GENERALIZED FELLER PROCESSES AND MARKOVIAN LIFTS OF STOCHASTIC VOLTERRA PROCESSES: THE AFFINE CASE

CHRISTA CUCHIERO AND JOSEF TEICHMANN

Abstract. We consider stochastic (partial) differential equations appearing as Markovian lifts of affine Volterra processes with jumps from the point of view of the generalized Feller property which was introduced in, e.g., [9]. In particular we provide new existence, uniqueness and approximation results for Markovian lifts of affine rough volatility models of general jump diffusion type. We demonstrate that in this Markovian light the theory of stochastic Volterra processes becomes almost classical.

1. Introduction

In the realm of the recent discovery of rough volatility models (see e.g., [17, 20, 15]), stochastic Volterra processes have been studied extensively, for instance in [11, 3, 12] and the references therein. It is well known that Markovian lifts of such processes exist, but it has rarely been detailed how these lifts look like in general, with the notable exception of [21] and the recent preprint [1]. If at all, these lifts have been constructed by solving first the stochastic Volterra equations and identifying a state space a posteriori, on which respective Markov properties show up. Neither maximum principles, nor strongly continuous semigroups, nor Feller properties, nor approximation properties, nor Kolmogorov backward equations have been considered in this context so far. The reason is of course that one enters the world of SPDEs where each of the previously mentioned concepts needs particular care and is often cumbersome due to the lack of local compactness of the underlying state spaces. Also the question whether all those considerations are restricted to Brownian driven processes has not yet been investigated due to technical difficulties in writing down (probabilistically) weak or martingale solution concepts with jump drivers that are not standard Poisson random measures.

With this article we aim to bridge this gap and introduce a functional analytic setting which clearly shows that it is an advantage to consider Markovian lifts right from the beginning without solving the Volterra equation in the first place. As a particular important example we deal with Volterra processes whose kernels are Laplace transforms of (signed) measures (see Section 5.1). Their Markovian lifts can be represented in terms of signed measure valued processes which can be treated in a systematic way comparable to Feller processes on locally compact state spaces. In particular, we demonstrate that many results in the realm of stochastic Volterra processes appear transparent and clearly structured in this framework. Indeed, thanks to the generalized Feller theory introduced in [9] many properties and theorems can be formulated abstractly while in the literature so far often rather concrete and example-like specifications have been considered. The passage to these generalized Feller processes is necessary, since no standard Feller formulation is

2010 Mathematics Subject Classification. 60H15, 60J25.

Key words and phrases. variation of constants formula, stochastic partial differential equations, affine processes, stochastic Volterra processes, rough volatility models.

The authors are grateful for the support of the ETH Foundation, and the first author thanks the Forschungsinstitut für Mathematik for its support of a research stay at ETH Zürich in fall 2016.
available due to the lack of local compactness, as demonstrated in Example 5.1. This is in contrast to superprocesses, see, e.g. [8], which take values in spaces of nonnegative measures which are locally compact when the underlying space is compact.

We do also believe that the presented theory of generalized Feller semigroups and processes is a natural framework for analyzing SPDEs valid also far beyond Markovian lifts of stochastic Volterra processes: it opens the door to treat for instance large deviations, long term behaviors, or asymptotic results in a generic framework.

This first article takes as guiding example affine Volterra processes on $\mathbb{R}_+$ of the form

\begin{equation}
V_t = h(t) + \int_0^t K(t - s)dX_s,
\end{equation}

where $h$ is a deterministic function, $K$ a deterministic kernel and $X$ a semimartingale whose characteristics depend linearly on the state of the Volterra process. In a subsequent article the general (non-affine) case, including properly formulated martingale problems and strong existence results, as well as other versions of Markovian lifts will be considered. Here we focus on Markovian lifts where the kernel $K$ can be represented by

\begin{equation}
K(t) = \langle g, S_t^* \nu \rangle
\end{equation}

with $(S_t^*)_{t \geq 0}$ a strongly continuous semigroup acting on a Banach space $Y^*$, $\nu \in Y^*$ (or in a slightly bigger space), $g \in Y$ with $Y$ the pre-dual of $Y^*$ and pairing $\langle \cdot, \cdot \rangle$. This abstract setting includes the above mentioned signed measure valued lifts as well as what we call “forward process lifts”, where a variant of the latter has also been considered in [1] in the special case of the Volterra Heston model. This richness already indicates that our almost axiomatic approach does not only simplify arguments but also unifies several concepts and branches of the literature. In particular, it provides due to its simple structure new existence and uniqueness results for (1.1), and leads to so far unknown approximation schemes of any order. For instance it will be easy to construct higher order weak approximation schemes like Ninomiya-Victoir schemes with precise and optimal convergence rates.

Inspired from Hawkes processes, see e.g. [19, 6], and in contrast to some recent literature, e.g. [21, 9, 2, 1], we do not consider the Brownian motion driven case as the simplest one but rather the case when the stochastic driver is a finite activity jump process (as it is often true and useful in the theory of stochastic processes): Brownian driven stochastic Volterra equations or stochastic Volterra equations with more complicated jump structures are then easy and well described limits (in the sense of generalized Feller processes) of processes with finitely many jumps on compact time intervals, from which then weak solutions can be constructed. Another difference to most of the literature is that we do not transform the processes in question into semimartingales to use stochastic calculus but we rather work directly with their Markovian structure. We also want to point out that with representation (1.2) we go well beyond standard assumptions on Volterra kernels like complete monotonicity, in particular we do not need resolvents of the first kind which do not always exist.

Let us outline in the sequel one of our lines of ideas in the above addressed case of kernels that are Laplace transforms of signed measures: we consider the vector space of finite, signed Borel measures $Y^* := \mathcal{M}(\mathbb{R}_+)$ with its weak-$*$-topology, see [23] or any other text book on functional analysis, and look at the following simple linear homogenous equation thereon

\begin{equation}
d\lambda_t(dx) = -x\lambda_t(dx)dt - w\nu(dx)\lambda_t(\mathbb{R}_+)dt,
\end{equation}
where the initial value $\lambda_0$ is a signed measure, $w \geq 0$ a constant and $\nu$ a signed measure such that its Laplace transform

$$\int_0^\infty \exp(-tx)\nu(dx) \geq 0$$

is finite and non-negative for all $t > 0$, and belongs to $L^2_{\text{loc}}(\mathbb{R}_+, \mathbb{R}_+)$. The representation of the kernel in (1.2) is here

$$(1.3) \quad K(t) = \langle g, S_t^* \nu \rangle = \langle 1, \exp(-t \cdot) \nu \rangle = \int_0^\infty \exp(-tx)\nu(dx)$$

i.e. the evaluation of the element $\exp(-t \cdot) \nu$ in the dual space $\mathcal{M}(\mathbb{R}_+)$ of $\mathcal{Y} = C_0(\mathbb{R}_+, \mathbb{R})$ on the constant function 1. Under these circumstances one can easily prove two facts: first the solution defines a generalized Feller process (see [9] and the subsequent sections for details on this notion) on the state space of all signed measures, and, second, there is a closed invariant subspace $\mathcal{E}$ consisting of initial measures $\lambda_0$ where the total mass $\langle 1, \lambda_t \rangle = \lambda_t(\mathbb{R}_+)$ remains non-negative along its evolution in time for all $w \geq 0$. Of course the total mass satisfies a one-dimensional linear Volterra equation, whence the set $\mathcal{E}$ of measures $\lambda_0$ can be characterized by resolvents, see Section 4 and Section 5.1 for all details. Next we construct on this very state space $\mathcal{E}$ all sorts of jump diffusion processes with general characteristics by Poisson approximations applying an approximation theorem for generalized Feller semigroups (see Theorem 3.2). This is parallel to classical work on affine processes, see [10], where general existence is obtained by pure jump approximations.

We can finally show that the SPDE

$$(1.4) \quad d\lambda_t(dx) = -x\lambda_t(dx)dt + \nu(dx)dX_t$$

for all semimartingales $X$ whose characteristics depend linearly on $\lambda_t(\mathbb{R}_+)$ admits an analytically mild and probabilistically weak solution and that the associated Cauchy problem has a solution given in terms of a generalized Feller semigroup acting on a well specified set of functions. In particular, by the variation of constants formula and (1.2) it is easily seen that the total mass $\langle 1, \lambda_t \rangle = \lambda_t(\mathbb{R}_+)$ solves (1.1) with $h(t) = \int_0^\infty \exp(-tx)\lambda_0(dx)$.

This previous example highlights our results which can be summarized as follows:

- we provide a full solution theory for univariate stochastic Volterra equations (1.1) driven by jump-diffusions. In particular for a large class of functionals we can provide semigroup like expectation operations and solutions of the respective Kolmogorov equations.
- we provide approximation schemes, (generalized) maximum principles (in certain cases), weak solutions concepts for Markovian lifts of jump-diffusion driven stochastic Volterra equations (1.1).
- we provide a solution theory of the corresponding generalized Riccati differential equations on the pre-dual space $\mathcal{E}^*$ and the respective affine transform formulas.
- we can go beyond standard assumptions on kernels of stochastic Volterra equations such as complete monotonicity.
- we provide an abstract theory of Markovian lifts which is not restricted to the above introduced univariate case. Indeed, by considering as $\mathcal{Y}^*$ the space of $\mathbb{R}^d$ valued measures on $E$ for some locally compact space $E$, (1.4) can be generalized to the following multivariate case

$$d\lambda^l_t(dx) = -f^l(x)\lambda^l_t(dx)dt + \sum_{j=1}^d \nu^l_j(dx)dX^j_t,$$
for \( l \in \{1, \ldots, d\} \) in a straightforward way. This is due to the completely general character of Proposition 3.3. Here, \( f^l \) are functions from \( E \rightarrow \mathbb{R}_+ \) satisfying appropriate conditions, and \( X^j \) are semi-martingales whose characteristics depend linearly on \( \lambda \). Note that this allows for Volterra process specifications of the form

\[
V^l_t = h^l(t) + \int_0^t \sum_{j=1}^d K^l_j(t-s)dX^j_s,
\]

for kernels \( K^l_j \) given by \( K^l_j(t) = \int_E g^l(x)e^{-f^l(x)(t)}\nu_j(dx) \) for some bounded continuous \( g^l \in C_b(E; \mathbb{R}) \).

- Heston-like or Barndorff-Nielsen-Shephard-like structures can be analogously constructed from the abstract setting again by choosing appropriate Banach spaces \( Y \). For instance the Heston case can be treated with \( Y = C_b(\mathbb{R}_+; \mathbb{R}^2) \) and the two equations

\[
d\lambda^1_t(dx) = -x\lambda^1_t(dx)dt + \nu(dx)\sqrt{\langle 1, \lambda^1_t \rangle}dB^1_t
\]

and

\[
d\lambda^2_t(dx) = -x\lambda^2_t(dx)dt + \delta_0(dx)\sqrt{\langle 1, \lambda^1_t \rangle}dB^2_t
\]

with a correlated Brownian motions \( B^1 \) and \( B^2 \), i.e. \( dB^1_t, dB^2_t = \rho dt \) as respective limits of finite activity processes.

- Covariance matrix (or cone-) valued processes can be constructed by considering symmetric matrix (or cone-) valued measures as respective state space. Here particular geometric restrictions appear in the diffusion driven case which will be considered in subsequent work.

The remainder of the article is as follows: in Section 1.1 we introduce some notation and review certain functional analytic concepts. In Section 2 we deal with generalized Feller processes as introduced in, e.g., [9]. In the following Section 3 we show simple approximation theorems for generalized Feller semigroups beyond standard Trotter-Kato type theorems, and we prove a crucial existence result for pure jump processes with finite but unbounded intensity. In Sections 4, 5.1 and 5.2 we apply the presented theory to SPDEs which are lifts of affine Volterra processes.

1.1. Notation and some functional analytic notions. For the background in functional analysis we refer to the excellent textbook [23] as main reference and to the equally excellent books [13, 22] for the background in strongly continuous semigroups. We emphasize however that only very basic knowledge in functional analysis and strongly continuous semigroups is required.

We shall often apply the following notations: let \( Y \) be a Banach space and \( Y^* \) its dual space, i.e. the space of linear continuous functionals with the strong dual norm

\[
\|\lambda\|_{Y^*} = \sup_{\|y\| \leq 1} |\langle y, \lambda \rangle|,
\]

where \( \langle y, \lambda \rangle := \lambda(y) \) denotes the evaluation of the linear functional \( \lambda \) at the point \( y \in Y \). Since cones \( \mathcal{E} \) of \( Y^* \) will be our statespaces, we denote the polar cones in pre-dual notation, i.e.

\[
\mathcal{E}_* = \{y \in Y \mid \langle y, \lambda \rangle \leq 0 \text{ for all } \lambda \in \mathcal{E}\}.
\]

We denote spaces of bounded linear operators from Banach spaces \( Y_1 \) to \( Y_2 \) by \( L(Y_1, Y_2) \) with norm

\[
\|A\|_{L(Y_1, Y_2)} := \sup_{\|y\|_{Y_1} \leq 1} \|Ay\|_{Y_2}.
\]
If $Y_1 = Y_2$ we only write $\| \cdot \|_{L(Y_1)}$. On $Y^*$ we shall usually consider beside the strong topology (induced by the strong dual norm) the weak-\$-topology, which is the weakest locally convex topology making all linear functionals $\langle y, \cdot \rangle$ on $Y^*$ continuous. Let us recall the following facts:

- The weak-\$-topology is metrizable if and only if $Y$ is finite dimensional: this is due to Baire’s category theorem since $Y^*$ can be written as a countable union of closed sets, whence at least one has to contain an open set, which in turn means that compact neighborhoods exist, i.e. a strictly finite dimensional phenomenon.
- Norm balls $K_R$ of any radius $R$ in $Y^*$ are compact with respect to the weak-\$-topology, which is the Banach-Alaoglu theorem.
- These balls are metrizable if and only if $Y$ is separable: this is true since $Y$ can be isometrically embedded into $C(K_1)$, where $y \mapsto \langle y, \cdot \rangle$, for $y \in Y$. Since $Y$ is separable, its embedded image is separable, too, which means – by looking at the algebra generated by $Y$ in $C(K_1)$ – that $C(K_1)$ is separable, which is the case if and only if $K_1$ is metrizable.

Since we do not need metrizability of $K_R$, we do not assume that our Banach spaces $Y$ are separable.

A family of linear operators $(P_t)_{t \geq 0}$ on a Banach space $Y$ with $P_tP_s = P_{t+s}$ for $s, t \geq 0$ and with $P_0 = I$ where $I$ denotes the identity is called strongly continuous semigroup if $\lim_{t \to 0} P_t y = y$ holds true for every $y \in Y$. We denote its generator usually by $A$ which is defined as $\lim_{t \to 0} \frac{P_t y - y}{t}$ for all $y \in \text{dom}(A)$, i.e. the set of elements where the limit exists. Notice that $\text{dom}(A)$ is left invariant by the semigroup $P$ and that its restriction on the domain equipped with the operator norm

$$\|y\|_{\text{dom}(A)} := \sqrt{\|y\|^2 + \|Ay\|^2}$$

is again a strongly continuous semigroup.

Moreover, as already used in the introduction, $\mathcal{M}(E)$ denotes the space of signed finite measures on $E$ with respect to the Borel $\sigma$-algebra and $\mathcal{M}_+(E)$ the space of finite nonnegative measures.

## 2. Generalized Feller Semigroups and Processes

Feller semigroups have proved to be useful in the context of locally compact state spaces. Generalized Feller semigroups serve the same purpose on state spaces which are not locally compact, which is a typical infinite dimensional phenomenon. Local compactness is replaced by adjoining a proper weight function to a state space $X$, which measures explosion. Interestingly even in the locally compact case new results relevant for the theory of affine processes beyond canonical state spaces, see [10], appear. We refer to [9] for all necessary details and proofs, and also to the references therein. Notice that even in the infinite dimensional setting there are notable examples of local compactness, see Remark [1.4] but we point out that our cases are not locally compact in their natural topology, see Example [5.1].

### 2.1. Definitions and results

First we introduce weighted spaces and state a central Riesz representation result. The underlying space $X$ here is a completely regular Hausdorff topological space.

**Definition 2.1.** A function $\varrho : X \to (0, \infty)$ is called admissible weight function if the sets $K_R := \{x \in X : \varrho(x) \leq R\}$ are compact for all $R > 0$.

An admissible weight function $\varrho$ is necessarily lower semicontinuous and bounded from below. We call the pair $X$ together with an admissible weight function $\varrho$ a weighted space. A weighted space is $\sigma$-compact.
For completeness we shall put definitions for general Banach space valued functions, although in the sequel we shall only deal with \(\mathbb{R}\)-valued functions: let \(Z\) be a Banach space with norm \(\|\cdot\|_Z\). The vector space

\[
B^0(X; Z) := \left\{ f : X \to Z : \sup_{x \in X} \varrho(x)^{-1} \|f(x)\|_Z < \infty \right\}
\]

of \(Z\)-valued functions \(f\) equipped with the norm

\[
\|f\|_\varrho := \sup_{x \in X} \varrho(x)^{-1} \|f(x)\|_Z,
\]

is a Banach space itself. It is also clear that for \(Z\)-valued bounded continuous functions the continuous embedding \(C_b(X; Z) \subset B^0(X; Z)\) holds true, where we consider the supremum norm on bounded continuous functions, i.e. \(\sup_{x \in X} \|f(x)\|\).

**Definition 2.2.** We define \(B^0(X; Z)\) as the closure of \(C_b(X; Z)\) in \(B^0(X; Z)\). The normed space \(B^0(X; Z)\) is a Banach space.

If the range space \(Z = \mathbb{R}\), which from now on will be the case, we shall write \(B^0(X)\) for \(B^0(X; \mathbb{R})\) and analogously \(B^0(X)\).

We consider elements of \(B^0(X)\) as continuous functions whose growth is controlled by \(\varrho\). More precisely we have by [9, Theorem 2.7] that \(f \in B^0(X)\) if and only if \(f|_{K_R} \in C(K_R)\) for all \(R > 0\) and

\[
\lim_{R \to \infty} \sup_{x \in X \setminus K_R} \varrho(x)^{-1} \|f(x)\| = 0.
\]

Additionally, by [9, Theorem 2.8] it holds that for every \(f \in B^0(X)\) with \(\sup_{x \in X} f(x) > 0\), there exists \(z \in X\) such that

\[
\varrho(x)^{-1} f(x) \leq \varrho(z)^{-1} f(z) \quad \text{for all } x \in X,
\]

which emphasizes the analogy with spaces of continuous functions vanishing at \(\infty\) on locally compact spaces.

Let us now state the following crucial representation theorem of Riesz type:

**Theorem 2.3** (Riesz representation for \(B^0(X)\)). For every continuous linear functional \(\ell : B^0(X) \to \mathbb{R}\) there exists a finite signed Radon measure \(\mu\) on \(X\) such that

\[
\ell(f) = \int_X f(x) \mu(dx) \quad \text{for all } f \in B^0(X).
\]

Additionally

\[
\int_X \varrho(x)|\mu|(dx) = \|\ell\|_{L(B^0(X), \mathbb{R})},
\]

where \(|\mu|\) denotes the total variation measure of \(\mu\). \(\mathbb{R}\).

We shall next consider strongly continuous semigroups on \(B^0(X)\) spaces and recover very similar structures as well known for Feller semigroups on the space of continuous functions vanishing at \(\infty\) on locally compact spaces.

**Definition 2.4.** A family of bounded linear operators \(P_t : B^0(X) \to B^0(X)\) for \(t \geq 0\) is called generalized Feller semigroup if

- (i) \(P_0 = I\), the identity on \(B^0(X)\),
- (ii) \(P_{t+s} = P_t P_s\) for all \(t, s \geq 0\),
- (iii) for all \(f \in B^0(X)\) and \(x \in X\), \(\lim_{t \to 0^+} P_t f(x) = f(x)\),
- (iv) there exist a constant \(C \in \mathbb{R}\) and \(\varepsilon > 0\) such that for all \(t \in [0, \varepsilon]\),
  \[\|P_t\|_{L(B^0(X))} \leq C.\]
- (v) \(P_t\) is positive for all \(t \geq 0\), that is, for \(f \in B^0(X), f \geq 0\), we have \(P_t f \geq 0\).

We obtain due to the Riesz representation property the following key theorem:
Theorem 2.5. Let \((P_t)_{t \geq 0}\) satisfy (i) to (iv) of Definition 2.4. Then, \((P_t)_{t \geq 0}\) is strongly continuous on \(B^\theta(X)\), that is,

\[
\lim_{t \to 0^+} \|P_t f - f\|_\theta = 0 \quad \text{for all } f \in B^\theta(X) .
\]

One can also establish a positive maximum principle in case that the semigroup \(P_t\) grows around 0 like \(\exp(\omega t)\) for some \(\omega \in \mathbb{R}\) with respect to the operator norm on \(B^\theta(X)\). Indeed, the following theorem proved in [9, Theorem 3.3] is a reformulation of the Lumer-Philips theorem for pseudo-contraction semigroups using a generalized positive maximum principle which is formulated in the sequel.

Theorem 2.6. Let \(A\) be an operator on \(B^\theta(X)\) with domain \(D\), and \(\omega \in \mathbb{R}\). \(A\) is closable with its closure \(\overline{A}\) generating a generalized Feller semigroup \((P_t)_{t \geq 0}\) with \(\|P_t\|_{L(B^\theta(X))} \leq \exp(\omega t)\) for all \(t \geq 0\) if and only if

(i) \(D\) is dense,

(ii) \(A - \omega I\) has dense image for some \(\omega_0 > \omega\), and

(iii) \(A\) satisfies the generalized positive maximum principle, that is, for \(f \in D\) with \((\rho^{-1} f) \vee 0 \leq \rho(z)^{-1} f(z)\) for some \(z \in X\), \(A f(z) \leq \omega f(z)\).

The following example provides the simplest generalized Feller process being of pure jump type with a bounded generator.

Example 2.7. Let \((X, \rho)\) be a weighted space and \(\mu(x, \cdot)\) for \(x \in X\) be a probability kernel on \(X\), i.e. measurable with respect to both variables, such that

\[
\int \rho(y) \mu(x, dy) \leq M \rho(x), \quad \text{for all } x \in X.
\]

for some constant \(M\). Define now an operator \(A\) acting on \(B^\theta(X)\) by

\[
Af(x) := \int (f(y) - f(x)) \mu(x, dy).
\]

Then \(A\) generates a generalized Feller semigroup. Indeed,

\[
\frac{A f(x)}{\rho(x)} = \frac{\int \frac{\rho(y)}{\rho(x)} \rho(y) \mu(x, dy) - f(x)}{\rho(x)}
\]

yields that \(\|A f\|_\theta \leq \|f\|_\theta (M + 1)\) for all measurable functions \(f : X \to \mathbb{R}\) such that \(\|f\|_\theta < \infty\). Hence, if \(A : B^\theta(X) \to B^\theta(X)\), then the operator \(A\) is bounded and generates a strongly (even uniformly) continuous semigroup on \(B^\theta(X)\), implying conditions (i) to (iv) of Definition 2.4. Since \(A\) satisfies the classical maximum principle, the semigroup is positive and whence \(A\) generates a generalized Feller semigroup.

If \(A\) does not map \(B^\theta(X)\) to itself, we can instead look at the space of all measurable functions in \(B^\theta(X)\), i.e. the vector space of all measurable functions \(f : X \to \mathbb{R}\) such that \(\|f\|_\theta < \infty\), where \(A\) again defines a positive, strongly continuous semigroup by the same argument. In both cases we obtain a Markov process on \(X\) generated by \(A\).

Note that \(A\) also satisfies the generalized positive maximum principle on \(B^\theta(X)\). This allows to transfer it to all limits of operators of type \(A\) (acting on the same dense set of \(B^\theta(X)\) and satisfying the maximum principle with the same constant \(\omega\)) so that actually a large class of operators satisfies the generalized positive maximum principle.

The next theorem is a perturbation result in the spirit of [14, Corollary 7.2] but without the assumption that the semigroup is pseudocontractive. This will be needed later for our constructions and is in contrast to classical Feller theory where one always stays in the setting of contractive semigroups. The proof follows [22, Chapter 3, Theorem 1.1].
Theorem 2.8. Let \( A \) be a generator of a generalized Feller semigroup \((P_t)_{t \geq 0}\) on \( B^\theta(X) \) with \( \|P_t\|_{L(B^\theta(X))} \leq M \exp(\omega t) \) for all \( t \geq 0 \) and some \( M \geq 1, \omega \in \mathbb{R} \). Let \( B \) be a bounded generator of a generalized Feller semigroup on \( B^\theta(X) \). Then \( A + B \) is a generator of a generalized Feller semigroup on \( B^\theta(X) \) whose operator norm is bounded by \( M \exp(\omega t + M\|B\|_{L(B^\theta(X))}t) \) for \( t \geq 0 \). If \( M = 1 \), then the semigroup generated by \( A + B \) satisfies the assumptions of Theorem 2.7 and the generalized maximum principle with constant \( \omega + \|B\|_{L(B^\theta(X))} \).

Proof. By standard semigroup perturbation theory we obtain that \( A + B \) is actually a generator of a strongly continuous semigroup \( \tilde{P} \), see [22, Chapter 3, Theorem 1.1]: indeed we can choose an equivalent norm \( \| \cdot \| \) on \( B^\theta(X) \) such that \( \tilde{P}_t \) is pseudocontractive, i.e. \( |P_t f| \leq \exp(\omega t)|f| \) and \( \|f\|_{\infty} \leq |f| \leq M\|f\|_{\infty} \) for \( f \in B^\theta(X) \). Then we can conclude by a standard Chernoff product formula that
\[
\tilde{P}_t f = \lim_{n \to \infty} (P_{t/n} \exp(Bt/n))^n f
\]
for all \( f \in B^\theta(X) \) by stability. Whence the result and the growth bound
\[
|\tilde{P}_t| \leq \exp((\omega + |B|)t)
\]
for \( t \geq 0 \). Expressing this formula in the original norm yields the assertion. \( \square \)

Remark 2.9. The semigroup \( \tilde{P} \) constructed in the proof of Theorem 2.8 is the unique solution of the variation of constants equations
\[
\tilde{P}_t f = P_t f + \int_0^t P_{t-s} B \tilde{P}_s f ds
\]
for \( t \geq 0 \) by basic perturbation theory as stated in [22, Chapter 3, Proposition 1.2].

If furthermore \( |B g(x)| \leq \tilde{\omega} g(x) \), for \( x \in X \), and for some \( \tilde{\omega} \), then the growth bound can be given in terms of \( \tilde{\omega} \). Indeed, the iteration scheme \( V_t^{(n)} f = P_t f \) and
\[
V_t^{(n+1)} f = \int_0^t P_{t-s} B V_s^{(n)} f ds
\]
for \( t \geq 0 \) and \( f \in B^\theta(X) \) yields a uniformly on compacts and bounded sets converging series \( \sum_{n \geq 0} V_t^{(n)} f \), for \( f \in B^\theta(X) \), with limit \( \tilde{P}_t f \), the same holds true for \( g \) itself (see [22, Chapter 3, proof of Proposition 1.2]). Similarly as in [22, Chapter 3, Equation (1.6)], we obtain by induction
\[
V_t^{(n)} g(x) \leq M \exp(\omega t) \frac{M^n \tilde{\omega}^{n+1} n!}{n!} g(x)
\]
for \( t \geq 0 \) and for \( x \in X \). This means
\[
\|\tilde{P}_t\|_{L(B^\theta(X))} \leq M \exp(\omega t + M\tilde{\omega} t)
\]
for \( t \geq 0 \).

Remark 2.10. Limits of generators of the additive form as of Theorem 2.8 are very general: in a stochastic setting \( A \) can for instance be of transport type (see e.g. Section 5.2 generating a generalized Feller semigroup (corresponding to a deterministic process) and \( B \) is of (bounded) pure jump type similarly as in Example 2.7). Considering limits one can obtain 'almost all' operators of Lévy-Khintchine form.

We now give a version of the Kolmogorov extension theorem, which has not be shown so far. Notice that standard versions of the extension theorem, i.e. for Borel spaces or in case of inner-regular measures on topological spaces, see [4], do not apply in this case. Still the proof is quite standard, see [5].
Theorem 2.11. Let \((P_t)_{t \geq 0}\) be a generalized Feller semigroup with \(P_{t1} = 1\) and \(P_{t0} \leq M \exp(\omega t)g\) for all \(t \geq 0\). Then there exists a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a family of random variables \((\lambda_t)_{t \geq 0}\) for any initial value \(\lambda_0 \in X\) such that

\[
\mathbb{E}_{\lambda_0}[f(\lambda_t)] = P_t f(\lambda_0)
\]

for \(t \geq 0\) and every \(f \in \mathcal{B}^0(X)\).

Proof. We choose a functional analytic way for the proof and split it in two steps: first we prove the statement for a countable subset of points in time, in a second step we lift this result to the general case.

\[
\lambda_0 \in X \text{ and let } I \subseteq \mathbb{R}_+ \text{ be countable. By standard constructions we obtain a consistent family of probability measures } \mu^J \text{ on } X^J \text{ for all finite subsets } J \subseteq I, \text{ i.e. for the canonical projections } \pi^{J \to J_i} : X^J \to X^{J_i} \text{ the push forward of } \mu^J \text{ is } \mu^{J_i}, \text{ for all finite sets } J_1 \subseteq J_2 \subseteq I. \text{ The topological space }
\]

\[
\ell^1(X; I) := \{x = (x_i)_{i \in I} \in X^I \text{ such that } \tilde{g}(x) < \infty \},
\]

where \(\tilde{g}(x) := \sum_i M^{-1} \exp(-\omega_i - i)g(x_i)\), is \(\sigma\)-compact by construction and together with \(\tilde{g}\) a weighted space. Notice here that \(\tilde{g}\) is lower semicontinuous on \(X\). Denote by \(D = \text{Cyl}(\ell^1(X; I)) \subseteq \mathcal{B}^0(\ell^1(X; I))\) the point-separating, dense subalgebra of bounded cylindrical functions on \(\ell^1(X, I)\). Take \(f \in D\), then \(f = g \circ \pi^{J_\to J}\) for some bounded continuous \(g : X^J \to \mathbb{R}\) and we define a linear functional \(k : D \to \mathbb{R}\)

\[
k(f) := \int_{X^J} g(x) \mu^J(dx).
\]

By consistency this is well defined and by

\[
|k(f)| \leq C \|f\|_{\tilde{g}}
\]

for a constant \(C\) independent of the cardinality of \(J\) a bounded linear functional. Hence we obtain by the Riesz representation (see Theorem 2.3) a probability measure \(\mu^f\) on \(\ell^1(X, I)\), which satisfies

\[
\int_X g(\pi^{J \to \{t\}}(x)) \mu^J(dx) = P_t(g)(\lambda_0)
\]

by construction for all \(g \in \mathcal{B}^0(X)\). For the second step, we define now \(\Omega = X^{\mathbb{R}_+}\) equipped with the (Baire) \(\sigma\)-algebra \(\mathcal{F}\) generated by projections \(\Omega \to \ell^1(X, I)\) for all possible countable \(I \subseteq \mathbb{R}_+\) and obtain thereon a probability measure \(\mathbb{P}\) which satisfies the desired (2.8). Finally define \(\lambda_t : \Omega \to X\) as projection on the \(t\)-th coordinate, for \(t \geq 0\).

We end this subsection with a classical theorem on path properties which is well known for Feller semigroups.

Theorem 2.12. Let \((P_t)_{t \geq 0}\) be a generalized Feller semigroup and let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space together with a family of random variables \((\lambda_t)_{t \geq 0}\) chosen for a fixed initial value \(\lambda_0 \in X\) such that

\[
\mathbb{E}_{\lambda_0}[f(\lambda_t)] = P_t f(\lambda_0)
\]

for \(t \geq 0\) and every \(f \in \mathcal{B}^0(X)\). For every countable family \((f_n)_{n \geq 0}\) of functions in \(\mathcal{B}^0(X)\) we can choose a version of the processes \((f_n(\lambda_t))_{t \geq 0}\) such that the trajectories are càdlàg for all \(n \geq 0\).

Remark 2.13. Notice that we do not speak – in this general context – about a version of \((\lambda_t)_{t \geq 0}\) itself. In case of separability \(\mathcal{B}^0(X)\) we can actually find a càdlàg version of \(\lambda\) itself by choosing as point separating family \(f_n\) a countable dense subset of \(\mathcal{B}^0(X)\).
Proof. Notice that for every generalized Feller semigroup the process
\[ Y_t^\alpha := \exp(-at) R(\alpha)f(\lambda_t) \]
for \( t \geq 0 \) is actually a nonnegative supermartingale for \( f \geq 0 \) in \( B^0(X) \). Here
\[ R(\alpha) := \int_0^\infty \exp(-as) P_s ds \]
denotes the resolvent of \( (P_t)_{t \geq 0}, \) defined for large enough \( \alpha \), say \( \alpha > \omega \) for some \( \omega \geq 0 \). Indeed by the Markov property
\[
E[Y_t^\alpha|F_s] = E\left[\exp(-at) \int_s^\infty \exp(-au) P_u f(\lambda_t) du \right|F_s]
\]
\[
= \exp(-at) \int_0^\infty \exp(-au) P_{s+t-u} f(\lambda_s) du
\]
\[
= \exp(-as) \int_s^\infty \exp(-au) P_u f(\lambda_s) du
\]
\[ \leq Y_s^\alpha, \]
for \( 0 \leq s \leq t \), where \( (F_t)_{t \geq 0} \) denotes the natural filtration of \( (\lambda_t)_{t \geq 0} \) made right continuous. Hence \( (Y_t^\alpha)_{t \geq 0} \) is a supermartingale and \( t \mapsto E[Y_t^\alpha] \) is continuous. This implies together with the right continuity of the filtration the existence of a càglàd version. Consider now the countable set of functions
\[ \mathcal{H} := \{aR(\alpha) f_n \mid n \in \mathbb{N}, \alpha > \omega \}. \]
Since \( (Y_t^\alpha)_{t \geq 0} \) has a càglàd version, this translates also to each of the processes \( (h(\lambda_t))_{t \geq 0} \) for \( h \in \mathcal{H} \). Moreover, as \( \|aR(\alpha) f - f\|_0 \to 0 \) for \( \alpha \to \infty \), also \( (f_n(\lambda_t))_{t \geq 0} \) has càglàd trajectories. \( \square \)

Remark 2.14. This theorem allows to formulate a martingale problem in this general context. Indeed, let \( A \) be the generator of a generalized Feller semigroup. Then, for every \( f \in \text{dom}(A) \) we can choose versions for \( (f(\lambda_t))_{t \geq 0} \) and \( (Af(\lambda_t))_{t \geq 0} \) which are càglàd. Hence \( \int_0^t Af(\lambda_s) ds \) is in particular well-defined and the Markov property implies that
\[
\left( f(\lambda_t) - \int_0^t Af(\lambda_s) ds \right)_{t \geq 0}
\]
is actually a (càglàd) martingale. This will be investigated further in a subsequent article.

2.2. Dual spaces of Banach spaces. The most important playground for our theory will be closed subsets of duals of Banach spaces, where the weak++-topology appears to be \( \sigma \)-compact due to the Banach-Alaoglu theorem. Assume that \( \mathcal{E} \subset Y^* \) is a closed subset of the dual space \( Y^* \) of some Banach space \( Y \) where \( Y^* \) is equipped with its weak++-topology. Consider a lower semicontinuous function \( g: \mathcal{E} \to (0, \infty) \) and denote by \( (\mathcal{E}, g) \) the corresponding weighted space. We have the following approximation result (see [9, Theorem 4.2]) for functions in \( B^0(\mathcal{E}) \) by cylindrical functions. Set
\[
\mathcal{Z}_N := \{g(\langle \cdot, y_1 \rangle, \ldots, \langle \cdot, y_N \rangle) : g \in C_0^\infty(\mathbb{R}^N) \}
\]
and \( y_j \in Y, j = 1, \ldots, N \),
\[
\text{where } \langle \cdot, \cdot \rangle \text{ denotes the pairing between } Y^* \text{ and } Y. \text{ We denote by } \mathcal{Z} := \bigcup_{N \in \mathbb{N}} \mathcal{Z}_N \text{ the set of bounded smooth continuous cylinder functions on } \mathcal{E}. \text{ We can prove the following theorem beyond any separability assumptions on } Y. \]
Theorem 2.15. The closure of $\mathcal{Z}$ in $B^{\theta}(\mathcal{E})$ coincides with $B^{\theta}(\mathcal{E})$, whose elements appear to precisely the functions $f \in B^{\theta}(\mathcal{E})$ which satisfy (2.3) and that $f|_{K_{R}}$ is weak-$*$-continuous for any $R > 0$.

Proof. Since $\mathcal{Z}|_{|\rho|\leq R}$ is a point separating algebra we can apply the Stone Weierstrass theorem on the compact sets $\{\rho \leq R\}$ to obtain density of the restrictions in $C_{b}(\{\rho \leq R\})$ for any $R \geq 0$. Then we can apply [9] Theorem 2.7.

Remark 2.16. Of course we can consider also subsets of general cylindrical functions to serve the same purpose (we just need a Stone-Weierstrass theorem to be applicable), i.e. the subset should be point separating and an algebra. This will play an important role in the case of affine processes where we can consider the linear span of all products of Fourier basis elements $\exp(\langle \cdot, y \rangle)$ for $y \in Y$.

In the following we will give a theorem telling when the semigroup of a Markov process is actually a generalized Feller semigroup. For the dense subset appearing in this theorem we take in practice the set of cylindrical function $\mathcal{Z}$ introduced above. For its formulation we need the following assumptions.

Assumption 2.17. Let $(\lambda_{t})_{t \geq 0}$ denote a time homogeneous Markov process on some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_{t})_{t \geq 0}, P)$ with values in $\mathcal{E}$.

Then we assume that

(i) there are constants $C$ and $\varepsilon > 0$ such that

\[ (2.10) \quad E_{\lambda_{0}}[\theta(\lambda_{t})] \leq C\theta(\lambda_{0}) \quad \text{for all } \lambda_{0} \in \mathcal{E} \text{ and } t \in [0, \varepsilon]. \]

(ii) \[ (2.11) \quad \lim_{\varepsilon \to 0^{+}} E_{\lambda_{0}}[f(\lambda_{t})] = f(\lambda_{0}) \quad \text{for any } f \in B^{\theta}(\mathcal{E}) \text{ and } \lambda_{0} \in \mathcal{E}. \]

(iii) for all $f$ in a dense subset of $B^{\theta}(\mathcal{E})$, the map $\lambda_{0} \mapsto E_{\lambda_{0}}[f(\lambda_{t})]$ lies in $B^{\theta}(\mathcal{E})$.

Remark 2.18. Of course inequality (2.10) implies that $|E_{\lambda_{0}}[f(\lambda_{t})]| \leq C\theta(\lambda_{0})$ for all $f \in B^{\theta}(\mathcal{E})$, $\lambda_{0} \in \mathcal{E}$ and $t \in [0, \varepsilon]$.

Theorem 2.19. Suppose Assumptions (2.17) hold true. Then $P_{t}f(\lambda_{0}) := E_{\lambda_{0}}[f(\lambda_{t})]$ satisfies the generalized Feller property and is therefore a strongly continuous semigroup on $B^{\theta}(\mathcal{E})$.

Proof. This follows from the arguments of [9] Section 5. □

3. Approximation theorems

In order to establish existence of Markovian solutions for general generators $A$ we could either directly apply Theorem 2.6 where we have to assume that the generator $A$ satisfies on a dense domain $D$ a generalized positive maximum principle and that for at least one $\omega > 0$ the range of $A - \omega$ is dense, or we approximate a general generator $A$ by (pure jump) generators $A^{n}$ and apply the following (well known) approximation theorems:

Theorem 3.1. Let $(P_{t}^{n})_{n \in \mathbb{N}, t \geq 0}$ be a sequence of strongly continuous semigroups on a Banach space $Z$ with generators $(A^{n})_{n \in \mathbb{N}}$ such that there are uniform (in $n$) growth bounds $M \geq 1$ and $\omega \in \mathbb{R}$ with

\[ (3.1) \quad \|P_{t}^{n}\|_{L(Z)} \leq M \exp(\omega t) \]

for $t \geq 0$. Let furthermore $D \subset \cap_{n} \text{dom}(A^{n})$ be a dense subspace with the following three properties:
and by (3.1), (i), (iii) and (ii), we can estimate

We shall apply the following well known formula, see [22], for

Proof. This yields uniform convergence of

Theorem 3.2. Let \( (P^n_t)_{n \in \mathbb{N}, t \geq 0} \) be a sequence of strongly continuous semigroups on a Banach space \( Z \) with generators \( (A^n)_{n \in \mathbb{N}} \) such that there are uniform (in \( n \)) growth bounds \( M \geq 1 \) and \( \omega \in \mathbb{R} \) with

\[
\|P^n_t\|_{L(Z)} \leq M \exp(\omega t)
\]

for \( t \geq 0 \). Let furthermore \( D \subset \cap_n \text{dom}(A^n) \) be a subset with the following two properties:

(i) The linear span \( \text{span}(D) \) is dense.
(ii) There is a norm \( \| \cdot \|_D \) on \( \text{span}(D) \) such that for each \( f \in D \) and for \( t > 0 \) there exists a sequence \( a_{nm}^t \), possibly depending on \( f \),
\[
\| A^n P^m_t f - A^m P^m_t f \| \leq a_{nm}^t \| f \|_D
\]
holds true for \( n, m \) and for \( 0 \leq u \leq t \), with \( a_{nm}^t \to 0 \) as \( n, m \to \infty \).

Then there exists a strongly continuous semigroup \( (P^\infty_t)_{t \geq 0} \) with the same growth bound on \( Z \) such that \( \lim_{n \to \infty} P^n_t f = P^\infty_t f \) for all \( f \in Z \) uniformly on compacts in time. If in addition for each \( n \in \mathbb{N} \), \( (P^n_t)_{t \geq 0} \) is a generalized Feller semigroup, then this property transfers also to the limiting semigroup.

**Proof.** Again we apply the following well known formula for \( P^m_t f \) and \( P^n_t f \) operators.

\[
P^m_t f - P^n_t f = \int_0^t P^n_s (A^n - A^m) P^m_{t-s} f ds.
\]

By (3.1), (i), (iii) and (ii), we can estimate
\[
\| P^m_t f - P^n_t f \| \leq \int_0^t \| P^n_s (A^n - A^m) P^m_{t-s} f \| ds
\]
\[
\leq \int_0^t M \exp(\omega s) \| (A^n - A^m) P^m_{t-s} f \| ds
\]
\[
\leq \int_0^t M \exp(\omega s) a_{nm}^t \| f \|_D ds
\]

This yields uniform convergence of \( P^m_t f \) on compact intervals in time (with respect to the norm of \( Z \)) for all \( f \in D \) by assumption. Hence we obtain bounded linear operators \( P^\infty_t \) satisfying the semigroup property and the same growth bound on \( Z \) with constants \( \omega \) and \( M \), in particular \( \lim_{n \to \infty} P^n_t f = P^\infty_t f \) for all \( f \in Z \) by stability of growth bounds. Strong continuity follows by the respective stability estimates, i.e.
\[
\| P^\infty_t f - f \| \leq \| P^\infty_t f - P^\infty_t g \| + \| P^\infty_t g - P^n_t g \| + \| P^n_t g - g \| + \| g - f \|
\]
for \( g \in \text{span}(D) \), where the first term is small due to \( \| P^\infty_t f \| \leq M \exp(\omega t) \) for \( t \geq 0 \).

For generalized Feller semigroups, the only additional property is positivity, which of course remains in the limit. \( \square \)

Our first application of Theorem 3.1 is the next proposition that extends Example 2.7 towards unbounded limits.

**Proposition 3.3.** Let \((X, \rho)\) be a weighted space with weight function \( \rho \geq 1 \). Consider an operator \( A \) on \( \mathcal{B}^\rho(X) \) with dense domain \( \text{dom}(A) \) generating a generalized Feller semigroup \( (P_t)_{t \geq 0} \) on \( \mathcal{B}^\rho(X) \) which leaves \( \mathcal{B}^{\sqrt{\rho}}(X) \subset \mathcal{B}^\rho(X) \) invariant and satisfies \( \| P_t \|_{L(B^\rho(X))} \leq M \exp(\omega t) \) for some constants \( M \geq 1 \) and \( \omega \in \mathbb{R} \). Consider furthermore a family of finite measures \( \mu(x, \cdot) \) for \( x \in X \) on \( X \) with the following properties:

- For all \( x \in X \)
\[
(3.2) \quad \int \rho(y) \mu(x, dy) \leq M \rho^2(x),
\]
as well as
\[
(3.3) \quad \int \sqrt{\rho(y)} \mu(x, dy) \leq M \rho(x),
\]
and

\[ \int \mu(x, dy) \leq M \sqrt{\varrho(x)}, \]

hold true for some constant \( M \).

\[ \int \left| \frac{\varrho(y) - \varrho(x)}{\varrho(x)} \right| \mu(x, dy) \leq \tilde{\omega}, \]

for all \( x \in X \).

Define an operator \( B \) acting on \( \mathcal{B}^0(X) \) by

\[ Bf(x) := \int (f(y) - f(x))\mu(x, dy). \]

Then \( A + B \) generates a generalized Feller semigroup \( (P^n_t)_{t \geq 0} \) on \( \mathcal{B}^0(X) \) satisfying \( \| P^n_t \|_{L(\mathcal{B}^0(X))} \leq M \exp(\omega + M\tilde{\omega}t) \).

**Proof.** We apply Theorem 3.1 with \( D = \text{dom}(A) \cap \mathcal{B}^{\sqrt{\varrho}}(X) \). We construct a sequence of pure jump operators which generate by Example 2.7 generalized Feller semigroups on \( \mathcal{B}^0(X) \) and on \( \mathcal{B}^{\sqrt{\varrho}}(X) \):

\[ B^n f(x) := \int (f(y) - f(x)) \frac{n}{\varrho(x) \vee n} \mu(x, dy). \]

Indeed, they are bounded with respect to both norms: by (3.2) and (3.4), we have

\[ \frac{B^n f(x)}{\varrho(x)} \leq \int \frac{(f(y) - f(x))}{\varrho(x) \vee n} \mu(x, dy) \]

\[ \leq \| f \|_2 Mn \]

and by (3.3) and (3.4) for \( f \in \mathcal{B}^{\sqrt{\varrho}}(X) \)

\[ \frac{B^n f(x)}{\sqrt{\varrho(x)}} \leq \int \frac{\sqrt{\varrho(y) - \varrho(x)} - f(x) - f(y)}{\sqrt{\varrho(x) \vee n}} \mu(x, dy) \]

\[ \leq \| f \|_{\sqrt{\varrho}} Mn (\sqrt{n} + 1). \]

Consider now the operators \( A^n := A + B^n \), on the one hand with domain \( \text{dom}(A) \) and on the other hand with domain \( D = \text{dom}(A) \cap \mathcal{B}^{\sqrt{\varrho}}(X) \). By Theorem 2.3 they generate generalized Feller semigroups \( P^n \) on \( \mathcal{B}^0(X) \) and \( \mathcal{B}^{\sqrt{\varrho}}(X) \) respectively, the latter because \( (P_t) \) leaves \( \mathcal{B}^{\sqrt{\varrho}}(X) \) invariant by assumption. Since the domain of the generator is anyhow left invariant, Condition (i) of Theorem 3.1 is satisfied with \( D = \text{dom}(A) \cap \mathcal{B}^{\sqrt{\varrho}}(X) \). Moreover, by (3.3) and (3.4) we have for \( f \in D \)

\[ \left| \frac{B^n f(x) - B^m f(x)}{\varrho(x)} \right| \leq \| f \|_{\sqrt{\varrho}} \int \frac{\varrho(y) - \varrho(x)}{\varrho(x) \vee n} \frac{n}{\varrho(x) \vee m} |\mu(x, dy)| + \| f \|_{\sqrt{\varrho}} \int \frac{\varrho(y) - \varrho(x)}{\varrho(x) \vee n} \frac{n}{\varrho(x) \vee m} |\mu(x, dy)| \]

for some \( a_{nm} \to 0 \) as \( n, m \to \infty \). Hence we also have that

\[ \| Af + B^n f - (Af + B^m f) \|_{\sqrt{\varrho}} \leq a_{nm} \| f \|_{\sqrt{\varrho}} \]

for all \( f \in D \), implying that Condition (iii) of Theorem 3.1 is satisfied.
We finally have to check whether the growth bounds on \( B^\theta(X) \) and \( D \subseteq B^{\sqrt{\tau}}(X) \) are uniform in \( n \). We here apply, Condition 3.5 and the following immediate consequence of Condition 3.5:

\[
\int \frac{\sqrt{g(y)} - \sqrt{g(x)}}{\sqrt{g(x)}} \mu(x, dy) \leq \int \frac{\sqrt{g(y)} - \sqrt{g(x)}}{\sqrt{g(x)}} \left[ \sqrt{g(y)} + \sqrt{g(x)} \right] \mu(x, dy) = \int \frac{g(y) - g(x)}{g(x)} \mu(x, dy) \leq \tilde{\omega}, \quad \text{for all } x \in X.
\]

These conditions imply that \( |B^n g(x)| \leq \tilde{\omega} g(x) \) and \( |B^n \sqrt{g(x)}| \leq \tilde{\omega} \sqrt{g(x)} \) for all \( n \in \mathbb{N} \) so that Remark 2.9 leads to a uniform growth bounds in \( n \) for the semigroups \( B^\theta(X) \) and \( D \subseteq B^{\sqrt{\tau}}(X) \), respectively, and hence the approximation Theorem 3.3 can readily be applied and leads to a generalized Feller semigroup \( P_t^\infty \) with generator \( A + B \) whose growth bound satisfies \( \|P_t^\infty\|_{L(B^\theta(X))} \leq M \exp((\omega + M\tilde{\omega})t). \)

\[ \square \]

Remark 3.4. In contrast to classical Feller theory also processes with unbounded jump intensities can be constructed easily if \( g \) is unbounded on \( X \). The general character of the proposition allows to build general processes from simple ones by perturbation.

Corollary 3.5. Let \( X \subseteq \mathbb{R}^d \) be closed subset and let \( g(x) := 1 + \|x\|^2 \) be a weight function. Let furthermore \( A \) be the generator of a generalized Feller semigroup \( (P_t)_{t \geq 0} \) on \( B^\theta(X) \) which leaves \( B^{\sqrt{\tau}}(X) \subseteq B^\theta(X) \) invariant and satisfies \( \|P_t\|_{L(B^\theta(X))} \leq M \exp(\omega t) \) for some constants \( M \geq 1 \) and \( \omega \in \mathbb{R} \). Let

\[
\mu(x, d\xi) := \sum_{i=1}^d x_i \mu^i(d\xi)
\]

for some possibly signed measures \( \mu^i \) with bounded support such that \( \mu(x, d\xi) \) defines a family of finite (positive) measures on \( X \) and \( x + \text{supp}(\mu(x, \cdot)) \in X \) for all \( x \in X \). Then

\[
f \mapsto (x \mapsto Af(x) + \int (f(x + \xi) - f(x)) \mu(x, d\xi))
\]

for \( f \in \text{dom}(A) \cap B^{\sqrt{\tau}}(X) \) generates a generalized Feller semigroup on \( B^\theta(X) \).

Proof. Substituting \( y = x + \xi \), one easily verifies that Conditions 3.2 - 3.3 of Proposition 3.3 are satisfied since the supports of \( \mu^i \) are bounded. All other requirements are met as well and we can conclude. \[ \square \]

4. Lifting Stochastic Volterra processes

In the subsequent sections our main goal will be to treat the following types of SPDEs

\[
d\lambda_t = A^* \lambda_t dt + \nu dX_t,
\]

\[ (4.1) \]

on spaces \( E \subseteq Y^* \) as introduced in Section 2.2 where \( A^* \) is the generator of a strongly continuous semigroup \( S^* \) on \( Y^* \), \( \nu \in Y^* \) (or in a slightly larger space denoted by \( Z^* \) in the sequel), \( g \in Y \) and \( X \) a real valued Itô-semimartingale whose differential characteristics depend linearly on \( \langle g, \lambda_t \rangle \), which will turn out to be the solution of the Volterra equation with kernel \( \langle 1, 2 \rangle \).

Remark 4.1. As indicated in the introduction, it will be easy to consider vector valued structures, with \( M \) a semimartingale whose characteristics depend – instead of the \( \mathbb{R} \)-valued pairing \( \langle g, \lambda_t \rangle \) – linearly on a projection of \( \lambda_t \) onto some space, finite
or infinite dimensional. For the sake of the first exposition we stay one dimensional here.

In the following we summarize the main ingredients of our setting.

**Assumption 4.2.** Throughout this section we shall work under the following conditions:

(i) We consider Banach spaces $Z$ and $Y$ such that $Z \subset Y$ and $Z$ embeds continuously into $Y^*$, and their duals $Y^* \subset Z^*$ with their respective weak-\$^\ast$-topology.

(ii) We are given an admissible weight function $\varrho = 1 + \varrho_0$ in the sense of Section 4.4 on $Y^*$ such that

\[ \varrho_0(\lambda) = \|\lambda\|^2_{Y^*}, \quad \lambda \in Y^*, \]

where $\|\cdot\|_{Y^*}$ denotes the norm on $Y^*$.

(iii) We are given a closed convex cone $E \subset Y^*$ such that $\langle E, g \rangle$ is a weighted space in the sense of Section 2. This will serve as statespaces of $(\mathcal{E}, \mathcal{F})$.

(iv) We assume that a semigroup $S^t$ with generator $\mathcal{A}^*$ acts in a strongly continuous way on $Y^*$ and $Z^*$, with respect to the respective norm topologies.

(v) We assume that $\lambda \mapsto S^t \lambda$ is weak-*continuous on $Y^*$ and on $Z^*$ for every $t \geq 0$ (considering the weak-*topology on both the domain and the image space).

(vi) We suppose that the (pre-) adjoint operator of $\mathcal{A}^*$, denoted by $\mathcal{A}$ and domain $\text{dom}(\mathcal{A}) \subset Z \subset Y$, generates a strongly continuous semigroup on $Z$ with respect to the respective norm topology but not necessarily with respect to $Y$.

**Remark 4.3.** We could allow more general weight functions $\varrho$ but it is not necessary for our purposes here and the formulation of their abstract properties is cumbersome.

**Remark 4.4.** Notice the following curious equivalence: the weight function $\varrho$ is continuous with respect to the weak-*topology on $\mathcal{E}$ if and only if the space $\mathcal{E}$ is locally compact. Indeed if $\varrho$ is continuous on $\mathcal{E}$, then of course the topology on $\mathcal{E}$ is locally compact since every point has a compact neighborhood of type $\{ \varrho \leq R \}$ for some $R > 0$. On the other hand if the topology on $\mathcal{E}$ is locally compact, then for every point $\lambda_0 \in \mathcal{E}$ there is a convex, compact neighborhood $V \subset \mathcal{E}$ such that $\varrho(\lambda) - \varrho(\lambda_0)$ is bounded on $V$ by a number $k > 0$, whence by convexity $|\varrho(s(\lambda - \lambda_0) + \lambda_0) - \varrho(\lambda_0)| \leq sk$ for $\lambda - \lambda_0 \in s(V - \lambda_0)$ and $s \in [0, 1]$. This in turn means that $\varrho$ is continuous at $\lambda_0$.

This case appears in the following well known case: $Y$ is the space of continuous functions on a compactum $E$, and $\mathcal{E}$ is the space of non-negative Borel measures thereon, then $\mathcal{E}$ is locally compact in the weak-*topology. In this case the total variation norm of a non-negative measures is nothing but the total mass, i.e. a linear functional on it. The one-point compactification in this case is called Watanabe topology, see [8].

**Remark 4.5.** A prototypical example presented in Section 5.1 is given by $Y = C_0(\mathbb{R}, \mathbb{R})$ with supremum norm and $Z$ the space of functions $g \in Y$ such that $(x \mapsto xg(x)) \in Y$ together with the operator norm on it, i.e.

\[ \|g\| = \sqrt{\|g\|^2 + \|xg(x)\|^2} \]

for $g \in Z$. The semigroup is given by $S_t f(x) = \exp(-xt)f(x)$ for $t \geq 0$, $f \in Y$ and $x \geq 0$. All above requirements on spaces and semigroups are then satisfied.
To analyze (4.1) and to construct $E$ we first consider the following linear deterministic equation

$$d\lambda_t = A^* \lambda_t dt - w \nu \langle g, \lambda_t \rangle dt$$

(4.2)

for $\lambda_0 \in Y^*$, some fixed $g \in Y$, a real number $w > 0$ and $\nu \in Z^*$.

Under the subsequent assumptions on $S^*$ and $\nu \in Z^*$ we can guarantee that it can be solved on the space $Y^*$ for all times in the mild sense with respect to the dual norm $\| \cdot \|_{Y^*}$ by a standard Picard iteration method.

**Assumption 4.6.** We assume that

(i) $S^*_t \nu \in Y^*$ for all $t > 0$ even though $\nu$ does not necessarily lie in $Y^*$ itself, but only in $Z^*$;

(ii) $\int_0^t \| S^*_s \nu \|^2_{Y^*} ds < \infty$ for all $t > 0$.

As in (1.2), we define

$$K(t) := \langle g, S^*_t \nu \rangle,$$

(4.3)

which will correspond to the kernel in the Volterra equation (1.1) and define the $R \in L^2_{\text{loc}}(\mathbb{R}_+, \mathbb{R})$ as the resolvent of the second kind that satisfies

$$K * R = R * K = K - R.$$ 

(4.4)

This resolvent always exists and is unique (see [18, Theorem 3.1])

**Proposition 4.7.** Under Assumption 4.6, there exists a unique mild solution of (4.2) with values in $Y^*$. Additionally, the solution operator is a weak-$*$-continuous map $\lambda_0 \mapsto \lambda_t$, for each $t > 0$, and the solution satisfies

$$g(\lambda_t) \leq C g(\lambda_0), \quad \text{for all } \lambda_0 \in Y^* \text{ and } t \in [0, \varepsilon]$$

for some positive constants $C$ and $\varepsilon$.

**Remark 4.8.** The unique mild solution of Equation (4.2) satisfies

$$\lambda_t = S^*_t \lambda_0 - w \int_0^t S^*_{t-s} \nu \langle g, \lambda_s \rangle ds$$

for all $t \geq 0$. Pairing with $g \in Y$ yields a deterministic linear Volterra equation of the form

$$\langle g, \lambda_t \rangle = \langle g, S^*_t \lambda_0 \rangle - w \int_0^t \langle g, S^*_{t-s} \nu \rangle \langle g, \lambda_s \rangle ds$$

(4.5)

$$= \langle g, S^*_t \lambda_0 \rangle - w \int_0^t K(t-s) \langle g, \lambda_s \rangle ds,$$

where we used (4.3).

**Proof.** We prove first the completely standard convergence of the Picard iteration scheme with respect to the dual norm on $Y^*$. Define

$$\lambda^0_t = \lambda_0,$$

$$\lambda^{n+1}_t = S^*_t \lambda_0 - w \int_0^t S^*_{t-s} \nu \langle g, \lambda^0_s \rangle ds, \quad n \geq 0.$$
Then, by Assumption \[4.6\] (i) each \(\lambda^n_t\) lies \(Y^*.\) Consider now

\[
\|\lambda^{n+1}_t - \lambda^n_t\|_{Y^*} = w\int_0^t S^*_t \nu(g, \lambda^n_t - \lambda^{n-1}_t)ds Y^*,
\]

\[
\leq w\int_0^t \|S^*_t \nu(g, \lambda^n_t - \lambda^{n-1}_t)\|_Y ds Y^*,
\]

\[
\leq w\int_0^t \|S^*_t \nu\|_Y \|g\|_Y \|\lambda^n_t - \lambda^{n-1}_t\|_{Y^*} ds Y^*.
\]

Assumption \[4.6\] (ii) and an extended version of Gronwall’s inequality see [7, Lemma 15] then yield convergence of \((\lambda^n_t)_{n \in \mathbb{N}}\) to some \(\lambda_t\) with respect to the dual norm \(\| \cdot \|_{Y^*}\) uniformly in \(t\) on compact intervals. For details on strongly continuous semigroups and mild solutions see [22].

Having established the existence of a mild solution of (4.2) in \(Y^*\), consider now the linear, deterministic Volterra equation (4.5), which can be written as (4.6).

\[
\langle g, \lambda_t \rangle = \langle g, S^*_t \lambda_0 \rangle - \int_0^t R^w(t - s)\langle S^*_t \lambda_0, g \rangle ds Y^*,
\]

where \(R^w\) denotes the resolvent of \(wK(t) = w(g, S^*_t \nu)\) as introduced in (4.4). Since by assumption \(S^*_t\) is a weak-\(\ast\)-continuous solution operator, the map \(\lambda_0 \mapsto (t \mapsto \langle g, S^*_t \lambda_0 \rangle)\) is weak-\(\ast\)-continuous as a map from \(Y^*\) to \(C(\mathbb{R}_+, \mathbb{R})\) (with the topology of uniform convergence on compacts on \(C(\mathbb{R}_+, \mathbb{R})\)). From (4.6) we thus infer that \(\langle g, \lambda_t \rangle\) is weak-\(\ast\)-continuous for every \(t \geq 0\), which clearly translates to the solution map of Equation (4.2).

Finally we have to show that the stated inequality for \(\varrho(\lambda_t)\) holds true on small time intervals \([0, \varepsilon]\).

Observe first that for \(t \in [0, \varepsilon]\)

\[
\|S^*_t \lambda\|_{Y^*}^2 \leq C\|\lambda\|_{Y^*}^2
\]

for all \(\lambda \in Y^*\) just by the assumption that \(S^*_t\) is strongly continuous, for some constant \(C \geq 1\). Furthermore for \(t \in [0, \varepsilon]\)

\[
\|\lambda_t\|_{Y^*}^2 \leq 2(C\|\lambda_0\|_{Y^*}^2 + t \int_0^t \|wS^*_t \nu(g, \lambda_s)\|_{Y^*}^2 ds Y^*)
\]

\[
\leq 2(C\|\lambda_0\|_{Y^*}^2 + \varepsilon \int_0^t w^2\|S^*_t \nu\|_{Y^*}^2 \|g\|_{Y^*}^2 \|\lambda_s\|_{Y^*}^2 Y^*, ds Y^*).
\]

Consider now the kernel \(K'(t, s) = 2\varepsilon w^2\|S^*_t \nu\|_{Y^*} Y^* \|g\|_{Y^*}^2 1_{\{s \leq t\}}.\) We follow now the arguments of the proof of [3, Lemma 3.1]. Indeed \(K'\) is a Volterra kernel as defined in [18, Definition 9.2.1] and for any interval \([U, V) \subset \mathbb{R}_+\) we have by Young’s convolution inequality

\[
\|K'\|_{L^1([U, V))} \leq 2\varepsilon w^2 \|g\|_{Y^*}^2 \int_U^V \|S^*_s \nu\|_{Y^*}^2 ds Y^*.
\]

where \(\| \cdot \|_{L^1([U, V))}\) is defined in [18, Definition 9.2.2]. We can now literally take the proof of [3, Lemma 3.1] to deduce that the generalized Gronwall Lemma (see [18, Lemma 9.8.2]) can be applied. This yields for \(t \in [0, \varepsilon]\)

\[
\|\lambda_t\|_{Y^*}^2 \leq \|\lambda_0\|_{Y^*}^2 \cdot 2C(1 - \int_0^t R'(s)ds) \leq \|\lambda_0\|_{Y^*}^2 \cdot 2C(1 - \int_0^\varepsilon R'(s)ds),
\]

where \(R'\) denotes the resolvent of \(-K'\), which is nonpositive. This leads to the desired assertion due to the definition of \(\varrho\). From this inequality also uniqueness follows in a standard way.
It is of crucial importance to understand that there is actually a closed sub-cone \( E \subset Y^* \) left invariant by the solution of Equation (4.7). This cone will play the role as announced in Assumption 4.2 (iii) and can be described in terms of initial values \( \lambda_0 \) which give rise to nonnegative solutions of (4.7). Indeed, it will be the intersection of the following cones. Define for fixed \( w > 0 \)

\[
E^w := \{ \lambda_0 \in Y^* | \langle g, S^*_t \lambda_0 \rangle - \int_0^t R^w(t - s) \langle g, S^*_s \lambda_0 \rangle ds \geq 0 \text{ for all } t \geq 0 \}
\]

\[
= \{ \lambda_0 \in Y^* | \langle g, \lambda_t \rangle \geq 0 \text{ with } \lambda_t = S^*_t \lambda_0 - w \int_0^t S^*_{t-s} \nu \langle g, \lambda_s \rangle ds \text{ for all } t \geq 0 \},
\]

where \( R^w \) denotes the resolvent of \( wK \).

**Proposition 4.9.** Let Assumption 4.4 be in force and let \( w > 0 \) be fixed. The set \( E^w \) as defined in (4.7) is a weak-\( * \)-closed convex cone. The solution of (4.7) leaves \( E^w \) invariant and defines a generalized Feller semigroup on \( (E^w, \nu) \) by \( P_t \nu(\lambda_0) := f(\lambda_t) \) for all \( \lambda_0 \in E^w \) and \( t \geq 0 \).

**Proof.** The weak-\( * \)-closedness follows from the weak-\( * \)-continuity of \( S^* \) and the convex cone property is obvious. By uniqueness of solutions and Definition of \( E^w \) in (4.7), it is clear that for any \( \lambda_0 \in E^w \) also \( \lambda_t \in E^w \), whence we obtain invariance of \( E^w \). Since by Proposition 4.7 the solution operator is weak-\( * \)-continuous, we can conclude that \( \lambda_0 \mapsto f(\lambda_t) \) lies in \( B^\sigma(E^w) \) for a dense set of \( B^\sigma(E^w) \) by Theorem 2.15. Moreover, it satisfies the necessary bound \( 2.10 \) on \( E^w \) by (norm)-continuity of \( t \mapsto \lambda_t \). Hence all the conditions of Assumption 2.17 are satisfied and the solution operator therefore defines a generalized Feller semigroup \( (P_t) \) by Theorem 2.10.

We need an additional assumption assuring that the above defined state space contains the cone hull of \( S_t^* \nu \):

**Assumption 4.10.** Let \( \nu \) be such that \( S_t^* \nu \in E^w \) for all \( u > 0 \) and for all \( w > 0 \).

**Remark 4.11.** This condition is satisfied if \( K \) and \( R^w \) are nonnegative for all \( w > 0 \). Indeed, \( S_t^* \nu \in E^w \) if and only if

\[
\langle g, S^*_{t+u} \nu \rangle - \int_0^t R^w(t - s) \langle g, S^*_{s+u} \nu \rangle ds = K(t + u) - \int_0^t R^w(t - s) K(s + u) ds \geq 0.
\]

Since by the resolvent equation we have

\[
R^w(t + u) = wK(t + u) - \int_0^{t+u} R^w(t + u - s) K(s) ds
\]

\[
= wK(t + u) - \int_0^t R^w(t - s) K(s + u) ds - \int_0^u R^w(t + u - s) K(s) ds,
\]

it follows that

\[
K(t + u) - \int_0^t R^w(t - s) K(s + u) ds = \frac{1}{w} R^w(t + u) + \int_0^u R^w(t + u - s) K(s) ds.
\]

If \( K \) and \( R^w \) are nonnegative (4.8) is clearly satisfied.

The following lemma states that the scale of invariant spaces \( E^w \) is actually decreasing.
Lemma 4.12. Let Assumptions 4.6 and 4.10 be in force. For \( w_1 < w_2 \) we have \( E^{w_1} \supseteq E^{w_2} \). Additionally we have that for \( \lambda_0 \in E^{w_2} \) the unique solution of

\[
\lambda_t = S_t^* \lambda_0 - w_1 \int_0^t S_{t-s}^* \nu(g, \lambda_s) ds
\]

for all \( t \geq 0 \) actually lies in \( E^{w_2} \).

Proof. Fix \( w_1 < w_2 \) and \( \lambda_0 \in E^{w_2} \). Actually we can write

\[
\lambda_t = S_t^* \lambda_0 - w_1 \int_0^t S_{t-s}^* \nu(g, \lambda_s) ds = \underbrace{S_t^* \lambda_0 - w_2 \int_0^t S_{t-s}^* \nu(g, \lambda_s) ds}_{A} + \underbrace{(w_2 - w_1) \int_0^t S_{t-s}^* \nu(g, \lambda_s) ds}_{B}.
\]

By Proposition 4.9 the term \( A \) clearly lies in \( E^{w_2} \). The same holds true for the second term \( B \) due to Assumption 4.10 and the cone property of \( E^{w_2} \), since \((w_2 - w_1) > 0 \) and \( \langle g, \lambda_s \rangle \geq 0 \) for all \( s > 0 \). Thus \( \lambda_t \in E^{w_2} \). However, by definition of \( E^{w_1} \), \( \lambda_t \) also lies there, whence the conclusion. \( \square \)

Definition 4.13. Let Assumptions 4.6 and 4.10 be in force. Then we define the following weak-* closed cone in \( Y^* \)

\[
E = \cap_{w > 0} E^w.
\]

Theorem 4.14. Let Assumptions 4.6 and 4.10 be in force. Then the weak-* closed cone \( E \) contains all damped measures \( S_t^* \nu \) for \( t > 0 \) and is left invariant by the solution of

\[
(4.9) \quad dl_t = A^* \lambda_t dt - w \nu(g, \lambda_t) dt
\]

for all \( w > 0 \). This solution defines a generalized Feller semigroup on \( B^0(\mathcal{E}) \).

Proof. The first assertion follows from Assumption 4.10. By monotonicity of \( E^w \) and the fact that for solutions with respect to \( w_1 \)-equations actually lie in \( E^{w_2} \), as stated in the previous lemma, the invariance assertion and generalized Feller property follow together with Proposition 4.9. \( \square \)

Remark 4.15. By the generalized maximum principle as stated in Theorem 2.9 we can make the following assertion about the geometry of \( E \). Let \( \tilde{Y} \) be a dense set of the \( (\text{pre}) \)-polar cone \( E \), such that \( \tilde{Y} \subset \text{dom}(A) \) and \( \langle \nu, y \rangle \) is well defined for all \( y \in \tilde{Y} \). Consider the generator \( A^w \) of (4.10), which acts on functions \( f \mapsto f(\langle g, \lambda \rangle) \)

\[
A^w f(\langle g, \lambda \rangle) = f'(\langle g, \lambda \rangle)(\langle Ay, \lambda \rangle - w \nu(g, \lambda)).
\]

From Theorem 4.13 and Theorem 2.6 we know that for all \( w > 0 \), \( A^w \) satisfies the generalized maximum principle for a dense set of functions in \( B^0(\mathcal{E}) \). Applying this to \( f(x) = -x \) and some \( y^* \in \tilde{Y} \) we see that the maximum of \( \lambda \mapsto g^{-1} f(\langle y^*, \lambda \rangle) \) is achieved at some \( \lambda^* \in E \) such that \( \langle y^*, \lambda^* \rangle = 0 \). The generalized maximum principle thus implies that

\[
A f(\langle y^*, \lambda^* \rangle) = -w \nu(g, \lambda^*) \leq \omega f(\langle y^*, \lambda^* \rangle) = 0.
\]

From this we infer that

\[
w \leq \frac{\langle Ay^*, \lambda^* \rangle}{\langle g^*, \nu \rangle(g, \lambda^*)}.
\]

Since this has to hold for all \( w > 0 \), \( \langle y^*, \lambda^* \rangle = 0 \), necessarily implies that either \( \langle y^*, \nu \rangle = 0 \) or \( \langle g, \lambda^* \rangle = 0 \).
By the previous results we can now construct a generalized Feller process on $\mathcal{E}$ which jumps up by multiples of $S^*_t \nu$ for some $\varepsilon \geq 0$ and with an instantaneous intensity of size $(g, \lambda_t)$. We formulate these assertions in two propositions. For their formulation recall that $\mathcal{E}_c \subset Y$ denotes the (pre-)polar cone of $\mathcal{E}$.

**Proposition 4.16.** Let Assumptions 4.6 and 4.10 be in force. Moreover, let $\mu \in \mathcal{M}_+(\mathbb{R}_+)$ with finite second moment. Consider the SPDE

$$
(4.10) \quad d\lambda_t = A^* \lambda_t dt - w\nu(g, \lambda_t) dt + S^*_t \nu dN_t,
$$

where $(N_t)_{t \geq 0}$ is a pure jump process with compensator $F(\lambda_t, d\xi) = (g, \lambda_t) \mu(d\xi)$.

(i) Then for every $\lambda_0 \in \mathcal{E}$, $\varepsilon > 0$ and $w > 0$, the SPDE (4.10) has a solution in $\mathcal{E}$ given by a generalized Feller process associated to the generator of (4.10).

(ii) This generalized Feller process is also a probabilistically weak and analytically mild càglàd solution of (4.10), i.e.

$$
\lambda_t = S^*_t \lambda_0 - \int_0^t S^*_{t-s} w\nu(g, \lambda_s) ds + \int_0^t S^*_{t+s+\varepsilon} \nu dN_s,
$$

which justifies Equation (1.11), in particular for every initial value the process $N$ can be constructed on an appropriate probabilistic basis.

(iii) For every $\varepsilon > 0$ and $w > 0$, the corresponding Riccati equation $\partial_y y_t = R(y_t)$ with $R : \mathcal{E}_c \to \mathbb{R}$ given by

$$
(4.11) \quad R(y) = Ay - wg(y, \nu) + g \int_{\mathbb{R}_+} (\exp((y, S^*_s \nu \xi)) - 1) \mu(d\xi), \quad y \in \mathcal{E}_c,
$$

admits a unique global solution in the mild sense.

(iv) The affine transform formula holds true.

$$
\mathbb{E}_{\lambda_0} [\exp((y_0, \lambda_t))] = \exp((y_t, \lambda_0)),
$$

where $y_t$ solves $\partial_y y_t = R(y_t)$ in the mild sense with $R$ given by (4.11). Moreover $y_t \in \mathcal{E}_c$ for all $t \geq 0$.

**Proof.** To prove the first assertion we apply Proposition 3.3. By Theorem 4.14 the deterministic equation (4.11) has a mild solution on $\mathcal{E}$ which – by Assumption 4.6 – defines a generalized Feller semigroup $(P_t)_{t \geq 0}$ on $\mathcal{B}^0(\mathcal{E})$. The operator $A$ in Proposition 3.3 then corresponds to the generator of $(P_t)_{t \geq 0}$, i.e. the semigroup associated to the purely deterministic part of (4.10), which clearly satisfies the growth bound simply by strong continuity.

Note that by the same arguments as in Proposition 4.9 and by applying Theorem 2.19 we can prove that $(P_t)_{t \geq 0}$ also defines a generalized Feller semigroup on $\mathcal{B}^0(\mathcal{E})$. Indeed, by weak-$*$-continuity, we can conclude that $\lambda_0 \mapsto f(\lambda_t)$ lies in $\mathcal{B}^0(\mathcal{E})$ for a dense set of $\mathcal{B}^0(\mathcal{E})$ by Theorem 2.15 whence condition (iii) of Assumption 2.17 is satisfied. Moreover, (2.11) is satisfied by (norm)-continuity of $t \mapsto \lambda_t$ and the necessary bound (2.10) holds also for $\sqrt{t}$. For the latter observe first that

$$
||S^*_t \lambda||_{Y^*} \leq C||\lambda||_{Y^*} \quad \text{for all } 0 \leq t \leq T \text{ and } \lambda \in Y^*,
$$

since $S^*$ is strongly continuous on $Y^*$, for some $T > 0$. Using this we can estimate

$$
(4.12) \quad ||\lambda_t||_{Y^*} \leq ||\lambda_0||_{Y^*} + \int_0^t w||S^*_{t-s} \nu||_{Y^*} ||g||_{Y^*} ||\lambda_s||_{Y^*} ds.
$$

We now proceed similarly as in Proposition 4.7. Consider the kernel $K'(t, s) = w||S^*_{t-s} \nu||_{Y^*} ||g||_{Y^*} 1_{s \leq t}$, which is again a Volterra kernel in the sense of Definition 9.2.1. Moreover, for any interval $[U, V] \subset \mathbb{R}_+$ we have by Young’s convolution
inequality
\[ |||K'||||_{L^1(U,V)} \leq w\|g\|_Y \int_0^{V-U} \|S^*_\nu\|_{Y^*} ds \leq (V-U)w\|g\|_Y^2 \left( \int_0^{V-U} \|S^*_\nu\|_{Y^*}^2 ds \right)^{1/2}, \]
where \( ||| \cdot |||_{L^1(U,V)} \) is defined in [13, Definition 9.2.2]. We now can again literally take the proof of [3, Lemma 3.1] to deduce that the generalized Gronwall Lemma (see [13, Lemma 9.8.2]) can be applied. This yields for \( t \in [0,T] \)
\[ \|\lambda_1\|_{Y^*} \leq \|\lambda_0\|_{Y^*} C(1 - \int_0^t R'(s) ds), \]
where \( R' \) denotes the resolvent of \(-K'\), which is nonpositive. Hence, for \( t \in [0,T] \) we have by Jensen’s inequality
\[ \sqrt{\varrho(\lambda_0)} = \sqrt{1 + ||\lambda_0||^2_{Y^*}} \leq 1 + ||\lambda_0||_{Y^*} \leq 1 + ||\lambda_0||_C - \int_0^T R'(s) ds \]
\[ \leq C \sqrt{1 + ||\lambda_0||^2_{Y^*}} = C \sqrt{\varrho(\lambda_0)}, \]
where \( C \) depends on \( T \).

Finally, we need to verify (3.2) - (3.5), which read as follows
\[ \int \varrho(\lambda + S^*_\nu\xi)(g, \lambda) \mu(d\xi) \leq M\varrho(\lambda)^2, \]
\[ \int \sqrt{\varrho(\lambda + S^*_\nu\xi)(g, \lambda) \mu(d\xi)} \leq M \varrho(\lambda), \]
\[ \int (g, \lambda) \mu(d\xi) \leq M \sqrt{\varrho(\lambda)} \]
\[ \int \frac{|g(\lambda + S^*_\nu\xi) - \varrho(\lambda)(g, \lambda) \mu(d\xi)|}{\varrho(\lambda)} (g, \lambda) \mu(d\xi) \leq \tilde{w}. \]
By the second moment condition on \( \mu \) these conditions are clearly satisfied for some \( M \) and some \( \tilde{w} \). Proposition 3.3 now allows to conclude that \( A + B \) where \( B \) is given by
\[ Bf(\lambda) = \int (f(\lambda + S^*_\nu\xi) - f(\lambda))(g, \lambda) \mu(d\xi) \]
generates a generalized Feller semigroup \( \tilde{P} \) as asserted.

For (ii) we now construct the probabilistically weak and analytically mild solution directly from the properties of the generalized Feller process: take \( y \in \text{dom}(A) \) such that \( \langle y, \nu \rangle \) is well defined and consider the martingale
\[ M^y_t := \langle y, \lambda_t \rangle - \langle y, \lambda_0 \rangle - \int_0^t \langle Ay, \lambda_s \rangle - w\langle y, \nu \rangle \langle g, \lambda_s \rangle ds \]
(4.13)
\[ - \int_0^t \int (y, S^*_\nu\xi)(g, \lambda_s) \mu(d\xi) ds \]
for \( t \geq 0 \) (after appropriate regularization as it is possible due to Theorem 2.12 such that the integral term is well defined). Let now \( y \) be as above with the additional property that \( \langle y, S^*_\nu \rangle = 1 \). Define
\[ N_t := M^y_t + \int_0^t \int \langle y, S^*_\nu\xi \rangle (g, \lambda_s) \mu(d\xi) ds \]
(4.14)
for \( t \geq 0 \), which is a càglàd semimartingale. Then \( N \) does not depend on \( y \). Indeed, for all \( y_i \) with \( \langle y_i, S^*_\nu \rangle = 1 \), \( i = 1, 2 \), we clearly have
\[ \int_0^t \int (y_1 - y_2, S^*_\nu\xi)(g, \lambda_s) \mu(d\xi) ds = 0 \]
and $M^{y_1} - M^{y_2} = M^{y_1-y_2} = 0$ and as well. The latter follows from the fact that the martingale $M^y$ is constant if $\langle y, S^*_\nu \rangle = 0$. This is indeed true since the martingale’s quadratic variation vanishes in this case: we apply here the carré du champ formula

$$E[[M^y, M^y]] = E\left[ \int_0^t \left( Af^2(\lambda_s) - 2f(\lambda_s)Af(\lambda_s) \right) ds \right]$$

where $f(\lambda) = \langle y, \lambda \rangle$ with $y$ satisfying $\langle y, S^*_\nu \rangle = 0$. Moreover, by the definition of $N$ in (4.11) its compensator is given by $\int_0^t \xi(\lambda_s, g) \mu(d\xi)$ showing that $N$ has the desired properties.

By (4.14) we additionally obtain that

$$\langle y, \lambda_t \rangle = \langle y, \lambda_0 \rangle + \int_0^t \langle Ay, \lambda_s \rangle ds - \int_0^t \langle y, w \nu \rangle g, \lambda_s \rangle ds + \langle y, S^*_\nu \rangle N_t$$

This analytically weak form can be readily translated into a mild form by standard methods, which proves (ii).

Concerning (iii), note first that we have a unique mild solution to

(4.15)

$$\partial_y y_t = Ay_t - w g(y_t, \nu), \quad y_0 \in Y,$$

since this is the adjoint equation of (4.12). For the equation with jumps we proceed as in Proposition 4.7 via Picard iteration. Denote the semigroup associated to (4.15) by $S^w_t$ and define

$$y^0_t = y_0,$$

$$y^n_t = S^w_t y_0 + \int_0^t S^w_{t-s} g \left( \int_{\mathbb{R}^+} \left( \exp(\langle y^n_{s-1}, S^*_\nu \xi \rangle) - 1 \right) \mu(d\xi) \right) ds.$$ 

Moreover, for $t \in [0, \delta]$ for some $\delta > 0$ we have by local Lipschitz continuity of $x \mapsto \exp(x)$

$$\|y^{n+1}_t - y^n_t\|_Y \leq \| \int_0^t S^w_{t-s} g \left( \exp(\langle y^n_{s-1}, S^*_\nu \xi \rangle) - \exp(\langle y^n_{s-1}, S^*_\nu \xi \rangle) \right) \mu(d\xi) ds\|_Y$$

$$\leq \int_0^t C \| S^w_{t-s} g \|_Y \| y^{n-1}_t - y^{n-1}_t \|_Y \left( \int \| S^*_\nu \xi \|_Y \mu(d\xi) \right) ds.$$ 

By an extension of Gronwall’s inequality (see [7, Lemma 15] this yields convergence of $(y^n_t)_{n \in \mathbb{N}}$ with respect to $\| \cdot \|_Y$ and hence the existence of a unique local mild solution to (4.14) up to some maximal life time $t_+(y_0)$. That $t_+(y_0) = \infty$ for all $y_0 \in \mathcal{E}^*$ follows from the subsequent estimate

$$\|y_t\|_Y = \| S^w_t y_0 + \int_0^t S^w_{t-s} g \left( \int \exp(\langle y_s, S^*_\nu \xi \rangle) - 1 \right) \mu(d\xi) ds \|_Y$$

$$\leq \| S^w_t y_0 \|_Y + \int_0^t \| S^w_{t-s} g \|_Y \left( \int \exp(\langle y_s, S^*_\nu \xi \rangle) - 1 \right) \mu(dx) ds$$

$$\leq \| S^w_t y_0 \|_Y + \sup_{s \leq t} \| S^w_{t-s} g \|_Y \mu(\mathcal{R}_+),$$

where we used $| \exp(y_s, S^*_\nu \xi) - 1 | \leq 1$ for all $y \in \mathcal{E}^*$ in the last estimate.

To prove (iv), just note that by the existence of a generalized Feller semigroup the abstract Cauchy problem for initial value $\exp(y_0, \cdot)$ can be solved uniquely for $y_0 \in \mathcal{E}^*$. Indeed, $\mathbb{E}(\exp(y_0, \lambda_t))$ uniquely solves

$$\partial_y u(t, \lambda) = A u(t, \lambda), \quad u(0, \lambda) = \exp(y_0, \lambda).$$
where $A$ denotes the generator associated to (4.10). Setting $u(t, \lambda) = \exp(y_t, \lambda)$, we have
\[
\partial_t u(t, \lambda) = \exp(y_t, \lambda)R(y_t),
\]
where the right hand side is nothing else than $A\exp(y_t, \lambda)$, hence the affine transform formula holds true. This also implies that $y_t \in \mathcal{E}_*$ for all $t \geq 0$, simply because $E_\lambda[\exp((y_0, \lambda_t))] \leq 1$ for all $\lambda \in \mathcal{E}$.

The next statement is a refinement of Proposition 4.16 for the case $\varepsilon = 0$ since path properties change in general in this case.

**Proposition 4.17.** Let Assumptions 4.10 and 4.11 be in force. Moreover, let $\mu \in \mathcal{M}_+(\mathbb{R}_+)$ with finite second moment. Then the conclusions (i) and (ii) of Proposition 4.11 hold true for $\varepsilon = 0$, except that the probabilistically weak and analytically mild solution can only be chosen càglàd due to the possibly infinite mass of the measure $\nu$. Moreover, the affine transform formula for the generalized Feller process $(\lambda_t)_{t \geq 0}$ of (4.11) with $\varepsilon = 0$ is also satisfied, i.e.
\[
E_{\lambda_0}[\exp((y_0, \lambda_t))] = \exp((y_t, \lambda_0)),
\]
where $y_t$ solves $\partial_t y_t = R(y_t)$ in the mild sense with $R$ given by (4.11) with $\varepsilon = 0$.

**Proof.** To prove (i) we apply Theorem 3.2 and consider a sequence of generalized Feller semigroups $(P^n)_{n \in \mathbb{N}}$ with generators $A^n$ corresponding to the solution of (4.10) for $\varepsilon = \frac{1}{n}$, $n \in \mathbb{N}$. Let us first establish a uniform growth bound for this sequence. Note that for the solution of (4.10), we have due to Proposition 4.11 (ii) the following estimate for $t \in [0, T]$ for some fixed $T > 0$
\[
E[\|\lambda_t^*\|_{Y^*}^2] \leq 3\|S_{r+t}^\varepsilon\|_{Y^*}^2 + 3t \int_0^t w^2\|S_{t-s}^\varepsilon\|_{Y^*}^2 \cdot \|g\|_{Y^*}^2 E[\|\lambda_s^\varepsilon\|_{Y^*}^2] ds
+ 6E\int_0^t (S_{t-s}^\varepsilon \nu\xi dN_s - \int_0^t S_{t-s}^\varepsilon \nu\xi \mu(g, \lambda_s^\varepsilon) ds)^2 \]
+ 6E\int_0^t \int S_{t-s}^\varepsilon \nu\xi \mu(g, \lambda_s^\varepsilon) ds)^2 \]
\[
\leq C_0\|\lambda_0\|_{Y^*}^2 + 3t \int_0^t w^2\|S_{t-s}^\varepsilon\|_{Y^*}^2 \cdot \|g\|_{Y^*}^2 E[\|\lambda_s^\varepsilon\|_{Y^*}^2] ds
+ 6E\int \xi^2 \mu(d\xi) \int_0^t \|S_{t-s}^\varepsilon\|_{Y^*}^2 \cdot \|g\|_{Y^*}^2 E[\|\lambda_s^\varepsilon\|_{Y^*}^2] ds
+ 6t \int \xi^2 \mu(d\xi) \int_0^t \|S_{t-s}^\varepsilon\|_{Y^*}^2 \cdot \|g\|_{Y^*}^2 E[\|\lambda_s^\varepsilon\|_{Y^*}^2] ds
\]
\[
\leq C_0\|\lambda_0\|_{Y^*}^2 + 3t \int_0^t w^2\|S_{t-s}^\varepsilon\|_{Y^*}^2 \cdot \|g\|_{Y^*}^2 E[\|\lambda_s^\varepsilon\|_{Y^*}^2] ds
+ 6t \int \xi^2 \mu(d\xi) \int_0^t \|S_{t-s}^\varepsilon\|_{Y^*}^2 \cdot \|g\|_{Y^*}^2 E[\|\lambda_s^\varepsilon\|_{Y^*}^2] ds
\]
where $C_0$ and $C_2$ depend on $T$. We use $\|S_{t}^\varepsilon\|_{Y^*} \leq C_0\|\lambda_0\|^2$ for $t \in [0, T]$, as well as $\|S_{t-s}^\varepsilon\|_{Y^*} \leq \tilde{C}\|S_{t-s}^\varepsilon\|_{Y^*}$ for some constant $\tilde{C}$ and all $\varepsilon \in (0, 1]$ due to strong
continuity. Exactly by the same arguments as in the proof of Proposition 4.17, we thus obtain for $t \in [0, T]$ for some fixed $T$

$$
\mathbb{E}[\|\lambda t\|^2_{\mathcal{F}_T}] \leq \tilde{C}(\|\lambda 0\|^2_{\mathcal{F}_0} + 1)(1 - \int_0^t R'(s), ds),
$$

where $R'$ denotes the resolvent of $-C_2 S_{t-\nu}^\nu\nu y \nu$. Hence, $\mathbb{E}[\phi(\lambda t)] \leq C\phi(\lambda 0)$ for $t \in [0, T]$. From this the desired uniform growth bound $\|P_t\|_{L(E \rightarrow E)} \leq M exp(\omega t)$ for some $M \geq 1$ and $\omega \in \mathbb{R}$ follows.

For the set $D$ as of Theorem 3.2 we here choose Fourier basis elements of the form

(4.16) 
$$
f_y : \mathcal{E} \rightarrow [0, 1]; \lambda \mapsto \exp(\langle y, \lambda \rangle)
$$

such that $y \in \mathcal{E}$ and $\lambda \mapsto \exp(\langle y, \lambda \rangle)$ lies in $\cap_{n \geq 1} \text{dom}(A^n)$, whose span is of course dense, whence (i) of Theorem 3.2. We now equip span$(D)$ with the uniform norm $\| \|_{\infty}$ and verify Condition (ii), i.e. we check

(4.17) 
$$
\|A^n P^n f_y - A^m P^n f_y\|_{\epsilon} \leq \|f_y\|_{\infty} \alpha_{nm}
$$

for all $0 \leq u \leq t$ with $a_{nm} \rightarrow 0$ as $n, m \rightarrow \infty$, and possibly depending on $y$. Note that

$$
A^n f_y(\lambda) = \langle R^n(y), \lambda \rangle f_y(\lambda),
$$

where $R^n$ corresponds to (4.11) for $\epsilon = \frac{1}{n}$. As $P^n$ leaves $D$ invariant for all $n \in \mathbb{N}$ by Proposition 4.16 (iv), we have

$$
\frac{|A^n P^n f_y(\lambda) - A^m P^n f_y(\lambda)|}{\|\phi(\lambda)\|_{\epsilon}} = \frac{f_y(\lambda)}{\|\phi(\lambda)\|_{\epsilon}} \int_{\mathbb{R}_+} \exp(\langle y^{m}_0, S^\nu_{t-s} \nu \xi \rangle) |\exp(\langle y^{m}_0, (S^\nu_{t-s} - S^\nu_{t-s} \nu) \xi \rangle) - 1| \mu(d\xi)
$$

where $y^{m}_0$ denotes the solution of $\partial_t y^{m}_u = R^n(y^{m}_u)$ at time $u$ with $y_0 = y$ and $\tilde{a}_{nm}$ is chosen uniformly for all $u \leq t$ and tends to 0 as $n, m \rightarrow \infty$. This is possible since for the chosen initial values $y$ we obtain that $\|y^{m}_u\|_Z$ is bounded on compact intervals in time uniformly in $m$ also with respect to the $Z$ norm. Indeed

$$
\|y\|_{Z} = \|S^w_{t} y_0 \| + \int_0^t \int_{\mathbb{R}_+} S^w_{t-s} g \left( \int \exp(\langle y_s, S^\nu_{t-s} \nu \xi \rangle) - 1 \right) \mu(d\xi) \right) ds \|z\|
$$

$$
\leq \|S^w_{t} y_0 \|_Z + \int_0^t \int_{\mathbb{R}_+} S^w_{t-s} g \left( \int \exp(\langle y_s, S^\nu_{t-s} \nu \xi \rangle) - 1 \right) \mu(dx) \right) ds
$$

$$
\leq \|S^w_{t} y_0 \|_Z + \mu(\mathbb{R}_+) \int_0^t \int_{\mathbb{R}_+} S^w_{t-s} g \|z\| du.
$$

From this we infer (4.17) with $a_{nm} = C\tilde{a}_{nm}$ for some constant $C$. The conditions of Theorem 3.2 are therefore satisfied and we obtain a generalized Feller semigroup whose generator corresponds to (4.10) with $\epsilon = 0$.

For the second assertion we proceed as in the proof of Proposition 4.16 the proof of the existence of $N$ can be transferred verbatim. However, when obtaining the mild formulation one looses the existence of right limits due to possible lack of finite mass of $\nu$.

Concerning the last assertion, the affine transform formula follows simply from the convergence of the semigroups $P^n$ as asserted in Theorem 3.2 by setting $y_\ell = \lim_{\epsilon \rightarrow 0} y_\ell^\epsilon$, where $y_\ell^\epsilon$ solves $\partial_t y_\ell^\epsilon = R^\epsilon(y_\ell^\epsilon)$ in the mild sense with $R^\epsilon$ given by (4.11).

Since $\exp(\langle y_0, \lambda \rangle)$ is then also the unique solution of the abstract Cauchy problem for initial value $\exp(\langle y_0, \lambda \rangle)$, i.e. it solves

$$
\partial_t u(t, \lambda) = A u(t, \lambda), \quad u(0, \lambda) = \exp(\langle y_0, \lambda \rangle),
$$
Theorem 4.18. Let Assumptions 4.10 and 4.6 be in force.

\[ m(4.20) \]

for some Brownian motion \( B \) of \( N \) where \( \text{exp}(y_t, \lambda) = \text{exp}(y_t, \lambda)R(y_t). \)

We are now ready to state our main theorem, namely an existence and uniqueness result for equations of the type

\[ d\lambda_t = A^*\lambda_t dt + \nu dX_t, \]

where \( (X_t)_{t \geq 0} \) is an Itô semimartingale of the form

\[ X_t = \int_0^t \beta(g, \lambda_s)ds + \int_0^t \sigma(y, \lambda_s)dB_t + \int_0^t \int \xi(\mu^y(d\xi, ds) - \langle g, \lambda_s \rangle m(d\xi)ds, \]

for some Brownian motion \( B \) and random jump measure \( \mu^X \) with \( \beta, \sigma \in \mathbb{R} \) and \( m(d\xi) \) is a Lévy measure on \( \mathbb{R}_{++} \) admitting a second moment.

**Theorem 4.18.** Let Assumptions 4.10 and 4.7 be in force.

(i) Then the stochastic partial differential equation (4.18) admits a unique Markovian solution \( (\lambda_t)_{t \geq 0} \) in \( \mathcal{E} \) given by a generalized Feller semigroup on \( B^0(\mathcal{E}) \) whose generator takes the form

\[ A f(y) = f(y)(\langle Ay, \lambda \rangle + \mathcal{R}(\langle y, \nu \rangle \langle g, \lambda \rangle) \]

on the set \( D \) defined in (4.16) where \( \mathcal{R} : \mathbb{R} \to \mathbb{R} \) is given by

\[ \mathcal{R}(u) = \frac{1}{2} u^2 + \int_{\mathbb{R}_+} \left( \exp(u\xi) - 1 - u\xi \right) m(d\xi). \]

(ii) This generalized Feller process is also a probabilistically weak and analytically mild càglàd solution of (4.18) i.e.

\[ \lambda_t = S^*_t \lambda_0 + \int_0^t S^*_{t-s} \nu dX_s, \]

in particular for every initial value the semimartingale \( X \) can be constructed on an appropriate probabilistic basis.

(iii) The affine transform formula is satisfied, i.e.

\[ \mathbb{E}_{\lambda_0} [\exp(y_t, \lambda_t)] = \exp(y_0, \lambda_0), \]

where \( y_t \) solves \( \partial_t y_t = R(y_t) \) in the mild sense with \( R : \mathcal{E} \to \mathbb{R} \) given by

\[ R(y) = Ay + \mathcal{R}(\langle y, \nu \rangle)g \]

with \( \mathcal{R} \) defined in (4.21). Furthermore, \( y_t \in \mathcal{E} \) for all \( t \geq 0 \).

(iv) For all \( \lambda_0 \in \mathcal{E} \), the corresponding jump diffusion stochastic Volterra equation, i.e.

\[ \langle g, \lambda_t \rangle = \langle g, S^*_t \lambda_0 \rangle + \int_0^t \langle g, S^*_{t-s} \nu \rangle dX_s = \langle g, S^*_t \lambda_0 \rangle + \int_0^t K(t-s)dX_s \]

admits a unique (probabilistically) weak solution.

**Proof.** To prove (i), we proceed again by applying Theorem 3.2. To this end consider the following sequence of SPDEs

\[ d\lambda^n_t = A^*\lambda^n_t dt + \beta \nu(g, \lambda^n_t)dt - n \sigma^2 \nu(g, \lambda^n_t)dt + \nu dN^n_{1,t} \]

\[ - \int_{\{t \geq s\}} \xi m(d\xi) \nu(g, \lambda^n_t)dt + \nu dN^n_{2,t} \]

where the compensator of the jump process \( N^n_{1,t} \) is given by \( \sigma^2 n^2 \langle g, \lambda^n_t \rangle \delta_{\frac{1}{n}}(d\xi) \) and of \( N^n_{2,t} \) by \( \int_{\{t \geq s\}} \xi m(d\xi) \nu(g, \lambda^n_t) \). In terms of (4.10), this corresponds to \( \varepsilon = 0, N = \ldots \)
\[ N_1 + N_2, w = -\beta + n + \int_{\{\xi > \frac{1}{\delta}\}} \xi m(d\xi) \text{ and } \mu^n(d\xi) := \sigma^2 n^2 \delta_{\frac{1}{\delta}}(d\xi) + m(1_{\{\xi \geq \frac{1}{\delta}\}} d\xi). \]

From Proposition [4.17] we thus know that these SPDEs admit a solution in terms of generalized Feller semigroups \( P^n \) with associated generators \( A^n \). Let us now establish a uniform growth bound for this sequence. Due to Proposition [4.17] we can establish the following estimate for \( t \in [0, T] \) for some fixed \( T > 0 \)

\[
\mathbb{E}[\| \lambda^n_t \|_{Y^*}] \leq 3\| S^n_t \lambda_0 \|_{Y^*}^2 + 3t \int_0^t \| S^n_{t-s} \nu \|_{Y^*}^2 \| g \|_{Y^*}^2 \mathbb{E}[\| \lambda^n_s \|_{Y^*}] ds \\
+ 3\mathbb{E}\left[ \int_0^t (S^n_{t-s} \nu dN^n_s - \int_0^t \int S^n_{t-s} \nu \xi \mu^n(d\xi | g, \lambda^n_s) ds)^2 \right] \\
\leq 3C_0\| \lambda_0 \|_{Y^*}^2 + 3t \int_0^t \| S^n_{t-s} \nu \|_{Y^*}^2 \| g \|_{Y^*}^2 \mathbb{E}[\| \lambda^n_s \|_{Y^*}] ds \\
+ 3\mathbb{E}\left[ \xi^2 \mu^n(d\xi) \int_0^t \| S^n_{t-s} \nu \|_{Y^*}^2 \langle g, \lambda^n_s \rangle ds \right] \\
\leq 3C_0\| \lambda_0 \|_{Y^*}^2 + 3(\sigma^2 + \int \xi^2 m(d\xi)) \int_0^t \| S^n_{t-s} \nu \|_{Y^*}^2 ds \\
+ 3\| g \|_{Y^*}^2 (t + \sigma^2 + \int \xi^2 m(d\xi)) \int_0^t \| S^n_{t-s} \nu \|_{Y^*}^2 \mathbb{E}[\| \lambda^n_s \|_{Y^*}] ds.
\]

Exactly as in Proposition [4.17], this now yields the uniform growth bound. For the set \( D \) as of Theorem [5.2] we consider again the Fourier basis elements as in [1.10], whose span is of course dense, whence (1). We equip span\( (D) \) again with the uniform norm \( \| \cdot \|_{\infty} \) and obtain a contraction property there. To verify Condition (ii) of Theorem [5.2] we check [4.17] in the present setting. Note that \( A^n \) takes the form

\[ A^n f_y(\lambda) = f_y(\lambda)(\langle Ay, \lambda \rangle + \beta g, \lambda \langle y, \nu \rangle + \langle y, \lambda \rangle \int (\exp(\langle y, \nu \xi \rangle) - 1 - \langle y, \nu \xi \rangle) \mu^n(d\xi)) \]

on \( D \) and from the last assertion of Proposition [4.17] we know that \( P^n \) leaves \( D \) invariant for all \( n \in \mathbb{N} \). By this invariance it thus suffices to check

\[ ||A^n f_y - A^n f_y||_0 \leq ||f_y||_{\infty} a_{nm} \]

for \( y \in K \), where \( K \) is a bounded set with respect to the \( Z \) norm. To this end consider for \( f_y \) for \( y \in K \),

\[
\left| A^n f_y(\lambda) - A^n f_y(\lambda) \right| \\
= \left| f_y(\lambda)(\langle g, \lambda \rangle + \beta g, \lambda \langle y, \nu \rangle + \langle y, \lambda \rangle \int (\exp(\langle y, \nu \xi \rangle) - 1 - \langle y, \nu \xi \rangle) (\mu^n(d\xi) - \mu^m(d\xi))) \right| \\
= \left| f_y(\lambda)(\langle g, \lambda \rangle \int (\xi^2 (y, \nu)^2 + o((y, \nu)^2 \xi^2))(\mu^n(d\xi) - \mu^m(d\xi))) \right| \\
= \left| f_y(\lambda)(\langle g, \lambda \rangle \left( | \int o((y, \nu)^2 \xi^2) (n^2 \delta_{\frac{1}{\delta}}(d\xi) - m^2 \delta_{\frac{1}{\delta}}(d\xi)) | \right) + | \int (\xi^2 (y, \nu)^2 + o((y, \nu)^2 \xi^2))(1_{\{\xi \geq \frac{1}{\delta}\}} - 1_{\{\xi \geq \frac{1}{\delta}\}}) m(d\xi) | \right) \\
= a_{nm} \]

From this we infer [4.23], where \( a_{nm} = C(a_{nm}^1 + a_{nm}^2) \) for some constant \( C \), for all \( y \in K \). The conditions of Theorem [5.2] are therefore satisfied and we obtain a generalized Feller semigroup whose generator is given by [4.20] on \( D \).
Concerning the second assumption, let us proceed similarly as in the proof of Proposition 4.16 (ii). Take \( y \in \text{dom}(A) \) such that \( \langle y, \nu \rangle \) is well-defined. Consider the local martingale

\[
M_t^y = \langle y, \lambda_t \rangle - \langle y, \lambda_0 \rangle - \int_0^t \langle Ay, \lambda_s \rangle - \beta \langle y, \nu \rangle \langle g, \lambda_s \rangle ds.
\]  

(4.24)

Let now \( y \) be such that \( \langle y, \nu \rangle = 1 \) and denote \( M_t^y \) for such \( y \) simply by \( M_t \). By the carré du champs formula the predictable quadratic variation of \( M_t \) is given by

\[
\langle M, M \rangle_t = \int_0^t (Af^2(\lambda_s) - 2f(\lambda_s)Af(\lambda_s))ds = \left( \sigma^2 + \int \xi^2 m(d\xi) \right) \int_0^t \langle g, \lambda_s \rangle ds
\]

for \( f(\lambda) = \langle y, \lambda \rangle \) and all \( \langle y, \nu \rangle = 1 \), which shows in particular that \( M_t \) is independent of the particular representative \( y \). Moreover, due to (4.22)

\[
\sum_{s \leq t} (\Delta M_s)^2 = \lim_{n \to \infty} \left( \frac{1}{n^2} K^n_1 + \sum_{s \leq t} (\Delta N^n_{2,s})^2 \right) = \lim_{n \to \infty} \sum_{s \leq t} (\Delta N^n_{2,s})^2
\]

where \( K^n_1 \) denotes the number of jumps of \( N^n_1 \) up to time \( t \). Hence, the compensator of \( \sum_{s \leq t} (\Delta M_s)^2 \) is \( \int_0^t \xi^2 m(d\xi) \langle g, \lambda_s \rangle ds \) and we deduce by the unique decomposition of \( M_t \) into a continuous and purely discontinuous local martingale \( M^c_t \) and \( M^d_t \), that \( \langle M^c_t, M^c_t \rangle = \sigma^2 \int_0^t \langle g, \lambda_s \rangle ds \). Therefore \( M_t \) can be written as

\[
M_t = M^c_t + M^d_t = \int \sigma \langle g, \lambda_s \rangle dB_s + \int_0^t \int \xi \eta (d\xi, ds) - \langle g, \lambda_s \rangle m(d\xi) ds,
\]

for some Brownian motion \( B \) and random measure \( \eta \). Define now \( X \) as

\[
X_t = \int_0^t \beta \langle g, \lambda_s \rangle ds + M_t.
\]

Then \( X_t \) is of form (4.19). Furthermore note that \( M^y_t \) of (4.24) for general \( y \in \text{dom}(A) \) such that \( \langle y, \nu \rangle \) is well-defined is given by \( M^y_t = \langle y, \nu \rangle M_t \) which can be seen again from the carré du champs formula, yielding \( \langle M^y_t, M^y_t \rangle = \langle y, \nu \rangle \langle M_t, M_t \rangle \). We thus deduce from (4.24)

\[
\langle y, \lambda_t \rangle = \langle y, \lambda_0 \rangle + \int_0^t \langle Ay, \lambda_s \rangle + \langle y, \nu \rangle X_t.
\]

The mild formulation then follows from this analytically weak solution by standard arguments.

The third assertion can be shown exactly along the lines of the proof of the last assertion of Proposition 4.17.

The last claim is any easy consequence of statement (ii).

\[\square\]

5. Concretely specifications

The goal of this section is to concretely specify the abstract setting introduced in Section 4 when we deal with specific kernels of form (4.3).
5.1. Lifting rough Volterra processes to measures. As already outlined in the introduction, we here consider the case of kernels given as Laplace transforms of some (signed) measure $\nu$ on $\mathbb{R}_+$, i.e.

$$K(t) = \int_0^\infty e^{-xt} \nu(dx),$$

such that $K(t) < \infty$ for all $t > 0$ and $K \in L^2_{\text{loc}}(\mathbb{R}_+, \mathbb{R})$.

In case that $\nu$ is a nonnegative measure, we are in the important setting of completely monotone kernels, in to which for instance the fractional kernels appearing in rough volatility models of the form $K(t) = \Gamma(\alpha)^{-1}t^{\alpha-1}$ for $\alpha \in (\frac{1}{2}, 1)$ fall. To cast this into the framework considered in Section 4, in particular Assumption 4.2, let $Y$ be the space of bounded continuous functions on the extended real half-line $\mathbb{R}_+ := \mathbb{R}_+ \cup \{\infty\}$. We here compactify $\mathbb{R}_+$, in order to make $C_0(\mathbb{R}_+)$ separable.

Its dual space $Y^*$ is the space of the finite (signed) regular Borel measures on the extend real half-line $\mathbb{R}_+$. As above we consider the weak-$\ast$-topology on $Y^*$ and a weight function

$$g(\lambda) = 1 + |\lambda|^2(\mathbb{R}_+),$$

where $|\lambda|$ denotes the total variation norm of $\lambda$. Moreover, $Z$ is the space of functions $g \in Y$ such that

$$(x \mapsto x g(x)) \in Y$$

together with the operator norm on it, i.e. $\|g\| = \sqrt{\|g\|^2 + \sup_{x \geq 0} |x g(x)|^2}$ for $g \in Z$.

Its dual $Z^*$ is the space of regular Borel measures $\nu$ on $\mathbb{R}_+$ that satisfy

$$\int_0^\infty \left(\frac{1}{x} \wedge 1\right) \nu(dx) < \infty.$$ 

As semigroup $(S^*_t)_{t \geq 0}$ acting on $Z^*$ and $Y^*$ we consider the multiplication semigroup

$$S^*_t \lambda = e^{-t} \lambda$$

which leaves $Z$, $Y$, $Y^*$ and $Z^*$ invariant and acts in a strongly continuous way thereon with respect to the strong norms. The linear operators on the dual spaces are weak-$\ast$-continuous. Whence, as stated in Remark 4.3, this semigroup then satisfies all requirements of Assumption 4.2 (iv) to (vi).

In terms of (4.3), $K$ can now be written as

$$K(t) = \int_0^\infty e^{-xt} \nu(dx) = (1, S^*_t \nu),$$

where the pairing is here $\langle y, \lambda \rangle = \int_0^\infty y(x) \lambda(dx)$. The element $g \in Y$ appearing in (4.3) is here simply the constant function $1$. Note also that the requirements of Assumption 4.4 are satisfied as well. Indeed $S^*_t \nu \in Y^*$ for all $t > 0$ and

$$\int_0^t \|S^*_s \nu\|^2_{Y^*} ds = \int_0^t (\int_0^\infty e^{-xs} \nu(dx))^2 ds = \int_0^t K(s)^2 ds < \infty$$

for all $t > 0$, since $K$ is assumed to lie in $L^2_{\text{loc}}(\mathbb{R}_+, \mathbb{R})$.

As in (4.7), we define for fixed $w > 0$

$$\mathcal{E}^w := \{\lambda_0 \in Y^* \mid (1, S^*_t \lambda_0) - \int_0^t R^w(t-s)(1, S^*_s \lambda_0) ds \geq 0 \text{ for all } t \geq 0\}$$

$$= \{\lambda_0 \in Y^* \mid \lambda_t \geq 0 \text{ with } \lambda_t(dx) = e^{-xt} \lambda_0(dx) - w \int_0^t e^{-(t-s)x} \nu(dx) \lambda_s ds \text{ for all } t \geq 0\},$$

where $R^w$ denotes the resolvent of $wK(t)$ and $\lambda = (1, \lambda)$ the total mass of $\lambda$. 


Example 5.1. Notice that the cone $\mathcal{E}^w$ is not locally compact with respect to the weak-$*$-topology in contrast to cases of non-negative measures. In the simplest case $K = 1$ a Laplace transform

$$g(t) = \int_0^\infty \exp(-xt)\lambda_0(dx)$$

lies in $\mathcal{E}^w$ if and only if $g(t) \geq 0$ and $g'(t) \geq 0$ for $t \geq 0$. Choose a positive measure $\lambda_0$ whose support does not contain $[1, 2]$. Then the signed measures $\lambda_n = \lambda_0(dx) + 1_{[1, 2]}\frac{1}{x} \sin(n\pi x)dx \in \mathcal{E}^w$, for $n \geq 1$, but $\lambda_n \to \lambda_0$ in the weak-$*$-topology. On the other hand the total variation norm of $\lambda_n$ does not converge to the total variation norm of $\lambda_0$. Hence by Remark 4.14 the cone $\mathcal{E}^w$ is not locally compact.

Assumption 4.10 now reads as follows:

**Assumption 5.2.** Let $\nu$ be such that $\exp(-t)\nu \in \mathcal{E}^w$ for all $t > 0$ and for all $w > 0$.

**Remark 5.3.** As stated in Remark 4.11 Assumption 5.2 is satisfied if $R^w \geq 0$ for all $w > 0$ and $K \geq 0$, which is satisfied if $K$ is completely monotone.

The state space that we consider for the stochastic processes in the sequel is as of (4.13), namely $\mathcal{E} = \cap_{w>0} \mathcal{E}^w$.

**Remark 5.4.** Comparing $\mathcal{E}$ with the state spaces considered in [1] (e.g. Equation (4.6)), we see that the conditions there translate to

$$\{\lambda_0 \in Y^* \mid t \mapsto \langle 1, S_t^* \lambda_0 \rangle = \int_0^\infty e^{-tx} \lambda_0(dx) \in \mathcal{G}_K \},$$

where $\mathcal{G}_K$ is defined in [1] Equation (2.5)]. From [1] Theorem A.2 we see that this describes a subspace of $\mathcal{E}$ in case a resolvent of the first kind exists for the kernel $K$.

It is interesting to consider the case when $\nu$ is a finite measure: in this case it is sufficient to consider the finite dimensional, invariant subspace of measures $\lambda \ll \nu$. We identify it via $\lambda \mapsto \frac{d\lambda}{d\nu}$ with some $\mathbb{R}^N$ and we denote $\text{supp}(\nu) = \{0 \leq x_1 < \cdots < x_N \}$.

**Proposition 5.5.** Under Assumptions 5.2 the state space $\mathcal{E}$ is a well defined closed cone given as the intersection of $\cap_{w>0} \mathcal{E}^w$ where for $w > 0$

$$\mathcal{E}^w = \{\lambda_0 \in \mathbb{R}^N \mid (e^{tA^w} \lambda_0, 1)_{\mathbb{R}^N} \geq 0 \text{ for all } t \geq 0\}$$

with $A^w = \text{diag}(-x_1, \ldots, -x_N) - w(\nu(x_1)1, \ldots, \nu(x_N)1)^	op$. Here, $1 \in \mathbb{R}^N$ is the (column) vector consisting of ones.

**Proof.** Note that the deterministic equation as of (4.2) has here the following form

$$d\lambda_t(x_i) = -x_i \lambda_t(x_i)dt - w(\nu(x_1)1, \ldots, \nu(x_N)1)\sum_{j=1}^N \lambda_t(x_j)dt, \quad i \in \{1, \ldots, N\}.$$ 

This can be written as $d\lambda_t = A^w\lambda_t dt$ so that the assertion follows from the definition of $\mathcal{E}^w$. □

**Example 5.6.** There are simple two dimensional examples when $\mathcal{E}^0 \neq \mathcal{E}^w$ for $w > 0$: indeed take $x_0 = 0$ and $x_1 = 1$ and $\nu(dx) = \delta_0(dx) + \delta_1(dx)$. In this case

$$\mathcal{E}^0 = \text{cone}(e_2, e_1 - e_2)$$

whereas

$$\mathcal{E}^w = \text{cone}(e_1 + e_2, e_1 - e_2).$$
Here $e_i, i = 1, 2$ denotes the canonical basis vectors of $\mathbb{R}^2$. Observe here that $\mathcal{E}^w$ does not depend on $w$ for $w > 0$, but it changes when $w \to 0$.

Let us finally come to the SPDE formulation which will in particular lead to rough affine volatility models. Consider the following measure valued SPDE

$$d\lambda_t(dx) = -x\lambda_t(dx)dt + \nu(dx)dX_t,$$

(5.2)

where $(X_t)_{t \geq 0}$ is an Itô semimartingale of the form

$$X_t = \int_0^t \beta \lambda_s ds + \int_0^t \sigma \sqrt{\lambda_s} dB_t + \int_0^t \xi(\mu_X(d\xi, ds) - \lambda_s m(d\xi))ds,$$

(5.3)

for some Brownian motion $B$ and random jump measure $\mu_X$ with $\beta, \sigma \in \mathbb{R}$ and $m(d\xi)$ is a Lévy measure on $\mathbb{R}_{++}$ admitting a second moment. Recall here that $\lambda = \langle 1, \lambda \rangle$.

As a corollary of Theorem 4.18 we now obtain the following result.

**Theorem 5.7.** Let Assumption 5.2 be in force.

(i) Then the stochastic partial differential equation (5.2) admists a unique Markovian solution $(\lambda_t)_{t \geq 0}$ in $\mathcal{E}$ given by a generalized Feller semigroup on $\mathcal{B}^\infty(\mathcal{E})$ whose generator takes the form

$$Af(\lambda) = f'(\lambda)\int_0^\infty -xy(x)\lambda(dx) + \mathcal{R}(\langle y, \nu \rangle) \lambda$$

(5.4)

on the set $D$ defined in (4.10) where $\mathcal{R} : \mathbb{R}^- \to \mathbb{R}$ is given by

$$\mathcal{R}(u) = \beta u + \frac{1}{2} \sigma^2 u^2 + \int_{\mathbb{R}^+} (\exp(u\xi) - 1 - u\xi) m(d\xi).$$

(5.5)

(ii) This generalized Feller process is also a probabilistically weak and analytically mild càg solution of (5.2) i.e.

$$\lambda_t(dx) = e^{-tx} \lambda_0(dx) + \int_0^t e^{(t-s)x} \nu(dx)dX_s,$$

in particular for every initial value the semimartingale $X$ can be constructed on an appropriate probabilistic basis.

(iii) The affine transform formula is satisfied, i.e.

$$\mathbb{E}_{\lambda_0}[\exp(\langle y_t, \lambda_t \rangle)] = \exp(\langle y_t, \lambda_0 \rangle),$$

where $y_t$ solves $\partial_t y_t = R(y_t)$ in the mild sense with $R : \mathcal{E} \to \mathbb{R}$ given by

$$R(y)(x) = -xy(x) + \mathcal{R}(\langle y, \nu \rangle)$$

with $\mathcal{R}$ defined in (4.21). Furthermore, $y_t \in \mathcal{E}$ for all $t \geq 0$.

(iv) For all $\lambda_0 \in \mathcal{E}$, the corresponding jump diffusion stochastic Volterra equation, i.e.

$$\lambda_t = \int_0^t e^{-tx} \lambda_0(dx) + \int_0^t K(t-s)dX_s$$

with $X$ given by (5.3) admits a unique (probabilistically) weak solution in $\mathbb{R}_+$. 
5.2. Lifting rough Volterra processes to forward curves. There are several ways to lift stochastic Volterra processes to Markovian processes: in the previous section the focus was on the completely monotone nature of many Volterra kernels. The lift that we treat here considers instead the stochastic Volterra process together with its conditional expectations from future values. As in the Heath-Jarrow-Morton (HJM) case, this yields shift semigroups. This Markovian lift also falls in the realm of Section 4 and can be considered from the perspective of generalized Feller semigroups. Indeed, here $K$ (satisfying Conditions 5.8 below) is represented by

$$K(t) = \int_0^\infty S^*_t K(x)\delta_0(dx),$$

where $S^*_t$ denotes the shift semigroup. In order meet all the conditions of Section 4, we here assume the following:

**Assumption 5.8.**

(i) $K$ is assumed to lie in $L^2_{loc}(\mathbb{R}_+, \mathbb{R})$.

(ii) there exists $L \in AC(\mathbb{R}_+, \mathbb{R})$, where $AC(\mathbb{R}_+, \mathbb{R})$ denotes the space of real-valued absolutely continuous functions on $\mathbb{R}_+$, such that $K = L'$.

(iii) There exists some strictly positive weight function $v > 0$ such that

$$\int_0^\infty |K(x)|^2 v(x)dx = \int_0^\infty |L'(x)|^2 v(x)dx < \infty.$$

In terms of the underlying Banach spaces we we let $Y$ and in turn also $Y^*$ be Hilbert spaces, similarly as in Filipovic [16], namely

$$Y = \{y \in AC(\mathbb{R}_+, \mathbb{R}) | \int_0^\infty |y'(x)|^2 v(x)dx < \infty\}$$

for the specified strictly positive weight function $v > 0$. We endow $Y$ with the scalar product

$$\langle y, \lambda \rangle_v = y(0)\lambda(0) + \int_0^\infty y'(x)\lambda'(x)v(x)dx$$

and denote the associated norm via $\| \cdot \|_v$.

**Assumption 5.9.** We furthermore assume that the shift semigroup $S^*$ acts in a strongly continuous way on $Y^*$, which can be easily expressed in terms of the function $v$, and that $S^*_t K \in Y^*$ for $t > 0$.

**Remark 5.10.** One can choose $v(x) = \exp(\alpha x)$ for $x \geq 0$, which proved to be useful in case of term structures of interest rates.

The weight function, in the sense weighted spaces, is again given by

$$\varrho(\lambda) = 1 + \|\lambda\|_v^2.$$

Moreover, for $Z$ we consider the subspace of functions vanishing at 0 and whose first (weak) derivative lies in $Y$ (with the corresponding operator Hilbert space norm), then $K \in Z^*$ can be viewed as a weak derivative of a function $L \in Y^*$ with norm

$$\|K\|_{Z^*}^2 := \int_0^\infty |L'(x)|^2 v(x)dx.$$ 

In particular the value $\infty$ at 0 is possible. By condition (ii) and (iii) of Assumption 5.8 $K \in Z^*$. Notice also the following inclusions

$$Z \subset Y = Y^* \subset Z^*.$$ 

In terms of (4.3), $K \in Z^*$ can now be written as

$$K(t) = \langle S^*_t K, 1 \rangle_v = \langle S^*_t K \rangle(0).$$

The element $g \in Y$ appearing in (4.3) is here again simply the constant function 1.
Note that the shift semigroup \((S_t^*)_{t \geq 0}\) acting on \(Z^*\) and \(Y^*\) satisfies all requirements of Assumption 4.2 (iv) to (v). Concerning (vi), the (pre)adjoint \((S_t^*)_{t \geq 0}\) acts on \(Z^*\) and \(Y^*\) via the negative shift defined as follows
\[
S_t^*y(x) = \begin{cases} 
y(x - t) & \text{if } x - t \geq 0, \\
0 & \text{else.}
\end{cases}
\]
Notice that the pre-dual semigroup \(S^*\) is only acting on \(Z\) in a strongly continuous way. Furthermore the requirements of Assumption 4.6 are met as well.

As in (4.7), we define for fixed \(w > 0\)
\[
E^w := \{\lambda_0 \in Y^* | (1, S_t^*\lambda_0) - \int_0^t R^w(t - s)(1, S_s^*\lambda_0) ds \geq 0 \text{ for all } t \geq 0\}
\]
(5.6)

where \(R^w\) denotes the resolvent of \(wK(t)\). Note that \(\langle 1, \lambda \rangle = \lambda(0)\). Again the state space will be.
\[
E = \cap_{w > 0} E^w.
\]

Assumption 4.10 now reads as follows:

**Assumption 5.11.** Let \(K\) be such that \(K(t + \cdot) \in E^w\) for all \(t > 0\) and for all \(w > 0\).

The SPDE that we consider in the present setting is
\[
d\lambda_t(x) = \frac{d}{dx} \lambda_t(x) dt + K(x) dX_t,
\]
where \((X_t)_{t \geq 0}\) is an Itô semimartingale of the form
\[
X_t = \int_0^t \beta \lambda_s(0) ds + \int_0^t \sigma \sqrt{\lambda_s(0)} dB_s + \int_0^t \int \xi \left(\mu^X(d\xi, ds) - \lambda_s(0) m(d\xi) ds, \right.
\]
for some Brownian motion \(B\) and random jump measure \(\mu^X\) with \(\beta, \sigma \in \mathbb{R}\) and \(m(d\xi)\) is a Lévy measure admitting a second moment. As a corollary of Theorem 4.18 we now obtain the following result.

**Theorem 5.12.** Let Assumptions 5.8, 5.9 and 5.11 be in force.

(i) Then the stochastic partial differential equation (5.7) admists a unique Markovian solution \((\lambda_t)_{t \geq 0}\) in \(E\) given by a generalized Feller semigroup on \(B^2(E)\) whose generator takes the form
\[
Af_y(\lambda) = f_y(\lambda)(\langle -\frac{d}{dx} y, \lambda \rangle_v + \mathcal{R}(\langle y, K \rangle_v) \lambda(0)
\]
on the set \(D\) defined in (4.10) where \(\mathcal{R} : \mathbb{R}_- \to \mathbb{R}\) is given by
\[
\mathcal{R}(u) = \beta u + \frac{1}{2} \sigma^2 u^2 + \int_{\mathbb{R}_+} (\exp(u\xi) - 1 - u\xi) m(d\xi).
\]

(ii) This generalized Feller process is also a probabilistically weak and analytically mild càg solution of (5.7) i.e.
\[
\lambda_t(x) = \lambda_0(t + x) + \int_0^t K(t - s + x) dX_s
\]
in particular for every initial value the semimartingale \(X\) can be constructed on an appropriate probabilistic basis.
The affine transform formula is satisfied, i.e.
\[ E_{\lambda_0} \left[ \exp\left( \langle y_0, \lambda_t \rangle \right) \right] = \exp\left( \langle y_t, \lambda_0 \rangle \right) \]
where \( y_t \) solves \( \partial_t y_t = R(y_t) \) in the mild sense with \( R : \mathcal{E}_* \to \mathbb{R} \) given by
\[ R(y)(x) = -\frac{d}{dx}y(x) + R(\langle y, \nu \rangle_v) \]
with \( R \) defined in (4.21). Furthermore, \( y_t \in \mathcal{E}_* \) for all \( t \geq 0 \).

For all \( \lambda_0 \in \mathcal{E}_* \), the corresponding jump diffusion stochastic Volterra equation, i.e.
\[ \lambda_t(0) = \lambda_0(t) + \int_0^t K(t-s)dX_s \]
admits a unique (probabilistically) weak solution in \( \mathbb{R}_+ \).

References

[1] E. Abi Jaber and O. E. Euch. Markovian structure of the Volterra Heston model. arXiv:1803.00477, 2018.
[2] E. Abi Jaber and O. E. Euch. Multi-factor approximation of rough volatility models. arXiv:1801.10359, 2018.
[3] E. Abi Jaber, M. Larsson, and S. Pulido. Affine Volterra processes. arXiv:1708.08796, 2017.
[4] C. D. Aliprantis and K. C. Border. Infinite dimensional analysis. Springer, Berlin, third edition, 2006. A hitchhiker’s guide.
[5] V. S. Borkar. Probability theory. Universitext. Springer-Verlag, New York, 1995. An advanced course.
[6] A. Boumezoued. Population viewpoint on Hawkes processes. Advances in Applied Probability, 48(2):463–480, 2016.
[7] R. C. Dalang. Extending the martingale measure stochastic integral with applications to spatially homogeneous SPDEs. Electron. J. Probab., 4:no. 6, 29 pp. 1999.
[8] D. A. Dawson and E. Perkins. Superprocesses at Saint-Flour. Probability at Saint-Flour. Springer, Heidelberg, 2012.
[9] P. Dörsek and J. Teichmann. A Semigroup Point of View on Splitting Schemes for Stochastic (Partial) Differential Equations. arXiv:1011.2651 2010.
[10] O. E. Euch and M. Rosenbaum. Perfect hedging in rough Heston models. arXiv:1703.05049, 2017.
[11] K.-J. Engel and R. Nagel. One-parameter semigroups for linear evolution equations, volume 194 of Graduate Texts in Mathematics. Springer-Verlag, Berlin, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.
[12] S. Ethier and T. G. Kurtz. Markov processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
[13] D. Filipović. Consistency Problems for Heath-Jarrow-Morton Interest Rate Models, volume 1760 of Lecture Notes in Mathematics. Springer Verlag, Berlin, 2001.
[14] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. SSRN:2509457, 2014.
[15] G. Gripenberg, S.-O. Londen, and O. Staffans. Volterra integral and functional equations, volume 34 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1990.
[16] A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. Biometrika, 58(1):83–90, 1971.
[17] O. E. Euch and M. Rosenbaum. The characteristic function of rough Heston models. arXiv:1609.05177, 2016.
[18] J. Myktyuk and T. Salisbury. Uniqueness for Volterra-type stochastic integral equations. arXiv:1502.05515, 2015.
[19] A. Pap. Semigroups of linear operators and applications to partial differential equations, volume 44 of Applied Mathematical Sciences. Springer-Verlag, New York, 1983.
[23] H. H. Schaefer and M. P. Wolff. *Topological vector spaces*, volume 3 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1999.

Vienna University, Oskar-Morgenstern-Platz 1, A-1090 Vienna and ETH Zürich, Rämistrasse 101, CH-8092 Zürich