COINTEGRATION IN CONTINUOUS TIME FOR FACTOR MODELS

FRED ESPEN BENTH AND ANDRE SÜSS

Abstract. We develop cointegration for multivariate continuous-time stochastic processes, both in finite and infinite dimension. Our definition and analysis are based on factor processes and operators mapping to the space of prices and cointegration. The focus is on commodity markets, where both spot and forward prices are analysed in the context of cointegration. We provide many examples which include the most used continuous-time pricing models, including forward curve models in the Heath-Jarrow-Morton paradigm in Hilbert space.

1. Introduction

We aim at developing a formalism to the concept of *cointegration* in continuous time. Cointegration has since the seminal paper of Engle and Granger [19] become a very popular concept for stochastic modelling of dependent time series of data, in particular in economics. For example, the price series of two financial assets can be non-stationary, while one may find that a linear combination of these is stationary. Cointegration provides a framework for analysing and modelling time series that explains such observable features in data.

Although there has been a huge development in continuous-time financial models over the last decades, the literature on cointegration for continuous time stochastic processes and its application to finance is relatively scarce. A non-exhaustive list of papers in this stream of research include Comte [15], Duan and Pliska [17], Duan and Theriault [18], Nakajima and Ohashi [34], Paschke and Prokopczuk [36], Benth and Koekebakker [10], and recently Farkas et al. [21]. In the present paper we formalise ideas on cointegration in continuous time for factor processes, and extend these to cointegration for stochastic processes with infinite dimensional state space. The latter will provide a theoretical framework for studying cointegration in forward and futures markets, say.

Comte [15] presents an in-depth analysis on classical cointegration and its extension to continuous-time models, where continuous-time autoregressive moving average processes (CARMA) play a central role. Duan and Pliska [17] analyse a

Date: April 24, 2018.

Key words and phrases. cointegration, infinite dimensional stochastic processes, polynomial processes, forward prices, commodity markets.

F. E. Benth acknowledges financial support from the project FINEWSTOCH funded by the Norwegian Research Council. A. Süss acknowledges partial financial support by the grant MTM 2015-65092-P, Secretaria de estado de investigacion, desarrollo e innovacion, Ministerio de Economia y Competitividad, Spain. Two anonymous referees are thanked for their careful reading and constructive criticism of a former version of this paper, leading to a significantly improved presentation.

specific cointegrated asset price model, and show that pricing options will not be influenced by cointegration. Their paper has triggered many theoretical and empirical studies, including Nakajima and Ohashi [34], Paschke and Prokopczuk [36], Benth and Koekebakker [10] and Farkas et al. [21]. Duan and Theriault [18] extend cointegration to continuous-time forward price models. Benth and Koekebakker [10], and more recently Benth [6] focus on the relationship between cointegration in spot and forward markets, and propose cointegration models for forward markets. Contrary to the conclusions of Duan and Pliska [17], these two papers argue that in commodity markets the pricing measure may preserve cointegration. We refer to Back and Prokopczuk [4] for a review of modelling and pricing in commodity markets.

Starting with spot price models, we discuss a framework for cointegration based on factor models. Our concept makes use of a set of stochastic processes, which we call factors, which explains the dynamics of prices via a linear transformation. This yields a vector-valued price dynamics, for which one can introduce the concept of cointegration. The following example is frequently referred to in the text, and explains our ideas in a simple setting.

Example 1. *Consider the classical spot price model for two commodity markets, given by the two-factor model;*

$$S_i(t) = X_i(t) + X_3(t), i = 1, 2. \tag{1}$$

We assume $X_3(t) = \mu t + \sigma B_3(t)$ *being a drifted Brownian motion and* $X_i(t)$ *being two Ornstein-Uhlenbeck processes,*

$$dX_i(t) = -\alpha_i Y_i(t)\, dt + \eta_i\, dB_i(t),$$

with constants $\alpha_i > 0, \eta_i > 0, i = 1, 2$*. Here,* (B_1, B_2, B_3) *is a trivariate Brownian motion, possibly correlated. This model was proposed in the univariate case by Lucia and Schwartz* [32] *for electricity spot prices and extended to cross-commodity markets by Paschke and Prokopczuk* [35] *for oil markets (see also Duan and Pliska* [17] *and Benth and Koekebakker* [10] *for general analysis). For example,* (S_1, S_2) *can model the joint spot price dynamics in the coal and electricity market, or in two different electricity markets. Since the bivariate Ornstein-Uhlenbeck process* (X_1, X_2) *admits a limiting Gaussian distribution, the price difference process* $S_1(t) - S_2(t)$ *will have a limiting distribution. On the other hand, each marginal price process* S_i *is non-stationary since the drifted Brownian motion* X_3 *is (unless* $\mu = \sigma = 0$*). According to Duan and Pliska* [17]*, the processes* S_1 *and* S_2 *are cointegrated. We notice that the definition of the bivariate price process* (S_1, S_2) *involves three factor processes* X_1*,* X_2 *and* X_3*, and a linear combination of these. Introducing the matrix*

$$\mathcal{P} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \tag{2}$$

we represent the vector $\mathbf{S} := (S_1, S_2)^\top$ *as* $\mathbf{S}(t) = \mathcal{P}(X_1(t), X_2(t), X_3(t))^\top$*. We assume that all elements* $\mathbf{x} \in \mathbb{R}^n$*,* $n \in \mathbb{N}$*, are coloumn vectors, and* $\mathbf{x}^\top$ *is the transpose of* $\mathbf{x}$*. Cointegration is achieved since there exists a vector* $\mathbf{c} = (1, -1)^\top \in \mathbb{R}^2$ *such that the process* $\mathbf{c}^\top \mathbf{S} = X_1 - X_2$ *admits a limiting distribution.*

Based on multivariate spot price models of the form introduced in this example, we analyse forward prices derived from processes with certain affinity properties. In this context, polynomial processes (see Filipović and Larsson [24]) constitute an important case, along with the more specific CARMA processes. We present results on the cointegration relationship between spot and forward markets, with a particular view towards applications in commodity markets.

Our analysis of spot and forward markets motivates the definition of cointegration for stochastic processes in infinite dimensions. We introduce a concept for modelling cointegrated forward curves, following the HJM-paradigm (see Heath, Jarrow and Morton [29]) of modelling forward prices directly rather than explaining these via spot models (we refer to Benth, Šaltytė Benth and Koekebakker [7] for an extensive analysis of forward modelling in energy markets). Since forward curves can be modelled as stochastic processes in Hilbert space of real-valued functions on $\mathbb{R}_+$ (see Benth and Krühner [11, 12]), we concentrate our analysis on formulating cointegration via linear operators on Hilbert spaces. We show how cointegration in Hilbert space can be related to the finite dimensional case. It turns out that product Hilbert spaces provide a natural framework for modelling, and we give several examples including infinite dimensional factor processes capturing stationary and non-stationary effects as well as non-Gaussianity. We also include a discussion of some recent empirical studies on forward gas markets by Geman and Liu [28] viewed in our cointegration context.

The results of this paper are presented as follows: in Section 2 we define and analyse cointegration for multivariate spot price models based on factor processes. The question of forward pricing in cointegrated spot markets is analysed in Section 3, where we give a description of cointegration of forwards. Finally, in Section 4 we introduce cointegration for Hilbert-space valued stochastic processes, and apply this to cross-commodity forward prices modelled within the HJM-approach.

2. Cointegration for factor models

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space equipped with a right-continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$ where $\mathcal{F}_t$ contains all sets of $\mathcal{F}$ of probability zero (i.e., satisfying the *usual conditions*). Let $\{\mathbf{S}(t)\}_{t\geq 0} \in \mathbb{R}^d$ be $d \geq 2$ asset prices in a given market, where we define

$$\mathbf{S}(t) = \mathcal{P}\mathbf{X}(t)\,, t \geq 0 \tag{3}$$

for an adapted stochastic process $\{\mathbf{X}(t)\}_{t\geq 0} \in \mathbb{R}^n$ and $\mathcal{P} \in \mathbb{R}^{d\times n}$. The matrix $\mathcal{P}$ is hereafter referred to as the *pricing matrix*, and $\mathbf{X}$ the *factor process* of the market. We assume that the number of factors n is at least equal to the number of assets d, i.e., $n \geq d$. **Moreover, it is also assumed that the coordinate processes are linearly independent, e.g., that $\mathbb{P}(\mathbf{x}^\top\mathbf{X}(t) = 0\,, t \geq 0) = 0$ for any $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} \neq 0$. In particular, this assumption means that we cannot have any coordinate processes being identically equal to zero, nor that a coordinate process is simply a scaling of another.** We reserve the notation $\{\mathbf{e}_i\}_{i=1}^k$ for the canonical basis vectors in $\mathbb{R}^k$, where the dimension $k \in \mathbb{N}$ will be clear from the context.

Definition 2. *$\mathcal{P}$ is minimal if $\mathbf{e}_i \notin ker(\mathcal{P})$ for all $i = 1, \ldots, n$.*

The definition expresses that all factors $X_i(t), i = 1, \ldots, n$ are represented in $\mathbf{S}(t)$. Indeed,

$$\mathbf{X}(t) = \sum_{j=1}^{n} X_j(t)\mathbf{e}_j,$$

where $X_j(t) = \mathbf{X}^\top(t)\mathbf{e}_j$. Hence,

$$\mathbf{S}(t) = \sum_{j=1}^{n} X_j(t)\mathcal{P}\mathbf{e}_j.$$

If $\mathbf{e}_i \in \ker(\mathcal{P})$ for given $i \in \{1,\ldots,n\}$, then $\mathbf{S}$ will not depend on X_i. Opposite, if $\mathbf{S}$ is not depending on X_i, then $\mathcal{P}\mathbf{e}_i = 0$. Under minimality of $\mathcal{P}$, we have all necessary factors to define the price dynamics $\mathbf{S}$. Notice that the definition *does not* **say that all factors are present in each price coordinate $S_j(t), j = 1,\ldots,d$. On the other hand, under minimality, we have removed any redundant factor process and use only those factors $\mathbf{X}$ necessary to model the dynamics of $\mathbf{S}$.**

We restrict our considerations to minimal pricing matrices $\mathcal{P}$. **Furthermore, we also make the assumption that $\mathcal{P}$ has full rank, rank$(\mathcal{P}) = d$. The full rank condition on $\mathcal{P}$ is equivalent to $\mathbb{P}(\mathbf{x}^\top\mathbf{S}(t) = 0\,, t \geq 0) = 0$ for any $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} \neq 0$. In other words, we assume that we have a market of d linearly independent asset price dynamics.**

Assuming the price process $\mathbf{S}$ is defined by (3) with $\mathcal{P}$ being a minimal pricing matrix having full rank, we define cointegration as follows: Let $P_{\mathbf{X}}(t,\cdot)$ denote the probability distribution of $\mathbf{X}(t)$ defined on $\mathcal{B}(\mathbb{R}^n)$, the Borel σ-algebra on $\mathbb{R}^n$. For any $\mathbf{x} \in \mathbb{R}^n$, we denote by $P_{\mathbf{x}^\top\mathbf{X}}(t,\cdot)$ the probability distribution of the real-valued random variable $\mathbf{x}^\top\mathbf{X}(t)$. Furthermore, denote by $\Psi_{\mathbf{X}}(t,\mathbf{z})$ the characteristic function of $\mathbf{X}(t)$, defined for $\mathbf{z} \in \mathbb{R}^n$ as

$$\Psi_{\mathbf{X}}(t,\mathbf{z}) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\mathbf{z}^\top\mathbf{X}(t)}\right].$$

Sometimes, one is using the cumulant function instead of the characteristic function, where the cumulant $\kappa_{\mathbf{X}}(t,\mathbf{z}) := \log\Psi_{\mathbf{X}}(t,\mathbf{z})$ with log denoting the distinguished logarithm (see e.g. Sato [38]). We see that $\Psi_{\mathbf{X}}(t,\mathbf{z}) = \widehat{P}_{\mathbf{X}}(t,\mathbf{z})$, where $\widehat{P}_{\mathbf{X}}(t,\mathbf{z})$ is the Fourier transform of the distribution $P_{\mathbf{X}}(t,\cdot)$.

Definition 3 (Definition of cointegration). *We say that $\mathbf{S}$ is cointegrated if there exists $\mathbf{c} \in \mathbb{R}_0^d := \mathbb{R}^d\backslash\{\mathbf{0}\}$ and a probability distribution $\mu_{\mathbf{c}}$ on $\mathcal{B}(\mathbb{R})$ such that $P_{\mathbf{c}^\top\mathcal{P}\mathbf{X}}(t,\cdot)$ converges to $\mu_{\mathbf{c}}$ when $t \to \infty$. We call $\mathbf{c}$ a cointegration vector for $\mathbf{S}$.*

In the definition of cointegration, the convergence **is understood to be** *weak convergence of probability measures* **(see Definition 2.2 in Sato [38]). In the literature, this is frequently expressed as $\mathbf{c}^\top\mathcal{P}\mathbf{X}(t)$ converges** *in distribution* **to $\mu_{\mathbf{c}}$ when $t \to \infty$, denoted by $\mathbf{c}^\top\mathcal{P}\mathbf{X}(t) \xrightarrow{d} \mu_{\mathbf{c}}$.**

Our definition of cointegration means that there exists a linear combination of the price process vector $\{\mathbf{S}(t)\}_{t\geq 0}$ which admits a limiting probability distribution, where the linear combination is represented by the cointegration vector $\mathbf{c}$. We recall that $\mathbf{c}^\top\mathbf{S}(t) = \mathbf{c}^\top\mathcal{P}\mathbf{X}(t)$, and thus $P_{\mathbf{c}^\top\mathcal{P}\mathbf{X}}(t,\cdot)$ is the probability distribution of $\mathbf{c}^\top\mathbf{S}(t)$, which must converge to a probability distribution when time tends to infinity in order to achieve cointegration.

Definition 3 excludes the case $\mathbf{c} = \mathbf{0}$. Obviously, $P_{\mathbf{0}^\top\mathcal{P}\mathbf{X}}(t,\cdot) = \delta_{\mathbf{0}}(\cdot)$, with $\delta_{\mathbf{0}}$ being the Dirac measure at zero, which is trivially cointegrated for any choice of $\mathcal{P}$ and $\mathbf{X}$. From a practical viewpoint, we are obviously not interested in this degenerate case. Remark in Definition 3 that $\mathbb{R}_0^d$ denotes $\mathbb{R}^d$ with $\mathbf{0}$ excluded, a notation we will use frequently. If we choose

$n = d$ and $\mathcal{P} = I$, the $d \times d$ identity matrix, we have $\mathbf{S} = \mathbf{X}$. Thus the definition of cointegration can also be directly applied for the factor process.

As the next result shows, cointegration can be characterized by convergence of characteristic functions when time tends to infinity:

Proposition 4 (Cumulant characterisation of cointegration)**.** *If S is cointegrated with cointegration vector $\mathbf{c} \in \mathbb{R}_0^d$, then $\lim_{t\to\infty} \Psi_{\mathbf{X}}(t, z\mathcal{P}^\top \mathbf{c}) = \Psi_{\mu_{\mathbf{c}}}(z)$ uniformly (in $z \in \mathbb{R}$) on any compact set, with $\Psi_{\mu_{\mathbf{c}}}$ being the characteristic function of the distribution $\mu_{\mathbf{c}}$. Opposite, if there exists a $\mathbf{c} \in \mathbb{R}_0^d$ and a complex-valued function $z \mapsto \Psi_{\mathbf{c}}(z)$ on $\mathbb{R}$ which is continuous at $z = 0$ such that $\lim_{t\to\infty} \Psi_{\mathbf{X}}(t, z\mathcal{P}^\top \mathbf{c}) = \Psi_{\mathbf{c}}(z)$ for every $z \in \mathbb{R}$, then $\mathbf{S}$ is cointegrated with cointegration vector $\mathbf{c}$.*

Proof. By Definition 3 of cointegration we have that $P_{\mathbf{c}^\top \mathcal{P}\mathbf{X}}(t, \cdot) \to \mu_{\mathbf{c}}$ for a probability distribution $\mu_{\mathbf{c}}$. It is well-known (see e.g., Sato [38, Prop. 2.5 (vi)]) that this implies $\Psi_{\mathbf{X}}(t, z\mathcal{P}^\top \mathbf{c}) \to \Psi_{\mu_{\mathbf{c}}}(z)$ as $t \to \infty$ on any compact set. This shows the first part.

If for a given $\mathbf{c} \in \mathbb{R}_0^d$, there exists a $\Psi_{\mathbf{c}}$ such that $\Psi_{\mathbf{X}}(t, z\mathcal{P}^\top \mathbf{c}) \to \Psi_{\mathbf{c}}(z)$ for every $z \in \mathbb{R}$, then Sato [38, Prop. 2.5 (viii)] ensures the existence of a probability distribution $\mu_{\mathbf{c}}$ with characteristic function $\Psi_{\mathbf{c}}$ as long as $z \mapsto \Psi_{\mathbf{c}}(z)$ is continuous at $z = 0$. But this means that $P_{\mathbf{c}^\top \mathcal{P}\mathbf{X}}(t, \cdot) \to \mu_{\mathbf{c}}$. The result follows. □

If $\mathbf{c}$ is a cointegration vector for $\mathbf{S}$, then by appealing to Definition 3, $\mathcal{P}^\top \mathbf{c} \in \mathbb{R}_0^d$ is a cointegration vector for the factor process $\mathbf{X}$. Opposite, if $\mathbf{a} \in \mathbb{R}_0^n$ is a cointegration vector for $\mathbf{X}$, that is, there exists a distribution μ such that $P_{\mathbf{a}^\top \mathbf{X}}(t, \cdot) \to \mu$, then for any $\mathbf{c} \in \mathbb{R}_0^d$ for which $\mathcal{P}^\top \mathbf{c} = \mathbf{a}$ becomes a cointegration vector for $\mathbf{S}$. We note that such a $\mathbf{c}$ may fail to exist, so even if $\mathbf{X}$ admits a cointegration vector, it may not give rise to a cointegration vector for $\mathbf{S}$. If $d = n$, then $\mathbf{c} := \mathcal{P}^{-\top}\mathbf{a}$ since $\mathcal{P}$ is invertible due to the full rank assumption. However, the typical situation is that $d < n$, and then the linear system $\mathcal{P}^\top \mathbf{c} = \mathbf{a}$ is over-determined and in general will not possess a solution.

We have the following convenient definition:

Definition 5. *Denote by $\mathcal{C}_{\mathbf{X}}$ the set of all cointegration vectors for $\mathbf{X}$ and $\mathcal{C}_{\mathbf{S}}$ the set of all cointegration vectors for $\mathbf{S}$.*

We note from the discussion above that if $\mathbf{c} \in \mathcal{C}_{\mathbf{S}}$, then $\mathcal{P}^\top \mathbf{c} \in \mathcal{C}_{\mathbf{X}}$. Hence, $\mathcal{P}^\top \mathcal{C}_{\mathbf{S}} \subset \mathcal{C}_{\mathbf{X}}$. For many specifications of $\mathcal{P}$, this inclusion is strict, telling that the set of cointegration vectors for $\mathbf{S}$ is restricted compared to the range of cointegration possibilities given by the vector $\mathbf{X}$. But if $\mathbf{a} \in \mathcal{C}_{\mathbf{X}}$ is in the image of $\mathcal{P}^\top$, then we have the existence of a unique $\mathbf{c} \in \mathbb{R}_0^d$ such that $\mathcal{P}^\top \mathbf{c} = \mathbf{a}$, that is, $\mathbf{c} \in \mathcal{C}_{\mathbf{S}}$. Uniqueness of $\mathbf{c}$ follows from the fact that the $n \times d$ matrix $\mathcal{P}^\top$ has full column rank d. In particular, if $\mathcal{C}_{\mathbf{X}} \subset \text{Range}(\mathcal{P}^\top)$, then $\mathcal{P}^\top \mathcal{C}_{\mathbf{S}} = \mathcal{C}_{\mathbf{X}}$.

We define a *cointegrated pricing system*:

Definition 6. *If $\mathcal{P} \in \mathbb{R}^{d\times n}$ is a pricing matrix, i.e., minimal and with rank$(\mathcal{P}) = d$, and $\mathbf{c} \in \mathbb{R}_0^d$ is such that $\mathcal{P}^\top \mathbf{c} \in \mathcal{C}_{\mathbf{X}}$, we say that $(\mathcal{P}, \mathbf{c})$ is a cointegrated pricing system for the factor process $\mathbf{X}$.*

If $(\mathcal{P}, \mathbf{c})$ is a cointegrated pricing system for the factor process $\mathbf{X}$, we can define a system of prices $\mathbf{S}(t) = \mathcal{P}\mathbf{X}(t)$ which becomes cointegrated for the vector $\mathbf{c}$, according to Definition 3. To a given $\mathbf{a} \in \mathcal{C}_{\mathbf{X}}$, there may exist many cointegrated pricing systems $(\mathcal{P}, \mathbf{c})$; indeed all possible combinations of pricing matrices $\mathcal{P}$ and vectors $\mathbf{c}$ such that $\mathcal{P}^\top \mathbf{c} = \mathbf{a}$.

Let us return to Example 1 with spot price dynamics given by (1). It is straightforward to verify that $\mathcal{P}$ is a pricing matrix and that the 2-dimensional Ornstein-Uhlenbeck process (X_1, X_2) satisfies,

$$(X_1(t), X_2(t)) \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, C)$$

where $\mathcal{N}_m$ denotes the m-variate Gaussian distribution, and the covariance matrix $C \in \mathbb{R}^{2\times 2}$ is

$$C = \begin{bmatrix} \frac{\eta_1^2}{2\alpha_1} & \rho\frac{\eta_1\eta_2}{\alpha_1+\alpha_2} \\ \rho\frac{\eta_1\eta_2}{\alpha_1+\alpha_2} & \frac{\eta_2^2}{2\alpha_2} \end{bmatrix}.$$

Here, ρ is the correlation between B_1 and B_2. Hence, since (X_1, X_2) has a limiting distribution, we find from the non-stationarity of X_3 that $\mathcal{C}_{\mathbf{X}} = \{\mathbf{a} \in \mathbb{R}_0^3 \,|\, a_3 = 0\}$. In particular, $\mathcal{C}_{\mathbf{X}} \cup \{\mathbf{0}\}$ is a vector space with basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. We remark that in general, $\mathcal{C}_{\mathbf{X}} \cup \{\mathbf{0}\}$ does not need to be a vector space. If we for example substitute X_1 and X_2 with two stationary stochastic processes which are not jointly stationary, we have that a linear combination of the two may fail to be stationary even though they are marginally stationary. We find further that $\mathcal{C}_{\mathbf{S}} \cup \{\mathbf{0}\}$ is the vector space spanned by the vector $(1, -1)^\top$. Finally, the range of $\mathcal{P}^\top$ is spanned by the two row vectors $(1, 0, 1)^\top$ and $(0, 1, 1)^\top$ of $\mathcal{P}$. Thus, if $\mathbf{a} \in \mathcal{C}_{\mathbf{X}}$, then $\mathbf{a}$ is in the range of $\mathcal{P}^\top$ *only* when $\mathbf{a} = k(1, -1, 0)^\top$ for $k \in \mathbb{R}_0$. Therefore, $\mathcal{P}^\top\mathcal{C}_{\mathbf{S}} \subset \mathcal{C}_{\mathbf{X}}$, with a strict inclusion in this case. From these considerations, we also see that there exist many pricing systems $(\mathcal{P}, \mathbf{c})$, indeed, for a fixed $\mathcal{P}$ we have a continuum of $\mathbf{c} \in \mathcal{C}_{\mathbf{S}}$. But we may also choose different pricing matrices. For example, if

$$\mathcal{P} = \begin{bmatrix} a & b & w \\ u & v & 1 \end{bmatrix}.$$

for any $a, b, u, v, w \in \mathbb{R}$ such that $\mathcal{P}$ is minimal and non-degenerate, we can use $\mathbf{c} = (1, -w)^\top$ to define a pricing system, where $\mathbf{c}^\top\mathcal{P}\mathbf{X}(t) = (a - uw)X_1(t) + (b - vw)X_2(t)$. Such a pricing matrix $\mathcal{P}$ is relevant when modelling two commodities that do not share the same denominator. For example, gas and coal typically have different energy units than power, and we will have a conversion factor (heat rate) between them modelled by w in the present context.

The particular example discussed above motivates some further analysis of the set $\mathcal{C}_{\mathbf{X}}$. In many situations, as in the example, we can single out a subset of factors from $\mathbf{X}$ which has a limit in distribution, i.e., $\mathbf{X}^m(t) := (X_1(t), \ldots, X_m(t))^\top$ with $m \leq n$ for which $P_{\mathbf{X}^m}(t, \cdot) \to \mu^m$ for a probability distribution μ^m on $\mathbb{R}^m$ as $t \to \infty$. Then $\mathcal{C}_{\mathbf{X}}^m \subset \mathcal{C}_{\mathbf{X}}$, where

$$\mathcal{C}_{\mathbf{X}}^m := \{\mathbf{a} \in \mathbb{R}_0^n \,|\, a_{m+1} = \ldots = a_n = 0\}.$$

Remark that we do not in general have equality between $\mathcal{C}_{\mathbf{X}}^m$ and $\mathcal{C}_{\mathbf{X}}$ as there may be cointegration between some of the factors $X_{m+1}, \ldots, X_n$ that may not hold jointly with the first m factors. For convenience, we have assumed that the subset of factors which has a limiting distribution consists of the first m. Since we may re-label the factors, this assumption is of course without loss of generality. We observe that $\mathcal{C}_{\mathbf{X}}^m \cup \{\mathbf{0}\}$ is a vector space, and that in the case $m = n$, we trivially have $\mathcal{C}_{\mathbf{X}}^m = \mathcal{C}_{\mathbf{X}} = \mathbb{R}_0^n$. When $m < n$, any $(\mathcal{P}, \mathbf{c})$ such that $\mathcal{P}^\top\mathbf{c} \in \mathcal{C}_{\mathbf{X}}^m$ will be a cointegrated pricing system for $\mathbf{X}$. We observe that these considerations are in line with the example above, where $\mathcal{C}_{\mathbf{X}}^2 = \mathcal{C}_{\mathbf{X}}$ since the two first factors have jointly a limiting distribution, while the last factor is non-stationary. We have the following general result:

Lemma 7. *Suppose $P_{\mathbf{X}^{n-1}}(t,\cdot)$ has a limiting distribution, while $P_{X_n}(t,\cdot)$ does not have a limiting distribution. If X_n is independent of $\mathbf{X}^{n-1}$, then $\mathcal{C}_{\mathbf{X}} = \mathcal{C}_{\mathbf{X}}^{n-1}$.*

Proof. Let $\mathbf{c} \in \mathcal{C}_{\mathbf{X}}$ with $c_n \neq 0$. By independence, we find for $z \in \mathbb{R}$

$$\begin{aligned}\Psi_{\mathbf{c}^\top \mathbf{X}}(t,z) &= \mathbb{E}[\mathrm{e}^{\mathrm{i}z\mathbf{c}^\top \mathbf{X}(t)}] \\ &= \mathbb{E}[\mathrm{e}^{\mathrm{i}z(c_1X_1(t)+\cdots+c_{n-1}X_{n-1}(t))}]\mathbb{E}[\mathrm{e}^{\mathrm{i}zc_nX_n(t)}] \\ &= \Psi_{\mathbf{X}^{n-1}}(t,z(c_1,\ldots,c_{n-1})^\top)\Psi_{X_n}(t,zc_n)\,.\end{aligned}$$

For every z, $\Psi_{\mathbf{X}^{n-1}}(t,z(c_1,\ldots,c_{n-1})^\top)$ will have a limit, while there exists a Borel set A_0 with positive Lebesgue measure such that $\Psi_{X_n}(t,x)$ does not have a limit for every $x \in A_0$ (this could be the whole of the real line, or some subset with infinite Lebesgue measure). But then for all $z \in A_0/c_n$ we have that $\Psi_{X_n}(t,zc_n)$ does not have a limit, and in conclusion $\Psi_{\mathbf{c}^\top\mathbf{X}}(t,z)$ does not have a limit for every $z \in \mathbb{R}$ as $t \to \infty$. This violates the assumption that $\mathbf{c} \in \mathcal{C}_{\mathbf{X}}$ with $c_n \neq 0$. Thus, $c_n = 0$, showing the claim. □

Remark that in Example 1, the non-stationary drifted Brownian motion is not necessarily independent of the two other factors, showing that the assumption of independence is sufficient, but not necessary.

Notice that if $\mathbf{X}^m$ admits a stationary limit, and $(X_{m+1},...,X_n)$ is dependent on X^m, we may have non-trivial $\mathbf{c} \in \mathcal{C}_{\mathbf{X}} \backslash \mathcal{C}_{\mathbf{X}}^m$. Indeed, consider $n = 3$ and the processes

$$X_1(t) = \int_0^t \exp(-\alpha_1(t-s))\,dB_1(s)\,,$$

and

$$X_i(t) = \mu t + \int_0^t \exp(-\alpha_i(t-s))\,dB_i(s), i = 2,3\,,$$

for constants $\mu, \alpha_i > 0, i = 2,3$ and a trivariate Brownian motion (B_1,B_2,B_3) being correlated. Then $\mathbf{X} = (X_1,X_2,X_3)^\top$ have dependent coordinates, and for any vector $\mathbf{c} = (a,b,-b)^\top \in \mathbb{R}_0^3, a,b \in \mathbb{R}$, we find that

$$\mathbf{c}^\top\mathbf{X}(t) = a\int_0^t \mathrm{e}^{-\alpha_1(t-s)}\,dB_1(s) + b\left(\int_0^t \mathrm{e}^{-\alpha_2(t-s)}\,dB_2(s) - \int_0^t \mathrm{e}^{-\alpha_3(t-s)}\,dB_3(s)\right)$$

which will converge in distribution to a normally distributed random variable with zero mean as $t \to \infty$. Hence, $\mathbf{c} \in \mathcal{C}_{\mathbf{X}}$. Here, X_1 has a limit in distribution, while $X_i, i = 2,3$ both will have a mean μt and thus there does not exist any limiting distribution. This is an example with $m = 1$ and $n = 3$. We remark that the example is slightly pathological, as we could have assumed $n = 4$ with $X_4(t) = \mu t$, and defined $\widetilde{X}_i(t) = \int_0^t \exp(-\alpha_i(t-s))\,dB_i(s), i = 2,3$. Then, with $\mathbf{X} := (X_1,\widetilde{X}_2,\widetilde{X}_3,X_4)^\top$ we are back to the situation with $m = n-1 = 3$ and X_4 being (trivially) independent of $\mathbf{X}^3 = (X_1,\widetilde{X}_2,\widetilde{X}_3)^\top$.

We have the following remark, which gives a practical consequence of our considerations so far:

Remark 8. *In a practical application we can model a system of d commodity price dynamics with cointegration as follows: first, we assume that we have m factor processes which jointly admit a limiting distribution, and $n-m$ non-stationary* *processes, with $n \geq d$. Then we know that any $(\mathcal{P}, \mathbf{c})$ such that $\mathcal{P}^\top \mathbf{c} \in \mathcal{C}_{\mathbf{X}}^m$ will be a cointegrated pricing system. This provides us with a constraint on the possible specifications of $(\mathcal{P}, \mathbf{c})$ which can be used in the next step on specifying parametric models for the factor processes and estimating on data. As long as we know that $\mathbf{X}^m$ admits a limiting distribution, we can characterize a set of admissible pricing systems $(\mathcal{P}, \mathbf{c})$ before any further specification and estimation on data.*

In the anaysis so far we have exclusively thought of the price dynamics $\mathbf{S}$ in (3) as being on an arithmetic form. However, commonly one models cointegration on the logarithm of prices, $\ln \mathbf{S} := (\ln S_1, \ldots, \ln S_d)^\top$. If we suppose that $\ln \mathbf{S}$ satisfies

$$\ln \mathbf{S}(t) = \mathcal{P}\mathbf{X}(t),\, t \geq 0, \tag{4}$$

we can repeat the analysis above for a *geometric* price dynamics. Energy markets like gas and power have frequently experienced negative prices, and hence an arithmetic price dynamics may be attractive.

2.1. **Particular model specifications.** Recalling Example 1, we may for the cross-commodity spot price dynamics (1) assume a general (non-stationary) dynamics X_3 and a bivariate process $\mathbf{Y} := (X_1, X_2)^\top$ with the property that $P_{\mathbf{Y}}(t, \cdot) \to P_\infty(\cdot)$ for some probability distribution P_∞ on $\mathbb{R}^2$. Then, we find for any $\mathbf{c} = (k, -k)^\top \in \mathcal{C}_{\mathbf{S}}$ that the characteristic function of the random variable $\mathbf{c}^\top \mathcal{P}\mathbf{X}(t)$ is

$$\Psi_{\mathbf{c}^\top \mathcal{P}\mathbf{X}}(t, z) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}zk(X_1(t) - X_2(t))}\right] = \Psi_{\mathbf{X}}(t, z\mathcal{P}^\top \mathbf{c}) = \widehat{P}_{\mathbf{Y}}(t, z\mathbf{c}^\top) \to \widehat{P}_\infty(z\mathbf{c}^\top)\,,$$

for every $z \in \mathbb{R}$ as $t \to \infty$. But then by Proposition 4, $\mathbf{c}^\top \mathcal{P}\mathbf{X}(t)$ has a limiting distribution $\mu_{\mathbf{c}}$ with characteristic function $\widehat{\mu}_{\mathbf{c}}(z) = \widehat{P}_\infty(z\mathbf{c}^\top)$. This shows that we may significantly go beyond the dynamics discussed in (1) that preserves cointegration, and has a marginal structure with a (long term) non-stationary factor and a (short term) factor modelling the "stationary" variations. In this Subsection we discuss various other particular specifications of these factors beyond classical Ornstein-Uhlenbeck models.

The so-called Lévy stationary (LS) processes provides us with a flexible class of stationary models which can be applied as dynamics for $\mathbf{X}^m = (X_1, \ldots, X_m^\top)$, $m \in \mathbb{N}$. To this end, assume

$$\mathbf{X}^m(t) = \int_{-\infty}^t G(t-s)\, d\mathbf{L}(s), \tag{5}$$

where $\mathbf{L} = (L_1, \ldots, L_k)^\top$ is a two-sided square integrable k-dimensional Lévy process with zero mean and $u \mapsto G(u)$ is a measurable mapping from $\mathbb{R}_+$ into the space of $m \times k$ matrices with elements $g_{ij} \in L^2(\mathbb{R}_+), i = 1, \ldots, m, j = 1, \ldots, k$. We remark that the assumption on the g_{ij}'s ensures that $\mathbf{X}^m$ is a well-defined mean zero square integrable stochastic process with values in $\mathbb{R}^m$. LS processes form a subclass of the more general Lévy semistationary processes considered in e.g. Barndorff-Nielsen, Benth and Veraart [5].

As the following Lemma shows, $\mathbf{X}^m = (X_1, \ldots, X_m)^\top$ is strictly stationary:

Lemma 9. *The process $\mathbf{X}^m = (X_1, \ldots, X_m)^\top$ defined in (5) is a strictly stationary process, that is, for any $\tau \geq 0, r \in \mathbb{N}$, and $0 \leq t_1 < \cdots < t_t < \infty$, the $m \times r$-dimensional random matrices $(\mathbf{X}^m(t_1+\tau), ..., \mathbf{X}^m(t_r+\tau))$ and $(\mathbf{X}^m(t_1), ..., \mathbf{X}^m(t_r))$ have the same probability distribution.*

Proof. Let $t_\ell, \ell = 1, \dots, r$ be an increasing sequence of times on $\mathbb{R}$, and notice that $(\mathbf{X}^m(t_1), \mathbf{X}^m(t_2), \dots, \mathbf{X}^m(t_r))^\top \in \mathbb{R}^{mr}$. By the independent increment property of Lévy processes, it follows that with ψ denoting the cumulant of $\mathbf{L}(1)$ and $\mathbf{z}^\top = ((z_1^1, \dots, z_1^m), (z_2^1, \dots, z_2^m), \dots, (z_r^1, \dots, z_r^m)) \in \mathbb{R}^{mr}$,

$$
\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\mathbf{z}^\top (\mathbf{X}^m(t_1),\dots,\mathbf{X}^m(t_r))^\top}\right] &= \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\mathbf{z}_1^\top \int_{-\infty}^{t_1} G(t_1-s)\,d\mathbf{L}(s)+\cdots \mathrm{i}\mathbf{z}_r^\top \int_{-\infty}^{t_1} G(t_r-s)\,d\mathbf{L}(s)}\right] \\
&\quad \times \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\mathbf{z}_2^\top \int_{t_1}^{t_2} G(t_2-s)\,d\mathbf{L}(s)+\cdots \mathrm{i}\mathbf{z}_r^\top \int_{t_1}^{t_2} G(t_r-s)\,d\mathbf{L}(s)}\right] \\
&\quad \times \cdots \times \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\mathbf{z}_r^\top \int_{t_{r-1}}^{t_r} G(t_r-s)\,d\mathbf{L}(s)}\right] \\
&= \exp\left(\int_{-\infty}^{t_1} \psi\left(G(t_1-s)^\top \mathbf{z}_1 + \cdots + G(t_r-s)^\top \mathbf{z}_r\right)\,ds\right) \\
&\quad \times \exp\left(\int_{t_1}^{t_2} \psi\left(G(t_2-s)^\top \mathbf{z}_2 + \cdots + G(t_r-s)^\top \mathbf{z}_r\right)\,ds\right) \\
&\quad \times \cdots \times \exp\left(\int_{t_{r-1}}^{t_r} \psi(G(t_r-s)^\top \mathbf{z}_r)\,ds\right).
\end{aligned}
$$

Here we have used the notation $\mathbf{z}_i := (z_i^1, \dots, z_i^m)^\top$. Thus, after a change of variables, we see that the characteristic function of $(\mathbf{X}^m(t_1), ..., \mathbf{X}^m(t_r))$ depends on $(t_2 - t_1, t_3 - t_2, ..., t_r - t_{r-1})$ only, and we can conclude that the probability distribution of $(\mathbf{X}^m(t_1+\tau), ..., \mathbf{X}^m(t_r+\tau))$ equals that of $(\mathbf{X}^m(t_1), ..., \mathbf{X}^m(t_r))$ for any $\tau > 0$. Strict stationarity follows. □

Remark that any linear combination of $X_1, \dots, X_m$ is strictly stationary whenever $(X_1, \dots, X_m)$ is strictly stationary. If the real-valued process $\{U(t)\}_{t\geq 0}$ is a strictly stationary process, we have that its probability distribution $P_U(t, \cdot)$ satisfies $P_U(t+\tau, \cdot) = P_U(t, \cdot)$ for all $t \geq 0$, for any given $\tau \geq 0$. Hence, $P_U(t, \cdot) \equiv P_U(\cdot)$, that is, it is independent of time t. This implies trivially that $P_U(t, \cdot) \to P_U(\cdot)$ when $t \to \infty$, and moreover, the characteristic function of $U(t)$ is also independent of t. Hence, for any pricing system $(\mathcal{P}, \mathbf{c})$, where $\mathcal{P}^\top \mathbf{c} \in \mathcal{C}_{\mathbf{X}}^m$, we have that $\mathbf{c}^\top \mathcal{P}\mathbf{X}(t) = (\mathbf{e}_1^\top \mathcal{P}^\top \mathbf{c})X_1(t) + \cdots + (\mathbf{e}_m^\top \mathcal{P}^\top \mathbf{c})X_m(t)$, i.e., a linear combination of $X_1, \dots, X_m$, which is a strictly stationary process. We note in passing that if the characteristic function of a stochastic process $V(t)$ is independent of t, it holds that $P_V(t, \cdot) = P_V(\cdot)$. This implies stationarity in the sense that the probability distribution of $V(t)$ is invariant of time, however, it does not necessarily imply *strict* stationarity. In Benth [6], cointegration models based on LS processes were proposed and analysed.

An example of an LS process is given by

$$X_i(t) = \eta_i \int_{-\infty}^t \mathrm{e}^{-\alpha_i(t-s)}\,dB_i(s),$$

for $i = 1, \dots, m$ with $\mathbf{B} = (B_1, \dots, B_m)^\top$ being a two-sided m-dimensional Brownian motion (possibly correlated) and $\alpha_1, \dots, \alpha_m$, $\eta_1, \dots, \eta_m$ positive constants. Then it holds that the distribution function of $(X_1, \dots, X_m)$ is time invariant and equal to $\mathcal{N}_m(\mathbf{0}, C)$, with covariance matrix $C \in R^{m\times m}$ having diagonal elements $\eta_i^2/(2\alpha_i), i = 1, \dots, m$ and off-diagonal elements $\rho_{ij}\eta_i\eta_j/(\alpha_i + \alpha_j)$ for ρ_{ij} being the correlation coefficient between B_i and B_j, $i \neq j$. In fact, this example is a particular case of so-called continuous-time autoregressive moving average (CARMA) processes, as we discuss next.

For $p \in \mathbb{N}$, define the matrix $A \in \mathbb{R}^{p \times p}$ as

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \because & \cdot & \cdot & \cdot & \dots & \cdot \\ \because & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & 0 & \dots & 1 \\ -\alpha_p & -\alpha_{p-1} & -\alpha_{p-2} & -\alpha_{p-3} & \dots & -\alpha_1 \end{bmatrix}, \tag{6}$$

for positive constants α_k, $k = 1, \dots, p$. Consider the p-dimensional Ornstein-Uhlenbeck process

$$d\mathbf{Y}(t) = A\mathbf{Y}(t)\, dt + \mathbf{e}_p\, dL(t). \tag{7}$$

Here we recall that $\{\mathbf{e}_i\}_{i=1}^p$ are the p canonical basis vectors in $\mathbb{R}^p$ and L is a (two-sided) real-valued square-integrable Lévy process with zero mean. Following Brockwell [14], we define a CARMA(p, q) process Z for $q < p, p, q \in \mathbb{N}$ by

$$Z(t) = \mathbf{b}^\top \mathbf{Y}(t), \tag{8}$$

for $\mathbf{b} \in \mathbb{R}^p$, where $\mathbf{b} = (b_0, b_1, \dots, b_q, 0, \dots, 0)^\top$ and $b_q = 1$. We observe that for $q = 0$, $\mathbf{b} = \mathbf{e}_1$ and we say in this case that Z is a continuous-time autoregressive process of order p (a CAR(p)-process in short). We suppose that the p eigenvalues of A have negative real part, which yields that Z is strictly stationary with

$$Z(t) = \int_{-\infty}^t \mathbf{b}^\top \mathrm{e}^{A(t-s)} \mathbf{e}_p\, dL(s).$$

Thus, with $G(s) := \mathbf{b}^\top \exp(A(s))\mathbf{e}_p$, a CARMA$(p, q)$-process is an example of a real-valued LS-process.

Indeed, CARMA-processes constitute a rich class of models that can play the role as factors in a cointegration model (see Comte [15] **for an extensive analysis of cointegration based on CARMA processes). Let us discuss this a bit closer, relating to a large body of literature where such models have been applied in a commodity market context.** To this end, let $\mathbf{X}$ be an n-dimensional process, and $\mathbf{X}^m = (X_1, \dots, X_m)$ for $m < n$ be an m-dimensional CARMA-process. A simple way to define such a process is as follows: given an m-dimensional two-sided square integrable Lévy process $\mathbf{L} = (L_1, \dots, L_m)^\top$ with zero mean. For $i = 1, \dots, m$, let X_i be as in (8), that is, a CARMA(p_i, q_i)-process driven by L_i and with matrix $A_i \in \mathbb{R}^{p_i \times p_i}$ having eigenvalues with negative real part. In the notation of LS-processes in (5), this means that the $m \times m$-matrix-valued function $G(u)$ has diagonal elements $g_{ii}(u) := \mathbf{b}_i^\top \exp(A_i(s))\mathbf{e}_{p_i}$ and off-diagonal elements being zero. By Lemma 9, $\mathbf{X}^m$ is an m-dimensional strictly stationary process. Benth and Koekebakker [10] consider such models in the context of cointegration. We remark in passing that CARMA-processes have been applied to model commodities like oil and power (see e.g. Paschke and Prokopczuk [36] and Benth, Klüppelberg, Müller and Vos [9]). Multivariate CARMA-processes going beyond the simple specification we consider here have been proposed and analysed by Marquardt and Stelzer [33]. Their definition will yield an LS-process (5) with the matrix-valued function G having non-zero off-diagonal elements. Thus, we do not only have dependency through the Lévy processes, but also functional dependencies between the coordinates in vector-valued CARMA process. Such multivariate CARMA processes is further studied by Schlemm and Stelzer [39] and Kevei [31]. Taking these extensions into account, we have a rich class of stationary processes available for cointegration modelling.

It is well-known (see e.g. Benth and Šaltytė Benth [8] and Benth, Šaltytė Benth and Koekebakker [7]) that a CARMA(p,q)-process on a discrete time scale will define an ARMA(p,q) time series. Furthermore, as is demonstrated in Aadland, Benth and Koekebakker [1], the process $X(t)=\int_0^t Z(s)\,ds$, where Z is a CAR(p)-process, becomes a non-stationary process. Hence, it may serve as a non-stationary factor process in modelling the price dynamics $\mathbf{S}$. Indeed, we can use a set of $n-m$ dependent CAR-processes to define non-stationary processes $X_{m+1},\dots,X_n$ in this way. As Aadland, Benth and Koekebakker [1] show, these processes will become integrated autoregressive times series on a discrete time scale. Aadland, Benth and Koekebakker [1] model cointegration in a freight rate market using CAR-processes, both for the stationary and the non-stationary processes.

3. Forward pricing under cointegration

Denote the forward prices at time $t\geq 0$ of contracts delivering the underlying assets $\mathbf{S}=(S_1,\dots,S_d)^\top$ at time $T\geq t$ by $\mathbf{F}(t,T):=(F_1(t,T),\dots,F_d(t,T))^\top\in\mathbb{R}^d$. The price vector of the d assets are defined by $\mathbf{S}(t)\in\mathbb{R}^d$ in (3) with $\mathcal{P}$ being minimal and of full rank. Thus, we suppose an arithmetic model for the spot market. Assume $\mathbb{Q}\sim\mathbb{P}$ is a pricing measure such that $\mathbf{X}(t)\in\mathbb{R}^n$ is $\mathbb{Q}$-integrable for all $t>0$. Then, the forward price vector $\mathbf{F}(t,T)$ is defined as (see Benth, Šaltytė Benth and Koekebakker [7]),

$$\mathbf{F}(t,T)=\mathbb{E}_{\mathbb{Q}}[\mathbf{S}(T)\,|\,\mathcal{F}_t]. \tag{9}$$

Hence, by the definition of $\mathbf{S}$ we find that

$$\mathbf{F}(t,T)=\mathcal{P}\mathbb{E}_{\mathbb{Q}}[\mathbf{X}(T)\,|\,\mathcal{F}_t]\,. \tag{10}$$

To proceed with our analysis, the **introduce the following convenient property of X:**

Definition 10. *The stochastic process $\{\mathbf{X}(t)\}_{t\geq 0}$ is said to be linearity-generating with respect to $\mathbb{Q}$, or $\mathbb{Q}$-LG for short, if there exist measurable deterministic functions $(t,T)\mapsto\mathcal{A}(t,T)\in\mathbb{R}^{n\times n}$ and $(t,T)\mapsto\mathbf{a}(t,T)\in\mathbb{R}^d$ such that*

$$\mathbb{E}_{\mathbb{Q}}[\mathbf{X}(T)\,|\,\mathcal{F}_t]=\mathcal{A}(t,T)\mathbf{X}(t)+\mathbf{a}(t,T)$$

for $0\leq t\leq T<\infty$.

We remark that our definition of X being linearity-generating with respect to $\mathbb{Q}$ is inspired by Gabaix [26] (see also Filipović, Larsson and Trolle [25] for an extensive analysis of linearity-generating processes). In Gabaix [26], linearity-generating processes are defined tailormade to applications in interest rate theory, whereas in our context the above definition of $\mathbb{Q}$-LG processes are convenient.

A trivial example of a $\mathbb{Q}$-LG process $\mathbf{X}$ is the case when $\mathbf{X}$ is an n-dimensional $\mathbb{Q}$-Brownian motion $\mathbf{B}=(B_1,\dots,B_n)^\top$. Then $\mathbf{a}=0$, and $\mathcal{A}$ is the covariance matrix with elements $\rho_{ij}t$ for $\rho_{ii} = 1$ and ρ_{ij} being the correlation between B_i and B_j, $i \neq j$. A less trivial example is provided by Ornstein-Uhlenbeck processes. We show next the linearity-generating property for $\mathbb{Q}$-semimartingales which are polynomial processes (see e.g. Cuchiero, Keller-Ressel and Teichmann [16] and Filipović and Larsson [24] for a definition and analysis of polynomial processes).

Proposition 11. *Assume the $\mathbb{Q}$-dynamics of $\mathbf{X}$ is a polynomial process in $\mathbb{R}^n$. Then $\mathbf{X}$ is $\mathbb{Q}$-integrable and $\mathbb{Q}$-LG, with the functions $(t,T) \mapsto \mathbf{a}(t,T)$ and $(t,T) \mapsto \mathcal{A}(t,T)$ being homogeneous, i.e., $\mathcal{A}(t,T) = \mathcal{A}(T-t)$ and $\mathbf{a}(t,T) = \mathbf{a}(T-t)$ (with a slight abuse of notation).*

Proof. A polynomial process has finite moments (Lemma 2.17 in Cuchiero, Keller-Ressel and Teichmann [16]), and thus $\mathbf{X}$ is $\mathbb{Q}$-integrable. Following the definition of a polynomial process (see e.g. Cuchiero, Keller-Ressel and Teichmann [16] and Filipović and Larsson [24]), we know that for the generator $\mathcal{G}$ of $\mathbf{X}$, there exists a matrix $G \in \mathbb{R}^{n\times n}$ and a vector $\mathbf{b} \in \mathbb{R}^n$ such that $\mathcal{G}\mathbf{x} = G\mathbf{x}+\mathbf{b}$. This holds true since $\mathbf{x}$ is a first order polynomial and the generator is preserving the order when applied to polynomials. Therefore, from the martingale problem of polynomial processes,

$$\mathbb{E}_{\mathbb{Q}}[\mathbf{X}(T)\,|\,\mathcal{F}_t] = \mathbf{X}(t) + \int_t^T \big(G\mathbb{E}_{\mathbb{Q}}[\mathbf{X}(s)\,|\,\mathcal{F}_t] + \mathbf{b}\big)\,ds.$$

and thus,

$$\mathbb{E}_{\mathbb{Q}}[\mathbf{X}(T)\,|\,\mathcal{F}_t] = \mathrm{e}^{G(T-t)}\mathbf{X}(t) + \int_t^T \mathrm{e}^{G(T-s)}\mathbf{b}\,ds.$$

The result follows. □

We remark in passing that the class of polynomial processes has a much richer structure than really needed for the linearity-generating property with respect to $\mathbb{Q}$. The generator of a polynomial process preserves the order of any polynomial, while to be $\mathbb{Q}$-LG only requires that the generator preserves the first order polynomials.

As an example, consider an Ornstein-Uhlenbeck process in $\mathbb{R}^n$ with $\mathbb{Q}$-dynamics

$$d\mathbf{X}(t) = (\boldsymbol{\mu} + C\mathbf{X}(t))\,dt + \Sigma\,d\mathbf{W}(t).$$

Here, $\boldsymbol{\mu} \in \mathbb{R}^n$, $C \in \mathbb{R}^{n\times n}$, $\Sigma \in \mathbb{R}^{n\times m}$ and $\mathbf{W}$ is an m-dimensional Brownian motion. A direct calculation reveals that for $t \leq T$,

$$\mathbf{X}(T) = \mathrm{e}^{C(T-t)}\mathbf{X}(t) + \int_t^T \mathrm{e}^{C(T-s)}\boldsymbol{\mu}\,ds + \int_t^T \mathrm{e}^{C(T-s)}\Sigma\,d\mathbf{W}(s),$$

and thus

$$\mathbb{E}_{\mathbb{Q}}[\mathbf{X}(T)\,|\,\mathcal{F}_t] = \mathrm{e}^{C(T-t)}\mathbf{X}(t) + \int_0^{T-t} \mathrm{e}^{Cs}\boldsymbol{\mu}\,ds.$$

In conclusion, $\mathbb{Q}$-LG property holds with $\mathcal{A}(T-t) = \exp(C(T-t))$ and $\mathbf{a}(T-t) = \int_0^{T-t}\exp(Cs)\boldsymbol{\mu}\,ds$. Note that both $\mathcal{A}$ and $\mathbf{a}$ are homogeneous in time. Whenever C is an invertible matrix, we find

$$\mathbf{a}(T-t) = C^{-1}(\mathrm{e}^{C(T-t)} - I)\boldsymbol{\mu}$$

where $I \in \mathbb{R}^{n\times n}$ is the identity matrix.

We have the following simple result:

Corollary 12. *If $\{\mathbf{X}(t)\}_{t\geq 0}$ is $\mathbb{Q}$-integrable and $\mathbb{Q}$-LG process in $\mathbb{R}^n$, then*
$$\mathbf{F}(t,T) = \mathcal{P}\mathcal{A}(t,T)\mathbf{X}(t) + \mathcal{P}\mathbf{a}(t,T)\,.$$
Moreover, if $\mathcal{P}\mathcal{A}(t,T) = \widetilde{\mathcal{A}}(t,T)\mathcal{P}$ for some $\widetilde{\mathcal{A}}(t,T) \in \mathbb{R}^{d\times d}$, $0 \leq t \leq T < \infty$, then $\mathbf{F}(t,T) = \widetilde{\mathcal{A}}(t,T)\mathbf{S}(t) + \mathcal{P}\mathbf{a}(t,T)$ (i.e., the forward price vector is linearity-generating in the asset price $\mathbf{S}$.)

Proof. This is trivial from the definition of linearity-generating processes. □

We note that in the case $d = n$, we have forward prices which are linearity-generating in the underlying spot when $\mathcal{P}$ and $\mathcal{A}(t,T)$ commutes for all $t \leq T$.

Let us next turn to the question of cointegration in the forward market. As $t \leq T < \infty$, it is natural to switch to the Musiela parametrization, and express forward prices in terms of *time to maturity* $x := T - t$ rather than *time of maturity* T. I.e., introduce the random fields $\mathbf{f}(t,x)$ for $x \geq 0$ by
$$\mathbf{f}(t,x) := \mathbf{F}(t,t+x)\,. \tag{11}$$
Hence, we find in the case of $\{\mathbf{X}(t)\}_{t\geq 0}$ being $\mathbb{Q}$-integrable and $\mathbb{Q}$-LG that
$$\mathbf{f}(t,x) = \mathcal{P}\mathcal{A}(t,t+x)\mathbf{X}(t) + \mathbf{a}(t,t+x)\,.$$
The following Proposition holds:

Proposition 13. *Fix $x \geq 0$, and suppose that $\mathbf{X}$ is $\mathbb{Q}$-integrable and $\mathbb{Q}$-LG, with $\mathcal{A}$ and $\mathbf{a}$ homogeneous (e.g., $\mathcal{A}(t,T) = \mathcal{A}(T-t)$ and $\mathbf{a}(t,T) = \mathbf{a}(T-t)$). Then $t \mapsto \mathbf{f}(t,x)$ is cointegrated if there exists a vector $\mathbf{c} \in \mathbb{R}_0^d$ such that $\mathbf{c}^\top\mathcal{P}\mathcal{A}(x) \in \mathcal{C}_{\mathbf{X}}$, or, equivalently, $(\mathcal{P}\mathcal{A}(x),\mathbf{c})$ is a cointegrated pricing system.*

Proof. This follows readily from the definitions and the fact that for homogeneous $\mathcal{A}$ and $\mathbf{a}$, $\mathcal{A}(t,t+x) = \mathcal{A}(x)$ and $\mathbf{a}(t,t+x) = \mathbf{a}(x)$. □

Remark 14. *We emphasise that $x \geq 0$ is fixed in Proposition 13. This means that it is the dynamics of the forward contracts with fixed time to maturity that is cointegrated. This can be viewed as a roll-over contract, where one fixes the time to maturity and "rolls over" the position when time progresses. The actual forward price dynamics will in general not be cointegrated as it will depend on $\mathcal{A}(t,T)$ and $\mathbf{a}(t,T)$, which varies with time t. Benth and Koekebakker [10] make a similar observation for a more particular HJM-type cointegrated forward price model. If $x = 0$, or equivalently $t = T$, we are back to the spot price case. Propositions 13 and 11 show that polynomial processes can be used to build cointegrated forward price models.*

Consider the case when the $\mathbb{Q}$-dynamics of $\mathbf{X}(t) \in \mathbb{R}^3$ is such that $X_3(t)$ is a non-stationary process and (X_1, X_2) admits a limiting distribution. From previous considerations **(see discussion following Definition 6)** we then have that $\mathcal{C}_{\mathbf{X}} = \{\mathbf{a} \in \mathbb{R}_0^3 \,|\, a_3 = 0\}$. In the context of Example 1, for any pricing matrix $\mathcal{P} \in \mathbb{R}^{2\times 3}$ and $\mathbf{c} \in \mathbb{R}_0^2$, we find that $\mathbf{c}^\top\mathcal{P} \in \mathcal{C}_{\mathbf{X}}$ if and only if $\mathbf{c}^\top\mathcal{P}\mathbf{e}_3 = 0$ (e.g, the third coordinate of $\mathbf{c}^\top\mathcal{P}$ is equal to zero). With p_{ij} denoting the ijth element of $\mathcal{P}$, we find that $\mathbf{c}^\top\mathcal{P} \in \mathcal{C}_{\mathbf{X}}$ if and only if $c_1p_{13} + c_2p_{23} = 0$. Minimality of $\mathcal{P}$ means that $\mathcal{P}\mathbf{e}_i \neq (0,0)^\top$ for $i = 1,2,3$, and in particular for $i = 3$ we find $(p_{13},p_{23}) \neq (0,0)$. Thus, we find that $(\mathcal{P},\mathbf{c})$ is a cointegrated pricing system if and only if either $c_2, p_{13} \neq 0$ and $c_1/c_2 = -p_{23}/p_{13}$ or $c_1, p_{23} \neq 0$ and $c_2/c_1 = -p_{13}/p_{23}$. If $\mathbf{X}$ is $\mathbb{Q}$-LG with a matrix $\mathcal{A}(t,T) = \mathcal{A}(T-t) \in \mathbb{R}^{3\times 3}$ satisfying $\mathbf{e}_1^\top\mathcal{A}(x)\mathbf{e}_3 = \mathbf{e}_2^\top\mathcal{A}(x)\mathbf{e}_3 = 0$, then $\mathbf{c}^\top\mathcal{P}\mathcal{A}(x) \in \mathcal{C}_{\mathbf{X}}$ for any cointegrated pricing system $(\mathcal{P},\mathbf{c})$.

As a particular case of the above, consider the factor process

$$d\mathbf{X}(t) = \left(\boldsymbol{\mu} + \begin{bmatrix} C & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} \mathbf{X}(t)\right) dt + \Sigma\, d\mathbf{W}(t) \tag{12}$$

with $\Sigma, C \in \mathbb{R}^{2\times 2}$, $\boldsymbol{\mu} \in \mathbb{R}^3$ and $\mathbf{0} = (0,0)^\top$. Further, $\mathbf{W}$ is assumed to be a trivariate $\mathbb{Q}$-Brownian motion. Here, (X_1, X_2) will be a bivariate OU process with mean-reversion matrix C and noise vector $(\mathbf{e}_1^\top \Sigma\, d\mathbf{W}(t), \mathbf{e}_2^\top \Sigma\, d\mathbf{W}(t))^\top$, which admits a limiting distribution whenever C has eigenvalues with negative real part. The process X_3 is a drifted Brownian motion. Then,

$$\mathbb{E}_{\mathbb{Q}}[\mathbf{X}(T)\,|\,\mathcal{F}_t] = \mathcal{A}(T-t)\mathbf{X}(t) + \mathbf{a}(T-t),$$

where

$$\mathcal{A}(T-t) = \begin{bmatrix} e^{C(T-t)} & \mathbf{0} \\ \mathbf{0}' & 1 \end{bmatrix}, \tag{13}$$

and $\mathbf{a}(T-t) = \int_0^{T-t} \mathcal{A}(y)\boldsymbol{\mu}\, dy$. Thus, for $(\mathcal{P}, \mathbf{c})$ being a cointegrated pricing system, the forward prices will also be cointegrated. We see that we obtain cointegration both for the spot (under $\mathbb{Q}$) and the forward prices for rather general models of $\mathbf{X}$, including a full corrrelation structure between the three noises $\mathbf{W}$ and flexible mean reversion matrix C.

Note that if $\mathcal{A}$ is not homogeneous, that is, $\mathbf{X}$ is $\mathbb{Q}$-LG for a non-homogeneous $\mathcal{A}$, then we may lose cointegration in the forward process. In the case $\mathbf{a}$ is non-homogeneous, we may recover cointegration as long as $\mathcal{A}$ is homogeneous by considering the "de-trended" forward price vector $\bar{\mathbf{f}}(t,x) := \mathbf{f}(t,x) - \mathcal{P}\mathbf{a}(t,x)$. In that case, $\bar{\mathbf{f}}(t,x)$ is cointegrated whenever $\mathbf{c}^\top \mathcal{P}\mathcal{A}(x) \in \mathcal{C}_{\mathbf{X}}$. Indeed, this is a relevant case for commodity markets with seasonally varying prices. For example, in power markets, where prices are highly influenced by weather conditions, it may appear that $\mathbf{a}$ is not homogeneous. Indeed, the factor model (12) can be used as a model for spot prices with $\boldsymbol{\mu}$ being time dependent, i.e. $t \mapsto \boldsymbol{\mu}(t)$ for some measurable real-valued function being bounded on compacts. Then $\mathbf{a}(t,T) = \int_t^T \mathcal{A}(T-s)\boldsymbol{\mu}(s)\, ds$ is not in general homogeneous. Typically, $\boldsymbol{\mu}(t)$ models a seasonal mean price, towards which the stationary part of $\mathbf{X}$ mean reverts (see e.g. Benth, Šaltyte Benth and Koekebakker [7] for models of this type with seasonality).

3.1. **General LS-processes.** In general, LS-process will not be $\mathbb{Q}$-LG. In this subsection we analyse forward pricing based on LS-processes.

For $x \geq 0$, we define the $\mathbb{R}^m$-valued random field $\widetilde{\mathbf{X}}^m(t,x)$ by

$$\widetilde{\mathbf{X}}^m(t,x) := \int_{-\infty}^t G(t-s+x)\, d\mathbf{L}(s), \tag{14}$$

with G and $\mathbf{L}$ being as in the definition of the LS-process in (5). We assume that this is the $\mathbb{Q}$-dynamics of $\widetilde{\mathbf{X}}^m(t,x)$. In particular, for $x = 0$, we are back to $\mathbf{X}^m(t)$ as in (5) (but now considered as a dynamics with respect to $\mathbb{Q}$). Moreover, following the proof of Lemma 9, the stochastic process $t \mapsto \widetilde{\mathbf{X}}^m(t,x)$ is strictly stationary for every $x \geq 0$. It is simple to see that

$$\mathbb{E}_{\mathbb{Q}}[\mathbf{X}^m(T)\,|\,\mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}[\widetilde{\mathbf{X}}^m(T,0)\,|\,\mathcal{F}_t] = \widetilde{\mathbf{X}}^m(t, T-t), \tag{15}$$

by appealing to the independent increment property of Lévy processes. Hence, assuming a factor process $\mathbf{X}$ which is $\mathbb{Q}$-integrable, where $\mathbf{X}^m$ is given by an LS-process as in (5) with respect to the probability $\mathbb{Q}$, we find that

(16)
$$\mathbf{f}(t,x) = \mathcal{P}\left(\widetilde{X}_1(t,x),\ldots,\widetilde{X}_m(t,x),\mathbb{E}_{\mathbb{Q}}[X_{m+1}(T)\,|\,\mathcal{F}_t],\ldots,\mathbb{E}_{\mathbb{Q}}[X_n(T)\,|\,\mathcal{F}_t]\right)^\top .$$

We see that any $\mathbf{c} \in \mathbb{R}^d_0$ such that $\mathbf{c}^\top\mathcal{P} \in \mathcal{C}^m_{\mathbf{X}}$ implies that $\mathbf{c}^\top\mathbf{f}(t,x)$ becomes a linear combination of $\widetilde{X}_1(t,x),\ldots,\widetilde{X}_m(t,x)$, and therefore strictly stationary. Hence, $\mathbf{c}$ will be a cointegration vector for $\mathbf{f}(t,x)$.

The classes of CARMA-processes and their multivariate extensions discussed in the previous section provide a rich class of LS-processes that can be used for modelling cointegrated forward prices under the Musiela parametrization.

3.2. **Factor models of geometric type.** Classically, pricing models in finance have been geometric. In our context, we recall from (4) that this means a spot price dynamics $\mathbf{S}$ of the form $\ln\mathbf{S}(t) = \mathcal{P}\mathbf{X}(t)$. The forward price vector $\mathbf{F}(t,T) = (F_1(t,T),\ldots,F_d(t,T))^\top$ will be given by

$$F_i(t,T) = \mathbb{E}_{\mathbb{Q}}\left[\exp(\mathbf{e}_i^\top\mathcal{P}\mathbf{X}(T))\,|\,\mathcal{F}_t\right] \tag{17}$$

for $t \leq T$ and $i = 1,\ldots,d$. We recall that $\{\mathbf{e}_i\}_{i=1}^d$ are the canonical basis vectors in $\mathbb{R}^d$, thus $\mathbf{e}_i^\top\mathcal{P}\mathbf{X}(T)$ is the ith coordinate of the vector $\mathcal{P}\mathbf{X}(t)$, i.e., $\ln S_i(t)$. We are naturally led to define the following class of factor processes:

Definition 15. *A process $\mathbf{X}$ is called exponentially $\mathbb{Q}$-affine if for every $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z}^\top\mathbf{X}(T)$ has finite exponential moment under $\mathbb{Q}$ and there exist measurable mappings $(t,T) \mapsto \boldsymbol{\alpha}(t,T;\mathbf{z}) \in \mathbb{R}^n$ and $(t,T) \mapsto a(t,T;\mathbf{z}) \in \mathbb{R}$ such that*

$$\mathbb{E}_{\mathbb{Q}}[\exp(\mathbf{z}^\top\mathbf{X}(T))\,|\,\mathcal{F}_t] = \exp\left(\boldsymbol{\alpha}(t,T;\mathbf{z})^\top\mathbf{X}(t) + a(t,T;\mathbf{z})\right)$$

for all $t \leq T$.

We remark that our definition of exponentially $\mathbb{Q}$-affine processes is closely related to the time-inhomogeneous affine processes introduced by Filipović [23, Definition 2.1]**. Our definition is tailormade to the purpose of forward pricing, restricting the affine property to $z \in \mathbb{R}^n$ and claiming additional exponential integrability of the process under $\mathbb{Q}$.**

For exponentially $\mathbb{Q}$-affine factor processes, we have:

Proposition 16. *If $\mathbf{X}$ is an n-dimensional exponentially $\mathbb{Q}$-affine factor process, then $\mathbf{F}(t,T), t \leq T$ has coordinates*

$$F_i(t,T) = \exp\left(\boldsymbol{\alpha}(t,T;\mathcal{P}^\top\mathbf{e}_i)^\top\mathbf{X}(t) + a(t,T;\mathcal{P}^\top\mathbf{e}_i)\right)$$

for $i = 1,\ldots,d$.

Proof. This follows immediately from the definition of exponentially affinity and (17). □

For example, if $\mathbf{X}$ is given by (12) (under $\mathbb{Q}$), it will be exponentially $\mathbb{Q}$-affine, as the following Lemma shows:

Lemma 17. *Suppose that $\mathbf{X}$ is the factor process in $\mathbb{R}^3$ defined in* (12)*. Then $\mathbf{X}$ is exponentially $\mathbb{Q}$-affine, with $\boldsymbol{\alpha}(t,T;\mathbf{z}) = \mathcal{A}(T-t)^\top\mathbf{z}$ and*

$$a(T-t;\mathbf{z}) = \int_0^{T-t} \mathbf{z}^\top\mathcal{A}(s)\boldsymbol{\mu} + \mathbf{z}^\top\mathcal{A}(s)\Sigma C\Sigma^\top\mathcal{A}(s)^\top\mathbf{z}\,ds.$$

Here, $\mathcal{A}$ is defined in (13) *and C is the 3×3 covariance matrix of* $\mathbf{W}$.

Proof. It holds that

$$\mathbf{X}(T) = \mathcal{A}(T-t)\mathbf{X}(t) + \int_0^{T-t} \mathcal{A}(s)\boldsymbol{\mu}\, ds + \int_t^T \mathcal{A}(T-s)\Sigma\, d\mathbf{W}(s),$$

where $\mathcal{A}(s)$ is defined in (13). As the stochastic integral on the right hand side is a Wiener integral, it is a Gaussian random variable and hence $\mathbf{z}^\top \mathbf{X}(t)$ has finite exponential moment for every $\mathbf{z} \in \mathbb{R}^n$. By the independent increment property of Brownian motion and the $\mathbf{X}(t)$ being $\mathcal{F}_t$-measurable, we find

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[\exp(\mathbf{z}^\top \mathbf{X}(T)) \,|\, \mathcal{F}_t\right] &= \exp\left(\mathbf{z}^\top \mathcal{A}(T-t)\mathbf{X}(t) + \int_0^{T-t} \mathbf{z}^\top \mathcal{A}(s)\boldsymbol{\mu}\, ds\right) \\ &\quad \times \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_t^T \mathbf{z}^\top \mathcal{A}(T-s)\Sigma\, d\mathbf{W}(s)\right)\right] \\ &= \exp\left(\mathbf{z}^\top \mathcal{A}(T-t)\mathbf{X}(t) + \int_0^{T-t} \mathbf{z}^\top \mathcal{A}(s)\boldsymbol{\mu}\, ds\right) \\ &\quad \times \exp\left(\int_t^T \mathbf{z}^\top \mathcal{A}(T-s)\Sigma C \Sigma^\top \mathcal{A}(T-s)^\top \mathbf{z}\, ds\right).\end{aligned}$$

The result follows. □

We remark in passing that one can easily extend the above Lemma to higher dimensions than 3. Observe that both $\boldsymbol{\alpha}$ and a are homogeneous, i.e., depending only on the time to maturity $T-t$. Let $(\mathcal{P}, \mathbf{c})$ for $\mathbf{c} \in \mathbb{R}_0^2$ and $\mathcal{P} \in \mathbb{R}^{2\times 3}$ be a cointegrated pricing system (under $\mathbb{Q}$), i.e., $\mathbf{c}^\top \mathcal{P} \in \mathcal{C}_{\mathbf{X}}$. We have then

$$\ln f_i(t,x) = \mathbf{e}_i^\top \mathcal{P}\mathcal{A}(x)\mathbf{X}(t) + a(x; \mathcal{P}^\top \mathbf{e}_i)$$

for $i = 1, 2$. Moreover, as we have seen earlier, $\mathbf{c}^\top \mathcal{P}\mathcal{A}(x) \in \mathcal{C}_{\mathbf{X}}$, and therefore the logarithmic forward prices $f_1(t,x)$ and $f_2(t,x)$ are cointegrated for the cointegration vector $\mathbf{c}$.

Next, let us focus on general LS-processes as factors in a geometric model. Suppose that $\mathbf{X}^m$ has $\mathbb{Q}$-dynamics defined as in (5). We find the following:

Proposition 18. *Assume $\mathbf{X}^m$ is an m-dimensional process with $\mathbb{Q}$-dynamics as in* (5). *Let $t \le T$. If $\mathbf{z}^\top \mathbf{X}^m(t)$ has finite exponential moment under $\mathbb{Q}$ for $\mathbf{z} \in \mathbb{R}^m$, then*

$$\begin{aligned}&\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{z}^\top \int_{-\infty}^T G(T-s)\, d\mathbf{L}(s)\right) \,|\, \mathcal{F}_t\right] \\ &\quad = \exp\left(\mathbf{z}^\top \int_{-\infty}^t G(T-s)\, d\mathbf{L}(s) + \int_0^{T-t} \psi_{\mathbb{Q}}(G(s)^\top \mathbf{z})\, ds\right)\end{aligned}$$

where $\psi_{\mathbb{Q}}$ is the cumulant of $\mathbf{L}(1)$ under $\mathbb{Q}$.

Proof. Since $\int_{-\infty}^t G(T-s)\, d\mathbf{L}(s)$ is $\mathcal{F}_t$-measurable, and the Lévy process $\mathbf{L}$ has independent increments, it follows that

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{z}^\top \int_{-\infty}^T G(T-s)\, d\mathbf{L}(s)\right) \,|\, \mathcal{F}_t\right]$$

$$= \exp\left(\mathbf{z}^\top \int_{-\infty}^t G(T-s)\,d\mathbf{L}(s)\right)\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\left(\int_t^T \mathbf{z}^\top G(T-s)\,d\mathbf{L}(s)\right)\right)\right]$$
$$= \exp\left(\mathbf{z}^\top \int_{-\infty}^t G(T-s)\,d\mathbf{L}(s) + \int_t^T \psi_{\mathbb{Q}}\left(G(T-s)^\top \mathbf{z}\right)\,ds\right).$$

and the result follows. □

As we will show next, there exist cases where a cointegrated spot model fails to imply cointegrated forward prices. To this end, express the factor process as $\mathbf{X} = (\mathbf{X}^m, \widehat{\mathbf{X}}) \in \mathbb{R}^n$, for $\widehat{\mathbf{X}}$ being a process in $\mathbb{R}^{n-m}$, $n > m \in \mathbb{N}$. We further suppose that $\mathbf{X}^m$ is an LS-process under $\mathbb{Q}$, as in (5). Consider a cointegrated pricing system $(\mathcal{P}, \mathbf{c})$, that is, $\mathcal{P}^\top\mathbf{c} \in \mathcal{C}_{\mathbf{X}}$, and introduce the following representation of the $d \times n$-matrix $\mathcal{P}$: let $\mathcal{P}^m \in \mathbb{R}^{d\times m}$ and $\widehat{\mathcal{P}} \in \mathbb{R}^{d\times(n-m)}$ be such that $\mathcal{P} = [\mathcal{P}^m \quad \widehat{\mathcal{P}}]$. Then for $i = 1,\ldots,d$,

$$\mathbf{e}_i^\top \mathcal{P}\mathbf{X}(T) = \mathbf{e}_i^\top \mathcal{P}^m\mathbf{X}^m(T) + \mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(T).$$

Assume that $\mathbf{e}_i^\top \mathcal{P}^m\mathbf{X}^m(T)$ and $\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(T)$ have finite exponential moment under $\mathbb{Q}$, and that they are conditionally independent with respect to $\mathcal{F}_t$ for all $t \le T$. Then it holds for $i = 1,\ldots,d$ and $t \le T$ that

$$\begin{aligned}
F_i(t,T) &= \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{e}_i^\top \mathcal{P}\mathbf{X}(T)\right) \mid \mathcal{F}_t\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{e}_i^\top \mathcal{P}^m\mathbf{X}^m(T)\right) \mid \mathcal{F}_t\right]\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(T)\right) \mid \mathcal{F}_t\right] \\
&= \exp\left(\mathbf{e}_i^\top \mathcal{P}^m \int_{-\infty}^t G(T-s)\,d\mathbf{L}(s) + \int_0^{T-t} \psi_{\mathbb{Q}}\left(G(s)^\top \mathcal{P}^{m,\top}\mathbf{e}_i\right)\,ds\right) \\
&\qquad \times \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(T)\right) \mid \mathcal{F}_t\right]
\end{aligned}$$

where we used Proposition 18 with $\mathbf{z} = \mathcal{P}^{m,\top}\mathbf{e}_i$ in the last equality. We find that

$$\begin{aligned}
\ln f_i(t,x) = \mathbf{e}_i^\top \mathcal{P}^m \int_{-\infty}^t G(t-s+x)\,d\mathbf{L}(s) + \int_0^x \psi_{\mathbb{Q}}\left(G(s)^\top \mathcal{P}^{m,\top}\mathbf{e}_i\right)\,ds \\
+ \ln \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(t+x)\right) \mid \mathcal{F}_t\right],
\end{aligned}$$

for $i = 1,\ldots,d$ and $x \ge 0$. The last term is nonlinear in the vector $\widehat{\mathbf{X}}$, and $\mathbf{c}$ may fail to be a cointegration vector for $\ln \mathbf{f}(t,x)$, even in the case when $\mathcal{P}^\top\mathbf{c} \in \mathcal{C}^m_{\mathbf{X}}$.

However, typically in applications, $\widehat{\mathbf{X}} = \mathbf{U}$ for a Lévy process $\mathbf{U}$ in $\mathbb{R}^{n-m}$. In the simplest case, $\mathbf{U}(t) = \boldsymbol{\mu}t + \Sigma\,d\mathbf{W}(t)$ for $\boldsymbol{\mu} \in \mathbb{R}^{n-m}$, Σ an $(n-m)\times(n-m)$ volatility matrix and $\mathbf{W}$ a $\mathbb{Q}$-Brownian motion in $\mathbb{R}^{n-m}$. Suppose that $\mathbf{U}$ is independent of $\mathbf{L}$. Then it follows that $\mathbf{e}_i^\top \mathcal{P}^m\mathbf{X}^m(T)$ and $\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(T)$ are conditionally independent with respect to $\mathcal{F}_t$ for all $t \le T$. Moreover, $\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(T)$ have finite exponential moment under $\mathbb{Q}$ when $\mathbf{U}$ is a drifted Brownian motion as exemplified above. Without any loss of generality, we assume that the coordinates W_i, $i = 1,\ldots n-m$, of $\mathbf{W}$ are independent. Denoting $\kappa_{\mathbb{Q}}$ the cumulant function of $\mathbf{U}$, we find by resorting to the independent increment property of Lévy processes that

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(t+x)\right) \mid \mathcal{F}_t\right] = \exp\left(\mathbf{e}_i^\top \widehat{\mathcal{P}}\widehat{\mathbf{X}}(t) + x\kappa_{\mathbb{Q}}(\widehat{\mathcal{P}}^\top\mathbf{e}_i)\right).$$

In this case we have that

$$\ln \mathbf{f}(t,x) = \mathbf{c}^\top \mathcal{P} \begin{pmatrix} \widetilde{\mathbf{X}}^m(t,x) \\ \widehat{\mathbf{X}}(t) \end{pmatrix} + \mathbf{h}(x) \tag{18}$$

with $\mathbf{h}(x) \in \mathbb{R}^d$ having coordinates

$$h_i(x) = \int_0^x \psi_{\mathbb{Q}} \left(G(s)^\top \mathcal{P}^{m,\top} \mathbf{e}_i \right) \, ds + x\kappa_{\mathbb{Q}}(\widehat{\mathcal{P}}^\top \mathbf{e}_i)). \tag{19}$$

After a simple modification of Lemma 9, we know that $\widetilde{\mathbf{X}}^m(t,x)$ is a strictly stationary process in $\mathbb{R}^m$. In this case, any $\mathbf{c} \in \mathbb{R}^d_0$ such that $\mathcal{P}^\top \mathbf{c} \in \mathcal{C}^m_{\mathbf{X}}$ is a cointegration vector for $\ln \mathbf{f}(t,x)$. Thus, the cointegration vector for the spot yields cointegration of the forwards as well.

Throughout this section we have assumed a factor process specified directly under the pricing measure $\mathbb{Q}$ in our analysis of cointegration for forward markets. Indeed, we have supposed a cointegrated spot model *under the pricing measure* $\mathbb{Q}$ **rather than under the market probability $\mathbb{P}$. In practice, the situation may be that one has a cointegrated spot model** *under the market probability* $\mathbb{P}$**, and introduces a pricing measure $\mathbb{Q}$ to price forwards on the spot prices. The next step is to analyse possible cointegration of the forward prices. We will not pursue this discussion further here, but leave it for future research.**

4. Cointegration for Hilbert-valued stochastic processes

We want to generalize the concept of cointegration to Hilbert-valued stochastic processes. Recall from above that we defined cointegration in finite dimensions by the following scheme of linear mappings for a cointegrated pricing system $(\mathcal{P}, \mathbf{c})$ and a factor process $\mathbf{X} \in \mathbb{R}^n$:

$$\mathbf{X} \in \mathbb{R}^n \xrightarrow{\mathcal{P}} \mathbf{S}(t) \in \mathbb{R}^d \xrightarrow{\mathbf{c}} \mathbf{c}^\top \mathbf{S}(t) \in \mathbb{R}.$$

I.e., for a cointegrated pricing system $(\mathcal{P}, \mathbf{c})$, we use a linear operator $\mathcal{P}$ to map the factor vector from the *factor space* $\mathbb{R}^n$ to the *price space* $\mathbb{R}^d$, and next the linear operator $\mathbf{c}$ to map the price vector from the price space to the real line, which we can think of as the *cointegration space*. We lift this to Hilbert-valued stochastic processes:

Let F, P and C be three separable Hilbert spaces, denoting the *factor*, *price* and *cointegration* space, resp. We denote $\langle \cdot, \cdot \rangle_i$, the inner product with associated norm $| \cdot |_i$, for $i = \mathsf{F}, \mathsf{P}, \mathsf{C}$. Assume $\mathcal{P} \in L(\mathsf{F}, \mathsf{P})$, which we call the *price operator* and $\mathcal{C} \in L(\mathsf{P}, \mathsf{C})$ with $\mathcal{C} \neq 0$ the *cointegration operator*. For $\{X(t)\}_{t \geq 0}$ being an F-valued predictable process, we define the price process

$$Y(t) = \mathcal{P}X(t) \, , t \geq 0 \tag{20}$$

which becomes a P-valued predictable process.

Definition 19. *We say that $(\mathcal{P}, \mathcal{C})$ is a cointegrated pricing system if the C-valued stochastic process $\{\mathcal{C}\mathcal{P}X(t)\}_{t \geq 0}$ admits a limiting distribution. We say that the price process $Y(t) = \mathcal{P}X(t)$ for given $\mathcal{P} \in L(\mathsf{F}, \mathsf{P})$ is cointegrated if there exists a $\mathcal{C} \in L(\mathsf{P}, \mathsf{C})$, $\mathcal{C} \neq 0$, such that $\mathcal{C}Y(t)$ admits a limiting distribution.*

Obviously, if Y in (20) is cointegrated, then $(\mathcal{P},\mathcal{C})$ is a cointegrated pricing system for the given cointegration operator $\mathcal{C}$.

We recall from infinite dimensional stochastic analysis (see e.g. Peszat and Zabczyk [37]) that the distribution of a C-valued random variable Z is defined as the image measure P_Z on the Borel sets $\mathcal{B}(\mathsf{C})$ of C, that is $P_Z(A)=\mathbb{P}(Z\in A)$ for $A\in\mathcal{B}(\mathsf{C})$. The definition of cointegration demands the existence of a probability measure P_∞ on $\mathcal{B}(\mathsf{C})$ such that $P_{\mathcal{C}\mathcal{P}X(t)}\to P_\infty$ when $t\to\infty$, where the limit is in the sense of probability measures, e.g., for every bounded measurable function $g:\mathsf{C}\to\mathbb{R}$, it holds for $t\to\infty$

$$\int_{\mathsf{C}} g(u)\,P_{\mathcal{C}\mathcal{P}X(t)}(du)\to\int_{\mathsf{C}} g(u)P_\infty(du).$$

Denote the cumulant functional of X by $\Psi_X(t,v)\,,v\in\mathsf{F}$, defined as

$$\Psi_X(t,v):=\log\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle v,X(t)\rangle_{\mathsf{F}}}\right],$$

where log is the distinguished logarithm (see e.g. Sato [38]). We have the following equivalent characterization of cointegration:

Proposition 20. *Y is cointegrated if and only of there exists a $\mathcal{C}\in L(\mathsf{P},\mathsf{C})$, $\mathcal{C}\neq 0$, and a cumulant function $\Psi_{\mathcal{C}}$ such that*

$$\lim_{t\to\infty}\Psi_X(t,\mathcal{P}^*\mathcal{C}^*u)=\Psi_{\mathcal{C}}(u)$$

for all $u\in\mathsf{C}$.

Proof. It holds that $\langle u,\mathcal{C}\mathcal{P}X(t)\rangle_{\mathsf{C}}=\langle\mathcal{P}^*\mathcal{C}^*u,X(t)\rangle_{\mathsf{F}}$ for every $u\in\mathsf{C}$. Thus,

$$\log\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u,\mathcal{C}\mathcal{P}X(t)\rangle_{\mathsf{C}}}\right]=\Psi_X(t,\mathcal{P}^*\mathcal{C}^*u).$$

If $(\mathcal{P},\mathcal{C})$ is a cointegrated pricing system, then $\Psi_{\mathcal{C}}$ is the cumulant function of P_∞. Opposite, the existence of a cumulant function $\Psi_{\mathcal{C}}$ as the limit of $\Psi_X(t,\mathcal{P}^*\mathcal{C}^*u)$ yields the existence of a P_∞ having $\Psi_{\mathcal{C}}$ as its cumulant. The result follows. □

We next connect cointegration in Hilbert space to cointegration in finite dimensions, as considered in the previous sections:

Proposition 21. *Let $(\mathcal{P},\mathcal{C})$ be a cointegrated pricing system (for the factor process X in F). Assume that $\dim(\ker\mathcal{C}^\perp)=:d<\infty$ and $\dim(\ker\mathcal{P}^\perp)=:n<\infty$. Then for every $\mathcal{T}\in\mathsf{C}^*$, there exist $\mathbf{c}_{\mathcal{T}}\in\mathbb{R}^d,\overline{\mathcal{P}}\in\mathbb{R}^{d\times n}$ and an $\mathbb{R}^n$-valued factor process $\mathbf{X}(t)$ such that*

$$\mathcal{T}\mathcal{C}\mathcal{P}X(t)=\mathbf{c}_{\mathcal{T}}^\top\overline{\mathcal{P}}\mathbf{X}(t),$$

and where the real-valued process $t\mapsto\mathbf{c}_{\mathcal{T}}^\top\overline{\mathcal{P}}\mathbf{X}(t)$ admits a limiting distribution.

Proof. For any $u\in\mathsf{P}$, $u^\perp:=u-\mathrm{Proj}_{\ker\mathcal{C}}u\in\ker\mathcal{C}^\perp$ and for an ONB $\{h_i\}_{i=1}^d$ in $\ker\mathcal{C}^\perp$ we find,

$$u^\perp=\sum_{i=1}^d\langle u^\perp,h_i\rangle_{\mathsf{P}}h_i=\sum_{i=1}^d\langle u,h_i\rangle_{\mathsf{P}}h_i-\sum_{i=1}^d\langle\mathrm{Proj}_{\ker\mathcal{C}}u,h_i\rangle_{\mathsf{P}}h_i=\sum_{i=1}^d\langle u,h_i\rangle_{\mathsf{P}}h_i.$$

But then, since $\mathcal{C}u=\mathcal{C}u^\perp$ for every $u\in\mathsf{P}$, it follows

$$\mathcal{C}\mathcal{P}X(t)=\sum_{i=1}^d\langle\mathcal{P}X(t),h_i\rangle_{\mathsf{P}}\mathcal{C}h_i.$$

Next, for any $v \in \mathsf{F}$, we have that $v^\perp := v - \mathrm{Proj}_{\ker \mathcal{P}} v \in \ker \mathcal{P}^\perp$, and for an ONB $\{f_j\}_{j=1}^n$ in $\ker \mathcal{P}^\perp$ it holds that

$$v^\perp = \sum_{j=1}^n \langle v, f_j \rangle_{\mathsf{F}} f_j.$$

Since $\mathcal{P} v = \mathcal{P} v^\perp$ for any $v \in \mathsf{F}$, we derive

$$\mathcal{P} X(t) = \sum_{j=1}^n \langle X(t), f_j \rangle_{\mathsf{F}} \mathcal{P} f_j.$$

From this we find

$$\mathcal{C}\mathcal{P} X(t) = \sum_{i=1}^d \sum_{j=1}^n \langle X(t), f_j \rangle_{\mathsf{F}} \langle \mathcal{P} f_j, h_i \rangle_{\mathsf{P}} \mathcal{C} h_i. \tag{21}$$

Define the $d \times n$-matrix $\overline{\mathcal{P}} := \{\langle \mathcal{P} f_j, h_i \rangle_{\mathsf{P}}\}_{i=1,\ldots,d, j=1,\ldots,n}$ and the $\mathbb{R}^n$-valued factor process $\mathbf{X}(t) := (\langle X(t), f_1 \rangle_{\mathsf{F}}, \ldots, \langle X(t), f_n \rangle_{\mathsf{F}})^\top$. Finally, we introduce $\mathbf{c}_\mathcal{T} := (\mathcal{T}\mathcal{C} h_1, \ldots, \mathcal{T}\mathcal{C} h_d)^\top \in \mathbb{R}^d$, and the representation of $\mathcal{T}\mathcal{C}\mathcal{P} X(t)$ follows.

Note that for any $\theta \in \mathbb{R}$,

$$\theta \mathcal{T}\mathcal{C}\mathcal{P} X(t) = \langle X(t), \mathcal{P}^* \mathcal{C}^* \mathcal{T}^* \theta \rangle_{\mathsf{F}}.$$

Therefore, by the assumption that $\mathcal{C}\mathcal{P} X(t)$ admits a limiting distribution in combination with Proposition 20, there exists a function $\mathbb{R} \ni \theta \mapsto \Psi_{\mathcal{T}\mathcal{C}}(\theta) \in \mathbb{C}$ given by

$$\Psi_{\mathcal{T}\mathcal{C}}(\theta) = \lim_{t \to \infty} \Psi_X(t, \mathcal{P}^* \mathcal{C}^* \mathcal{T}^* \theta) = \Psi_\mathcal{C}(\mathcal{T}^* \theta).$$

The function $\Psi_{\mathcal{T}\mathcal{C}}$ is a cumulant function, since $\mathcal{T}$ is a continuous linear operator (see e.g. Sato [38, Prop. 2.5 (viii)]). The Proposition follows. □

Notice that for any $\mathcal{T} \in \mathsf{C}^*$, $\mathcal{T}\mathcal{C} \in \mathsf{P}^*$, and whenever $\mathcal{T}\mathcal{C} \neq 0$ we can interpret $(\mathcal{P}, \mathcal{T}\mathcal{C})$ as a cointegrated pricing system with cointegration space $\mathbb{R}$. This holds for general cointegrated pricing system $(\mathcal{P}, \mathcal{C})$ and not only those for which the complement space of the kernels of $\mathcal{C}$ and $\mathcal{P}$ are finite.

We see from the proof of Proposition 21 that only $\mathbf{c}_\mathcal{T}$ is depending on $\mathcal{T}$, which explains the subscript. Given the factor process $\mathbf{X}$ in Proposition 21, it follows that $(\overline{\mathcal{P}}, \mathbf{c}_\mathcal{T})$ is a cointegrated pricing system whenever $\mathbf{c}_\mathcal{T} \neq 0$, and indeed, $\mathbf{c}_\mathcal{T}$ is in that case a cointegration vector for the price vector $\mathbf{S}(t) = \overline{\mathcal{P}}\mathbf{X}(t)$. The vector $\mathbf{c}_\mathcal{T}$ is further depending on $\mathcal{C}$, naturally. The pricing matrix $\overline{\mathcal{P}}$ depends on the basis functions $\{f_j\}_{j=1}^n$ in $\ker \mathcal{P}^\perp$ and $\{h_i\}_{i=1}^d$ in $\ker \mathcal{C}^\perp$. Thus, it depends on the cointegrated pricing system $(\mathcal{P}, \mathcal{C})$. The finite-dimensional factor process $\mathbf{X}$ depends on the basis $\{f_j\}_{j=1}^n$, and thus on the pricing operator $\mathcal{P}$. **Remark that we here dispense with the minimality condition on $\overline{\mathcal{P}}$ and the assumption $n \geq d$ made in the finite dimensional case.**

From (21) it follows that $\mathcal{C}\mathcal{P} X(t)$ is a process with values in the finite-dimensional subspace $\mathrm{span}\{\mathcal{C} h_1, \ldots, \mathcal{C} h_d\}$ of C when $\ker \mathcal{C}^\perp$ and $\ker \mathcal{P}^\perp$ have finite dimension. Thus, we may introduce the following definition:

Definition 22. *A cointegrated pricing system $(\mathcal{P}, \mathcal{C})$ has a finite dimensional realization (FDR) if, for $n, d \in \mathbb{N}$, there exist an $\mathbb{R}^n$-valued factor process $\mathbf{X}$, a $d \times n$ pricing matrix $\overline{\mathcal{P}}$ and a $c \in \mathsf{C}^{\times d}$ such that $\mathcal{C}\mathcal{P} X(t) = c^\top \overline{\mathcal{P}} \mathbf{X}(t)$.*

In view of Proposition 21, we have an FDR when $\ker \mathcal{C}^{\perp}$ and $\ker \mathcal{P}^{\perp}$ are finite dimensional. In this case, $c = (\mathcal{C}h_1, \dots, \mathcal{C}h_d)^{\top}$ with $\{h_i\}_{i=1}^d$ being the ONB of $\ker \mathcal{C}^{\top}$. If $(\mathcal{P}, \mathcal{C})$ is a general cointegrated pricing system which has an FDR, then for any $\mathcal{T} \in \mathsf{C}^*$ we find that

$$\mathcal{T}\mathcal{C}\mathcal{P}X(t) = \mathbf{c}_{\mathcal{T}}^{\top}\overline{\mathcal{P}}\mathbf{X}(t),$$

for $\mathbf{c}_{\mathcal{T}} := \mathcal{T}c = (\mathcal{T}c_1, \dots \mathcal{T}c_d)^{\top} \in \mathbb{R}^d$. Hence, $(\overline{\mathcal{P}}, \mathcal{T}c)$ will be a finite dimensional cointegrated pricing system for the factor process $\mathbf{X}$ whenever $\mathbf{c}_{\mathcal{T}} \neq 0$. We remark that Definition 22 does not really rest on the fact that there exist any limiting distribution, but as we work with cointegration in this paper, we focus on cointegrated pricing systems, that is, pricing systems $(\mathcal{P}, \mathcal{C})$ for which $\mathcal{C}\mathcal{P}X(t)$ admits a limiting distribution.

We recall that we have not assumed any minimality of the pricing matrix $\overline{\mathcal{P}}$ in the above considerations. From the proof of Proposition 21 we see that the $d \times n$-matrix $\overline{\mathcal{P}}$ has elements $\langle \mathcal{P}f_j, h_i \rangle_{\mathsf{P}}$, and minimality is achieved as long as this matrix has full rank. However, the next proposition shows that we must take into account a possible finite-dimensionality of the factor process X as well. Indeed, another situation where we may have an FDR is when the factor process has a finite-dimensional state space:

Proposition 23. *Assume that the factor process $\{X(t)\}_{t \geq 0}$ takes values in $\mathsf{F}_n \subset \mathsf{F}$, where $\dim(\mathsf{F}_n) := n < \infty$ for $n \in \mathbb{N}$. Then any cointegrated pricing system $(\mathcal{P}, \mathcal{C})$ has a finite dimensional realization, with $\overline{\mathcal{P}} = Id$ (the identity matrix on $\mathbb{R}^n$), $\mathbf{X}(t) := (\langle X(t), f_1 \rangle_{\mathsf{F}}, \dots, \langle X(t), f_n \rangle_{\mathsf{F}})^{\top} \in \mathbb{R}^n$ for an ONB $\{f_j\}_{j=1}^n$ of F_n and $c = (\mathcal{C}\mathcal{P}f_1, \dots, \mathcal{C}\mathcal{P}f_n)^{\top} \in \mathsf{C}^{\times n}$.*

Proof. If $X(t) \in \mathsf{F}_n$, then $X(t) = \sum_{j=1}^n \langle X(t), f_j \rangle_{\mathsf{F}} f_j$ and therefore

$$\mathcal{C}\mathcal{P}X(t) = \sum_{j=1}^n \langle X(t), f_j \rangle_{\mathsf{F}} \mathcal{C}\mathcal{P}f_j = c^{\top}\mathrm{Id}\mathbf{X}(t).$$

The result follows. □

This result indicates strongly the possible non-uniqueness of the FDR, since depending on the pricing operator $\mathcal{P}$, one may specify a different $\overline{\mathcal{P}}$ than the identity matrix, and thus also different c. It also shows that the question of minimality depends on $\mathcal{C}, \mathcal{P}$ and the possible finite dimensionality of X.

Let us now focus on the case where F and P can be represented as product spaces, e.g., when $\mathsf{F} = \mathsf{H}^{\times n}$ and $\mathsf{P} = \mathsf{K}^{\times d}$ for two separable Hilbert spaces H and K. We denote the inner product as usual by $\langle \cdot, \cdot \rangle_i$ with corresponding norms $| \cdot |_i$, where the subscript indicates the space, here $i = \mathsf{H}, \mathsf{K}$. The inner product on the product space F is then given by $\langle u, v \rangle_{\mathsf{F}} = \sum_{j=1}^n \langle u_j, v_j \rangle_{\mathsf{H}}$ for $u = (u_1, \dots, u_n) \in \mathsf{F}$ and $v = (v_1, \dots, v_n) \in \mathsf{F}$ (and likewise for P).

To make an example, suppose we have given a factor process $X \in \mathsf{H}^{\times n}$ and a pricing operator $\mathcal{P}$ given as an $d \times n$-matrix of operators $\mathcal{P} = \{\mathcal{P}_{ij}\}_{i=1,\dots,d, j=1,\dots,n}$ with $\mathcal{P}_{ij} \in L(\mathsf{H}, \mathsf{K})$. Then the pricing vector will be $Y(t) = \mathcal{P}X(t)$, which is a $\mathsf{K}^{\times d}$-valued stochastic process. Indeed, we have that $Y = (Y_1, \dots, Y_d)^{\top}$ with

$$Y_i(t) = \sum_{j=1}^n \mathcal{P}_{ij} X_j(t)$$

for $i = 1, \ldots, d$. In analogy with Example 1, we assume that $(X_1, \ldots, X_{n-1})^\top \in \mathsf{H}^{\times(n-1)}$ admits a limiting distribution, while X_n may be non-stationary. We observe that any $\mathcal{C} = (\mathcal{C}_1, \ldots, \mathcal{C}_d)^\top$ with $\mathcal{C}_i \in L(\mathsf{K}, \mathsf{C})$ will be such that $\mathcal{C} \in L(\mathsf{K}^{\times d}, \mathsf{C})$. Under the condition $\sum_{i=1}^d \mathcal{C}_i \mathcal{P}_{in} = 0$ we find $\mathcal{C}Y(t) = \sum_{i=1}^d \sum_{j=1}^{n-1} \mathcal{C}_i \mathcal{P}_{ij} X_j(t)$, that is, a C-valued stochastic process not depending on X_n but only on X_j for $j = 1, \ldots, n-1$. This provides us with a simple example of a cointegrated pricing system.

A way to generate a system of factor processes $X \in \mathsf{H}^{\times n}$ can be as follows: consider an $\mathbb{R}^m$-valued stochastic process $\{\mathbf{Z}(t)\}_{t \geq 0}$ and $\mathcal{A} \in L(\mathbb{R}^m, \mathsf{H}^{\times n})$. For $b \in \mathsf{H}^{\times n}$, define the factor process

$$X(t) = \mathcal{A}\mathbf{Z}(t) + b.$$

We remark that $\mathcal{A}$ can be represented as an $n \times m$-matrix with elements in H. Indeed, the columns of this matrix will be given by the action of $\mathcal{A}$ on the canonical basis vectors in $\mathbb{R}^m$. If H is some space of functions on $\mathbb{R}_+$, we may relate the factor process X to the affine models of forward prices from the previous section, i.e., the affine forward models provide a class of factors in an infinite dimensional framework. The existence of a limiting distribution of one or more of the factors $X_j, j = 1, \ldots, n$ can be traced back to the process $\mathbf{Z}$. Indeed, this simplified case relates us back to the models considered in Section 3, for example the polynomial processes in Proposition 11.

4.1. **A discussion of cross-commodity forward markets.** Let us now focus specifically on commodity forward markets, and start with a discussion on cross-commodity models. Suppose we have d forward markets, with forward price dynamics denoted by $f_i(t,x)$, $i = 1, \ldots, d$ and $x \in \mathbb{R}_+$ being time to maturity. We are aiming at a d-dimensional model of the forward curve dynamics $t \mapsto f(t, \cdot) = (f_1(t, \cdot), \ldots, f_d(t, \cdot))^\top$. We choose H to be a Hilbert space of real-valued measurable functions on $\mathbb{R}_+$. Following the analysis in Benth and Krühner [12], a convenient choice of such a space could be the so-called Filipović space of absolutely continuous functions (see Appendix A for a definition).

Based on the analysis in Benth and Krühner [12] (see also Benth and Krühner [11]), the forward price dynamics $\{f(t, \cdot)\}_{t \geq 0}$ can be expressed as a $\mathsf{H}^{\times d}$-valued stochastic process

$$df(t, \cdot) = \partial f(t, \cdot)\, dt + \beta(t, f(t, \cdot))\, dt + \sigma(t, f(t, \cdot))\, dL(t) \tag{22}$$

where L is a V-valued square-integrable Lévy process with zero mean and V being a separable Hilbert space. We use the notation ∂ for the $d \times d$ matrix-operator

$$\partial = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & \cdots & 0 \\ 0 & \frac{\partial}{\partial x} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{\partial}{\partial x} \end{bmatrix}, \tag{23}$$

with $\partial/\partial x$ being the derivative operator on the functions in H. We assume that this operator is a densely defined unbounded operator on H which is the generator of a C_0-semigroup (the shift semigroup). This holds if we choose H to be the Filipović space, say. Further, the measurable mappings $\sigma : \mathbb{R}_+ \times \mathsf{H}^{\times d} \to L(\mathsf{V}, \mathsf{H}^{\times d})$ and $\beta : \mathbb{R}_+ \times \mathsf{H}^{\times d} \to \mathsf{H}^{\times d}$ are assumed to satisfy the Lipschitz conditions stated in Peszat and Zabczyk [37, Section 9.2] such that there exists a unique mild predictable cadlag solution to (22).

The function β models the risk premium in this cross-commodity model of forward curves. We note in passing that (22) is formulated under $\mathbb{P}$, and to ensure an arbitrage-free dynamics there must exist a probability $\mathbb{Q} \sim \mathbb{P}$ such that the $\mathbb{Q}$-dynamics of f is

$$df(t,\cdot) = \partial f(t,\cdot)\,dt + dM(t)$$

where M is a $\mathsf{H}^{\times d}$-valued (local) $\mathbb{Q}$-martingale (see Benth & Krühner [11]). We will not pursue the existence of such a $\mathbb{Q}$ in further detail here.

We may view the cross-commodity forward model (22) in our contegration context by choosing the factor process X to be equal to the price vector process f. Thus, we have $\mathsf{H} = \mathsf{K}$ and $n = d$, with a pricing matrix $\mathcal{P}$ simply being the identity operator in $\mathsf{H}^{\times d}$. In particular, we let $\mathsf{P} = \mathsf{H}^{\times d}$, i.e., the pricing space is the product space. In many markets, prices are naturally varying over seasons. For example in power markets, prices are typically higher in heating and cooling seasons. Such a behaviour may be modelled into β. Further, many commodities are based on extinguishable resources, with oil and gas as prime examples. For such commodities, one may expect non-stationarity effects in prices. Other sources of non-stationarity are technological changes and inflation. Such non-stationarity could possibly be modelled in the β, as well, for example by adding dependency on additional (non-stationary) stochastic factors Z, e.g., assuming a drift of the form $\beta(t, f(t,\cdot), Z(t))$. The additional factors Z may be Hilbert-valued processes.

Cointegration in this context could be formulated as follows: There is an operator $\mathcal{C} \in L(\mathsf{H}^{\times d}, \mathsf{H})$, $\mathcal{C} \neq 0$, such that the H-valued stochastic process $t \mapsto g(t,\cdot) := \mathcal{C}f(t,\cdot)$ admits a limiting distribution. In many applications one is interested in the spread between two or more forward markets, and it is natural to consider linear combinations of the forward curves, which again will be an element in the space of (marginal) forward curves. This gives a rationale for choosing $\mathsf{C} = \mathsf{H}$. If we assume the rather strong condition that ∂ commutes with $\mathcal{C}$ in the sense that $\mathcal{C}\partial = \frac{\partial}{\partial x}\mathcal{C}$ on $\mathrm{Dom}(\partial)$, we find the stochastic dynamics of g to be

$$dg(t,\cdot) = \frac{\partial}{\partial x} g(t,\cdot)\,dt + \mathcal{C}\beta(t, f(t,\cdot), Z(t))\,dt + \mathcal{C}\sigma(t, f(t,\cdot))\,dL(t).$$

Thus,

$$g(t,\cdot) = \mathcal{S}(t)g_0(\cdot) + \int_0^t \mathcal{S}(t-s)\mathcal{C}\beta(s, f(s,\cdot), Z(s))\,ds + \int_0^t \mathcal{S}(t-s)\mathcal{C}\sigma(s, f(s,\cdot))\,dL(s),$$

where $\mathcal{S}$ is the C_0-semigroup generated by $\partial/\partial x$ on H (the shift semigroup, also called the translation semigroup), and $g(0,\cdot) = \mathcal{C}f(0,\cdot) =: g_0(\cdot) \in \mathsf{H}$. The existence of a limiting distribution is closely linked to properties of the C_0-semigroup along with β and σ. Specializing to $L = W$, a Wiener process, and H being the Filipović space, we may resort to Tehranchi [40] for sufficient conditions for the existence of an invariant measure of g. In particular, these conditions will include the time-homogeneity and Lipschitzianity of β and σ. We remark in passing that Tehranchi [40] treats HJM models, which has a nonlinearity in the drift satisfying a no-arbitrage condition with the volatility σ. In our context we will have a simplified situation where this drift condition is not needed.

As a specific case, we could consider the highly dependent power forward markets in Germany and France. In Germay, there has been a gradual increase of renewable power generation from photovoltaic and wind, and we let $Z(t)$ be a real-valued stochastic process measuring the total generation of such. Since the amount of sunshine over the day is varying with season, and so is the average wind speed, one has that Z is likely to vary seasonally. Moreover, with the "Energiewende" still in place, the process will likely show an increasing trend, at least on a short term horizon. Hence, Z may be thought of as a non-stationary stochastic process. Assume now that $\beta(t, f(t, .), Z(t)) = (\beta_1 Z(t), \beta_2 Z(t))^\top$, for β_1, β_2 **two non-zero constants**, which is an $\mathbb{R}^2$-valued stochastic process, and thus trivially in $\mathsf{H}^{\times 2}$. Further, we let the volatility be constant, in the sense that $\sigma(s, f(s, \cdot)) = \Sigma \in L(\mathsf{V}, \mathsf{H}^{\times 2})$. Under this specification, we choose $\mathcal{C}^\top := (\beta_2, -\beta_1) \neq 0$, which will commute with $\partial/\partial x$, and we find for $g(t, \cdot) := \beta_2 f_1(t, \cdot) - \beta_1 f_2(t, \cdot)$

$$g(t, \cdot) = \mathcal{S}(t) g_0(\cdot) + \int_0^t \mathcal{S}(t-s)\mathcal{C}^\top \Sigma \, dL(s). \tag{24}$$

The cointegration process g will be an Ornstein-Uhlenbeck process with unbounded operator $\partial/\partial x$ and volatility $\mathcal{C}^\top \Sigma$. Invariant measures for Lévy-driven Ornstein-Uhlenbeck processes are thoroughly discussed in Applebaum [3] (see also references therein). Although Tehranchi [40] considers more general HJM-models with Gaussian noise, one can apply his methods to conclude that g in (24) admits a limiting distribution if we choose H to be the Filipović space (see Appendix A). We remark in passing that Tehranchi [40] makes use of the fact that the shift semigroup $\mathcal{S}(t)$ is a strict contraction on a convenient subspace of the Filipović space.

So far we have only considered arithmetic forward models. To introduce a geometric model, of the form $F(t, x) := \exp(f(t, x))$, with f defined by the dynamics (22) and $\exp(f) := (\exp(f_1), \dots, \exp(f_d))$, we must impose additional structure on the Hilbert space H. Indeed, it has to be closed under exponentiating, that is, for any $h \in \mathsf{H}$, it must hold that $\exp h \in \mathsf{H}$. If H is a Banach algebra under pointwise multiplication, this holds true, since in that case we have $|h^n|_{\mathsf{H}} \leq |h|_{\mathsf{H}}^n$ and thus $|\exp h|_{\mathsf{H}} \leq \exp |h|_{\mathsf{H}} < \infty$. We remark that after an appropriate scaling of the norm in the Filipović space, it becomes a Banach algebra (see Benth and Krühner [11]).

4.2. **A three-factor example.** We end this section with a concrete example adopted from Benth [6]. Let H be a Hilbert space of real-valued measurable functions on $\mathbb{R}_+$. Consider a three-factor process $X = (X_1, X_2, X_3)^\top \in \mathsf{H}^{\times 3}$ given by $X_3(t) = L(t)$ where L is an $\mathbb{R}$-valued Lévy process and for $x \in \mathbb{R}_+$,

$$X_k(t, x) = h_k(t, x) + \int_0^t g_k(t + x - s) \, dU_k(s), k = 1, 2. \tag{25}$$

Here, for $k = 1, 2$, U_k are $\mathbb{R}$-valued Lévy processes with zero mean and finite variance, and $h_k(t, \cdot), g_k \in \mathsf{H}$. In the next lemma, we state conditions such that $\{X_k(t)\}_{t \geq 0}$ becomes an H-valued stochastic process.

Lemma 24. *Suppose that the shift semigroup $\{\mathcal{S}(t)\}_{t\geq 0}$ is bounded on H, i.e., $\mathcal{S}(t) \in L(\mathsf{H})$ for all $t \geq 0$. If $\int_0^t |g_k(s+\cdot)|_{\mathsf{H}}^2 \, ds < \infty$ for every $t \geq 0$, then $\{X_k(t)\}_{t\geq 0}$ defined in* (25) *is an H-valued stochastic process. Its cumulant is*

$$\log \mathbb{E}\left[\exp\left(\mathrm{i}(h, X_k(t))_{\mathsf{H}}\right)\right] = \mathrm{i}(h, h_k(t))_{\mathsf{H}} + \int_0^t \psi_{U_k}\left((h, g_k(s + \cdot))_{\mathsf{H}}\right) ds,$$

for $h \in \mathsf{H}$ and ψ_{U_k} the cumulant of $U_k(1)$.

Proof. Fix $t \geq 0$. By assumption, it holds that $g_k(t-s+\cdot) = \mathcal{S}(t-s)g_k(\cdot) \in \mathsf{H}$ for all $s \in [0, t]$. From Peszat and Zabczyk [37], the stochastic integral $\int_0^t g_k(t-s+\cdot) \, dU_k(s)$

is well-defined and defines an element in H if $\int_0^t |g_k(t-s+\cdot)|^2_{\mathsf{H}}\,ds<\infty$, which holds by assumption. Thus, $\{X_k(t)\}_{t\geq 0}$ is an H-valued stochastic process.

We have that the operator $G(t-s)(h)=(h,g_k(t-s+\cdot))_{\mathsf{H}}$ is a linear functional on H. Moreover, by the Cauchy-Schwartz inequality,

$$\begin{aligned}\int_0^t G^2(t-s)(h)\,ds &= \int_0^t (h,g_k(t-s+\cdot))^2_{\mathsf{H}}\,ds\\ &\leq |h|^2_{\mathsf{H}}\int_0^t |g_k(s+\cdot)|^2_{\mathsf{H}}\,ds.\end{aligned}$$

Hence, by the integrability assumption on the norm of g_k, $s\to G(t-s)(h)$ is U_k-integrable on $[0,t]$, and by linearity we find

$$(h,\int_0^t g_k(t-s+\cdot)\,dU_k(s))_{\mathsf{H}}=\int_0^t (h,g_k(t-s+\cdot))_{\mathsf{H}}\,dU_k(s)=\int_0^t G_k(t-s)(h)\,dU_k(s).$$

Hence,

$$\begin{aligned}\log\mathbb{E}\left[\exp\left(\mathrm{i}(h,\int_0^t g_k(t-s+\cdot)\,dU_k(s))_{\mathsf{H}}\right)\right] &= \log\mathbb{E}\left[\mathrm{i}\int_0^t G(t-s)(h)\,dU_k(s)\right]\\ &= \int_0^t \psi_{U_k}\left(G(s)(h)\right)\,ds.\end{aligned}$$

Since U_k is a zero mean square integrable Lévy process, its cumulant becomes

$$\psi_{U_k}(z)=-\frac{1}{2}\sigma^2z^2+\int_{\mathbb{R}}(\mathrm{e}^{\mathrm{i}zy}-1-\mathrm{i}zy)\,\ell(dy)$$

for $\sigma\geq 0$ a constant and ℓ the Lévy measure (see Applebaum [2]). We have

$$|\mathrm{e}^{\mathrm{i}zy}-1-\mathrm{i}zy|=|(\mathrm{i}z)^2\int_0^y\int_0^x \mathrm{e}^{\mathrm{i}zu}\,du\,dx|\leq\frac{1}{2}z^2y^2,$$

and therefore $\psi_{U_k}(G(s)(h))$ is integrable on $[0,t]$ whenever $G(s)(h)\in L^2([0,t])$, which holds by assumption after appealing to the Cauchy-Schwartz inequality, as argued above. □

In the Lemma 24 above we assumed that the function $[0,t]\ni s\mapsto|g_k(s+\cdot)|_{\mathsf{H}}\in\mathbb{R}_+$ is in $L^2([0,t])$. As the shift operator $\mathcal{S}(t)$ is assumed continuous, a sufficient condition for this to hold is that $s\mapsto\|\mathcal{S}(s)\|_{\mathrm{op}}\in L^2([0,t])$. Whenever the family of shift operators defines a strongly continuous semigroup, say, this holds true. If in addition $\{\mathcal{S}(t)\}_{t\geq 0}$ is exponentially stable, we have that $s\mapsto\|\mathcal{S}(s)\|_{\mathrm{op}}\in L^2(\mathbb{R}_+)$. If $s\mapsto|g_k(s+\cdot)|_{\mathsf{H}}\in L^2(\mathbb{R}_+)$ and $h_k(t)$ has a limit in H as $t\to\infty$, it follows from Lemma 24 that $\{X_k\}_{t\geq 0}$ admits a limiting distribution in H.

Introduce next the pricing operator $\mathcal{P}\in L(\mathsf{H}^{\times 3},\mathsf{H}^{\times 2})$ simply as

$$\mathcal{P}=\left[\begin{array}{ccc}\mathrm{Id} & 0 & \mathrm{Id}\\ 0 & \mathrm{Id} & \mathrm{Id}\end{array}\right], \tag{26}$$

where Id is the identity operator on H. If we assume H to be a Banach algebra, we can define the exponential forward price dynamics for a bivariate commodity market by

$$F(t):=\exp\left(\mathcal{P}X(t)\right). \tag{27}$$

Following the analysis in Benth [6], we can choose $h_k(t)$ to ensure an arbitrage-free dynamics (see Proposition 2 in [6]). One can also think of h_k as a model for the market price of risk/risk premium in the forward market.

In this bivariate cross commodity forward price model, we see that $\ln F(t) = \mathcal{P}X(t)$, and thus for any $\mathcal{C} \in L(\mathsf{H}^{\times 2}, \mathsf{C})$, $\mathcal{C} \neq 0$, we have

$$\mathcal{C} \ln F(t) = \mathcal{C}_1 X_1(t) + \mathcal{C}_2 X_2(t) + (\mathcal{C}_1 + \mathcal{C}_2) X_3(t).$$

Here we have represented the operator $\mathcal{C}$ in matrix form, i.e.,

$$\mathcal{C} = \begin{bmatrix} \mathcal{C}_1 & \mathcal{C}_2 \end{bmatrix}$$

for $\mathcal{C}_i \in L(\mathsf{H}, \mathsf{C}), i = 1, 2$. Letting $\mathcal{C}_2 = -\mathcal{C}_1$, we find $\mathcal{C} \ln F(t) = \mathcal{C}_1(X_1(t) - X_2(t))$. In the next lemma, we state sufficient conditions for $\mathcal{C}_1(X_1(t) - X_2(t))$ to admit a limiting distribution in H, which thus yield sufficient conditions for having a cointegrated model.

Lemma 25. *Assume that the shift operator $\mathcal{S}(t)$ is bounded in H for all $t \geq 0$ and $|g_k(s+\cdot)|_{\mathsf{H}} \in L^2(\mathbb{R}_+)$ for $k = 1, 2$. If $h_\infty := \lim_{t\to\infty}(h_1(t) - h_2(t))$ exists in H, then $\mathcal{C}_1(X_1(t) - X_2(t))$ admits a limiting distribution in C. This limiting distribution has cumulant*

$$\lim_{t\to\infty} \log \mathbb{E}\left[\exp\left(\mathrm{i}(h, \mathcal{C}_1(X_1(t) - X_2(t)))_{\mathsf{C}}\right)\right]$$
$$= \mathcal{C}_1 h_\infty + \int_0^\infty \psi_U\left((\mathcal{C}_1^* h, g_1(s+\cdot))_{\mathsf{C}}, -(\mathcal{C}_1^* h, g_2(s+\cdot))_{\mathsf{C}}\right) ds$$

where ψ_U is the cumulant of the bivariate Lévy process $U = (U_1, U_2)$ and $h \in \mathsf{C}$.

Proof. We find, following Lemma 24, that the processes $\int_0^t g_k(t - s + \cdot)\, dU_k(s)$ in H both admit a limiting distribution. Moreover, by using the same argument for marginal integrability as in the proof of Lemma 24, we find that $s \mapsto \psi_U((\mathcal{C}_1^* h, g_1(s+\cdot))_{\mathsf{H}}, -(\mathcal{C}_1^* h, g_2(s+\cdot))_{\mathsf{H}})$ is integrable on $\mathbb{R}_+$ for any $h \in \mathsf{C}$. The result follows. □

A special case is to choose $\mathsf{C} = \mathsf{H}$ and $\mathcal{C}_1 = \mathrm{Id}$. Thus, $X_1(t) - X_2(t)$ admits in particular a limiting distribution when the conditons in Lemma 25 are fullfilled.

Geman and Liu [28] perform an empirical analysis of cointegration between the gas forward markets at Henry Hub (US) and National Balancing Point (UK). They introduce various measures on the forward curves to study how integrated the markets are. More specifically, it is proposed to measure the distance between the average of the respective forward curves, or simply the distance between the implied spot prices (closest maturity forwards), or the distance between some geometric weighted average of forward prices. In our context, the latter two distance measures can be expressed as $|\mathcal{C}_1 X_1(t) - \mathcal{C}_1 X_2(t)|$ with $\mathsf{C} = \mathbb{R}$ and $\mathcal{C}_1 \in \mathsf{H}^*$. For example, in the case of closest forwards (or spot), we choose $\mathcal{C}_1 = \delta_0$, the evaluation operator at zero, assuming that this is continuous on H. A weighted geometric average of the curve, on the other hand, can be translated into a weighted sum of log-prices over different maturities, which gives rise to a linear operator $\mathcal{C}_1$ being a weighted sum of evaluation maps δ_x for different x. The average of the forward curve is not possible to represent via a linear operator $\mathcal{C}_1$ in a geometric model. However, if we choose to work with an arithmetic model, this would simply become an integral operator on the curves in H.

In view of the results in Section 3, one can find spot models that leads to cointegration of forward prices with given time to maturity. In the context of Geman and Liu [28], measuring the difference of the average of the forward curves at given maturity-times could lead to stationarity and thus the conclusion that the markets are cointegrated. However, Geman and Liu [28] do not find evidence for cointegration of the two gas forward markets in Henry Hub and National Balancing Point. This could be explained by a possible term structure of the risk premium (which can be traced back in the β function above) and thus the need for more sophisticated choices of operators $\mathcal{C}$ to reveal a potential cointegration.

Appendix A. The Filipović space

We present the Filipović space following Filipović [22]: Let $w:\mathbb{R}_+\to\mathbb{R}_+$ be a monotonely increasing function with $w(0)=1$ and $\int_0^\infty w^{-1}(x)\,dx<\infty$. Introduce the Filipović space, denoted H_w, as the space of absolutely continuous functions $f:\mathbb{R}_+\to\mathbb{R}$ for which

$$|f|_w^2 := f^2(0)+\int_0^\infty w(x)(f'(x))^2\,dx<\infty,$$

where f' is the weak derivative of f. With the inner product

$$(f,g)_w = f(0)g(0)+\int_0^\infty w(x)f'(x)g'(x)\,dx$$

for $f,g\in\mathsf{H}_w$, H_w becomes a separable Hilbert space. The shift operator $\mathcal{S}(t)$: $f\mapsto f(t+\cdot)$ for $t\geq 0$ defines a C_0-semigroup on H_w which is uniformly bounded. **Moreover, Benth and Krühner** [11] **show that the shift operator is quasi-contractive**. The generator of $\mathcal{S}(t)$ is the derivative operator. The evaluation map $\delta_x: f\mapsto f(x)$ is a linear functional on H_w. Finally, from Benth and Krühner [11], H_w becomes a Banach algebra after appropriate rescaling of the norm $|\cdot|_w$, that is, if $f,g\in H_w$, then $fg\in H_w$ and $\|fg\|_w\leq\|f\|_w\|g\|_w$ with $\|\cdot\|_w := c|\cdot|_w$ for a suitable constant $c>0$ depending on $\int_0^\infty w^{-1}(x)\,dx$.

References

[1] R. Aadland, F. E. Benth and S. Koekebakker (2018). Multivariate modeling and analysis of regional ocean freight rates. *Transp. Research Part E*, **113**, pp. 194-221.

[2] D. Applebaum (2009). *Lévy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge.

[3] D. Applebaum (2015). Infinite dimensional Ornstein-Uhlenbeck processes driven by Lévy processes. *Probab. Surveys*, **12**, pp. 33-54.

[4] J. Back and M. Prokopczuk (2013). Commodity price dynamics and derivative valuation: A review. *Intern. J. Theor. Appl. Finance*, **16**(6), 1350032

[5] O. E. Barndorff-Nielsen, F. E. Benth and A. Veraart (2013). Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. *Bernoulli*, **19**, pp. 803–845.

[6] F. E. Benth (2016). Cointegrated commodity markets and pricing of derivatives in a non-Gaussian framework. In *Advanced Modelling in Mathematical Finance. Festschrift in honour of Ernst Eberlein.* J. Kallsen and A. Papapantoleon (eds.). Springer Verlag, pp. 477–496.

[7] F. E. Benth, J. Šaltytė Benth, and S. Koekebakker (2008).*Stochastic Modelling of Electricity and Related Markets*, World Scientific, Singapore.

[8] F. E. Benth and J. Šaltytė Benth (2013). *Modeling and Pricing in Financial Markets for Weather Derivatives*, World Scientific, Singapore.

[9] F. E. Benth, G. Müller, C. Klüppelberg and L. Vos (2014). Futures pricing in electricity markets based on stable CARMA spot models. *Energy Econ.*, **44**, pp. 392–406.

[10] F. E. Benth and S. Koekebakker (2015). Pricing of forwards and other derivatives in cointegrated commodity markets. *Energy Econ.*, **52**, pp 104–117.

[11] F. E. Benth and P. Krühner (2014). Representation of infinite dimensional forward price models in commodity markets. *Comm. Math. Statistics*, **2**(1), pp. 47–106.

[12] F. E. Benth and P. Krühner (2015). Derivatives pricing in energy markets: an infinite dimensional approach. *SIAM J. Financial Math.*, **6**, pp. 825–869.

[13] F. E. Benth and S. Ortiz-Latorre (2014). A pricing measure to explain the risk premium in power markets. *SIAM J. Financial Math.*, **5**, pp. 685–728.

[14] P. J. Brockwell (2001). Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.*, **53**, pp. 113–124.

[15] F. Comte (1999). Discrete and continuous time cointegration. *J. Econometrics*, **88**, pp. 207–226.

[16] C. Cuchiero, M. Keller-Ressel, and J. Teichmann (2012). Polynomial processes and their applications to mathematical finance. *Finance & Stoch.*, **16**, pp. 711–740.

[17] J.-C. Duan and S. R. Pliska (2004). Option valuation with cointegrated asset prices. *J. Econ. Dynam. & Control*, **28**, pp. 727–754.

[18] J.-C. Duan and A. Theriault (2007). Cointegration in crude oil components and the pricing of crack spread options. Working paper, University of Toronto.

[19] R. F. Engle and C. W. J. Granger (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica*, **55**(2), pp. 251–276.

[20] A. Eydeland and K. Wolyniec (2003). *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. John Wiley & Sons, Hoboken, New Jersey.

[21] W. Farkas, E. Gourier, R. Huitema and C. Necula (2017). A two-factor cointegrated commodity price model with an application to spread option pricing. *J. Banking & Finance*, **77**, pp. 249–268.

[22] D. Filipović (2001). *Consistency Problems for Heath-Jarrow-Morton Interest Rate Models*, Lecture Notes in Mathematics 1760, Springer-Verlag, Berlin.

[23] D. Filipović (2005). Time-inhomogeneous affine processes. *Stoch Proc. Applic.*, **115**(4), pp. 639–659.

[24] D. Filipović and M. Larsson (2016). Polynomial diffusions and applications in finance. *Finance & Stoch.*, **20**, pp. 931–972.

[25] D. Filipović, M. Larsson and A. Trolle (2017). On the relation between linearity-generating processes and linear-rational models. *Swiss Finance Institute Research Paper*, No. 16–23. Available at SSRN: https://ssrn.com/abstract=2753484.

[26] X. Gabaix (2007). Linearity-generating processes: a modelling tool yielding closed forms for asset prices. *NBER Working paper*, 13430, September.

[27] H. Geman (2005). *Commodities and Commodity Derivatives*. John Wiley & Sons, Chichester.

[28] H. Geman and B. Liu (2015). Are world natural gas markets moving toward integration? Evidence from the Henry Hub and National Balancing Point forward curves. *J. Energy Markets*, **8**(2), pp. 47–65.

[29] D. Heath, R. Jarrow and A. Morton (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, **60**(1), pp. 77–105.

[30] I. Karatzas and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus*. Springer Verlag, New York.

[31] P. Kevei (2018). Asymptotic moving average representation of high-frequency sampling of multivariate CARMA processes. *Ann. Inst. Stat. Math.*, **70**(2), pp. 467–487.

[32] J. Lucia, and E. S. Schwartz (2002). Electricity prices and power derivatives: evidence from the Nordic Power Exchange. *Rev. Derivatives Res.*, **5**(1), pp. 5–50.

[33] T. Marquardt and R. Stelzer (2007). Multivariate CARMA processes. *Stoch. Process. Appl.*, **117**, pp. 96–120.

[34] K. Nakajima and K. Ohashi (2012). A cointegrated commodity pricing model. *J. Futures Markets*, **32**(11), pp. 995–1033.

[35] R. Paschke and M. Prokopczuk (2009). Integrating multiple commodities in a model of stochastic price dynamics. *J. Energy Markets*, **2**(3), pp. 47–68.

[36] R. Paschke and M. Prokopczuk (2010). Commodity derivatives valuation with autoregressive and moving average components in the price dynamics. *J. Banking & Finance*, **34**, pp. 2742–2752.

[37] S. Peszat and J. Zabczyk (2007). *Stochastic Partial Differential Equations with Lévy Noise*. Cambridge University Press, Cambridge.

[38] K.-I. Sato (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.

[39] E. Schlemm and R. Stelzer (2012). Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. *Bernoulli*, **18**, pp. 46–63.

[40] M. Tehranchi (2005). A note on invariant measures for HJM models. *Finance Stoch.*, **9**(3), pp. 389–398.

Fred Espen Benth, Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, N–0316 Oslo, Norway

E-mail address: `fredb@math.uio.no`

Andre Süss, Centre for Advanced Study, Drammensveien 78, N-0271 Oslo, Norway

E-mail address: `suess.andre@web.de`