Risky Options Simplified

Martin Schweizer
Technische Universität Berlin
Fachbereich Mathematik, MA 7–4
Straße des 17. Juni 136
D – 10623 Berlin
Germany

Abstract: We study a general version of a quadratic approach to the pricing of options in an abstract financial market. The resulting price is the expectation of the option's discounted payoff under the variance-optimal signed martingale measure, and we give a very simple proof of this result. A conjecture of G. Wolczyńska essentially says that this measure coincides with the minimal signed martingale measure in a certain class of models. We show by a counterexample that this conjecture is false.

Key words: option pricing, mean-variance hedging, variance-optimal martingale measure, minimal martingale measure, risky options

1991 Mathematics Subject Classification: 90A09

JEL Classification Numbers: G10

(International Journal of Theoretical and Applied Finance 2 (1999), 59–82)

This version: 25.08.1998

0. Introduction

This paper studies a quadratic approach to option pricing in a general financial market. Starting with Föllmer/Sondermann (1986), several authors have used a quadratic criterion to determine optimal hedging and/or pricing rules; among others, we mention Bouleau/Lamberton (1989), Schweizer (1991, 1995, 1996), Schäl (1994), Bouchaud/Sornette (1994), Monat/Stricker (1995) and Aurell/Simdyankin (1998). We follow here the ideas of Bouchaud/Sornette (1994) and Aurell/Simdyankin (1998), but present a general version of their approach. The underlying idea is as follows. Given a contingent claim or risky option, one considers self-financing trading strategies and measures the risk of a strategy by the variance of its shortfall against the claim at the terminal date. Minimizing this variance yields an optimal strategy ϑ^*, say, and the option price h^* is then defined by the condition that the shortfall for the strategy ϑ^* with initial capital h^* should have expectation zero. In a model with i.i.d. returns, Aurell/Simdyankin (1998) showed by lengthy calculations that this price can be computed as the expectation of the option's discounted payoff with respect to a certain signed measure. In subsequent papers, Wolczyńska (1998) essentially conjectured that this measure coincides with the minimal signed martingale measure for the underlying asset's price process, and Hammarlid (1998) provided an argument in support of her conjecture.

The present paper makes two main contributions. We first develop the basic approach in an abstract L^2-framework encompassing both discrete and continuous time models and show how the central pricing result can be obtained by a very simple duality argument. This could actually also be deduced from slight modifications of previous results in Schweizer (1995, 1996, 1998) on approximating contingent claims in L^2 by trading gains, but we provide here direct proofs to keep the paper as self-contained as possible. We then settle both Wolczyńska's conjecture and a natural extension of it. On the positive side, we show that both conjectures are always true in one- and two-period models or for a binomial model with arbitrary time horizon. For incomplete N-period models with $N > 2$, however, the conjectures are false in general. We provide under a mild technical assumption a necessary and sufficient condition on the first two moments of the returns' distribution for the extended conjecture to be true and show that this condition fails in typical realistic models. We also give an explicit example of a quaternary 3-period model where the original Wolczyńska conjecture is false.

The paper is structured as follows. Section 1 introduces the basic terminology and gives a precise formulation of the pricing approach. Section 2 proves a general representation of the resulting price in terms of the variance-optimal signed martingale measure; this is defined as the solution of a dual optimization problem. Section 3 provides additional results on this measure and the optimal strategy if the underlying price process has the Markov property. Section 4 explains Wolczyńska's conjecture and its extension and presents positive and negative results on the latter. Finally, section 5 contains the counterexample to Wolczyńska's original conjecture.

1. Setup and problem

This section presents in an abstract setup the basic problem under consideration. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $L^2 = L^2(\Omega, \mathcal{F}, P)$ the space of all square-integrable real random variables with scalar product $(U, Z) = E[UZ]$ and norm $\|U\| = \sqrt{E[U^2]}$. For any subset $\mathcal{U}$ of L^2, we denote by $\mathcal{U}^\perp := \{Z \in L^2 \,|\, (Z, U) = 0 \text{ for all } U \in \mathcal{U}\}$ the orthogonal complement and by $\bar{\mathcal{U}}$ the closure of $\mathcal{U}$ in L^2. Fix $b \in L^2$ with $b > 0$ P-a.s., let $\mathcal{G}$ be a fixed subset of L^2 and set $\mathcal{A} := \mathbb{R}b + \mathcal{G} = \{a = hb + g \,|\, h \in \mathbb{R}, g \in \mathcal{G}\}$.

The pair $(\mathcal{G}, b)$ represents a general financial market in the following sense. An element g of $\mathcal{G}$ models the total *gains from trade* resulting from a self-financing trading strategy with initial capital 0, and b is interpreted as the final value of some *riskless bond* with initial value 1. "Riskless" as translated by $b > 0$ means that the bond is always worth some money at the end. $\mathcal{A}$ consists of those random payoffs which are *strictly attainable* in the sense that one can obtain them as final wealth of some self-financing strategy with some initial capital. We always assume that

$$\mathcal{G} \text{ is a linear subspace of } L^2; \tag{1.1}$$

this corresponds to a financial market without frictions like transaction costs, constraints or other nonlinear restrictions on strategies. Square-integrability gives us a nice Hilbert space structure and the existence of means and variances. For simplicity, we also assume that

$$b \text{ is deterministic, i.e., a non-random constant;} \tag{1.2}$$

see Schweizer (1998) for generalizations to the case of random b.

Example 1 (finite discrete time). For our prime example, we consider a frictionless market where one stock and a riskless bond are traded at a finite number of dates. We index these trading dates by $k = 0, 1, \ldots, N$ for a fixed $N \in \mathbb{N}$ and denote by $S_k > 0$ and $B_k > 0$ the stock and bond prices, respectively, at time k. More formally, let $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\ldots,N}$ be a filtration on $(\Omega, \mathcal{F})$, i.e., $\mathcal{F}_k \subseteq \mathcal{F}_{k+1} \subseteq \mathcal{F}$ are σ-algebras on Ω. Intuitively, $\mathcal{F}_k$ describes the information available at time k, and so the stochastic processes $S = (S_k)_{k=0,1,\ldots,N}$ and $B = (B_k)_{k=0,1,\ldots,N}$ must be adapted to $\mathbb{F}$; this means that S_k and B_k must be $\mathcal{F}_k$-measurable ("observable at time k") for each k. We always take $B_0 = 1$ for simplicity, denote by

$$X_k := \frac{S_k}{B_k} \qquad \text{for } k = 0, 1, \ldots, N$$

the discounted stock prices and suppose that X is square-integrable, i.e., $X_k \in L^2$ for each k. If each B_k is deterministic, then $b := B_N$ clearly satisfies (1.2). For later use, we define the *return factors* Y_k by $S_k = S_{k-1} Y_k$ for $k = 1, \ldots, N$.

We next explain how to model trading in the assets S and B. Intuitively, a trading strategy prescribes at each instant how many shares of stock and how many bonds we hold in our portfolio. Because trading should be *self-financing* in the sense that expenses for stock purchases must be paid by income from bond sales and vice versa, a strategy is completely described by its initial capital h and its stock holdings ϑ_k at each date k; see Proposition 1.1.3 of Lamberton/Lapeyre (1996). To exclude clairvoyance, ϑ_k must for each k be chosen at date $k-1$ on the basis of the information then available. Hence each ϑ_k must be $\mathcal{F}_{k-1}$-measurable, and this is equivalent to saying that the process $\vartheta = (\vartheta_k)_{k=1,\dots,N}$ must be predictable with respect to $I\!\!F$. Observe that our notation means that ϑ_k shares are actually held on $(k-1,k]$ so that our ϑ_k corresponds to the quantity φ_{k-1} in Aurell/Simdyankin (1998).

Let Θ be a linear space of $I\!\!F$-predictable processes ϑ and denote by $\Delta X_j := X_j - X_{j-1}$ the increment of X over $(j-1,j]$. For $\vartheta \in \Theta$, the *gains process* $G(\vartheta)$ is defined by

$$G_k(\vartheta) := B_k \sum_{j=1}^{k} \vartheta_j \Delta X_j = \sum_{j=1}^{k} \vartheta_j (S_j - rS_{j-1}) \frac{B_k}{B_j} = \sum_{j=1}^{k} \psi_j (Y_j - r) \frac{B_k}{B_j} \qquad \text{for } k = 0,1,\dots,N,$$

where $\psi_k := \vartheta_k S_{k-1}$ describes the *amount* held in shares on $(k-1,k]$. Like ϑ, ψ is predictable. A Θ-*strategy* is any pair $(h,\vartheta) \in I\!\!R \times \Theta$ and its *value process* is

$$V_k(h,\vartheta) := hB_k + G_k(\vartheta) = B_k \left(h + \sum_{j=1}^{k} \psi_j (Y_j - r) \frac{1}{B_j} \right) \qquad \text{for } k = 0,1,\dots,N. \tag{1.3}$$

$V(h,\vartheta)$ describes the wealth evolution of the self-financing strategy associated to (h,ϑ).

In this example, we have

$$\mathcal{G} := \left\{ V_T(0,\vartheta) \,\middle|\, \vartheta \in \Theta \right\} = V_T(0,\Theta) = G_T(\Theta) = \left\{ G_T(\vartheta) \,\middle|\, \vartheta \in \Theta \right\}; \tag{1.4}$$

this satisfies (1.1) because Θ is a linear space and it only remains to impose conditions on Θ to ensure that $\mathcal{G} \subseteq L^2$. One way to do this is to assume (1.2) and to consider

$$\Theta_{\mathrm{S}} := \left\{ \text{all predictable processes } \vartheta = (\vartheta_k)_{k=1,\dots,N} \text{ with } \vartheta_k \Delta X_k \in L^2 \text{ for } k = 1,\dots,N \right\};$$

this has been used in Schweizer (1995, 1996). Another possible choice for Θ under (1.2) is

$$\Theta_{\mathrm{AS}} := \left\{ \vartheta \in \Theta_{\mathrm{S}} \,\middle|\, \vartheta_k = g_k(S_{k-1}) \text{ with measurable functions } g_k \text{ on } I\!\!R \text{ for } k = 1,\dots,N \right\}.$$

This space has been used by Aurell/Simdyankin (1998) and Wolczyńska (1998); the corresponding strategies are “Markovian” in the sense that the choice of stock holdings may only depend on the currently observable stock prices.

Remark. We emphasize at this point that the choice of Θ in Example 1 becomes crucial later on. The duality arguments in section 2 only use that Θ is linear and thus work for both

$\Theta_{\rm S}$ and $\Theta_{\rm AS}$. But the results on Wolczyńska's conjecture and its extension differ for $\Theta = \Theta_{\rm S}$ and for $\Theta = \Theta_{\rm AS}$. The original conjecture was formulated for $\Theta = \Theta_{\rm AS}$ and then turns out to be false; see the counterexample in section 5. If one decides to take the larger space $\Theta = \Theta_{\rm S}$ and then examines the conjectures, they both are true; see the remark following Theorem 10.

Example 2 (i.i.d. returns). As a special case of Example 1, consider the situation where $B_0 = 1$, $S_0 > 0$ are fixed initial values, $B_k = r^k$ for some $r > 0$ and the return factors $Y_1, \dots, Y_N$ are i.i.d. > 0 under P and square-integrable. By independence, the process S is then also square-integrable, and so is X because B is deterministic. This model has been used by Aurell/Simdyankin (1998) and Wolczyńska (1998), partly under the additional assumption that the Y_k only take finitely many values. There is no explicit mention of a filtration in Aurell/Simdyankin (1998) or Wolczyńska (1998), but it is clear from their arguments that they use $\mathcal{F}_k = \sigma(S_0, S_1, \dots, S_k) = \sigma(Y_1, \dots, Y_k) = \sigma(X_0, X_1, \dots, X_k)$, i.e., the filtration generated by S, Y or X. Observe that (1.3) corresponds to (7) of Aurell/Simdyankin (1998).

Because B_k is deterministic, $\Delta X_k = \frac{S_k}{B_k} - \frac{S_{k-1}}{B_{k-1}} = \frac{S_{k-1}}{B_k}(Y_k - r)$ implies by the independence of $Y_1, \dots, Y_N$ that

$$E[\Delta X_k | \mathcal{F}_{k-1}] = \frac{S_{k-1}}{B_k}(E[Y_k] - r) \tag{1.5}$$

and

$$\mathrm{Var}[\Delta X_k | \mathcal{F}_{k-1}] = \frac{S_{k-1}^2}{B_k^2} \mathrm{Var}[Y_k]. \tag{1.6}$$

The so-called *mean-variance tradeoff process* $\widehat{K}$ is therefore given by

$$\widehat{K}_\ell := \sum_{j=1}^{\ell} \frac{(E[\Delta X_j | \mathcal{F}_{j-1}])^2}{\mathrm{Var}[\Delta X_j | \mathcal{F}_{j-1}]} = \sum_{j=1}^{\ell} \frac{(E[Y_j] - r)^2}{\mathrm{Var}[Y_j]} = \ell \frac{(E[Y_1] - r)^2}{\mathrm{Var}[Y_1]} \qquad \text{for } \ell = 0, 1, \dots, N,$$

since the Y_j are identically distributed. In particular, $\widehat{K}$ is *deterministic* in this example, hence also bounded uniformly in ℓ and ω.

Example 3 (continuous time). To illustrate the generality of our formulation, we briefly explain how to incorporate a continuous-time model into our framework. Let $T \in (0, \infty]$ be a fixed time horizon and $X = (X_t)_{0 \le t \le T}$ an $\mathbb{R}^d$-valued semimartingale with respect to P and a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ on $(\Omega, \mathcal{F})$. Let Θ be the space of all $\mathbb{R}^d$-valued $\mathbb{F}$-predictable X-integrable processes $\vartheta = (\vartheta_t)_{0 \le t \le T}$ such that the stochastic integral process $G(\vartheta) := \int \vartheta \, dX$ is in the space $\mathcal{S}^2$ of semimartingales. Then we could take $b \equiv 1$ and $\mathcal{G} := G_T(\Theta)$; this space has been studied by Delbaen/Monat/Schachermayer/Schweizer/Stricker (1997). For a different choice of Θ, see also Gouriéroux/Laurent/Pham (1998). This ends the example.

Let us now return to our abstract framework and consider a general *contingent claim* or *risky option*. This is a random variable $H \in L^2$; it describes a financial derivative by specifying its net payoff $H(\omega)$ at the terminal time in each state ω. A typical example in the framework of Example 1 is a European call where $H = (S_N - K)^+$ for some $K > 0$. Our goal is to determine a price for H at time 0 and to that end, we use an abstract version of an approach suggested by Bouchaud/Sornette (1994) and taken up again by Aurell/Simdyankin (1998). If we sell H at time 0 for an amount h, use h as initial capital and work with a self-financing strategy whose outcome is $g \in \mathcal{G}$, we end up with a net final wealth of

$$hb + g - H =: \text{wealth balance}$$

in the terminology of Aurell/Simdyankin (1998). We define

$$\text{profit}(h, g, H) := E\,[\text{wealth balance}] = hb + E[g - H] =: hb + \text{gain}(g, H) \tag{1.7}$$

and

$$\text{risk}(g, H) := \text{Var}\,[\text{wealth balance}] = \text{Var}\,[g - H]\,,$$

where we have used (1.2). To obtain a price for H, we now first

$$\text{minimize risk}(g, H) \text{ over all } g \in \bar{\mathcal{G}} \tag{1.8}$$

to get an optimal element $g^* \in \bar{\mathcal{G}}$. The $(\mathcal{G}, b)$*-price* $h^*_{\mathcal{G}}$ for H is then defined by requiring that

$$\text{profit}\left(h^*_{\mathcal{G}}, g^*, H\right) = 0 \tag{1.9}$$

which implies by (1.7) that

$$h^*_{\mathcal{G}} = -\frac{1}{b}\text{gain}(g^*, H) = E\left[\frac{H}{b} - \frac{g^*}{b}\right]. \tag{1.10}$$

We shall compute $h^*_{\mathcal{G}}$ quite generally in the next section.

Remarks. 1) Observe that the minimization in (1.8) runs over $g \in \bar{\mathcal{G}}$ and not only over $g \in \mathcal{G}$. This is important because it ensures the existence of a solution to (1.8) even if $\mathcal{G}$ is not closed in L^2. Moreover, it is easy to verify that

$$\inf_{g \in \mathcal{G}} \text{Var}[g - H] = \inf_{g \in \bar{\mathcal{G}}} \text{Var}[g - H]$$

so that passing from $\mathcal{G}$ to $\bar{\mathcal{G}}$ does not really change anything. We shall later provide sufficient conditions for $\mathcal{G}$ to be closed in L^2 so that we can then replace $\bar{\mathcal{G}}$ by $\mathcal{G}$ in (1.8).

2) Since all our concepts depend on the financial market $(\mathcal{G}, b)$ under consideration, it would be more accurate to index with $(\mathcal{G}, b)$ instead of $\mathcal{G}$ only. We omit b for ease of notation, but keep $\mathcal{G}$ to indicate in examples the dependence on the choice of Θ.

2. $(\mathcal{G}, b)$-pricing made easy

In this section, we give a simple general formula for the $(\mathcal{G}, b)$-price $h^*_{\mathcal{G}}$ for H by considering the dual problem of (1.8). We first recall some terminology introduced in Schweizer (1998).

Definition. We say that $(\mathcal{G}, b)$ *admits no approximate profits in* L^2 if $\bar{\mathcal{G}}$ does not contain b.

With the preceding interpretations, this notion is very intuitive: It says that one cannot approximate (in the L^2-sense) the riskless payoff b by a self-financing strategy with initial wealth 0. This is a *no-arbitrage condition* on the financial market; loosely speaking, it should be impossible to turn nothing into something without incurring costs.

Definition. A *signed* $(\mathcal{G}, b)$*-martingale measure* is a signed measure Q on $(\Omega, \mathcal{F})$ with $Q[\Omega] = 1$, $Q \ll P$ with $\frac{1}{b}\frac{dQ}{dP} \in L^2$ and

$$E_Q\left[\frac{g}{b}\right] = \left(\frac{1}{b}\frac{dQ}{dP}, g\right) = 0 \qquad \text{for all } g \in \mathcal{G}. \tag{2.1}$$

We denote by $I\!P^2_s(\mathcal{G})$ the convex set of all signed $(\mathcal{G}, b)$-martingale measures. An element $\widetilde{P}_{\mathcal{G}}$ of $I\!P^2_s(\mathcal{G})$ is called *b-variance-optimal* if

$$\left\|\frac{1}{b}\frac{d\widetilde{P}_{\mathcal{G}}}{dP}\right\| \le \left\|\frac{1}{b}\frac{dQ}{dP}\right\| \qquad \text{for all } Q \in I\!P^2_s(\mathcal{G}).$$

Our first result links the above two definitions and gives elementary facts for later use.

Proposition 1. *Assume (1.1). Then:*

1) $(\mathcal{G}, b)$ *admits no approximate profits in* L^2 *if and only if* $I\!P^2_s(\mathcal{G}) \neq \emptyset$.

2) *If* $(\mathcal{G}, b)$ *admits no approximate profits in* L^2*, then* $\bar{\mathcal{A}} = I\!Rb + \bar{\mathcal{G}}$.

3) *If* $(\mathcal{G}, b)$ *admits no approximate profits in* L^2*, then the b-variance-optimal signed* $(\mathcal{G}, b)$*-martingale measure* $\widetilde{P}_{\mathcal{G}}$ *exists, is unique and satisfies*

$$\frac{1}{b}\frac{d\widetilde{P}_{\mathcal{G}}}{dP} \in \bar{\mathcal{A}}. \tag{2.2}$$

Proof. These results are basically well known from Delbaen/Schachermayer (1996) and Schweizer (1998), but we include a proof for completeness.

1) An element Q of $I\!P_s^2(\mathcal{G})$ can be identified with a continuous linear functional Ψ on L^2 satisfying $\Psi = 0$ on $\mathcal{G}$ and $\Psi(b) = 1$ by setting $\Psi(U) = E\left[\frac{1}{b}\frac{dQ}{dP}U\right] = \left(\frac{1}{b}\frac{dQ}{dP}, U\right)$. Hence 1) is clear from the Hahn-Banach theorem.

2) Any $g \in \bar{\mathcal{G}}$ is the limit in L^2 of a sequence (g_n) in $\mathcal{G}$; hence $hb + g_n = a_n$ is a Cauchy sequence in $\mathcal{A}$ and thus converges in L^2 to a limit $a \in \bar{\mathcal{A}}$ so that $hb + g = a \in \bar{\mathcal{A}}$. This gives the inclusion "$\supseteq$" in general. For the converse, we use the assumption that $(\mathcal{G}, b)$ admits no approximate profits in L^2 to obtain from part 1) a signed $(\mathcal{G}, b)$-martingale measure Q. The random variable $Z := \frac{1}{b}\frac{dQ}{dP}$ is then in $\mathcal{G}^\perp$ and satisfies $(Z, b) = Q[\Omega] = 1$. For any $a \in \bar{\mathcal{A}}$, there is a sequence $a_n = h_n b + g_n$ in $\mathcal{A}$ converging to a in L^2. Since $h_n b + g_n \in I\!Rb + \mathcal{G}$ for all n, we conclude that

$$h_n = (h_n b + g_n, Z) = (a_n, Z)$$

converges in $I\!R$ to $(a, Z) =: h$. Therefore $g_n = a_n - h_n b$ converges in L^2 to $g := a - hb$, and since this limit is in $\bar{\mathcal{G}}$, we have $a = hb + g \in I\!Rb + \bar{\mathcal{G}}$ which proves the inclusion "$\subseteq$".

3) Existence and uniqueness of $\widetilde{P}_{\mathcal{G}}$ are clear once we observe that we have to minimize $\|Z\|$ over the closed convex set $\mathcal{Z} := \left\{ Z = \frac{1}{b}\frac{dQ}{dP} \,\middle|\, Q \in I\!P_s^2(\mathcal{G}) \right\}$ which is non-empty thanks to 1). For any $Z_0 \in \mathcal{Z}$, the projection $\widetilde{Z}$ of Z_0 in L^2 on $\bar{\mathcal{A}}$ is again in $\mathcal{Z}$; in fact, one easily verifies that $\widetilde{\Psi}(U) := \left(\widetilde{Z}, U\right)$ is 0 on $\mathcal{G}$ and has $\widetilde{\Psi}(b) = 1$. Since part 2) tells us that $\widetilde{Z} = \widetilde{h}b + \widetilde{g}$ with $\widetilde{g} \in \bar{\mathcal{G}}$, we obtain $\left(Z, \widetilde{Z}\right) = \widetilde{h} = \left(\widetilde{Z}, \widetilde{Z}\right)$ for all $Z \in \mathcal{Z}$ and therefore

$$\|Z\|^2 = \left\|\widetilde{Z}\right\|^2 + \left\|Z - \widetilde{Z}\right\|^2 \geq \left\|\widetilde{Z}\right\|^2 \qquad \text{for all } Z \in \mathcal{Z}.$$

Hence we conclude that $\frac{1}{b}\frac{d\widetilde{P}_{\mathcal{G}}}{dP} = \widetilde{Z}$ is in $\bar{\mathcal{A}}$.

q.e.d.

Example 1 (finite discrete time). Consider again the situation of Example 1. Because $\frac{g}{b} = \sum_{j=1}^{N} \vartheta_j \Delta X_j$ for any $g \in \mathcal{G}$ by (1.4) and the definition of $G(\vartheta)$, (2.1) reduces to

$$E\left[\frac{dQ}{dP}\sum_{j=1}^{N}\vartheta_j \Delta X_j\right] = 0 \qquad \text{for all } \vartheta \in \Theta, \tag{2.3}$$

and we write $I\!P_s^2(\Theta)$, $\widetilde{P}_\Theta$ as shorthand for $I\!P_s^2\big(G_T(\Theta)\big)$, $\widetilde{P}_{G_T(\Theta)}$, respectively. In this context, we also speak of *signed* Θ*-martingale measures* instead of signed $(\mathcal{G}, b)$-martingale measures. Note that although $Q[\Omega] = 1$, we may have $Q[A] \leq 0$ so that each $Q \in I\!P_s^2(\Theta)$ is only a "pseudo-probability" in the terminology of Aurell/Simdyankin (1998). If $\Theta = \Theta_{\mathrm{S}}$, condition (2.3) is equivalent to

$$0 = E\left[\frac{dQ}{dP}I_A(X_k - X_{k-1})\right] = E_Q\left[I_A \frac{1}{B_k}(S_k - rS_{k-1})\right]$$

for $k = 1, \dots, N$ and all $A \in \mathcal{F}_{k-1}$. If Q is a probability measure (i.e., ≥ 0), this means that

$$E_Q[X_k - X_{k-1} | \mathcal{F}_{k-1}] = 0 \qquad \text{for } k = 1, \dots, N$$

or that X is a Q-martingale; this explains the terminology. For $\Theta = \Theta_{\mathrm{AS}}$ and a probability measure $Q \in \mathbb{P}_s^2(\Theta_{\mathrm{AS}})$, (2.3) is analogously equivalent to

$$E_Q[S_k - rS_{k-1} | S_{k-1}] = 0 \qquad \text{for } k = 1, \dots, N;$$

compare (34) – (37) of Aurell/Simdyankin (1998). Note that this latter relation does not give a martingale property for X because $\mathcal{H}_k := \sigma(S_k)$, $k = 0, 1, \dots, N$, is not a filtration (these σ-algebras are not increasing with k). This ends the example.

Theorem 2. *Assume (1.1), (1.2) and that $(\mathcal{G}, b)$ admits no approximate profits in L^2. For any contingent claim H, the $(\mathcal{G}, b)$-price is then given by*

$$h^*_{\mathcal{G}} = \widetilde{E}_{\mathcal{G}}\left[\frac{H}{b}\right],$$

where $\widetilde{E}_{\mathcal{G}}$ denotes expectation with respect to the b-variance-optimal signed $(\mathcal{G}, b)$-martingale measure $\widetilde{P}_{\mathcal{G}}$.

Proof. Because b is deterministic by (1.2), we have

$$\begin{aligned} E\left[\left(\frac{H}{b} - h - \frac{g}{b}\right)^2\right] &= \mathrm{Var}\left[\frac{H}{b} - \frac{g}{b}\right] + \left(E\left[\frac{H}{b} - \frac{g}{b}\right] - h\right)^2 \\ &= \frac{1}{b^2}\mathrm{risk}(g, H) + \left(h - E\left[\frac{H}{b} - \frac{g}{b}\right]\right)^2 \end{aligned}$$

for any pair $(h, g) \in \mathbb{R} \times \bar{\mathcal{G}}$. Since both terms on the right-hand side are nonnegative and the first one does not depend on h, it is clear that minimizing the left-hand side over $(h, g) \in \mathbb{R} \times \bar{\mathcal{G}}$ is achieved by first solving (1.8) for g^* and then choosing

$$h^* = E\left[\frac{H}{b} - \frac{g^*}{b}\right]$$

to make the last term vanish. Hence finding h^* is equivalent to finding the constant $\widetilde{h}$ in

$$(\widetilde{h}, \widetilde{g}) := \mathop{\arg\min}_{(h,g) \in \mathbb{R} \times \bar{\mathcal{G}}} E\left[\left(\frac{H}{b} - h - \frac{g}{b}\right)^2\right] = \frac{1}{b^2} \mathop{\arg\min}_{(h,g) \in \mathbb{R} \times \bar{\mathcal{G}}} E\left[(H - hb - g)^2\right]$$

by (1.2). But of course $\widetilde{h}b + \widetilde{g}$ is simply the projection in L^2 of H on $I\!Rb + \bar{\mathcal{G}} = \bar{\mathcal{A}}$ by Proposition 1, and since $\left(H - \widetilde{h}b - \widetilde{g}, \frac{1}{b}\frac{d\widetilde{P}_{\mathcal{G}}}{dP}\right) = 0$ due to (2.2), we obtain

$$h^*_{\mathcal{G}} = h^* = \widetilde{h} = \left(\widetilde{h}b + \widetilde{g}, \frac{1}{b}\frac{d\widetilde{P}_{\mathcal{G}}}{dP}\right) = \left(H, \frac{1}{b}\frac{d\widetilde{P}_{\mathcal{G}}}{dP}\right) = \widetilde{E}_{\mathcal{G}}\left[\frac{H}{b}\right],$$

where the third equality uses that $\widetilde{P}_{\mathcal{G}} \in I\!P^2_s(\mathcal{G})$.

q.e.d.

Theorem 2 is a general version of the result (28) of Aurell/Simdyankin (1998). If we specialize our model to the framework of Example 1, we obtain an extension of the latter result in several directions. Apart from square-integrability, our model for S is completely general, and so is the contingent claim H; we can therefore deal with arbitrary path-dependent derivatives. More importantly, though, the proof of Theorem 2 is very simple and transparent; by exploiting the geometric structure of the problem, we can avoid the lengthy and model-specific computations of Aurell/Simdyankin (1998) and Aurell/Życzkowski (1996).

To conclude this section, we present some results on the closedness of $\mathcal{G}$ in L^2 in the framework of Example 1. Note that all assumptions of Corollary 4 are satisfied in Example 2, the case of i.i.d. returns.

Proposition 3. *Consider the situation of Example 1 with B_N deterministic so that (1.2) holds. Let $\Theta \subseteq \Theta_{\mathrm{S}}$ be a linear space and suppose that Θ is stable under P-a.s. convergence in the following sense: If we have a sequence $(\vartheta^n)_{n\in I\!N}$ in Θ such that $\lim_{n\to\infty} \vartheta^n_k = \psi_k$ P-a.s. for $k = 1, \dots, N$, then ψ is also in Θ. If the mean-variance tradeoff process $\widehat{K}$ is bounded (uniformly in ℓ and ω), then $G_N(\Theta)$ is closed in L^2.*

Proof. Because

$$G_N(\Theta) = \left\{ b\sum_{j=1}^{N} \vartheta_j \Delta X_j \,\middle|\, \vartheta \in \Theta \right\} = \left\{ \sum_{j=1}^{N} \vartheta_j \Delta X_j \,\middle|\, \vartheta \in \Theta \right\}$$

due to (1.2), this is basically a consequence of the proof for Theorem 2.1 in Schweizer (1995). If we go through that argument, we see that the assertion follows once we can show that ϑ^∞ constructed in that proof is again in Θ. But for $k = 1, \dots, N$, we have

$$\vartheta^\infty_k = I_{\{\mathrm{Var}[\Delta X_k|\mathcal{F}_{k-1}]>0\}} \frac{1}{\sqrt{\mathrm{Var}[\Delta X_k|\mathcal{F}_{k-1}]}} \lim_{n\to\infty}\left(\vartheta^n_k \sqrt{\mathrm{Var}[\Delta X_k|\mathcal{F}_{k-1}]}\right),$$

where each ϑ^n is in Θ and the limit is in L^2. The proof of Theorem 2.1 in Schweizer (1995) also shows that the value of ϑ_k^∞ on $\{\mathrm{Var}[\Delta X_k|\mathcal{F}_{k-1}] = 0\}$ does not influence $\sum_{j=1}^N \vartheta_j^\infty \Delta X_j$, and so we can obtain ϑ_k^∞ as a P-a.s. limit of $(\vartheta_k^{n_\ell})_{\ell \in \mathbb{N}}$ by passing to a subsequence. Stability of Θ under P-a.s. convergence thus implies that ϑ^∞ is in Θ, and this completes the proof.

q.e.d.

Corollary 4. *Consider the situation of Example 1 with B_N deterministic so that (1.2) holds. If the mean-variance tradeoff process $\widehat{K}$ is bounded (uniformly in ℓ and ω), then $G_N(\Theta_{\mathrm{S}})$ and $G_N(\Theta_{\mathrm{AS}})$ are both closed in L^2. Moreover, $\mathbb{P}_s^2(\Theta_{\mathrm{S}}) \subseteq \mathbb{P}_s^2(\Theta_{\mathrm{AS}})$ is non-empty.*

Proof. The first assertion follows immediately from Proposition 3 because Θ_{S} and Θ_{AS} are both obviously stable under P-a.s. convergence. The second is proved in Schweizer (1995) by showing that the minimal signed Θ-martingale measure $\widehat{P}$ is in $\mathbb{P}_s^2(\Theta)$; see also section 4.

q.e.d.

3. Additional results on $\widetilde{P}_{\Theta_{\mathrm{S}}}$ in the Markovian case

In this section, we consider the framework of Example 1 and provide more precise structural results on $\widetilde{P}_{\Theta_{\mathrm{S}}}$ for the case where X is a Markov process under P. Intuitively, predictions about the future evolution of X then only depend on the current value of X. More precisely, we assume that

$$E\left[F_k(X_k, X_{k+1}, \ldots, X_N)|\mathcal{F}_k\right] = E\left[F_k(X_k, X_{k+1}, \ldots, X_N)|X_k\right]$$

for any k and any measurable function F_k on $\mathbb{R}^{N-k+1}$ such that $F_k(X_k, X_{k+1}, \ldots, X_N)$ is integrable. A special case is any model with i.i.d. returns and B deterministic as in Example 2. We choose $\Theta = \Theta_{\mathrm{S}}$ because even in a Markovian framework, one should start by allowing as many strategies as possible for use.

We start by recalling from Schweizer (1995, 1996) the explicit expressions for $\widetilde{P}_{\Theta_{\mathrm{S}}}$ and for the optimal strategy ϑ^* for (1.8). These results were obtained for $\Theta = \Theta_{\mathrm{S}}$, but without assuming that X is a Markov process under P. We first define the predictable process $\beta = (\beta_k)_{k=1,\ldots,N}$ via backward induction by

$$\beta_k := \frac{E\left[\Delta X_k \prod_{j=k+1}^N (1 - \beta_j \Delta X_j) \middle| \mathcal{F}_{k-1}\right]}{E\left[\Delta X_k^2 \prod_{j=k+1}^N (1 - \beta_j \Delta X_j)^2 \middle| \mathcal{F}_{k-1}\right]}, \qquad k = 1, \ldots, N; \tag{3.1}$$

by convention, an empty product equals 1. According to Theorem 5 of Schweizer (1996), $\widetilde{P}_{\Theta_{\mathrm{S}}}$ is then given by

$$\frac{d\widetilde{P}_{\Theta_{\mathrm{S}}}}{dP} = \frac{1}{\widetilde{c}_{\mathrm{S}}} \prod_{j=1}^{N}(1 - \beta_j \Delta X_j) \qquad \text{with } \widetilde{c}_{\mathrm{S}} = E\left[\prod_{j=1}^{N}(1 - \beta_j \Delta X_j)\right], \tag{3.2}$$

provided that $I\!P^2_s(\Theta_{\mathrm{S}}) \neq \emptyset$. If in addition $\widehat{K}$ is bounded, the optimal strategy ϑ^* for (1.8) can also be given explicitly: If we set

$$\varrho_k := \frac{E\left[\frac{H}{B_N}\Delta X_k \prod\limits_{j=k+1}^{N}(1 - \beta_j \Delta X_j)\middle|\mathcal{F}_{k-1}\right]}{E\left[\Delta X_k^2 \prod\limits_{j=k+1}^{N}(1 - \beta_j \Delta X_j)^2\middle|\mathcal{F}_{k-1}\right]} \qquad \text{for } k = 1, \ldots, N, \tag{3.3}$$

then

$$\begin{aligned} \vartheta^*_k &= \varrho_k - \beta_k\left(\widetilde{E}_{\Theta_{\mathrm{S}}}\left[\frac{H}{B_N}\right] + \frac{1}{B_{k-1}}G_{k-1}(\vartheta^*)\right) \\ &= \varrho_k - \beta_k\left(\widetilde{E}_{\Theta_{\mathrm{S}}}\left[\frac{H}{B_N}\right] + \sum_{j=1}^{k-1}\vartheta^*_j \Delta X_j\right) \qquad \text{for } k = 1, \ldots, N \end{aligned} \tag{3.4}$$

by Theorem 2.4 of Schweizer (1995). Due to their recursive character, these formulae are rather hard to evaluate in general, but we shall presently see that their structure simplifies considerably in a Markovian framework.

Our first result shows that if X is Markovian under P, then β from (3.1) only depends on the current state of X. The same is true for ϱ from (3.3) if H is a function of S_N.

Lemma 5. *Consider the situation of Example 1 and assume that B is deterministic. Suppose that X is a Markov process under P. Then $\beta_k = b_k(S_{k-1})$ for some measurable functions b_k on $I\!R$ for $k = 1, \ldots, N$. If H has the form $H = f(S_N)$, then we also have $\varrho_k = g_k(S_{k-1})$ for $k = 1, \ldots, N$.*

Proof. We first prove the assertion for β by backward induction. For $k = N$, the Markov property of X yields

$$\beta_N = \frac{E[\Delta X_N|\mathcal{F}_{N-1}]}{E[\Delta X_N^2|\mathcal{F}_{N-1}]} = \frac{E[\Delta X_N|X_{N-1}]}{E[\Delta X_N^2|X_{N-1}]} = b^0_N(X_{N-1}) = b_N(S_{N-1}),$$

since B_N is deterministic. If we now have $\beta_j = b_j(S_{j-1})$ for $j = k+1, \ldots, N$, then $\Delta X_k \prod\limits_{j=k+1}^{N}(1 - \beta_j \Delta X_j)$ is a function of $X_k, X_{k+1}, \ldots, X_N$, and so (3.1) and the Markov property of X imply that $\beta_k = b_k(S_{k-1})$. This proves the assertion for β, and if $H = f(S_N)$, a similar argument yields $\varrho_k = g_k(S_{k-1})$ for all k.

q.e.d.

A first consequence is that X is again a Markov process under $\widetilde{P}_{\Theta_S}$. A precise formulation needs some care with conditional expectations since $\widetilde{P}_{\Theta_S}$ is in general only a signed measure.

Proposition 6. *Consider the situation of Example 1 and assume that B is deterministic. If X is a Markov process under P, then X is also Markovian under $\widetilde{P}_{\Theta_S}$ in the sense that for any k and any bounded measurable function F_k on $I\!R^{N-k+1}$, we have*

$$E\left[\widetilde{Z}_N F_k(X_k, X_{k+1}, \dots, X_N)\middle|\mathcal{F}_k\right] = \widetilde{Z}_k f_k(X_k) \tag{3.5}$$

for some measurable functions f_k on $I\!R$, where $\widetilde{Z}_k := E\left[\frac{d\widetilde{P}_{\Theta_S}}{dP}\middle|\mathcal{F}_k\right]$ denotes the density process of $\widetilde{P}_{\Theta_S}$ with respect to P.

Proof. By (3.2), we have for any k

$$\widetilde{Z}_N = \prod_{j=1}^{k}(1-\beta_j\Delta X_j)\prod_{j=k+1}^{N}(1-\beta_j\Delta X_j),$$

and due to Lemma 5, the second factor only depends on $X_k, X_{k+1}, \dots, X_N$. By the Markov property of X under P, the left-hand side of (3.5) thus equals $f_k^{(1)}(X_k)\prod_{j=1}^{k}(1-\beta_j\Delta X_j)$ for some measurable function $f_k^{(1)}$. Again using the Markov property of X under P, we also get

$$\widetilde{Z}_k = f_k^{(2)}(X_k)\prod_{j=1}^{k}(1-\beta_j\Delta X_j),$$

and so the assertion follows.

q.e.d.

Remark. If $\widetilde{P}_{\Theta_S}$ is *equivalent* to P (and thus in particular a probability measure), the Bayes rule yields for any bounded $\mathcal{F}_N$-measurable H that

$$\widetilde{E}_{\Theta_S}[H|\mathcal{F}_k] = \frac{1}{\widetilde{Z}_k}E\left[\widetilde{Z}_N H\middle|\mathcal{F}_k\right].$$

In that case, (3.5) can be rewritten as

$$\widetilde{E}_{\Theta_S}\left[F_k(X_k, X_{k+1}, \dots, X_N)\middle|\mathcal{F}_k\right] = f_k(X_k) \qquad \text{for each } k.$$

By choosing $F_k(X_k, X_{k+1}, \dots, X_N) := I_{\{X_{k+1} \in A\}}$ for arbitrary sets $A \in \mathcal{F}$, this implies that X is also a Markov process under $\widetilde{P}_{\Theta_{\mathrm{S}}}$.

It would be pleasant if the optimal strategy ϑ^* for a contingent claim of the form $H = f(S_N)$ were in Θ_{AS}. This is not quite the case, but we can obtain for ϑ^* a generalized Markovian structure if we add the process $G(\vartheta^*)$ as a second state variable.

Proposition 7. *Consider the situation of Example 1 and assume that B is deterministic. Suppose that X is a Markov process under P and that the mean-variance tradeoff process $\widehat{K}$ is bounded (uniformly in ℓ and ω). For a contingent claim of the form $H = f(S_N)$, the optimal strategy ϑ^* for (1.8) can then be written as*

$$\vartheta^*_k = f_k\big(S_{k-1}, G_{k-1}(\vartheta^*)\big) \qquad \text{for } k = 1, \dots, N$$

for some measurable functions f_k on $I\!R^2$.

Proof. Since

$$\vartheta^*_k = \varrho_k - \beta_k \left(\widetilde{E}_{\Theta_{\mathrm{S}}} \left[\frac{H}{B_N} \right] + \frac{1}{B_{k-1}} G_{k-1}(\vartheta^*) \right)$$

by (3.4) and since β_k and ϱ_k are functions of S_{k-1} only by Lemma 5, the assertion follows.

q.e.d.

4. On Wolczyńska's conjecture and an extension

In this section, we present some results related to a conjecture raised by Wolczyńska (1998). Throughout the section, we consider a model in finite discrete time as in Example 1 with a deterministic bond process B. Before stating the conjecture and a natural extension of it, we have to introduce some notation.

Let us first define a signed measure $\widehat{P}$ by

$$\frac{d\widehat{P}}{dP} := \prod_{j=1}^{N} \frac{1 - \alpha_j \Delta X_j}{1 - \alpha_j E[\Delta X_j | \mathcal{F}_{j-1}]} \tag{4.1}$$

with

$$\alpha_j := \frac{E[\Delta X_j | \mathcal{F}_{j-1}]}{E\left[\Delta X_j^2 \middle| \mathcal{F}_{j-1}\right]} \qquad \text{for } j = 1, \dots, N.$$

It is shown in Schweizer (1995) that $\widehat{P}$ is a signed Θ_{S}-martingale measure if the mean-variance tradeoff process $\widehat{K}$ is bounded; this implies in particular that $I\!P^2_s(\Theta_{\mathrm{S}}) \neq \emptyset$. $\widehat{P}$ is called the *minimal signed* Θ_{S}*-martingale measure*. Because $\Theta_{\mathrm{AS}} \subseteq \Theta_{\mathrm{S}}$, we have $I\!P^2_s(\Theta_{\mathrm{AS}}) \supseteq I\!P^2_s(\Theta_{\mathrm{S}})$, hence also $\widehat{P} \in I\!P^2_s(\Theta_{\mathrm{AS}})$.

Example 2 (i.i.d. returns). If we consider the special case where $Y_1, \dots, Y_N$ are i.i.d. under P, we can obtain $\widehat{P}$ much more explicitly than in (4.1). In fact, (1.5) and (1.6) yield

$$\alpha_j = \gamma \frac{1}{X_{j-1}} \qquad \text{with } \gamma := \frac{E\left[\frac{Y_j}{r} - 1\right]}{E\left[\left(\frac{Y_j}{r} - 1\right)^2\right]} \tag{4.2}$$

and therefore

$$\alpha_j \Delta X_j = \gamma \left(\frac{Y_j}{r} - 1 \right)$$

and

$$\alpha_j E[\Delta X_j | \mathcal{F}_{j-1}] = \gamma \left(\frac{m}{r} - 1 \right) \qquad \text{with } m := E[Y_1].$$

Plugging these expressions into (4.1) leads to

$$\frac{d\widehat{P}}{dP} = \prod_{j=1}^{N} \frac{1 - \gamma \left(\frac{Y_j}{r} - 1 \right)}{1 - \gamma \left(\frac{m}{r} - 1 \right)} = \frac{1}{\big(r - \gamma(m - r)\big)^N} \prod_{j=1}^{N} \big(r - \gamma(Y_j - r)\big). \tag{4.3}$$

If we take $r = 1$ and write $\mu = E[Y_j - 1]$, $\sigma^2 = \mathrm{Var}[Y_j - 1]$, then $\gamma = \frac{\mu}{\sigma^2 + \mu^2}$ and (4.3) becomes

$$\frac{d\widehat{P}}{dP} = \prod_{j=1}^{N} \frac{\sigma^2 + \mu^2 - \mu(Y_j - 1)}{\sigma^2} \tag{4.4}$$

which agrees with (3.5) of Wolczyńska (1998).

For later use, we also give an explicit result about the structure of $\widetilde{P}_{\Theta_{\mathrm{AS}}}$. Because $b = B_N$ is deterministic and $G_N(\Theta_{\mathrm{AS}})$ is closed in L^2 by Corollary 4, $\bar{\mathcal{A}} = I\!R + G_N(\Theta_{\mathrm{AS}}) = V_T(I\!R, \Theta_{\mathrm{AS}})$ and so (2.2) implies that

$$\frac{d\widetilde{P}_{\Theta_{\mathrm{AS}}}}{dP} = bV_T\big(\widetilde{h}, \widetilde{\vartheta}\big) = b^2 \widetilde{h} + bG_N\big(\widetilde{\vartheta}\big) \qquad \text{for some } \widetilde{h} \in I\!R \text{ and } \widetilde{\vartheta} \in \Theta_{\mathrm{AS}}. \tag{4.5}$$

Being in Θ_{AS}, $\widetilde{\vartheta}$ has the form $\widetilde{\vartheta}_k = \widetilde{g}_k(S_{k-1})$ for $k = 1, \dots, N$. Combining this with (1.3) and using that B is deterministic allows us to rewrite (4.5) as

$$\frac{d\widetilde{P}_{\Theta_{\mathrm{AS}}}}{dP} = \widetilde{c} + \sum_{j=1}^{N} g_j(S_{j-1})(Y_j - r) \tag{4.6}$$

for some $\widetilde{c} \in I\!R$ and some measurable functions g_j on $I\!R_+$; in fact, $\widetilde{c} = B_N^2 \widetilde{h}$ and $g_j(s) = \frac{B_N^2}{B_j} s \widetilde{g}_j(s)$. This ends the example.

With the above terminology, the conjecture of Wolczyńska (1998) is then:

(C) If $B \equiv 1$ and $Y_1, \dots, Y_N$ are i.i.d. under P and take only finitely many values, then $\widetilde{E}_{\Theta_{\mathrm{AS}}}[h(S_N)] = \widehat{E}[h(S_N)]$ for all measurable functions h on $I\!R$.

An equivalent formulation is

(C) If $B \equiv 1$ and $Y_1, \dots, Y_N$ are i.i.d. under P and take only finitely many values, then $\widetilde{P}_{\Theta_{\mathrm{AS}}} = \widehat{P}$ on the σ-algebra $\sigma(S_N)$ generated by S_N.

Somewhat more generally, one might also conjecture that

(EC) If B is deterministic and $Y_1, \dots, Y_N$ are i.i.d. under P, then $\widetilde{P}_{\Theta_{\mathrm{AS}}} = \widehat{P}$.

Of course, (EC) implies (C), but not vice versa. This section deals with the conjecture (EC) and the conjecture (C) is the subject of the next section. We start with a positive result.

Theorem 8. *Consider the situation of Example 2 so that B is deterministic and X has i.i.d. returns. If $N \in \{1,2\}$, then $\widetilde{P}_{\Theta_{\mathrm{AS}}} = \widehat{P}$ so that the conjectures (EC) and (C) are true for $N = 1$ and $N = 2$.*

Proof. 1) Because X has i.i.d. returns, we know from section 1 that the mean-variance tradeoff process $\widehat{K}$ is deterministic. By Corollary 4.2 of Schweizer (1995), this implies that $\widehat{P} = \widetilde{P}_{\Theta_{\mathrm{S}}}$ and so it is enough to show that $\widetilde{P}_{\Theta_{\mathrm{AS}}}$ and $\widetilde{P}_{\Theta_{\mathrm{S}}}$ coincide. This part of the argument holds for any $N \in I\!N$.

2) To finish the proof, we now show that $\Theta_{\mathrm{AS}} = \Theta_{\mathrm{S}}$ for $N \in \{1,2\}$; this implies of course that $\widetilde{P}_{\Theta_{\mathrm{AS}}} = \widetilde{P}_{\Theta_{\mathrm{S}}}$. Clearly, we only have to show that $\Theta_{\mathrm{S}} \subseteq \Theta_{\mathrm{AS}}$, and by the definitions, this amounts to proving that each $\vartheta \in \Theta_{\mathrm{S}}$ can be written as

$$\vartheta_k = g_k(S_{k-1}) \qquad \text{for } k = 1, \dots, N$$

with measurable functions g_k on $I\!R$. Now if ϑ is in Θ_{S}, then each ϑ_k is measurable with respect to $\mathcal{F}_{k-1} = \sigma(S_0, S_1, \dots, S_{k-1})$ and thus a function of $S_0, S_1, \dots, S_{k-1}$ in general. But for $N \in \{1,2\}$, we need only consider the cases $k = 1$ and $k = 2$, and then we have $\mathcal{F}_0 = \sigma(S_0)$ and $\mathcal{F}_1 = \sigma(S_0, S_1) = \sigma(S_1)$, since S_0 is deterministic. This shows that ϑ_1 and ϑ_2 are functions of S_0 and of S_1, respectively, and thus completes the proof.

q.e.d.

Theorem 8 generalizes the results of Aurell/Simdyankin (1998) who showed by rather laborious calculations that $\widetilde{P}_{\Theta_{\mathrm{AS}}} = \widehat{P}$ if either $N = 1$ and S_1 takes a finite number of values or $N \in \{1,2\}$ and X follows a binomial process as in the Cox/Ross/Rubinstein (1979) model.

Here we obtain the same result for an arbitrary distribution of Y_1 with finite second moment. We can also generalize the result for the binomial model.

Proposition 9. *The conjectures (EC) and (C) are true for any $N \in I\!N$ if X is given by the binary Cox/Ross/Rubinstein (1979) model.*

Proof. We first observe that due to the binary structure of that model, $I\!P^2_s(\Theta_{\rm S})$ contains just one element P^* given by the classical CRR prescription so that $\widetilde{P}_{\Theta_{\rm S}} = \widehat{P} = P^*$. Moreover, the time-homogeneous structure of the CRR model (or, put differently, the fact that we have a recombining binary tree with constant parameters) implies that P^* is already determined by the condition that

$$E^*[S_k - rS_{k-1}|S_{k-1}] = 0 \qquad \text{for } k = 1, \ldots, N.$$

But this means that $I\!P^2_s(\Theta_{\rm AS}) \supseteq I\!P^2_s(\Theta_{\rm S})$ also contains P^* as its sole element, and so we also have $\widetilde{P}_{\Theta_{\rm AS}} = P^*$, hence $\widetilde{P}_{\Theta_{\rm AS}} = \widehat{P}$.

q.e.d.

Let us now examine the case where $N > 2$. Since $\mathcal{F}_2 = \sigma(S_0, S_1, S_2)$ will in general be strictly larger that $\sigma(S_2)$, we expect in general a strict inclusion $\Theta_{\rm AS} \subset \Theta_{\rm S}$. Hence the argument used in the proof of Theorem 8 no longer works and it is not too surprising that the situation for (EC) also changes. In our next results, we exclude the case where X happens to be a martingale under P because in that case we trivially have $\widetilde{P}_{\Theta_{\rm AS}} = \widetilde{P}_{\Theta_{\rm S}} = \widehat{P} = P$.

Theorem 10. *Consider the situation of Example 2 so that B is deterministic and X has i.i.d. returns. Suppose also that X is not a martingale under P and that the support of the distribution (under P) of Y_1 contains an interval. If $N > 2$, then $\widetilde{P}_{\Theta_{\rm AS}} = \widehat{P}$ if and only if*

$$E\left[Y_1^2\right] = rE[Y_1]. \tag{4.7}$$

In particular, the conjecture (EC) is false in that case unless (4.7) happens to hold.

Proof. 1) We first show that (4.7) is necessary because this argument also illuminates where the condition comes from. Suppose that $\widetilde{P}_{\Theta_{\rm AS}} = \widehat{P}$. Then the explicit representation (4.3) and the structural result (4.6) imply that

$$\text{const.} \prod_{j=1}^{N} \big(r - \gamma(Y_j - r)\big) = \widetilde{c} + \sum_{j=1}^{N} g_j(S_{j-1})(Y_j - r) \qquad P\text{-a.s.} \tag{4.8}$$

Because $Y_1, \ldots, Y_N$ are i.i.d., we can view (4.8) as an identity between two polynomials in the variables $Y_1, \ldots, Y_N$. Because the support of the distribution of Y_1 contains an interval, we can conclude that all coefficients of these polynomials must coincide, and so comparing the coefficients of Y_N yields the new identity

$$\text{const.} \prod_{j=1}^{N-1} \big(r - \gamma(Y_j - r)\big) = g_N(S_{N-1}) \qquad P\text{-a.s.} \tag{4.9}$$

But the right-hand side of (4.9) depends on $Y_1, \dots, Y_{N-1}$ only via the product $S_{N-1} = S_0 \prod_{j=1}^{N-1} Y_j$ and so the same must be true for the left-hand side. In particular, all linear terms in Y_j must vanish and multiplying out shows that this implies that $\gamma r(1+\gamma) = 0$. Since X is not a P-martingale, $\gamma \neq 0$. Hence we must have $\gamma = -1$, i.e.,

$$E\left[\left(\frac{Y_1}{r} - 1\right)^2\right] = -E\left[\frac{Y_1}{r} - 1\right]$$

by (4.2), and this is equivalent to (4.7).

2) Conversely, suppose now that (4.7) holds. Then $\gamma = -1$ and so (4.3) simplifies to

$$\frac{d\widehat{P}}{dP} = \text{const.} \prod_{j=1}^{N} Y_j = \text{const.}\, S_N = \text{const.}\, X_N,$$

because B is deterministic. Choosing $\widetilde{\vartheta}_k \equiv \frac{\text{const.}}{B_N}$ and $\widetilde{h} := \frac{\text{const.}\, X_0}{B_N}$ therefore yields

$$\frac{d\widehat{P}}{dP} = \text{const.}\, X_0 + \sum_{j=1}^{N} \text{const.}\, \Delta X_j = B_N \widetilde{h} + G_N\big(\widetilde{\vartheta}\,\big)$$

with $\widetilde{\vartheta}$ obviously in Θ_{AS}, and so part 3) of Proposition 1 implies that $\widehat{P} = \widetilde{P}_{\Theta_{\text{AS}}}$. This completes the proof.

q.e.d.

Remarks. 1) In realistic models satisfying the assumptions of Theorem 10, condition (4.7) will not be satisfied. In fact, (4.7) implies that

$$rE[Y_1] = E\left[Y_1^2\right] \geq (E[Y_1])^2$$

by Jensen's inequality so that $E[Y_1](r - E[Y_1]) \geq 0$. Since we want $Y_1 \geq 0$ for nonnegative stock prices, we conclude that (4.7) can only hold if we have

$$0 \leq E[Y_1] \leq r,$$

and this means that the discounted stock price X follows a supermartingale under the original measure P. Under the assumptions of Theorem 10, (EC) therefore fails in the realistic case where the growth rate $E[Y_1]$ of the stock exceeds the riskless interest rate r.

2) As a by-product, the proof of Theorem 8 shows that the failure of the conjecture (EC) is due to the restrictive choice $\Theta = \Theta_{\mathrm{AS}}$. If one allows strategies in Θ_{S} instead of only the "Markovian" ones from Θ_{AS}, (EC) takes the form $\widetilde{P}_{\Theta_{\mathrm{S}}} = \widehat{P}$, and we know from Corollary 4.2 of Schweizer (1995) that this is true for any $N \in I\!N$ in the case of i.i.d. returns.

Theorem 10 makes it clear that for $N > 2$, we must expect the conjecture (EC) to be false in general. The assumption that the support of the distribution of Y_1 contains an interval is not very restrictive, but excludes of course all models where S takes only finitely many values. In the rest of this section, we therefore examine this case more carefully.

So consider an N-period model with i.i.d. returns $Y_1, \ldots, Y_N$ and suppose that each Y_i can take M distinct values $y_1, \ldots, y_M$ with positive probability. As in (4.8), the conjecture (EC) can be written as

$$L(Y_1, \ldots, Y_N) := \text{const.} \prod_{j=1}^{N} \big(r - \gamma(Y_j - r)\big) = \widetilde{c} + \sum_{j=1}^{N} g_j(S_{j-1})(Y_j - r) =: R(Y_1, \ldots, Y_N). \tag{4.10}$$

This must hold P-a.s., hence for every possible realization of $Y_1, \ldots, Y_N$, and so we can read (4.10) as an identity between the functions L and R defined on $\{y_1, \ldots, y_M\}^N$. More precisely, we are given L and have to find functions $g_1, \ldots, g_N$ such that (4.10) holds. Each g_j is completely described by the set of its possible values $g_j(S_{j-1}) = g_j\left(S_0 \prod_{\ell=1}^{j-1} Y_\ell\right)$. Since each Y_ℓ takes values in $\{y_1, \ldots, y_M\}$ and the argument S_{j-1} of g_j is symmetric in $Y_1, \ldots, Y_{j-1}$, the number of possible values of $g_j(S_{j-1})$ is at most the number $K(j-1, M) = \binom{j-1+M-1}{j-1}$ of ordered $(j-1)$-tuples one can form from M elements. Hence each g_j gives us at most $K(j-1, M)$ free variables we can choose, and so the function R is determined by at most

$$1 + \sum_{j=1}^{N} K(j-1, M) = 1 + K(N-1, M+1)$$

parameters. In fact, 1 stands for the constant $\widetilde{c}$ and the second equality follows from an easy combinatorial argument.

The function L is of course also symmetric in its N arguments and so its range can contain up to $K(N, M)$ elements. Heuristically, we thus have from (4.10) about $K(N, M)$

equations for $1 + K(N-1, M+1)$ variables, and so we expect that (4.10), hence (EC), will typically fail as soon as

$$1 + K(N-1, M+1) < K(N, M).$$

If we ignore the summand 1, then

$$\frac{K(N,M)}{K(N-1,M+1)} = \frac{\binom{N+M-1}{N}}{\binom{N+M-1}{N-1}} = \frac{M}{N}$$

shows that (EC) is likely to fail as soon as $M > N$, i.e., if we have few time steps and many possible outcomes at each step.

This heuristic argument has several points in its favour. It fits together with Theorem 10 where we formally have $M = \infty$, and it may also explain why Aurell/Simdyankin (1998) did not obtain a contradiction to (C) with their numerical experiments (they took $M = 3$, $N = 10$ and $M = 4$, $N = 5$, respectively). But most importantly, it tells us where to look for a counterexample: since

$$K(3,4) = 20 > 16 = 1 + K(2,5),$$

we should study a quaternary 3-period model.

Example 4. Consider a model with i.i.d. returns and B deterministic as in Example 2. More specifically, we take $S_0 = B_0 = 1$, $r = 1$ (hence $B \equiv 1$ and $X \equiv S$), $N = 3$ and assume that Y_1, Y_2, Y_3 are i.i.d. under P with values in $\left\{\frac{1}{2}, 1, 2, 4\right\}$ (so that $M = 4$). To avoid degeneracy, each of these values should be taken with positive probability, but the actual probabilities are for the moment irrelevant.

To describe a strategy $\vartheta \in \Theta_{\mathrm{AS}}$ in this model, we have to specify the possible values of $g_j(S_{j-1}) = \vartheta_j S_{j-1}$ for $j = 1, 2, 3$. To that end, we index according to the possible (date, price) pairs $\big(j, S_j(\omega)\big)$ for $j = 0, 1, 2$. Since S_0 can only take the value 1, we write $\xi_0 = g_1(1)$. S_1 has 4 possible values $\frac{1}{2}, 1, 2, 4$ and we write $\xi_{11} = g_2(\frac{1}{2}), \dots, \xi_{14} = g_2(4)$. Finally, S_2 can take the 7 values $\frac{1}{4}, \frac{1}{2}, 1, 2, 4, 8, 16$ and we write $\xi_{21} = g_3(\frac{1}{4}), \dots, \xi_{27} = g_3(16)$. Thus ξ_{jk} is the amount in stock chosen at time j if S_j is in the k-th of its possible states at that time.

Theorem 11. *Consider the framework of Example 4 and assume that $E[Y_1 - 1] \neq 0$ and $E\left[(Y_1 - 1)^2\right] \neq -E[Y_1 - 1]$. Then (EC) is false.*

Proof. Observe that our assumptions mean that $\gamma \neq 0$ (or equivalently that X is not a martingale under P) and that $\gamma \neq -1$. We now assume that (4.10) holds and show that this leads to a contradiction. First of all, $L(1,1,1) = \text{const.}$ and $R(1,1,1) = \widetilde{c}$ yields

$$\text{const.} = \widetilde{c}.$$

$L(1,4,1) = \widetilde{c}(1-3\gamma)$ and $R(1,4,1) = \widetilde{c} + 3g_2(1)$ yields

$$g_2(1) = -\widetilde{c}\gamma,$$

$L(2,1,1) = \widetilde{c}(1-\gamma)$ and $R(2,1,1) = \widetilde{c} + g_1(1)$ yields

$$g_1(1) = -\widetilde{c}\gamma$$

and $L(2,2,1) = \widetilde{c}(1-\gamma)^2$, $R(2,2,1) = \widetilde{c} + g_1(1) + g_2(2)$ then gives

$$g_2(2) = -\widetilde{c}\gamma(1-\gamma).$$

From $L(1,4,2) = \widetilde{c}(1-\gamma)(1-3\gamma)$ and $R(1,4,2) = \widetilde{c} + 3g_2(1) + g_3(4)$, we then obtain

$$g_3(4) = -\widetilde{c}\gamma(1-3\gamma)$$

and therefore

$$R(2,2,2) = \widetilde{c} + g_1(1) + g_2(2) + g_3(4) = \widetilde{c}\left((1-\gamma)^2 - \gamma(1-3\gamma)\right) = \widetilde{c}\left(1 - 3\gamma + 4\gamma^2\right).$$

But

$$L(2,2,2) = \widetilde{c}(1-\gamma)^3 = \widetilde{c}\left(1 - 3\gamma + 4\gamma^2 - \gamma^2(1+\gamma)\right),$$

and so $R(2,2,2) = L(2,2,2)$ implies that $\gamma^2(1+\gamma) = 0$, hence $\gamma \in \{-1, 0\}$. This contradicts our assumption and so we conclude that (EC) must be false.

q.e.d.

5. A counterexample to Wolczyńska's conjecture

Theorem 11 shows that the extension (EC) of Wolczyńska's conjecture (C) is not true in general; even in the case of i.i.d. returns, we may have $\widetilde{P}_{\Theta_{\mathrm{AS}}} \neq \widehat{P}$. However, the restrictions of these two measures to $\sigma(S_N) \subseteq \mathcal{F}$ could still coincide so that (C) could still be true. A very recent paper of Hammarlid (1998) claims indeed that this is the case. Unfortunately, Hammarlid's arguments are not always completely clear and they also contain an error. In fact:

Theorem 12. *Consider the framework of Example 4 with $P[Y_1 = y] = \frac{1}{4}$ for $y \in \left\{\frac{1}{2}, 1, 2, 4\right\}$. Then Wolczyńska's conjecture (C) is false.*

Proof. The idea of the proof is very simple: we just compute the distribution of S_3 under $\widetilde{P}_{\Theta_{\mathrm{AS}}}$ by solving (1.8) – (1.10) for the claims $H_y = I_{\{S_3 = y\}}$ for y from the set of the 10 possible values of S_3. We then compare the result to the distribution of S_3 under $\widehat{P}$.

So fix $y \in \left\{\frac{1}{8}, \frac{1}{4}, \dots, 32, 64\right\}$ and consider the claim $H_y = I_{\{S_3=y\}}$. To determine $\widetilde{P}_{\Theta_{\mathrm{AS}}}[S_3 = y] = \widetilde{E}_{\Theta_{\mathrm{AS}}}[H_y]$, we have to solve the minimization problem

$$\text{(5.1)} \qquad \text{minimize } E\left[\big(H_y - h - G_3(\vartheta)\big)^2\right] \text{ over all } (h, \vartheta) \in I\!R \times \Theta_{\mathrm{AS}}.$$

Since $b = B_3 = 1$, the proof of Theorem 2 tells us that the optimal $h^* \in I\!R$ coincides with $\widetilde{E}_{\Theta_{\mathrm{AS}}}[H_y]$. To rewrite the objective function in (5.1), we use the parametrization of ϑ in terms of the ξ_{ij} introduced in Example 4. If we write $\mathcal{S}_i = \{s_{ij} \mid j = 1, \dots, n_i\}$ for the set of the n_i possible values of $S_i(\omega)$, then we have

$$\vartheta_i \Delta X_i = g_i(S_{i-1})(Y_i - 1) = \sum_{j=1}^{n_{i-1}} \xi_{i-1,j} I_{\{S_{i-1}=s_{i-1,j}\}}(Y_i - 1)$$

and therefore with $n_0 = 1$, $n_1 = 4$, $n_2 = 7$

$$H_y - h - G_3(\vartheta) = H_y - h - \xi_0(Y_1 - 1) - \sum_{j=1}^{4} \xi_{1,j} I_{\{S_1=s_{1,j}\}}(Y_2 - 1) - \sum_{j=1}^{7} \xi_{2,j} I_{\{S_2=s_{2,j}\}}(Y_3 - 1).$$

If we set $x := (h, \xi_0, \xi_{11}, \dots, \xi_{14}, \xi_{21}, \dots, \xi_{27})^{\mathrm{tr}} \in I\!R^{13}$, then $E\left[\big(H_y - h - G_3(\vartheta)\big)^2\right]$ can be viewed as a quadratic function $f(x)$ of x and so finding its minimum is achieved by setting its gradient with respect to x equal to 0 and solving for x. This yields the following system of 13 equations:

$$\text{(5.2)} \qquad \begin{aligned} &E[H_y - h - G_3(\vartheta)] = 0, \\ &E\left[\big(H_y - h - G_3(\vartheta)\big)(Y_1 - 1)\right] = 0, \\ &E\left[\big(H_y - h - G_3(\vartheta)\big) I_{\{S_1=s_{1,j}\}}(Y_2 - 1)\right] = 0 \qquad \text{for } j = 1, \dots, 4, \\ &E\left[\big(H_y - h - G_3(\vartheta)\big) I_{\{S_2=s_{2,j}\}}(Y_3 - 1)\right] = 0 \qquad \text{for } j = 1, \dots, 7 \end{aligned}$$

by differentiating f with respect to $h, \xi_0, \xi_{1,j}, \xi_{2,j}$, respectively. By setting

$$\begin{aligned} z_y := \Big(&E[H_y], E[H_y(Y_1 - 1)], E\left[H_y I_{\{S_1=s_{11}\}}(Y_2 - 1)\right], \dots, E\left[H_y I_{\{S_1=s_{14}\}}(Y_2 - 1)\right], \\ &E\left[H_y I_{\{S_2=s_{21}\}}(Y_3 - 1)\right], \dots, E\left[H_y I_{\{S_2=s_{27}\}}(Y_3 - 1)\right]\Big)^{\mathrm{tr}} \in I\!R^{13}, \end{aligned}$$

we can rewrite (5.2) as a linear equation $Ax = z_y$ with a 13×13-matrix A. To compute A and z_y for given y, we use the fact that due to our choice of the P-distribution of Y_1, all trajectories ω have the same probability $\left(\frac{1}{4}\right)^3 = \frac{1}{64}$. Hence computing probabilities and expectations essentially amounts to counting trajectories with desired properties, and multiplying everything by 256 to obtain integers yields

$$256A = \begin{pmatrix}
256 & 224 & 56 & 56 & 56 & 56 & 14 & 28 & 42 & 56 & 42 & 28 & 14 \\
224 & 656 & -28 & 0 & 56 & 168 & -7 & -7 & 7 & 49 & 56 & 56 & 42 \\
56 & -28 & 164 & 0 & 0 & 0 & -7 & 0 & 14 & 42 & 0 & 0 & 0 \\
56 & 0 & 0 & 164 & 0 & 0 & 0 & -7 & 0 & 14 & 42 & 0 & 0 \\
56 & 56 & 0 & 0 & 164 & 0 & 0 & 0 & -7 & 0 & 14 & 42 & 0 \\
56 & 168 & 0 & 0 & 0 & 164 & 0 & 0 & 0 & -7 & 0 & 14 & 42 \\
14 & -7 & -7 & 0 & 0 & 0 & 41 & 0 & 0 & 0 & 0 & 0 & 0 \\
28 & -7 & 0 & -7 & 0 & 0 & 0 & 82 & 0 & 0 & 0 & 0 & 0 \\
42 & 7 & 14 & 0 & -7 & 0 & 0 & 0 & 123 & 0 & 0 & 0 & 0 \\
56 & 49 & 42 & 14 & 0 & -7 & 0 & 0 & 0 & 164 & 0 & 0 & 0 \\
42 & 56 & 0 & 42 & 14 & 0 & 0 & 0 & 0 & 0 & 123 & 0 & 0 \\
28 & 56 & 0 & 0 & 42 & 14 & 0 & 0 & 0 & 0 & 0 & 82 & 0 \\
14 & 42 & 0 & 0 & 0 & 42 & 0 & 0 & 0 & 0 & 0 & 0 & 41
\end{pmatrix}.$$

To illustrate how these figures are obtained, let us explain how to get $a_{93} = 14$. Since a_{93} is the coefficient of ξ_{11} in $256E\left[\big(h + G_3(\vartheta)\big) I_{\{S_2 = s_{23} = 1\}}(Y_3 - 1)\right]$, the above representation of $h + G_3(\vartheta)$ yields

$$\begin{aligned}
a_{93} &= 256E\left[I_{\left\{S_1 = s_{11} = \frac{1}{2}\right\}}(Y_2 - 1) I_{\{S_2 = s_{23} = 1\}}(Y_3 - 1)\right] \\
&= 256\frac{1}{64}\sum_{\omega\in\Omega} I_{\left\{S_1(\omega) = \frac{1}{2}, S_2(\omega) = 1\right\}}\big(Y_2(\omega) - 1\big)\big(Y_3(\omega) - 1\big).
\end{aligned}$$

The condition $S_1(\omega) = \frac{1}{2}, S_2(\omega) = 1$ forces $Y_1(\omega) = \frac{1}{2}, Y_2(\omega) = 2$ and leaves $Y_3(\omega)$ unrestricted. Since the sum over all possible values of $Y_3 - 1$ is $\frac{7}{2}$, we get $a_{93} = 4(2-1)\frac{7}{2} = 14$. All other entries of A are obtained in a similar way, and z_y is computed in the same manner.

For each $y \in \mathcal{S}_3$, we then obtain $x_y^* = A^{-1} z_y$ as the solution of (5.1), and $\widetilde{E}_{\Theta_{\mathrm{AS}}}[H_y]$ is the first coordinate of x_y^*. If we compute z_y for all 10 elements of $\mathcal{S}_3$ and stack the resulting column vectors one beside the other, we obtain a 13×10-matrix Z. The first row of the 13×10-matrix $A^{-1}Z$ then consists of the 10 numbers

$$\widetilde{E}_{\Theta_{\mathrm{AS}}}[H_y] = \widetilde{P}_{\Theta_{\mathrm{AS}}}[S_3 = y] \qquad \text{for } y \in \mathcal{S}_3. \tag{5.3}$$

Computation yields

$$256Z = \begin{pmatrix}
4 & 12 & 24 & 40 & 48 & 48 & 40 & 24 & 12 & 4 \\
12 & 28 & 44 & 58 & 48 & 30 & 12 & -2 & -4 & -2 \\
0 & 0 & 0 & 12 & 16 & 16 & 14 & 2 & -2 & -2 \\
0 & 0 & 12 & 16 & 16 & 14 & 2 & -2 & -2 & 0 \\
0 & 12 & 16 & 16 & 14 & 2 & -2 & -2 & 0 & 0 \\
12 & 16 & 16 & 14 & 2 & -2 & -2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 12 & 4 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 & 24 & 8 & 0 & -4 & 0 \\
0 & 0 & 0 & 0 & 36 & 12 & 0 & -6 & 0 & 0 \\
0 & 0 & 0 & 48 & 16 & 0 & -8 & 0 & 0 & 0 \\
0 & 0 & 36 & 12 & 0 & -6 & 0 & 0 & 0 & 0 \\
0 & 24 & 8 & 0 & -4 & 0 & 0 & 0 & 0 & 0 \\
12 & 4 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},$$

and this allows us to obtain the distribution of S_3 under $\widetilde{P}_{\Theta_{\mathrm{AS}}}$ explicitly via (5.3).

For comparison, we next compute the values of $\widehat{P}[S_3 = y]$ for $y \in \mathcal{S}_3$. Since $E[Y_1 - 1] = \frac{7}{8}$ and $\mathrm{Var}[Y_1 - 1] = \frac{115}{64}$, (4.4) yields

$$\frac{d\widehat{P}}{dP} = \prod_{j=1}^{3} \frac{164 - 56(Y_j - 1)}{115}.$$

In particular, Y_1, Y_2, Y_3 are again i.i.d. under $\widehat{P}$ with

$$\widehat{P}\left[Y_1 = \frac{1}{2}\right] = \frac{48}{115}, \quad \widehat{P}[Y_1 = 1] = \frac{41}{115}, \quad \widehat{P}[Y_1 = 2] = \frac{27}{115}, \quad \widehat{P}[Y_1 = 4] = -\frac{1}{115}.$$

This allows us to compute $\widehat{P}[S_3 = y]$ explicitly and leads to the following table:

y	$\widehat{P}[S_3 = y]$	$\widetilde{P}_{\Theta_{\mathrm{AS}}}[S_3 = y]$
$\frac{1}{8}$	0.0727160	0.0690573
$\frac{1}{4}$	0.186335	0.178837
$\frac{1}{2}$	0.281869	0.274829
1	0.250399	0.254933
2	0.150788	0.159894
4	0.0505288	0.0572169
8	0.00866935	0.00857501
16	−0.00135711	−0.00269784
32	0.0000532588	−0.000670921
64	−0.000000657516	0.0000263986

Since the last two columns do not agree, we have $\widetilde{P}_{\Theta_{\mathrm{AS}}} \neq \widehat{P}$ on $\sigma(S_3)$ and this shows that Wolczyńska's conjecture (C) is false in this example.

q.e.d.

Remark. The above table illustrates a well-known drawback of our pricing approach. Both $\widehat{P}$ and $\widetilde{P}_{\Theta_{\mathrm{AS}}}$ are genuinely signed measures; although the column entries sum to 1, some are negative and lead to negative prices for some nonnegative payoffs. An alternative approach with a risk loading that may mitigate this problem is developped in Schweizer (1998).

6. Conclusion

This paper extends the option pricing approach of Bouchaud/Sornette (1994) and Aurell/Simdyankin (1998) to a general L^2-context. We prove in a very simple way that the resulting option price is the expectation of the option's discounted payoff under the variance-optimal signed martingale measure. We also show by a counterexample that Wolczyńska's conjecture is false in general.

References

E. Aurell and S. I. Simdyankin (1998), "Pricing Risky Options Simply", *International Journal of Theoretical and Applied Finance 1*, 1–23

E. Aurell and K. Życzkowski (1996), "Option Pricing & Partial Hedging: Theory of Polish Options", *preprint, Royal Institute of Technology, Stockholm*

J.-P. Bouchaud and D. Sornette (1994), "The Black-Scholes Option Pricing Problem in Mathematical Finance: Generalization and Extensions for a Large Class of Stochastic Processes", *Journal de Physique (France) I, 4*, 863–881

N. Bouleau and D. Lamberton (1989), "Residual Risks and Hedging Strategies in Markovian Markets", *Stochastic Processes and their Applications 33*, 131–150

J. C. Cox, S. A. Ross and M. Rubinstein (1979), "Option Pricing: A Simplified Approach", *Journal of Financial Economics 7*, 229–263

F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer and C. Stricker (1997), "Weighted Norm Inequalities and Hedging in Incomplete Markets", *Finance and Stochastics 1*, 181–227

F. Delbaen and W. Schachermayer (1996), "The Variance-Optimal Martingale Measure for Continuous Processes", *BERNOULLI 2*, 81–105; Amendments and Corrections, *BERNOULLI 2*, 379–380

H. Föllmer and D. Sondermann (1986), "Hedging of Non-Redundant Contingent Claims", *in: W. Hildenbrand and A. Mas-Colell (eds.), Contributions to Mathematical Economics, North-Holland*, 205–223

C. Gouriéroux, J. P. Laurent and H. Pham (1998), "Mean-Variance Hedging and Numéraire", *Mathematical Finance 8*, 179–200

O. Hammarlid (1998), "On Minimizing Risk in Incomplete Markets Option Pricing Models", *International Journal of Theoretical and Applied Finance 1*, 227–233

D. Lamberton and B. Lapeyre (1996), "Introduction to Stochastic Calculus Applied to Finance", *Chapman & Hall*

P. Monat and C. Stricker (1995), "Föllmer-Schweizer Decomposition and Mean-Variance Hedging of General Claims", *Annals of Probability 23*, 605–628

M. Schäl (1994), "On Quadratic Cost Criteria for Option Hedging", *Mathematics of Operations Research 19*, 121–131

M. Schweizer (1991), "Option Hedging for Semimartingales", *Stochastic Processes and their Applications 37*, 339–363

M. Schweizer (1995), "Variance-Optimal Hedging in Discrete Time", *Mathematics of Operations Research 20*, 1–32

M. Schweizer (1996), "Approximation Pricing and the Variance-Optimal Martingale Measure", *Annals of Probability 24*, 206–236

M. Schweizer (1998), "From Actuarial to Financial Valuation Principles", *preprint, Technical University of Berlin*

G. Wolczyńska (1998), "An Explicit Formula for Option Pricing in Discrete Incomplete Markets", *International Journal of Theoretical and Applied Finance 1*, 283–288