{Z}_n} |q_1(z) - q_2(z)|}{\sum_{z \in \mathbb{Z}_n} |q_1(z)|} < \text{eps}$ THEN $Q_n^{ess} : = Q_n^{ess} / \{q_2\}$;
5. i := i + 1;
6. UNTIL i == 5000;

the application of which gives the quantization set consisting of 25 unnormalized densities.

It should be underlined that functions $\varrho$ representing the beliefs about $Y$, described by recursion 7, are not necessarily densities. They are unnormalized densities. An example of the evolution of beliefs is presented on the figure 4.

The number of nodes of the portfolio values, which is 41, multiplied by the number of the densities in the density quantization set equal to 25 gives 1025 arguments of $\hat{V}(t, \cdot, \cdot)$.

Searching for the optimal consumption and investment for a given point of time $t$, the current wealth $\bar{x}$ and the beliefs about the unobserved factors driving the volatility on the market, we discretize the space of investors deci-
We were looking for the maximum expected value of the value function $\hat{V}(t+1, x^c, \pi(t+1), \rho_{t+1})$ among $c \in \{0.00\bar{x}, 0.05\bar{x}, 0.10\bar{x}, \ldots, 0.95\bar{x}, 1.00\bar{x}\}$ and $\pi \in \{0.00\bar{x}, 0.05\bar{x}, 0.10\bar{x}, \ldots, 0.95\bar{x}, 1.00\bar{x}\}$.

The four main components of the solution of the numerical optimization were presented: the consumption, the investment in the risky asset and the risk-free asset and the value of the optimal portfolio. Assuming that the investor possesses 6 units of the initial capital we generated his wealth paths along the estimated consumption and investment policies. We perform 1000 simulations of the risky asset prices and we construct the corresponding consumption and investment paths according to the approximate optimal policy.

The results of the simulation suggest that the optimally behaving investor consume about half of the initial wealth during the first two periods and then gradually decreases the consumption rate (see figure 5). In the period $t = 6$ the investor is left with less then 1 units of wealth. In the last three periods he consumes only about 0.1 units of his portfolio, almost irrespective of the market situation (i.e. independent of the evolution of the risky asset prices).

Conversely, his investment decisions are much more dependent on the market trends and on the time of decision making. His stronger beliefs about
Figure 6: Distribution of the optimal investment into the risky asset for different points of time

Note: histograms for 1000 simulation of the risky asset dynamics. $T - n$ means $n$ time periods before the time horizon $T$.
Source: own calculation

unfavorable volatility (about $y$) decrease his propensity to invest into the risky asset. The investor allocates quite a small part of his wealth in the risky project $S_1$ only in the first two periods (see figure 6). The closer to the terminal time $T$ the more he is inclined to allocate his current wealth into the risk-free asset $S_0$ (see figure 7). As the figure 8 indicates, the terminal wealth reaches almost 0 units.

Even though the set of beliefs seems to be only roughly approximated by 30 density functions, the simulation reflects the behavior of a rational investor in a quite reasonable and intuitive way. However, the computational burden of the numerical procedure makes it difficult to support investment decisions on a market for multiple assets unless the complexity of the stochastic volatility part of the model is substantially reduced. In fact the minimal error of the estimate of the value function $V$ that can be attained in the numerical example and given by the constant $C_{n,M_0}$ multiplied by the factor $\delta_{1-\delta^T}$ (see theorem 3.2) is equal to 0.81 which is about 12% of the utility of the initial wealth.

However, the proposed quantization framework follows the promising av-
Figure 7: Distribution of the optimal investment into the risk-free asset for different points of time

Note: histograms for 1000 simulation of the risky asset dynamics. $T - n$ means $n$ time periods before the time horizon $T$.
Source: own calculation

Figure 8: Distribution of the optimal value of portfolio for different points of time

Note: histograms for 1000 simulation of the risky asset dynamics. $T - n$ means $n$ time periods before the time horizon $T$.
Source: own calculation
enue of research on the portfolio optimization under the partial observation of
the stochastic volatility. It shows that the nonlinear filtering can be tractable
even in the dynamic programming setting. From the financial point of view,
the incomplete information is a very important driver of the market dynamics
and its influence on the economic agents’ optimal choice is worse studying.
Certainly, a more efficient numerical implementation of the algorithm is re-
quired.

5. Conclusions

We solved the dynamic optimal consumption/investment strategy under
partial observation of economic factors determining asset price movements
by means of the numerical method. The special variation of the quantization
technique reduced the originally infinite dimensional problem to the finite
state space exercise. We successfully developed the algorithm to solve the
consumption / investment problem on the stylized market for 2 tradable
assets and 1 unhedgable economic factor.

The extension of the paper can be devoted to a study of the stability of the
algorithm solving numerically the optimal portfolio choice with consumption.
The first important question remains how the results of the simulation change
between two different sets of randomly generated asset returns. The second
one could be related to the size of quantization set that would in practice
give an assumed accuracy of numerical optimal strategies.

We transformed the partial observation to the full observation problem,
assuming that the shocks affecting prices and economic factors are indepen-
dents. However, existing methods of stochastic volatility estimation [11] do
not need shocks to be orthogonal. It would be interesting to develop a model
of consumption/investment with imperfect observation and correlation of
volatility and idiosyncratic shocks in the asset price movements.

Appendix A. Filters

The following lemma is the most crucial in obtaining filters of partially
observed processes. It states that $Y$s and $R^i$’s are independent in $(\mathcal{F}^{IR}, \tilde{P})$.
After disentangling $Y$ from $R^i$, the evolution of the variable $Y$ can be refor-
mulated to a recursive equation of measures on $(\Omega, \mathcal{F}^{IR})$.

Lemma Appendix A.1. The processes $Y$ and $R^i$ are independent in $\mathcal{F}^{IR}$
under $\tilde{P}$.
**Proof:** Let us take two Borel functions \( f_1 \) and \( f_2 \). Then

\[
E \left[ f_1(Y(t))f_2(R^l(t)) \middle| \mathcal{F}_{t-1}^{IR} \right] = \frac{E \left[ \Lambda_t^{-1}f_1(Y(t))f_2(R^l(t)) \middle| \mathcal{F}_{t-1}^{IR} \right]}{E \left[ \Lambda_t^{-1} \middle| \mathcal{F}_{t-1}^{IR} \right]} = \frac{E \left[ \lambda_t^{-1}f_1(Y(t))f_2(R^l(t)) \middle| \mathcal{F}_{t-1}^{IR} \right]}{E \left[ \lambda_t^{-1} \middle| \mathcal{F}_{t-1}^{IR} \right]} \tag{A.1}
\]

since \( \Lambda_t^{-1} = \lambda_t^{-1}\Lambda_{t-1}^{-1} \) and \( \Lambda_{t-1}^{-1} \) is \( \mathcal{F}_{t-1}^{Y} \)-measurable. For brevity \( \tilde{y} := Y(t-1) \).

Let us take the denominator in (A.1) and

\[
E \left[ \lambda_t^{-1} \middle| \mathcal{F}_{t-1}^{IR} \right] = E \left[ \frac{\det(\sigma(Y(t)))\phi(R^l(t))}{\phi(\sigma^{-1}(Y(t))[R^l(t) - \mu(Y(t))])} \frac{\det(\sigma_Y)(Y(t))}{\psi(\sigma_Y^{-1}[Y(t) - \tilde{y} - \alpha(\tilde{y} - \bar{y})])} \right] \mathcal{F}_{t-1}^{IR}
\]

\[
= E \left[ \frac{\phi(\mu(Y(t)) + \sigma(Y(t))\epsilon(t))}{\det(\sigma^{-1}(Y(t)))\phi(\epsilon(t))} \frac{\psi(\tilde{y} - \alpha(\tilde{y} - \bar{y}) + \sigma_Y\xi(t))}{\det(\sigma_Y^{-1})\psi(\xi(t))} \right] \mathcal{F}_{t-1}^{IR}.
\]

Put

\[
p_1 := \frac{\psi(\tilde{y} - \alpha(\tilde{y} - \bar{y}) + \sigma_Y\xi(t))}{\det(\sigma_Y^{-1})\psi(\xi(t))}.
\]

Then

\[
E \left[ \lambda_t^{-1} \middle| \mathcal{F}_{t-1}^{IR} \right] = E \left[ p_1E \left[ \frac{\phi(\mu(Y(t)) + \sigma(Y(t))\epsilon(t))}{\det(\sigma^{-1}(Y(t)))\phi(\epsilon(t))} \frac{\sigma(\mathcal{F}_{t-1}^{IR}, \mathcal{F}_{t-1}^{Y})}{\mathcal{F}_{t-1}^{IR}} \right] \mathcal{F}_{t-1}^{IR} \right].
\]

The variable \( \epsilon(t) \) is independent of \( \mathcal{F}_{t-1}^{IR} \) so we can integrate with respect to the density \( \phi \)

\[
E \left[ \lambda_t^{-1} \middle| \mathcal{F}_{t-1}^{IR} \right] = E \left[ p_1 \int_{\mathbb{R}^N} \frac{\phi(\mu(Y(t)) + \sigma(Y(t))\epsilon)}{\det(\sigma^{-1}(Y(t)))\phi(\epsilon)} \phi(\epsilon) d\epsilon \middle| \mathcal{F}_{t-1}^{IR} \right].
\]
By changing variables with the diffeomorphism \( \Theta(\dot{\epsilon}) = \sigma^{-1}(Y(t))\dot{\epsilon} \), with a
given fixed \( \omega \), as follows \( \int_{\Theta(\mathbb{R}^N)} G(\dot{\epsilon})d\dot{\epsilon} = \int_{\mathbb{R}^N} (\Theta \circ G)(\dot{\epsilon}) \det(\Theta'(\dot{\epsilon}))d\dot{\epsilon} \) we get

\[
E \left[ \lambda_t^{-1} | F_{t-1}^{IR} \right] = E \left[ p_1 \int_{\mathbb{R}^N} \phi(\mu(Y(t)) + \dot{\epsilon})d\dot{\epsilon} \bigg| F_{t-1}^{IR} \right] 
= E \left[ \frac{\psi(\tilde{y} - \alpha(\tilde{y} - y) + \sigma_Y \xi(t))}{\det(\sigma^{-1}_Y)\psi(\xi(t))} \bigg| \sigma(F_{t-1}^{IR}, F_Y) \bigg| F_{t-1}^{IR} \right] 
= E \left[ \int_{\mathbb{R}^K} \frac{\psi(\tilde{y} - \alpha(\tilde{y} - y) + \sigma_Y \xi)}{\det(\sigma^{-1}_Y)\psi(\xi)} \psi(\hat{\xi})d\hat{\xi} \bigg| F_{t-1}^{IR} \right] = 1,
\]

observing that \( \xi(t) \) is independent of \( F_{t-1}^{IR} \).

Let us take

\[
E \left[ \lambda_t^{-1} f_1(Y(t))f_2(R^l(t)) \bigg| F_{t-1}^{IR} \right] = 
= E \left[ p_1 f_1(Y(t))f_2(R^l(t)) \frac{\phi(\mu(Y(t)) + \sigma(Y(t))\dot{\epsilon})}{\det(\sigma^{-1}(Y(t)))\phi(\dot{\epsilon})} \bigg| F_{t-1}^{IR} \right] 
= E \left[ p_1 f_1(Y(t))E \left[ f_2(R^l(t)) \frac{\phi(\mu(Y(t)) + \sigma(Y(t))\dot{\epsilon})}{\det(\sigma^{-1}(Y(t)))\phi(\dot{\epsilon})} \bigg| \sigma(F_{t-1}^{IR}, F_Y) \bigg| F_{t-1}^{IR} \right] \right].
\]

A similar argument about independence of \( \dot{\epsilon}(t) \) and \( F_{t-1}^{IR} \) gives

\[
E \left[ \lambda_t^{-1} f_1(Y(t))f_2(R^l(t)) \bigg| F_{t-1}^{IR} \right] = 
= E \left[ f_1(Y(t))p_1 \int_{\mathbb{R}^N} f_2(\mu(Y(t)) + \sigma(Y(t))\dot{\epsilon}) \frac{\phi(\mu(Y(t)) + \sigma(Y(t))\dot{\epsilon})}{\det(\sigma^{-1}(Y(t)))} d\dot{\epsilon} \bigg| F_{t-1}^{IR} \right] 
and after the diffeomorphic change of variables

\[
= E \left[ f_1(Y(t))p_1 E \left[ f_2(R^l(t)) \bigg| F_{t-1}^{IR} \right] \bigg| F_{t-1}^{IR} \right] 
= E \left[ f_1(Y(t)) \frac{\psi(\tilde{y} - \alpha(\tilde{y} - y) + \sigma_Y \xi(t))}{\det(\sigma^{-1}_Y)\psi(\xi(t))} E \left[ f_2(R^l(t)) \bigg| F_{t-1}^{IR} \right] \bigg| F_{t-1}^{IR} \right] 
= E \left[ \int_{\mathbb{R}^K} f_1(\tilde{y} - \alpha(\tilde{y} - y) + \sigma_Y \xi) \frac{\psi(\tilde{y} - \alpha(\tilde{y} - y) + \sigma_Y \xi)}{\det(\sigma^{-1}_Y)\psi(\xi)} d\xi \cdot E \left[ f_2(R^l(t)) \bigg| F_{t-1}^{IR} \right] \bigg| F_{t-1}^{IR} \right] 
= E \left[ E \left[ f_1(Y(t)) \bigg| F_{t-1}^{IR} \right] \cdot E \left[ f_2(R^l(t)) \bigg| F_{t-1}^{IR} \right \bigg| F_{t-1}^{IR} \right] 
= E \left[ f_1(Y(t)) \bigg| F_{t-1}^{IR} \right] \cdot E \left[ f_2(R^l(t)) \bigg| F_{t-1}^{IR} \right].
\]

28
Taking \( f_1(z) = 1_A(z) \) and \( f_2(z) = 1_B(z) \) for \( A, B \in \mathcal{F}_t^{IR} \) we obtained
\[
P \left( \{ R'(t) \in A \} \cap \{ Y(t) \in B \} \right| \mathcal{F}_{t-1}^{IR} ) = \mathbb{E} \left[ f_1(R'(t))f_2(Y(t)) \right| \mathcal{F}_t^{IR} ] = \\
= \mathbb{E} \left[ f_1(R'(t)) \right| \mathcal{F}_t^{IR} ] \mathbb{E} \left[ f_2(Y(t)) \right| \mathcal{F}_t^{IR} ] = P \left( R'(t) \in A \right| \mathcal{F}_{t-1}^{IR} ) P \left( Y(t) \in B \right| \mathcal{F}_{t-1}^{IR} ) .
\]

**Proof:** [Theorem 2.1] For an arbitrary Borel function \( f \) let us take
\[
\int_{\mathbb{R}^K} f(z) \varrho_t(z)dz = \mathbb{E} \left[ \Lambda_t f(Y(t)) \right| \mathcal{F}_t^{IR} ] = \\
= \mathbb{E} \left[ \Lambda_{t-1} \frac{\phi \left( \sigma^{-1}(Y(t)) \left[ R^{(l)}(t) - \mu(Y(t)) \right] \right) \psi \left( \sigma_Y^{-1}[Y(t) - \bar{y} - \alpha(\bar{y} - \bar{y})] \right) f(Y(t)) \right| \mathcal{F}_t^{IR} ] \\
= \mathbb{E} \left[ \Lambda_{t-1} \mathbb{E} \left[ \frac{\phi \left( \sigma^{-1}(Y(t)) \left[ R^{(l)}(t) - \mu(Y(t)) \right] \right) \psi \left( \sigma_Y^{-1}[Y(t) - \bar{y} - \alpha(\bar{y} - \bar{y})] \right) f(Y(t)) \right| \mathcal{F}_t^{IR} , \mathcal{F}_{t-1}^{IR} ] \right| \mathcal{F}_t^{IR} ] \\
= \mathbb{E} \left[ \Lambda_{t-1} \int_{\mathbb{R}^K} \frac{\phi \left( \sigma^{-1}(z) \left[ R^{(l)}(t) - \mu(z) \right] \right) \psi \left( \sigma_Y^{-1}[z - \bar{y} - \alpha(\bar{y} - \bar{y})] \right) f(z)\psi(z)dy \right| \mathcal{F}_t^{IR} ] .
\]

The inner process is the function of \( Y(t - 1) \) and \( \Lambda_{t-1} \) so we can integrate with respect to \( \varrho_{t-1}(z) \) in the following
\[
\int_{\mathbb{R}^K} f(z) \varrho_t(z)dz = \int_{\mathbb{R}^K} \left( \int_{\mathbb{R}^K} \frac{\Phi(z,R'(t))}{\phi(R'(t))} \Psi(z,y)f(z)dz \right) \varrho_{t-1}(y)dy \\
= \int_{\mathbb{R}^K} f(z) \frac{\Phi(z,R'(t))}{\phi(R'(t))} \left( \int_{\mathbb{R}^K} \Psi(z,y)\varrho_{t-1}(y)dy \right) dz .
\]

Since \( f \) is the arbitrary function the identity \[7\] holds.

**Appendix B. The constant \( v_N \) in the error estimate in lemma 3.3**

Changing the limits of integration in the integral
\[
\int_{[M_0, +\infty)^N} |x|^{-p}dx
\]
to the polar ones with

\[ F(w, \alpha_1, \ldots, \alpha_N) = \begin{bmatrix}
    w \cos \alpha_1 \cos \alpha_2 \cos \alpha_3 \ldots \cos \alpha_{N-1} \\
    w \sin \alpha_1 \cos \alpha_2 \cos \alpha_3 \ldots \cos \alpha_{N-1} \\
    w \sin \alpha_2 \cos \alpha_3 \ldots \cos \alpha_{N-1} \\
    \vdots \\
    w \sin \alpha_{N-2} \cos \alpha_{N-1} \\
    w \sin \alpha_{N-1}
\end{bmatrix} \]

we obtain

\[
\int_{[0,\infty)^{N}} \left[ x \right]^{-p} \, dx = \int_{[0,\infty)^{N} \times [0,2\pi]} \left( \frac{r^{-p}}{\det F'} \right) \, dw \, \omega_1 \ldots \omega_{N-1} = \\
\int_{[0,\infty)^{N}} w^{-p} w^{N-1} \int_{[0,2\pi]} 1 \, \omega_1 \int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} \cos \omega_2 \, \omega_2 \ldots \int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} (\cos \omega_{N-1})^{N-2} \, d\omega_{N-1}.
\]

Let us denote \( I_n = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n t \, dt \). Note that \( I_0 = \pi \) and \( I_1 = 2 \). Integrating by parts

\[
\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos t \cos^{n-1} t \, dt = \sin t \cos^{n-1} t \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2 t \cos^{n-2} t \, dt \\
I_n = 0 + n I_{n-1} - n I_n,
\]

thus

\[
n = 2k \quad \Rightarrow \quad I_n = \frac{n!!}{(n+1)!!} \pi, \\
n = 2k + 1 \quad \Rightarrow \quad I_n = 2 \frac{n!!}{(n+1)!!}.
\]

Then

\[
v_N = \frac{2\pi}{N - p} \prod_{j=2}^{N-1} \int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} (\cos t)^{j-1} \, dt \\
= \frac{2\pi}{N - p} \prod_{j=1}^{N-2} 2^{j \mod 2} \pi^{(j+1) \mod 2} \frac{j!!}{(j+1)!!} \\
= \begin{cases} 
    \frac{2\pi}{(N-p)(N-2)!!} (2\pi)^{\frac{N-2}{2}}, & \text{if } N = 2k \\
    \frac{2\pi}{(N-p)(N-2)!!} 2^{\frac{N-3}{2}} \pi^{\frac{N-3}{2}+1}, & \text{if } N = 2k + 1
\end{cases}
\]

(B.2)
Acknowledgements

The paper is an extension of the part of my doctoral thesis completed under supervision of Prof. Lukasz Stetter. I would like to thank the participants of the Third General AMAMEF Conference “Advances in Mathematical Finance”, May 5-10, 2008, Pitesti, Romania, for their valuable comments to the preliminary results which were presented there. I also express my gratefulness to Tom Hurd and the Fields Institute in Toronto for invitation to the postdoctoral program and the financial support that contributed a lot to essentially improving and finishing the paper. Many thanks to my wife Ania for the interminable support in my research.

References

[1] F. Baghery, B. Øksendal A maximum principle for stochastic control with partial information. Stochastic Analysis and Applications 25 (2007) 705–717.

[2] O.E. Barndorff-Nielsen, E. Nicolato, N. Shephard, Some recent developments in stochastic volatility modeling, Quantitative Finance 2(1) (2002) 11–23.

[3] V. Beneš, I. Karatzas, D. Ocone, H. Wang, Control with Partial Observations and an Explicit Solution of Mortensen’s Equation, Applied Mathematical Optimisation, 49 (2002) 217–239.

[4] A. Bensoussan, Stochastic Control of Partially Observable Systems, Cambridge University Press, Cambridge, 1992.

[5] D. Bertsekas, Dynamic Programming and Stochastic Control, New York: Academic Press, 1976.

[6] S. Brendle, Portfolio Selection under Partial Observation and Filtering, preliminary version of working paper, Prinston University (2005) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=590362.

[7] R. Carmona, M. Ludkovski, Convenience yield model with partial observations and exponential utility, Princeton working paper (2004).
[8] M. Corsi, H. Pham, W. Runggaldier, *Numerical approximation by quantization of control problems in finance under partial observations*, Mathematical modeling and numerical methods in finance, in A. Bensousan, Q. Zhang (Eds.), Handbook of Numerical analysis (special volume), 2008.

[9] P. Del Moral, J. Jacod, Ph. Protter, *The Monte-Carlo methods for filtering with discrete time observations*, Probability Theory and Related Fields 120 (2001) 346–368.

[10] R. Desai, T. Lele., F. Viens, *A Monte-Carlo method for portfolio optimization under partially observed stochastic volatility*, IEEE International Conference on Computational Intelligence for Financial Engineering, Proceedings 2003.

[11] G. Durham, *Monte Carlo methods for estimating, smoothing and filtering one- and two-factor stochastic volatility models*, Journal of Econometrics 133 (2007) 273–305.

[12] R. Elliot, H. Miao, *Stochastic Volatility Model with Filtering*, Stochastic Analysis and Applications 24 (2006) 661–683.

[13] G. Halaj, *Risk-based decisions on the asset structure of a bank under economic information*, Applied Mathematical Finance 15(4) (2008) 305–329.

[14] V. Krishnamurthy, F. Le Gland, *the Discretization of Continuous-Time Filters and Smoothers for HMM Parameter Estimation*, IEEE Transactions on Information Theory 42(3) (1996) 593–605.

[15] P. Lakner, *Utility maximization with partial information*, Stochastic Processes and their Application 56 (1995) 247–273.

[16] G. Pagès, H. Pham, J. Printemps, *An Optimal Markovian Quantization Algorithm for Multidimensional Stochastic Control Problems*, Stochastics and Dynamics 4(4) (2004) 501–545.

[17] G. Pagès, H. Pham, J. Printemps, *Optimal quantization methods and applications to numerical problems in finance*, in: S. Rachev (Ed.), Handbook on Numerical Methods in Finance, Birkhauser, Boston, 2004, pp. 253–298.
[18] G. Pagès, H. Pham, *Optimal quantization methods for nonlinear filtering with discrete-time observations*, Bernoulli 11(5) (2005) 893–932.

[19] H. Pham, M.-C. Quenez, *Optimal Portfolio in Partially Observed Stochastic Volatility Models*, The Annals of Applied Probability 11(1) (2001) 210–238.

[20] W. Runggaldier, L. Stettner, *Approximation of Discrete Time Partially Observed Control Problems*, Applied Mathematics Monographs CNR, Pisa, 1994.

[21] J. Sass, U.G. Haussmann, *Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain*, Finance and Stochastics 8 (2004) 553–577.

[22] A. Shapiro, Y. Wardi, *Convergence Analysis of Gradient Descent Stochastic Algorithms*, Journal of optimization theory and applications 91(2) (1996) 439–454.

[23] H. Sørensen, *Simulated Likelihood Approximations for Stochastic Volatility Models*, Scandinavian Journal of Statistics 30(2) (2003) 257–276.

[24] L. Stettner, *Risk-Sensitive Portfolio Optimization With Completely and Partially Observed Factors*, IEEE Transactions on Automatic Control 49(3) (2004) 457–464.