ORIGINAL ARTICLE

Swing options in commodity markets: a multidimensional Lévy diffusion model

**Marcus Eriksson · Jukka Lempa ·
Trygve Kastberg Nilssen**

Abstract We study valuation of swing options on commodity markets when the commodity prices are driven by multiple factors. The factors are modeled as diffusion processes driven by a multidimensional Lévy process. We set up a valuation model in terms of a dynamic programming problem where the option can be exercised continuously in time. Here, the number of swing rights is given by a total volume constraint. We analyze some general properties of the model and study the solution by analyzing the associated HJB-equation. Furthermore, we discuss the issues caused by the multidimensionality of the commodity price model. The results are illustrated numerically with three explicit examples.

1 Introduction

The purpose of this paper is to propose and analyze a model for valuation of a swing option, see, e.g. Burger and Graebler (2007), written on multiple commodities when the

M. Eriksson (✉)
Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, 0316 Oslo, Norway
e-mail: mkerikss@math.uio.no

J. Lempa
Institute for Economics, Oslo and Akershus University College, P.O. Box 4, St. Olavs plass,
0130 Oslo, Norway
e-mail: jukka.lempa@hioa.no

T. K. Nilssen
Department of Economics and Business Administration, University of Agder, Serviceboks 422,
4604 Kristiansand, Norway
e-mail: trygve.k.nilssen@uia.no

commodity spot prices are driven by multiple, potentially non-Gaussian factors. More precisely, the model is formulated as a dynamic programming problem in continuous time. The holder of the option is contracted an amount of a given commodity that can be purchased for a fixed price during the lifetime of the contract. The purchases can be done (that is, the option can be exercised) continuously in time such that contracted rate constraints are fulfilled. This form of contract originates from electricity markets, where they are called flexible load contracts, see, e.g. Bjorgan et al. (2000), Lund and Ollmar (2003). However, this model setting can also fit a traditional swing option with a high number of swing rights and possible exercise times. For example, we can think of a situation in an electricity market where contract is written for a year and holder can exercise on the hour-ahead market. This results into over 8,000 possible exercise times, which makes, in particular, Monte–Carlo methods virtually intractable.

During recent years, there has been a lot of activity on analysis of swing options. Being essentially a multi-strike American or Bermudan option, a natural way to approach swing options is via an optimal multiple stopping problem. In the recent papers (Carmona and Touzi 2008; Bender 2011), the theory of optimal multiple stopping is developed in continuous time using sophisticated martingale theory. To compute option prices numerically, they develop the appropriate Monte–Carlo methodology. Other methodology for swing option pricing includes forests of trees (Jaillet et al. 2004; Hambly et al. 2009; Wahab et al. 2010; Edoli et al. 2012) or stochastic meshes (Marshall and Mark Reesor 2011), multi-stage stochastic optimization (Haarbrücker and Kuhn 2009), (quasi-)variational inequalities (Dahlgren 2005; Kiesel et al. 2010) and PDE approaches (Kjaer 2008; Benth et al. 2012; Lund and Ollmar 2003). Fundamentally, all of these methods are based on the dynamic programming principle.

As the main contribution of this paper, we develop a valuation model for multicommodity swing options inspired by Benth et al. (2012). In Benth et al. (2012), the valuation problem was studied in the case of a single commodity driven by a one-factor Gaussian price process. In this paper, we generalize the results of Benth et al. (2012) to cover multiple contracted commodities with prices driven by multiple factors. From an applications point of view, this is an important generalization, since there is a substantial body of literature supporting the usage of multi-factor models for commodity prices. Moreover, we allow also for non-Gaussian factors, which are favored, for example, in electricity price models, see, e.g. Benth et al. (2007), Hambly et al. (2009). We model the factors as a multi-dimensional Lévy diffusion and the underlying commodity prices are obtained by a linear mapping of the factors. This makes our model more tractable yet keeping it still very flexible as it allows us to take, for example, heat rates and spreads into account in a natural way. Our study is also related to Keppo 2004, where a similar model is used to study the hedging of swing options. We also refer to Kjaer (2008), where swing option pricing is considered under a non-Gaussian multi-factor price model. However, the analysis of Kjaer (2008) is restricted to a modification of the so-called Deng model (see Deng 2000), which is a particular mean-reverting model. In our paper, we set up and analyze a class of models where the underlying factor prices follow a general Lévy diffusion. The existing mathematical literature on swing options is mostly concerned with the pricing of a swing option. In addition to pricing, we also address the question of how to exercise a swing option optimally. From the analytical point of view, we identify using the HJB-equation an

optimal exercise policy and characterize it in an intuitive way using the notion of marginal lost option value. We also present a numerical analysis of the problem including a numerical scheme based on the finite difference method.

The reminder of the paper is organized as follows. In Sect. 2 we propose our model for the valuation of swing options. In Sect. 3 we analyze some general properties of the value function. Section 4 is devoted to the derivation of necessary and sufficient conditions for a function to coincide with the value function. We illustrate our results with explicit examples in Sect. 5, which are solved numerically in Sect. 6. Finally, we conclude in Sect. 7.

2 The valuation model

2.1 The price dynamics

As we mentioned in the introduction, the prices of the commodities are driven by multiple factors. Throughout the study, the number of commodities is m and the number of driving factors is n. The factor dynamics X are modeled by an n-dimensional Lévy diffusion. To make a precise statement, let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ be a complete filtered probability space satisfying the usual conditions, where $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ is the filtration generated by X. We assume that the factor process X are given as a strongly unique solution of the Itô equation

$$dX(t) = \alpha(t, X(t))dt + \bar{\bar{\sigma}}(t, X(t))dW_t + \int_{\mathbb{R}^l} \bar{\bar{\gamma}}(t, X(t), \xi)N(dt, d\xi), \quad (2.1)$$

where $W = (W^1, \ldots, W^{n_b})$ is an n_b-dimensional, potentially correlated, Brownian motion satisfying $d\langle W^i, W^j \rangle_t = \rho_{ij}dt$ with $\rho_{ij} \in [-1, 1]$ for all i, j. Furthermore, $N = (N^1, \ldots, N^{n_l})$ denotes an n_l-dimensional Poisson random measure with Lévy measure ν given by the independent Poisson processes $\eta^1, \ldots, \eta^{n_l}$. Here, $\nu(\{0\})$ is the unit measure concentrated on zero and it is finite. The coefficients $\alpha : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$, $\bar{\bar{\sigma}} : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n_b}$ and $\bar{\bar{\gamma}} : [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n_l} \longrightarrow \mathbb{R}^{n \times n_l}$ are assumed to be sufficiently well behaving Lipschitz continuous functions to guarantee that the Itô equation (2.1) has a unique strong solution—see Applebaum (2009), pp. 365–366. The motivation to model the randomness using Brownian and finite activity jump noise comes from electricity prices. In this framework, the jump process models the spiky behavior in the prices whereas the Brownian motion takes care of the small fluctuations.

Using the factor dynamics X, we define the m-dimensional price process $t \mapsto P_t := P(X_t)$ via the linear transformation

$$P(x) = Bx, \quad (2.2)$$

where $x \in \mathbb{R}^n$ and B is a constant $m \times n$ matrix with $rank(B) = m \leq n$. In other words, there exists constants (b_{ij}) such that $P^i(x) = \sum_{j=1}^n b_{ij}x^j$ for all $i = 1, \ldots, m$, that is, the commodity prices are linear combinations of the driving factors. The component

P^i models the time evolution of the price of the ith commodity and this price is driven by the n factors, i.e. the n-dimensional Lévy diffusion X given as the solution of the Itô equation (2.1). Since the price is linear as a function of the factors it is easy to change the model into a price model for spreads. Furthermore, the matrix B in (2.2) can be interpreted as a constant weight between the different factors X affecting the price. That allows us to take, for example, heat rates into account in our model.

In the definition of the factor dynamics, we assumed that the jump-diffusion X and the driving Brownian motion and Lévy process have all different dimensions. For notational convenience, we assume in what follows that these dimensions are the same, i.e. $n = n_b = n_l$. We point out that the following analysis holds with obvious modifications also in the case where these dimensions are different.

2.2 The valuation model

The swing option written on the price process $P = P(X)$ gives the right to purchase the given amount M of the commodities i over the time period $[0, T]$. In addition to the global constraint M, the purchases are also subject to a local constraint $\bar{u}$ which corresponds to the maximal number of swing rights that can be exercised on a given time. Since the swing option can be exercised in continuous time, the local constraint is the maximum rate at which the option can be exercised. To formalize this, let $\mathcal{U}^i = \mathcal{U}_{M^i, \bar{u}^i}$ be the set of $\mathbb{F}$-measurable, real-valued processes $u^i = u^i(X)$ satisfying the constraints

$$u_t^i \in [0, \bar{u}^i], \quad \int_0^T u_s^i ds \leq M^i,$$

for all $i = 1, \ldots, m$ and $t \in [0, T]$. Here, the elements $\bar{u} = (\bar{u}^i) \in \mathbb{R}^m$ and $M = (M^i) \in \mathbb{R}^m$. The $\mathbb{R}^m$-valued process Z defined as

$$Z_t^i = \int_0^t u_s ds, \tag{2.3}$$

where $i = 1, \ldots, m$, keeps track of the amount purchased of commodity i up to time t. In what follows, we call Z the total volume and denote the product $\bigotimes_{i=1}^m \mathcal{U}^i$ as $\mathcal{U}$. The integral representation for Z^i in (2.3) is well defined due to the local constraint.

Our valuation model is in line with a flexible load contract where the holder has no obligation to exercise. These type of contracts are traded in Scandinavian electricity markets. A practically interesting extension would be to consider the case the holder in obliged to purchase a certain volume of the commodity, that is, there is a strictly positive lower limit on Z_T. The authors are currently investigating this extension.

Denote the set $\mathcal{S} := [0, T] \times \bigotimes_{i=1}^m [0, M_i] \times \mathbb{R}^m$ and define the affine function $A : \mathbb{R}^m \to \mathbb{R}^m$ as

$$A(x) = Qx + K, \qquad (2.4)$$

where $Q = (q_{ij})$ is an $m \times m$ matrix and $K \in \mathbb{R}^m$. Define the expected present value of the total exercise payoff $J : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$ given by the rate $u \in \mathcal{U}$ from time t up to the terminal time T (or, the performance functional of u) as

$$J(t, z, p, u) = \mathbf{E}\left[\int_t^T e^{-r(s-t)} \sum_{l=1}^{m} A^l(P_s) u_s^l ds \;\middle|\; Z_t = z, X_t = x \right], \qquad (2.5)$$

where $r > 0$ is the constant discount factor. We point out that function J is defined explicitly as a function of the factors X. This corresponds to that the holder of the contract observes the underlying factor and bases her exercise decisions of this information. Furthermore, we remark that this framework covers essentially call- and put-like payoffs, where the strike prices are given by the constant vector K. Now, the value function $V : \mathcal{S} \to \mathbb{R}$ is defined as

$$V(t, z, p) = \sup_{u \in \mathcal{U}} J(t, z, p, u). \qquad (2.6)$$

We denote an optimal rate as u^*.

We make some remarks on the valuation problem (2.6). The dimension of the decision variable u is the same as the dimension of the price. That is, we can exercise the option for each price component, which corresponds to different commodities, with a different decision variable. Furthermore, we defined the function A such that it takes values in $\mathbb{R}^m$. This is done for notational convenience. Suppose that we have an m-dimensional price process but the decision variable u is k-dimensional with $k \leq m$. This corresponds to the case where m commodities are bundled into k baskets and the holder can exercise the option on the baskets. Formally this is done by defining the affine function as $A : \mathbb{R}^m \to \mathbb{R}^k$. This will not affect the form of the value function.

3 Some general properties

In this section we study some general properties of the valuation problem (2.6). We split the analysis in two cases, depending on whether $M^i \geq \bar{u}^i T$ or $M^i < \bar{u}^i T$ for a given commodity i. In the latter case, the limit M^i imposes an effective constraint on the usage of the option in the sense that the amount M^i is dominated by the amount that can be purchased if the option is exercised on full rate over the entire time horizon. This case, i.e. the case when an effective volume constraint is present, is the interesting one from the practical point of view. It is also substantially more difficult to analyze mathematically as we will see later. Before considering this case, we study the complementary case when the effective volume constraint is absent. This will give us a point of reference in the other case.

3.1 Without an effective volume constraint

We consider first the case where $M^i \geq \bar{u}^i T$ for a given commodity i. The total volume constraint for the commodity i is now superfluous, since it is possible for the holder to exercise the option at full rate throughout the lifetime of the contract. In the absence of an effective volume constraint for the commodity i, an optimal exercise rule is given by the next proposition.

Proposition 3.1 *Assume that $M^i \geq \bar{u}^i T$ for a given commodity i. Then an optimal exercise rate u^{*i} for the commodity i reads as*

$$
u_t^{*i} = \begin{cases} \bar{u}^i, & \text{if } A^i(P_t) > 0, \\ 0, & \text{if } A^i(P_t) \leq 0, \end{cases}
$$

for all $t \in [0, T]$.

Proof Let $u \in \mathcal{U}$ and $t \in [0, T]$. First, we observe that $u^{*i} \in \mathcal{U}^i$. Furthermore, we find that

$$
\mathbf{E}\left[\int_t^T e^{-r(s-t)} \sum_{l=1}^m A^l(P_s) u_s^l ds \,\middle|\, Z_t = z, \, X_t = x \right]
$$

$$
= \mathbf{E}\left[\int_t^T e^{-r(s-t)} A^i(P_t) u_s^i \mathbf{1}_{\{A^i(P_t)\leq 0\}} ds \,\middle|\, Z_t = z, \, X_t = x \right]
$$

$$
+ \mathbf{E}\left[\int_t^T e^{-r(s-t)} A^i(P_t) u_s^i \mathbf{1}_{\{A^i(P_t)>0\}} ds \,\middle|\, Z_t = z, \, X_t = x \right]
$$

$$
+ \mathbf{E}\left[\int_t^T e^{-r(s-t)} \sum_{\substack{l=1 \\ l\neq i}}^m A^l(P_t) u_s^l ds \,\middle|\, Z_t = z, \, X_t = x \right]
$$

$$
\leq \mathbf{E}\left[\int_t^T e^{-r(s-t)} A^i(P_t) u_s^{i\,*} ds \,\middle|\, Z_t = z, \, X_t = x \right]
$$

$$
+ \mathbf{E}\left[\int_t^T e^{-r(s-t)} \sum_{\substack{l=1 \\ l\neq i}}^m A^l(P_t) u_s^l ds \,\middle|\, Z_t = z, \, X_t = x \right]. \tag{3.1}
$$

Now, take supremum over all u^l on the left hand side and supremum over u^l, $l \neq i$, on the right hand side of (3.1). Since the functional J is linear in u, the same inequality still holds and, consequently, the conclusion follows. □

Proposition 3.1 states that in the absence of an effective volume constraint for commodity i, it is optimal to exercise the option whenever the payoff $A^i(P_t)$ is positive, i.e. when $(QP_t)^i \geq -K^i$. This is a natural result, since the holder does not have to worry of running out of the option over the planning horizon.

Remark 3.2 Since $\bar{u}$ is a constant vector, we find using Proposition 3.1 that the value function does not depend on z^i in the absence of an effective volume constraint for commodity i. In other words, we observe immediately that in the absence of an effective volume constraint for commodity i, the marginal value $\frac{\partial V}{\partial z^i}(t, z, p) = 0$. From the financial point of view, this is also a very natural result. Indeed, if the holder uses the option on a commodity with no effective volume constraint, the option will not lose value.

To close the subsection, we discuss how the dimension of the range of the function A affects the value given by (2.6). For simplicity, assume that there is no effective volume constraint for any of the commodities and that the function $A : \mathbb{R}^m \to \mathbb{R}^m$ is of the form

$$A(x) = \mathrm{diag}(q_1, \ldots, q_m) \cdot x + K, \tag{3.2}$$

for $K \in \mathbb{R}^m$. Using Proposition 3.1, we know that the optimal exercise rule for the valuation problem specified by the payoff structure (3.2) is

$$u_t^{*l} = \begin{cases} \bar{u}^l & \text{if } A^l(P_t) > 0 \\ 0 & \text{if } A^l(P_t) \leq 0, \end{cases} \tag{3.3}$$

for all $t \in [0, T]$ and $l = 1, \ldots, m$. Formally, we can decrease the dimension of the range of A from m, for example, as follows. Take $m' < m$ and define the $m' \times m$-matrix $\hat{Q} = (\hat{q}_{ij})$ such that each q_i occurs only once and on exactly one column of $\hat{Q}$ and the other elements are zero. In financial terms, this means that the commodities i are bundled into m' pairwise disjoint baskets with weights q_i. Then the option gives exercise rights on each of these baskets with separate exercise rates. Now, let the function $\hat{A} : \mathbb{R}^m \to \mathbb{R}^{m'}$ be $\hat{A}(x) = \hat{Q} \cdot x + \hat{K}$ with $\hat{K} \in \mathbb{R}^{m'}$ such that

$$\hat{K}_i = \sum_{\substack{j=1 \\ \hat{q}_{ij} \neq 0}}^{m} K_j, \tag{3.4}$$

for all $i = 1, \ldots, m'$. Using the same reasoning as in Proposition 3.1 we find that the optimal exercise rule for the valuation problem (2.6) given by $\hat{A}$ is

$$\hat{u}_t^l = \begin{cases} \hat{\bar{u}}^l & \text{if } \hat{A}^l(P_t) > 0 \\ 0 & \text{if } \hat{A}^l(P_t) \leq 0, \end{cases} \tag{3.5}$$

for all $t \in [0, T]$ and $l = 1, \ldots, m'$. Denote the value for m-dimensional (m'-dimensional) problem as V $(\hat{V})$. Furthermore, denote the m'-dimensional total volume variable as $\hat{Z}$ and assume that all maximal exercise rates coincide: $\bar{u}^l = \hat{\bar{u}}^{l'} = \bar{u}$ for all $l = 1, \ldots, m$ and $l' = 1, \ldots, m'$. Then, due to the structure of matrix $\hat{Q}$, we find using (3.4) that

$$
V(t, z, p) = \mathbf{E} \left[\int_t^T e^{-r(s-t)} \sum_{l=1}^m A^l(P_s) \bar{u} \mathbf{1}_{\{q_l P_s^l + K_l > 0\}} ds \; \middle| \; Z_t = z, \; X_t = x \right]
$$

$$
\geq \mathbf{E} \left[\int_t^T e^{-r(s-t)} \sum_{i=1}^{m'} \left(\sum_{\substack{j=1 \\ \hat{q}_{ij} \neq 0}}^m (\hat{q}_{ij} P_s^j + \hat{K}_j) \bar{u} \mathbf{1}_{H_s^i} \right) ds \; \middle| \; \hat{Z}_t = \hat{z}, \; X_t = x \right]
$$

$$
= \hat{V}(t, \hat{z}, p), \tag{3.6}
$$

where the events

$$
H_s^i = \left\{ \sum_{\substack{j=1 \\ \hat{q}_{ij} \neq 0}}^m (\hat{q}_{ij} P_s^j + \hat{K}_j) > 0 \right\},
$$

and the cumulative variable $\hat{Z}$ is defined analogously to (3.4). Summarizing, we have shown that by bundling commodities i into mutually disjoint baskets and, thus, reducing the dimension of the exercise rate process u, we lower the value of the option. This is, again, a natural result, since the bundling of commodities lowers flexibility of option contract in the sense that the holder must exercise the option at the same rate for all commodities in the same basket. This is in contrast to the case with separate commodities, where the exercise rates can be decided individually for each commodity.

3.2 With an effective volume constraint

In this section, we consider the case where $M_i \leq \bar{u}_i T$, in other words, the case when the total volume constraint is less than the maximal amount of commodity that can be acquired over the lifetime of the option. From a practical point of view, this is the interesting case. It is also substantially more difficult to analyze, since in this case we cannot find an optimal exercise policy explicitly as in Proposition 3.1. Instead we find the value function as the solution to the HJB-equation and an optimal exercise policy is obtained as a biproduct.

Our first task is to write the conditional expectation in (2.6) such that it depends explicitly on Z. This will be helpful in the later analysis. To this end, define the process Y as $Y_t = e^{-rt} \sum_{l=1}^m A^l(P(X_t)) Z_t^l$. Then the Itô formula yields

$$dY_t = -re^{-rt} \sum_{l=1}^{m} A^l(P(X_t))Z_t^l dt \qquad (3.7)$$

$$+ \sum_{i=1}^{n} e^{-rt} \sum_{l=1}^{m} \sum_{j=1}^{m} Z_t^l \frac{\partial A^l(P(X_t))}{\partial p^j} \frac{\partial P^j(X_t)}{\partial x^i}(\alpha_i(t, X_t)dt + \sigma_i(t, X_t)dW_t)$$

$$+ e^{-rt} \sum_{l=1}^{m} A^l(P(X_t))u_t^l dt$$

$$+ \frac{1}{2} \sum_{i,k=1}^{n} (\sigma\sigma^T)_{ik} e^{-rt} \sum_{l=1}^{m} Z_t^l \frac{\partial^2 A^l(P(X_t))}{\partial x^i \partial x^k} d\langle W^i, W^k \rangle_t$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-rt} \sum_{l=1}^{m} Z_t^l \left\{ A^l(P(X_{t^-} + \gamma^{(k)}(t, X_{t^-}, \xi^k))) - A^l(P(X_{t^-})) \right\}$$

$$N^k(dt, d\xi^k).$$

Since A is affine and P is linear, we have that

$$A^l(P(X_{t^-} + \gamma^{(k)}(t, X_{t^-}, \xi^k))) - A^l(P(X_{t^-}))$$

$$= \sum_{v=1}^{m} q_{lv} P^v(X_{t^-} + \gamma^{(k)}(t, X_{t^-}, \xi^k)) + K^l - q_{lv} P^v(X_{t^-}) - K^l$$

$$= \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(t, X_{t^-}, \xi^k)). \qquad (3.8)$$

Furthermore, we observe that for all i, j, k, l,

$$\frac{\partial A^l(P(X_t))}{\partial p^j} = q_{lj}, \quad \frac{\partial P^j(X_t)}{\partial x^i} = b_{ji}, \qquad (3.9)$$

$$\frac{\partial^2 A^l(P(X_t))}{\partial x^i \partial x^k} = 0, \qquad (3.10)$$

where q_{lj} and b_{ji} are the corresponding matrix elements given by the definitions (2.4) and (2.2), respectively. Using (3.9), denote

$$c_{li} = \sum_{j=1}^{m} q_{lj} b_{ji}, \qquad (3.11)$$

for all i, l. Then substitution of (3.8), (3.10) and (3.11) into (3.7) yields

$$dY_t = -re^{-rt} \sum_{l=1}^{m} A^l(P(X_t))Z_t^l dt + \sum_{i=1}^{n} e^{-rt} \sum_{l=1}^{m} Z_t^l c_{li}(\alpha_i(t, X_t)dt + \sigma_i(t, X_t)dW_t)$$

$$+ e^{-rt} \sum_{l=1}^{m} A^l(P(X_t))u_t^l dt$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-rt} \sum_{l=1}^{m} Z_t^l \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(t, X_{t-}, \xi^k)) N^k(dt, d\xi^k). \tag{3.12}$$

Since

$$N(dt, d\xi) = \tilde{N}(dt, d\xi) + v(d\xi)dt, \tag{3.13}$$

where v is the Lévy measure, we find that (3.12) can be written as

$$dY_t = -re^{-rt} \sum_{l=1}^{m} A^l(P(X_t))Z_t^l dt + \sum_{i=1}^{n} e^{-rt} \sum_{l=1}^{m} Z_t^l c_{li}(\alpha_i(t, X_t)dt + \sigma_i(t, X_t)dW_t)$$

$$+ e^{-rt} \sum_{l=1}^{m} A^l(P(X_t))u_t^l dt$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-rt} \sum_{l=1}^{m} Z_t^l \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(t, X_{t-}, \xi^k)) v^k(d\xi^k)dt$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-rt} \sum_{l=1}^{m} Z_t^l \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(t, X_{t-}, \xi^k)) \tilde{N}^k(dt, d\xi^k).$$

By integrating this from t to T, we obtain

$$e^{-rT} \sum_{l=1}^{m} A^l(P(X_T))Z_T^l - e^{-rt} \sum_{l=1}^{m} A^l(P(X_t))Z_t^l$$

$$= \int_t^T \left[-re^{-rs} \sum_{l=1}^{m} A^l(P(X_s))Z_s^l + e^{-rs} \sum_{i=1}^{n} \sum_{l=1}^{m} Z_s^l c_{li}\alpha_i(s, X_s) \right.$$

$$\left. + \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-rs} \sum_{l=1}^{m} Z_s^l \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) v^k(d\xi^k) \right] ds$$

$$+ \int_t^T e^{-rs} \sum_{l=1}^{m} A^l(P(X_s))u_s^l ds + \int_t^T e^{-rs} \sum_{i=1}^{n} \sum_{l=1}^{m} Z_s^l c_{li}\sigma_i(s, X_s)dW_s$$

$$+ \int_t^T \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-rs} \sum_{l=1}^{m} Z_s^l \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) \tilde{N}^k(ds, d\xi^k). \tag{3.14}$$

Consider first the Brownian integral $\int_t^T e^{-rs} \sum_{i=1}^n \sum_{l=1}^m Z_s^l c_{li} \sigma_i(s, X_s) dW_s$. Each of the integrands is of the form $e^{-rs} Z_s^l c_{li} \sigma_{ij}(s, X_s) dW_s^j$. By definition of Z_t^l, we know that $0 \leq Z_t^l \leq \bar{u}^l t$. Since Z_t^l is nondecreasing, it follows that $(Z_t^l)^2 \leq (\bar{u}^l t)^2 \leq (\bar{u}^l T)^2$. Hence,

$$
\mathbf{E}\left[\int_0^T e^{-2rs} (Z_s^l)^2 c_{li}^2 \sigma_{ij}^2(s, X_s) ds \;\middle|\; Z_0 = z, \; X_0 = x\right]
$$

$$
\leq (\bar{u}^l T c_{li})^2 \mathbf{E}\left[\int_0^T \sigma_{ij}^2(s, X_s) ds \;\middle|\; Z_0 = z, \; X_0 = x\right] < \infty.
$$

Using a martingale representation theorem, see, e.g. Applebaum (2009), Thrm. 5.3.6, we conclude that

$$
t \mapsto \int_0^t e^{-rs} \sum_{i=1}^n \sum_{l=1}^m Z_s^l c_{li} \sigma_i(s, X_s) dW_s
$$

is a martingale with respect to $\mathbb{F}$. Using the same argument, we find that the process

$$
t \mapsto \int_0^t \sum_{k=1}^n \int_{\mathbb{R}} e^{-rs} \sum_{l=1}^m Z_s^l \sum_{v=1}^m q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) \tilde{N}^k(ds, d\xi^k)
$$

is a also a martingale with respect to $\mathbb{F}$. Consequently, the conditional expectation with respect to $\mathcal{F}_t$ is zero for the last two terms in (3.14).

By multiplying (3.14) with e^{rt} on both sides, substituting into (2.6) and using the martingale properties, we find

$$
V(t, z, p) = \sup_{u \in \mathcal{U}} \mathbf{E}\left[e^{-r(T-t)} \sum_{l=1}^m A^l(P(X_T)) Z_T^l - \sum_{l=1}^m A^l(P(X_t)) Z_t^l \right.
$$

$$
- \int_t^T \left(-re^{-r(s-t)} \sum_{l=1}^m A^l(P(X_s)) Z_s^l + e^{-r(s-t)} \sum_{i=1}^n \sum_{l=1}^m Z_s^l c_{li} \alpha_i(s, X_s) \right.
$$

$$
\left. + \sum_{k=1}^n \int_{\mathbb{R}} e^{-r(s-t)} \sum_{l=1}^m Z_s^l \sum_{v=1}^m q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) v^k(d\xi^k) \right) ds
$$

$$
\left. \vphantom{\int} \quad Z_t = z, \; X_t = x \right]. \tag{3.15}
$$

Using that the measurability of $\sum_{l=1}^{m} A^l(P(X_t))Z_t^l$, we can express the value function (2.6) as

$$
\begin{aligned}
V(t, z, p) = &-\sum_{l=1}^{m} A^l(p)z^l + \sup_{u \in \mathcal{U}} \mathbf{E}\left[e^{-r(T-t)} \sum_{l=1}^{m} A^l(P(X_T))Z_T^l \right.\\
&- \int_t^T \left(-re^{-r(s-t)} \sum_{l=1}^{m} A^l(P(X_s))Z_s^l + e^{-r(s-t)} \sum_{i=1}^{n} \sum_{l=1}^{m} Z_s^l c_{li}\alpha_i(s, X_s) \right.\\
&\left. + \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-r(s-t)} \sum_{l=1}^{m} Z_s^l \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k))v^k(d\xi^k) \right) ds \Bigg|\\
&\left. Z_t = z, X_t = x \right].
\end{aligned}
$$

We now have an explicit dependence on Z in the value function, which will be useful in the proof of the following proposition. We point out that we can assume that we have an effective volume constraint in all commodities i, since the complementary case is already covered by Proposition 3.1.

Proposition 3.3 *In the presence of an effective volume constraint, i.e. when $M^i < \bar{u}^i T$, the marginal value $V_{z^i}(t, z, p) \leq 0$ for all i.*

Proof Let $u^\varepsilon = (u_1^\varepsilon, \ldots, u_m^\varepsilon) \in \mathcal{U}^\varepsilon$ be processes giving rise to admissible exercise policies $Z_t^\varepsilon = (Z_t^1 + \varepsilon, \ldots, Z_t^m + \varepsilon)$ at time t. Let $u \in \mathcal{U}$ be the processes giving rise to admissible exercise policies Z_t at time t. Since the exercise policies Z_t^ε arising from u^ε are admissible and must satisfy the effective volume constraint we have that $\mathcal{U}^\varepsilon \subseteq \mathcal{U}$. Also, for an arbitrary admissible $s \mapsto Z_s^\varepsilon$ on $(t, T]$, define an associated $\check{Z}$ as

$$
\check{Z}_s = Z_s^\varepsilon - \varepsilon, \tag{3.16}
$$

for all $s \in (t, T]$. With this in mind, we proceed by expressing the marginal value as

$$
\begin{aligned}
\frac{\partial V(t, z, p)}{\partial z^j} = &-A^j(p)\\
&+ \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left\{ \sup_{u^\varepsilon} \mathbf{E}\left[e^{-r(T-t)} \left(A^j(P(X_T))Z_j^\varepsilon(T) + \sum_{\substack{l=1\\l \neq j}}^{m} A^l(P(X_T))Z^l(T) \right) \right.\right.\\
&- \int_t^T \left[-re^{-r(s-t)} \left(A^j(P(X_s))Z_j^\varepsilon(s) + \sum_{\substack{l=1\\l \neq j}}^{m} A^l(P(X_s))Z^l(s) \right) \right.
\end{aligned}
$$

$$+ e^{-r(s-t)} \sum_{i=1}^{n} \left(Z_j^{\varepsilon}(s) c_{ji} + \sum_{\substack{l=1 \\ l \neq j}}^{m} Z^l(s) c_{li} \right) \alpha_i(s, X_s)$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-r(s-t)} \left(Z_j^{\varepsilon}(s) + \sum_{\substack{l=1 \\ l \neq j}}^{m} Z^l(s) \right)$$

$$\times \left. \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s^-}, \xi^k)) v^k(d\xi^k) \right] ds \, \bigg| \, Z_t = z, \ X_t = x \Bigg]$$

$$- \sup_u \mathbf{E} \Bigg[e^{-r(T-t)} \sum_{l=1}^{m} A^l(P(X_T)) Z^l(T)$$

$$- \int_t^T \left[-r e^{-r(s-t)} \sum_{l=1}^{m} A^l(P(X_s)) Z^l(s) + e^{-r(s-t)} \sum_{i=1}^{n} \sum_{l=1}^{m} Z^l(s) c_{li} \alpha_i(s, X_s) \right.$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-r(s-t)} \sum_{l=1}^{m} Z^l(s)$$

$$\times \left. \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s^-}, \xi^k)) v^k(d\xi^k) \right] ds \, \bigg| \, Z_t = z, \ X_t = x \Bigg] \Bigg\}. \tag{3.17}$$

By collecting the terms containing Z^{ε} in $\sup_{u^{\varepsilon}}$ and taking out the jth term in the supremum over u, we obtain

$$\frac{\partial V(t, z, p)}{\partial z^j} = -A^j(p)$$

$$+ \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left\{ \sup_{u^{\varepsilon}} \mathbf{E} \Bigg[e^{-r(T-t)} A^j(P(X_T)) Z_j^{\varepsilon}(T) \right.$$

$$- \int_t^T \left[-r e^{-r(s-t)} A^j(P(X_s)) Z_j^{\varepsilon}(s) + e^{-r(s-t)} \sum_{i=1}^{n} Z_j^{\varepsilon}(s) c_{ji} \alpha_i(s, X_s) \right.$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-r(s-t)} Z_j^{\varepsilon}(s) \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s^-}, \xi^k)) v^k(d\xi^k) \right] ds \, \bigg|$$

$$\left. Z_t = z, \ X_t = x \right] + \sup_{u^{\varepsilon}} I_1 - \sup_u I_1$$

$$- \sup_u \mathbf{E} \left[e^{-r(T-t)} A^j (P(X_T)) Z^j(T) \right.$$

$$- \int_t^T \left[-re^{-r(s-t)} A^j (P(X_s)) Z^j(s) + e^{-r(s-t)} \sum_{i=1}^n Z^j(s) c_{ji} \alpha_i(s, X_s) \right.$$

$$+ \sum_{k=1}^n \int_{\mathbb{R}} e^{-r(s-t)} Z^j(s) \sum_{v=1}^m q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) \nu^k(d\xi^k) \bigg] ds \bigg|$$

$$\left. \left. Z_t = z, \ X_t = x \right] \right\}, \qquad (3.18)$$

where

$$I_1 := \mathbf{E} \left[e^{-r(T-t)} \sum_{\substack{l=1 \\ l \neq j}}^m A^l (P(X_T)) Z^l(T) \right.$$

$$- \int_t^T \left[-re^{-r(s-t)} \sum_{\substack{l=1 \\ l \neq j}}^m A^l (P(X_s)) Z^l(s) + e^{-r(s-t)} \sum_{i=1}^n \sum_{\substack{l=1 \\ l \neq j}}^m Z^l(s) c_{li} \alpha_i(s, X_s) \right.$$

$$+ \sum_{k=1}^n \int_{\mathbb{R}} e^{-r(s-t)} \sum_{\substack{l=1 \\ l \neq j}}^m Z^l(s)$$

$$\left. \times \sum_{v=1}^m q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) \nu^k(d\xi^k) \bigg] ds \bigg| Z_t = z, \ X_t = x \right]. \qquad (3.19)$$

Furthermore, define

$$I_{\varepsilon} := \mathbf{E} \left[e^{-r(T-t)} A^j (P(X_T)) Z_j^{\varepsilon}(T) \right.$$

$$- \int_t^T \left[-re^{-r(s-t)} A^j (P(X_s)) Z_j^{\varepsilon}(s) + e^{-r(s-t)} \sum_{i=1}^n Z_j^{\varepsilon}(s) c_{ji} \alpha_i(s, X_s) \right.$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-r(s-t)} Z_j^{\varepsilon}(s)$$

$$\times \left. \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) v^k(d\xi^k) \right] ds \, \Bigg| \, Z_t = z, \, X_t = x \Bigg], \quad (3.20)$$

and

$$I_0 := \mathbf{E} \Bigg[e^{-r(T-t)} A^j(P(X_T))(Z^j(T) + \varepsilon)$$

$$- \int_t^T \Bigg[-re^{-r(s-t)} A^j(P(X_s))(Z^j(s)+\varepsilon) + e^{-r(s-t)} \sum_{i=1}^{n} (Z^j(s) + \varepsilon) c_{ji}\alpha_i(s, X_s)$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-r(s-t)} \left(Z^j(s) + \varepsilon \right)$$

$$\times \left. \sum_{v=1}^{m} q_{lv} P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) v^k(d\xi^k) \right] ds \, \Bigg| \, Z_t = z, \, X_t = x \Bigg]. \quad (3.21)$$

Then we can write (3.18) as

$$\frac{\partial V(t, z, p)}{\partial z^j} = -A^j(p)$$

$$+ \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Bigg\{ \sup_{u^{\varepsilon}} I_1 - \sup_{u} I_1 + \sup_{u^{\varepsilon}} I_{\varepsilon} - \sup_{u} I_0 + \varepsilon \mathbf{E} \Bigg[e^{-r(T-t)} A^j(P(X_T))$$

$$- \int_t^T \Bigg[-re^{-r(s-t)} A^j(P(X_s)) + e^{-r(s-t)} \sum_{i=1}^{n} c_{ji}\alpha_i(s, X_s)$$

$$+ \sum_{k=1}^{n} \int_{\mathbb{R}} e^{-r(s-t)} \sum_{v=1}^{m} q_{lv}$$

$$\times P^v(\gamma^{(k)}(s, X_{s-}, \xi^k)) v^k(d\xi^k) \Bigg] ds \, \Bigg| \, Z_t = z, \, X_t = x \Bigg] \Bigg\}. \quad (3.22)$$

Since $U^{\varepsilon} \subseteq U$ we have that $\sup_{u^{\varepsilon}} I_1 - \sup_u I_1 \leq 0$. By (3.16) there is an injective map between each functional I_{ε} and I_0 for arbitrary Z^{ε} such that $I_{\varepsilon} \hookrightarrow I_0$, hence $\sup_{u^{\varepsilon}} I_{\varepsilon} - \sup_u I_0 \leq 0$. Consequently,

$$\frac{\partial V(t, z, p)}{\partial z^j} \leq -A^j(p) + \mathbf{E}\left[e^{-r(T-t)} A^j(P(X_T)) \right.$$

$$-\int_t^T \left[-re^{-r(s-t)} A^j(P(X_s)) + e^{-r(s-t)} \sum_{i=1}^n c_{ji}\alpha_i(s, X_s) \right.$$

$$left. + \sum_{k=1}^n \int_{\mathbb{R}} e^{-r(s-t)} \sum_{v=1}^m q_{lv} \, right.$$

$$\left. \times \, P^v(\gamma^{(k)}(s, X_{s-}, \xi^k))\nu^k(d\xi^k) \right] ds \, \Big| \, Z_t = z, \, X_t = x \right]. \quad (3.23)$$

By applying the Itô formula to the process $s \mapsto e^{-rs} A^j(P(X_s))$ and taking conditional expectation with respect to $\mathcal{F}_s$, we find that the right-hand side of (3.23) is zero. □

In the case of a one-dimensional decision variable, i.e. an option of one commodity, this states an intuitively obvious result, namely that in the presence of an effective volume constraint, the usage of the option will lower its value. Note that if $M > \bar{u}T$ there is an ε such that $\mathcal{U}^\varepsilon = \mathcal{U}$ and the map (3.16) is bijective. Hence, we obtain the result of Remark 3.2.

4 The HJB-equation

In the previous section, we studied the dynamic programming problem (2.6) first in the absence of an effective volume constraint for commodity i. We showed that in this case the optimal exercise rule can be determined explicitly and that the option does not lose value if used for this commodity. We also considered the problem in the presence of an effective volume constraint and showed that in this case it loses value when used. In this section, we determine an optimal exercise rule in the presence of an effective volume constraint. To this end, we first derive the associated HJB-equation.

For the reminder of the paper, we change the notation on the value function. Since $rank(B) = m$, from now on we may write V explicitly as a function of the factors X instead of the price $P(x) = Bx$, that is, we write $V(t, z, x)$ instead of $V(t, z, p)$ where the domain of V is modified accordingly.

4.1 Necessary conditions

We derive now the HJB-equation of the problem (2.6). To this end, assume that value V exists. Then the Bellman principle of optimality yields

$$V(t, z, x) = \sup_{u \in \mathcal{U}} \mathbf{E} \left[\int_t^w e^{-r(s-t)} \sum_{l=1}^m A^l(P(X_s)) u_s^l ds \right.$$

$$\left. + e^{-r(w-t)} V(w, Z_w, X_w) \, \middle| \, Z_t = z, \; X_t = x \right], \qquad (4.1)$$

for all times $0 \le t < w \le T$. Rewrite the Eq. (4.1) as

$$\sup_{u \in \mathcal{U}} \mathbf{E} \left[\int_t^w e^{-rs} \sum_{l=1}^m A^l(P(X_s)) u_s^l ds + e^{-rw} V(w, Z_w, X_w) \right.$$

$$\left. - e^{-rt} V(t, Z_t, X_t) \, \middle| \, Z_t = z, \; X_t = x \right] = 0. \qquad (4.2)$$

Furthermore, assume that $V \in \mathcal{C}^{1,1,2}(\mathcal{S})$. Then we obtain by the Itô formula

$$e^{-rw} V(w, Z_w, X_w) - e^{-rt} V(t, Z_t, X_t) = \int_t^w d(e^{-rs} V(s, Z_s, X_s)$$

$$= \int_t^w \left[e^{-rs} V_s(s, Z_s, X_s) - r e^{-rs} V(s, Z_s, X_s) + e^{-rs} \sum_{i=1}^n V_{x^i}(s, Z_s, X_s) \alpha_i(s, X_s) \right] ds$$

$$+ \int_t^w e^{-rs} \frac{1}{2} \sum_{i,j}^n (\sigma \sigma^T)_{ij} V_{x^i x^j}(s, Z_s, X_s) d\langle W^i, W^j \rangle_s$$

$$+ \int_t^w e^{-rs} \sum_{l=1}^m V_{z^l}(s, Z_s, X_s) u_s^l ds$$

$$+ \int_t^w e^{-rs} \sum_{i=1}^n \sum_{l=1}^m V_{x^i}(s, Z_s, X_s) \sigma_i(s, X_s) dW_s$$

$$+ \int_t^w \sum_{k=1}^n \int_{\mathbb{R}} e^{-rs} \left[V(s, Z_s, X_s + \gamma^{(k)}(s, X_s, \xi^k)) - V(s, Z_s, X_s) \right]$$

$$N^k(ds, d\xi^k). \qquad (4.3)$$

Here, $\sigma_i dW_s \equiv \sum_j \sigma_{ij} dW_s^j$ and

$$x + \gamma^k = (x^1 + \gamma_1^k, \dots, x^n + \gamma_n^k),$$

where γ_j^k is the jk-th element in the matrix $\bar{\bar{\gamma}}$. By compensating the Poissonian stochastic integral in (4.3), we find under suitable L^2-assumptions on σ and V_{x_i},

see Applebaum (2009), Thrm. 5.3.6, that the Brownian and compensated Poissonian integrals in (4.3) are martingales. Thus the Eq. (4.2) yields

$$
\begin{aligned}
0 = \sup_{u \in \mathcal{U}} \mathbf{E} \Bigg[\int_t^w e^{-r(s-t)} \Bigg[& V_s(s, Z_s, X_s) + \sum_{i=1}^n V_{x^i}(s, Z_s, X_s) \alpha_i(s, X_s) \\
& + \frac{1}{2} \sum_{i,j}^n (\sigma \sigma^T)_{ij} V_{x^i x^j}(s, Z_s, X_s) \rho_{ij} - r V(s, Z_s, X_s) \\
& + \sum_{k=1}^n \int_{\mathbb{R}} \Big[V(s, Z_s, X_s + \gamma^{(k)}(s, X_s, \xi^k)) - V(s, Z_s, X_s) \Big] \nu^k(d\xi^k) \\
& + \sum_{l=1}^m \Big(A^l(P(X_s)) + V_{Z^l}(s, Z_s, X_s) \Big) u_s^l \Bigg] ds \ \Bigg| \ Z_t = z, \ X_t = x \Bigg].
\end{aligned}
\tag{4.4}
$$

Define the integro-differential operator $\mathcal{L}$ on $\mathcal{C}^{1,1,2}(\mathcal{S})$ as

$$
\begin{aligned}
\mathcal{L}F(t, z, x) = F_t(t, z, x) &+ \sum_{i=1}^n \alpha_i(t, x) F_{x^i}(t, z, x) + \frac{1}{2} \sum_{i,j}^n (\sigma \sigma^T)_{ij} F_{x^i x^j}(t, z, x) \rho_{ij} \\
&+ \sum_{k=1}^n \int_{\mathbb{R}} \Big[F(s, z, x + \gamma^{(k)}(s, x, \xi^k)) - F(s, z, x) \Big] \nu^k(d\xi^k),
\end{aligned}
\tag{4.5}
$$

and rewrite (4.4) as

$$
\begin{aligned}
0 = \sup_{u \in \mathcal{U}} \mathbf{E} \Bigg[\frac{1}{w-t} \int_t^w e^{-r(s-t)} \Bigg[& (\mathcal{L} - r) V(s, Z_s, X_s) \\
& + \sum_{l=1}^m \Big(A^l(P(X_s)) + V_{z^l}(s, Z_s, X_s) \Big) u_s^l \Bigg] ds \ \Bigg| \ Z_t = z, \ X_t = x \Bigg].
\end{aligned}
$$

Under appropriate conditions on V, see, e.g. Fleming and Soner (2006), we can pass to the limit $w \downarrow t$ and obtain the HJB-equation

$$
(\mathcal{L} - r) V(t, z, x) + \sup_u \left\{ \sum_{l=1}^m (A^l(P(x)) + V_{z^l}(t, z, x)) u^l(t) \right\} = 0,
\tag{4.6}
$$

where the u varies over the set of $\mathbb{R}^m$-valued functions defined on $[0, T]$ satisfying the conditions

$$0 \le u^l(t) \le \bar{u}^l, \quad \int_0^t u^l(s) = z^l, \quad \int_0^T u^l(s)ds \le M^l,$$

for all $l = 1, \ldots, m$ and $t \in [0, T]$.

We observe from the Eq. (4.6) that the sign of quantity $A^l(P(x)) + V_{z^l}(t, z, x), l = 1, \ldots, m$, determines whether the option should be exercised or not. From economic point of view, this quantity has a natural interpretation. Indeed, for a given commodity l, the function $A^l(P(\cdot))$ gives the instantaneous exercise payoff whereas the function V_{z^l} measures the marginal lost option value. If the payoff dominates the lost option value for a given point (t, z, x) and commodity l, the option should exercised at the full rate. That is, for each commodity l, the option should exercised according to the rule

$$\hat{u}_t^l = \begin{cases} \bar{u}^l & \text{if } A^l(P(x)) > -V_{z^l}(t, z, x), \\ 0 & \text{if } A^l(P(x)) \le -V_{z^l}(t, z, x). \end{cases}$$

We also point out that this rule is in line with the case when there is no effective volume constraint. In this case, the marginal lost option value is zero and, consequently, the option is used every time it yields a positive payoff. In particular, we find that the presence of an effective volume constraint postpones the optimal exercise of the option for a given commodity l.

4.2 Sufficient conditions

In this subsection we consider sufficient conditions for a given function to coincide with the value function (2.6). These conditions are given by the following verification theorem.

Theorem 4.1 *Assume that a function $F : \mathcal{S} \longrightarrow \mathbb{R}$ satisfies the following conditions:*

(i) $F(T, \cdot, \cdot) \equiv 0$, $F \in \mathcal{C}^{1,1,2}(\mathcal{S})$,

(ii) $(\mathcal{L} - r)F(t, z, x) + \sum_{l=1}^m (A^l(P(x)) + F_{z^l}(t, z, x))u_t^l \le 0$ for all $(t, z, x) \in \mathcal{S}$ and $u \in \mathcal{U}$, where $\mathcal{L}$ is defined in (4.5),

(iii) The processes

a) $\theta \mapsto \int_0^\theta e^{-rs} \sum_{i=1}^n F_{x^i}(s, Z_s, X_s)\sigma_i(s, X_s)dW_s$,

b) $\theta \mapsto \int_0^\theta \sum_{k=1}^n \int_{\mathbb{R}} e^{-rs} [F(s, Z_s, X_s + \gamma^{(k)}(s, \xi^k)) - F(s, Z_s, X_s)] \tilde{N}^k(ds, d\xi^k)$,

are martingales with respect to $\mathbb{F}$.

Then F dominates the value V. In addition, if there exist an admissible $\mathring{u}$ such that

$$(\mathcal{L} - r)F(t, z, x) + \sup_u \left[\sum_{l=1}^m (A^l(p(x)) + F_{z^l}(t, z, x))u_t^l \right]$$

$$= (\mathcal{L} - r)F(t, z, x) + \sum_{l=1}^m (A^l(p(x)) + F_{z^l}(t, z, x))\mathring{u}_t^l = 0, \qquad (4.7)$$

for all $(t, z, x) \in \mathcal{S}$, then $\mathring{u} = u^*$ and the function F coincides with the value V.

Proof Let $u \in \mathcal{U}$ and $t \in [0, T]$. By applying the Itô formula to the process $t \mapsto e^{-rt} F(t, Z_t, X_t)$, we find in the same way as in (4.3) that

$$
e^{-rT} V(T, Z_T, X_T) - e^{-rt} F(t, Z_t, X_t) = \int_t^T d(e^{-rs} F(s, Z_s, X_s))
$$

$$
= \int_t^T \left[e^{-rs} F_s(s, Z_s, X_s) - r e^{-rs} F(s, Z_s, X_s) + e^{-rs} \sum_{i=1}^n F_{x^i}(s, Z_s, X_s) \alpha_i(s, X_s) \right] ds
$$

$$
+ \int_t^T e^{-rs} \frac{1}{2} \sum_{i,j}^n (\sigma \sigma^T)_{ij}(s, X_s) F_{x^i x^j}(s, Z_s, X_s) d\langle W^i, W^j \rangle_s
$$

$$
+ \int_t^T e^{-rs} \sum_{l=1}^m F_{z^l}(s, Z_s, X_s) u_s^l ds
$$

$$
+ \int_t^T e^{-rs} \sum_{i=1}^n \sum_{l=1}^m F_{x^i}(s, Z_s, X_s) \sigma_i(s, X_s) dW_s
$$

$$
+ \int_t^T \sum_{k=1}^n \int_{\mathbb{R}} e^{-rs} \left[F(s, Z_s, X_s + \gamma^{(k)}(s, X_s, \xi^k)) - F(s, Z_s, X_s) \right] N^k(ds, d\xi^k).
$$

By using the assumption (i), definition of the operator $\mathcal{L}$ and the Eq. (3.13), we obtain the equality

$$
0 = e^{-rt} F(t, Z_t, X_t) + \int_t^T e^{-rs} (\mathcal{L} - r) F(s, Z_s, X_s) ds
$$

$$
+ \int_t^T e^{-rs} \sum_{l=1}^m F_{z^l}(s, Z_s, X_s) u_s^l ds
$$

$$
+ \int_t^T e^{-rs} \sum_{i=1}^n F_{x^i}(s, Z_s, X_s) \sigma_i dW
$$

$$
+ \int_t^T \sum_{k=1}^n \int_{\mathbb{R}} e^{-rs} \left[F(s, Z_s, X_s + \gamma^{(k)}(s, X_s, \xi^k)) - F(s, Z_s, X_s) \right] \tilde{N}^k(ds, d\xi^k).
$$

Conditioning up to time t and the assumption (iii) yields

$$0 = e^{-rt} F(t, Z_t, X_t) + \mathbf{E}\left[\int_t^T e^{-rs} (\mathcal{L} - r) F(s, Z_s, X_s) ds \,\bigg|\, Z_t = z, \ X_t = x \right]$$

$$+ \mathbf{E}\left[\int_t^T e^{-rs} \sum_{l=1}^m F_{z^l}(s, Z_s, X_s) u_s^l ds \,\bigg|\, Z_t = z, \ X_t = x \right].$$

By assumption (ii), we get

$$0 \le e^{-rt} F(t, Z_t, X_t) - \mathbf{E}\left[\int_t^T e^{-rs} \sum_{l=1}^m A^l (P(X_s)) u_s^l ds \,\bigg|\, Z_t = z, \ X_t = x \right].$$

$$\tag{4.8}$$

for all $u \in \mathcal{U}$. Thus, the first claim follows. Now, if there exist an admissible $\mathring{u}$ such that (4.7) holds, then we would get equality in (4.8), i.e. for all ω

$$0 = e^{-rt} F(t, Z_t, X_t) - \mathbf{E}\left[\int_t^T e^{-rs} \sum_{l=1}^m A^l (P(X_s)) \mathring{u}_s^l ds \,\bigg|\, Z_t = z, \ X_t = x \right].$$

The conclusion follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5 Examples

To illustrate our theory, we consider in this section three examples. These examples illustrate two main issues. First, we compare two one-factor spot price models (Examples 1 and 2) with underlying Ornstein–Uhlenbeck factor dynamics. The first model has a single Brownian driver whereas the other is driven by a sum of a Brownian motion and a compound Poisson process. To illustrate the effect of the jumps, the parameters of the factor dynamics are fixed such that the volatilities and the long term means are matched. In the third example, we study a two-factor model for the spot price with underlying Ornstein–Uhlenbeck factor dynamics. As we will observe, the boundary conditions in the factor price dimensions are a delicate matter in this case. We point out that the spot price dynamics in the examples are somewhat simplified and are designed for illustrative purposes. In all these examples, the aim is to find

$$V(t, z, x) = \sup_{u \in \mathcal{U}} \mathbf{E}\left[\int_t^T e^{-r(s-t)} (P(X_s) - K) u_s ds \,\bigg|\, Z_t = z, \ X_t = x \right]. \tag{5.1}$$

Here, $P(X)$ is the spot price of the commodity. The boundary conditions in the x-direction are found by the same arguments as in Benth et al. (2012). The terminal

condition is

$$V(T, z, x) = 0, \tag{5.2}$$

for all z and x. This follows directly from the definition of the value function. Furthermore, the boundary condition in the z-direction, i.e. $z = M$, is

$$V(t, M, x) = 0, \tag{5.3}$$

for all t and x. This follows from the fact that when $z = M$ the only exercise rule available in $\mathcal{U}$ is the trivial one. The conditions (5.2) and (5.3) hold for all three examples below.

5.1 Example 1

Let the factor dynamics X be given by

$$dX_s = \kappa(\mu - X_s)ds + \sigma dW_s, \quad X_t = x \tag{5.4}$$

and $P(x) = x$, where $s > t$. Then it is well known that at time $s > t$, the solution

$$X_s = (x - \mu)e^{-\kappa(s-t)} + \mu + \sigma \int_t^s e^{-\kappa(s-v)}dW_v, \quad X_t = x. \tag{5.5}$$

Furthermore,

$$X_s \sim \mathcal{N}\left(\mu + e^{-\kappa(s-t)}(x - \mu), \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa(s-t)})\right).$$

With this specification, the value function (5.1) is given as a solution to the HJB-equation

$$V_t(t, z, x) + \kappa(\mu - x)V_x(t, z, x) + \frac{1}{2}\sigma V_{xx}(t, z, x)$$
$$-rV(t, z, x) + \sup_{u \in \mathcal{U}}[(x - K + V_z(t, z, x))u_t] = 0, \tag{5.6}$$

with boundary conditions conditions in x-direction given by

$$V(t, z, x_{max}) = \bar{u} \int_t^\tau e^{-r(s-t)}[(x_{max} - \mu)e^{-\kappa(s-t)} + \mu]ds, \tag{5.7}$$

and

$$V(t, z, x_{min}) = \bar{u} \int_{\theta}^{T} e^{-r(s-t)} [(x_{min} - \mu)e^{-\kappa(s-t)} + \mu] ds,$$

(5.8)

where

$$\tau := t + \frac{M - z}{\bar{u}}$$

(5.9)

and

$$\theta := T - \frac{M - z}{\bar{u}}.$$

(5.10)

We remark that this is similar to the example in Benth et al. (2012), Appendix A. However, we consider an arithmetic OU-process whereas in Benth et al. (2012) the dynamics are given by an *exponential* OU-process.

5.2 Example 2

To illustrate the effect of the jumps in the factor dynamics, we add in this example a compound Poisson process to the factor dynamics defined in (5.4) and match the expectation and volatility with Example 1. More precisely, consider the factor dynamic given by the Itô equation

$$dX_s = \kappa(\tilde{\mu} - X_s)ds + \tilde{\sigma}dW_s + dY_s,$$

where the compound Poisson process Y_s has Lévy measure $\nu(dy) = f\alpha e^{-\alpha y}\mathbf{1}_{\{y \geq 0\}}dy$ with $f, \alpha > 0$. This equation can be written as

$$dX_s = \kappa \left(\mu + \int_0^\infty y\nu(dy) - X_s \right) ds + \tilde{\sigma}dW_s + \int_{\mathbb{R}} y\tilde{N}(dy, ds),$$

(5.11)

where the compensator

$$\int_0^\infty y\nu(dy) = \frac{f}{\alpha}.$$

With this specification, the value function (5.1) is given as a solution to the HJB-equation

$$0 = V_t(t, z, x) + \kappa(\tilde{\mu} - x)V_x(t, z, x) + \frac{1}{2}\tilde{\sigma} V_{xx}(t, z, x)$$

$$+ \int_0^\infty [V(t, z, x + y) - V(t, z, x)] f \alpha e^{-\alpha y} dy$$

$$- rV(t, z, x) + \sup_{u \in \mathcal{U}}[(x - K + V_z(t, z, x))u_t], \qquad (5.12)$$

with boundary conditions in x-direction given by

$$V(t, z, x_{max}) = \bar{u} \int_t^\tau e^{-r(s-t)}[(x_{max} - \tilde{\mu})e^{-\kappa(s-t)} + \tilde{\mu} + \frac{f}{\kappa\alpha}(1 - e^{-\kappa(s-t)})]ds,$$

$$\tag{5.13}$$

and

$$V(t, z, x_{min}) = \bar{u} \int_\theta^T e^{-r(s-t)}[(x_{min} - \tilde{\mu})e^{-\kappa(s-t)} + \tilde{\mu} + \frac{f}{\kappa\alpha}(1 - e^{-\kappa(s-t)})]ds,$$

$$\tag{5.14}$$

where τ and θ is defined as in Example 1. The solution X_s to (5.11) is given by

$$X_s = \left(x - \frac{f}{\alpha} - \tilde{\mu}\right)e^{-\kappa(s-t)} + \tilde{\mu} + \frac{f}{\alpha}$$

$$+ \tilde{\sigma} \int_t^s e^{-\kappa(s-v)} dW_v + \int_t^s \int_{\mathbb{R}} e^{-\kappa(s-v)} y \tilde{N}(dy, dv). \qquad (5.15)$$

It is easy to compute from the expression above that

$$\mathbf{E}[X_s] = \left(x - \frac{f}{\alpha} - \tilde{\mu}\right)e^{-\kappa(s-t)} + \tilde{\mu} + \frac{f}{\alpha},$$

and

$$\text{Var}(X_s) = \mathbf{E}[(X_s - \mathbf{E}[X_s])^2] = \frac{1}{2\kappa}(1 - e^{-2\kappa(s-t)})\left(\tilde{\sigma}^2 + \frac{2f}{\alpha^2}\right).$$

To match the volatility and the long term mean in Examples 1 and 2, we solve the equations above for $\tilde{\mu}$ and $\tilde{\sigma}$, when the expectation and variance is equal to that in Example 1. It follows that

$$\tilde{\mu} = \mu - \frac{f}{\alpha},$$

and

$$\tilde{\sigma} = \sqrt{\sigma^2 - \frac{2f}{\alpha^2}}.$$

Then the mean and total volatility in Examples 1 and 2 will be the same. This is good for comparison reasons, which will be discussed more in the next section.

5.3 Example 3

The purpose of this example is illustrate the results when the price is driven by multiple factors. To this end, let $\beta^1, \beta^2, \lambda_1, \mu^1, f, \kappa, K$ be non-negative constants. Consider the two-factor model

$$dX(t) = \alpha(t, X(t))dt + \bar{\bar{\sigma}}(t, X(t))dW(t) + \int_{\mathbb{R}^2} \bar{\bar{\gamma}}(t, X(t), \xi)N(dt, d\xi),$$

where $dW_t = (dW_t^1, 0)$, $N(dt, d\xi) = (0, N^2(dt, d\xi^{(2)}))$ and the Lévy measure is $\nu(dy) = (0, f\kappa e^{-\kappa y}\mathbf{1}_{\{y \geq 0\}}dy)$. Here, f is the jump frequency and κ is the parameter of the exponentially distributed jumps. Furthermore,

$$\alpha(t, X_t) = (\mu^1 - \beta^1 X_t^1, -\beta^2 X_t^2),$$
$$\bar{\bar{\sigma}}(t, X(t)) = \begin{pmatrix} \lambda^1 & 0 \\ 0 & 0 \end{pmatrix},$$
$$\bar{\bar{\gamma}}(t, X(t)) = \begin{pmatrix} 0 & 0 \\ 0 & \xi^2 \end{pmatrix}.$$

In component form we have,

$$dX_v^1 = (\mu^1 - \beta^1 X_v^1)dv + \lambda^1 dW_v^1, \quad X_t^1 = x^1,$$

and

$$dX_v^2 = -\beta^2 X_v^2 dv + \int_R \xi^2 N^2(dv, d\xi^2), \quad X_t^2 = x^2.$$

The solutions can be written as

$$X_s^1 = e^{-\beta^1(s-t)}x^1 + \int_t^s \mu^1 e^{-\beta^1(s-v)} + \int_t^s \lambda^1 e^{-\beta^1(s-v)}dW_v^1, \qquad (5.16)$$

and

$$X_s^2 = e^{-\beta^2(s-t)}x^2 + \int_t^s \int_{\mathbb{R}} e^{-\beta^2(s-v)}\xi^2 \tilde{N}^2(d\xi^2, dv)$$

$$+ \int_t^s \int_{\mathbb{R}} e^{-\beta^2(s-v)}\xi^2 v^2(dv, d\xi^2). \tag{5.17}$$

To set up the valuation model, we define the spot price function $P : \mathbb{R}^2 \to \mathbb{R}$ as $P(x) = x^1 + x^2$. Furthermore, let Z be as in (2.3) with $m = 1$. The payoff is of call option type, i.e. $A(p) = p - K$. Then, the value function (2.6) reads as

$$V(t, z, x) = \sup_{u \in \mathcal{U}} \mathbf{E}\left[\int_t^T e^{-r(s-t)}(P(X_t) - K)u_s ds \,\middle|\, Z_t = z,\ X_t = x\right].$$

From Proposition 3.1, we see that in the absence of an effective final volume constraint the optimal exercise policy u^* is given by

$$u_t^* = \begin{cases} \bar{u} & \text{if } P(X_t) > K, \\ 0 & \text{if } P(X_t) \le K. \end{cases}$$

for all $t \in [0, T]$. Hence, it is optimal to use the option whenever the swing yields a positive payoff. This is in line with Benth et al. (2012), in which no jumps are considered.

Consider now the case with an effective volume constraint. The value function can be written in the component form as

$$V(t, z, x^1, x^2) = \sup_{u \in \mathcal{U}} \mathbf{E}\left[\int_t^T e^{-r(s-t)}(X_s^1 + X_s^2 - K)u_s ds \,\middle|\, Z_t = z,\ X_t = x\right]. \tag{5.18}$$

This function is given as the solution to the HJB-equation

$$V_t(t, z, x^1, x^2) + (\mu^1 - \beta^1 x^1)V_{x^1}(t, z, x^1, x^2) - \beta^2 x^2 V_{x^2}(t, z, x^1, x^2)$$

$$+ \int_{\mathbb{R}} \left(V(t, z, x^1, x^2 + \xi^2) - V(t, z, x^1, x^2)\right) v^2(d\xi^2)$$

$$+ \frac{1}{2}\lambda^2 V_{x^1 x^1}(t, z, x^1, x^2) - rV(t, z, x^1, x^2)$$

$$+ \sup_{u \in \mathcal{U}}[(x^1 + x^2 - K + V_z(t, z, x^1, x^2))u_t] = 0. \tag{5.19}$$

With boundary conditions in t- and z-direction

$$V(T, z, x^1, x^2) = 0 \quad \text{and} \quad V(t, M, x^1, x^2) = 0,$$

and boundary conditions in x-direction

$$V(t, z, x^1_{min}, x^2), \tag{5.20}$$

$$V(t, z, x^1_{max}, x^2). \tag{5.21}$$

To solve the problem (5.19)–(5.21), we need to find the boundary conditions (5.20)–(5.21). We assume that we only have positive finite jumps, i.e. $x^2 \geq 0$ and that $0 < x^1_{min} << \mu^1$. That is, the problem is solved in the first quadrant in the $x^1 x^2$-plane.

Remark 5.1 The reason for choosing these spatial boundaries is due to the properties of the HJB-equation. In the x^1-direction we have diffusion, which requires boundary conditions at both ends. However, in the x^2-direction we have transport in the positive direction, because the coefficient in front of V_{x^2} is negative and that PDE is solved backward in time, and therefore no boundary condition is needed at $x^2 = x^2_{max}$. At $x^2 = 0$ the derivative in x^2-direction vanishes, thus no boundary condition is needed.

In what follows, the calculations rely on the fact that the underlying factor dynamics are Ornstein–Uhlenbeck processes. By plugging in the processes X^1 and X^2 given by (5.16) and (5.17), respectively, into the value function (5.18) and rearranging the terms we obtain

$$
\begin{aligned}
V(t, z, x^1, x^2) = \sup_{u \in \mathcal{U}} \Bigg\{ & \mathbf{E}\Bigg[\int_t^T e^{-r(s-t)} \left(x^1 e^{-\beta^1(s-t)} + x^2 e^{-\beta^2(s-t)} \right. \\
& + \frac{f}{\kappa\beta^2}(1 - e^{-\beta^2(s-t)}) + \frac{\mu^1}{\beta^1}(1 - e^{-\beta^1(s-t)}) - K \Bigg) u_s ds \; \Bigg| \; Z_t = z, \; X_t = x \Bigg] \\
& + \mathbf{E}\Bigg[\int_t^T e^{-r(s-t)} \left(\int_t^s \lambda^1 e^{-\beta^1(s-v)} dW_v \right. \\
& \left. + \int_t^s \int_{\mathbb{R}} e^{-\beta^2(s-v)} \xi^2 \tilde{N}^2(d\xi^2, dv) \right) u_s ds \; \Bigg| \; Z_t = z, \; X_t = x \Bigg] \Bigg\}.
\end{aligned}
\tag{5.22}
$$

To compute $V(t, z, x^1_{max}, x^2)$, we plug $x^1 = x^1_{max}$ into (5.22). Since the volatilities of the processes X^1 and X^2 are not state dependent, we can, by choosing x^1_{max} sufficiently large, expect the trajectories of the process $X^1 + X^2$ to be decreasing until the maturity since both the processes X^1 and X^2 tend towards their long term means, μ^1 and 0, respectively. Then it is optimal to start to exercise the option immediately with maximum rate until $z = M$ since $x^1_{max} + x^2$ is much larger than the long time expectation.

This argument holds since we only consider positive jumps, i.e. $x_{max}^1 + x^2 \geq x_{max}^1$. Thus, if we start the process $X^1 + X^2$ in (x_{max}^1, x^2), we can define an optimal control as

$$u_s = \bar{u} \mathbf{1}_{s \in \left[t, t + \frac{M-z}{\bar{u}}\right]}(s). \tag{5.23}$$

Then we get

$$
\begin{aligned}
V(t, z, x_{max}^1, x^2) = \bar{u} \int_t^{t + \frac{M-z}{\bar{u}}} & e^{-r(s-t)} \left(x_{max}^1 e^{-\beta^1(s-t)} + x^2 e^{-\beta^2(s-t)} \right. \\
& \left. + \frac{f}{\kappa \beta^2}(1 - e^{-\beta^2(s-t)}) + \frac{\mu^1}{\beta^1}(1 - e^{-\beta^1(s-t)}) - K \right) ds.
\end{aligned}
\tag{5.24}
$$

Here, we used the fact that the control u as defined in (5.23) is deterministic. This enables us to use the Fubini theorem and the martingale property for the second expectation in (5.22).

On the contrary, when we start the process X^1 at x_{min}^1, assumed to be sufficiently small, X^1 will increase until maturity. It is thus tempting to wait as long as possible before we use the control, cf. equation (4.6) in Benth et al. (2012). However, since x^2 can be very large and the jump frequency is state independent, we are unable to draw this conclusion. To deal with this issue, we proceed as follows. First, we assume that the value function is continuous for all $(x^1, x^2) \in \mathbb{R}^2$ and both x_{min}^1 and x_{max}^2 are finite. Using this we consider the *deterministic part* of the process $X^1 + X^2$ starting at (x_{min}^1, x^2) at time t, that is, the first expectation in the value function (5.22). We observe that the integrand is a continuous function in time and will assume a maximum (and minimum) on $[t, T]$. Suppose it has its maximum at a time t_{max}. Furthermore, assume that

The deterministic part will dominate the whole process at t_{max}.

Then, due to continuity, the deterministic part will dominate the process on an interval (t_1, t_2) that contains t_{max}. We then choose a control defined as

$$u_s = \bar{u} \mathbf{1}_{s \in [t_1, t_2]}(s).$$

We substitute this into the expression (5.22). This is a deterministic control so, again, we can use the Fubini theorem and the martingale property to get rid of the conditional expectations. Define

$$
\begin{aligned}
J(t_1, t_2) := \bar{u} \int_{t_1}^{t_2} & e^{-r(s-t)} \left[x_{min}^1 e^{-\beta^1(s-t)} + x^2 e^{-\beta^2(s-t)} + \frac{f}{\kappa \beta^2}(1 - e^{-\beta^2(s-t)}) \right. \\
& \left. + \frac{\mu^1}{\beta^1}(1 - e^{-\beta^1(s-t)}) - K \right] ds.
\end{aligned}
\tag{5.25}
$$

Then

$$V(t, z, x_{min}^1, x^2) = \max_{t_1, t_2 \in [t, T]} J(t_1, t_2) \tag{5.26}$$

which is found by solving the (deterministic) maximization problem:

$$\max_{t_1, t_2 \in [t, T]} \bar{u} \int_{t_1}^{t_2} e^{-r(s-t)} \left[x_{min}^1 e^{-\beta^1(s-t)} + x^2 e^{-\beta^2(s-t)} + \frac{f}{\kappa \beta^2}(1 - e^{-\beta^2(s-t)}) \right.$$
$$\left. + \frac{\mu^1}{\beta^1}(1 - e^{-\beta^1(s-t)}) - K \right] ds,$$

subject to

$$\bar{u}(t_2 - t_1) \le M - z. \tag{5.27}$$

We solve the limits t_1, t_2 numerically by using elementary calculus methods. That is, to solve the maximization problem, define

$$g(s, t) := e^{-r(s-t)} \left[x_{min}^1 e^{-\beta^1(s-t)} + x^2 e^{-\beta^2(s-t)} + \frac{f}{\kappa \beta^2}(1 - e^{-\beta^2(s-t)}) \right.$$
$$\left. + \frac{\mu^1}{\beta^1}(1 - e^{-\beta^1(s-t)}) - K \right].$$

This is the integrand in (5.25). By differentiating (5.25) with respect to t_1 and t_2, we obtain the first order necessary conditions

$$\frac{\partial J(t_1, t_2)}{\partial t_2} = \bar{u} g(t_2, t) = 0, \tag{5.28}$$

and

$$\frac{\partial J(t_1, t_2)}{\partial t_1} = -\bar{u} g(t_1, t) = 0. \tag{5.29}$$

Furthermore, at the boundary where $t_2 = t_1 + \frac{M-z}{\bar{u}}$, we find

$$\frac{\partial J(t_1, t_2)}{\partial t_1} = \bar{u} \left(g(t_1 + \frac{M - z}{\bar{u}}, t) - g(t_1, t) \right) = 0. \tag{5.30}$$

To conclude, we solve t_1 and t_2 from these three equations, substitute these into (5.25) and see which gives the highest value still satisfying the constraint (5.27).

6 Numerical experiments

In this section we will present the numerical solutions of the HJB-equations from the three examples in Sect. 5. All equations have the boundary conditions that the option value is zero when $t = T$ and $z = M$. In addition we truncate the boundary in infinity, and the truncated boundaries requires boundary conditions. In Example 1 we solve (5.6) with the boundary condition (5.7) and (5.8) on the truncated boundary. Similarly in Example 2 we solve (5.12) with the boundary condition (5.13) and (5.14) on the truncated boundary. And finally in Example 3 we solve (5.19) with the boundary condition (5.24) and (5.26) on the truncated boundary. Below we specify the model parameters, which are chosen for the purpose of illustration, and the discretization parameters of the numerical scheme.

6.1 Numerical scheme

In Examples 1 and 2 the HJB equations are PDEs defined over the variables t, z and x, while the HJB equation in Example 3 is defined over the variables t, z, x^1 and x^2. The equations are solved with finite difference methods (FDM). We use a first order Euler scheme in t-direction. The z- and x^2-directions are handled explicitly with first order upwind schemes, while the x or x^1-directions are handled implicitly with a second order central difference scheme. We start the time stepping at $t = T$ and go backward in time until $t = 0$.

The domain is discretized with a uniform grid in the $t-$, x^2- and z-directions whereas in the x- or x^1-direction we use an adaptive grid. The integral term is approximated with numerical integration, more precisely, the rectangle method with second order midpoint approximations. The truncated boundary in the direction of jump, i.e. the x-direction in Example 2 and the x^2- direction in Example 3, causes some problems for the approximation of the integral, which is supposed to have upper limits at infinity. This problem is solved by linearly extrapolating the option price outside the truncated domain, and integrating up to a level where we get sufficiently accurate approximation of the integral.

6.2 Numerical examples

In this subsection we study the numerical examples from the previous section. They are all motivated by some swing options traded in the Scandinavian electricity market, which are called "Brukstidskontrakt". Such a contract gives the owner the right to buy a certain amount of electricity for her own selection of hours during 1 year. More precisely, for each of the 8,760 h in 1 year the holder of the contract must choose whether or not to use the contract. In our example we set the portion to 50 %, i.e. the holder must choose 4,380 h.

In practice the contracts are usually paid in advance and not for each time it is exercised, so the strike price will be $K = 0$ in all the examples. We also use $T = 1$, $M = \frac{1}{2}$ and $\bar{u} = 1$. In the following we specify the parameters of the three examples. The discretization parameters are chosen to get accurate solutions and the model

parameters are chosen to illustrate the numerical schemes ability to solve the PDEs. The model parameters in Examples 1 and 2 are similar to those in Benth et al. (2012). The model parameters in Example 3 are chosen to look reasonable and are not calibrated to the market.

Example 1 The model is specified by $\kappa = 0.014, \mu = 40, \sigma = 2.36$. The truncated domain is defined by $x_{min} = 18.7, x_{max} = 61.3$. The discretization parameters are $\Delta t = \Delta z = \frac{1}{1000}$. The grid in $x-$direction is adaptive and consists of 671 grid points, with higher grid point density around μ and lower density near the truncated boundaries, x_{min} and x_{max}.

Example 2 The model is specified by $\kappa = 0.014, \alpha = 0.4, f = 0.04, \mu = 39.9, \sigma = 2.3387$. With these parameters the model in Examples 1 and 2 have the same mean and volatility. The grid and the truncated domain is as in Example 1.

Example 3 The model is specified by $\beta^1 = 0.014, \mu = 40, \sigma = 2.36, \beta^2 = 0.04, \kappa = 0.014, f = 0.04$. The truncated domain is defined by $x^1_{min} = 17.2, x^1_{max} = 62.8, x^2_{min} = 0, x^2_{max} = 9$. The discretization parameters are $\Delta t = \frac{1}{3200}, \Delta z = \frac{1}{3198}, \Delta x^2 = \frac{x^2_{max}}{40}$. In the x^1-direction we use an adaptive of 1,200 grid points, with higher grid point density around μ and lower density near the truncated boundary in x^1-direction, i.e. $x^1 = x^1_{min}$ and $x^1 = x^1_{max}$. On the truncated boundary condition in x^1-direction we need to calculate (5.24) and solve (5.26). With the parameters in our example, t_1 and t_2 in (5.26) turn out to be $t_1 = t$ and $t_2 = t + \frac{1}{2} - z$. The truncated boundary in x^2-direction needs no boundary condition due to the nature of the PDE. We also tried adaptive grid in the x^2-direction, but this did not seem to improve the accuracy.

In the following we visualize the numerical solution of these three examples. The two things we are most interested in are the option prices and the trigger prices. The trigger prices are also referred to as exercise curves, and they tell us when to exercise and when to hold. The option price is a function of two variables for each point in time in Examples 1 and 2, and can then be visualized in a 3D plot. However, the option price in Example 3 is a function of three variables and is therefore harder to visualize, even for a fixed point in time, but we present it in 3 plots. The trigger prices are in Benth et al. (2012) presented as exercise curves (see more in figures below). These prices are presented similarly for the results of Examples 1 and 2. But for Example 3 the trigger price is actually a 3D surface for each point in time. We solve this by projecting the surface down to two different planes. We could also have made 3D surface plots of the trigger price in this example, but we think 2D plots are more instructive.

Figure 1a, b show the option price at $t = 0.5$ for Examples 1 and 2, respectively. We see that the two plots a quite similar, and that the option price increases with increasing x-values. This makes sense from an economical point of view, since one would expect that a higher spot price results in a higher option price. Mathematically we see it from the value function (5.1) and the fact that the solution functions (5.5)
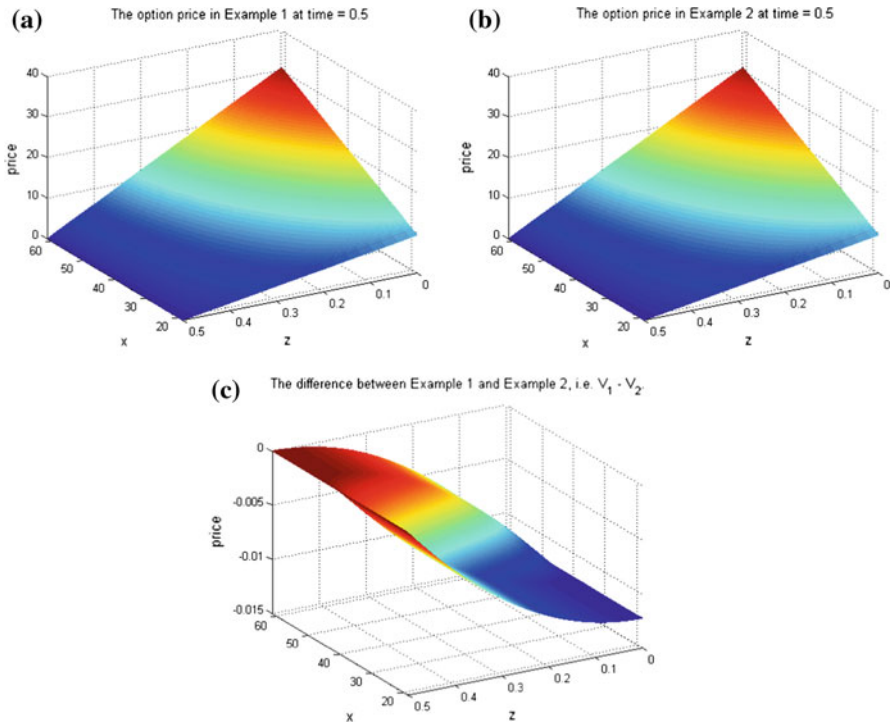
Fig. 1 Option prices in Examples 1 and 2. **a** Option price Example 1. **b** Option price Example 2. **c** Difference between option prices in Examples 1 and 2

and (5.15) are increasing functions of x. Furthermore, the option price decreases with increasing z-values which is in accordance with Proposition 3.3, stating that whenever using the option it loses value, which makes economical sense.

Figure 1c shows the difference in option price between Example 1 and 2 at the same time level. More precisely Fig. 1c shows the price from Example 2 minus the price from Example 1. We see that this difference is negative, which means that the option price is a little higher when we assume an underlying jump process. This is reasonable because of the distributional differences of the underlying processes. Indeed, the Gaussian OU-process in Example 1 has a symmetric distribution whereas the non-Gaussian OU-process in Example 2 has a positively skewed distribution. This positive skewness, which increases value in financial markets, is caused by the fact that we only have positive jumps.

In Fig. 2 we show the exercise curves for Examples 1 and 2 at time = 0.5. The red curve corresponds to Example 2 and the black curve corresponds to Example 1. We observe that the red curve lies more to the right than the black curve. This means the the optimal trigger price is higher in the presence of positive jumps, that is, positive jumps postpone the optimal exercise. This is an intuitive result, since the positivity of jumps implies that the distribution of the prices is positively skewed. Therefore the holder of the contract has an incentive to wait longer to exercise.
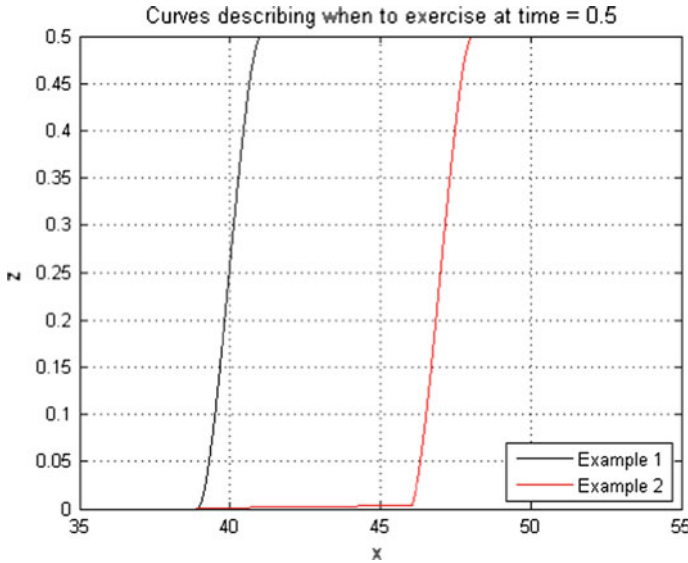
Fig. 2 Exercise curves for Examples 1 and 2 at time = 0.5. Exercise when at points to the *right* of the curve and hold when at points to the *left* of the curve

In Fig. 3a–c we show the option price of Example 3 at $t = 0.5$. The function is plotted as a function of x^1 and z for three values of x^2. We see that for each value of x^2 the plot looks similar to the plots in Fig. 1a, b, only the level of the surfaces changes a little.

In Fig. 4a–b we visualize the trigger price for $t = 0.5$ in Example 3. Figure 4a shows the exercise surface projected down to the price,z-plane (where price $= x^1 + x^2$), for various values of x^2. Figure 4b shows the exercise surface projected down to the x^1x^2-plane, for various values of z.

It is worth noting that the slope of the curves in Fig. 4b is approximately $-\frac{\beta_1}{\beta_2} = -0.35$. For example if we study the red curve and remove the point where $x^2 = 0$, and make a linear least squares approximation of it, it will have a slope of -0.3517. This means that the negative ratio of the mean reversion speeds approximates the slope of the exercise curves in the x^1x^2-plane. This is plausible from an economical point of view for the following reason. If the mean reversion speed is smaller for x^1 than for x^2, the holder will exploit a deviation from the long term mean earlier for process X^2 than X^1 by exercising the option. That is, she would require a higher contribution from x^1 to the price than from x^2 before exercising. This is because a high value of x^2 is likely to reduce more quickly and therefore it is beneficial to exercise with a lower value of x^2 in relation to x^1. On the contrary, if x^1 has a high price it is more likely to stay high longer. In this case, the holder might wait for even higher prices. A similar reasoning can be done for the opposite case when the mean reversion speed is bigger for x^1 than for x^2.

Notice that at $t = 0.5$ and for any value of z, the holder should always hold if the values of x^1 and x^2 are small enough. The reason for this is the following: if the
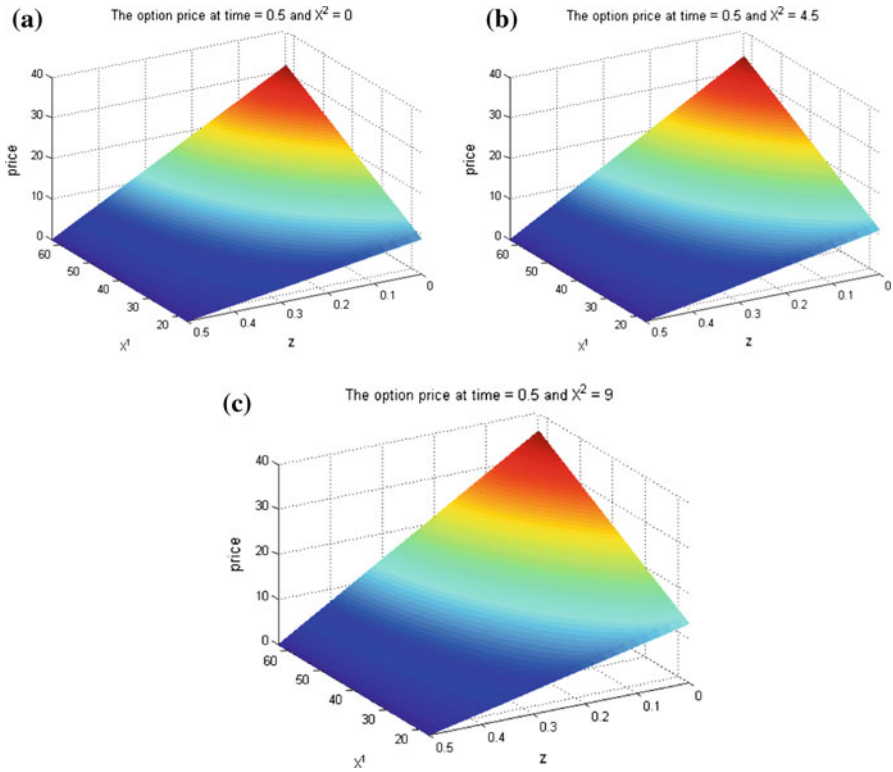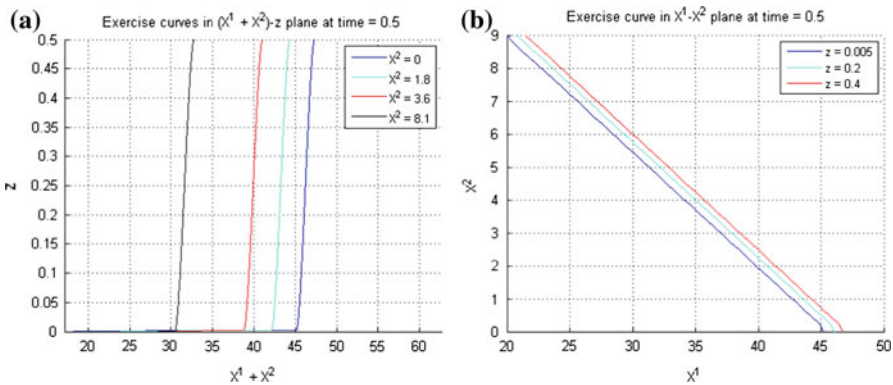
line $z = 0.001$. This is due to numerical error/instability. For finer grids the exercise curves will be closer to $z = 0$ at for small values of $P = X^1 + X^2$ and $t = 0.5$. Similar effects can be observed for other values of t. The solution to this problem is either higher resolution on the grid, which requires more memory on the computer, or more accurate numerical schemes. The curves in Fig. 4b bends a little in the lower right corner, and this is due to the mentioned instability in the numerical scheme.

6.3 Numerical accuracy in Example 3

The analysis of the numerical scheme in Examples 1 and 2 is similar to that of Benth et al. (2012). To analyse Example 3 we study the Courant–Friedrichs–Lewy (CFL) condition. The CFL number of the HJB equation in Example 3 is

$$
\begin{aligned}
C &= \frac{\Delta t}{\Delta x^2} \beta^2 x^2 + \frac{\Delta t}{\Delta z} \bar{u} \\
&= \frac{\frac{1}{3200}}{\frac{9}{40}} 0.04 x^2 + \frac{\frac{1}{3200}}{\frac{1}{3198}} 1 \\
&= \frac{1}{3200} \left(1.6 \frac{x^2}{9} + 3198 \right).
\end{aligned}
$$

A necessary condition for convergence is that the CFL number $C \le C_{max}$, and in our case with $x^2 \in (0, 9)$ we see that the CFL number $C \le 1$. As mentioned we have observed some small instabilities in the numerical solution. We have also tried with larger values of Δz compared to Δt, but the reported discretization parameters seem to give most accurate solutions. In order to establish convergence an implicit scheme should be developed, but this is not done in this work.

In the following we will present some evidence that the numerical solution converges to the correct solution of the HJB equation. We attempt to evaluate both the numerical scheme and the calculated boundary conditions using (5.21). This is done by trying to see how well they fit for extreme values of x^1, i.e. $x^1 >> \mu$. We have used a very fine grid to solve the HJB equation where V is required to be linear in x^1-direction at the boundaries. This is similar to Benth et al. (2012), where it is shown that this type of inaccurate boundary condition gives quite accurate solutions. Now this solution can be compared to the values we get from Eq. (5.24).

The difference between the numerical solution of the HJB equation and the value calculated by (5.24) is illustrated in Fig. 5. We see that the difference is relatively small compared to the option value. For this example with $t = 0.5$ and $z = 0.4$ it is about 6 orders of magnitude lower than the option value. This is sufficiently small for us to trust the numerical solver. The difference may come from all of the following five sources: assumptions that the control is as described in (5.23), numerical inaccuracy/instability of the scheme, truncation in x^1 direction, truncation in x^2 direction, extrapolation in x^2 direction and linear boundary condition on the PDE solver.
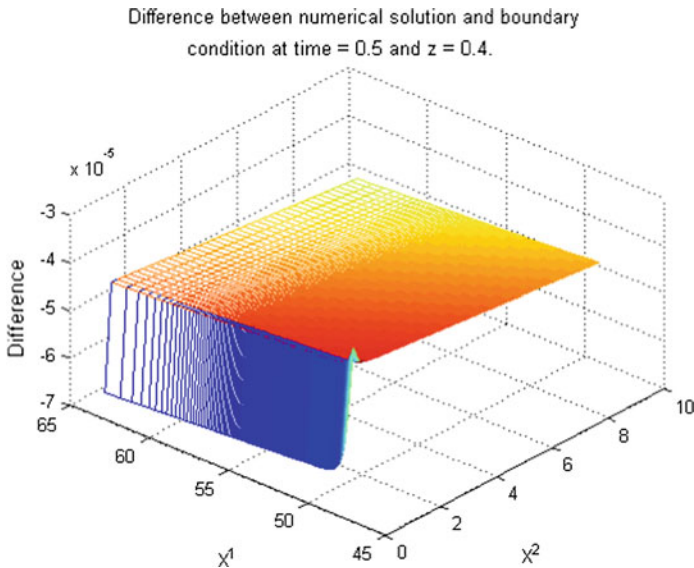
Fig. 5 Difference between the numerical solution and (5.21) for large values of x^1

7 Conclusions

In this paper, we developed and analyzed a valuation model for swing options on multi-commodity markets. The model is formulated as a dynamic programming problem, where the underlying dynamic structure is given by a multi-dimensional Lévy diffusion. This process models the price evolution of the commodities. The commodity prices are driven by a multi-dimensional Brownian motion and a multi-dimensional compound Poisson process. The introduction of the compound Poisson process is important since it allows non-Gaussian price evolution. This is important, in particular, on electricity markets, see, e.g. Benth et al. (2007); Hambly et al. (2009); Kjaer (2008). Furthermore, this model allows us to take into account jumps in price processes, which is also important on electricity markets.

From a analytical point of view, this study provides a multi-dimensional generalization of the analysis in Benth et al. (2012). First, we analyze the model in the absence of an effective volume constraint. Along the lines of Benth et al. (2012), we find that in this case the option does not loose value if used. Moreover, we prove that in the presence of an effective volume constraint for a given commodity, the usage of the option for this commodity will lower the value of the option. This is a intuitively appealing from the economical point of view. To tackle the problem of finding an optimal exercise rule and the price of the option, we analyze the pricing problem using the Bellman principle of optimality and derive the associated HJB-equation. In Sect. 4, we obtained an optimal exercise rule which states that if the immediate exercise payoff dominated the lost option value for a given commodity, then the option on this commodity should be exercised at a full rate. In particular, we conclude that this optimal exercise rule

is a bang–bang rule. We also provide a verification theorem, which states conditions under which a given function coincides with the value function.

In addition we illustrate the results with three examples which we study numerically. We set up a straightforward FDM scheme to solve the associated HJB-equations. The numerical experiments seems to give reasonable results from both mathematical and economical points of view. In the last of our examples we have also given evidence for convergence of the numerical solution.

References

Applebaum D (2009) Lévy processes and stochastic calculus, 2nd edn. Cambridge University Press, Cambridge

Bender C (2011) Primal and dual pricing of multiple exercise options in continuous time. SIAM J Financ Math 2(1):562–586

Benth FE, Kallsen J, Meyer-Brandis T (2007) A non-Gaussian Ornstein–Uhlenbeck process for electricity spot price modeling and derivative pricing. Appl Math Financ 14:153–169

Benth FE, Lempa J, Nilssen TK (2012) On the optimal exercise of swing options in electricity markets. J Energy Mark 4(4):3–28

Bjorgan R, Song H, Liu C-C, Dahlgren R (2000) Pricing flexible electrcity contracts. IEEE Trans Electr Syst 15(2):477–482

Burger M, Graebler B, Schindlmayr G (2007) Managing energy risk. Wiley Finance, New York

Carmona R, Touzi N (2008) Optimal multiple stopping and valuation of swing options. Math Financ 18(2):239–268

Dahlgren M (2005) A continuous time model to price commodity based swing options. Rev Deriv Res 8(1):27–47

Deng S (2000) Stochastic models of energy commodity prices and their applications: mean-reversion with jumps and spikes. Power working paper 073, University of California Energy Institute

Edoli E, Fiorenzani S, Ravelli S, Vargiolu T (2012) Modeling and valuing make-up clauses in gas swing valuation. Energy Econ. doi:10.1016/j.eneco.2011.11.019

Fleming WH, Soner M (2006) Controlled Markov processes and viscosity solutions, 2nd edn. Springer, New York

Haarbrücker G, Kuhn D (2009) Valuation of electricity swing options by multistage stochastic programming. Automatica 45:889–899

Hambly B, Howison S, Kluge T (2009) Modelling spikes and pricing swing options in electricity markets. Quant Financ 9(8):937–949

Jaillet P, Ronn M, Tompadis S (2004) Valuation of commodity based swing options. Manag Sci 14(2):223–248

Keppo J (2004) Pricing of electricity swing contracts. J Deriv 11:26–43

Kiesel R, Gernhard J, Stoll S-O (2010) Valuation of commodity based swing options. J Energy Mark 3(3):91–112

Kjaer M (2008) Pricing of swing options in a mean reverting model with jumps. Appl Math Financ 15(5):479–502

Lund A-C, Ollmar F (2003) Analyzing flexible load contracts (preprint)

Marshall TJ, Mark Reesor R (2011) Forest of stochastic meshes: a new method for valuing high-dimensional swing options. Oper Res Lett 39:17–21

Wahab MIM, Yin Z, Edirisinghe NCP (2010) Pricing swing options in the electricity markets under regime-switching uncertainty. Quant Financ 10(9):975–994