Optional decomposition and Lagrange multipliers

H. Föllmer[1], Yu.M. Kabanov[2]

[1] Institut für Mathematik, Humboldt Universität, Unter den Linden 6, D-10099 Berlin, Germany
[2] Central Economics and Mathematics Institute of the Russian Academy of Sciences, Moscow, and Laboratoire de Mathématiques, Université de Franche-Comté, 16 Route de Gray, F-25030 Besançon Cedex, France[*]

Abstract. Let $\mathcal{Q}$ be the set of equivalent martingale measures for a given process S, and let X be a process which is a local supermartingale with respect to any measure in $\mathcal{Q}$. The optional decomposition theorem for X states that there exists a predictable integrand φ such that the difference $X - \varphi \cdot S$ is a decreasing process. In this paper we give a new proof which uses techniques from stochastic calculus rather than functional analysis, and which removes any boundedness assumption.

1 Introduction

Let S be an $\mathbf{R}^d$-valued right-continuous semimartingale given on a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_t), P)$ with the usual assumptions. We denote by $\mathcal{Q}$ the set of all probability measures Q such that $Q \sim P$ and S is a local martingale with respect to Q. For a predictable process φ such that the stochastic integrals with respect to S are well defined, we denote by $\varphi \cdot S$ a right-continuous version of the process defined by $(\varphi \cdot S)_t = \int_0^t \varphi_s \, dS_s$.

Theorem 1 *Assume that $\mathcal{Q} \neq \emptyset$. Let X be a right-continuous process which is a local supermartingale with respect to any $Q \in \mathcal{Q}$. Then there exist an increasing right-continuous adapted process C with $C_0 = 0$ and a predictable integrand φ such that $X = X_0 + \varphi \cdot S - C$.*

Note that the conclusion of the theorem can be reformulated as follows: *There exists a predictable process φ such that the difference $X - \varphi \cdot S$ is a decreasing process*. Note also that if X is bounded from below, then φ is *admissible* in the sense that the stochastic integral $\varphi \cdot S$ is bounded from below.

In contrast to the standard Doob–Meyer decomposition, the process C is in general not predictable but only optional, and it is not uniquely determined. On the other hand, the decomposition in Theorem 1 is "universal" in the sense that it holds simultaneously for any probability measure $Q \in \mathcal{Q}$.

The existence of such an "optional decomposition" was shown by El Karoui and Quenez in [5] for a special class of models; see also [11] and the references given there. Kramkov [12] proved existence of an optional decomposition in a general semimartingale context, but under the assumptions that S is locally bounded and $X \geq 0$. The aim of this note is to prove the theorem in full generality, and to give an interpretation of the integrand values as Lagrange multipliers for some optimization problem with constraints. We follow a probabilistic approach in the spirit of [5] which uses methods of stochastic calculus rather than functional analysis and exploits the specific structure of the set of local densities.

In [5] and [11], optional decompositions arise in the context of incomplete financial markets. There, the process S describes the stochastic price fluctuation of some underlying financial assets; note that our results remove previous assumptions of local boundedness and thus permit the inclusion of models with unbounded jumps as they appear, e.g., in [4] and [1]. The process X is defined in terms of essential suprema of conditional expectations of a given $\mathcal{F}_T$-measurable contingent claim $H \geq 0$ over the class of all equivalent martingale measures, i.e., X is a right-continuous version of the process given by

$$X_t = \text{ess.sup}_{Q \in \mathcal{Q}} E_Q[H \,|\, \mathcal{F}_t].$$

It follows that X is a supermartingale with respect to any $Q \in \mathcal{Q}$. The point of the optional decomposition is to identify X as the value process of a strategy of "superhedging". The integrand φ specifies the amounts invested in the underlying assets. This strategy induces a perfect hedge $X_T = H$, and it generates an increasing process $C = X - \varphi \cdot S$ of cumulative side payments. Thus, the strategy always stays on the safe side. Such an approach to the problem of hedging in incomplete markets may seem rather "extreme". In fact, in various incomplete market models such as the multinomial extension of the binomial model, the superhedging strategy can be identified with the unique strategy of perfect replication in an associated extremal complete model which is no longer equivalent to the initial measure P. In models based on Lévy processes studied in [4], the strategy of superhedging for a call option simply reduces to the trivial strategy of just holding the underlying asset. Nevertheless, there are good reasons to investigate the structure of superhedging strategies and of the corresponding optional decompositions, not only from a mathematical but also from an applied financial point of view. For example, superhedging strategies appear as building stones in the construction of strategies which maximize the probability of a successful hedge $X_T \geq H$ under some given constraint on the initial portfolio value; see

[7]. In [6] the technique of superhedging is applied to models where volatility is stochastic but respects some a priori bounds. Further examples can be found in [13]. For an application of the optional decomposition to the arbitrage pricing theory for large financial markets see [9].

2 The discrete time version

In order to illustrate the basic idea of our construction we begin by proving the following discrete time version of Theorem 1 which does not need advanced stochastic calculus:

Theorem 2 *Let $S = (S_n)$ be an adapted process with values in $\mathbf{R}^d$ on $(\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n), P)$, $0 \leq n \leq N$. Let $\mathcal{Q}$ be the set of all probability measures equivalent to P such that S is a local martingale with respect to Q and assume that $\mathcal{Q} \neq \emptyset$. Let $X = (X_n)$ be a process which is a local supermartingale with respect to any $Q \in \mathcal{Q}$. Then there exist an increasing adapted process C with $C_0 = 0$ and a predictable process φ such that $X = X_0 + \varphi \cdot S - C$.*

The structure of our proof is the following. We have to show the existence of a predictable process φ such that $\Delta X_n - \varphi_n \Delta S_n \leq 0$ where $\Delta X_n := X_n - X_{n-1}$. It is easy to reduce this problem to one period; details are given after Lemma 3. The one-stage problem is first treated in the particular situation where the initial σ-algebra is trivial. In this case, the set $\mathcal{Q}$ is given by all probability measures $Q \sim P$ such that $E_Q|\eta| < \infty$, $E_Q\eta = 0$ for a given $\mathbf{R}^d$-valued random variable η (corresponding to ΔS_n). For a given scalar random variable ξ (corresponding to ΔX_n), the supermartingale assumption means that $E_Q\xi \leq 0$ for all $Q \in \mathcal{Q}$. We need to show that there is a vector $\lambda^* \in \mathbf{R}^d$ (corresponding to $-\varphi_n$) such that $\xi + \lambda^*\eta \leq 0$ a.s. In Lemma 1 we restate the problem in terms of the joint distribution of (η, ξ) and show that λ^* does exist and can be chosen from a set $\mathscr{L}$ of Lagrange multipliers for an associated optimization problem. In Lemma 2 we consider the general case where the initial σ-algebra is no longer trivial. In this case, the proof consists in combining the construction of Lagrange multipliers with a measurable selection argument.

We shall use the notation $m(f)$ for the integral $\int f dm$ and π for the natural projection of $\mathbf{R}^{d+1}$ to the first d coordinates, i.e.

$$\pi(x) := (x^1, \ldots, x^d)$$

for $x = (x^1, \ldots, x^d, x^{d+1}) \in \mathbf{R}^{d+1}$.

In view of measurable selection problems appearing in the proof of Lemma 2 it is convenient to work with the Polish space $C(\bar{\mathbf{R}}^{d+1})$ of all continuous real-valued functions on the one-point compactification of $\mathbf{R}^{d+1}$.

Lemma 1 *Let m be a probability measure on $\mathbf{R}^{d+1}$. Let G be the set of all functions $g \in C(\bar{\mathbf{R}}^{d+1})$ with $g > 0$, $m(g) = 1$, $m(|x|g) < \infty$. Assume that $G^0 := G \cap \{g : m(\pi g) = 0\} \neq \emptyset$ and $m(x^{d+1}g) \leq 0$ for all $g \in G^0$. Then*

(*a*) *we have*

$$\inf_{\lambda \in \mathbf{R}^d} \sup_{g \in G}(m(x^{d+1}g) + \lambda m(\pi g)) = \sup_{g \in G^0} m(x^{d+1}g); \qquad (2.1)$$

(*b*) *the set* $\mathscr{L}$ *of all* $\lambda^* \in \mathbf{R}^d$ *such that*

$$\sup_{g \in G}(m(x^{d+1}g) + \lambda^* m(\pi g)) = \inf_{\lambda} \sup_{g \in G}(m(x^{d+1}g) + \lambda m(\pi g)) \qquad (2.2)$$

is nonempty;
(*c*) *for any* $\lambda^* \in \mathscr{L}$ *we have* $x^{d+1} + \lambda^* \pi(x) \leq 0$ *m-a.s.*

Proof. Let us consider the following optimization problem:

$$\text{maximize } f(g) := m(x^{d+1}g) \qquad (2.3)$$

under the constraints

$$m(\pi g) = 0, \qquad (2.4)$$

$$g \in G. \qquad (2.5)$$

Let f^* be the optimal value; $f^* \leq 0$ by the hypothesis. Following a well-known argument, we show now that for this problem there exists a Lagrange multiplier $\lambda^* \in \mathbf{R}^d$ removing the equality constraint, i.e.

$$\sup_{g \in G}(m(x^{d+1}g) + \lambda^* m(\pi g)) = f^*. \qquad (2.6)$$

Without loss of generality we assume that the components of $\pi(x)$ are linear independent elements of $L^0(m)$; otherwise the problem can be reduced to a lower dimension. By our assumption $G^0 \neq \emptyset$ there exists $g^0 \in G$ satisfying the constraint (2.4); for this g^0 and any $\lambda \in \mathbf{R}^d$ we have $m(\lambda \pi g^0) = 0$. Thus, if $\lambda \pi \leq 0$ m-a.s. then $\lambda \pi = 0$ m-a.s. and hence, by the assumed linear independence of the components of π, we have $\lambda = 0$.

Define the nonempty convex set

$$\Xi := \{(y_1, y_2) \in \mathbf{R}^1 \times \mathbf{R}^d : y_1 < m(x^{d+1}g), \ y_2 = m(\pi g) \text{ for some } g \in G\}.$$

The point $(f^*, 0)$ does not belong to Ξ. Hence, by the separation theorem there exists a nonzero vector $\Lambda = (\lambda_1, \lambda_2) \in \mathbf{R}^1 \times \mathbf{R}^d$ such that

$$\lambda_1 y_1 + \lambda_2 y_2 \leq \lambda_1 f^* \qquad (2.7)$$

for all (y_1, y_2) in the closure of Ξ. Since y_1 can be a negative number with arbitrary large absolute value, $\lambda_1 \geq 0$. If λ_1 were equal to zero, then we would have that for all $g \in G$, hence for all Borel functions $g \geq 0$ with $m(g) < \infty$ and $m(|x|g) < \infty$,

$$m(\lambda_2 \pi g) = \lambda_2 m(\pi g) \leq 0.$$

This means that $\lambda_2 \pi \leq 0$ m-a.s. As we observed, this inequality holds only when $\lambda_2 = 0$. But $\Lambda \neq 0$. Hence $\lambda_1 > 0$.

Put $\lambda^* := \lambda_2/\lambda_1$. The inequality (2.7) implies that

$$\sup_{g \in G}(m(x^{d+1}g) + \lambda^* m(\pi g)) \le f^*.$$

On the other hand, for any λ we have

$$\sup_{g \in G}(m(x^{d+1}g) + \lambda m(\pi g)) \ge \sup_{g \in G, \, m(\pi g)=0}(m(x^{d+1}g) + \lambda m(\pi g)) = f^* \qquad (2.8)$$

and so we have shown (2.6).

We infer from (2.6) and (2.8) that $\lambda^* \in \mathscr{L}$ and that (2.1) holds and this proves (a) and (b). For $\lambda^* \in \mathscr{L}$ it follows from (2.6) that $m((x^{d+1} + \lambda^*\pi)g) \le 0$ for all Borel functions $g \ge 0$ with $m(g) < \infty$ and $m(|x|g) < \infty$; this property implies (c).   □

Lemma 2 *Let $\mathscr{G}$ be a sub-σ-algebra of $\mathscr{F}$. Let ξ and η be random variables taking values in $\mathbf{R}$ and $\mathbf{R}^d$, respectively. Assume that $E(\xi z \mid \mathscr{G}) \le 0$ for all random variables $z > 0$ with $E(z|\mathscr{G}) = 1$, $E(|\eta z||\mathscr{G}) < \infty$, $E(|\xi z||\mathscr{G}) < \infty$, and $E(\eta z|\mathscr{G}) = 0$. Suppose that there is at least one such z. Then there exists a $\mathscr{G}$-measurable d-dimensional random variable λ^* such that $\xi + \lambda^*\eta \le 0$ P-a.s.*

Proof. Without loss of generality we assume that $(\Omega, \mathscr{F}, P)$ is a complete probability space, that $\mathscr{G}$ contains all null sets from $\mathscr{F}$, and that the random variables z appearing in the assumption are $\sigma\{\mathscr{G}, \xi, \eta\}$-measurable. Let $m(\omega, dx)$ be a regular conditional distribution of the $d+1$-dimensional random variable (η, ξ) with respect to $\mathscr{G}$. Then the hypothesis of the lemma can be formulated as follows:

for any strictly positive $\mathscr{G} \otimes \mathscr{B}^{d+1}$-measurable function g on $\Omega \times \mathbf{R}^{d+1}$ such that

$$\int g(\omega, x)m(\omega, dx) = 1, \qquad (2.9)$$

$$\int |x|g(\omega, x)m(\omega, dx) < \infty \quad P\text{-}a.s., \qquad (2.10)$$

$$\int \pi(x)g(\omega, x)m(\omega, dx) = 0 \quad P\text{-}a.s. \qquad (2.11)$$

we have

$$\int x^{d+1}g(\omega, x)m(\omega, dx) \le 0 \quad P\text{-}a.s., \qquad (2.12)$$

and there exists at least one strictly positive function g^0 satisfying (2.9) – (2.11).

To prove the lemma it is sufficient to find a $\mathscr{G}$-measurable random variable λ^* with values in $\mathbf{R}^d$ such that $x^{d+1} + \lambda^*\pi(x) \le 0$ $m(\omega, dx)$-a.s. for almost all ω. To this end let us show that there exists a $\mathscr{G}$-measurable set Γ with $P(\Gamma) = 1$ such that for any $\omega \in \Gamma$ we have the following property:

for any strictly positive function $g \in C(\bar{\mathbf{R}}^{d+1})$ such that

$$m(\omega, g) := \int g(x)m(\omega, dx) = 1, \qquad (2.13)$$

$$m(\omega, |x|g) := \int |x|g(x)m(\omega, dx) < \infty, \tag{2.14}$$

$$m(\omega, \pi g) := \int \pi(x)g(x)m(\omega, dx) = 0 \tag{2.15}$$

we have

$$m(\omega, x^{d+1}g) := \int x^{d+1}g(x)m(\omega, dx) \le 0. \tag{2.16}$$

Indeed, the set

$$B := \{(\omega, g) \in \Omega \times C(\bar{\mathbf{R}}^{d+1}) : \ g > 0, \ (2.13) - (2.15) \text{ hold but } (2.16) \text{ does not}\}$$

is $\mathscr{G} \otimes \mathscr{B}(C(\bar{\mathbf{R}}^{d+1}))$-measurable. Denote by Γ the complement of the projection of B onto Ω. By the measurable selection theorem, see e.g. [3], III.44-45, B admits a measurable selector, i.e. there exists a $\mathscr{G}$-measurable mapping $F : \Omega \to C(\bar{\mathbf{R}}^{d+1})$ such that $(\omega, F(\omega)) \in B$ for all $\omega \in \Gamma^c$ (recall that $(\Omega, \mathscr{G}, P)$ is assumed to be complete). Notice that the scalar function $F(\omega, x) := F(\omega)(x)$ being $\mathscr{G}$-measurable in ω and continuous in x is $\mathscr{G} \otimes \mathscr{B}^{d+1}$-measurable. Put $\tilde{g}(\omega, x) := g^0(\omega, x)$ for $\omega \in \Gamma$ and $\tilde{g}(\omega, x) := F(\omega, x)$ for $\omega \in \Gamma^c$. If $P(\Gamma^c)$ were not equal to zero, the function $\tilde{g}$ would violate the assumption $(2.9) - (2.12)$.

By our assumption, there exists a null set $N \in \mathscr{G}$ such that for all $\omega \notin N$ there exists a strictly positive $\mathscr{B}^{d+1}$-measurable function, namely, $g^0(\omega, .)$ such that the relations $(2.13) - (2.16)$ hold. But from Lemma 3 below it follows that for each such ω there is a function $g_1^0(\omega, .) \in C(\bar{\mathbf{R}}^{d+1})$ with the same properties. Thus, the hypotheses of Lemma 1 are satisfied for $m(\omega, .)$ when $\omega \notin N \cup \Gamma^c$. It follows that for any such ω there exists a Lagrange multiplier $\lambda^* \in \mathbf{R}^d$ from the nonempty set $\mathscr{L}(\omega)$ defined as in (b) of Lemma 1. It remains to show that one can choose representatives from the sets $\mathscr{L}(\omega)$ in a measurable way. To this end we notice that

$$\Phi : (\omega, g, \lambda) \mapsto m(\omega, x^{d+1}g) + \lambda m(\omega, \pi g)$$

is a $\mathscr{G} \otimes \mathscr{B}(C(\bar{\mathbf{R}}^{d+1})) \otimes \mathscr{B}^{d+1}$-measurable function, and that the sets

$$\Delta_N := \{(\omega, g) \in \Omega \times C(\bar{\mathbf{R}}^{d+1}) : \ g > 0, \ m(\omega, g) = 1, \ m(\omega, |x|g) \le N\}$$

and

$$\Delta := \{(\omega, g) \in \Omega \times C(\bar{\mathbf{R}}^{d+1}) : \ g > 0, \ m(\omega, g) = 1, \ m(\omega, |x|g) < \infty\}$$

belong to $\mathscr{G} \otimes \mathscr{B}(C(\bar{\mathbf{R}}^{d+1}))$; see [2], Lemma 2.5. Denote by $G_N(\omega)$ and $G(\omega)$ the ω-sections of Δ_N and Δ. Put $\tilde{\Phi}_N(\omega, g, \lambda) := \Phi(\omega, g, \lambda)$ if $(\omega, g) \in \Delta_N$ and $-\infty$ otherwise. Clearly, $\tilde{\Phi}_N$ is $\mathscr{G} \otimes \mathscr{B}(C(\bar{\mathbf{R}}^{d+1}))$-measurable. It follows that for any fixed λ the function

$$\phi_N(\omega, \lambda) := \sup_{g \in G_N(\omega)} \Phi(\omega, g, \lambda) = \sup_{g \in C(\bar{\mathbf{R}}^{d+1})} \tilde{\Phi}_N(\omega, g, \lambda)$$

is $\mathscr{G}$-measurable in ω; see e.g. the proof of IV.33 in [3]. For any fixed ω, the function $\phi_N(\omega, \cdot)$, as a supremum of linear functions, is convex in λ and, being

bounded, it is continuous in this variable. Hence, ϕ_N is $\mathscr{G} \otimes \mathscr{B}^d$-measurable. It follows that the function ϕ defined by $\phi(\omega, \lambda) := \sup_{g \in G(\omega)} \Phi(\omega, g, \lambda)$ has the same property. Thus, the set

$$\{(\omega, \lambda^*) : \lambda^* \in \mathscr{L}(\omega)\} = \{(\omega, \lambda^*) : \phi(\omega, \lambda^*) = \inf_{\lambda} \phi(\omega, \lambda)\}$$

belongs to $\mathscr{G} \otimes \mathscr{B}^d$. We have shown that the projection of this set onto Ω has full measure. Applying again the measurable selection theorem we obtain the existence of a $\mathscr{G}$-measurable selector $\lambda^*(\omega)$. □

Lemma 3 *Let m be a probability measure on $(\mathbf{R}^n, \mathscr{B}^n)$ and let g be a strictly positive $\mathscr{B}^n$-measurable function on $\mathbf{R}^n$ with $m(g) = 1$ and $m(|x|g) < \infty$. Then there exists a function $g_1 \in C(\bar{\mathbf{R}}^n)$ such that $g_1 > 0$, $m(g_1) = 1$, $m(|x|g_1) < \infty$, and $m(xg) = m(xg_1)$.*

Proof. Let $a := m(xg)$. Set $T_a(x) := x - a$, $m^a := T_a^{-1}m$, and $g^a(x) := g(x + a)$. Since $m^a(xg^a) = 0$, the problem is reduced to the case $a = 0$. But the latter property means that $\tilde{m}(dx) := g(x)m(dx)$ is an equivalent martingale measure for the identity mapping, and the existence of another equivalent martingale with density $g_1 \in C(\bar{\mathbf{R}}^n)$ follows, e.g., from Th. 2.1 in [2]. □

Reduction to a one-stage problem. Any measure $Q \sim P$ has the form $Q = Z_N P$ with $Z_N = z_1 z_2 \ldots z_N$ where z_n is a strictly positive $\mathscr{F}_n$-measurable random variable such that $E(z_n \mid \mathscr{F}_{n-1}) = 1$. In discrete time the class of local martingales coincides with the class of generalized martingales, see [14], Ch. 7. Hence, $Q \in \mathcal{Q}$ iff $E(|\Delta S_n| z_n \mid \mathscr{F}_{n-1}) < \infty$ and $E(\Delta S_n z_n \mid \mathscr{F}_{n-1}) = 0$. If the process X is a generalized Q-supermartingale then $E(X_n z_n \mid \mathscr{F}_{n-1}) \leq X_{n-1}$. After these remarks the result follows from the application of Lemma 2 for each n (with $\mathscr{G} = \mathscr{F}_{n-1}$, $\xi = \Delta X_n$, $\eta = \Delta S_n$, and $\varphi_n = -\lambda_n^*$). □

3 Proof of Theorem 1

Let us consider the continuous time case. Without loss of generality we may assume that $P \in \mathcal{Q}$. Under P, the $d + 1$-dimensional process $W := (S, X)$ is a special semimartingale. Denoting by μ the jump measure of W and by ν the compensator of μ we can write the canonical decomposition of W in the form

$$W = W_0 + W^c + x * (\mu - \nu) + D \tag{3.1}$$

where $W^c = (S^c, X^c)$ is a continuous local martingale with the covariance process C and where D is a predictable process of locally bounded variation, see [8], II.2.38. Since S is a local martingale and X is a local supermartingale under P, the first d components of D vanish, and the last component has the form $-U$ where U is a predictable increasing process with $U_0 = 0$. Thus, (3.1) can be rewritten as

$$S = S_0 + S^c + \pi(x) * (\mu - \nu), \tag{3.2}$$

$$X = X_0 + X^c + x^{d+1} * (\mu - \nu) - U, \tag{3.3}$$

where π is the projection of $\mathbf{R}^{d+1}$ to the first d coordinates. Moreover,

$$\int \pi(x)\nu(\{t\}, dx) = 0, \quad \int x^{d+1}\nu(\{t\}, dx) = -\Delta U_t. \tag{3.4}$$

and $(|x|^2 \wedge |x|) * \nu$ belongs to the set $\mathscr{A}_{loc}^+$ of locally integrable increasing processes or, equivalently, $(|x|^2 \wedge |x|) * \nu_t < \infty$ for finite t, see [8], II.2.29.

Notice that, in the notation of [8], the triplet of predictable characteristics of the semimartingale W is given by (B, C, ν) (with truncation function $h(x) = xI_{\{|x| \leq 1\}}$) where $B = D - xI_{\{|x|>1\}} * \nu$ and ν, C, D are defined as above.

According to Proposition II.2.9 in [8] one can choose a "good" version of the characteristics of W with respect to some predictable reference process $A \in \mathscr{A}_{loc}^+$, i.e., a version such that $U = u \cdot A$, $C^{ij} = c^{ij} \cdot A$, $\nu(\omega, dt, dx) = K(\omega, t, dx)dA_t(\omega)$ where u is a predictable process, c is a predictable process with values in the set of all non-negative symmetric $(d+1) \times (d+1)$ matrices, and $K(\omega, t, dx)$ is a transition kernel from $(\Omega \times \mathbf{R}_+, \mathscr{P})$ into $(\mathbf{R}^{d+1}, \mathscr{B}^{d+1})$ with the properties II.2.11 in [8].

Now we describe the special properties which are induced by our assumptions on the behavior of X under a measure $\tilde{P} \in \mathcal{Q}$. Let $\mathscr{P}$ denote the predictable σ-field on $\Omega \times \mathbf{R}_+$. Since $\tilde{P} \sim P$, the general Girsanov theorem [8], III.3.24 in connection with [8], III.5.7 simplifies as follows: There exists a predictable $\mathbf{R}^{d+1}$-valued process β and a positive $\mathscr{P} \otimes \mathscr{B}^{d+1}$-measurable function $Y = Y(\omega, t, x)$ such that

$$H_\infty := \beta' c\beta \cdot A_\infty + (1 - \sqrt{Y})^2 * \nu_\infty +$$

$$+ \sum_{s \geq 0} (\sqrt{1 - a_s} - \sqrt{1 - \widehat{Y}_s})^2 < \infty,$$

$\{0 < a < 1\} = \{0 < \widehat{Y} < 1\}$, $\{a = 1\} = \{\widehat{Y} = 1\}$ where

$$a_s := \nu(\{s\}, \mathbf{R}^{d+1}), \quad \widehat{Y}_s := \int Y(s, x)\nu(\{s\}, dx),$$

and the triplet of predictable characteristics $(\tilde{B}, \tilde{C}, \tilde{\nu})$ under $\tilde{P}$ has the form:

$$\tilde{B} = B + c\beta \cdot A + xI_{\{|x| \leq 1\}}(Y - 1) * \nu, \tag{3.5}$$

$$\tilde{C} = C, \quad \tilde{\nu} = Y\nu.$$

The integrals in (3.5) exist in the usual sense. Being a special semimartingale with respect to $\tilde{P}$, the process W admits the canonical decomposition

$$W = W_0 + \tilde{W}^c + x * (\mu - Y\nu) + \tilde{D},$$

where $(|x|^2 \wedge |x|)Y * \nu_t < \infty$ for finite t, and $\tilde{B} = \tilde{D} - xI_{\{|x|>1\}}Y * \nu$.

Since S remains a local martingale with respect to $\tilde{P}$, the first d components of $\tilde{D}$ vanish, i.e.

$$\sum_{j=1}^{d+1} c^{ij}\beta^j \cdot A + x^i(Y-1)*\nu = 0, \quad i \le d. \tag{3.6}$$

The condition that X remains a local supermartingale with respect to $\tilde{P}$ can be written in the following way:

$$\sum_{j=1}^{d+1} c^{d+1,j}\beta^j \cdot A + x^{d+1}(Y-1)*\nu - U \quad \text{is a decreasing process.} \tag{3.7}$$

The above relations make sense because for finite t we have

$$|c\beta| \cdot A_t < \infty, \tag{3.8}$$

$$|x(Y-1)| * \nu_t < \infty. \tag{3.9}$$

In terms of "intensities" the conditions (3.6) and (3.7) take the form

$$\sum_{j=1}^{d+1} c_t^{ij}(\omega)\beta_t^j + \int x^i(Y(\omega,t,x)-1)K(\omega,t,dx) = 0, \quad i \le d, \tag{3.10}$$

$$\sum_{j=1}^{d+1} c_t^{d+1,j}(\omega)\beta_t^j + \int x^{d+1}(Y(\omega,t,x)-1)K(\omega,t,dx) \le u_t \tag{3.11}$$

$P \otimes A$-a.e.

For a point $(\omega,t) \in \{a = 0\}$ we define the set $\mathscr{L}_{\omega,t}$ of all pairs $(\beta,Y) \in \mathbf{R}^{d+1} \times C_+(\bar{\mathbf{R}}^{d+1})$ such that

$$\int (|x|^2 \wedge |x|)Y(x)K(\omega,t,dx) < \infty, \quad \int(\sqrt{Y(x)}-1)^2 K(\omega,t,dx) < \infty, \tag{3.12}$$

and

$$\sum_{j=1}^{d+1} c_t^{ij}(\omega)\beta^j + \int x^i(Y(x)-1)K(\omega,t,dx) = 0, \quad i \le d; \tag{3.13}$$

for $(\omega,t) \in \{0 < a < 1\}$ we include in the definition of $\mathscr{L}_{\omega,t}$ also the constraint

$$0 < \int Y(x)\nu(\omega,\{t\},dx) < 1 \tag{3.14}$$

while for $(\omega,t) \in \{a = 1\}$ we add the constraint

$$\int Y(x)\nu(\omega,\{t\},dx) = 1. \tag{3.15}$$

Lemma 4 *There is a set $\Gamma \in \mathscr{P}$ with $(P \otimes A)(\Gamma^c) = 0$ such that for $(\omega,t) \in \Gamma$ we have*

$$\sum_{j=1}^{d+1} c_t^{d+1,j}(\omega)\beta^j + \int x^{d+1}(Y(x)-1)K(\omega,t,dx) \le u_t(\omega) \tag{3.16}$$

for all $(\beta,Y) \in \mathscr{L}_{\omega,t}$.

Proof. Without loss of generality we can assume that

$$\mathbf{1}'c\mathbf{1} \cdot A_\infty + A_\infty \leq \text{const}$$

where $\mathbf{1}$ is the (column) vector with unit coordinates.

Let us consider in $\Omega \times \mathbf{R}_+ \times \mathbf{R}^{d+1} \times C(\bar{\mathbf{R}}^{d+1})$ the subset

$$
\Delta \quad := \quad \Big\{ (\omega, t, \beta, Y) : \ |\beta| \leq k_1,
$$
$$
\int (|x|^2 \wedge |x|) Y(x) K(\omega, t, dx) + \int (\sqrt{Y(x)} - 1)^2 K(\omega, t, dx) \leq k_2,
$$
$$
Y(x) > 0, \ (3.10) \text{ holds}, \ (3.11) \text{ fails} \Big\} \cap \Delta_1 \cap \Delta_2
$$

where

$$
\Delta_1 := \begin{cases}
\Omega \times \mathbf{R}_+ \times \mathbf{R}^{d+1} \\
\quad \times C(\bar{\mathbf{R}}^{d+1}), & (\omega, t) \notin \{0 < a < 1\}, \\
\Omega \times \mathbf{R}_+ \times \mathbf{R}^{d+1} \\
\quad \times \Big\{ Y : \ 0 < \int Y(x)\nu(\omega, \{t\}, dx) < 1 \Big\}, & (\omega, t) \in \{0 < a < 1\},
\end{cases}
$$

$$
\Delta_2 := \begin{cases}
\Omega \times \mathbf{R}_+ \times \mathbf{R}^{d+1} \times C(\bar{\mathbf{R}}^{d+1}), & (\omega, t) \notin \{a = 1\}, \\
\Omega \times \mathbf{R}_+ \times \mathbf{R}^{d+1} \\
\quad \times \Big\{ Y : \ \int Y(x)\nu(\omega, \{t\}, dx) = 1 \Big\}, & (\omega, t) \in \{a = 1\}.
\end{cases}
$$

Note that Δ is measurable with respect to the σ-algebra $\bar{\mathscr{P}} \otimes \mathscr{B}^{d+1} \otimes \mathscr{B}(C(\bar{\mathbf{R}}^{d+1}))$ where $\bar{\mathscr{P}}$ is the completion of $\mathscr{P}$ with respect to $P \otimes A$.

If the claim of the lemma is false then for some constants k_1 and k_2 the projection of Δ onto $\Omega \times \mathbf{R}_+$ has a positive measure $P \otimes A$. Applying the measurable selection theorem as in the proof of Lemma 2 we can construct a predictable process β and a positive $\mathscr{P} \otimes \mathscr{B}^{d+1}$-measurable function Y such that the relation (3.11) is violated on a set of positive $P \otimes A$-measure and $H_\infty \leq \text{const}$. Theorem 12 in [10] (see also [8], Lemma III.5.30) implies that there exists a probability measure with these parameters β and Y, and this is a contradiction. $\quad\square$

Lemma 5 *There exist a predictable process λ^* with values in $\mathbf{R}^d$ and a set $\Gamma \in \mathscr{P}$ with $(P \otimes A)(\Gamma^c) = 0$ such that for all $(\omega, t) \in \Gamma$ we have*

$$
\sum_{j=1}^{d+1} c_t^{d+1,j}(\omega)\beta^j + \int x^{d+1}(Y(x) - 1)K(\omega, t, dx) +
$$

$$
+ \sum_{i=1}^{d} \lambda^{*i} \Big(\sum_{j=1}^{d+1} c_t^{ij}(\omega)\beta^j + \int x^i(Y(x) - 1)K(\omega, t, dx) \Big) \leq u_t(\omega) \qquad (3.17)
$$

for any $(\beta, Y) \in \mathscr{L}_{\omega,t}$.

Proof. As in Lemma 1, we first show that for any (ω, t) such that the set $\mathscr{L}_{\omega,t}$ is nonempty there exists a vector λ^* depending on (ω, t) which is a Lagrange multiplier of the following optimization problem:

$$\text{maximize } f(z) \tag{3.18}$$

under the constraints

$$l(z) = 0, \tag{3.19}$$

$$z \in G, \tag{3.20}$$

where f and l denote the functions defined by the left-hand sides of (3.11) and (3.10), and where we put $G := \{z = (\beta, Y) \in \mathbf{R}^{d+1} \times C(\bar{\mathbf{R}}^{d+1}) : Y(x) > 0, (3.12) \text{ holds}\}$. Let f^* be the optimal value; $f^* \leq u_t$ by (3.11). It is sufficient to consider only the case when the components of l are linearly independent. Define the nonempty convex set

$$\Xi := \{(y_1, y_2) : y_1 < f(z), y_2 = l(z) \text{ for some } z \in G\} \subset \mathbf{R}^{d+1}.$$

By the separation theorem there exists a nonzero vector $\Lambda = (\lambda_1, \lambda_2) \in \mathbf{R}^{d+1}$ such that

$$\lambda_1 y_1 + \lambda_2 y_2 \leq \lambda_1 f^* \tag{3.21}$$

for all (y_1, y_2) from the closure of Ξ. Clearly, $\lambda_1 \geq 0$. The only problem is to show that λ_1 is not equal to zero. Indeed, if $\lambda_1 = 0$ then $\lambda_2 l(z) \leq 0$ for all $z \in G$. In particular, taking $z = (\beta, Y)$ with $\beta = 0$ and arbitrary $Y > 0$ which satisfies (3.12) we have

$$\sum_{i=1}^{d} \lambda_2^i \int x^i (Y(x) - 1) K(\omega, t, dx) \leq 0. \tag{3.22}$$

Suppose that $(\omega, t) \in \{a = 0\}$. There is a $\mathscr{B}^{d+1}$-measurable function $\varepsilon(x)$ with values in $]0, 1[$ such that the functions $Y(x) = 1 \pm \varepsilon(x)$ satisfy (3.12); hence the same holds for all functions $Y(x) = 1 + \alpha(x)\varepsilon(x)$ where $\alpha \geq -1$ and bounded. This implies the identity

$$\sum_{i=1}^{d} \lambda_2^i x^i = 0 \quad K(\omega, t, .) - \text{a.e.} \tag{3.23}$$

It follows that $\lambda_2 l(z) = 0$ for all $z \in G$, the situation which we excluded.

If (ω, t) is in $\{0 < a < 1\}$ or in $\{a = 1\}$ we need to include the additional constraint. Now the measure $K(\omega, t, .)$ is finite and by (3.4) the relation (3.22) has the form

$$\sum_{i=1}^{d} \lambda_2^i \int x^i Y(x) K(\omega, t, dx) \leq 0. \tag{3.24}$$

As in Lemma 1 we get that $\sum \lambda_2^i x^i = 0$ (a.e. with respect to $K(\omega, t, .)$) and again $\lambda_2 l(z) \leq 0$ for all $z \in G$ which is impossible. Hence, a Lagrange multiplier exists if $\mathscr{L}_{\omega,t}$ is nonempty.

Making use of Lemma 3 we conclude that the set of the Lagrange multipliers is nonempty for $P \otimes A$-almost all (ω, t). Taking a measurable (predictable) selector we get the result. $\quad \square$

Now we easily accomplish the proof of Theorem 1. For the sake of clarity we first consider the case $d = 1$. By Lemma 5 there exists a scalar predictable process λ^* such that

$$c_t^{21} \beta^1 + c_t^{22} \beta^2 + \lambda_t^* (c_t^{11} \beta^1 + c_t^{12} \beta^2) + \int [x^2(Y-1) + \lambda_t^* x^1 (Y-1)] K(t, dx) \le u_t,$$

or

$$(c_t^{21} + \lambda_t^* c_t^{11}) \beta^1 + (c_t^{22} + \lambda_t^* c_t^{12}) \beta^2 + \int [(x^2 + \lambda_t^* x^1) Y(x) - (x^2 + \lambda_t^* x^1)] K(t, dx) \le u_t$$

where β^1, β^2 are arbitrary numbers, Y is any positive function which satisfies the integrability conditions. It follows that

$$c_t^{21} + \lambda_t^* c_t^{11} = 0, \quad c_t^{22} + \lambda_t^* c_t^{12} = 0, \tag{3.25}$$

$$x^2 + \lambda_t^* x^1 \le 0, \quad -\int (x^2 + \lambda_t^* x^1) K(t, dx) \le u_t. \tag{3.26}$$

From the Galchouk–Kunita–Watanabe decomposition we have that $X^c = g \cdot S^c + N^c$ where $g = c^{12}(c^{11})^{\oplus}$ ($\oplus$ denotes the pseudoinverse) and N^c is a continuous local martingale with $\langle N^c, S^c \rangle = 0$. Considering densities of the form $\mathscr{E}(v \cdot N^c)$ with bounded predictable integrands v where $\mathscr{E}$ denotes the Doléans exponential, it follows as in the proof of the optional decomposition in [5] that $N^c = 0$. Thus, X can be written as follows:

$$X = X_0 + (-\lambda^*) \cdot S + (x^2 + \lambda^* x^1) * (\mu - \nu) - U. \tag{3.27}$$

It follows from (3.26) that $(x^2 + \lambda^* x^1) * (\mu - \nu)$ is a process of locally bounded variation which is dominated by U. Hence, (3.27) is the optional decomposition in the scalar case.

Let us now consider the general case with arbitrary $d \ge 1$. There exists a predictable process λ^* with values in $\mathbf{R}^d$ such that the relation (3.16) holds, and so we have

$$\sum_{j=1}^{d+1} \left(c_t^{d+1,j} + \sum_{i=1}^{d} \lambda_t^{*i} c_t^{ij} \right) \beta^j + \int \left[\left(x^{d+1} + \sum_{i=1}^{d} \lambda_t^{*i} x^i \right) Y(x) \right.$$

$$\left. - \left(x^{d+1} + \sum_{i=1}^{d} \lambda_t^{*i} x^i \right) \right] K(t, dx) \le u_t.$$

As above it follows that

$$c_t^{d+1,j} + \sum_{i=1}^{d} \lambda_t^{*i} c_t^{ij} = 0, \quad i \le d, \tag{3.28}$$

$$x^{d+1} + \sum_{i=1}^{d} \lambda_t^{*i} x^i \le 0, \qquad - \int \left(x^{d+1} + \sum_{i=1}^{d} \lambda_t^{*i} x^i \right) K(t, dx) \le u_t. \qquad (3.29)$$

The Galchouk–Kunita–Watanabe decomposition of X^c with respect to S^c has again the form $X^c = g \cdot S^c$ where g is a predictable process such that

$$c^{d+1,j} - \sum_{i=1}^{d} g^i c^{ij} = 0, \quad i \le d; \qquad (3.30)$$

as an integrand in this decomposition, we can take any predictable function satisfying (3.28).

The resulting representation

$$X = X_0 + (-\lambda^*) \cdot S + (x^{d+1} + \lambda^* \pi(x)) * (\mu - \nu) - U \qquad (3.31)$$

is the desired optional decomposition. $\square$

References

1. Barndorff-Nielsen O.: Processes of normal inverse gaussian type. Finance Stochast. 2, 39–66 (1998)
2. Dalang R. C., Morton A., Willinger W.: Equivalent martingale measures and no-arbitrage in stochastic securities market model. Stochast. Stochast. Rep. **29**, 185–201 (1990)
3. Dellacherie C., Meyer P.-A.: Probabilité et potentiel. Paris: Hermann 1976
4. Eberlein E., Jacod J.: On the range of options prices. Finance Stochast. **1**, 131–140 (1997)
5. El Karoui N., Quenez M.-C.: Dynamic programming and pricing of contingent claims in an incomplete market. SIAM J. Control Optimizat. **33**(1) 27–66 (1995)
6. El Karoui N., Jeanblanc-Picque M., Shreve S.: Robustness of the Black and Scholes formula. Math. Finance (To appear)
7. Föllmer H., Leukert P.: Hedging with maximal probability of success. (To appear)
8. Jacod J., Shiryayev A. N.: Limit theorems for stochastic processes. Berlin, Heidelberg, New York: Springer 1987
9. Kabanov Yu. M., Kramkov D. O.: Asymptotic arbitrage on large financial markets. Finance Stochast. **2** (1998, in press)
10. Kabanov Yu. M., Liptser R. Sh., Shiryayev A. N.: Absolute continuity and singularity of locally absolute continuous probability distributions. Part I. Math. USSR Sbornik, **39**, 2 (1981)
11. Karatzas I.: Lectures on the Mathematics of Finance. CRM Monograph Series. American Mathematical Society, Providence, 1996
12. Kramkov D. O.: Optional decomposition of supermartingales and hedging in incomplete security markets. Probab. Theory Relat. Fields **105**, 459–479 (1996)
13. Kramkov D. O., Volkov S. N.: On the methodology of option hedging. Rev. Appl. Indust. Math. (To appear)
14. Shiryayev A. N.: Probability. Berlin, Heidelberg, New York: Springer 1984