Duality and the well-posedness of a martingale problem

ANDREJ DEPPERSCHMIDT, ANDREAS GREVEN, PETER PFAFELHUBER

May 26, 2023

Abstract
For two Polish state spaces $E_X$ and $E_Y$, and an operator $G_X$, we obtain existence and uniqueness of a $G_X$-martingale problem provided there is a bounded continuous duality function $H$ on $E_X \times E_Y$ together with a dual process $Y$ on $E_Y$ which is the unique solution of a $G_Y$-martingale problem. For the corresponding solutions $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$, duality with respect to a function $H$ in its simplest form means that the relation $E_x [H(X_t, y)] = E_y [H(x, Y_t)]$ holds for all $(x, y) \in E_X \times E_Y$ and $t \geq 0$. While duality is well-known to imply uniqueness of the $G_X$-martingale problem, we give here a set of conditions under which duality also implies existence without using approximating sequences of processes of a different kind (e.g. jump processes to approximate diffusions) which is a widespread strategy for proving existence of solutions of martingale problems. Given the process $(Y_t)_{t \geq 0}$ and a duality function $H$, to prove existence of $(X_t)_{t \geq 0}$ one has to show that the r.h.s. of the duality relation defines for each $y$ a measure on $E_X$, i.e. there are transition kernels $(\mu_t)_{t \geq 0}$ from $E_X$ to $E_X$ such that $\mathbb{E}_y [H(x, Y_t)] = \int \mu_t(x, dx') H(x', y)$ for all $(x, y) \in E_X \times E_Y$ and all $t \geq 0$.

As examples, we treat resampling and branching models, such as the Fleming-Viot measure-valued diffusion and its spatial counterparts (with both, discrete and continuum space), as well as branching systems, such as Feller’s branching diffusion. While our main result as well as all examples come with (locally) compact state spaces, we discuss the strategy to lift our results to genealogy-valued processes or historical processes, leading to non-compact (discrete and continuum) state spaces. Such applications will be tackled in forthcoming work based on the present article.

1 Introduction
A general method for constructing a class of time-homogeneous Markov processes on a Polish state space $E$ with measurable paths is by using martingale problems, which we briefly recall.

Given a linear operator $G$ on a domain $\mathcal{D}$ which is a subspace of measurable, real-valued functions on $E$, and an initial law $\mathbb{P}_0 \in M_1(E)$, the set of probability measures on $E$, we say that the distribution $\mathbb{P}$ of an $E$-valued progressively measurable stochastic process $Z$ solves the martingale problem for $(G, \mathcal{D}, \mathbb{P}_0)$, if $\mathbb{P}(Z_0 \in \cdot) = \mathbb{P}_0(\cdot)$ and

\[
\left( f(Z_t) - \int_0^t G f(Z_s) \, ds \right)_{t \geq 0}
\] (1.1)
is a \( \mathbb{P} \)-martingale (with respect to the filtration generated by \( X \)) for all \( f \in D \). By a martingale problem for \((G, D, z)\) for a \( z \in E \), we mean the martingale problem with initial measure \( \mathbb{P}_0 = \delta_z \).

The martingale problem for \((G, D, \mathbb{P}_0)\) is called well-posed, if a solution exists and is unique. We say that the martingale problem for \((G, D)\) is well-posed if the martingale problem for \((G, D, \mathbb{P}_0)\) is well-posed for all \( \mathbb{P}_0 \in M_1(E) \).

**Remark 1.1** (Path regularity). Recall that a solution of a martingale problem must have a modification with measurable paths to ensure existence of the integral in (1.1) [EK86, Section 4.3], and therefore has a progressively measurable modification [KS91, Proposition 1.12]. So, without losing generality, the above definition uses a strong form of uniqueness and therefore has a progressively measurable modification [KS91, Proposition 1.12]. So, with identification with measurable paths to ensure existence of the integral in (1.1) [EK86, Section 4.3], ensures the existence of a càdlàg modification of the solution of the \((G_X, \mathcal{H}_X)\) martingale problem provided the compact containment condition holds.

Duality, which we recall below, is a technique often used to show uniqueness of solutions of a martingale problem. For existence however, a typical strategy is to construct a tight sequence \( Z^1, Z^2, \ldots \) of approximating processes (typically some pure jump Markov processes), to prove tightness of the laws and to show that every limit point solves the martingale problem. The main goal of the paper is to use duality also for existence of solutions of martingale problems; see Theorem 2.1. This approach avoids approximations with processes of a different nature than the solutions of the martingale problem. Note however that we also provide in Corollary 2.7 a method to obtain solutions by approximations where the existence (and uniqueness) of the approximating sequences themselves is obtained using duality.

Two processes \( X \) and \( Y \) with Polish state spaces \( E_X \) and \( E_Y \), which arise as solutions of martingale problems \((G_X, D_X)\) respectively \((G_Y, D_Y)\), are said to be dual with respect to a bounded, continuous function \( H : E_X \times E_Y \rightarrow \mathbb{R} \), if

$$
\mathbb{E}_{\mathbb{P}_0}[H(X_t, y)] = \int_{E_X} \mathbb{E}_y[H(x, Y_t)] \mathbb{P}_0(dx), \quad \mathbb{P}_0 \in M_1(E_X), \ y \in E_Y, \tag{1.2}
$$

where \( \mathbb{E}_{\mathbb{P}_0}[\cdot] \) and \( \mathbb{E}_y[\cdot] \) denote the expectations with respect to the initial conditions \( X_0 \sim \mathbb{P}_0 \) and \( Y_0 = y \), respectively. In particular, properties of \( X \) can be read off from properties of \( Y \) and vice versa. (We note that more general notions of duality exist, where one or both sides of (1.2) contain an exponential penalty term, usually called Feynman-Kac-term; see (2.5) below. Also, the boundedness of \( H \) can be relaxed in which case some additional integrability conditions have to be checked; see Remark 2.4.)

Usually, (1.2) is proved as follows (cf. (4.39)–(4.42) in Chapter 4 in [EK86]): If \( G_X \) and \( G_Y \) are operators with domains \( D_X \supseteq \mathcal{H}_X \) := \{\( H(\cdot, y) : y \in E_Y \)\} and \( D_Y \supseteq \mathcal{H}_Y \) := \{\( H(x, \cdot) : x \in E_X \)\}, respectively, and if \( X \) and \( Y \) are solutions of the corresponding martingale problems then (1.2) is equivalent to

$$
G_XH(\cdot, y)(x) = G_YH(x, \cdot)(y), \quad x \in E_X, y \in E_Y. \tag{1.3}
$$
In order to see that this suffices for (1.2), take a probability space where \(X\) and \(Y\) are independent and conclude from (1.3) that

\[
\frac{d}{ds} \mathbb{E}[H(X_s, Y_{t-s})] = \mathbb{E}[G_X H(\cdot, Y_{t-s})(X_s)] - \mathbb{E}[G_Y H(X_s, \cdot)(Y_{t-s})] = 0. \tag{1.4}
\]

In addition, (1.3) is necessary for (1.2) since for \(x \in E_X, y \in E_Y\)

\[G_X H(\cdot, y)(x) - G_Y H(x, \cdot)(y) = \lim_{h \to 0} \frac{1}{h} \left( \mathbb{E}_{\delta_x}[H(X_t, y)] - H(x, y) - \mathbb{E}_{\delta_y}[H(x, Y_t)] + H(x, y) \right) = 0.\]

A classical result addresses the uniqueness of the martingale problem for \((G_X, D_X, \mathbb{P}_0)\); see e.g. Proposition 4.4.7 and Remark 4.4.8 in [EK86]. If \(E_X\) and \(E_Y\) are Polish, \(\mathcal{H}_X\) is separating on the space of probability measures on \(E_X\), and if for every \(y \in E_Y\), there exists a solution \(Y\) of the martingale problem for \((G_Y, \mathcal{H}_Y, \mathbb{P})\), and if (1.2) holds for all \(x \in E_X\) with \(\mathbb{P}_0 = \delta_x\) and \(y \in E_Y\), then uniqueness of the martingale problem for \((G_X, \mathcal{H}_X, x)\) holds. Also uniqueness of the more general martingale problems for \((G_X, \mathcal{H}_X, \mathbb{P}_0)\) with random initial conditions \(\mathbb{P}_0 \in \mathcal{M}_1(E_X)\) holds. The reason is that the duality relation (1.2) and separability of \(\mathcal{H}_X\) specify the one-dimensional distributions of \(X\) uniquely, and therefore, by [EK86, Theorem 4.4.2], uniqueness of the martingale problem follows.

Duality is also very useful if \(Y\) is a much simpler process than \(X\), because questions concerning the behaviour of \(X\) can be translated to questions about \(Y\). For example, duality can be used to show the Feller property of \(X\), or to determine its longtime behaviour and characterize equilibria. Duality was the key tool for studying interacting particle systems such as the voter model and the contact process [Lig85], but also for measure-valued processes such as the Fleming-Viot process (which is dual to some form of coalescent process; see also Examples [1, 4] and [4]), and the Dawson-Watanabe superprocess (which is dual to the solution of a deterministic process given by a log-Laplace equation) [Daw93, Eth00]. For a general reference on duality for Markov processes including various sorts of applications see [EK86, JK14] and references therein.

The idea to use duality for the existence of a solution of a martingale problem was motivated by constructions appearing in the literature. To the best of our knowledge, the first examples appear in [EP94, Eva97], where duality is used to show existence of the continuum space version of interacting Fisher-Wright diffusions on the discrete hierarchical group, by lifting the duality relation from the corresponding discrete case. This has been studied in on \(\mathbb{Z}^1\) and \(\mathbb{R}^1\) with other methods in [KS88, Shi94, MT95]. We believe that our approach provides proofs of these results (when formulated differently) as well. The approach using duality is also used in [GdHKK14] to construct a spatial Cannings model, and in [BEV13, EVY20] for the construction of a model with locally constant population size in a spatial continuum. For a branching process, Dynkin gave in [Dyn93] – what he called – a direct construction, which can be viewed as a construction based on the deterministic dual (as opposed to the construction via particle approximations in [Daw93] for example).

We give here a systematic approach to the existence problem together with some examples. Let us briefly describe the idea for showing existence by using a dual process; see Theorem 2.1 for all details: We are given the \((G_X, D_X, \mathbb{P}_0)\) martingale problem for which we want to establish
well-posedness. We look both for a Markov process $Y$ and a function $H$ for which the relation \[ (1.3) \] holds. Then we define the operator $P_t$ on $\mathcal{H}_X$ by setting $(P_t H(\cdot), y)(x) := \mathbb{E}_y[H(X_t, y)]$, which defines an operator on $\mathcal{H}_X$. Then $P_t$ inherits the semigroup property $P_t \circ P_s = P_{t+s}$ from the semigroup of the dual process $Y$. The semigroup $(P_t)_{t \geq 0}$ will be the semigroup of some process $X$, provided there is a probability measure $\mathbb{P}_x$ (with expectation $\mathbb{E}_x$) and for each $t \geq 0$ a random variable $X_t$ such that \[ \mathbb{E}_x[H(X_t, y)] = (P_t H(\cdot, y))(x) := \mathbb{E}_y[H(x, Y_t)] . \] (1.5) Then $(P_t)_{t \geq 0}$ is a Markov semigroup and we have existence of a solution of the martingale problem for $(G_X, \mathcal{H}_X, x)$ provided some additional measurability property holds. Moreover, if duality is derived from the operator criterion it also implies uniqueness. Altogether we obtain well-posedness of the martingale problem for $(G_X, \mathcal{H}_X, x)$ for each $x \in E_X$. From that we obtain the well-posedness of the $(G_X, \mathcal{H}_X, \mathbb{P}_0)$ martingale problem for any $\mathbb{P}_0 \in \mathcal{M}_1(E_X)$. At least on compact state spaces, the existence of a càdlàg modification is immediate.

The main requirement in applying our main result, namely Theorem 2.1 is to find (the distribution of) $X_t$ satisfying (1.5). For this, we provide two general approaches, one based on the Riesz-Markov Theorem in Proposition 2.6 which requires compact state spaces. In various applications, relaxing the assumption of compactness of $E_X$ is the main challenge. An approach in this direction is Proposition 2.8 which requires $E_Y$ to be a set of functions on some compact set.

For the construction of a solution of a martingale problem using duality we give several examples. Since our motivation came from [Eva97], we also discuss here resampling systems with our approach. Namely in Examples 1-5 we show how our results can be used for the (spatial) Fleming-Viot process (with mutation) and the Cannings model, as well as the spatial Fleming-Viot process from [BEV13]. In addition, we adapt arguments from [Dyn93] and [Bez11] in order to show existence in a continuous state branching model; see Example 6. We also give an example how to use a Feynman-Kac term, by using the duality of the Feller branching diffusion to a Kingman coalescent; see Example 7.

In future work, we want to systematize the approach to be able to construct genealogy-valued processes based on martingale problems as introduced in [GPW13, DGP12, GSW16] and which could be generalized to genealogy-valued Fleming-Viot models with recombination using arguments of the present paper to construct and characterize these new processes. Compare here also Section 5 for more details. Another possibility is to use the approach to construct continuum space dynamics, which was also the original motivation in [EF96] and this is taken up in work of Etheridge and coathors on $\mathbb{R}^d$ [EVY20] and in [GSW16] and subsequently on the continuum space hierarchical group extending [GdHKK14].

**Remark 1.2** (Other methods for showing existence). Let us discuss two more options to show existence of solutions of a $(G_X, D_X, \mathbb{P}_0)$ martingale problem without using a tight sequence of approximating processes: the positive maximum principle and the Girsanov transform. For the former, consider locally compact $E_X$. Here, if (i) $G_X$ satisfies the positive maximum principle (i.e. if $f \in D_X$ and $x_0 \in E$ such that $\sup_x f(x) = f(x_0) \geq 0$, then $G_X f(x_0) \leq 0$) and (ii) is conservative (i.e. there is $f_1, f_2, \cdots \in D_X$ with $f_n \xrightarrow{\text{a.s.}} 1$ and $G_{f_n} \xrightarrow{\text{a.s.}} 0$ boundedly
pointwise), existence follows (see e.g. [EK86, Theorem 4.5.4 and Remark 4.5.5]). However, we note that the positive maximum principle is very often not straightforward to verify, for example in systems with infinitely many components.

The Cameron-Martin-Girsanov theorem is another way to show existence of solutions of the $(G_X, \mathcal{D}_X, \mathbb{P}_0)$ martingale problem for, given (i) existence of a process $Z$, (ii) a mean-1-martingale $M \geq 0$ and (iii) a proof that $(M \cdot \mathbb{P})_t Z$ (here $M \cdot \mathbb{P}$ denotes the probability measure with density $M$ with respect to $\mathbb{P}$) solves the $(G_X, \mathcal{D}_X, \mathbb{P}_0)$ martingale problem. However, it might here be necessary to prove existence of the process $Z$ by some other methods, for instance by again using approximation techniques or the positive maximum principle. $\triangle$

For future reference we introduce in the following remark the notation used throughout the paper. The reader might skip it and return to it if the notation that we use is not familiar.

**Remark 1.3** (Notation and some basic concepts). Throughout, let $(E, r)$ be a complete and separable metric space. Also, let $C_b(E)$ and $\mathcal{B}(E)$ be the spaces of real-valued, continuous and bounded respectively bounded measurable functions. With a slight abuse of notation, we also write $\mathcal{B}(E)$ for the set of Borel-measurable subsets of $E$. On $C_b(E)$, we use the supremum norm $\|\cdot\|$ and equip $C_b(E)$ with the bounded pointwise (bp)-topology where $f_n \to f$ iff $\sup_n \|f_n\| < \infty$ and $f_n \to f$ pointwise. We denote by $\mathcal{M}(E)$ ($\mathcal{M}_1(E)$) the space of (probability) Radon measures on $E$. If $E$ is locally compact, we denote by $\widetilde{C}(E) \subseteq C_b(E)$ the set of continuous functions vanishing at infinity. For $E$-valued random variables $Y, Z$, we write $Y \sim Z$ or $Y \sim \mathcal{L}(Z)$ if $Y$ and $Z$ have the same distribution.

We say that $\Pi \subseteq C_b(E)$ is separating (on $\mathcal{M}_1(E)$) if for all $\mu, \nu \in \mathcal{M}_1(E)$, $(\int f \, d\mu = \int f \, dv$ for all $f \in \Pi) \Rightarrow \mu = \nu$ holds, and convergence determining (in $\mathcal{M}_1(E)$) if, for all $\mu, \mu_1, \mu_2, \ldots$, $(\int f \, d\mu_n \underset{n \to \infty}{\longrightarrow} \int f \, d\mu$ for all $f \in \Pi) \Rightarrow (\mu_n \overset{\ast}{\longrightarrow} \mu)$ holds.

Recall that a semigroup $(P_t)_{t \geq 0}$ on a vector space $\mathcal{D} \subseteq C_b(E)$ is a family of bounded linear functions $P_t : \mathcal{D} \to \mathcal{B}(E)$ such that $P_t(P_{s} f) = P_{t+s} f$ for all $t, s \geq 0$ and $f \in \mathcal{D}$ with $P_{t} f \in \mathcal{D}$. The operator $P_t$ (or the semigroup $(P_t)_{t \geq 0}$) is a contraction if $\|P_t f\| \leq \|f\|$ (for all $t \geq 0$). It is positive if $P_t f \geq 0$ for $f \geq 0$. It is conservative if $P_t 1 = 1$. A semigroup $(P_t)_{t \geq 0}$ is called strongly continuous if $P_t f \underset{t \to 0}{\longrightarrow} f$ for all $f \in C_b(E)$. If a conservative, positive, strongly continuous contraction semigroup satisfies $P_t f \in C_b(E)$ for $f \in C_b(E)$ and $t \geq 0$, we call $(P_t)_{t \geq 0}$ a $C_b(E)$-Feller semigroup. If the same holds for locally compact $E$ with $\widetilde{C}(E)$ instead of $C_b(E)$, then we say that $(P_t)_{t \geq 0}$ is a $\widetilde{C}(E)$-Feller semigroup. The generator of a semigroup $(P_t)_{t \geq 0}$ is given by $Gf(x) = \lim_{t \to 0} \frac{1}{t}(P_tf(x) - f(x))$, whenever the limit exists boundedly pointwise. The set $\mathcal{D}(G)$ of functions for which the limit exists boundedly pointwise is referred to as the domain of the generator $G$.

Recall that with any time-homogeneous Markov process $X = (X_t)_{t \geq 0}$ on a state space $E$ we can associate a semigroup $P = (P_t)_{t \geq 0}$ with $P_t f(x) = \mathbb{E}_x[f(X_t)]$ satisfying the Chapman-Kolmogorov equations $P_s P_t f = P_{s+t} f$ for $s, t \geq 0$. This semigroup is a positive, conservative contraction. We say that $X$ is a Feller process if its semigroup is Feller (with respect to either $C_b(E)$ or $\widetilde{C}(E)$). $\triangle$
2 Results

We will first present in Theorem 2.1 the general result on the well-posedness of a martingale problem using duality in Section 2.1. Then in Section 2.2 we will discuss how to check the assumptions appearing in Theorem 2.1. In Section 2.3 we show how our results can be applied to processes whose generators consist of sums of generator terms each of which corresponds to different mechanisms of the process and which we can characterize by a martingale problem for which we have a duality. Proofs or arguments for results are found in Section 3. Several examples are treated in Section 4. Finally, in the Outlook-Section 5 we discuss how the restrictions of our results to compact state spaces can be used for non-compact and in particular non locally compact cases by checking additional conditions.

2.1 The principal result

Theorem 2.1 below is our main result for showing existence of solutions of martingale problems. We will say that two processes $X$ and $Y$ (with state spaces $E_X$ and $E_Y$) are in $H$-duality (for some $H : E_X \times E_Y \to \mathbb{R}$) with potential $\beta : E_Y \to \mathbb{R}$ if

$$
\mathbb{E}_x[H(X_t,y)] = \mathbb{E}_y[H(x,Y_t)\exp\left(-\int_0^t \beta(Y_r)dr\right)]. \tag{2.1}
$$

Note that – in contrast to the introduction – we are dealing with the slightly more complex situation because (2.1) involves (in contrast to (1.2)) an extra term on the right-hand-side, often referred to as a Feynman-Kac term), denoted here by $\beta$. In various applications which we present in Section 4, we will have $\beta = 0$; in Example 7 we treat a case for $\beta \neq 0$. The proof of the following result is given in Section 3.

**Theorem 2.1** (A semigroup property and existence by duality). Let $E_X, E_Y$ be Polish, $H : E_X \times E_Y \to \mathbb{R}$ bounded and continuous, and $G_Y : \mathcal{H}_Y \to C_b(E_Y), \beta \in C_b(E_Y)$. Define $\mathcal{H}_X := \{H(\cdot,y) : y \in E_Y\}$ and $\mathcal{H}_Y := \{H(x,\cdot) : x \in E_X\}$.

(i) Suppose that for each $y \in E_Y$ there is an $E_Y$-valued Markov process $Y$ with a strongly continuous semigroup, which is the unique solution of the $(G_Y, \mathcal{H}_Y, y)$-martingale problem. Then, the family $(P_t)_{t \geq 0}$, defined on the closure of span($\mathcal{H}_X$), given by

$$
P_tH(\cdot,y) := \mathbb{E}_y[H(\cdot,Y_t)\exp\left(-\int_0^t \beta(Y_r)dr\right)], \tag{2.2}
$$

is a semigroup. Assume that its generator $G_X$ has domain $\mathcal{D}_X \supseteq \mathcal{H}_X$ and satisfies

$$
G_XH(\cdot,y)(x) = G_YH(x,\cdot)(y) + \beta(y)H(x,y), \quad x \in E_X, y \in E_Y. \tag{2.3}
$$

(ii) In addition, assume that span($\mathcal{H}_X$) is separating on $\mathcal{M}_1(E_X)$ and there exists a family $(\mu_t)_{t \geq 0}$ of probability kernels from $E_X$ to $E_X$ such that, for all $\Gamma \in \mathcal{B}(E_X)$,

$$(t, x) \mapsto \mu_t(x, \Gamma) \text{ is } \mathcal{B}([0, \infty) \times E_X) - \mathcal{B}([0, 1]) \text{ measurable}, \tag{2.4}$$
and for all \( y \in E_Y \) and \( t \geq 0 \) the kernel representability condition

\[
P_t H(\cdot, y) = \int_{E_X} \mu_t(\cdot, dx') H(x', y)
\]

(2.5)

holds. Then, for each \( x \in E_X \), there exists a Markov process \( X = (X_t)_{t \geq 0} \) starting in \( x \) and having transition kernels \( (\mu_t)_{t \geq 0} \), i.e. the right hand side of (2.5) equals \( \mathbb{E}_X[H(X_t, y)] \). In particular, \( X \) and \( Y \) are in duality w.r.t. \( H \) and potential \( \beta \). Moreover, the process \( X \) is the unique solution of the martingale problem for \((G_X, H_X, x)\) and the martingale problem for \((G_X, H_X)\) is well-posed.

(iii) Finally, if \( \text{span}(H_X) \) is convergence determining, then \( X \) is \( C_b(E_X) \)-Feller.

We note that Theorem 2.1 is concerned with martingale problems for \((G_X, H_X)\) and does not make any statements about existence of solutions of the martingale problem for \((G_X, D_X)\) at this point. The uniqueness of the solution is of course immediate. The step from \( H_X \) to \( D_X \) is an application of general theory; see Section 4.3 in [EK86]. Using Proposition 4.3.1 of [EK86] we obtain the following corollary to Theorem 2.1.

Corollary 2.2 (Well-posedness of martingale problems). Assume that the bp-closures of \( (f, G_X f) : f \in H_X \) and of \((f, G_X f) : f \in D_X \) agree,

\[
\text{(2.6)}
\]

and that the assumptions of Theorem 2.1 are satisfied. Then the martingale problem for \((G_X, D_X)\) is well-posed.

In the case of a locally compact state space \( E_X \), recall from Theorem 4.2.7 in [EK86] that \( \hat{C}(E_X) \)-Feller semigroups generate strong Markov processes with càdlàg paths. We give the corresponding result in our case only for compact state spaces, since Theorem 2.1(ii) in general only gives the \( C_b(E_X) \)-Feller property.

Corollary 2.3 (Path regularity). Let \( E_X \) be compact and let the assumptions of Theorem 2.1 be satisfied and assume that \( X \) is the process obtained in Theorem 2.1(a). If \( \text{span}(H_X) \) is convergence determining and \( Y \) is \( C_b(E_Y) \)-Feller, then \( X \) has a modification with càdlàg paths.

Proof. By Theorem 2.1(ii), \( X \) is \( C(E_X) \)-Feller, since \( E_X \) is compact. Then, the result follows from Corollary 4.3.7 in [EK86]. \( \square \)

Remark 2.4 (More general choices of \( H \)). In Theorem 2.1 we have assumed that \( H \) is a bounded and continuous function. By inspection of its proof one can see that the boundedness assumption is only used in a calculation that uses a Fubini argument. Thus, if \( H \) is nonnegative or if the condition

\[
\mathbb{E}_y[H(x_t, Y_t)|\exp\left(\int_0^t \beta(Y_s)ds\right)] < \infty, \quad y \in E_Y, \ t \geq 0,
\]

(2.7)
is fulfilled, then the assertions of Theorem 2.1 remain true for unbounded \( H \). \( \triangle \)
2.2 Checking the conditions of Theorem 2.1

In this subsection we discuss how to check the conditions of Theorem 2.1. First, let us note that in Theorem 2.1 it is not a restriction to assume that $Y$ is Markovian because the Markov property of $Y$ follows from the well-posedness of the martingale problem $(G_Y, \mathcal{H}_Y)$; see e.g. Theorem 4.4.2 in [EK86]. Thus, if we have a process $Y$ and a function $H$ satisfying the generator relation (2.3), it remains to check the assumptions (2.4) and (2.5).

In the following proposition we provide sufficient conditions for the measurability assumption (2.4). The proof can be found in Section 3.

**Proposition 2.5** (Sufficient conditions for (2.4)). Let $H, H_X, Y = (Y_t)_{t \geq 0}$ Markov with a strongly continuous semigroup be as in Theorem 2.1, and let $(\mu_t)_{t \geq 0}$ be as in (2.5). If $\text{span}(H_X)$ is convergence determining, then the following assertions hold:

(i) The mapping $(t, x) \mapsto \mu_t(x, \cdot) \in \mathcal{M}_1(E_X)$ is continuous.

(ii) For all $x \in E_X$, $\Gamma \in \mathcal{B}(E_X)$, the mapping $(t, x) \mapsto \mu_t(x, \Gamma)$ is measurable, i.e. (2.4) holds.

The key condition that remains to be checked is (2.5). For this, recall the following version of the Riesz-Markov theorem: If $E$ is compact and $P : C(E) \to C(E)$ is linear, positive (i.e. $f \geq 0$ implies $Pf \geq 0$) and $P1 = 1$, there is $\mu \in \mathcal{P}(E)$ such that $Pf = \int f d\mu$. We will use this theorem in two ways (always for fixed $x \in E_X$ and $t \geq 0$, and denoting the left hand side of (2.5) by $P_x^t y$).

First (see Proposition 2.6), set $E = E_X$ and assume that $\text{span}(H_X)$ is dense in $C(E_X)$. Then, if the map on $\text{span}(H_X)$, given by $H(\cdot, y) \mapsto P_x^t y$, is positive, we can extend this to $C(E_X)$ and find the corresponding $\mu_t(x, \cdot)$.

Second, (see Proposition 2.8), assume that $E_Y$ is a vector space which is dense in $C(E_U)$ for some compact $E_U$, and assume that $y \mapsto P_x^t y$ is a positive linear form. Then, we find some $U$-valued random variable (due to the Riesz-Markov theorem) with $P_x^t y = \mathbb{E}[y(U)]$. Now, if it is possible to find $X$ with $\mathbb{E}[H(X, y)] = \mathbb{E}[y(U)]$, we can take $\mu_t(x, \cdot)$ as the distribution of $X$.

We note that many measure-valued processes of interest lead to compact (or locally compact) state spaces. However, recall that the historical processes (see [DP91]) for branching models and for Fleming-Viot processes have state spaces which are not locally compact. The same is typically true for genealogy-valued processes; see Remark 5 and [GPW13, DGP12, GSW16]. Strategies how to use Theorem 2.1 in such situations are discussed in Section 5.

Now, we give two conditions which can be used to show (2.5).

**Proposition 2.6** (A way to check condition (2.5) for compact $E_X$). Let $E_X$, $E_Y$, $H$, $Y$ and $P_t$ be as in Theorem 2.1(i). Assume that

(i) $E_X$ is compact;

(ii) $\text{span}(H_X) \subseteq C(E_X)$ is a convergence determining algebra (i.e. it is closed under multiplication) containing $1$;

(iii) the semigroup of $Y$ is $C_b(E_Y)$-Feller;
(iv) for all \( t \geq 0 \) and \( x \in E_X \), the linear map \( P^*_t : \text{span}(H_X) \rightarrow \mathbb{R} \), given by

\[
P^*_t H(\cdot, y) := P_t H(\cdot, y)(x) = \mathbb{E}_x \left[ H(x, Y_t) \exp \left( \int_0^t \beta(Y_s) \, ds \right) \right]
\]

(2.8)

is positive with \( P^*_t 1 = 1 \).

Then, there is a unique continuous extension of \((P_t)_{t \geq 0}\) to \( C(E_X) \), which is again a positive linear form. Moreover, there is a family of probability kernels \((\mu_t)_{t \geq 0}\) from \( E_X \) to \( E_X \) such that (2.5) holds.

The verification of the point (iv) is based on using the properties of \( H(\cdot, y) \) and the form of the states of the dual process \( Y \). In the examples from Section 4 we will e.g. apply moment problems for checking (iv). In general, there are more applications of the proposition than one might think as we will see in the section on examples.

In some cases, verification of (iv) is possible by using approximate dual processes, as in the following corollary to Proposition 2.6. It will be used in Example 5. Again the proof can be found in Section 3.

**Corollary 2.7** (Approximating duals for (iv) of Proposition 2.6). Let \( E_X, E_Y, H, \beta, Y \) and \( P_t \) be as in Theorem 2.1(i) and assume that (i)–(iii) of Proposition 2.6 hold. In addition, let \( X^1, X^2, \ldots \) be Markov processes (with state space \( E_X \)), and \( Y, Y^1, Y^2, \ldots \) be Markov processes (with state space \( E_Y \)) such that \( X^n, Y^n \) are in \( H \)-duality with potential \( \beta \) (see (2.1)), \( n = 1, 2, \ldots \), and \( Y^n \Rightarrow Y \). Then, (iv) of Proposition 2.6 holds. In particular, (2.5) of Theorem 2.1 holds.

In many situations it is necessary to work with function-valued duals, for instance in population genetics, if we deal with measure-valued processes (measures on some type space \( I \) which is often compact) and if the mechanisms include mutation and selection. Then duality functions are functions on the space of the samples on \( \mathcal{U} = I^\mathbb{N} \) for some compact type space \( I \) and hence \( \mathcal{U} \) is compact. We now give a second condition for verifying (2.5), which applies in the situation where \( E_Y \) is a space of continuous functions on some compact set.

**Proposition 2.8** (Another way to check condition (2.5)). Let \( E_X, E_Y, H, \beta, Y \) and \( P_t \) be as in Theorem 2.7(i) with \( E_Y \) being a set of continuous functions to be specified below. Assume that

(i) the semigroup of \( Y \) is \( C_b(E_Y) \)-Feller;

(ii) there exists a compact metric space \( E_U \) so that \( E_Y \subseteq C(E_U) \) is a vector space containing \( I \), which is dense (with respect to the sup-norm) in \( C(E_U) \);

(iii) for all \( t \geq 0 \) and \( x \in E_X \), the linear map \( Q^*_t : \text{span}(E_Y) \rightarrow \mathbb{R} \), given by

\[
Q^*_t y := P_t H(\cdot, y)(x) = \mathbb{E}_x \left[ H(x, Y_t) \exp \left( \int_0^t \beta(Y_s) \, ds \right) \right]
\]

(2.9)

is a positive linear form with \( Q^*_t 1 = 1 \);
(iv) for any $E_U$-valued random variable $U$ (with $E_U$ from (ii)), there is an $E_X$-valued random variable $X$ such that

$$\mathbb{E}[H(X,y)] = \mathbb{E}[y(U)] \quad \text{for all} \ y \in E_Y. \quad (2.10)$$

Then, there is a family of probability kernels $(\mu_t)_{t \geq 0}$ from $E_X$ to $E_X$ such that $(2.5)$ holds.

In Example 3, we will apply Proposition 2.8 via the following corollary which contains easier to check conditions. Recall measure-valued processes on $M(I)$ and that duality processes in this case are function-valued with functions depending on samples from $U = I^U$.

**Corollary 2.9 (How to check (2.10)).** Let $Y^y$ denote the stochastic process, distributed according to $Y$ with initial value $y$. In the situation of Proposition 2.8, we can replace (iv) by one of the following conditions.

(iv') Suppose $\beta = 0$ and there is a subset $F \subseteq \{f : E_U \to E_U \text{ measurable}\}$ with $E_Y \circ F := \{y \circ f : y \in E_Y, f \in F\} \subseteq E_Y$ and furthermore that $Y$ and $F$ are such that for all $x \in E_X$ and $f \in F$ (with $G_Y$ the generator of $Y$) $H(x,y \circ f) = H(x,y)$ and

$$G_Y H(x, \cdot)(y) = G_Y H(x, \cdot)(y \circ f). \quad (2.11)$$

Then, for any $E_U$-value random variable $U$ with $U \sim f(U)$ for all $f \in F$, there is an $E_X$-valued random variable $X$ such that $(2.10)$ holds.

(iv'') If $O \subseteq E_U$ is such that there are $y_1, y_2, \ldots \in E_Y$ such that $y_n \xrightarrow{n \to \infty} O$ boundedly pointwise and $H(x, y_n^t) \xrightarrow{n \to \infty} 0$ in probability for all $t \geq 0$. Then, for any $O$-valued random variable $U$, there is an $E_X$-valued random variable $X$ such that $(2.10)$ holds.

**Remark 2.10 (Using (iv') and (iv'')).** We will use condition (iv') in Example 3. We do not provide an example for using (iv''), but note that this result paves the way to deal with non-compact $E_X$, provided that there is a compactification $E_U$ of $O := E_X$. In this case, we can use $E_U$ as a state space of $X$, but show that it never leaves $E_X$ using a sequence $y_1, y_2, \ldots \in E_Y$ as in Corollary 2.9.

2.3 **Combination of mechanisms**

The above results develop considerable strength due to the possibility to extend the theory further to *sums of operators*, each of which correspond to processes whose existence and uniqueness is already verified. Using Trotter’s product formula [EK84, Corollary 1.6.7] we will show that if $E_X$ is compact and the generator can be written as a sum of operators corresponding to different mechanisms, then it suffices to check the assumptions of Theorem 2.1 for each mechanism separately via Propositions 2.8 and 2.6, provided we have existence of the dual process corresponding to the sum. More general state spaces will be discussed briefly in the Outlook-Section 5.

**Theorem 2.11 (Trotter’s formula, combination of mechanisms).** Let $E_X, E_Y, H, H_X, H_Y$ be as in Theorem 2.7 and assume $\beta = 0$. In addition, let $G_Y^{(1)}, \ldots, G_Y^{(m)}$ satisfy the conditions for $G_Y$ in Theorem 2.7(i), giving rise to Markov processes $Y^{(1)}, \ldots, Y^{(m)}$ and semigroups $P^{(1)}, \ldots, P^{(m)}$. Assume that
(a) (i) and (ii) of Proposition 2.6 hold (in particular, $E_X$ is compact);

(b) (iii) and (iv) of Proposition 2.6 hold for each $Y(1), \ldots, Y(m)$, and the corresponding semi-groups $P(1), \ldots, P(m)$;

(c) the $(G_X^{(1)} + \cdots + G_X^{(m)}, \mathcal{H}_Y)$-martingale problem is well-posed with solution $Y$, $\mathcal{H}_Y$ is a core, and the semigroup of $Y$ is $C_b(E_Y)$-Feller.

Then, there is a family of probability kernels $(\mu_t)_{t \geq 0}$ from $E_X$ to $E_X$, such that

(2.4) and (2.5) hold. In particular, for each $x \in E_X$ there exists a Markov process $X = (X_t)_{t \geq 0}$ starting in $x$ with transition kernels $(\mu_t)_{t \geq 0}$ such that $X$ is Feller and the unique solution of the $(G_X^{(1)} + \cdots + G_X^{(m)}, \mathcal{H}_X)$-martingale problem.

If in Theorem 2.11 for some reasons we already know that some of the mechanisms involved have actually unique solutions of the corresponding martingale problems then we can avoid checking the conditions for that mechanism. For an application of the strategy described in the following remark see Example 3.

**Remark 2.12** (Well-posed mechanisms). For Theorem 2.11(b), assume that $Y(1), \ldots, Y(m)$ are Feller. If for some $i \in \{1, \ldots, m\}$ the $G_X^{(i)}$-martingale problem is well-posed and in $H$-duality with $Y(i)$, then the semigroup of $X(i)$ satisfies (iv) of Proposition 2.6 by construction.

**Remark 2.13** (Deterministic solutions of martingale problems with first order operators). The corollary is applied in Example 3 in the case that $G_X^{(1)}$ has a simple structure. Recall the conditions and notation of Theorem 2.1 and assume that the operator $G_X$ is a first order operator, i.e. $\text{span}(\mathcal{H}_X)$ is closed under multiplication and

$$G_X\Phi^2 - 2\Phi G_X\Phi = 0 \text{ for all functions } \Phi \in \text{span}(\mathcal{H}_X). \quad (2.12)$$

Then duality guarantees existence and uniqueness of the corresponding martingale problem. Furthermore the solutions are deterministic. In particular the duality relation of the processes reads as follows:

$$\mathbb{E}_x[H(X_t, y)] = H(X_t, y) = \mathbb{E}_x[H(x, Y_t) \exp\left(\int_0^t \beta(Y_s) \, ds\right)]. \quad (2.13)$$

Since $\text{span}(\mathcal{H}_X)$ is separating, the transition kernels $(\mu_t)_{t \geq 0}$ in (2.5) must satisfy

$$\mu_t(x, \cdot) = \delta_{F_t(x)}(\cdot), \quad (2.14)$$

where $F_t(x)$ is the solution of the initial value problem corresponding to $G_X$, satisfying

$$\frac{d}{dt}H(F_t(x), y) = d(F_t(x), y) \quad \text{and} \quad F_0(x) = x. \quad (2.15)$$

Note that the measurability of $(t, x) \mapsto F_t(x)$ is guaranteed by (2.13) because $Y$ is a solution of a martingale problem.
3 Proofs

3.1 Proof of Theorem 2.1

Proof of Theorem 2.1(i). For the semigroup property of \((P_t)_{t \geq 0}\), observe that by construction (i.e. linearity), and dominated convergence, using Fubini,

\[
P_t P_s H(\cdot, y) = \mathbb{E}_y \left[ P_t H(\cdot, Y_s) \exp \left( - \int_0^s \beta(Y_r) dr \right) \right] = \mathbb{E}_y \left[ \mathbb{E}_{Y_s} \left[ H(\cdot, Y_t) \exp \left( - \int_0^t \beta(Y_r) dr \right) \right] \exp \left( - \int_s^t \beta(Y_r) dr \right) \right] = \mathbb{E}_y \left[ H(\cdot, Y_{t+s}) \exp \left( - \int_0^{t+s} \beta(Y_r) dr \right) \right] = P_{t+s} H(\cdot, y).
\]

(3.1)

For its generator \(G_X\) and each \(y \in E_Y\) we have

\[
G_X H(\cdot, y)(x) = \lim_{h \to 0} \frac{1}{h} \mathbb{E}_y \left[ H(x, Y_h) \exp \left( \int_0^h \beta(Y_s) ds \right) - H(x, y) \right] = \lim_{h \to 0} \frac{1}{h} \mathbb{E}_y \left[ H(x, Y_h) \left( \exp \left( \int_0^h \beta(Y_s) ds \right) - 1 \right) + H(x, Y_h) - H(x, y) \right] = \beta(y) H(x, y) + G_Y H(x, \cdot)(y),
\]

where we have used the strong continuity of the semigroup of \(Y\), i.e. \(Y_h \xrightarrow{h \to 0} y\). This shows that (2.3) holds. □

Proof of Theorem 2.1(ii). (ii) By Theorem 4.1.1 in [EK86], there exists a Markov process \(X\) with transition functions \((\mu_t)_{t \geq 0}\), provided that \((\mu_t)_{t \geq 0}\) is a family of probability distributions satisfying (2.4), \(\mu_0(x, \cdot) = \delta_x(\cdot)\) and

\[
\mu_{t+s}(x, \cdot) = \int_{E_X} \mu_t(x, dx') \mu_s(x', \cdot), \quad s, t \geq 0, \quad x \in E_X.
\]

(3.3)

First, by (2.5), there exists a transition kernel \(\mu_0\) such that for all \(y \in E_Y\)

\[
H(x, y) = \int_{E_X} \mu_0(x, dx') H(x', y).
\]

(3.4)

Since \(\text{span}(\mathcal{H}_X)\) is separating on \(M_1(E_X)\), this implies \(\mu_0(x, dx') = \delta_x(dx')\).
In order to show (3.3), observe that, by (2.5), the semigroup-property of (i), and Fubini,
\[
\int \mu_t(x, dx')H(x', y) = P_tH(x)(x) = P_t(P_tH(x))(x)
\]
\[
= P_tE_y[H(., Y_t) \exp \left(- \int_0^t \beta(Y_r)dr \right)](x)
\]
\[
= E_y[(P_tH(., Y_t)(x)) \exp \left(- \int_0^t \beta(Y_r)dr \right)]
\]
\[
= \int \mu_t(x, dx')E_y[H(x', Y_t) \cdot \exp \left(- \int_0^t \beta(Y_r)dr \right)]
\]
\[
= \int \mu_t(x, dx') \int \mu_t(x', dx'')H(x'', y).
\]
(3.5)

Since span(\(\mathcal{H}_X\)) is separating, we have shown (3.3) and we have constructed a Markov process \(X\) with \(X_0 = x\) and
\[
E_x[H(X_t, y)] = \int \mu_t(x, dx')H(x', y) = E_y[H(x, Y_t) \exp \left(\int_0^t \beta(Y_s)ds \right)].
\]
(3.6)

We now show that \(X\) is the unique solution of the martingale problem for \((G_X, \mathcal{H}_X, x)\). Uniqueness follows directly from Proposition 4.4.7 in [EK86]. If we can show that \(X\) has a progressively measurable modification, then by Proposition 4.1.7 in [EK86] \(X\) is a solution of the martingale problem for \((\mathcal{H}_X, G_X, x)\), i.e. existence follows. For existence of a progressively measurable modification, it remains to show \(\mu_t(x, \cdot) \overset{h \to 0}{\longrightarrow} \delta_x(\cdot)\) for all \(x \in E_X\); see Theorem II.2.6 in [Doo53], together with [KS91, Proposition 1.12]. This follows via duality from the Assumption (2.5) on the dual process. □

Proof of Theorem 2.1(iii). For the \(C_b(E_X)\)-Feller property, using strong continuity of the semigroup of \(Y\) and (3.6), we obtain \(E_x[H(X_t, y)] \overset{t \to 0}{\longrightarrow} H(x, y)\) for all \(y \in E_Y\). If span(\(\mathcal{H}_X\)) is convergence determining, this implies \(X_t \overset{t \to 0}{\longrightarrow} x\) and therefore \(E_x[f(X_t)] \overset{t \to 0}{\longrightarrow} f(x)\) for all \(f \in C_b(E_X)\). This shows that the semigroup of \(X\) is strongly continuous. In order to show continuity of \(x \mapsto E_x[f(X_t)]\) for \(f \in C_b(E_X)\) and \(t \geq 0\), let \(x, x_1, x_2, \ldots \in E_x\) such that \(x_n \overset{n \to \infty}{\longrightarrow} x\) and write \(X_t^x\) for a random variable distributed according to \((X_t)_t, \mathbb{P}_x\). Using dominated convergence and continuity of \(H\) we obtain
\[
E[H(X_t^{x_n}, y)] = E_y[H(x_n, Y_t) \exp \left(\int_0^t \beta(Y_s)ds \right)]
\]
\[
\overset{n \to \infty}{\longrightarrow} E_y[H(x, Y_t) \exp \left(\int_0^t \beta(Y_s)ds \right)] = E[H(X_t^x, y)],
\]
(3.7)
for all \(y\). This shows that \(X_t^{x_n} \overset{n \to \infty}{\longrightarrow} X_t^x\), since span(\(\mathcal{H}_X\)) is convergence determining. Therefore, for \(f \in C_b(E_X)\), \(x \mapsto E[f(X_t^x)]\) is continuous and bounded, i.e. \(X\) is \(C_b(E_X)\)-Feller. □
3 PROOFS

3.2 Proof of Propositions 2.5, 2.6, 2.8 and Corollary 2.9

Proof of Proposition 2.5 We first show (i) and then we prove that (ii) is a consequence of (i). Since \( \text{span} (\mathcal{H}_X) \) is convergence determining, we only have to show that for all \( y \in E_Y \) the mapping

\[
(t, x) \mapsto \int_{E_X} \mu_t(x, dx') H(x', y) = \mathbb{E}_x \left[ H(x, Y_t) \exp \left( \int_0^t \beta(Y_s) ds \right) \right]
\]

is continuous. Continuity in \( x \) follows from boundedness of \( H \) and dominated convergence. Continuity in \( t \) follows from the Markov property and the strong continuity of the semigroup of \( Y \).

To see that (i) implies (ii) let \( f_1, f_2, \ldots \in C_0(E_X) \) be such that \( f_n \xrightarrow{n \to \infty} 1 \) boundedly pointwise, then \( (t, x) \mapsto \mu_t(x, \Gamma) = \lim_{n \to \infty} \int \mu_t(x, dx') f_n(x') \) is measurable as a limit of continuous functions.

Proof of Proposition 2.6 First, \( \text{span}(\mathcal{H}_X) \) is dense in \( C(E_X) \) due to (i), (ii) and the Stone-Weierstrass theorem. Second, \( P^t_\Gamma \) is continuous (since the semigroup of \( Y \) is \( C_0(E_Y) \)-Feller by (iii)) on \( \text{span}(\mathcal{H}_X) \), so there is a unique extension of \( P^t_\Gamma \) to \( C(E_X) \). By continuity, this extension also satisfies (iv), i.e. it is a positive linear form. By the Riesz-Markov theorem, there is a unique measure \( \mu_t(x, \cdot) \) such that

\[
P^t_\Gamma f = \int_{E_X} \mu_t(x, dx') f(x'), \quad f \in C(E_X).
\]

Since \( P^t_\Gamma 1 = 1 \) by (iv), we know that \( \mu_t(x, \cdot) \in M_1(E_X) \). Applying (3.9) to \( \mathcal{H}_X \) and using (2.8) we obtain

\[
\mathbb{E}_x \left[ H(x, Y_t) \exp \left( \int_0^t \beta(Y_s) ds \right) \right] = \int_{E_X} \mu_t(x, dx') H(x', y),
\]

which is precisely (2.5). Since \( x \mapsto \mu_t(x, \cdot) \) is continuous by the last display, we conclude that \( \mu_t \) is a probability kernel from \( E_X \) to \( E_X \).

Proof of Corollary 2.7 Let \( P^x \) be as in Theorem 2.1(i). By the approximate duality, we can write for (2.8) using the convergence \( Y_n \xrightarrow{\text{bp}} Y \)

\[
P^t_\Gamma H(., y) = \lim_{n \to \infty} \mathbb{E}_x \left[ H(x, Y^n_t) \exp \left( \int_0^t \beta(Y^n_s) ds \right) \right] = \lim_{n \to \infty} \mathbb{E}_x [H(X^n_t, y)].
\]

This calculation shows that \( P^t_\Gamma \) from Proposition 2.6 is positive (as a limit of positive maps) and \( P^t_\Gamma 1 = 1 \) (as a limit of maps with the same property). Hence, Proposition 2.6(iv) holds.

Proof of Proposition 2.8 Let \( t \geq 0 \) and \( x \in E_X \). First, since \( E_Y \) is dense in \( C(E_U) \) and the semigroup of \( Y \) is \( C_0(E_Y) \)-Feller, we can extend \( Q^t_\Gamma \) to \( C(E_U) \). This then gives a positive linear form on \( C(E_U) \). By the Riesz-Markov theorem (recall that \( E_U \) is compact), we find a probability
measure $\nu_t(x, \cdot) \in \mathcal{M}(E_U)$ such that $Q_t^* f = \int \nu_t(x, du)f(u)$ for all $f \in C(E_U)$. For $y \in E_Y$, this amounts to

$$\int \nu_t(x, du)y(u) = \mathbb{E}_y\left[H(x, Y_t) \exp\left(\int_0^t \beta(Y_s) \, ds\right)\right].$$

(3.12)

Now, let $U$ have the distribution $\nu_t(x, \cdot)$. By assumption, we find $X$ so that $\mathbb{E}[H(x, y)] = \mathbb{E}[y(U)]$ for all $y \in E_Y$. Denoting the distribution of $X$ by $\mu_t(x, \cdot)$, we obtain

$$\int \mu_t(x, dx')H(x', y) = \int \nu_t(x, du)y(u)$$

and (2.5) follows.

Proof of Corollary 2.9. From the proof of Proposition 2.8, we see that we only need to find $X$ such that (2.10) holds for $U \sim \nu_t(x, \cdot)$, where $\nu_t$ is from (3.12).

Under the assumptions of (iv’), we claim that for all $f \in \mathcal{F}$ and $y \in E_Y$

$$\mathbb{E}[H(x, Y^{\nu_t}_t)] = \mathbb{E}[H(x, Y^y_t)].$$

(3.13)

Indeed, by assumption, using the semigroup $(S_t)_{t \geq 0}$ of $Y$,

$$\mathbb{E}[H(x, Y^{\nu_t}_t)] = H(x, y \circ f) + \int_0^t S_s G_Y H(x, y \circ f) \, ds = \mathbb{E}[H(x, Y^y_t)].$$

(3.14)

Then, we have for such all $U$ with $U \sim \nu_t(x, \cdot)$ and for all $y \in E_Y$,

$$\mathbb{E}[y(f(U))] = \int \nu_t(x, du)(y \circ f)(u) = \mathbb{E}[H(x, Y^{\nu_t}_t)] = \mathbb{E}[H(x, Y^y_t)] = \mathbb{E}[y(U)].$$

(3.15)

Hence, $U \sim f(U)$ for all $f \in \mathcal{F}$, so we require to show (2.10) only for such $U$.

For (iv”), we write for $U \sim \nu_t(x, \cdot)$

$$\mathbb{P}(U \in O^c) = \lim_{n \to \infty} \mathbb{E}[\nu_n(U)] = \lim_{n \to \infty} \int \nu_t(x, du)\nu_n(u)$$

$$= \lim_{n \to \infty} \mathbb{E}[H(x, Y^{\nu_n}_t)] \exp\left(\int_0^t \beta(Y^{\nu_n}_s) \, ds\right) = 0.$$  

(3.16)

Hence, $U$ has values in $O$, almost surely, and we need to show (2.10) only for such $O$.  

3.3 Proof of Theorem 2.11

To facilitate reading, we restrict ourselves to the case $m = 2$. We will use Trotter’s product formula for the semigroups $(Q_t^{(1)})_{t \geq 0}$, $(Q_t^{(2)})_{t \geq 0}$ and $(Q_t)_{t \geq 0}$, given by

$$Q_t^{(i)} f(y) = \mathbb{E}_y[f(Y^{(i)}_t)], \quad i = 1, 2, \quad Q_t f(y) = \mathbb{E}_y[f(Y_t)].$$
These are strongly continuous contraction semigroups on $C_0(\mathbb{E}_X)$ with generators $G_Y^{(1)}, G_Y^{(2)}$ and $G_Y^{(1)} + G_Y^{(2)}$, respectively. In addition, $\mathcal{H}_X$ is a core for $G_Y^{(1)} + G_Y^{(2)}$ by assumption. From Trotter’s formula, we see that

$$Q_t f = \lim_{n \to \infty} (n^n Q_t f), \text{ with } (n^n)Q_t = \left( Q_{t/2n}^{(2)} Q_{t/2n}^{(1)} \cdots Q_{t/2n}^{(2)} Q_{t/2n}^{(1)} \right) \text{ and } 2n \text{ factors.} \hspace{1cm} (3.17)$$

From (a) and (b) and Propositions 2.5 and 2.6 we know that (2.4) and (2.5) hold for $i = 1, 2$. So, as Theorem 2.1 shows, there are $E_X$-valued processes $X^{(1)}$ and $X^{(2)}$ with semigroups $P^{(1)}$ and $P^{(2)}$, respectively, given by

$$P_i^{(1)} H(\cdot, y)(x) = \mathbb{E}_x [H(X_t^i, y) | \mathbb{E}_x] = Q_i^{(1)} H(x, \cdot)(y), \hspace{1cm} i = 1, 2.$$

As a next step, we define

$$P_i H(\cdot, y)(x) := Q_i H(x, \cdot)(y) = \lim_{n \to \infty} \left( P_i^{(2)} P_i^{(1)} P_i^{(2)} \cdots P_i^{(2)} P_i^{(1)} P_i^{(2)} \right) H(x, \cdot)(y) = \lim_{n \to \infty} \left( P_i^{(2)} \cdots P_i^{(2)} P_i^{(1)} P_i^{(2)} \cdots P_i^{(2)} P_i^{(1)} \right) H(\cdot, y)(x).$$

By Proposition 2.6, $P_i^{(1)}$ and $P_i^{(2)}$ can be extended to $C(\mathbb{E}_X)$ and by (b) are positive with $P_i^{(1)}1 = P_i^{(2)}1 = 1$. So, we see that $P_i$ can be continuously extended on $C(\mathbb{E}_X)$ with $P1 = 1$ and by the Riesz-Markov Theorem, for every $x \in \mathbb{E}_X$ and $t \geq 0$, there is a Markov kernel $\mu_i(x, \cdot)$ from $\mathbb{E}_X$ to $E_X$ such that (3.5) holds. In addition, (2.4) holds since $(t, x) \mapsto \mu_i(x, \cdot)$ is measurable as a limit of continuous functions; see Proposition 2.5. Hence, all conclusions of Theorem 2.1(ii) follow.

\[ \square \]

4 Examples

In this section we give several examples how the above results can be applied. We will distinguish between the compact and locally compact case. Example 1 is the Fleming-Viot measure-valued diffusion (without mutation and selection), which is a process taking values in $M_1([0, 1])$. As an extension, we consider the Cannings model in Example 2 with the same state space, but càdlàg paths with jumps. In Example 3, we add a spatial component, which gives an application of Theorem 2.1. In Example 4 (Fleming-Viot process with mutation) and Example 5 (spatial A-Fleming-Viot process), we use function-valued duals. Turning to the case of locally compact state spaces, we treat in Example 6 the continuous state branching process and in Example 7 as a special case the Feller branching process using a different duality, referred to as Feynman-Kac duality, and $\beta \neq 0$.

4.1 Compact state spaces – resampling systems

Population models with a constant population size do not only arise frequently in population genetics, but are also frequently analysed using dual processes. Their large-population-limits come as solutions of stochastic differential equations, such as the Wright-Fisher diffusion, measure-valued diffusions, or more complex approaches, such as historical or tree-valued processes;
and we say that $u \in \mathbb{R}$ for all $x \in M([0,1])$ (the set of finite Borel-measures on $[0,1]$) and $u \in [0,1]$

$$\frac{\partial F(x)}{\partial x}[u] = \lim_{\varepsilon \downarrow 0} \frac{F(x + \varepsilon \delta_u) - F(x)}{\varepsilon}$$

exists and $(u, x) \mapsto \frac{\partial F(x)}{\partial x}[u]$ is continuous. \hfill (4.1)

In the obvious way, if it exists, the second derivative is defined as

$$\frac{\partial^2 F(x)}{\partial x \partial x}[u, v] := \frac{\partial}{\partial x} \left( \frac{\partial F(x)}{\partial x}[u] \right)[v]$$

and we say that $F$ is twice continuously differentiable if $(u, v, x) \mapsto \frac{\partial^2 F(x)}{\partial u \partial x}[u, v]$ is continuous. We set

$$\mathcal{D}_X = \{ F : M_1([0,1]) \rightarrow \mathbb{R} : F \text{ is twice continuously differentiable} \}.$$ \hfill (4.3)

### 4.1.1 Resampling systems with compact state spaces and particle-valued duals

**Example 1** (Fleming-Viot process). For $F \in \mathcal{D}_X$, we define

$$G_X F(x) := \int_{[0,1]} \int_{[0,1]} \frac{\partial^2 F(x)}{\partial x \partial x}[u, v](x(du)\delta_u(dv) - x(du)x(dv)).$$ \hfill (4.4)

For $u = (u_1, \ldots, u_n) \in [0,1]^n$, $k = (k_1, \ldots, k_n) \in \mathbb{N}_0^n$ we write $u^k := u_1^{k_1} \cdots u_n^{k_n}$. Then for $F(x) := \langle x^{\otimes n}, u^k \rangle := \int x^{\otimes n}(du) u^k$, a straight-forward calculation shows that

$$G_X F(x) := \sum_{1 \leq i < j \leq n} \langle x^{\otimes n}, u^{\theta_{ij} k} - u^k \rangle,$$ \hfill (4.5)

where $\theta_{ij} k \in \mathbb{N}_0^{n-1}$ arises from $k = (k_1, \ldots, k_n)$ by replacing $k_{i,j}$ by $k_i + k_j$, and shifting the indices above $i \lor j$ down by one; see [Daw93, p. 31].

The dual process $Y$ is the partition-valued Kingman coalescent, i.e. a pure jump process in which every pair of partition elements coalesces at rate 1. More formally, we take

$$E_Y := \bigcup_{n=0}^{\infty} \mathcal{P}_n,$$ \hfill (4.6)

where $\mathcal{P}_n$ is the set of partitions of $\{1, \ldots, n\}$. For a partition $y$ we write $|y|$ for the number of its partition elements of $y$. A partition $y \in \mathcal{P}_n$ induces an equivalence relation “$\sim$” on $\{1, \ldots, n\}$
with \( i \sim j \) if \( i \) and \( j \) are in the same partition element of \( y \). We order the partition elements of \( y \in \mathcal{P}_n \) according to their smallest elements, so that we have a well-defined representation \( y = \{y_1, \ldots, y_{|y|}\} \). We also write \([y_k]\) for the number of elements of the \( k \)th partition element and we write \( y(i) = k \) if \( i \in y_k \), i.e. \( y(i) \) is the number of the partition element \( i \) is in. To define the duality function for \( y \in \mathcal{P}_n \) and \( u = (u_1, \ldots, u_{|y|}) \in [0,1]^{|y|} \), we set

\[
u^y := \prod_{k=1}^{|y|} u_k^{[y_k]} = \prod_{i=1}^n u_{y(i)}, \quad H(x,y) = \int x^{\otimes y}(du) \nu^y. \tag{4.7}\]

For the dynamics of the dual, for \( y \in \mathcal{P}_n \) with \(|y| = m \leq n \) and \( y' \in \mathcal{P}_m \), we write \( y' \circ y \in \mathcal{P}_n \) for the partition with \( i \sim j \) if \( y'(y(i)) = y'(y(j)) \), \( i, j = 1, \ldots, n \). In other words \( y' \circ y \) arises from \( y \) by merging partition elements of \( y \) according to partition \( y' \). For example, if \( y = \{\{1\}, \{2,3\}, \{4\}\} \) and \( y' = \{\{1,3\}, \{2\}\} \), then \( y' \circ y = \{\{1,4\}, \{2,3\}\} \). For \( 1 \leq i < j \leq m \) we define the partition \( \theta_{ij} \in \mathcal{P}_m \) by (note that \( i \in \{1, j-1\} \) and \( j = m \) is possible below)

\[
\theta_{ij} := \theta_{ij}^{(m)} := \{(i-1), \{i, j\}, \{i+1\}, \ldots, \{j-1\}, \{j+1\}, \ldots, \{m\}) \in \mathcal{P}_{m-1}. \tag{4.8}\]

We will typically omit the dependence of \( \theta_{ij} \) on \( m \) but it should be clear from the context what \( m \) is. For instance, the operation \( \theta_{ij} \circ y \) means that \( m = |y| \) and that \( i \)th and \( j \)th partition elements of \( y \) are merged.

With this notation the process \( Y \) can be defined as a process with transitions

\[
Y = y \to \theta_{ij} \circ y \text{ at rate 1 for all } 1 \leq i < j \leq |y|. \tag{4.9}\]

Thus, \( Y \) solves the martingale problem with the operator \( G_Y \) acting on the duality function as follows

\[
G_Y H(x,\cdot)(y) = \sum_{1 \leq i < j \leq |y|} \langle x^{\otimes (y|-i)}, u_{y(i)+y'} \rangle - \langle x^{\otimes j}, u' \rangle = \sum_{1 \leq i < j \leq |y|} \langle x^{\otimes j}, u_{y(i)+y'} - u' \rangle, \tag{4.10}\]

hence \( G_Y H(y,\cdot)(x) = G_Y H(x,\cdot)(y) \) by \( \text{(4.5)} \). We note that \( Y \) is a Feller process and writing \( \mathbb{P}_m \) for the distribution with initial condition \( \{\{1\}, \ldots, \{m\}\} \),

\[
\mathbb{E}_y[f(Y_t)] = \mathbb{E}_y[f(Y_t \circ y)] \tag{4.11}\]

by the definition of \( Y \) (since the dynamics is on and not within the partition elements). For \( y \in \mathcal{P}_n \), we have

\[
\mathcal{P}_r H(\cdot,\cdot)(x) := \mathbb{E}_y[H(x,Y_t)] = \sum_y \mathbb{P}_y[Y_t = y'] \int_{\{0,1\}^{|y'|}} X_t^{\otimes (y'|)}(du_1, \ldots, du_{|y'|}) u_1^{[y_1']} \cdots u_{|y'|}^{[y_{|y'|}]}, \]

where the sum is over all \( y' \in \mathcal{P}_m, m \in \{1, \ldots, n-1\} \) which are possible outcomes (otherwise the probability is 0) of the process \( Y \) with initial condition \( y \).

To check the condition \( \text{(2.5)} \) of Theorem \( \text{2.1} \) we need to find a \( \mathcal{M}_1([0,1]) \)-valued random variable \( X_t \) such that for all \( y \in E_Y \),

\[
\mathbb{E}_x[X_t^{\otimes y}, u'] = \mathbb{E}_x[\int_{\{0,1\}^{|y'|}} X_t^{\otimes y}(du_1, \ldots, du_{|y'|}) u_1^{[y_1']} \cdots u_{|y'|}^{[y_{|y'|}]}, u'] = \mathbb{E}_y[(x^{\otimes y}), u'] = \mathbb{E}_y[(x^{\otimes y}), u^{y'}] =: m_y. \tag{4.12}\]
In order to find $X_t$, we first fix $m \in \mathbb{N}$. We need to find $[0,1]$-valued random variables $U_1, \ldots, U_m$ such that for $U = (U_1, \ldots, U_m)$ all $y \in E_Y$ with $|y| = m$

$$\mathbb{E}[U^y] = m_y. \quad (4.13)$$

By the multi-dimensional Hausdorff moment problem [BCR84, Proposition 6.11, p. 134], this is guaranteed given that, for all $k, \ell \in \mathbb{N}_0^m$ (and sum over all $p = (p_1, \ldots, p_m) \in \mathbb{N}_0^m$ with $p \leq \ell$ componentwise below) we have

$$\sum_{p_1=0}^{\ell_1} \cdots \sum_{p_m=0}^{\ell_m} (-1)^{p_1+\cdots+p_m} \binom{\ell_1}{p_1} \cdots \binom{\ell_m}{p_m} m_{y^{(k,p)}} \geq 0. \quad (4.14)$$

Here, for $k = (k_1, \ldots, k_m) \in \mathbb{N}_0^m$, $y = y^{(k)}$ is a partition with $|y| = m$ and $|y_1| = k_1, \ldots, |y_m| = k_m$.

Note that for $p, k \in \mathbb{N}_0^m$ and $y \in \mathcal{P}_m$ we have $u^{y^{(k,p)}} = u^{y^{(k)}} u^{y^{(p)}}$.

In order to show (4.14), we write for the left hand side

$$\sum_{p_1=0}^{\ell_1} \cdots \sum_{p_m=0}^{\ell_m} (-1)^{p_1+\cdots+p_m} \binom{\ell_1}{p_1} \cdots \binom{\ell_m}{p_m} \mathbb{E}_m[\langle x^{(y)}_t, u^{y^{(k,p)}} \rangle]$$

$$= \mathbb{E}_m[\langle x^{(y)}_t, u^{y^{(k)}} \rangle \sum_{p_1=0}^{\ell_1} \cdots \sum_{p_m=0}^{\ell_m} (-1)^{p_1} \binom{\ell_1}{p_1} \cdots \binom{\ell_m}{p_m} u^{y^{(p)}}] \quad (4.15)$$

$$= \mathbb{E}_m[\langle x^{(y)}_t, u^{y^{(k)}}(1-u)^{y^{(p)}} \rangle] \geq 0.$$

Hence we have shown the existence of $U_1, \ldots, U_m$ with (4.13). By a projective limit argument we can extend this construction to show existence of $U_1, U_2, \ldots$ such that (4.13) holds for any finite subset. The resulting sequence $U_1, U_2, \ldots$ is exchangeable. Hence, by de Finetti’s theorem there is a $\mathcal{M}_1([0,1])$-valued random variable $X_t$ such that $U_1, U_2, \ldots$ is independent given $X_t$. In particular we have

$$\mathbb{E}[U^y] = \mathbb{E}[(X^{(y)}_t, u^y)]. \quad (4.16)$$

But this is exactly (2.5). Moreover, (2.4) holds by Proposition 2.5. Hence, well-posedness of the $(G_X, \mathcal{H}_X)$-martingale problem follows. Since $Y$ is Feller, and $\mathcal{H}_X$ is convergence determining (since $[0,1]$ is compact), $X$ is Feller as well by Theorem 2.1(ii). By Proposition 2.3 there is a modification with almost surely càdlàg paths. Moreover, since $G_X$ is second order, the solution has a version with almost surely continuous paths; see Proposition 4.5 in [DGP12]. This process is usually referred to as the Fleming-Viot measure-valued process.

**Example 2** (Cannings model). For the Cannings model (without Fleming-Viot resampling), let $\Lambda$ be a finite measure on $[0,1]$ (implying no mass at 0) and $\Lambda^*(dr) = \frac{\Lambda(dr)}{r}$, $r \in (0,1]$. Here,

$$G_XF(x) = \int_{[0,1]} \Lambda^*(dr) \int_{\mathcal{E}_X} x(du) (F((1-r)x + rd_u) - F(x)). \quad (4.17)$$
We note that for $F(x) = \langle x^\otimes n, u^k \rangle$, we have

$$G_X F(x) = \sum_{j=2}^{n} \binom{n}{j} \lambda_{nj} \langle x^\otimes (n-j+1), u^\otimes j - \langle x^\otimes n, u^k \rangle \rangle$$

(4.18)

with $\theta_j k = (k_1, \ldots, k_{n-j}, k_{n-j+1} + \cdots + k_n)$ (with a slight abuse of notation for $\theta$ from Example 1) and

$$\lambda_{nj} = \int_0^1 r^{j}(1 - r)^{n-j} \Lambda^*(dr), \quad j \in \{2, \ldots, n\}.$$  (4.19)

For the dual process $Y$, we use the notation of Example 1. Again, the state space of $Y$ is $E_Y$ from (4.6), and the duality function $H$ is as in (4.7). Here, $Y$ is the partition valued $\Lambda$-coalescent, i.e. a pure jump process with the following dynamics: If the current state of the process consists of $n \geq 2$ blocks then each $j$-tuple merges into a single block at rate $\lambda_{nj}$. Let $y \in E_Y$ with $|y| = n$ and let $J \subset \{1, \ldots, n\}$. Generalizing the notation from (4.8) and (4.18), we write $\theta_{J}$ for the partition of $\{1, \ldots, n\}$ in which all elements of $J$ are put in one block and the other partition elements are singletons. Note that $\theta_{J}$ from (4.8) equals $\theta_{J}$ with $J = \{i, j\}$ and $\theta_{\emptyset}$ from (4.18) equals $\theta_{\emptyset}$. With this notation the process $Y$ can be defined as a process with transitions

$$Y = Y \eta \circ \theta_{J} \circ y$$

at rate $\lambda_{\emptyset, |J|}$ for $J \subset \{1, \ldots, |y|\}$.  (4.20)

In particular, $Y$ solves the martingale problem for

$$G_Y H(x, \cdot)(y) = \sum_{J \subset \{1, \ldots, |y|\}} \lambda_{|J|, |J|} \langle x^\otimes (|J|-1), u_{\theta_{\emptyset, |J|}}^\otimes |J| - \langle x^\otimes |y|, u^\otimes |y| \rangle \rangle = \sum_{J \subset \{1, \ldots, |y|\}} \lambda_{|J|, |J|} \langle x^\otimes |y|, u_{\theta_{\emptyset, |J|}}^\otimes - u^\otimes |y| \rangle.$$  (4.21)

Now, we can argue as in Example 1 that (4.11) also holds and that the proof that the multidimensional Hausdorff moment problem has a solution (see (4.15)) literally carries through. As a result it follows that $(G_X, H_X)$-martingale problem is well-posed. Again, $X$ is Feller since $Y$ is Feller and $\Lambda^*$ is convergence determining, and $X$ has càdlàg paths by Proposition 2.3. Here, $G_X$ is not second order. In particular, the paths are not continuous but are jump processes which even have countably many jumps, if $\Lambda^*(\{0, \delta\}) > 0$ for all $\delta > 0$.

4.1.2 Resampling systems with compact state spaces and function-valued duals

**Example 3** (Fleming-Viot process with mutation). Now, we add mutation to the Fleming-Viot process as introduced in Example 1 which requires a different dual process. More precisely, the mutation operator reads for $F \in D_X$ as in (4.3)

$$C^\text{mut}_X F(x) = \theta \int_0^1 \frac{\partial F}{\partial x} (u, d\beta(u, dv) - x(du)),  \quad (4.22)$$

where we refer to $\theta \geq 0$ as the *mutation rate*, and $\beta(\cdot, \cdot)$ is a stochastic kernel on $I$, denoting the mutation kernel.
The operator of the process we aim to show existence for is given by

\[ G_X = G_X^{\text{res}} + G_X^{\text{mut}}, \]

where \( G_X^{\text{res}} \) is given by the right hand side of (4.4) and \( G_X^{\text{mut}} \) is as in (4.22). We will use here, different from the previous examples, a function-valued dual process, i.e. a process with state space

\[ E_Y := \bigcup_{n=0}^{\infty} \Pi_n, \quad \Pi_n := C(I^n), \]

where \( \Pi_0 \) consist of all constants. Moreover, we view \( y \in \Pi_n \) as a continuous function with domain \( I^n \), depending only on \( n \) coordinates, i.e. \( E_Y \subseteq C(E_U) \) if we choose now \( E_U = I^n \). For \( y \in E_Y \), we write \(|y| = n \) if \( y \in C(I^n) \).

For the duality function, we set

\[ H(x,y) := \langle x^{\otimes N}, y \rangle := \langle x^{\otimes |y|}, y \rangle := \int x(du_1) \cdots x(du_{|y|}) y(u_1, \ldots, u_{|y|}). \]

So, in words, \( H(x,y) \) is computed by choosing elements \( u_1, \ldots, u_{|y|} \) independently from \( x \), and evaluating them according to the function \( y \). Setting

\[ \theta_{ij}(u_1, \ldots, u_n) = (u_1, \ldots, u_i, u_{i+1}, \ldots, u_{j-1}, u_j, u_{j+1}, \ldots, u_n) \]

(and note that \( i \in \{1, j-1\} \) and \( j = n \) is possible and again we abuse notation for \( \theta \) from Examples 1 and 2). Note that \( G_X \) takes the special form, when applied to \( H(.,y) \in \mathcal{H}_X \).

Next, we claim that \( \mathcal{H}_X \) is a convergence determining set of functions, and \( 1 \in \mathcal{H}_X \). For this, recall that by Le Cam’s theorem \([\text{LeC57}]\) (see also [LR16]), the set of functions \( \mathcal{H}_X \subseteq C_{b}(E_X) \) on a completely regular Hausdorff space \( E_X \) is convergence determining for Radon probability measures, if it is multiplicatively closed and induces the topology of \( E_X \). In our case, \( \mathcal{H}_X := \{ x \mapsto \langle x^{\otimes |y|}, y \rangle : y \in E_Y \} \subseteq C_{b}(E_X) \) is multiplicatively closed and for \( x, x_1, x_2, \ldots \in E_X \)

\[ (x_n \xrightarrow{n \to \infty} x) \iff (x_n^{\otimes |y|} \xrightarrow{n \to \infty} x^{\otimes |y|}) \iff ((x_n^{\otimes |y|}, y) \xrightarrow{n \to \infty} (x^{\otimes |y|}, y) \text{ for all } y \in E_Y). \]

Hence, \( \mathcal{H}_X \) induces the weak topology on \( E_X \) and Le Cam’s theorem implies that \( \mathcal{H}_X \) is convergence determining.

For the dynamics of the dual process, let \( Y \) be the Markov jump process, which jumps from \( Y_t = y \)

\[ \text{(i) for all } 1 \leq i < j \leq |y|, \]

\[ \text{to } y \circ \theta_{ij} \text{ at rate 1,} \]

noting that \(|y \circ \theta_{ij}| = |y| - 1;\]
(ii) for all $1 \leq k \leq |y|$ to $\beta_k y$ at rate $\theta$, 
\begin{equation}
\beta_k y(u_1, \ldots, u_n) := \int y(u_1, \ldots, u_{k-1}, v, u_{k+1}, \ldots, u_n) \beta(u_k, dv).
\end{equation}

where $\beta_k y$ is a second order operator, this solution has continuous paths

Since $Y$ is a pure jump process with bounded jump rates, $Y$ is the unique solution of the martingale problem $G_Y$ for
\begin{equation}
G_Y H(x, \cdot)(y) := G_Y H(\cdot, y)(x),
\end{equation}
where the right hand side is from (4.27). This already shows that (2.3) holds (with $\beta = 0$).

In order to apply Theorem 2.1 we start with (2.5) using Proposition 2.8 and check the assumptions (i)-(iv) made there. Since $Y$ is Feller, and $E_Y$ is dense in $C(E_U)$ as an algebra containing 1 due to the Stone-Weierstrass theorem, (i) and (iii) hold. For (iv), i.e. the positivity of $P_t$, let us have a closer look at the two possible transitions of $y$ from above. If $y \geq 0$, note that $y \circ \theta_t \geq 0$ and $\beta_t y \geq 0$. Writing $Y^y$ for the process $Y$ when started in $y$, and looking at the transitions of $Y^y$, it is clear that $y \mapsto Y^y_t$ is linear and $Y^y_t \geq 0$ as well, and consequently $y \mapsto \mathbb{E}[H(x, Y^y_t)]$ is a positive linear form. In addition, if $y = 1$, then $Y^y_t = 1$, so $\mathbb{E}[H(x, Y^1_t)] = 1$, which shows that all properties of (iii) hold.

For (iv), we will make use of the reformulation given in Corollary 2.9(iv') to verify (2.10). Therefore, we define the set of permutations on $\mathbb{N}$
\begin{equation}
\Sigma = \bigcup_{n=0}^{\infty} \Sigma_n, \quad \Sigma_n := \{\sigma : \mathbb{N} \to \mathbb{N} \text{ bijective}, \sigma|_{\{n,n+1,\ldots\}} = \text{id}\}.
\end{equation}
and the set of functions
\begin{equation}
\mathcal{F} = \{f_\sigma : E_U \to E_U, u \mapsto u_{\sigma} \text{ with } \sigma \in \Sigma, f \in C_b(U)\},
\end{equation}
where $u_\sigma = (u_{\sigma(1)}, u_{\sigma(2)}, \ldots)$. Since $x^{\text{Gale}}$ is exchangeable, we have that $H(x, y \circ f_\sigma) = H(x, y)$ as well as
\begin{equation}
G_Y^{\text{res}} H(x, \cdot)(y \circ f_\sigma) = G_Y^{\text{res}} H(x, \cdot)(y), \quad G_Y^{\text{mut}} H(x, \cdot)(y \circ f_\sigma) = G_Y^{\text{mut}} H(x, \cdot)(y).
\end{equation}
Let us turn to the proof of (2.10) for an $E_U$-valued random variable $U$ with $U \sim f(U)$ for all $f \in \mathcal{F}$ and $E_U$-valued random variable. In other words, $U$ is exchangeable and using the de Finetti’s theorem we obtain for such $U$ that there is an $E_X = M_1(I)$-valued random variable $X$ such that $U \sim X^{\text{Gale}}$ conditional on $X$. In other words, (2.10) holds. Hence, Proposition 2.8 gives (2.5).

For the measurability, Proposition 2.5 gives (2.4) since $\mathcal{H}_X$ is convergence determining. So, we have shown all assumptions in Theorem 2.1 and we obtained a Feller process $X$ as a solution of the $G_X$-martingale problem. Moreover, there exists a càdlàg modification of $X$ by Proposition 2.3. Again, since $G_X$ is a second order operator, this solution has continuous paths by Proposition 4.5 in [DGP12].
Example 4 (Interacting Fleming-Viot and Cannings). Here we add space and migration to the Fleming-Viot process or the Cannings process from Examples 1 and 2. The approach is based on Remark 2.12. For some countable, discrete Abelian group $G$, we assume that $a : G \times G \to \mathbb{R}$ is a transition kernel such that $a(\xi, \eta) = a(0, \eta - \xi)$ with $\sum_\xi a(0, \xi) < \infty$. Setting $E_x = (M_t(I))^{G}$ and $E_y = \bigcup_{m=0}^\infty G^m \times C(I^m)$, we use for $\xi \in \mathcal{G}^m$ the probability measure $p_\xi := x_{\xi_1} \otimes \cdots \otimes x_{\xi_n}$ to define the duality function

$$H(x; \xi, y) := \langle x_\xi, y \rangle := \int x_\xi(du)y(u).$$

The migration operator reads

$$G^{(1)}_x H(\cdot; \xi, y)(x) = \sum_{i=1}^{d} \sum_{\eta \in \mathcal{G}} a(\xi_i, \eta)(x_{\eta} - x_\xi, y),$$

where $\xi^i = (\xi_1, \ldots, \xi_{i-1}, \eta, \xi_{i+1}, \ldots)$, whereas resampling and mutation operator are joined in

$$G^{(2)}_x H(\cdot; \xi, y)(x) = \sum_{1 \leq i < j \leq |\xi|} 1_{\xi_i = \xi_j} \langle x_\xi, y \circ \theta_{ij} - y \rangle + \theta \sum_{1 \leq k \leq |\xi|} \langle x_\xi, \beta_k y - y \rangle.$$  

The dual process $Y$ is a system of delayed coalescing random walks. More precisely, the $\xi$-component moves according to the random walk kernel $a$, and the function-component $y$ follows coalescence of coordinates $i$ and $j$ at unit rate, if $\xi_i = \xi_j$, and changes to $\beta_k y$ at rate $\theta$. With $G^{(i)}_y H(x, \cdot)(\xi, y) := G^{(i)}_x H(\cdot; \xi, y)(x)$, $i = 1, 2$, note that the dynamics of the $\xi$-component happens according to $G^{(1)}_y$, and of the function-component $y$ according to $G^{(2)}_y$.

We sketch the application of Remark 2.12, i.e. we need to check conditions (a), (b), (c) of Theorem 2.11. As in Example 3, $\text{span}(\mathcal{H}_x)$ is a convergence determining algebra of functions containing 1, i.e. (a) holds. Then, the $(G^{(i)}_y, \mathcal{H}_y)$ martingale problems for $i = 1, 2$ are well-posed (for $G^{(1)}_y$, we obtain a system of random walks, for $G^{(2)}_y$, we obtain independent coalescence processes), as well as the $(G_y := G^{(1)}_y + G^{(2)}_y, \mathcal{H}_y)$ martingale problem, its solution being a Markov jump process, the delayed spatial coalescent with random walk kernel $a$. Since both dual processes have bounded rates, they are Feller and (iii) of Proposition 2.6 holds. For (iv) of Proposition 2.6 we have to show that the (well-defined) mass flow induced by the transition kernel $a$ as given through $G^{(1)}_x$ is in duality with independent random walks on $\mathcal{G}$. This is well-known; cf. [DGV95], and uses that $G^{(1)}_x$ is a first order operator; see Remark 2.13. In particular, this gives well-posedness of the $(G^{(1)}_x, \mathcal{H}_x)$ martingale problem and therefore existence of $\mu^{(1)}_t$, as indicated in Remark 2.12. For $G^{(2)}_x$, we see that a $\mathcal{G}$-indexed family of processes, distributed independently as solutions of the $(G_y, \mathcal{H}_y)$-martingale problem from Example 3 is the unique solution of the $(G^{(2)}_x, \mathcal{H}_x)$ martingale problem. Altogether, (b) of Theorem 2.11 holds. For (c), we note that $Y$ (with generator $G^{(1)}_y + G^{(2)}_y$ again is a Markov jump process with bounded jump rates, hence Feller. So, Theorem 2.11 gives well-posedness of the $(G_y, \mathcal{H}_y)$ martingale problem, and a modification with càdlàg paths. Again, since $G_x$ is a second order operator, this solution has continuous paths by Proposition 4.5 in [DGP12].
Remark 4.1 (Extensions of Example 3). The above construction reproduces in particular the existence result of the system of interacting Fleming-Viot processes on the discrete hierarchical group, discussed in [EF96] and gives existence on general discrete abelian groups in an alternative way to [DGV95].

In order to obtain the spatial Cannings model from [GdHK14], we would then only need to check that Theorem 2.1 is applicable to the non-spatial Cannings process, since the operator in that case is an integral operator and not a second order differential operator, the argument in Example 1 has to be adapted at the point of the calculation done in (4.15), which are based on the property (4.11) of the dual dynamics and has been detailed in Example 2 for the non-spatial Cannings model. Hence this provides the details for the argument in [GdHK14].

Example 5 (Spatial Lambda-Fleming-Viot process). In a series of papers, Etheridge and co-authors have introduced and studied the spatial Λ-Fleming-Viot process [BEV10, EFS17, EK19, EVY20]. The model has been studied in more detail and extended in [VW15, CDK19, LV22].

In early papers, the existence of the process via the solution of a martingale problem is actually obtained using duality via the same approach as in [Eva97]; see [BEV10]. Later papers show existence by using convergence of approximate models; see e.g. [EVY20]. Here, we will describe how existence of a solution of the martingale problem in its most basic form can be constructed using duality as an application of Corollary 2.7. Here, we will use a novel function-valued dual. Dual processes similar in spirit have been used, but they were not function-valued.

We will use the spatial Λ-Fleming-Viot model with type space [0, 1] excluding mutation and selection.

Fix $d \in \mathbb{N}$, set $\mathbb{E} := \mathbb{R}^d$, as well as the state space of the process $X$, which is

$$E_X := \{x \text{ measure on } \mathcal{B}(\mathbb{E} \times [0, 1]): \pi x = \lambda\},$$

where $\pi : \mathbb{E} \times [0, 1] \to \mathbb{E}$ is a projection and $\lambda$ is Lebesgue-measure on $\mathbb{E}$. (Some $x \in E_X$ models a population with constant density across $\mathbb{E}$, but the density of types $\kappa \in [0, 1]$ may vary.) Equipping $E_X$ with the vague topology makes it a compact metric space; see Lemma 1.1 in [VW15]. Following Theorem 3.4 of [Kal21], recall that for each $x \in E_X$, there is a Markov kernel from $\mathbb{E}$ to $[0, 1]$ such that

$$\int f(u, \kappa) x(du, d\kappa) = \int \lambda(du) \int w(u, d\kappa)f(u, \kappa) \text{ for all } f \in C_c(\mathbb{E} \times [0, 1]).$$

We start with an informal description of the process. Let $\mu$ be a $\sigma$-finite measure on $(0, \infty) \times (0, 1]$ such that

$$\int r^d p \mu(dr, dp) < \infty,$$

and $\Pi$ be a Poisson process on $[0, \infty) \times \mathbb{E} \times (0, \infty) \times (0, 1]$ with intensity measure $dt \otimes \lambda(dv) \otimes \mu(dr, dp)$. Then, for $(t, v, r, p) \in \Pi$, and if the current state of the process is $x$, set $K = \kappa$ with probability proportional to $x(B_t(v), d\kappa)$, $\kappa \in [0, 1)$, and the process changes to

$$x_{i, r, p, \kappa}(du, d\kappa) := 1_{|u-v| \leq r} x(du, d\kappa) + 1_{|u-v| < r}((1 - p)x(du, d\kappa) + p\lambda(du) \otimes \delta_K(d\kappa))$$

$$= \lambda(du)(p1_{|u-v| < r}\delta_K(d\kappa) + (1 - p)1_{|u-v| < r}w(u, d\kappa)).$$
This means that the offspring of one parent, chosen at random from \( B_r(v) \) replaces a fraction \( p \) of the total population within \( B_r(v) \). The offspring inherits the type of their parents. Since \( \lambda \otimes \mu \) is an infinite measure, there are infinitely many events within each time interval, and the issue of existence of such a process must be answered.

For the function-valued dual process, let

\[
E_Y = \bigcup_{n=1}^{\infty} \Pi_n, \quad \Pi_n := \left\{ y \in C_c(\mathbb{E}^n) : y \geq 0, \int y(u)\lambda(du) = 1 \right\}
\]

and set \(|y| = n\) if \( y \in \Pi_n \), i.e. \( \Pi_n \) consists of densities for samples of size \( n \) with compact support.

For the duality function, we use \( E_Y \) as well as the duality function (recall \( w \) from (4.39) and set \( u := (u_1, \ldots, u_{|y|}) \))

\[
H(x, y) = \int \bigotimes_{i=1}^{|y|} \lambda(du_i)y(u) \int \bigotimes_{i=1}^{|y|} w(u_i, dk_i)k_1 \cdots k_{|y|}.
\]

In words, this gives the probability of choosing a sample of \(|y|\) individuals of type 1, if they are sampled according to the density \( y \). On such functions, we are going to show existence of a process solving the martingale problem for the operator

\[
G_X H(\cdot, y)(x) = \int \lambda(dr) \int \mu(dp) \frac{1}{|B_r(0)|} \int \lambda(du')1_{|u-u'|<r} \int w(u', dk')
\]

\[
\cdot \int (x^{(0)}_{v,r,p,e}(du, dk) - x^{(0)}_{v,e}(du, dk))k_1 \cdots k_{|y|}y(u).
\]

In order to evaluate the right hand side and find a function-valued dual, observe that

\[
x^{(0)}_{v,r,p,e}(du, dk) - x^{(0)}_{e}(du, dk)
\]

\[
= \sum_{k=0}^{n} \sum_{I \subseteq \{1, \ldots, n\} |I| = k} \lambda(du_i) \left( \bigotimes_{i \in I} p 1_{|u_i-v|<r} \right) \left( \bigotimes_{i \in I} \delta_{v_i}(dk_i) - \bigotimes_{i \in I} w(u_i, dk_i) \right)
\]

\[
\cdot \prod_{i \notin I} (1 - p 1_{|u_i-v|<r}) \left( \bigotimes_{i \notin I} w(u_i, dk_i), \right.
\]

and we can write

\[
G_X H(\cdot, y)(x) = \sum_{k=1}^{n} \sum_{I \subseteq \{1, \ldots, n\} |I| = k} A_{n,I}(x, y),
\]
with \( k := |I| \) and some \( t \in I \) (plugging (4.44) in (4.43) for (4.46), evaluating the integrals with respect to \( \kappa_i, i \notin I \) for (4.47), and using \(- \prod_{i \in I} 1_{|u_i| < r} = (1 - \prod_{i \in I} 1_{|u_i| < r}) - 1\) for (4.48).

\[
\begin{align*}
A_{n,I}(x,y) &= \int \lambda(dv) \int \mu(dr, dp)p^k \frac{1}{|B_r(0)|} \int \lambda(du')1_{|v' - u'| < r} \int w(u', dk') \\
&\quad \cdot \left( \prod_{i=1}^n \lambda(du_i) \prod_{i \notin I} 1_{|u_i| < r} \left( \prod_{i \in I} \delta_{\kappa_i}(d\kappa_i) \right) - \prod_{i \notin I} w(u_i, d\kappa_i) \right) \\
&\quad \cdot \prod_{i \notin I} \left( 1 - p1_{|u_i| < r} \right) \int \bigotimes_{i \notin I} w(u_i, d\kappa_i) y(u) \prod_{i \in I} \kappa_i \\
&= \int \mu(dr, dp)p^k|B_r(0)| \cdot \left( \prod_{i \notin I} \lambda(du_i) \prod_{i \notin I} \left( 1 - p1_{|u_i| < r} \right) \right) \\
&\quad \cdot \left( \prod_{i \notin I} w(u', d\kappa') \kappa' \prod_{i \notin I} w(u_i, d\kappa_i) \prod_{i \in I} \kappa_i \right) \\
&+ \left( \prod_{i=1}^n \lambda(du_i) \frac{1}{|B_r(0)|^2} \int \lambda(dv) \int \lambda(du') \int 1_{|v' - u'| < r} 1_{|u| < r} y(u) \\
&\quad \cdot \prod_{i \notin I} \left( 1 - \prod_{i \in I} 1_{|u_i| < r} \right) \prod_{i \notin I} \left( 1 - p1_{|u_i| < r} \right) \right) \int \bigotimes_{i=1}^n w(u_i, d\kappa_i) \prod_{i=1}^n \kappa_i \\
&\quad - \prod_{i=1}^n \lambda(du_i) y(u) \int \bigotimes_{i=1}^n w(u_i, d\kappa_i) \prod_{i=1}^n \kappa_i. 
\end{align*}
\]

We interpret the right hand side saying that \( y \) jumps to \( y'_{I,r,p} + y''_{I,r,p} \) with \( y'_{I,r,p} \in \Pi_{n-k+1} \) (note that \( 1 \leq n - k + 1 \leq n \) since \( 1 \leq |I| = k \leq n \)) and \( y''_{I,r,p} \in \Pi_n \) at rate \( \mu(dr, dp)p^k(1-p)^{k-1}|B_r(0)| \), where

\[
\begin{align*}
y'_{I,r,p} &= \frac{1}{|B_r(0)|^2} \int \lambda(dv)1_{|v' - u'| < r} \int \bigotimes_{i \notin I} \lambda(du_i) y(u) \prod_{i \notin I} 1_{|u_i| < r} \left( 1 - p1_{|u_i| < r} \right) \\
&\quad \cdot \left( 1 - \prod_{i \in I} 1_{|u_i| < r} \right) \prod_{i \notin I} \left( 1 - p1_{|u_i| < r} \right), \\
y''_{I,r,p} &= \frac{1}{|B_r(0)|^2} \int \lambda(du') \int \lambda(dv)1_{|v' - u'| < r} 1_{|u| < r} y(u) \\
&\quad \cdot \left( 1 - \prod_{i \notin I} 1_{|u_i| < r} \right) \prod_{i \notin I} \left( 1 - p1_{|u_i| < r} \right). 
\end{align*}
\]
Note that by construction
\[\int \lambda(du') \bigotimes_{u' \in I} \lambda(du_i)y'_{1,r,p}(u', (u_i)_{i \in I}) + \int \bigotimes_{i=1}^n \lambda(du_i)y''_{1,r,p}(u) = 1 \tag{4.51}\]
and we can change variables in $y'_{1,r,p}$ (i.e. changing $u'$ with $u_i$) such that $y'_{1,r,p} + y''_{1,r,p}$ depend on the same variables $u_1, \ldots, u_n$.

Let us use these calculations in order to show existence and uniqueness of the $G_X$-martingale problem. We will use Theorem 2.1, in particular Corollary 2.7, using the dual process $Y = (Y_t)_{t \geq 0}$. We argue as follows: First, span($\mathcal{H}_X$) is a convergence determining algebra since $x_n \rightarrow x$ vaguely if and only if $\int y(u)x_n(du) \xrightarrow{n \rightarrow \infty} \int y(u)x(du)$ for all $y \in C_c(\mathbb{E})$; see also Lemma 1.1 in [EVY20]. Second, note that for $Y_t = y \in \Pi_j$, the dual process jumps to $y'_{1,r,p} + y''_{1,r,p} \in \Pi_n$ at rate $\int \mu(dr, dp) |B_r(0)| < \infty$, as we seen above. Note that this rate is bounded by the left hand side of (4.40), such that $Y$ is the unique solution of its martingale problem and (2.3) holds (with $\beta = 0$) for the function-valued pure jump process $Y$, which is $C_0(E_Y)$-Feller. So, (2.4) and (2.5) follow from Propositions 2.5 and 2.6 provided we can show Proposition 2.6(iv). For this, we use Corollary 2.7 and argue similarly as in the proof of Theorem 1.2 of [EVY20]. We use an approximating sequence of models $X^1, X^2, \ldots$ with duals $Y^1, Y^2, \ldots$, which arise by restricting for $X^n$ to reproduction events on $(-n, n)^d$ and some finite $\mu^p \leq \mu$ on $(0, \infty) \times (0, 1]$, such that $\mu^p \uparrow \mu$ as $n \rightarrow \infty$. For these dual pairs $(X^n, Y^n)$, the construction guarantees:

\begin{enumerate}[(i)]
    \item The dual processes converge, i.e. $Y^n \xrightarrow{n \rightarrow \infty} Y$, since $Y$ is a pure Markov jump process (with finite jump rate), and jumping intensities converge;
    \item The martingale problems for $X^1, X^2, \ldots$ are well-posed and unique solutions of the corresponding martingale problems, since $X^n$ is a pure Markov jump process with finite jump rate.
\end{enumerate}

From Corollary 2.7, we see that Proposition 2.6(iv) holds and thus, we have shown existence and uniqueness of the $(G_X, \mathcal{H}_X)$-martingale problem.

### 4.2 Locally-compact state spaces – branching systems

**Example 6** (Continuous state branching processes). For the construction of a superprocess, E. Dynkin uses in [Dyn93] what he calls the direct construction, which can be viewed as a duality argument. In fact, this approach is connected to Theorem 2.1 which we demonstrate now for simplicity for a non-spatial branching system.

The state space of the process that we wish to construct is $E_X = \mathbb{R}_+$. To define the operator let $b \in \mathbb{R}$, $c \in \mathbb{R}_+$ and let $N$ be a measure on $[0, \infty)$ with $\int_0^\infty (s \wedge s^2)N(ds) < \infty$ and $\int_0^\infty s^2N(ds) = 0$. We set
\[\mathcal{D}_X = C^2_c(\mathbb{R}_+), \tag{4.52}\]
where $C^2_c(\mathbb{R}_+)$ denotes the set of twice continuously differentiable real-valued functions on $\mathbb{R}_+$ with compact support. The operator for the process we aim to construct is given by (see [DL06])
eq. (5.23) for a more general case

$$G_X f(x) = bxf'(x) + cx f''(x) + x \int_0^\infty (f(x + s) - f(x) - sf'(x))N(ds).$$  (4.53)

Note that for $N = 0$, this is the generator of a Feller diffusion with drift. Let $H : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$, $H(x, y) = e^{-xy}$ and let $Y^'$ be the deterministic process satisfying $Y_0 = y$ and solving

$$\dot{Y} = -\Psi(Y) \text{ with } \Psi(y) = by + cy^2 + \int_0^\infty (e^{-sy} - 1 + sy)N(ds).$$  (4.54)

Here, $\Psi$ is usually referred to as the branching mechanism. The generator of $Y$ is given by

$$G_y e^{-xy} = -\Psi(y) \frac{\partial}{\partial y} e^{-xy} = x \Psi(y) e^{-xy} = \left( by + cy^2 + \int_0^\infty (e^{-sy} - 1 + sy)N(ds) \right) xe^{-xy}$$

$$= bxx + cy^2 e^{-xy} + x \int_0^\infty (e^{-s(y+y')} - e^{-sy} - s \frac{\partial}{\partial y} e^{-sy})N(ds).$$  (4.55)

Then, for $(2.5)$, we need to find a random variable $X_t$ such that, for $Y^t_i$ solving $(4.54)$ with $Y_0 = y$,

$$\mathbb{E}_x[e^{-Y^t_i}] = e^{-\Psi(y)}. \quad (4.56)$$

So, we need to see if $\psi(y)$ is the Laplace transform of some $\mathbb{R}_+$-valued random variable. This is equivalent to the following four conditions: (i) $\psi$ is continuous, (ii) $\psi$ is positive definite, (iii) $\psi \geq 0$ and (iv) $\psi(0) = 1$. See for instance [BCR84, Corollary 4.5, p. 114] for the case of finite measures and note that (iv) ensures that we have a probability measure. Clearly, (i), (iii) and (iv) are satisfied. Condition (ii) is equivalent to the requirement that $y \mapsto Y^y$ is negative definite; see [BCR84, Proposition 6.10, p. 133]. This, however, is proved in [Bez11, Proposition 3.2(v)], and hence, we have shown $(2.5)$. Finally, $(2.4)$ follows as in Proposition 2.5. Since $Y$ is Feller, $X$ is Feller as well. For path regularity, the compact containment condition for $X$ can be proved using a priori moment bounds to get compact containment for fixed times $t$ and then using Doob’s inequality to for the argument on the paths space. Then the existence of a càdlàg modification follows; compare with Remark 1.1.

Now we give an example with $\beta \neq 0$ in the duality relation (2.5). We prepare this example with two lemmas.

**Lemma 4.2.** Let $Y = (Y_t)_{t \geq 0}$ be a pure jump process with countable state space and denote by $y_0$ the start point of $Y$ and by $Y_k$ the state of $Y$ after the $k$th jump, $k = 1, 2, \ldots$. Moreover, the total jump rate of $Y$ in state $y'$ is denoted $\gamma(y')$ and the jump rate from $y'$ to $y''$ by $\gamma(y' \to y'')$. Then, for any $f$,

$$\mathbb{E}_x \left[ f(Y_t) \cdot \exp \left( \int_0^t \gamma(Y_s)ds \right) \right] = \sum_{n=0}^\infty \left( \sum_{y_1, \ldots, y_n} f(y_n) \prod_{k=0}^{n-1} \gamma(y_k \to y_{k+1}) \right), \quad (4.57)$$

where $\prod_{k=0}^{-1} := 1$, if the right hand side exists.
Example 7 (Feller’s branching diffusion). Here, we have $E_X = \mathbb{R}_+$. The operator and its domain are given by

$$
G_X f(x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} f(x), \quad \mathcal{D}_X = C^2_0([0, \infty)).
$$

Proof. Let $N_t$ be the number of jumps before time $t$. Then for $n \geq 1$ we can compute as follows

$$
\mathbb{E}[f(Y_t) \cdot \exp \left( \int_0^t \gamma(y_s) ds \right), N_t = n] = \sum_{y_1, \ldots, y_n} \int_0^t dt_1 \gamma(y_0) \ldots \int_0^{t_{n-1}} dt_n \gamma(y_{n-1}) \ldots e^{-\gamma(y_t) t_n} \ldots e^{-\gamma(y_1) t_1} f(y_n) = \frac{r^n}{n!} \sum_{y_1, \ldots, y_n} f(y_n) \prod_{k=0}^{n-1} \gamma(y_k \rightarrow y_{k+1}).
$$

An analogous equation holds for $n = 0$. Summing over $n$ gives the assertion. □

The following result is standard and formulated here for reference in the next example.

**Lemma 4.3** (Moments, Bernstein functions and Laplace transforms). Let $(m_y)_{y \geq 0}$ be a sequence of non-negative real numbers. Define $\psi : (0, \infty) \rightarrow \mathbb{R}$ by

$$
\psi(\lambda) = \sum_{y=0}^{\infty} \frac{(-\lambda)^y}{y!} m_y.
$$

Assume that for some $x > 0$ there is a function $\varphi$ so that $\psi(\lambda) = e^{-\varphi(\lambda)}$ for all $\lambda > 0$. If $\varphi$ admits the representation

$$
\varphi(\lambda) = a + b \lambda + \int_{(0, \infty)} (1 - e^{-dr}) \nu(dr),
$$

where $a, b \geq 0$ and $\nu$ is a measure on $(0, \infty)$ satisfying $\int_{(0, \infty)} (1 \land r) \nu(dr) < \infty$, then there exists a unique non-negative measure $\mu$ on $[0, \infty)$ so that

$$
\psi(\lambda) = \int_{(0, \infty)} e^{-\lambda r} \mu(dr).
$$

Proof. The assertion of the lemma follows by a combination of results from [SSV12]. By [SSV12, Theorem 3.2] the function $\varphi$ is a Bernstein function which by [SSV12, Theorem 3.7] is equivalent to the fact that $\psi$ is a completely monotone function. By [SSV12, Theorem 1.4] it must be a Laplace transform of a unique measure $\mu$ on $[0, \infty)$. □

**Example 7** (Feller’s branching diffusion). Here, we have $E_X = \mathbb{R}_+$. The operator and its domain are given by

$$
G_X f(x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} f(x), \quad \mathcal{D}_X = C^2_0([0, \infty)).
$$
The state space of the dual process is \( E_Y = \mathbb{N} \). For the duality function we choose – similar to Example 1 – \( H(x, y) = x^y \) and we let \( Y \) be the Markov jump process with generator

\[
G_Y f(y) = \left( \frac{y}{2} \right) (f(y - 1) - f(y)).
\]

(4.61)

For \( \beta(y) = \left( \frac{y}{2} \right) \), this gives

\[
G_X H(\cdot, y)(x) := G_Y H(\cdot, y) + \beta(y)H(x, y) = \left( \frac{y}{2} \right) x^{y-1} = \frac{1}{2} x^2 \frac{\partial^2}{\partial x^2} x^y,
\]

(4.62)

which we recognize as the generator of Feller’s branching diffusion on \([0, \infty)\). Hence, for \( x \in E_X \) and \( t \geq 0 \), in order to show (2.5), we need to find (the law of a random variable) \( X_t \) such that for all \( y \in E_Y \) we have

\[
\mathbb{E}_y[X_t^y] = m_y := \mathbb{E}_y \left[ x^t \exp \left( \int_0^t \left( \frac{Y_s}{2} \right) ds \right) \right] = \sum_{n=0}^{\infty} \frac{t^n}{n!} x^{y-n} \prod_{k=0}^{n-1} \left( \frac{y-k}{2} \right).
\]

(4.63)

where we have used Lemma 4.2 in the last step. (Note that the product of binomial coefficients is interpreted as 1 in cases \( y = 1 \) or \( n = 0 \).) In order to find \( X_t \), we will use Lemma 4.3 Setting \( m_0 = 1 \) we have a sequence \((m_y)_{y=0,1,...}\) and for \( \psi \) as in (4.58) we obtain

\[
\psi(\lambda) = \sum_{y=0}^{\infty} \frac{(-\lambda)^y}{y!} m_y = 1 + \sum_{y=0}^{\infty} \sum_{n=0}^{y-1} \frac{(-\lambda)^y t^n}{n!} x^{y-n} \prod_{k=0}^{n-1} \left( \frac{y-k}{2} \right)
\]

\[
= 1 + \sum_{y=1}^{\infty} \sum_{n=0}^{y-1} \frac{(-\lambda x)^y n!}{n!} \left( \frac{t}{2x} \right)^n \frac{(y-1)!}{(y-n-1)! (y-1)!}
\]

\[
= 1 + \sum_{y=1}^{\infty} \sum_{n=0}^{y-1} \frac{(-\lambda x)^y n!}{n!} \left( \frac{t}{2x} \right)^n \frac{(y-1)!}{(y-1)! (y-n-1) (y-1) (y-n-1)}
\]

\[
= 1 + \sum_{y=1}^{\infty} \sum_{n=0}^{y} \frac{(-\lambda x)^y n!}{n!} \left( \frac{t}{2x} \right)^n \frac{(y-1)!}{(y-1)! (y-n-1) (y-1) (y-n-1)}
\]

\[
= 1 + \sum_{n=0}^{\infty} \frac{(2x/t)^n}{n!} \frac{(-\lambda x) n!}{n!} \left( \frac{t}{2x} \right)^n \frac{(y-1)!}{(y-1)! (y-n-1) (y-1) (y-n-1)}
\]

\[
= 1 + \sum_{n=0}^{\infty} \frac{(2x/t)^n}{n!} \left( \frac{-\lambda x}{1+\lambda x/t} \right)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{-\lambda x}{1+\lambda x/t} \right)^n = \exp \left( -\frac{\lambda x}{1+\lambda x/t} \right).
\]

(4.64)

Apart from several elementary manipulations we have used that \( \sum_{k=0}^{\infty} (-a)^k \binom{t}{k} = \frac{e^{-at}}{1+at/2} \) in the first equality of the last line. Now the function \( \varphi(\lambda) = \frac{4}{1+\lambda x/t} \) can be written in the form (4.59) with \( a = b = 0 \) and \( \nu(dr) = (t/2)^{-2} \exp(-r/(t/2)) dr \). Indeed we have

\[
\int_{(0,\infty)} (1-e^{-dr}) (t/2)^{-2} e^{-r/(t/2)} dr = \frac{4}{t^2} (t/2 - \frac{t/2}{1+\lambda x/t}) = \varphi(\lambda).
\]

(4.65)

Now the existence of \( X_t \) or more precisely the existence and uniqueness of the corresponding laws follows by Lemma 4.3 and we have obtained \( (\mu_t)_{t \geq 0} \) as required in (2.4) and (2.5). We note...
5 Outlook on non-locally compact state spaces

Several of our results require $E_X$ to be compact. In particular, Proposition 2.6 is based on an application of the Riesz-Markov theorem, which works best for compact spaces. Also the proof of Theorem 2.11 uses Proposition 2.6. Hence in the cases of non-compact Polish state spaces we need to work with a suitable compactification.

Typical examples of state spaces $E_X$ which are not locally compact arise in models involving a continuum spatial component, genealogies or some function spaces. Examples where existence by duality was already obtained in the literature are [Eva97], [Dyn93] and [BEV10]. We briefly discuss (i) historical processes and (ii) genealogy-valued processes.

Let $G$ be a countably infinite abelian group. In (i), the state space is $\mathcal{M}(\mathcal{D}(\mathbb{R}, G))$, where $\mathcal{D}(\mathbb{R}, G)$ is the set of càdlàg paths on $G$; see [DP91, Daw93]. The idea is to associate with every individual alive at time $t$ its path of descend describing the geographical position of its ancestor at times $s \in [0, t]$ and extend the path before time 0 and after time $t$ as a constant path. Then the state space is a (locally finite) measure on the set of such paths and hence we have in general a non-locally compact state space.

For (ii), the state space is called $\mathcal{U}^G$, which is the set of (equivalence classes of) $G$-marked metric measure spaces, i.e. triples $(X, r, \mu)$, where $(X, r)$ is a metric space (coding for the genealogy) and $\mu \in \mathcal{M}(X \times G)$; see [GPW09, DGP12, GPW13, GSW16, DG23]. This leads to state spaces which are not $\sigma$-compact and not locally compact. In particular one needs to check tightness conditions to study convergence and path properties of stochastic processes.

In both cases, Theorem 2.1 is applicable, but checking (2.4) and (2.5) requires some additional work due to non-compactness of the state space. We note, that in studying such processes the technique of duality is very useful and applicable for our existence problem. We shall formulate below a criterion and a condition we need to verify in order to obtain the existence of a solution. To check this condition one needs to develop methods to verify that the paths of the process in the compactified state space remain in some subset whose preimage w.r.t. the embedding of the original space is contained in the original space itself. For fixed times $t$ this is known for genealogy-valued Fleming-Viot or Cannings models due to the so called strong duality. For all $t$, i.e. on the process level ongoing work in [GKW23] suggests that this issue will be resolved in the context of genealogy process by the construction of the ancestral web and its dual.

Let us now discuss the announced approach useful for dealing with general Polish state spaces $E_X$. The key is the following result, which reformulates and combines the strategies appearing in the literature e.g. in [KS01]. Note that we will be using this result for $\hat{E}$ compact.
Proposition 5.1 (How to treat general state spaces). Let $E, \hat{E}$ be Polish, $G : \mathcal{D} \subseteq \mathcal{B}(E) \to \mathcal{B}(E)$ and $\mathbb{P}_0 \in \mathcal{M}_1(E)$. Assume that $\Psi : E \to \hat{E}$ is injective and bi-measurable (i.e., $\Psi$ and $\Psi^{-1}$ are measurable). Set

$$\hat{\mathcal{D}} := \{ \tilde{g}_f \in \mathcal{B}(E) : f \in \mathcal{D}, \tilde{g}_f|_{\Psi(E)} = f \circ \Psi^{-1} \},$$

(5.1)

$$\hat{G} \tilde{g}_f(\tilde{x}) := \begin{cases} Gf(\Psi^{-1}(\tilde{x})), & \text{if } \tilde{x} \in \Psi(E), \\ 0, & \text{otherwise,} \end{cases}$$

(5.2)

$$\hat{\mathbb{P}}_0 := \Psi_\ast \mathbb{P}_0.$$  

(5.3)

(a) If $X$ solves the $(G, \mathcal{D}, \mathbb{P}_0)$ martingale problem, then $\Psi(X)$ solves the $(\hat{G}, \hat{\mathcal{D}}, \hat{\mathbb{P}}_0)$ martingale problem. If, in addition, $\Psi$ is continuous and $X$ has càdlàg (continuous) paths, then $\Psi(X)$ has càdlàg (continuous) paths as well.

(b) If $\hat{X}$ solves the $(\hat{G}, \hat{\mathcal{D}}, \hat{\mathbb{P}}_0)$ martingale problem and has paths in $\Psi(E)$, then $\Psi^{-1}(\hat{X})$ solves the $(G, \mathcal{D}, \mathbb{P}_0)$ martingale problem. If, in addition, $\Psi^{-1}$ is continuous and $\hat{X}$ has càdlàg (continuous) paths, then $\Psi^{-1}(\hat{X})$ has càdlàg (continuous) paths as well.

Note that (b) can be used in various ways following literature to develop criteria which additionally have to be checked for the existence of solutions of the martingale problem. Indeed if $E$ is locally compact, the above construction is well-known. In this case one can use the one-point compactification $\hat{E} := E \cup \{ \ast \}$ via $\Psi = \text{id}$; see e.g. Section 4.3 of [EK86].

An example for $\hat{E}$ in the case of not locally compact Polish space $E$ is as follows (see e.g. [KS01, Section 3] and [BK93]): Assume that there is $\mathcal{D}' \subseteq \mathcal{D}$ countable and separating such that

$$\text{bp-closure of } \{(g, Gg) : g \in \mathcal{D}' \} \supseteq \{(f, Gf) : f \in \mathcal{D} \}.$$  

(5.4)

Then consider the compact (in the product topology on $\mathbb{R}^\mathcal{D}$) set

$$\hat{E} = \prod_{g \in \mathcal{D}'} [-\sup |g|, +\sup |g|].$$

(5.5)

and use $\Psi : E_X \to \hat{E}_X$ via

$$\Psi(x) = (g(x))_{g \in \mathcal{D}'}.$$  

(5.6)

Since $\mathcal{D}'$ is separating, $\Psi$ is injective. If $\mathcal{D}' \subseteq C_b(E)$, $\Psi$ is continuous, and if $\mathcal{D}'$ is convergence determining then $\Psi^{-1}$ is continuous (on $\Psi(E)$).

The question is now how to work with $\hat{E}$. Suppose we want to use Proposition 2.6 or Theorem 2.1 writing $E_X$ and $\hat{E}_X$ for the state space of the process $X$. In case we want to use one of these results for showing (2.5), we can make use of Proposition 5.1 (assuming $\hat{E}_X$ is compact) and $\Psi : E_X \to \hat{E}_X$ is as in Proposition 5.1 as follows: We use $\mathcal{D}' \subseteq \{ H(\cdot, y) : y \in E_Y \}$. Then we can extend the duality w.r.t to function $H$ to a duality w.r.t. function $\hat{H}$ on $\hat{E}_X$, $\hat{H} : \hat{E}_X \times E_Y \to \mathbb{R}$, satisfying (i), (ii) of Theorem 2.1 and

$$\hat{H}(\Psi(x), y) = H(x, y), \quad x \in E_X, y \in E_Y.$$  

(5.7)
Then there is a family of transition kernels $(\tilde{\mu}_t)_{t \geq 0}$ from $\tilde{E}_X$ to $\tilde{E}_X$ satisfying (2.4) and (2.5) (which can be shown using e.g. Proposition 2.6 or Theorem 2.11 due to compactness of $\tilde{E}_X$), if additionally we have the following containment property of a solution starting in $E_X$:

$$\tilde{\mu}_t(\Psi(x), \Psi(E_X)) = 1 \text{ for all } t \geq 0, x \in E_X.$$ (5.8)

We define $\mu_t(x, \cdot) = \Psi^{-1} \tilde{\mu}_t(\Psi(x), \cdot)$, for all $t \geq 0, x \in E_X$ and $y \in E_Y$. Then, (2.4) holds for $(\mu_t)_{t \geq 0}$, since $\Psi$ and $\Psi^{-1}$ are measurable (on $\Psi(E_X)$); see Section 3 of [KS01]. Moreover, by using successively (5.7) and (2.5) on $\tilde{E}_X$, then the definition of $\mu_t$ and finally (5.7) again we obtain

$$\mathbb{E}_y[H(x, Y_t) \exp\left(\int_0^t \beta(Y_s) \, ds\right)] = \mathbb{E}_y[\tilde{H}(\Psi(x), Y_t) \exp\left(\int_0^t \beta(Y_s) \, ds\right)]$$
$$= \int \tilde{H}(x', y) \tilde{\mu}_t(\Psi(x'), dx')$$
$$= \int \tilde{H}(\Psi(x'), y) \mu_t(x, dx')$$
$$= \int H(x', y) \mu_t(x, dx')$$ (5.9)

which shows (2.5) for $(\mu_t)_{t \geq 0}$. We obtain the following corollary.

**Corollary 5.2.** Let $E_X, E_Y, H, G_Y, \beta, \mathcal{H}_X$ and $\mathcal{H}_Y$ be as in Theorem 2.1 and let Theorem 2.1(i) hold. In addition, let $E := E_X$ and $\tilde{E}$ be as in Proposition 5.1. If $\tilde{X}$ satisfies Proposition 5.1(ii) (in particular (5.8), then (2.4) and (2.5) hold for $\tilde{X}$, existence and uniqueness of a solution to the $G_X$-martingale problem follows.

As discussed in Remark 1.1 we need to check the regularity of paths separately. Here it means that we have to check that (5.8) holds as an additional condition and we use it in the dual process or an extension of it. The first step would be to establish (5.8) for fixed $t$ and to then in a second step exclude exceptional points of the paths. This can sometimes be done using the dual process $Y$ or rather its extension to a strong duality.

Then we find that the $(G_X, \mathcal{H}_X)$ martingale problem has a unique solution which has a càdlàg modification, since for general state spaces, Theorem 4.3.6 in [EK86] states the existence of a càdlàg modification of the $(G_X, \mathcal{H}_X)$ martingale problem provided the compact containment condition holds.

**Acknowledgements**

We thank Tom Kurtz for useful suggestions concerning the non-locally compact case and pointing out a reference. This research was supported by the DFG priority program SPP 1590 through the grants Pf672/8-1 to PP and Gr876/16-1,2 to AG. AG was also supported by the DFG-Grant Gr876/17-1. PP is partially funded by the Freiburg Center for Data analysis and Modeling (FDM).
Declaration of Generative AI and AI assisted technologies in the writing process

During the preparation of this work the author(s) used no generative AI tools.

References

[BCR84] Christian Berg, Jens Peter Reus Christensen, and Paul Ressel. *Harmonic analysis on semigroups*, volume 100 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984. Theory of positive definite and related functions.

[BEV10] N. Barton, A. Etheridge, and A. Véber. A new model for evolution in a spatial continuum. *Electron. J. Probab.*, 15:162–216, 2010.

[BEV13] N H Barton, A M Etheridge, and A Véber. Modelling evolution in a spatial continuum. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01):P01002, jan 2013.

[Bez11] Lucian Beznea. Potential-theoretical methods in the construction of measure-valued Markov branching processes. *J. Eur. Math. Soc. (JEMS)*, 13(3):685–707, 2011.

[BK93] A.G. Bhatt and R.L. Karandikar. Invariant measures and evolution equations for Markov processes. *Ann. Probab.*, 21:2246–2268, 1993.

[CDK19] Jonathan Chetwynd-Diggle and Aleksander Klimek. Rare mutations in the spatial lambda-fleming-viot model in a fluctuating environment and superbrownian motion. *arXiv preprint arXiv:1904.02044*, submitted January 2023.

[Daw93] Donald A. Dawson. Measure-valued Markov processes. In *École d’Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of *Lecture Notes in Math.*, pages 1–260. Springer, Berlin, 1993.

[DG23] Andrej Depperschmidt and Andreas Greven. Genealogy-valued Feller diffusion. *http://arxiv.org/abs/1904.02044*, submitted January 2023.

[DGP12] Andrej Depperschmidt, Andreas Greven, and Peter Pfaffelhuber. Tree-valued Fleming-Viot dynamics with mutation and selection. *Ann. Appl. Probab.*, 22(6):2560–2615, 2012.

[DGV95] Donald A. Dawson, Andreas Greven, and Jean Vaillancourt. Equilibria and quasiequilibria for infinite collections of interacting Fleming-Viot processes. *Trans. Amer. Math. Soc.*, 347(7):2277–2360, 1995.

[DL06] D. A. Dawson and Zenghu Li. Skew convolution semigroups and affine Markov processes. *Ann. Probab.*, 34(3):1103–1142, 2006.

[Doo53] J. L. Doob. *Stochastic processes*. John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, 1953.
REFERENCES

[DP91] Donald A. Dawson and Edwin A. Perkins. Historical processes. *Mem. Amer. Math. Soc.*, 93(454):iv+179, 1991.

[Dyn93] E. B. Dynkin. Superprocesses and partial differential equations. *Ann. Probab.*, 21(3):1185–1262, 1993.

[EF96] Steven N. Evans and Klaus Fleischmann. Cluster formation in a stepping-stone model with continuous, hierarchically structured sites. *Ann. Probab.*, 24(4):1926–1952, 1996.

[EFS17] Alison Etheridge, Nic Freeman, and Daniel Straulino. The Brownian net and selection in the spatial \( \Lambda \)-Fleming-Viot process. *Electron. J. Probab.*, 22:Paper No. 39, 36, 2017.

[EK86] Stewart N. Ethier and Thomas 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.

[EK19] Alison M. Etheridge and Thomas G. Kurtz. Genealogical constructions of population models. *Ann. Probab.*, 47(4):1827–1910, 2019.

[Eth00] Alison M. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.

[Eva97] Steven N. Evans. Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(3):339–358, 1997.

[EVY20] Alison M. Etheridge, Amandine Véber, and Feng Yu. Rescaling limits of the spatial \( \Lambda \)-Fleming-Viot process with selection. *Electron. J. Probab.*, 25:Paper No. 120, 89, 2020.

[GdHKK14] A. Greven, F. den Hollander, S. Kliem, and A. Klimovsky. Renormalisation of hierarchically interacting Cannings processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11(1):43–140, 2014.

[GKW23] Andreas Greven, Anton Klimovsky, and Anita Winter. Ancestral web and measure-valued path representations. in preparation, 2023.

[GPW09] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces (\( \Lambda \)-coalescent measure trees). *Probab. Theory Related Fields*, 145(1-2):285–322, 2009.

[GPW13] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Tree-valued resampling dynamics Martingale problems and applications. *Probab. Theory Related Fields*, 155(3-4):789–838, 2013.
[GSW16] Andreas Greven, Rongfeng Sun, and Anita Winter. Continuum space limit of the genealogies of interacting Fleming-Viot processes on Z. Electron. J. Probab., 21:Paper No. 58, 64, 2016.

[JK14] Sabine Jansen and Noemi Kurt. On the notion(s) of duality for Markov processes. Probab. Surv., 11:59–120, 2014.

[Kal21] Olav Kallenberg. Foundations of modern probability, volume 99 of Probability Theory and Stochastic Modelling. Springer, Cham, 2021. Third edition [of 1464694].

[KS88] N. Konno and T. Shiga. Stochastic partial differential equations for some measure-valued diffusions. Probab. Theory Related Fields, 79(2):201–225, 1988.

[KS91] I. Karatzas and S. E. Shreve. Brownian motion and stochastic calculus. Springer-Verlag, New York, 1991.

[KS01] Thomas G. Kurtz and Richard H. Stockbridge. Stationary solutions and forward equations for controlled and singular martingale problems. Electron. J. Probab., 6:no. 17, 52, 2001.

[LeC57] Lucien LeCam. Convergence in distribution of stochastic processes. Univ. Calif. Publ. Statist., 2:207–236, 1957.

[Lig85] Thomas M. Liggett. Interacting particle systems, volume 276 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, New York, 1985.

[LR16] Wolfgang Löhr and Thomas Rippl. Boundedly finite measures: separation and convergence by an algebra of functions. Electron. Commun. Probab., 21:Paper No. 60, 16, 2016.

[LV22] Apolline Louvet and Amandine Véber. Growth properties of the infinite-parent spatial lambda-fleming-viot process. arXiv preprint arXiv:2205.03937, 2022.

[MT95] C. Müller and R. Tribe. Stochastic p.d.e.’s arising from the long range contact and long range voter processes. Probab. Theory Related Fields, 102(4):519–545, 1995.

[Shi94] Tokuzo Shiga. Two contrasting properties of solutions for one-dimensional stochastic partial differential equations. volume 46, pages 415–437, 1994.

[SSV12] René L. Schilling, Renming Song, and Zoran Vondraček. Bernstein functions, volume 37 of De Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin, second edition, 2012. Theory and applications.

[VW15] A. Véber and A. Wakolbinger. The spatial Lambda-Fleming-Viot process: an event-based construction and a lookdown representation. Ann. Inst. Henri Poincaré Probab. Stat., 51(2):570–598, 2015.