

DISS. ETH NO. 25227

PROBABILITY MEASURE-VALUED
JUMP-DIFFUSIONS IN FINANCE
AND RELATED TOPICS

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
SARA SVALUTO-FERRO

MSc ETH Mathematics

born on 15.08.1989
citizen of Glattfelden (ZH), Switzerland

Prof. Dr. Martin Larsson examiner
Prof. Dr. Josef Teichmann co-examiner
Prof. Dr. Jan Kallsen co-examiner

2018

Abstract

The martingale problem is known to be a classical and very powerful technology for the study of Markov processes through the analytical properties of the corresponding generators. This approach permits for instance to easily model Markov processes featuring both diffusion and jump components. The first aim of the thesis is to develop concrete and easy-to-use tools allowing to employ this technology in the context of probability measure-valued jump-diffusions. This is mainly motivated by our second line of research, consisting in the study of the so-called polynomial jump-diffusions, which have found broad applications in mathematical finance. However, we will also show that the same tools can be used in much greater generality.

Given a linear operator \mathcal{G} playing the role of the generator, the martingale problem asks for a process $X := (X_t)_{t \geq 0}$, called *jump-diffusion*, such that

$$p(X_t) - \int_0^t \mathcal{G}p(X_s) ds$$

is a martingale for each test function p . Under very general conditions, the existence of such an X is essentially equivalent to \mathcal{G} satisfying the *positive maximum principle*. This very elegant property requests $\mathcal{G}p(x)$ to be nonpositive for each test function p attaining a nonnegative maximum at x .

A jump-diffusion X is called *polynomial* if its generator \mathcal{G} maps any polynomial to a polynomial of equal or lower degree. Polynomial jump-diffusions were introduced in [Cuchiero et al. \(2012\)](#), see also [Filipović and Larsson \(2016\)](#). Their most relevant property in a finite dimensional setting is the *moment formula*, stating how their conditional moments can be expressed in closed form by solving a system of linear ODEs, or equivalently, by computing a matrix exponential. Choosing the space of probability measures as state space for the process, one needs first to understand which notion of polynomial, or more generally of derivative, is convenient to use. With the choice that we will make, a first degree monomial is an expression of the form

$$\langle g, \nu \rangle := \int g(x) \nu(dx),$$

where g denotes a test function and is referred to as the coefficient of the monomial. In this setting, a generalized version of the moment formula still holds. Given a probability measure-valued polynomial jump-diffusion X , it provides in particular an expression for the conditional moments of a first degree monomial $\langle g, X_t \rangle$ in terms of a solution of a linear partial (integro) differential equation.

The starting point of this thesis is the simplest context for probability measure-valued polynomial jump-diffusions, namely where the underlying space is finite. In this setting, every probability measure can be identified with a point on the unit simplex. After that, we will use the know-how gained by working on the simplex, to study probability measure-valued jump-diffusions on locally compact Polish spaces, both in the polynomial and in the general case. The similarity between the structure obtained in the infinite dimensional setting and that in the finite dimensional one is strikingly high.

Sunto

Il problema di martingala è noto per essere uno strumento molto efficace per lo studio di processi di Markov attraverso l'indagine delle proprietà analitiche del corrispondente generatore. Questo approccio permette per esempio di modellizzare facilmente processi di Markov con un comportamento sia diffusivo che discontinuo. Lo scopo principale della tesi consiste nello sviluppare strumenti concreti e facili da utilizzare al fine di applicare queste tecniche nel contesto delle diffusioni con salti a valori nello spazio delle probabilità. La motivazione principale su cui si basa questa scelta risiede nella nostra seconda linea di ricerca, concernente lo studio delle cosiddette diffusioni polinomiali con salti, note per la vasta applicabilità in matematica finanziaria. Tuttavia, mostreremo come di fatto gli stessi strumenti possano essere impiegati anche in un contesto più generale.

Sia dato un operatore lineare \mathcal{G} facente ruolo di generatore. Un processo $X := (X_t)_{t \geq 0}$, chiamato diffusione con salti, risolve il problema di martingala relativo a \mathcal{G} se

$$p(X_t) - \int_0^t \mathcal{G}p(X_s) ds$$

è una martingala per ogni funzione test p . Essenzialmente, sotto condizioni molto generali l'esistenza di tale X è garantita se e solo se \mathcal{G} soddisfa il principio del massimo positivo. Quest'elegante proprietà richiede che $\mathcal{G}p(x)$ sia nonpositivo per ogni funzione test p avente un massimo nonnegativo in x .

Una diffusione con salti X è detta polinomiale se il suo generatore mappa polinomi in polinomi conservandone il grado. Le diffusioni polinomiali con salti sono state introdotte in [Cuchiero et al. \(2012\)](#), si veda anche [Filipović e Larsson \(2016\)](#). La loro proprietà più importante in un contesto finito dimensionale è la formula dei momenti, la quale stabilisce come i momenti condizionali possano essere espressi in forma chiusa risolvendo un sistema di equazioni differenziali ordinarie, o equivalentemente, calcolando l'esponenziale di una matrice. Al fine di lavorare con lo spazio delle misure di probabilità come spazio degli eventi, è necessario stabilire quale nozione di polinomio, o più in generale di derivata, è più conveniente usare. Con la scelta che faremo, un monomio di primo grado è un'espressione della forma

$$\langle g, \nu \rangle := \int g(x) \nu(dx),$$

dove g denota una funzione test a cui ci riferiamo come coefficiente del monomio. In questo contesto, una versione generalizzata della formula dei momenti può essere stabilita. Data una diffusione polinomiale con salti X a valori nello spazio

delle probabilità, essa fornisce un'espressione per i momenti condizionali di un monomio di primo grado $\langle g, X_t \rangle$ in termini di una soluzione di un'equazione (integrale) differenziale lineare alle derivate parziali.

Il punto di partenza di questa tesi è il contesto più semplice in cui una diffusione polinomiale con salti a valori nello spazio delle probabilità possa avere luogo, ossia quando lo spazio di base è costituito da un numero finito di punti. In queste condizioni, ogni misura di probabilità può essere identificata con un punto sul simpleso unitario. In seguito, useremo la conoscenza acquisita lavorando sul simpleso per studiare diffusioni con salti a valori nello spazio delle probabilità su spazi polacchi localmente compatti, sia nel caso polinomiale che in quello generale. La somiglianza tra la struttura ottenuta nel caso infinito-dimensionale e quella nel caso finito è straordinaria.

Acknowledgements

I would like to begin by expressing my gratitude to my PhD supervisor, Martin Larsson, who gave me the opportunity to work with many fascinating and nevertheless challenging topics. His ideas, enthusiasm, scientific rigor, and vast knowledge taught me the right way of tackling mathematical research. Especially, I want to thank him for always believing in me, for always pushing me beyond my limits to question more, and for all the fascinating discussions which definitively opened my mind a bit more.

My gratitude goes also to Christa Cuchiero for her enthusiasm and her always innovative ideas. She has provided very interesting external insights for my studies and taught me to think about the biggest picture. I benefited a lot from our collaboration and all the interesting discussions.

I would also like to thank Jan Kallsen and Josef Teichmann, who accepted to be part of my thesis committee.

I wish also to thank all my colleagues and friends in the stochastic finance, insurance mathematics, and probability theory group of ETH for all the coffees and the aperos, and particularly Michel Baes, Dániel Bálint, Matteo Burzoni, Thomas Cayé, Aditi Dandapani, Philippe Deprez, Leif Döring, John Ery, Tolupe Fadina, Andrea Gabrielli, Matteo Gambarà, Lukas Gonon, Xi Kleinsinger-Yu, Thomas Krabichler, Philipp Harms, Calypso Herrera, Sebastian Herrmann, Blanka Horvath, Vincenzo Ignazio, Chong Liu, Marvin Müller, Ariel Neufeld, Oleg Reichmann, Max Reppen, Martin Stefanik, and David Stefanovits.

My thanks goes also to all my friends from Ticino, who helped me to rightly mix work with fun. Among the others, my special thanks goes to Francesca Chiesa, Elena Vandea, Luisa Delmuè, Giorgia Fedeli, and all the other friends in Ticino, as well as all of them in Zurich, Angelo Abächerli, Marco Eigenmann, Chiara Salvini, Anna Rapp, Rade Bankovic, Marco Bernasconi, Daniele Farrace, Athos Fiori, Nicolò Frapolli, Elisa Righetti, and Nastassia Spadini. Thanks to all of you.

Last but not least, a big thank to all my family, which constantly supported me in every step of my life, and which is and will always be my reference point in this world of changes. A huge, special thank goes to my future husband Luca Papariello, for the love and patience he had all along of my studies and this work.

Contents

I Introduction	1
I.1 The finite dimensional setting	2
I.2 The infinite dimensional setting	4
II Polynomial jump–diffusions on the unit simplex	11
II.1 Preliminaries	13
II.2 Polynomial operators on compact state spaces	13
II.2.1 The positive maximum principle	14
II.2.2 Lévy type representation	15
II.2.3 Conic combinations of polynomial operators	17
II.3 Affine and polynomial jump sizes	17
II.4 The unit interval	19
II.5 Examples of polynomial operators on the unit interval	23
II.5.1 Examples with affine jump sizes	23
II.5.2 Constructions using conic combinations	26
II.5.3 Mixed affine jump sizes	27
II.6 The unit simplex	31
II.7 Applications	35
II.7.1 Market weights with jumps in stochastic portfolio theory	36
II.7.2 Valuation of defaultable zero–coupon bonds	37
Appendices	
II.A Proof of Theorem II.2.8	38
II.B Proof of Theorem II.2.10 and Lemma II.2.11	40
II.C Proof of Theorem II.3.3	41
II.D The unit interval: Proof of Lemma II.4.1 and Theorem II.4.3	42
II.E The unit simplex: Proof of Lemma II.6.1 and Theorem II.6.3	47
III Probability measure-valued jump-diffusions	55
III.1 Introduction	55
III.2 Preliminaries	57
III.2.1 Notation	57
III.2.2 Differentiation for functions of measure argument	58
III.2.3 Monomials and polynomials of measure argument	58
III.2.4 Cylindrical functions and cylindrical polynomials	62
III.3 The martingale problem	64

III.3.1	The positive maximum principle	65
III.3.2	Properties of $M_1(E)$ -solution	67
III.3.3	Continuity of $M_1(E)$ -solutions	70
III.4	Lévy type operators	70
III.5	Optimality conditions	76
III.5.1	Generalization of the classical Karush–Kuhn–Tucker conditions	76
III.5.2	An optimality condition beyond the finite dimensional setting	78
III.5.3	An application of Theorem III.5.3	80
Appendices		
III.A	Proof of Theorem III.4.2	82
III.B	Auxiliary Lemmas	84
IV	Probability measure-valued polynomial diffusions and other applications	87
IV.1	Introduction	87
IV.2	Polynomial operators	89
IV.3	Existence and uniqueness of polynomial diffusions on $M_1(E)$	91
IV.3.1	Moment formula and uniqueness in law	92
IV.3.2	Existence and well-posedness	94
IV.4	Applications of probability measure-valued polynomial diffusions	98
IV.4.1	The unit simplex	98
IV.4.1.1	The classical moment formula	100
IV.4.1.2	Approximation by diffusions on unit simplexes of increasing dimension	100
IV.4.2	The underlying space $E \subseteq \mathbb{R}^d$	101
IV.5	Examples	104
IV.5.1	The role of the parameters of a polynomial operators	105
IV.5.1.1	The role of B	105
IV.5.1.2	The role of the jumps term in B	105
IV.5.1.3	The role of α	106
IV.5.1.4	The role of τ	106
IV.5.1.5	When α and τ appear together	107
IV.5.2	The role of the parameters Q and N of Lévy type operators	108
IV.5.2.1	Probability measure-valued martingale	108
IV.5.2.2	Polynomial jump-diffusions	109
IV.5.2.3	Spatial jump-diffusions	110
IV.5.3	Particle systems with mean fields interaction	112
IV.5.3.1	Dynamics of the first order moments	113
IV.5.3.2	Particle systems and McKean-Vlasov equations	113
IV.5.3.3	A jump-diffusion with mean fields interactions	115
IV.5.3.4	Interactions through weighted empirical measures	115
IV.5.4	$M_1(E)$ -solutions taking values in a subset of $M_1(E)$	117
IV.5.4.1	Systems of finitely many particles	117
IV.5.4.2	Empirical measures of branching processes	118

Appendices

IV.A Proof of Theorem IV.2.2 and a generalization	120
IV.B Proof of Theorem IV.3.10	122
IV.C Auxiliary Lemmas	124

Bibliography	127
---------------------	------------

Chapter I

Introduction

The martingale problem is known to be a classical and very powerful technology for the study of Markov processes through the analytical properties of the corresponding generators. This approach permits for instance to easily model Markov processes featuring both diffusion and jump components. The first aim of the thesis is to develop concrete and easy-to-use tools allowing to employ this technology in the context of probability measure-valued jump-diffusions. This is mainly motivated by our second line of research, consisting in the study of the so-called polynomial jump-diffusions, which have found broad applications in mathematical finance. However, we will also show that the same tools can be used in much greater generality.

Given a linear operator \mathcal{G} playing the role of the generator, the martingale problem asks for a process $X := (X_t)_{t \geq 0}$, called *jump-diffusion*, such that

$$p(X_t) - \int_0^t \mathcal{G}p(X_s) ds$$

is a martingale for each test function p . Under very general conditions, the existence of such an X is essentially equivalent to \mathcal{G} satisfying the *positive maximum principle*. This very elegant property requests $\mathcal{G}p(x)$ to be nonpositive for each test function p attaining a nonnegative maximum at x .

A jump-diffusion X is called *polynomial* if its generator \mathcal{G} maps any polynomial to a polynomial of equal or lower degree. Polynomial jump-diffusions were introduced in [Cuchiero et al. \(2012\)](#), see also [Filipović and Larsson \(2016\)](#). Their most relevant property in a finite dimensional setting is the *moment formula*, stating how their conditional moments can be expressed in closed form by solving a system of linear ODEs, or equivalently, by computing a matrix exponential. Choosing the space of probability measures as state space for the process, one needs first to understand which notion of polynomial, or more generally of derivative, is convenient to use. With the choice that we will make, a first degree monomial is an expression of the form

$$\langle g, \nu \rangle := \int g(x) \nu(dx),$$

where g denotes a test function and is referred to as the coefficient of the monomial.

In this setting, a generalized version of the moment formula still holds. Given a probability measure-valued polynomial jump-diffusion X , it provides in particular an expression for the conditional moments of a first degree monomial $\langle g, X_t \rangle$ in terms of a solution of a linear partial (integro) differential equation.

The starting point of this thesis is the simplest context for probability measure-valued polynomial jump-diffusions, namely where the underlying space is finite. In this setting, every probability measure can be identified with a point on the unit simplex. After that, we will use the know-how gained by working on the simplex, to study probability measure-valued jump-diffusions on locally compact Polish spaces, both in the polynomial and in the general case. The similarity between the structure obtained in the infinite dimensional setting and that in the finite dimensional one is strikingly high.

1 The finite dimensional setting

Tractable families of Markov processes on the unit simplex play an important role in a host of applications. These include population genetics ([Etheridge, 2011](#); [Epstein and Mazzeo, 2013](#)), dynamic modeling of probabilities ([Gourieroux and Jasiak, 2006](#)), and mathematical finance, in particular stochastic portfolio theory ([Fernholz, 2002](#); [Fernholz and Karatzas, 2005](#)). In the context of polynomial diffusions (without jumps), the unit simplex has already appeared numerous times in the literature. In population genetics, prototypical diffusion processes on the unit simplex known as *Wright-Fisher diffusions*, or *Kimura diffusions* (see [Kimura \(1964\)](#)), arise naturally as infinite population limits of discrete Wright-Fisher models for allele prevalence in a population of fixed size; see [Etheridge \(2011\)](#) for a survey.

In finance, similar processes have appeared in [Gourieroux and Jasiak \(2006\)](#) under the name of *multivariate Jacobi processes*. All these diffusions turn out to be polynomial, and a full characterization is provided in [Filipović and Larsson \(2016, Section 6.3\)](#) by means of necessary and sufficient parameter restrictions on the drift and diffusion coefficients.

More generally, due to their inherent tractability, polynomial jump-diffusions have played a prominent and growing role in a wide range of applications in finance. Examples include interest rates ([Delbaen and Shirakawa, 2002](#); [Filipović et al., 2017](#)), credit risk ([Akerer and Filipović, 2016](#)), default risk ([Krabichler and Teichmann, 2017](#)), exchange rates ([Larsen and Sørensen, 2007](#)), stochastic volatility models ([Gourieroux and Jasiak, 2006](#); [Akerer et al., 2016](#)), life insurance liabilities ([Biagini and Zhang, 2016](#)), variance swaps ([Filipović et al., 2016](#)), dividend futures ([Filipović and Willems, 2017](#)), and commodities and electricity [Filipović et al. \(2017\)](#). Applications of polynomial jump-diffusions on the unit simplex also appear in the context of stochastic portfolio theory ([Cuchiero, 2017](#)).

In addition, polynomial jump-diffusions are highly flexible in that they allow for a wide range of state spaces – the unit simplex being one of them – and a multitude of possible jump and diffusion phenomena. This stands in contrast to the thoroughly studied and frequently used sub-class of affine processes. Any

affine jump-diffusion that admits moments of all orders is polynomial, but there are many polynomial jump-diffusions that are not affine. In particular, an affine process on a compact and connected state space is necessarily deterministic; see [Krühner and Larsson \(2018\)](#). Thus our interest in the unit simplex forces us to look beyond the affine class.

As most of the cited papers focus on the case without jumps, it is natural to ask what happens in the jump-diffusion case, where the literature is much less developed. This case is considered by [Cuchiero et al. \(2012\)](#) and [Filipović and Larsson \(2017\)](#), however without treating questions of existence, uniqueness, and parameterization for polynomial jump-diffusions on specific state spaces.

As mentioned before, our approach is based on the technology of the *martingale problem*. This is based on two ingredients: a linear operator \mathcal{G} whose domain consists of polynomials on the unit simplex Δ^d , and a point $x \in \Delta^d$. They will play the role of the generator and the initial condition of the corresponding jump-diffusion, defined as a càdlàg process X with values in Δ^d such that $X_0 = x$ and the process $(N_t^f)_{t \geq 0}$ given by

$$N_t^f := f(X_t) - \int_0^t \mathcal{G}f(X_s) ds$$

is a martingale for all polynomials f on Δ^d .

The key tool for dealing with the martingale problem is given by the *positive maximum principle*. The operator \mathcal{G} is said to satisfy this property if

$$f \text{ polynomial, } x \in \Delta^d, \quad f(x) = \max_{\Delta^d} f \geq 0 \quad \Rightarrow \quad \mathcal{G}f(x) \leq 0.$$

The positive maximum principle is essentially equivalent to the existence of a solution to the martingale problem for \mathcal{G} for all $x \in \Delta^d$. Even more, [Theorem II.2.3](#) illustrates that in the setting of polynomial operators this property is essentially enough to guarantee uniqueness in law of the solutions. This is remarkable since uniqueness is often difficult to establish.

From those considerations it is now clear that a generator that does not satisfy the positive maximum principle is not interesting for the study of jump-diffusions. For this reason, the next step consists in studying the structure that this property imposes on linear operators. In the spirit of [Courrège \(1965\)](#) and [Hoh \(1998\)](#), we prove that a linear operator satisfying the positive maximum principle is of *Lévy-type*, specified by a drift, diffusion, and jump triplet (b, a, ν) ; see [Theorem II.2.8](#).

With this result, we are now in the position to study polynomial jump-diffusions on the unit simplex by studying the properties of the corresponding triplet of parameters.

Our next goal in this finite dimensional setting is to provide a concrete, large, class of polynomial jump-diffusions on the unit simplex. The aim is to describe the properties of the elements of that class, in order to have them ready-to-use in practice. To do that, we restrict our attention to jump specifications with affine jump sizes, namely,

$$\nu(x, A) = \lambda(x) \int \mathbb{1}_{A \setminus \{0\}}(\gamma(x, y)) \mu(dy) \tag{1.1}$$

where $\lambda : E \rightarrow \mathbb{R}_+$ is a nonnegative measurable function, μ a Lévy measure and $\gamma = (\gamma_1, \dots, \gamma_d)$ is of the affine form

$$\gamma_i(x, y) = y_i^0 + y_i^1 x_1 + \dots + y_i^d x_d; \quad (1.2)$$

see Definition II.3.1. This is the most general specification in the class of jump kernels with polynomial dependence on the current state; see Theorem II.3.3. Under the structural hypothesis of affine jump sizes, we classify all polynomial jump-diffusions on the unit interval (i.e. the unit simplex in \mathbb{R}^2); see Theorem II.4.3. This classification is subsequently extended – under an additional assumption – to higher dimensions; see Theorem II.6.3. Referring to the unit interval for notational convenience, we can distinguish four types of jump-diffusions, in addition to the pure diffusion case without jumps:

Type 1: λ is constant and the support of $\nu(x, \cdot)$ is contained in $[-x, 1-x]$;

Type 2: λ is (essentially) a linear-rational function with a pole of order one at the boundary, and the process can only jump in the direction of the pole;

Type 3: λ is (essentially) a quadratic-rational function with a pole of order two in the interior of the state space. There is no jump activity at the pole, but an additional contribution to the diffusion coefficient.

Type 4: λ is a quadratic-rational function whose denominator has only complex zeros, and μ in (1.1) is of infinite variation.

This classification already gives an indication of the diversity of possible behavior, an impression which is strengthened in Section II.5, where we provide a number of examples both with and without affine jump sizes. On the one hand, these examples clearly show that without any structural assumptions like (1.1)–(1.2), a full characterization of *all* polynomial jump-diffusions on the simplex, or even the unit interval, is out of reach. On the other hand, the examples illustrate the richness and flexibility of the polynomial class. The findings in the finite dimensional setting are presented in Chapter II and will be published in Cuchiero et al. (2018a).

2 The infinite dimensional setting

The primary motivation for this part of the thesis stems from finding tractable classes of stochastic processes for *dynamic modeling of random probability measures*. The applications in this respect are rich and include population genetics, interacting particle systems, stochastic partial differential equations, statistical physics, non-parametric Bayesian inference, and mathematical finance, in particular stochastic portfolio theory.

The common denominator of these areas is modeling of systems (evolving in time) for which the spatial structure plays an essential role. For instance, in

population genetics the spatial motion of the individuals, corresponding to some Markov process on the underlying space E , is usually interpreted as the mutation between genetic types, constituting the space E . Since a continuum of types is natural (for several arguments in this direction see e.g. [Ethier and Kurtz \(1993\)](#)), we are automatically led to an uncountably infinite underlying space E . The way to deal with such a setting was pioneered by [Fleming and Viot \(1979\)](#) who introduced what is now called the *Fleming–Viot process* which is the infinite dimensional analog of the Wright–Fisher diffusion. While in population genetics the (random) distribution of the genetic types is the quantity of interest, in other areas, like physics or finance, it is the distribution of diffusive fluids or capital among companies with a “type space” that is usually \mathbb{R}^d or \mathbb{R}_+ for nonnegative capital. In general, dynamic modeling of random probability measures (beyond finite dimensional ones having the same finite support at each time) clearly necessitates an infinite dimensional setup with an underlying space E containing infinitely many points.

Our approach consists now in following step by step the construction provided in the finite dimensional setting. In particular, we will again define a jump-diffusion as a solution of a martingale problem, then show the (essential) equivalence between the positive maximum principle and the existence of solutions of the martingale problem, and finally characterize the structure induced by this property on linear operators.

The main object is thus again a linear operator, now denoted by L , whose domain consists of polynomials on the space $M_1(E)$ of probability measures on some locally compact underlying space E , endowed with the topology of weak convergence. The choice of the notion of derivative and the induced polynomial structure now becomes crucial. Our choice falls on the following natural notion of derivative, well-known since the work of [Fleming and Viot \(1979\)](#): for $p : M(E) \rightarrow \mathbb{R}$, we say that p is differentiable at μ in direction δ_x if

$$\partial_x p(\mu) := \lim_{\varepsilon \rightarrow 0} \frac{p(\mu + \varepsilon \delta_x) - p(\mu)}{\varepsilon} \quad (2.1)$$

exists, and we write $\partial p(\mu)$ for the map $x \mapsto \partial_x p(\mu)$. Note that for p regular enough, the map $\nu \mapsto \langle \partial p(\mu), \nu \rangle$ coincides with the Fréchet derivative of p at μ . This choice is in contrast with the notion of Lions’ derivative (see [Lions \(2007-2008\)](#) for the original video-taped lectures and [Cardaliaguet \(2012\)](#) for the corresponding transcriptions), which has often been used, for instance in the context of mean-field games; see e.g. [Carmona and Delarue \(2017\)](#) and the references given there. However, in the last part of the thesis we will illustrate that the machinery we are going to present provides a compact and efficient method to approach many different fields of application, see for example [Section IV.5.3](#) and [Section IV.5.4](#). This in particular shows that the derivative (2.1) is highly tractable and flexible.

The polynomial structure induced by this notion of derivative is very simple: a monomial of degree k consists of an expression of the form

$$\langle g, \nu^k \rangle := \int_{E^k} g(x_1, \dots, x_k) \nu(dx_1) \cdots \nu(dx_k), \quad (2.2)$$

where $g : E^k \rightarrow \mathbb{R}$ is a well-behaved symmetric continuous function and is referred to as the coefficient of the monomial.

Having now fixed the notion of derivative and the corresponding polynomial structure, we can define a probability measure-valued jump-diffusion as a solution of the martingale problem corresponding to L . As in the finite dimensional setting, the positive maximum principle is essentially equivalent to the existence of a corresponding jump-diffusion for all initial conditions, see Lemma III.3.5 and Lemma III.3.6. The similarity continue in Theorem III.4.2, stating that an operator L satisfying the positive maximum principle is necessarily a Lévy type operator, specified by a killing, drift, diffusion, and jump quadruplet (Γ, B, Q, N) . With this result we can thus conclude that the study of probability measure-valued jump-diffusions can again be approached by studying the corresponding quadruplet of parameters, completing the picture of the machinery that we are going to use to tackle the infinite dimensional setting.

Next, in order to guarantee that an operator satisfies the positive maximum principle, we need to dispose of (necessary) optimality conditions, namely, properties shared by all maps on $M_1(E)$ having a maximum at some point of the domain. In the finite dimensional setting such optimality conditions are very well studied, see for instance the classical Karush–Kuhn–Tucker conditions. In the infinite dimensional setting, and in particular on the space of probability measures, this is not the case. Theorem III.5.1 and Theorem III.5.3 provide the necessary conditions to cover a wide class of generators, and in particular all the examples presented in this thesis. It is interesting to observe that the condition provided by the second theorem describes a behavior that cannot appear if the underlying space (and thus in particular the support of the studied measures) consists of finitely many points. As we explain in Section IV.5.1.4 this condition is however essential in some very fundamental examples, for instance the probability measure-valued process δ_W for W denoting a brownian motion. The results on this general approach to the infinite dimensional setting are presented in Chapter III and will be published in Larsson and Svaluto-Ferro (2018).

The last, but not less important, unexploited knowledge gained from the unit simplex regards the polynomial property of the generator. Our goal here is to systematically characterize the class of probability measure-valued polynomial diffusions, i.e. the class of probability measure-valued continuous jump-diffusions whose generator maps a polynomial of measure argument (in sense of (2.2)) to a polynomial of measure argument of equal or lower degree. As in the finite dimensional setting, this leads to a moment formula, which provides an expression for the corresponding k -th moments in terms of a solution of a k -dimensional linear partial (integro) differential equation. In other words, the measure-valued Kolmogorov backward PIDE reduces in the case of probability measure-valued polynomial diffusions and a polynomial terminal condition to a k -dimensional linear PIDE.

Our systematic analysis automatically leads to Fleming–Viot type processes (Fleming and Viot, 1979; Ethier and Kurtz, 1993) which are – as already mentioned above – in their standard form certainly the most prominent examples of

probability measure-valued processes. Their popularity is in particular due to the fact that their moments can be explicitly calculated by solving a certain system of linear PDEs describing the evolution of their moment measures (see e.g. Dawson and Hochberg (1982) or Dawson (1993), Section 2.8). This feature is also key for the *method of duality* which has been widely applied in the study of infinite particle systems (see e.g. Liggett (2012)).

The standard way to construct Fleming–Viot processes is as weak limits of the empirical measures of certain particle systems, usually the Moran particle process (see e.g. (Dawson, 1993, Section 2.5)). This can be achieved either by adopting semigroup methods as in (Dawson, 1993, Section 2), or via martingale problem approaches as for instance in Ethier and Kurtz (1987, 1993). As explained before, our approach relies also on the martingale problem however without approximations but rather by directly characterizing the positive maximum principle of the respective generators. This in turn leads, as in the finite dimensional case, to a full characterization of probability measure-valued polynomial diffusions when the domain of the generator is assumed to be the space of all polynomials, see Theorem IV.3.10. The class which we obtain in this case is an extension of the so-called *Fleming–Viot process with weighted sampling* which is mentioned as an example in Dawson (1993), Section 5.7.8. Indeed, the constant sampling–replacement rate of the standard Fleming–Viot process is modified to allow for a type dependent rate.

As our method to prove existence of solutions of the martingale problem differs from the ones in the literature, the same is true also for our approach to duality and the computation of moment. Indeed, instead of considering function valued stochastic dual processes (see e.g. Dawson and Hochberg (1982)), our dual processes are solutions of *deterministic* PIDEs. Let us be more precise on this point. Observe that for each polynomial p of degree k there is a coefficient $g : E^k \rightarrow \mathbb{R}$ such that $p(\nu) = \langle g, \nu^k \rangle$ for all $\nu \in M_1(E)$ (see Corollary III.2.6; in the finite dimensional setting this coincides with the property that each polynomial on the unit simplex has a homogeneous representation). This leads to the definition of the k -th dual operator $L_k : D \rightarrow C(E^k)$ of L which is uniquely determined by

$$Lp(\nu) = \langle L_k g, \nu^k \rangle, \quad \nu \in M_1(E), \quad (2.3)$$

for every $p(\nu) = \langle g, \nu^k \rangle$ with $g \in D^{\otimes k}$ and D being a suitable dense subspace of $C(E)$. Theorem IV.3.2 then states that if L is sufficiently regular, then the conditional expectation of the k -moments of X_t is uniquely determined by the solution of the PIDE corresponding to L_k . Mathematically, given $g \in D^{\otimes k}$, a solution $u : \mathbb{R}_+ \times E^k \rightarrow \mathbb{R}$ of

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x) &= L_k u(t, \cdot)(x), & (t, x) &\in \mathbb{R}_+ \times E^k, \\ u(0, x) &= g(x), & x &\in E^k, \end{aligned} \quad (2.4)$$

and X being a polynomial jump-diffusion corresponding to L , the moment formula

$$\mathbb{E}[\langle g, X_T^k \rangle \mid \mathcal{F}_t] = \langle u(T - t, \cdot), X_t^k \rangle \quad (2.5)$$

holds true for all $T \geq t$. Moreover, if the moment formula holds for a sufficiently large set of coefficients g , then the law of X is uniquely determined by L and ν ; see Corollary [IV.3.4](#).

Formula [\(2.5\)](#) shows on the one hand the inherent tractability of probability measure-valued polynomial processes, in the sense that all moments can be computed by solving a linear PIDE. In the case of a finite state space this boils down to the system of ODEs known from the theory of finite dimensional polynomial diffusions. Even in the present case, when L_k is itself a polynomial operator, we can express the right hand side of [\(2.5\)](#) via matrix exponentials. On the other hand one can also turn this perspective around and view [\(2.5\)](#) as a stochastic representation of PIDEs of type [\(2.4\)](#) whose solutions can in turn be obtained by simulating X and applying Monte Carlo techniques. This is particularly appealing in high dimensions. See [Henry-Labordère et al. \(2014\)](#) for recent work in this direction. New numerical schemes for solving PIDEs thus naturally arise from the theory of probability measure-valued polynomial processes, through we do not pursue this idea in this thesis.

Now that the power of the moment formula has been explained, we can come back to the meaning of “sufficiently regular polynomial operator L ”. As in the finite dimensional setting, if the domain of L consists in the space of all polynomials, then the positive maximum principle is enough to guarantee that the moment formula holds true for all $g \in D^{\otimes k}$, see Theorem [IV.3.10](#). This in particular implies that the corresponding martingale problem is well-posed.

It turns out that restricting the domain of L to the subspace of polynomials whose coefficients are more regular allows for an enlargement of the class of linear operators considered by the theory. For instance, one can see by [\(2.3\)](#) that if we require the coefficients g to be sufficiently differentiable, then the operator L_k can be a differential (and thus in particular unbounded) operator. This reflects in an enlargement of the class of probability measure-valued polynomial diffusions covered by the theory, including for instance the probability measure-valued process δ_W for W denoting a brownian motion. Theorem [IV.3.7](#) provides a very general specification for L guaranteeing that it satisfies the positive maximum principle, and thus that the corresponding martingale problem has a solution for each initial condition. As one can imagine, reducing the amount of test functions for which the martingale problem has to hold, translates into an increase in difficulty of proving uniqueness, and thus in particular the conditions under which the moment formula holds true. We will however show, see e.g. Remark [IV.3.3](#) or Example [IV.4.4](#), that under simple conditions the moment formula still holds true for all $g \in D^{\otimes k}$ and the corresponding martingale problem is thus still well-posed.

The findings about probability measure-valued polynomial jump-diffusions are presented in Chapter [IV](#) and will be published in [Cuchiero et al. \(2018b\)](#).

The last part of the thesis is devoted to applications and examples of polynomial and non-polynomial probability measure-valued jump-diffusions. In Section [IV.4](#) we investigate polynomial diffusions for particular choices of state-spaces. The goal of Section [IV.5](#) is two-fold. On the one hand we present many different examples to explain how the various parameters influence the corres-

ponding probability measure-valued jump-diffusion (see Sections [IV.5.1–IV.5.2](#)). On the other hand we illustrate the flexibility of the proposed machinery and how it can be used to study a number of different problems including in particular particle systems with and without common noise and mean fields interactions (see Sections [IV.5.3–IV.5.4](#)).

Chapter II

Polynomial jump–diffusions on the unit simplex

As explained in the introduction, tractable families of Markov processes on the *unit simplex*, featuring both *diffusion and jump* components, are challenging to construct, yet play an important role in a host of applications. The present chapter addresses this challenge by specifying Markovian jump–diffusions on the unit simplex that are *polynomial*, meaning that the (extended) generator maps any polynomial to a polynomial of the same or lower degree.

Polynomial processes were introduced in [Cuchiero et al. \(2012\)](#), see also [Filipović and Larsson \(2016\)](#), and are inherently tractable. Indeed, any polynomial jump–diffusion

- (i) is an Itô semimartingale, meaning that its semimartingale characteristics are absolutely continuous with respect to Lebesgue measure. This justifies the name jump–diffusion in the sense of [Jacod and Shiryaev \(2003, Chapter III.2\)](#);
- (ii) admits explicit expressions for all moments in terms of matrix exponentials.

To analyze existence, uniqueness, and parametrization for polynomial jump–diffusions on the unit simplex, the technical difficulties associated with the diffusion case remain, arising from the fact that the unit simplex is a non-smooth stratified space ([Epstein and Mazzeo, 2013, Chapter 1](#)), and that the diffusion coefficient degenerates at the boundary. This complicates the analysis, and precludes the use of standard results regarding existence and regularity of solutions to the corresponding Kolmogorov backward equations. Additionally, in the jump case, the drift and diffusion interact with the (small) jumps orthogonal to the boundary, which leads to further mathematical challenges.

Allowing for jumps is not only of theoretical interest, but has practical relevance as well. A concrete illustration of this fact comes from stochastic portfolio theory ([Fernholz, 2002](#); [Fernholz and Karatzas, 2009](#)), where one is interested in the market weights $X_i = S_i / (S_1 + \dots + S_d)$ computed from the market capitalizations S_i , $i = 1, \dots, d$, of the constituents of a large stock index such as the S&P 500 or the MSCI World Index. The time evolution of the vector of

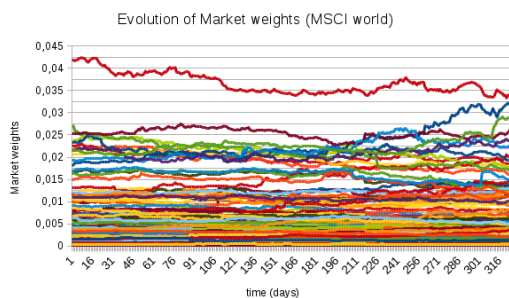


Figure 1: Market weights of the MSCI World Index, August 2006 – October 2007

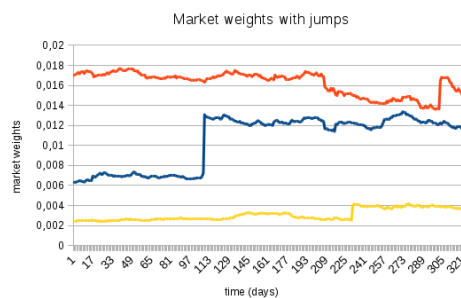


Figure 2: Example of market weights which exhibit jumps

market weights is thus a stochastic process on the unit simplex (see Figure 1, reprinted from Cuchiero et al. (2016)). To model the market weight process, polynomial diffusion models without jumps have been found capable of matching certain empirically observed features such as typical shape and fluctuations of capital distribution curves (Fernholz and Karatzas, 2005; Cuchiero, 2017; Cuchiero et al., 2016) when calibrated to *jump-cleaned* data. However, the absence of jumps is a deficiency of these models. Indeed, an inspection of market data shows that jumps do occur and are an important feature of the dynamics of the market weights. This is clearly visible in Figure 2 (reprinted from Cuchiero et al. (2016)) where, for illustrative purposes, three companies have been extracted from the MSCI World Index, whose market weights exhibit jumps in the period from August 2006 to October 2007. This application from stochastic portfolio theory underlines the importance of specifying jump structures within the polynomial framework. We elaborate on this in Section 7.1.

Another natural application of the results developed in this chapter arises in default risk modeling following the framework of Jarrow and Turnbull (1998) and Krabichler and Teichmann (2017). One is then interested in modeling a $[0, 1]$ -valued stochastic recovery rate which remains at level 1 for extended periods of time, while occasionally performing excursions away from 1. Polynomial jump-diffusion specifications turn out to be capable of producing such behavior, while at the same time maintaining tractability. Further details are given in Section 7.2.

The chapter is organized as follows. Section 1 summarizes some notation used throughout the chapter. Section 2 is concerned with polynomial operators on general compact state spaces and their associated martingale problems. Section 3 introduces affine jump sizes. In Section 4 we classify all polynomial jump-diffusions on the unit interval with affine jump sizes. It is followed by Section 5 which deals with examples. Section 6 treats the simplex in arbitrary dimension. Finally, Section 7 discusses applications in stochastic portfolio theory and default risk modeling. Most proofs are gathered in appendices.

1 Preliminaries

We denote by \mathbb{N} the natural numbers, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ the nonnegative integers, and \mathbb{R}_+ the nonnegative reals. The symbols $\mathbb{R}^{d \times d}$, \mathbb{S}^d , and \mathbb{S}_+^d denote the $d \times d$ real, real symmetric, and real symmetric positive semi-definite matrices, respectively. For any subset $E \subseteq \mathbb{R}^d$, we let as usual $C(E)$ denote the space of continuous functions on E . For any sufficiently differentiable function f we write ∇f for the gradient of f and $\nabla^2 f$ for the Hessian of f . Next, e_i stands for the i -th canonical unit vector, $|v|$ denotes the Euclidean norm of the vector $v \in \mathbb{R}^d$, δ_{ij} is the Kronecker delta, δ_x is the Dirac mass at x , and $\mathbf{1}$ is the vector whose entries are all equal to 1. We denote by $\text{Pol}(\mathbb{R}^d)$ the vector space of all polynomials on \mathbb{R}^d and $\text{Pol}_n(\mathbb{R}^d)$ the subspace consisting of polynomials of degree at most n . A *polynomial on E* is the restriction $p = q|_E$ to E of a polynomial $q \in \text{Pol}(\mathbb{R}^d)$. Its degree is given by

$$\deg p = \min\{\deg q : p = q|_E, q \in \text{Pol}(\mathbb{R}^d)\}.$$

We then let $\text{Pol}(E)$ denote the vector space of polynomials on E , and write $\text{Pol}_n(E)$ for those elements whose degree is at most n . We frequently use multi-index notation so that, for instance, $x^{\mathbf{k}} = x_1^{k_1} \cdots x_d^{k_d}$ for $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{N}_0^d$.

2 Polynomial operators on compact state spaces

Let $E \subseteq \mathbb{R}^d$ be a compact subset of \mathbb{R}^d that will play the role of the state space for a Markov process. Later we will specialize to the case where E is the unit interval or the unit simplex. In this chapter we are concerned with operators of the following type, along with solutions to the corresponding martingale problems.

Definition 2.1. *A linear operator $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ is called polynomial if*

$$\mathcal{G}(\text{Pol}_n(E)) \subseteq \text{Pol}_n(E) \quad \text{for all } n \in \mathbb{N}_0.$$

Given a linear operator $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ and a probability distribution ρ on E , a *solution to the martingale problem* for (\mathcal{G}, ρ) is a càdlàg process X with values in E defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(X_0 \in \cdot) = \rho$ and the process $(N_t^f)_{t \geq 0}$ given by

$$N_t^f := f(X_t) - \int_0^t \mathcal{G}f(X_s) ds \tag{2.1}$$

is a martingale with respect to the filtration $\mathcal{F}_t^X = \sigma(X_s : s \leq t)$ for every $f \in \text{Pol}(E)$. We say that *the martingale problem for \mathcal{G} is well-posed* if there exists a unique (in the sense of probability law) solution to the martingale problem for (\mathcal{G}, ρ) for any initial distribution ρ on E . If \mathcal{G} is polynomial, then X is called a *polynomial jump-diffusion*; this terminology is justified by Theorem 2.8 and the subsequent discussion.

2.1 The positive maximum principle

Definition 2.2. A linear operator $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ satisfies the positive maximum principle if $\mathcal{G}f(x_0) \leq 0$ holds for any $f \in \text{Pol}(E)$ and $x_0 \in E$ with $\sup_{x \in E} f(x) = f(x_0) \geq 0$.

Roughly speaking, the positive maximum principle is equivalent to the existence of solutions to the martingale problem. A typical result in this direction is Theorem 4.5.4 in [Ethier and Kurtz \(2005\)](#). For polynomial operators on compact state spaces more is true: we also get uniqueness.

Theorem 2.3. Let $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ be a polynomial operator. The martingale problem for \mathcal{G} is well-posed if and only if $\mathcal{G}1 = 0$ and \mathcal{G} satisfies the positive maximum principle.

Proof. The existence of a solution to the martingale problem for (\mathcal{G}, ρ) for any initial distribution ρ on E , is guaranteed by Theorem 4.5.4 and Remark 4.5.5 in [Ethier and Kurtz \(2005\)](#). To prove uniqueness in law, by compactness of E it is enough to prove that the marginal mixed moments of any solution X to the martingale problem for (\mathcal{G}, ρ) are uniquely determined by \mathcal{G} and ρ ; see Lemma 4.1 and Theorem 4.2 in [Filipović and Larsson \(2016\)](#). To this end, fix any $n \in \mathbb{N}$, let h_1, \dots, h_N be a basis of $\text{Pol}_n(E)$, and set $H = (h_1, \dots, h_N)^\top$. The operator \mathcal{G} admits a unique matrix representation $G \in \mathbb{R}^{N \times N}$ with respect to this basis, so that

$$\mathcal{G}p(x) = H(x)^\top G \vec{p},$$

where $p \in \text{Pol}_n(E)$ has coordinate representation $\vec{p} \in \mathbb{R}^N$, that is, $p(x) = H(x) \vec{p}$; cf. Section 3 in [Filipović and Larsson \(2016\)](#) and the proof of Theorem 2.7 in [Cuchiero et al. \(2012\)](#). Following the proof of Theorem 3.1 in [Filipović and Larsson \(2016\)](#) we use the definition of a solution to the martingale problem, linearity of expectation and integration, and the fact that polynomials on the compact set E are bounded, to obtain

$$\begin{aligned} \vec{p}^\top \mathbb{E}[H(X_T) | \mathcal{F}_t^X] &= \mathbb{E}[p(X_T) | \mathcal{F}_t^X] = p(X_t) + \mathbb{E} \left[\int_t^T \mathcal{G}p(X_s) ds | \mathcal{F}_t^X \right] \\ &= \vec{p}^\top H(X_t) + \vec{p}^\top G^\top \int_t^T \mathbb{E}[H(X_s) | \mathcal{F}_t^X] ds \end{aligned}$$

for any $t \leq T$ and any $\vec{p} \in \mathbb{R}^N$. For each fixed t this yields a linear integral equation for $\mathbb{E}[H(X_T) | \mathcal{F}_t^X]$, whose unique solution is $\mathbb{E}[H(X_T) | \mathcal{F}_t^X] = e^{(T-t)G^\top} H(X_t)$. Consequently,

$$\mathbb{E}[p(X_T) | \mathcal{F}_t^X] = \vec{p}^\top \mathbb{E}[H(X_T) | \mathcal{F}_t^X] = H(X_t)^\top e^{(T-t)G} \vec{p}, \quad (2.2)$$

which in particular shows that all marginal mixed moments are uniquely determined by \mathcal{G} and ρ , as required.

For the converse implication, observe that since every solution to the martingale problem for (\mathcal{G}, ρ) is conservative, the condition $\mathcal{G}1 = 0$ follows directly by

the martingale property of (2.1) with $f = 1$. The necessity of the positive maximum principle is standard; see for instance the proof of Lemma 2.3 in Filipović and Larsson (2016). \square

Remark 2.4. *Observe that a solution to the martingale problem is conservative by definition since it is supposed to take values in E . This is reflected by the condition $\mathcal{G}1 = 0$ of Theorem 2.3 and in the definition of Lévy type operator in the next section. Let us remark that the condition in Theorem 2.3, namely that the positive maximum principle and $\mathcal{G}1 = 0$ are satisfied, is equivalent to the maximum principle, that is, $\mathcal{G}f(x_0) \leq 0$ holds for any $f \in \text{Pol}(E)$ and any $x_0 \in E$ with $\sup_{x \in E} f(x) = f(x_0)$.*

Remark 2.5. *While existence of a solution to the martingale problem is equivalent to the maximum principle in very general settings, it is remarkable that in the case of polynomial operators on compact state spaces uniqueness also follows. Indeed, without the assumption that \mathcal{G} is polynomial, it is well-known that the maximum principle is not enough to guarantee uniqueness. For example, with $E = [0, 1]$ and $\mathcal{G}f(x) = \sqrt{x}(1-x)f'(x)$, the functions $X_t = (e^t - 1)^2 / (e^t + 1)^2$ and $X_t \equiv 0$ are two different solutions to the martingale problem for (\mathcal{G}, δ_0) . In the polynomial case, well-posedness is deduced from uniqueness of moments, which is a consequence of (2.2). Let us emphasize that (2.2) gives more than mere uniqueness: it gives an explicit formula for computing the moments via a matrix exponential. This tractability is crucial in applications, and was used as a defining property of this class of processes in Cuchiero et al. (2012).*

2.2 Lévy type representation

Definition 2.6. *An operator $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ is said to be of Lévy type if it can be represented as*

$$\begin{aligned} \mathcal{G}f(x) = & \frac{1}{2} \text{Tr} (a(x)\nabla^2 f(x)) + b(x)^\top \nabla f(x) \\ & + \int (f(x + \xi) - f(x) - \xi^\top \nabla f(x)) \nu(x, d\xi), \end{aligned} \tag{2.3}$$

where the right-hand side can be computed using an arbitrary representative of f , and the triplet (a, b, ν) consists of bounded measurable functions $a : E \rightarrow \mathbb{S}_+^d$ and $b : E \rightarrow \mathbb{R}^d$, and a kernel $\nu(x, d\xi)$ from E into \mathbb{R}^d satisfying

$$\sup_{x \in E} \int |\xi|^2 \nu(x, d\xi) < \infty, \quad \nu(x, \{0\}) = 0, \quad \nu(x, (E - x)^c) = 0 \quad \text{for all } x \in E. \tag{2.4}$$

Polynomial operators satisfying the positive maximum principle are always Lévy type operators, as is shown in Theorem 2.8 below. This parallels known results regarding operators acting on smooth and compactly supported functions, see Courrège (1965) or Böttcher et al. (2013, Theorem 2.21) for Feller generators, and also Hoh (1998). A crucial ingredient in the proof of Theorem 2.8 is the

classical Riesz-Haviland theorem, which we now state. A proof can be found in Haviland (1935) and (1936), or e.g. Marshall (2008).

Lemma 2.7 (Riesz-Haviland). *Let $K \subseteq \mathbb{R}^d$ be compact, and consider a linear functional $\mathcal{W} : \text{Pol}(K) \rightarrow \mathbb{R}$. Then the following conditions are equivalent.*

(i) $\mathcal{W}(f) = \int f(\xi)\mu(d\xi)$ for all $f \in \text{Pol}(K)$ and a Borel measure μ concentrated on K .

(ii) $\mathcal{W}(f) \geq 0$ for all $f \in \text{Pol}(K)$ such that $f \geq 0$ on E .

We now state Theorem 2.8 regarding the Lévy type representation of operators satisfying the positive maximum principle. The proof is given in Section A.

Theorem 2.8. *Consider a linear operator $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$. If $\mathcal{G}1 = 0$ and \mathcal{G} satisfies the positive maximum principle, then \mathcal{G} is a Lévy type operator.*

Suppose $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ is a linear operator with $\mathcal{G}1 = 0$ that satisfies the positive maximum principle, and let X be a solution to the associated martingale problem. Then X is a semimartingale, as can be seen by taking $f(x) = x_i$ in (2.1). We claim that its diffusion, drift, and jump characteristics (with the identity map as truncation function) are given by

$$\int_0^t a(X_s)ds, \quad \int_0^t b(X_s)ds, \quad \nu(X_{t-}, d\xi)dt,$$

where (a, b, ν) is the triplet of the Lévy type representation (2.3). To see this, first note that \mathcal{G} can be extended to C^2 functions on E using (2.3). Then, an approximation argument shows that N^f in (2.1) remains a martingale for such functions f . The claimed form of the characteristics of X now follows from Theorem II.2.42 in Jacod and Shiryaev (2003); see also Proposition 2.12 in Cuchiero et al. (2012). This justifies referring to X as a polynomial jump-diffusion. Since the martingale problem is well-posed by Theorem 2.3, such a polynomial jump-diffusion is a Markov process, and hence a polynomial process in the sense of Cuchiero et al. (2012).

The following lemma provides necessary and sufficient conditions on the triplet (a, b, ν) in order that \mathcal{G} be polynomial.

Lemma 2.9. *Let $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ be a Lévy type operator with triplet (a, b, ν) . Then \mathcal{G} is polynomial if and only if*

$$b_i \in \text{Pol}_1(E), \quad a_{ij} + \int \xi_i \xi_j \nu(\cdot, d\xi) \in \text{Pol}_2(E), \quad \int \xi^{\mathbf{k}} \nu(\cdot, d\xi) \in \text{Pol}_{|\mathbf{k}|}(E)$$

for all $i, j \in \{1, \dots, d\}$ and $|\mathbf{k}| \geq 3$.

Proof. This result is well-known, see for instance Cuchiero et al. (2012), and the proof is simple. Indeed, direct computation yields $0 = \mathcal{G}(1)(x)$, $b_i(x) = \mathcal{G}(e_i^\top(\cdot - x))(x)$, $a_{ij}(x) + \int \xi_i \xi_j \nu(x, d\xi) = \mathcal{G}(e_i^\top(\cdot - x)e_j^\top(\cdot - x))(x)$, and $\int \xi^{\mathbf{k}} \nu(x, d\xi) = \mathcal{G}((\cdot - x)^{\mathbf{k}})(x)$ for $|\mathbf{k}| \geq 3$. Thus, if \mathcal{G} is polynomial, one can show that the triplet satisfies the stated conditions. The converse implication is immediate from the observation that $\deg(pq) \leq \deg(p) + \deg(q)$ for any $p, q \in \text{Pol}(E)$. \square

2.3 Conic combinations of polynomial operators

Due to Theorem 2.3 and Theorem 2.8, every member of the set

$$\mathcal{K} := \{\mathcal{G} : \text{Pol}(E) \rightarrow C(E) : \mathcal{G} \text{ is polynomial} \\ \text{and its martingale problem is well-posed}\}$$

is of Lévy type (2.3). The set \mathcal{K} also possesses the following stability properties, which are useful for constructing examples of polynomial jump-diffusions; we do this in Section 5. The proofs of the following two results are given in Section B.

Theorem 2.10. *The set \mathcal{K} is a convex cone closed under pointwise convergence, in the sense that if $\mathcal{G}_n \in \mathcal{K}$ for $n \in \mathbb{N}$ and $\mathcal{G}f(x) := \lim_{n \rightarrow \infty} \mathcal{G}_n f(x)$ exists and is finite for all $f \in \text{Pol}(E)$ and $x \in E$, then $\mathcal{G} \in \mathcal{K}$.*

If an operator \mathcal{G} is the limit of \mathcal{G}_n as in Theorem 2.10, then its triplet (a, b, ν) can be expressed in terms of the triplets (a^n, b^n, ν^n) of the operators \mathcal{G}_n .

Lemma 2.11. *Suppose that $\mathcal{G}_n \in \mathcal{K}$, and let a^n , b^n , and $\nu^n(x, d\xi)$ be the coefficients of its Lévy type representation, for all $n \in \mathbb{N}$. Then*

$$\mathcal{G}f(x) := \lim_{n \rightarrow \infty} \mathcal{G}_n f(x)$$

exists and is finite for all $f \in \text{Pol}(E)$ and $x \in E$ if and only if the coefficients

$$b_i^n, \quad a_{ij}^n + \int \xi_i \xi_j \nu^n(\cdot, d\xi), \quad \int \xi^{\mathbf{k}} \nu^n(\cdot, d\xi)$$

converge pointwise as $n \rightarrow \infty$ for all $i, j \in \{1, \dots, d\}$ and $|\mathbf{k}| \geq 3$. In this case the triplet (a, b, ν) of the Lévy type representation of \mathcal{G} is given by

$$b_i(x) = \lim_{n \rightarrow \infty} b_i^n(x), \quad a_{ij}(x) = \lim_{n \rightarrow \infty} \left(a_{ij}^n(x) + \int \xi_i \xi_j \nu^n(x, d\xi) \right) - \int \xi_i \xi_j \nu(x, d\xi),$$

for all $x \in E$ and $i, j \in \{1, \dots, d\}$, where the kernel $\nu(x, d\xi)$ is uniquely determined by

$$\int \xi^{\mathbf{k}} \nu(x, d\xi) = \lim_{n \rightarrow \infty} \int \xi^{\mathbf{k}} \nu^n(x, d\xi), \quad |\mathbf{k}| \geq 3.$$

Remark 2.12. *The diffusion coefficient $a(x)$ is the limit of $a^n(x)$ if and only if the weak limit of $|\xi|^2 \nu^n(x, d\xi)$ exists and has no mass in zero. If the weak limit does have mass in zero, then this mass is equal to the difference between $\text{Tr}(a(x))$ and the limit of $\text{Tr}(a^n(x))$.*

3 Affine and polynomial jump sizes

Throughout this section we continue to consider a compact state space $E \subseteq \mathbb{R}^d$. In the absence of jumps it is relatively straightforward to explicitly write down a complete parametrization of polynomial diffusions on the unit interval or the

unit simplex; see Filipović and Larsson (2016). With jumps this is no longer the case. Indeed, examples in Section 5 illustrate the diversity of behavior that is possible even on the simplest nontrivial state space $[0, 1]$. Therefore, in order to make progress we will restrict attention to specifications whose jumps are of the following state-dependent type. Consider a jump kernel $\nu(x, d\xi)$ from E into \mathbb{R}^d satisfying (2.4).

Definition 3.1. *The jump kernel $\nu(x, d\xi)$ is said to have affine jump sizes if it is of the form*

$$\nu(x, A) = \lambda(x) \int \mathbb{1}_{A \setminus \{0\}}(\gamma(x, y)) \mu(dy) \quad (3.1)$$

where $\lambda : E \rightarrow \mathbb{R}_+$ is a nonnegative measurable function, $\gamma = (\gamma_1, \dots, \gamma_d)$ is of the affine form

$$\gamma_i(x, y) = y_i^0 + y_i^1 x_1 + \dots + y_i^d x_d, \quad (3.2)$$

and $\mu(dy)$ is a measure on $\mathbb{R}^{d(d+1)}$ satisfying $\int (|y|^2 \wedge 1) \mu(dy) < \infty$. Here we use the notation $y = (y_i^j : i = 1, \dots, d, j = 0, \dots, d) \in \mathbb{R}^{d(d+1)}$ for the vector of coefficients appearing in (3.2).

Remark 3.2. *By (2.4) and compactness of E , the measure $\mu(dy)$ can always be chosen compactly supported. In this case, all its moments of order at least two are finite.*

Intuitively, (3.1) means that the conditional distribution of the jump ΔX_t , given that it is nonzero and the location immediately before the jump is $X_{t-} = x$, is the same as the distribution of $\gamma(x, y)$ under $\mu(dy)$; at least when $\mu(dy)$ is a probability measure. The jump intensity is state-dependent and given by $\nu(x, \mathbb{R}^d) = \lambda(x) \mu(\{\gamma(x, \cdot) \neq 0\})$, which may or may not be finite.

Jump kernels with affine jump sizes can be used as building blocks to obtain a large class of specifications by means of Theorem 2.10. The jump kernels obtained in this way are of the form

$$\nu(x, d\xi) = \sum_k \nu_k(x, d\xi),$$

where each jump kernel $\nu_k(x, d\xi)$ has affine jump sizes. We refer to such specifications as having *mixed affine jump sizes*.

The affine form of $\gamma(x, y)$ is a particular case of the seemingly more general situation where $\gamma(x, y)$ is allowed to depend polynomially on the current state x . However, this would not actually lead to an increase in generality in the context of polynomial jump–diffusions. Indeed, at least in the case when E has nonempty relative interior in its affine hull, the following result shows that whenever jump sizes are polynomial, they are necessarily affine. The proof is given in Section C.

Theorem 3.3. *Assume that E has nonempty relative interior in its affine hull. Let $\nu(x, d\xi)$ be a jump kernel from E into \mathbb{R}^d of the form (3.1) and satisfying (2.4), where λ is nonnegative and measurable, γ is given by*

$$\gamma_i(x, y) = \sum_{|\mathbf{k}| \leq K} y_{\mathbf{k}}^i x^{\mathbf{k}}$$

for some $K \in \mathbb{N}_0$, and $\mu(dy)$ is a measure on $(\mathbb{R}^d)^{\dim \text{Pol}_K(\mathbb{R}^d)}$ with $\int (|y|^2 \wedge 1)\mu(dy) < \infty$. Assume also that $\nu(x, d\xi)$ satisfies

$$\int \xi^{\mathbf{k}} \nu(\cdot, d\xi) \in \text{Pol}_{|\mathbf{k}|}(E), \quad |\mathbf{k}| \geq 3, \quad (3.3)$$

and that E has nonempty interior. Then one can choose $\mu(dy)$ so that $y_{\mathbf{k}}^i = 0$ a.e. for all $i = 1, \dots, d$ and all $|\mathbf{k}| \geq 2$. That is, $\nu(x, d\xi)$ has affine jump sizes.

Remark 3.4. Note that if $\nu(x, d\xi)$ has affine jump sizes and satisfies (3.3), then the function λ can be expressed as the ratio of two polynomials of degree at most four,

$$\lambda(x) = \frac{\int |\xi|^4 \nu(x, d\xi)}{\int |\gamma(x, y)|^4 \mu(dy)},$$

at points x where the denominator is nonzero. At points x where the denominator vanishes, we have $\gamma(x, y) = 0$ for μ -a.e. y , whence $\nu(x, d\xi) = 0$ due to (3.1). Thus we may always take $\lambda(x) = 0$ at such points.

Remark 3.5. Jump specifications of the form (3.1) are convenient from the point of view of representing solutions X to the martingale problem for \mathcal{G} as solutions to stochastic differential equations driven by a Brownian motion and a Poisson random measure. Indeed, such a stochastic differential equation has the following form:

$$\begin{aligned} X_t = X_0 &+ \int_0^t b(X_s) ds + \int_0^t \sqrt{a(X_s)} dW_s \\ &+ \int_0^t \int_0^{\lambda(X_{s-})} \int \gamma(X_{s-}, y) (N(ds, du, dy) - ds du \mu(dy)), \end{aligned}$$

where $\sqrt{\cdot}$ denotes the matrix square root, W is a d -dimensional Brownian motion and $N(ds, du, dy)$ is a Poisson random measure on $\mathbb{R}_+^2 \times \text{supp}(\mu)$ whose intensity measure is $ds du \mu(dy)$. See also, for instance, [Dawson and Li \(2006, Section 5\)](#), regarding analogous representations of affine processes. Note that a representation of the form (3.1) always exists, even with $\lambda \equiv 1$, if one allows y to lie a suitable Blackwell space; see [Jacod and Shiryaev \(2003, Remark III.2.28\)](#). Thus, in view of [Theorem 3.3](#), our restriction to affine jump sizes in the sense of [Definition 3.1](#) is essentially equivalent to a polynomial dependence of $\gamma(x, y)$ on x , somewhat generalized by allowing a state dependent intensity $\lambda(x)$. Note also that once $\gamma(x, y)$ depends polynomially on x , there is no loss of generality to assume that y lies in an Euclidean space.

4 The unit interval

Throughout this section we consider the state space

$$E := [0, 1].$$

Our goal is to characterize all polynomial jump–diffusions on E with affine jump sizes. The general existence and uniqueness result Theorem 2.3, in conjunction with Lemma 2.9, leads to the following refinement of Theorem 2.8, characterizing those triplets (a, b, ν) that correspond to polynomial jump–diffusions. The proof is given in Section D.

Lemma 4.1. *A linear operator $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ is polynomial and its martingale problem is well–posed if and only if it is of form (2.3) and the corresponding triplet (a, b, ν) satisfies*

- (i) $a \geq 0$ and $\nu(x, d\xi)$ satisfies (2.4),
- (ii) $a(0) = a(1) = 0$, $b(0) - \int \xi \nu(0, d\xi) \geq 0$, and $b(1) - \int \xi \nu(1, d\xi) \leq 0$,
- (iii) $b \in \text{Pol}_1(E)$, $a + \int \xi^2 \nu(\cdot, d\xi) \in \text{Pol}_2(E)$, and $\int \xi^n \nu(\cdot, d\xi) \in \text{Pol}_n(E)$ for all $n \geq 3$.

Observe that condition (i) guarantees that \mathcal{G} is of Lévy Type.

Remark 4.2. *Condition (ii) implies that $\int |\xi| \nu(x, d\xi) < \infty$ for $x \in \{0, 1\}$. Intuitively, this means that the solution to the martingale problem for \mathcal{G} has a purely discontinuous martingale part which is necessarily of finite variation on the boundary of E .*

We now turn to the setting of affine jump sizes in the sense of Definition 3.1. We thus consider Lévy type operators \mathcal{G} of the form

$$\begin{aligned} \mathcal{G}f(x) &= \frac{1}{2}a(x)f''(x) + b(x)f'(x) \\ &\quad + \lambda(x) \int (f(x + \gamma(x, y)) - f(x) - \gamma(x, y)f'(x)) \mu(dy), \end{aligned} \tag{4.1}$$

where λ is nonnegative and measurable, and $\gamma(x, y)$ is affine in x . The main result of this section, Theorem 4.3 below, shows that the generator must be of one of five mutually exclusive types, which we now describe.

Type 0. Let $a(x) = Ax(1-x)$, $b(x) = \kappa(\theta-x)$, where $A \in \mathbb{R}_+$, $\kappa \in \mathbb{R}_+$, and $\theta \in [0, 1]$, and set $\lambda = 0$. Then \mathcal{G} is a polynomial operator whose martingale problem is well–posed. The solution X corresponds simply to the well–known Jacobi diffusion, which is the most general polynomial diffusion on the unit interval.

Type 1. Let $a(x) = Ax(1-x)$, $b(x) = \kappa(\theta-x)$, and $\lambda(x) = 1$, where $A \in \mathbb{R}_+$, $\kappa \in \mathbb{R}_+$, and $\theta \in [0, 1]$. Furthermore, writing $y = (y_1, y_2)$ we define $\gamma(x, y) = y_1(-x) + y_2(1-x)$ and let μ be a nonzero measure on $[0, 1]^2 \setminus \{0\}$. If the boundary conditions

$$\kappa\theta \geq \int y_2 \mu(dy) \quad \text{and} \quad \kappa(1-\theta) \geq \int y_1 \mu(dy)$$

are satisfied, then \mathcal{G} is a polynomial operator whose martingale problem is well–posed.

Note that the boundary conditions imply that $\int |\xi| \nu(x, d\xi) \leq 2 \int |y| \mu(dy)$ is bounded. Thus, the resulting process behaves like a Jacobi diffusion with summable jumps. The arrival intensity of the jumps is $\nu(x, E - x) = \mu(\{y : \gamma(x, y) \neq 0\})$, which may or may not be finite. Figure 3 illustrates the form of a , λ and the support $\gamma(x, y)$ under μ .

Type 2. Let $a(x) = Ax(1 - x)$, $b(x) = \kappa(\theta - x)$, and $\lambda(x) = \frac{1}{x}(1 + qx)\mathbb{1}_{\{x \neq 0\}}$ where $A \in \mathbb{R}_+$, $\kappa \in \mathbb{R}_+$, $\theta \in [0, 1]$, and $q \in [-1, \infty)$. Furthermore, define $\gamma(x, y) = -xy$ and let μ be a nonzero square-integrable measure on $(0, 1]$. Notice that y is scalar. If the boundary condition

$$\kappa(1 - \theta) \geq (1 + q) \int y \mu(dy)$$

is satisfied, then \mathcal{G} is a polynomial operator whose martingale problem is well-posed.

The boundary condition implies, if $q > -1$, that

$$\int |\xi| \nu(x, d\xi) \leq (1 + |q|) \int y \mu(dy)$$

is bounded. Thus, in this case, the solution X to the martingale problem for \mathcal{G} has summable jumps. If $q = -1$, the jumps need not be summable. The arrival intensity of the jumps is $\nu(x, E - x) = \lambda(x)\mu((0, 1])$ and hence, even if μ is a finite measure, the jump intensity is unbounded around $x = 0$. Moreover, due to the form of $\gamma(x, y)$, X can only jump to the left, and since $\nu(0, E) = 0$, X cannot leave $x = 0$ by means of a jump. Figure 4 illustrates the form of a , λ and the support $\gamma(x, y)$ under μ .

By reflecting the state space around the point $1/2$, we obtain a similar structure which we also classify as Type 2, where now the jump intensity is unbounded around $x = 1$. The diffusion and drift coefficients remain as before, while $\lambda(x) = \frac{1}{1-x}(1 + q(1-x))\mathbb{1}_{\{x \neq 1\}}$ for some $q \in [-1, \infty)$, the jump sizes are $\gamma(x, y) = (1-x)y$, and μ is a nonzero square-integrable measure on $(0, 1]$ as before. The boundary condition becomes $\kappa\theta \geq (1 + q) \int y \mu(dy)$.

Type 3. Let $x^* \in (0, 1)$; this will be a “no-jump” point. Let $b(x) = \kappa(\theta - x)$ and set

$$\lambda(x) = \frac{q_0 + q_1x + q_2x^2}{(x - x^*)^2} \mathbb{1}_{\{x \neq x^*\}},$$

where $\kappa \in \mathbb{R}_+$, $\theta \in [0, 1]$, and q_0, q_1, q_2 are real numbers such that the numerator of λ is nonnegative on E without zeros at x^* . Furthermore, define $\gamma(x, y) = -(x - x^*)y$, and let μ be a nonzero square-integrable measure on $(0, (x^* \vee (1 - x^*))^{-1}]$. Finally, let $a(x) = Ax(1 - x) + a^\nu \mathbb{1}_{\{x=x^*\}}$ where

$$a^\nu = (q_0 + q_1x^* + q_2(x^*)^2) \int y^2 \mu(dy).$$

If the boundary conditions

$$\kappa\theta \geq \frac{q_0}{x^*} \int y\mu(dy) \quad \text{and} \quad \kappa(1-\theta) \geq \frac{q_0 + q_1 + q_2}{1-x^*} \int y\mu(dy)$$

are satisfied, then \mathcal{G} is a polynomial operator whose martingale problem is well-posed.

If $q_0 + q_1x + q_2x^2 = Lx(1-x)$ for some constant $L \in \mathbb{R}_+$, the solution X to the martingale problem for \mathcal{G} may have non-summable jumps. If the numerator of $\lambda(x)$ is not of this form, then the boundary conditions imply that X has summable jumps. The arrival intensity of the jumps is

$$\nu(x, E-x) = \lambda(x)\mu\left(\left(0, \frac{1}{x^*} \wedge \frac{1}{1-x^*}\right]\right).$$

As a result, even if μ is a finite measure, the jump intensity has a pole of order two at $x = x^*$, which results in a contribution of size a^ν to the diffusion coefficient. Moreover, due to the form of $\gamma(x, y)$, the jumps of X are always in the direction of the “no-jump” point x^* . Although the jumps may overshoot x^* , they always serve to reduce the distance to x^* . In particular, since $\nu(x^*, E-x^*) = 0$, X cannot leave $x = x^*$ by means of a jump. Figure 5 illustrates the form of a , λ and the support $\gamma(x, y)$ under μ .

Type 4. Suppose $\alpha \in \mathbb{C} \setminus \mathbb{R}$ is a non-real complex number such that $|2\alpha - 1| < 1$ and let μ be a nonzero square-integrable measure on $[0, 1] \times [0, 1]$ such that

$$\int (y_1(-\alpha) + y_2(1-\alpha))^n \mu(dy) = 0, \quad n \geq 2, \quad (4.2)$$

and $\int y_1\mu(dy) = \int y_2\mu(dy) = \infty$. Let $b(x) = \kappa(\theta - x)$ and set

$$\lambda(x) = \frac{Lx(1-x)}{(x-\alpha)(x-\bar{\alpha})},$$

where $\kappa \in \mathbb{R}_+$, $\theta \in [0, 1]$, and $L > 0$. Furthermore, define $\gamma(x, y) = y_1(-x) + y_2(1-x)$ and let $a(x) = Ax(1-x)$ for some $A \in \mathbb{R}_+$. Then \mathcal{G} is a polynomial operator whose martingale problem is well-posed.

Having described five types of processes which are polynomial jump-diffusions due to the conditions of Lemma 4.1, we are now ready to state the converse result, namely that *all* polynomial jump-diffusions on $[0, 1]$ with affine jump sizes are necessarily of one of these types. The proof is given in Section D.

Theorem 4.3. *Let \mathcal{G} be a polynomial operator whose martingale problem is well-posed. If the associated jump kernel has affine jump sizes, then \mathcal{G} necessarily belongs to one of the Types 0-4.*

Remark 4.4. *Let us end this section with some remarks regarding Type 4. First, note that $\int y_1\mu(dy) = \int y_2\mu(dy) = \infty$ implies that $\mu(dy)$ cannot be a product*

measure since in this case $\int y_1 y_2 \mu(dy)$ would be infinite too, which however contradicts square integrability. Second, after passing to polar coordinates (r, φ) , condition (4.2) becomes

$$\int_{[0, \pi] \times \mathbb{R}_+} r^n e^{in\varphi} \mu_\alpha(d\varphi, dr) = 0, \quad n \geq 2, \quad (4.3)$$

where μ_α is the compactly supported measure given by

$$\mu_\alpha(A) := \int \mathbf{1}_A \left(\text{Arg}(y_1(-\alpha) + y_2(1-\alpha)), |y_1(-\alpha) + y_2(1-\alpha)| \right) \mu(dy)$$

for all measurable subsets $A \subseteq [0, \pi] \times \mathbb{R}_+$. It can then be shown that r and φ cannot be independent, i.e., μ_α cannot be a product measure. These observations indicate that natural attempts to find combinations of α and μ satisfying (4.2) do not work. In fact, it is unknown to us what a potential example of Type 4 might look like. Note also that Type 4 is distinct from all other types in the following respect. For Types 1–3, $\lambda \gamma^n(\cdot, y)$ is a polynomial on E (outside the “no-jump” point) of degree $n \geq 2$ for all $y \in \text{supp}(\mu)$, whereas for Type 4 this property holds true only for the integrated quantity $\lambda \int \gamma(\cdot, y)^n \mu(dy)$.

5 Examples of polynomial operators on the unit interval

In this section we present a number of examples that illustrate the diverse behavior of polynomial jump-diffusions on $[0, 1]$. While the diffusion case is simple – the Jacobi diffusions (Type 0) are the only possibilities – the complexity increases significantly in the presence of jumps. For instance, in Example 5.5 we obtain jump intensities with a countable number of poles in the state space.

5.1 Examples with affine jump sizes

Example 5.1. We start with a well-known example of a polynomial jump-diffusion on $[0, 1]$; see [Cuchiero et al. \(2012, Example 3.5\)](#). Consider the Jacobi process, which is the solution of the stochastic differential equation

$$dX_t = \kappa_0(\theta_0 - X_t)dt + \sigma \sqrt{X_t(1 - X_t)}dW_t, \quad X_0 = x_0 \in [0, 1],$$

where $\theta_0 \in [0, 1]$ and $\kappa_0, \sigma > 0$. This process can also be regarded as the unique solution to the martingale problem for $(\mathcal{G}, \delta_{x_0})$, with the Type 0 operator

$$\mathcal{G}f(x) := \frac{1}{2}\sigma^2 x(1-x)f''(x) + \kappa_0(\theta_0 - x)f'(x).$$

This example can be extended by adding jumps, where the jump times correspond to those of a Poisson process with intensity λ and the jump size is a function of

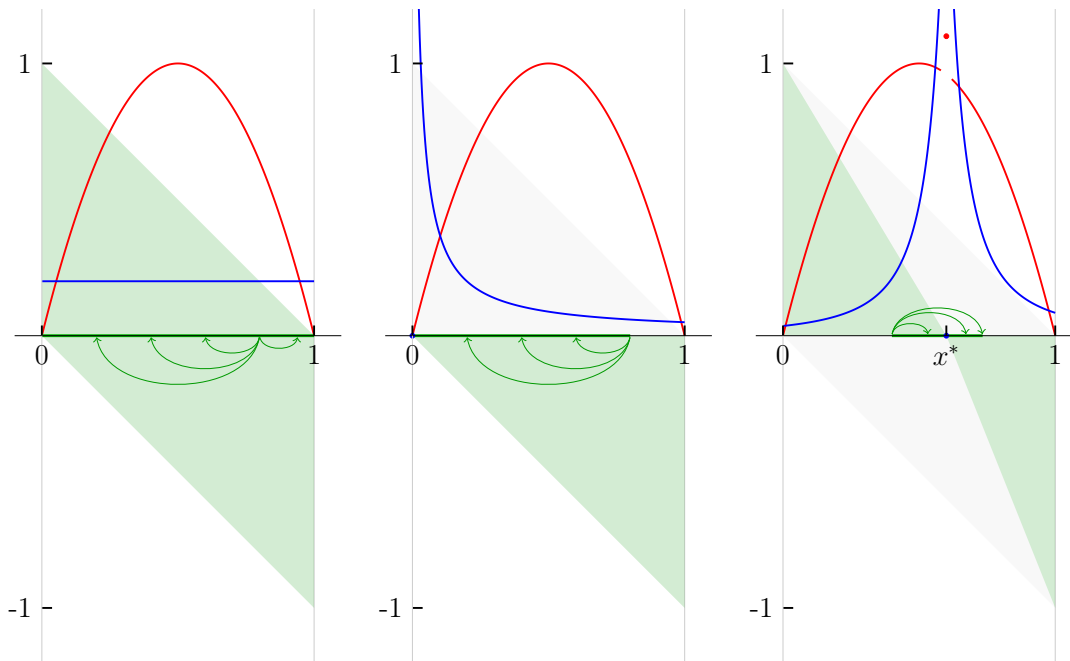


Figure 3

Figure 4

Figure 5

Figure 3: A representation of Type 1, where $\lambda(x) = 1$ (in blue), a is a polynomial of second degree vanishing on the boundaries (in red), and the support of $\nu(x, \cdot)$ is contained in $[-x, 1 - x]$ (in green).

Figure 4: A representation of Type 2, where λ has a pole of order 1 in $x = 0$ (in blue), a is a polynomial of second degree vanishing on the boundaries (in red), and the support of $\nu(x, \cdot)$ is contained in $[-x, 0]$ (in green) for all $x \in E$. This in particular implies that the distance to the “no-jump” point always decreases if a jump occurs. Note that in $x = 0$ there is no jump activity since $\lambda(0) = 0$ and thus $\nu(0, E) = 0$.

Figure 5: A representation of Type 3, where λ has a pole of order 2 in $x^* \in (0, 1)$ (in blue), a is a polynomial of second degree on $E \setminus \{x^*\}$ vanishing on the boundaries (in red), and the support of $\nu(x, \cdot)$ is contained in $[-2(x - x^*), 0]$, resp. $[0, -2(x - x^*)]$, (in green) for all $x \in E$. This in particular implies that the distance to the “no-jump” point x^* always decreases if a jump occurs. Note that in x^* there is no jump activity since $\lambda(x^*) = 0$, but there is an extra contribution to the diffusion coefficient at this point.

the process level. One can for instance specify that if a jump occurs, then the process is reflected in $1/2$. In this case the process would be the unique solution to the martingale problem for $(\mathcal{G}, \delta_{x_0})$, where

$$\mathcal{G}f(x) := \frac{1}{2}\sigma^2x(1-x)f''(x) + \kappa_0(\theta_0 - x)f'(x) + \lambda(f(1-x) - f(x)),$$

which is an operator of Type 1 with $A = \sigma^2$, $\kappa = \kappa_0 + 2\lambda$, $\theta = \frac{\kappa_0\theta_0 + \lambda}{\kappa_0 + 2\lambda}$, and $\mu = \lambda\delta_{(1,1)}$.

Example 5.2. *The following example features a simple state-dependent jump distribution. Consider a Lévy type operator \mathcal{G} whose jump kernel $\nu(x, d\xi)$ is chosen such that $x + \xi$ is uniformly distributed on $(\alpha(x), \beta(x))$, where $\alpha, \beta \in \text{Pol}_1(E)$ and $0 \leq \alpha(x) \leq \beta(x) \leq 1$ for all $x \in E$. This in particular implies that α and β can be written as*

$$\alpha(x) = \alpha_0(1 - x) + \alpha_1x \quad \text{and} \quad \beta(x) = \beta_0(1 - x) + \beta_1x$$

for some $0 \leq \alpha_0 \leq \beta_0 \leq 1$ and $0 \leq \alpha_1 \leq \beta_1 \leq 1$. Choosing the drift coefficient b suitably, the operator \mathcal{G} is then of Type 1 for μ being the pushforward of the uniform distribution on $(0, 1)$ under the map $z \mapsto (1 - z(\beta_1 - \alpha_1) - \alpha_1, z(\beta_0 - \alpha_0) + \alpha_0)$.

The solution to the corresponding martingale problem is a Jacobi process extended by adding jumps, where the jump times correspond to those of a Poisson process with unit intensity, and the jump's target point is uniformly distributed on $(\alpha(x), \beta(x))$, given that the process is located at x immediately before the jump.

Example 5.3. *Polynomial operators are not always easy to recognize at first sight. Consider a Lévy type operator \mathcal{G} whose diffusion and drift coefficients a and b are zero, and whose jump kernel $\nu(x, d\xi)$ is given by*

$$\nu(x, A) = \mathbb{1}_{\{x \neq 0\}} \frac{1 - x}{x} \int_0^1 \mathbb{1}_{A \setminus \{0\}}(-x \sin^2((x + z)\pi)) dz.$$

Despite the presence of the sine function, the operator \mathcal{G} satisfies all the conditions of Lemma 4.1. It is thus polynomial and its martingale problem is well-posed. In fact, this operator is of Type 2. Using the periodicity of the sine function, one can show that $\nu(x, d\xi)$ has affine jump sizes with $\lambda(x) = \frac{1-x}{x} \mathbb{1}_{\{x \neq 0\}}$, $\gamma(x, y) = -xy$, and μ being the pushforward of Lebesgue measure on $[0, 1]$ under the map $z \mapsto \sin^2(z\pi)$. The associated polynomial jump-diffusion is a martingale since $b = 0$. Moreover, the arrival intensity $\nu(x, E - x)$ of the jumps is given by $\frac{1-x}{x} \mathbb{1}_{\{x \neq 0\}}$, which is unbounded around zero.

Example 5.4. *The Dunkl process with parameter $n \in \mathbb{N}_0$ is a polynomial jump-diffusion on \mathbb{R} , see e.g. Cuchiero et al. (2012, Example 3.7), and can be characterized as the unique martingale whose absolute value is the Bessel process of dimension $1 + 2n$; see Gallardo and Yor (2006). The corresponding polynomial operator $\mathcal{G}^{\text{Dunkl}}$ is of Lévy type with diffusion and jump coefficients $a(x) = 2 + 2n\mathbb{1}_{\{x=0\}}$ and $b(x) = 0$, and jump kernel*

$$\nu(x, d\xi) = \mathbb{1}_{\{x \neq 0\}} \frac{n}{2x^2} \delta_{-2x}(d\xi).$$

The arrival intensity of its jumps is thus given by $\nu(x, \mathbb{R}) = \frac{n}{2x^2} \mathbb{1}_{\{x \neq 0\}}$, which is a rational function with a pole of second order in $x = 0$.

Observe that $\nu(x, d\xi)$ exhibits several similarities with jump kernels of operators of Type 3, such as the form of the arrival intensity of the jumps, and the extra contribution to the diffusion coefficient at the “no-jump” point $x = 0$. In fact, defining $\tilde{f} := f(\cdot + \frac{1}{2})$ and

$$\mathcal{G}f(x) = x(1-x)\mathcal{G}^{\text{Dunkl}}\tilde{f}(x-1/2),$$

we obtain a polynomial operator of Type 3 with “no-jump” point $x^* = 1/2$.

5.2 Constructions using conic combinations

We provide two examples illustrating the usefulness of Theorem 2.10 for combining operators with affine jump sizes to achieve specifications with interesting properties.

Example 5.5. We now construct a polynomial operator whose martingale problem is well-posed, such that the arrival intensity of the jumps is unbounded around infinitely many points, but finite for all $x \neq 1/2$.

Let \mathcal{G}_n , $n \geq 3$, be operators of Type 3 with “no-jump” points $x_n^* = \frac{1}{2} + \frac{1}{n}$. Let their diffusion coefficients be given by

$$a_n(x) = \frac{1}{3n^2}x_n^*(1-x_n^*)\mathbb{1}_{\{x=x_n^*\}},$$

the drift coefficients be 0, and the parameters of the jump kernels $\nu_n(x, d\xi)$ be given by

$$\lambda_n(x) = n^{-2}\frac{x(1-x)}{(x-x_n^*)^2}\mathbb{1}_{\{x \neq x_n^*\}}, \quad \gamma_n(x, y) = -y(x-x_n^*),$$

and μ be Lebesgue measure on $[0, 1]$. Note that for all $k \geq 2$ we have

$$\sum_{n=3}^{\infty} \left(a_n(x)\delta_{k2} + \int \xi^k \nu_n(x, d\xi) \right) = \frac{x(1-x)}{k+1} \sum_{n=3}^{\infty} n^{-2}(x_n^* - x)^{k-2} < \infty. \quad (5.1)$$

By Theorem 2.10 and Lemma 2.11 this implies that the operator $\mathcal{G} := \sum_{n=3}^{\infty} \mathcal{G}_n$ is again polynomial and its martingale problem is well-posed. In particular, \mathcal{G} is a Lévy type operator with coefficients $a(x) = \sum_{n=3}^{\infty} a_n(x)$ and $b(x) = 0$, and jump kernel $\nu(x, \cdot) := \sum_{n=3}^{\infty} \nu_n(x, \cdot)$. As a result, the arrival intensity of the jumps is given by

$$\nu(x, E-x) = \sum_{n=3}^{\infty} \lambda_n(x) = x(1-x) \sum_{n=3}^{\infty} \frac{1}{n^2(x-x_n^*)^2} \mathbb{1}_{\{x \neq x_n^*\}},$$

which is unbounded around each x_n^* but finite for all $x \neq 1/2$. At $x = 1/2$ the jump intensity is infinite. Figure 6 contains an illustration.

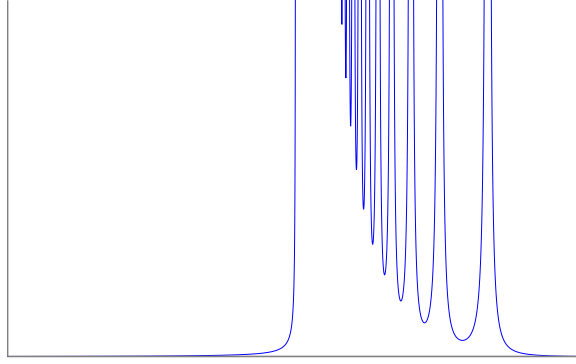


Figure 6: A graphical representation of arrival intensity of the jumps $\nu(x, E - x)$ appearing in Example 5.5.

Example 5.6. *This example shows that the operator of a polynomial diffusion, or equivalently an operator of Type 0, can always be written as the limit of “pure jump” polynomial operators, i.e. with zero diffusion coefficients. Consider the Jacobi diffusion with operator \mathcal{G} given by*

$$\mathcal{G}f(x) := Ax(1-x)f''(x) + \kappa(\theta-x)f'(x),$$

for some $A \in \mathbb{R}_+$, $\kappa \in \mathbb{R}_+$, and $\theta \in [0, 1]$. Let then \mathcal{G}_n be an operator of Type 2 and suppose that its diffusion coefficient a_n is zero, the drift coefficient is given by $b_n(x) = \kappa(\theta - x)$, and the parameters of the jump kernel $\nu_n(x, d\xi)$ are

$$\lambda_n(x) = n^2 \frac{A(1-x)}{x} \mathbb{1}_{\{x \neq 0\}}, \quad \gamma_n(x, y) = -yx, \quad \mu = \delta_{1/n}.$$

Observe that, trivially, we have $\lim_{n \rightarrow \infty} b_n(x) = \kappa(\theta - x)$. Also,

$$\lim_{n \rightarrow \infty} \left(a_n(x) + \int \xi^2 \nu_n(x, d\xi) \right) = Ax(1-x) \quad \text{and} \quad \lim_{n \rightarrow \infty} \int \xi^k \nu_n(x, d\xi) = 0, \quad k \geq 3.$$

By Lemma 2.11 we thus conclude that $\mathcal{G} = \lim_{n \rightarrow \infty} \mathcal{G}_n$ in sense of Theorem 2.10.

5.3 Mixed affine jump sizes

Consider now a Lévy type polynomial operator \mathcal{G} whose jump kernel has mixed affine jump sizes in the sense of Section 3, i.e.,

$$\nu(x, d\xi) = \sum_{\ell=1}^L \nu_\ell(x, d\xi), \tag{5.2}$$

where each kernel $\nu_\ell(x, d\xi)$ has affine jump sizes. Suppose the martingale problem for \mathcal{G} is well-posed, or equivalently, its triplet satisfies the conditions of Lemma 4.1. A natural question is now whether the individual kernels $\nu_\ell(x, d\xi)$ also satisfy the conditions of Lemma 4.1. If this were to be true, it would have

the pleasant consequence that \mathcal{G} could be represented as a sum of operators of Types 0–4. Unfortunately this is not the case, which we illustrate in Example 5.7 below. In fact, there exist kernels of the form (5.2) that cannot even be obtained as an infinite conic combination of the kernels appearing in Types 0–4.

Example 5.7. Consider a Lévy type operator \mathcal{G} , whose coefficients are given by $a(x) = 0$, $b(x) = 1 - 2x$, and whose jump kernel is given by (5.2) for $L = 2$, where $\nu_1(x, d\xi)$ and $\nu_2(x, d\xi)$ have affine jump sizes with parameters $\lambda_1(x) = \frac{1}{x(x+1)} \mathbb{1}_{\{x \neq 0\}}$, $\mu_1 = \delta_{(1,0)}$, and $\lambda_2(x) = \frac{2}{(1-x)(x+1)} \mathbb{1}_{\{x \neq 1\}}$, $\mu_2 = \delta_{(0,1/2)}$. Observe that

$$\gamma(x, y) = -x \quad \mu_1\text{-a.s.} \quad \text{and} \quad \gamma(x, y) = \frac{1}{2}(1-x) \quad \mu_2\text{-a.s.}$$

One can verify that \mathcal{G} satisfies all the conditions of Lemma 4.1, and is thus polynomial and its martingale problem is well-posed.

Assume now for contradiction that $\nu(x, d\xi) = \sum_{\ell=1}^{\infty} \tilde{\nu}_\ell(x, d\xi)$ for some kernels $\tilde{\nu}_\ell(x, d\xi)$ that satisfy the conditions of Lemma 4.1 for some coefficients $a_\ell(x)$ and $b_\ell(x)$. By Theorem 4.3, each $\tilde{\nu}_\ell(x, d\xi)$ then follows one of Types 0–4. Let $\tilde{\lambda}_\ell(x)$ and $\tilde{\mu}_\ell(x)$ be the parameters of the jump kernel $\tilde{\nu}_\ell(x, d\xi)$.

Since $\text{supp } \nu(x, \cdot) \subseteq \{-x, (1-x)/2\}$, we also have $\text{supp } \tilde{\nu}_\ell(x, \cdot) \subseteq \{-x, (1-x)/2\}$ for all $x \in E$, or equivalently,

$$\tilde{\mu}_\ell = \alpha_\ell \delta_{(1,0)} + \beta_\ell \delta_{(0,1/2)} \quad (5.3)$$

for some $\alpha_\ell, \beta_\ell \geq 0$. This already excludes that $\tilde{\nu}_\ell(x, d\xi)$ is of Type 3 or 4, and gives us that for all $x \in (0, 1)$

$$\tilde{\lambda}_\ell(x) = \begin{cases} \frac{q_\ell^\alpha(x)}{x} & \text{if } \beta_\ell = 0, \\ \frac{q_\ell^\beta(x)}{1-x} & \text{if } \alpha_\ell = 0, \\ c_\ell & \text{otherwise,} \end{cases}$$

for some $q_\ell^\alpha, q_\ell^\beta \in \text{Pol}_1(E)$ and $c_\ell \in \mathbb{R}_+$. In particular note that for all $x \in (0, 1)$ and $\ell \in \mathbb{N}$,

$$\alpha_\ell \tilde{\lambda}_\ell(x) = \alpha_\ell \frac{q_\ell^\alpha(x)}{x} \quad \text{and} \quad \beta_\ell \tilde{\lambda}_\ell(x) = \beta_\ell \frac{q_\ell^\beta(x)}{1-x} \quad (5.4)$$

and hence, since $\text{Pol}_1(E)$ is closed under pointwise convergence,

$$\sum_{\ell=1}^{\infty} \int \xi^n \tilde{\nu}_\ell(x, d\xi) = \frac{q^\alpha(x)}{x} (-x)^n + \frac{q^\beta(x)}{1-x} \left(\frac{1-x}{2} \right)^n,$$

for all $n \in \mathbb{N}$ and some $q^\alpha, q^\beta \in \text{Pol}_1(E)$. Since $\int \xi^n \nu(x, d\xi) = \sum_{\ell=1}^{\infty} \int \xi^n \tilde{\nu}_\ell(x, d\xi)$ by assumption, we obtain

$$\frac{-(-x)^{n-1} + ((1-x)/2)^{n-1}}{x+1} = -q^\alpha(x) (-x)^{n-1} + \frac{1}{2} q^\beta(x) \left(\frac{1-x}{2} \right)^{n-1},$$

for all $x \in (0, 1)$, $n \in \mathbb{N}$. The shortest way to see that this condition cannot be satisfied is to use that if two polynomials coincide on $(0, 1)$ they have to coincide on \mathbb{R} , too. But, choosing $x = -1$ we obtain

$$-q^\alpha(-1) + \frac{1}{2}q^\beta(-1) = \frac{n-1}{2}$$

for all $n \in \mathbb{N}$, which is clearly not possible.

Example 5.8. It is possible to show that operators with jump kernels of the form (5.2) can have intensities λ_ℓ with multiple poles of multiple order outside the state space. On the other hand, under some non-degeneracy conditions, they can only have a single pole of order at most 2 inside the state space. We develop this idea in more detail for the case when $\nu(x, \cdot)$ consists of finitely many atoms for all $x \in E$.

Let $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ be an operator of the form described in Lemma 4.1 and suppose that its jump kernel $\nu(x, d\xi)$ is supported on $\{\gamma_1(x), \dots, \gamma_L(x)\}$, where $\gamma_\ell \in \text{Pol}_1(E)$, $\ell = 1, \dots, L$, are pairwise distinct polynomials with $x + \gamma_\ell(x) \in E$ for all $x \in E$. As a result, we have

$$\nu(x, d\xi) = \sum_{\ell=1}^L \lambda_\ell(x) \delta_{\gamma_\ell(x)}(d\xi) \quad (5.5)$$

for some functions $\lambda_\ell : E \rightarrow \mathbb{R}_+$. For $n \geq 2$, set $r_n := \int \xi^n \nu(\cdot, d\xi) = \sum_{\ell=1}^L \lambda_\ell \gamma_\ell^n$, and recall that $r_n \in \text{Pol}_n(E)$ for all $n \geq 3$ and that r_2 is bounded on E . Using the nonnegativity of λ and the boundary conditions for a , one can then establish the following properties, which we state here without proof.

(i) If λ_ℓ has a pole at a point $x_0 \in E$, then $\gamma_\ell(x_0) = 0$. Moreover in this case, analogously to Types 2 and 3, if $x_0 \in \{0, 1\}$, the order of the pole is 1 and if $x_0 \in (0, 1)$, the order of the pole is 2. Note that nonnegativity of λ_ℓ and the fact that $\gamma_\ell \in \text{Pol}_1(E)$ imply that λ_ℓ can have a pole in at most one point of the state space.

(ii) $r_2 \in \text{Pol}_2(E \setminus \{x_1^*, \dots, x_L^*\})$, where x_ℓ^* denotes the zero of γ_ℓ , and we have

$$\lambda_\ell = \frac{q_\ell}{\gamma_\ell^2 \prod_{j \neq \ell} (\gamma_\ell - \gamma_j)} \mathbf{1}_{\{\gamma_\ell \neq 0\}}, \quad (5.6)$$

where $q_\ell \in \text{Pol}_{L+1}(E)$ and on $E \setminus \{x_1^*, \dots, x_L^*\}$ it is given by

$$q_\ell = \sum_{k=0}^{L-1} \left((-1)^k r_{L-k+1} \sum_{\substack{\ell_1 < \dots < \ell_k \\ \ell_1, \dots, \ell_k \neq \ell}} \gamma_{\ell_1} \cdots \gamma_{\ell_k} \right).$$

(iii) Since $a + \int \xi^2 \nu(\cdot, d\xi) \in \text{Pol}_2(E)$ by Lemma 4.1, we can conclude that

$$a(x) = Ax(1-x) + \sum_{\ell=1}^L \left(\frac{(-1)^{L-1} q_\ell(x)}{\prod_{j \neq \ell} \gamma_j(x)} \mathbf{1}_{\{x=x_\ell^*\}} \right), \quad (5.7)$$

for some $A \in \mathbb{R}_+$ and all $x \in E$.

Conversely, fix a sequence of polynomials r_{k+2} , $k = 0, \dots, L-1$, such that $r_{k+2} \in \text{Pol}_{k+2}(E)$ for all k . If for some affine functions $\gamma_1, \dots, \gamma_L$ as above, the functions λ_ℓ given by equation (5.6) satisfy (i) and are all nonnegative on E , one can conclude that for $\nu(x, d\xi)$ as in (5.5), a as in (5.7), and a suitably chosen $b \in \text{Pol}_1(E)$, the corresponding Lévy type operator is polynomial and its martingale problem is well-posed.

Remark 5.9. It is interesting to observe that Shur polynomials appear naturally in the context of Example 5.8. Indeed, by point (ii) we know that each $\lambda_\ell(x)$, and thus every moment $r_n(x)$ of the measures $\nu(x, \cdot)$, is uniquely determined by $\gamma_1, \dots, \gamma_L$ and r_2, \dots, r_{L+1} . More precisely for all $n > L+1$ we can write

$$r_n = \sum_{k=1}^L (-1)^{L-k} s_{\mu_{L,n,k}}(\gamma_1, \dots, \gamma_L) r_{k+1},$$

where $\mu_{L,n,k} = (\mu_{L,n,k}^1, \dots, \mu_{L,n,k}^L)$ is the partition given by

$$\mu_{L,n,k}^1 = n-L-1, \quad \mu_{L,n,k}^2 = \dots = \mu_{L,n,k}^{L-k+1} = 1, \quad \text{and} \quad \mu_{L,n,k}^{L-k+2} = \dots = \mu_{L,n,k}^L = 0,$$

and $s_{\mu_{L,n,k}}$ is the corresponding Shur polynomial.

We now propose two interesting applications of Example 5.8, showing that it can happen that $\lambda_1, \dots, \lambda_L$ have poles of high order and in several points outside the state space.

Example 5.10. Consider a kernel of the form described in (5.5) for

$$\gamma_1(x) = -x, \quad \gamma_2(x) = 1-x, \quad \gamma_3(x) = \frac{1}{3}(1-2x), \quad \gamma_4(x) = \frac{2}{3}(1-2x).$$

Defining λ_ℓ through expression (5.6) where we set $r_2(x) = 1$,

$$r_3(x) = \frac{1-2x}{2}, \quad r_4(x) = \frac{2x^2-2x+5}{18}, \quad r_5(x) = \frac{(2x-1)(5x^2-5x+1)}{6},$$

we obtain

$$\begin{aligned} \lambda_1(x) &= \frac{9}{2} \frac{(1-x)}{x(x+1)(2-x)}, & \lambda_2(x) &= \frac{9}{2} \frac{x}{(1-x)(x+1)(2-x)}, \\ \lambda_3(x) &= \frac{9}{(x+1)(2-x)}, & \lambda_4(x) &= \frac{9}{4} \frac{1}{(x+1)(2-x)}, \end{aligned}$$

for all $x \in E$. Note that the rational functions λ_ℓ satisfy point (i) of Example 5.8 and are all nonnegative on E . As a result, choosing the diffusion and drift coefficients suitably, \mathcal{G} is a polynomial operator whose martingale problem is well-posed. Observe that each λ_ℓ has a pole in $x = -1$ and $x = 2$.

Example 5.11. Consider a kernel of the form described in (5.5) for

$$\gamma_1(x) = -x, \quad \gamma_2(x) = \frac{1}{2}(1-x), \quad \gamma_3(x) = \frac{1}{3}(1-2x).$$

Defining λ_ℓ through expression (5.6) where we set

$$r_2(x) = 1, \quad r_3(x) = \frac{1-2x}{2}, \quad r_4(x) = \frac{10x^2-9x+3}{12},$$

we obtain

$$\lambda_1(x) = \frac{1}{x(x+1)^2}, \quad \lambda_2(x) = \frac{4(2x+1)}{(1-x)(x+1)^2}, \quad \lambda_3(x) = \frac{27x^2}{(1-2x)^2(x+1)^2},$$

for all $x \in E$. Note that the rational functions λ_ℓ satisfy point (i) of Example 5.8 and are all nonnegative on E . As a result, choosing the diffusion and drift coefficients suitably, \mathcal{G} is a polynomial operator whose martingale problem is well-posed. Observe that each λ_ℓ has a pole of second order in $x = -1$.

6 The unit simplex

Throughout this section the state space $E \subseteq \mathbb{R}^d$ is the unit simplex of dimension $d-1$, which we denote by

$$E := \Delta^d = \left\{ x \in \mathbb{R}_+^d : \sum_{i=1}^d x_i = 1 \right\}.$$

Similarly as in Section 4 our goal is to provide a characterization of polynomial jump-diffusions on E with affine jump sizes. Again, we combine Theorem 2.3 and Lemma 2.9 to specialize Theorem 2.8 to the state space E . The proof is given in Section E.

Lemma 6.1. *A linear operator $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ is polynomial and its martingale problem is well-posed if and only if it is of form (2.3) and the corresponding triplet (a, b, ν) satisfies*

- (i) $a(x) \in \mathbb{S}_+^d$ for all $x \in E$ and $\nu(x, d\xi)$ satisfies (2.4),
- (ii) $a_{ii}(x) = 0$ and $b_i(x) - \int \xi_i \nu(x, d\xi) \geq 0$ for all $x \in E \cap \{x_i = 0\}$,
- (iii) $a\mathbf{1} = 0$ and $b^\top \mathbf{1} = 0$,
- (iv) $b_i \in \text{Pol}_1(E)$, $a_{ij} + \int \xi_i \xi_j \nu(\cdot, d\xi) \in \text{Pol}_2(E)$, and $\int \xi^{\mathbf{k}} \nu(\cdot, d\xi) \in \text{Pol}_{|\mathbf{k}|}(E)$ for all $|\mathbf{k}| \geq 3$.

Observe that conditions (i) and (iii) guarantee that \mathcal{G} is of Lévy Type. This in particular ensures that the right-hand side of (2.3) can be computed using an arbitrary representative.

Remark 6.2. *Condition (ii) implies that $\int |\xi_i| \nu(x, d\xi) < \infty$ for all $x \in E \cap \{x_i = 0\}$. Analogously to the unit interval case, this give us some intuition about the behavior of the solution X on the boundary segment $x \in E \cap \{x_i = 0\}$. Indeed, even if the component orthogonal to the boundary of the purely discontinuous martingale part of X is necessarily of finite variation, the other components do not need to satisfy this property. Moreover, since $a(x) \in \mathbb{S}_+^d$, condition (ii) also implies that $a_{ij}(x) = 0$ for all $j \in \{1, \dots, d\}$.*

We now focus on the setting of affine jump sizes in the sense of Definition 3.1. We thus consider Lévy type operators \mathcal{G} of the form

$$\begin{aligned} \mathcal{G}f(x) &= \frac{1}{2} \operatorname{Tr} (a(x)\nabla^2 f(x)) + b(x)^\top \nabla f(x) \\ &\quad + \lambda(x) \int (f(x + \gamma(x, y)) - f(x) - \gamma(x, y)^\top \nabla f(x)) \nu(x, d\xi), \end{aligned} \quad (6.1)$$

where λ is nonnegative and measurable, and $\gamma(x, y)$ is affine in x . In order to describe the form of the jump sizes, let us introduce the set $(\Delta^d)^d$ which is given by

$$(\Delta^d)^d = \{y = (y^1, \dots, y^d) \in \mathbb{R}_+^{d \times d} : y^i \in \Delta^d \text{ for all } i \in \{1, \dots, d\}\}.$$

Type 0. For some $\alpha_{ij} \in \mathbb{R}_+$, $\alpha_{ij} = \alpha_{ji}$, $B \in \mathbb{R}^{d \times d}$ such that $B_{ij} \geq 0$ for $i \neq j$ and $B_{ii} = -\sum_{j \neq i} B_{ji}$, let

$$a_{ii}(x) = \sum_{i \neq j} \alpha_{ij} x_i x_j \quad \text{and} \quad a_{ij}(x) = -\alpha_{ij} x_i x_j \quad \text{for all } i \neq j, \quad (6.2)$$

$$b(x) = Bx, \quad (6.3)$$

and set $\lambda = 0$. Then \mathcal{G} is a polynomial operator whose martingale problem is well-posed. The solutions X are multivariate Jacobi-type diffusion processes which have been characterized in this form by Filipović and Larsson (2016, Section 6.3). In the special case where $\alpha_{ij} = \sigma^2$ for all i, j , they correspond to Wright-Fisher diffusions, which are also known under the name multivariate Jacobi process; see Gourieroux and Jasiak (2006).

Type 1. Let $\lambda(x) = 1$ and $a(x), b(x)$ be given by (6.2) and (6.3). For all $y \in (\Delta^d)^d$ set

$$\gamma(x, y) = \sum_{i=1}^d (y^i - e_i) x_i, \quad (6.4)$$

and let μ be a nonzero measure on $(\Delta^d)^d$. If the boundary conditions

$$B_{ij} - \int y_i^j \mu(dy) \geq 0$$

hold for all $i \neq j$, then \mathcal{G} is a polynomial operator whose martingale problem is well-posed.

Note that the boundary conditions imply that

$$\int |\xi| \nu(x, d\xi) \leq \sum_{i=1}^d \int |y^i - e_i| \mu(dy)$$

is bounded. Hence the resulting process behaves like a multivariate Jacobi-type diffusion process in the spirit of Filipović and Larsson (2016, Section 6.3), generalized to include summable jumps. The arrival intensity of the jumps is $\nu(x, E - x) = \mu(\{y : \gamma(x, y) \neq 0\})$, which may or may not be finite.

Type 2. Fix $i \in \{1, \dots, d\}$. Let $a(x), b(x)$ be given by (6.2) and (6.3), and let $\lambda(x) = \frac{q_1(x)}{x_i} \mathbb{1}_{\{x_i \neq 0\}}$ for some nonnegative $q_1 \in \text{Pol}_1(E)$ such that λ is not constant on $E \cap \{x_i \neq 0\}$. Furthermore, for $y \in \Delta^d$ we define

$$\gamma(x, y) = (y - e_i)x_i,$$

and let μ be a nonzero square-integrable measure on $\Delta^d \setminus \{e_i\}$. If the boundary conditions

$$B_{kj} - q_1(e_j) \int y_k \mu(dy) \geq 0 \quad (6.5)$$

hold for all $k \neq i$ and $j \neq k$, then \mathcal{G} is a polynomial operator whose martingale problem is well-posed.

If $q_1(x) = Lx_k$ for some $k \neq i$ and $L > 0$, the jumps need not to be summable. More precisely, we can have $\int |y_i - 1| \mu(dy) = \int |y_k| \mu(dy) = \infty$. Otherwise, if $q_1(x)$ is not proportional to x_k on E for any k , the expression $\int |\xi| \nu(x, d\xi)$ is bounded, and the solution X to the martingale problem for \mathcal{G} has thus summable jumps. Indeed,

$$\int |\xi_k| \nu(x, d\xi) \leq \sup_{x \in E} q_1(x) \int y_k \mu(dy),$$

which is bounded due to (6.5) and the existence of some $x \in E \cap \{x_k = 0\} \cap \{x_i \neq 0\}$ such that $q_1(x) \neq 0$; see Lemma E.1 in Section E for more details on the second point.

The arrival intensity of the jumps is $\nu(x, E - x) = \lambda(x)\mu(\Delta^d \setminus \{e_i\})$ and hence, even if μ is a finite measure, the jump intensity is unbounded around $x_i = 0$. Moreover, due to the form of $\gamma(x, y)$, X can only jump in the direction of the boundary segment $E \cap \{x_i = 0\}$, and since $\nu(x, E) = 0$ whenever $x_i = 0$, X cannot leave this boundary segment by means of a jump. Figure 7 illustrates the form of λ and the support of $\gamma(x, y)$ under μ .

Type 3. Let $i, j \in \{1, \dots, d\}$ be such that $i \neq j$, and fix some constant $c > 0$. Consider the hyperplane $\{cx_i = x_j\}$ which will be a “no-jump” region. Let b be given by (6.3) and set

$$\lambda(x) = \frac{q_2(x)}{(-cx_i + x_j)^2} \mathbb{1}_{\{cx_i \neq x_j\}}$$

for some $q_2 \in \text{Pol}_2(E)$ given by $q_2(x) = \sum_{k=1}^d (q_{ik}x_i x_k + q_{jk}x_j x_k)$, where $q_{k\ell} \in \mathbb{R}$ are chosen such that λ is nonnegative, and nonconstant on $\{cx_i \neq x_j\}$. Furthermore, define

$$\gamma(x, y) = y(-cx_i + x_j)(e_i - e_j)$$

and let μ be a nonzero square-integrable measure on $(0, \frac{1}{c} \wedge 1]$. Finally, let

$$a(x) = a^c(x) + a^\nu(x)A^\nu \mathbb{1}_{\{cx_i = x_j\}}$$

where a^c is of form (6.2), $A^\nu \in \mathbb{R}^{d \times d}$ is a symmetric matrix given by $A_{ii}^\nu = A_{jj}^\nu = 1$, $A_{ij}^\nu = -1$, and $A_{k\ell}^\nu = 0$ if $k \notin \{i, j\}$, and where

$$a^\nu(x) = q_2(x) \int y^2 \mu(dy).$$

If the boundary conditions

$$\sum_{\ell \neq k} B_{k\ell} x_\ell - \frac{q_2(x)}{(-cx_i + x_j)} \mathbb{1}_{\{cx_i \neq x_j\}} \int y(\delta_{ik} - \delta_{jk}) \mu(dy) \geq 0 \quad (6.6)$$

are satisfied for all $x \in E \cap \{x_k = 0\}$ and $k \in \{1, \dots, d\}$, then \mathcal{G} is a polynomial operator whose martingale problem is well-posed. Note in particular that for $k \notin \{i, j\}$, condition (6.6) coincides with $b_k \geq 0$ on $E \cap \{x_k = 0\}$.

If the numerator of λ is of the form $q_2(x) = 2q_{ij}x_ix_j$ for some $q_{ij} \in \mathbb{R}_+$, the solution X to the martingale problem for \mathcal{G} may have nonsummable jumps. If $q_2(x)$ is not of this form, then by similar reasoning as for Type 2, the boundary conditions imply that $\int y\mu(dy) < \infty$ and thus X has summable jumps. The arrival intensity of the jumps is

$$\nu(x, E - x) = \lambda(x)\mu\left(\left(0, \frac{1}{c} \wedge 1\right]\right).$$

As a result, even if μ is a finite measure, the jump intensity has a singularity of order two along $\{cx_i = x_j\}$, which results in a contribution of $a^\nu(x)A^\nu$ to the diffusion coefficient. Moreover, due to the form of $\gamma(x, y)$, the jumps of X are always in the direction of the “no-jump” hyperplane $\{cx_i = x_j\}$. Although the jumps may overshoot $\{cx_i = x_j\}$, they always serve to reduce the distance to $\{cx_i = x_j\}$. In particular, since $\nu(x, E - x) = 0$ for all $x \in E \cap \{cx_i = x_j\}$, X cannot leave $\{cx_i = x_j\}$ by means of a jump. Figure 8 illustrates the form of λ and the support of $\gamma(x, y)$ under μ .

In order to simplify the analysis, in particular in view of the arguments outlined in Remark 4.4, we do not consider operators corresponding to Type 4 on the unit simplex. A condition on the jump kernel excluding this class is given by the following assumption.

Assumption A. *The condition $\lambda\gamma_i(\cdot, y)^3 \in \text{Pol}_3(E)$ holds for all $i \in \{1, \dots, d\}$ and all $y \in \text{supp}(\mu)$.*

The polynomial property of \mathcal{G} implies that the integrated quantities

$$\lambda \int \gamma_i(\cdot, y)^3 \mu(dy)$$

lie in $\text{Pol}_3(E)$. Assumption A strengthens this by requiring that the functions $\lambda\gamma_i(\cdot, y)^3$ themselves lie in $\text{Pol}_3(E)$. This is a natural assumption, in particular in view of Types 1-3 on the unit interval. Moreover, it will turn out in the course of the proof of Theorem 6.3 that Assumption A implies under the condition of affine jump sizes and nonconstant λ that

$$\gamma(\cdot, y) = H(y)P_1$$

where H is a μ -measurable function and $P_1 \in \text{Pol}_1(E)$. Analogous to the unit interval, the “no-jump” region is the intersection of E with the hyperplane given

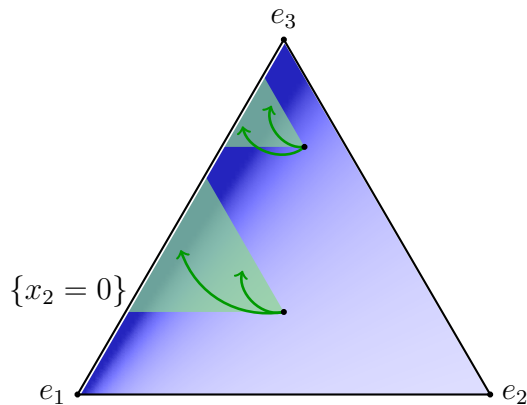


Figure 7

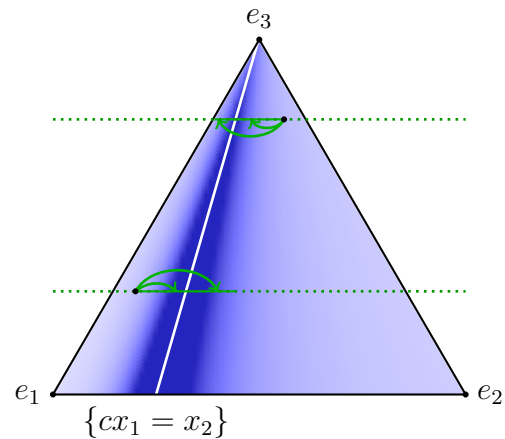


Figure 8

Figure 7: A representation of Type 2, where λ explodes on the boundary segment $\{x_2 = 0\}$ (in blue) and the support of $\nu(x, \cdot)$ is always contained in $(\Delta^3 - e_2)x_2$ (in green) for all $x \in E$. This in particular implies that the distance to the “no-jump” hyperplane $\{x_2 = 0\}$ always decreases if a jump occurs. Note that on $\{x_2 = 0\}$ there is no jump activity since $\nu(x, d\xi) = 0$ for all $x \in E \cap \{x_2 = 0\}$.

Figure 8: A representation of Type 3, where λ explodes on the hyperplane $\{cx_1 = x_2\}$ (in blue) and the support of $\nu(x, \cdot)$ is always contained in $\Delta^3 \cap \{x_3 = 0\}$ (in green) for all $x \in E$. Moreover the distance to the “no-jump” hyperplane $\{cx_1 = x_2\}$ always decreases if a jump occurs. Note that on $\{cx_1 = x_2\}$ there is no jump activity since $\nu(x, d\xi) = 0$ for all $x \in E \cap \{cx_1 = x_2\}$, but we know from the description of this type that there is an extra contribution to the diffusion matrix a .

by the zero set of a polynomial of first degree. The following theorem states the announced characterization of polynomial jump–diffusions with polynomial jump sizes under Assumption A. The proof is given in Section E.

Theorem 6.3. *Let \mathcal{G} be a polynomial operator whose martingale problem is well-posed. If the associated jump kernel has affine jump sizes and satisfies Assumption A, then \mathcal{G} necessarily belongs to one of the Types 0-3.*

7 Applications

In this section we outline two natural applications in finance of polynomial jump–diffusions on the unit simplex. The first application concerns stochastic portfolio theory, while the second application is in the area of default risk.

7.1 Market weights with jumps in stochastic portfolio theory

In the context of stochastic portfolio theory (SPT), polynomial diffusion models for the process of market weights have been found capable of matching certain empirically observed properties when calibrated to *jump-cleaned* data; see [Cuchiero \(2017\)](#); [Cuchiero et al. \(2016\)](#). This concerns the typical shape and dynamics of the capital distribution curves, but also features such as high volatility for low capitalized stocks. As mentioned in the introduction, a crucial deficiency of these models is the lack of jumps since they are present in typical market data; see [Figure 2](#).

We now demonstrate how the results of [Section 6](#) can be used to construct polynomial *jump*-diffusion models for the market weights. We focus on a concrete specification that extends the *volatility stabilized models* introduced by [Fernholz and Karatzas \(2005\)](#) by including jumps of Type 2. In the standard (diffusive) volatility stabilized model, the market weights follow a Wright-Fisher diffusion, which is a special case of Type 0 with parameters

$$\alpha_{ij} = 1 \quad \forall i \neq j \quad \text{and} \quad B = \frac{1 + \beta}{2} \mathbf{1}\mathbf{1}^\top - \frac{d(1 + \beta)}{2} \text{Id},$$

for some $\beta \geq 0$. These models have two key properties which are of particular relevance in SPT. First, the market weights remain a.s. in the relative interior of Δ^d , denoted by $\overset{\circ}{\Delta}^d$. Second, the model allows for relative arbitrage opportunities. We may preserve these features by adding jumps of Type 2. More precisely, we consider a model for the market weights $(X_t)_{t \geq 0}$ of the form

$$\begin{aligned} X_{t,i} = & \int_0^t \left(\frac{1 + \beta}{2} - \frac{d(1 + \beta)}{2} X_{s,i} \right) ds + \int_0^t \sqrt{X_{s,i}} (1 - X_{s,i}) dW_{s,i} \\ & - \sum_{i \neq j} \int_0^t X_{s,i} \sqrt{X_{s,i}} dW_{s,j} + \int_0^t \int \xi_i(\mu^X(d\xi, ds) - \nu(X_s, d\xi) ds), \end{aligned}$$

where μ^X denotes the integer-valued random measure associated to the jumps of X , and W a d -dimensional standard Brownian motion. The jump specification is given as a sum of Type 2 jumps,

$$\nu(x, A) = \sum_{i=1}^d \lambda_i(x) \int \mathbf{1}_A((y - e_i)x_i) \mu_i(dy),$$

where $\lambda_i(x) = \frac{q_i(x)}{x_i} \mathbf{1}_{\{x_i \neq 0\}}$ for some nonnegative $q_i \in \text{Pol}_1(E)$ such that λ is not constant on $E \cap \{x_i \neq 0\}$, and the measures μ_i are supported on $\Delta^d \setminus \{e_i\}$ and satisfy $\int |y| \mu_i(dy) < \infty$. Economically, this specification means that downward jumps occur with higher and higher intensity the closer the assets are to 0, and can therefore be used to model downward spirals in stock prices. We require that for all $j \neq k$,

$$\frac{\beta}{2} - \sum_{i \neq k} q_i(e_j) \int y_k \mu_i(dy) + q_k(e_j) \int (1 + \log(y_k) - y_k) \mu_k(dy) > 0,$$

which ensures that X remains in the relative interior $\overset{\circ}{\Delta}^d$. This can be proved similarly as in [Filipović and Larsson \(2016, Theorem 5.7\)](#). Furthermore, this model admits relative arbitrage opportunities. To see this, we argue that no equivalent probability measure can turn X into a martingale. Indeed, Lemma 5.6 in [Cuchiero \(2017\)](#) implies that, under any martingale measure, X must reach the relative boundary of Δ^d with positive probability on any time horizon, contradicting equivalence. Since no equivalent martingale measure exists for the market weights, the model admits relative arbitrage.

Clearly any other polynomial diffusion model on the simplex can be enhanced by jumps of this form, which yields a large class of tractable jump-diffusion models applicable in the realm of SPT.

7.2 Valuation of defaultable zero-coupon bonds

This section builds on discussions with Thomas Krabichler. Polynomial jump-diffusions on the unit interval can be brought to bear on default risk modeling. We consider the stochastic recovery rate framework of [Jarrow and Turnbull \(1998\)](#) and [Krabichler and Teichmann \(2017\)](#). For further references, see also [Filipović and Trolle \(2013\)](#), [Zheng \(2006\)](#), and [Jeanblanc et al. \(2009, Chapter 7\)](#) for an overview, as well as [Duffie \(2004\)](#) for the classical approach using affine processes. Note also that polynomial diffusion models for credit risk have appeared in [Akerer and Filipović \(2016\)](#).

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ with $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ be a filtered probability space satisfying the usual conditions. Here \mathbb{Q} is a risk neutral measure. Let $B = (B_t)_{t \geq 0}$ be the value of the risk-free bank account with initial value of one monetary unit. For any $t \leq T$, we denote by $\tilde{P}(t, T)$ the price at time t of a defaultable zero-coupon bond with maturity $T \geq 0$ and unit notional. Due to default risk, its actual payoff $\tilde{P}(T, T)$ at maturity is random and lies between zero and one. Under the premise that all discounted defaultable zero-coupon bond prices $\tilde{P}(t, T)/B(t)$ are true martingales under \mathbb{Q} , we get

$$\tilde{P}(t, T) = \mathbb{E}_{\mathbb{Q}} \left[\frac{B_t}{B_T} S_T \middle| \mathcal{F}_t \right],$$

where $S_t := \tilde{P}(t, t)$ is known as the *recovery rate*. Suppose now for simplicity that B and S are conditionally independent under \mathbb{Q} . Then

$$\tilde{P}(t, T) = \mathbb{E}_{\mathbb{Q}} \left[\frac{B_t}{B_T} \middle| \mathcal{F}_t \right] \mathbb{E}_{\mathbb{Q}} [S_T | \mathcal{F}_t] = P(t, T) \mathbb{E}_{\mathbb{Q}} [S_T | \mathcal{F}_t],$$

where $P(t, T)$ is the price of a non-defaultable zero-coupon bond with maturity $T \geq 0$ and unit notional.

Motivated by the typically long and complicated unwinding process after a default occurs, [Krabichler and Teichmann \(2017\)](#) drop the assumption that the recovery rate S is known when default happens. Excursions of S below 1 are interpreted as liquidity squeezes resulting in a delay of due payments, which may or may not turn into a default. In this framework, the risk-neutral recovery rate

S typically starts with a constant trajectory at level 1. Once the recovery has jumped below 1, it pursues an unsteady course. Downward moves of the recovery rate are self-exciting, as deterioration of the counterparty’s credit quality typically makes full recovery more unlikely. Nonetheless, S may return to 1 and remain there for some period of time.

A polynomial model for the recovery rate S can be constructed as follows. Let X be a polynomial jump–diffusion of Type 2 with “no-jump” point $x^* = 0$. Assume that $\kappa(1 - \theta) = (1 + q) \int y \mu(dy)$; this condition guarantees that if X reaches level 1, it can leave it only by means of a jump. More precisely, X persists at level 1 until its first jump, which occurs according to an $(1 + q)$ –exponentially distributed stopping time and a downward μ –distributed jump size. Moreover, since the jump intensity is the positive branch of a hyperbola with a pole in zero, downward jumps of X get more and more likely as the process approaches zero.

In view of the discussion above, polynomial transformations $S := p(X)$ of X , where $p \in \text{Pol}([0, 1])$ is increasing and satisfies $p([0, 1]) \subseteq [0, 1]$, are well-suited to describe the recovery rate. The polynomial property of X permits to express the forward recovery rate $F(t, T) = \mathbb{E}_{\mathbb{Q}}[S_T \mid \mathcal{F}_t]$ in closed form. We provide two concrete specifications, by choosing $p(x) = x$ and $p(x) = x^2$. In the first case, $S = X$, the moment formula (2.2) yields

$$F(t, T) = (1 - e^{-(T-t)\kappa})\theta + e^{-(T-t)\kappa} S_t.$$

In the second case, $S = X^2$, we find

$$F(t, T) = \frac{(\kappa(1 - e^{(T-t)G_2}) + G_2(1 - e^{-(T-t)\kappa}))\theta}{\kappa + G_2} + \frac{G_1(e^{(T-t)G_2} - e^{-(T-t)\kappa})}{\kappa + G_2} \sqrt{S_t} + e^{-(T-t)\kappa} S_t,$$

where $G_1 := A + 2\kappa\theta + \int y^2 \mu(dy)$ and $G_2 := -A - 2\kappa + q \int y^2 \mu(dy)$.

A Proof of Theorem 2.8

We assume that $\mathcal{G} : \text{Pol}(E) \rightarrow C(E)$ is a linear operator that satisfies the positive maximum principle and $\mathcal{G}1 = 0$.

Fix $x \in E$ and define the linear functionals $\mathcal{W}_{ij} : \text{Pol}(E - x) \rightarrow \mathbb{R}$ for $i, j = 1, \dots, d$ by

$$\mathcal{W}_{ij}(p) := \mathcal{G}(p(\cdot - x)e_i^\top(\cdot - x)e_j^\top(\cdot - x))(x),$$

as well as $\mathcal{W}_u : \text{Pol}(E - x) \rightarrow \mathbb{R}$ for $u \in \mathbb{R}^d$ by

$$\mathcal{W}_u(p) := \sum_{i,j=1}^d u_i u_j \mathcal{W}_{ij}(p) = \mathcal{G}(p(\cdot - x)(u^\top(\cdot - x))^2)(x).$$

Here and throughout the proof we view $u^\top(\cdot - x)$ as a polynomial on E to avoid the more cumbersome notation $u^\top(\cdot - x)|_E$.

If $p \geq 0$ on $E - x$, then $p(\cdot - x)(u^\top(\cdot - x))^2 \in \text{Pol}(E)$ is minimal at x , which by the positive maximum principle yields $\mathcal{W}_u(p) \geq 0$. The Riesz-Haviland theorem, Lemma 2.7, thus provides measures $\nu_u(x, d\xi)$ concentrated on $E - x$ such that

$$\mathcal{W}_u(p) = \int p(\xi)\nu_u(x, d\xi).$$

By polarisation we have $\mathcal{W}_{ij} = \frac{1}{2}(\mathcal{W}_{e_i+e_j} - \mathcal{W}_{e_i} - \mathcal{W}_{e_j})$, whence

$$\mathcal{W}_{ij}(p) = \int p(\xi)\nu_{ij}(x, d\xi), \quad \nu_{ij} = \frac{1}{2}(\nu_{e_i+e_j} - \nu_{e_i} - \nu_{e_j}).$$

The triplet (a, b, ν) is now defined at x by

$$a_{ij}(x) := \nu_{ij}(x, \{0\}), \quad b_i(x) := \mathcal{G}(e_i^\top(\cdot - x))(x), \quad (\text{A.1})$$

and

$$\nu(x, d\xi) := \frac{1}{|\xi|^2} \mathbb{1}_{\{\xi \neq 0\}} (\nu_{e_1}(x, d\xi) + \cdots + \nu_{e_d}(x, d\xi)).$$

Next, observe that

$$\begin{aligned} \int \xi_i \xi_j p(\xi) |\xi|^2 \nu(x, d\xi) &= \sum_{k=1}^d \int \xi_i \xi_j p(\xi) \nu_{e_k}(x, d\xi) \\ &= \sum_{k=1}^d \mathcal{G} (p(\cdot - x) e_i^\top(\cdot - x) e_j^\top(\cdot - x) (e_k^\top(\cdot - x))^2) (x) \\ &= \mathcal{G} (p(\cdot - x) e_i^\top(\cdot - x) e_j^\top(\cdot - x) |\cdot - x|^2) (x) \\ &= \mathcal{W}_{ij}(p|\cdot|^2) \\ &= \int p(\xi) |\xi|^2 \nu_{ij}(x, d\xi), \end{aligned}$$

for all $p \in \text{Pol}(E - x)$. By Weierstrass's theorem and dominated convergence, this actually holds for all $p \in C(E - x)$, whence $\mathbb{1}_{\{\xi \neq 0\}} \nu_{ij}(x, d\xi) = \xi_i \xi_j \nu(x, d\xi)$. Consequently,

$$\mathcal{W}_{ij}(p) = \int p(\xi) \xi_i \xi_j \nu(x, d\xi) + p(0) a_{ij}(x). \quad (\text{A.2})$$

Consider now any polynomial $p \in \text{Pol}(E - x)$, and choose a representative $q \in \text{Pol}(\mathbb{R}^d)$, $p = q|_{E-x}$. Note that q is of the form

$$q(\xi) = c_0 + \sum_{i=1}^d c_i \xi_i + \sum_{i,j=1}^d \xi_i \xi_j q_{ij}(\xi)$$

for some polynomials $q_{ij} \in \text{Pol}(\mathbb{R}^d)$. Let $p_{ij} := q_{ij}|_{E-x} \in \text{Pol}(E - x)$. Then the

linearity of \mathcal{G} , the fact that $\mathcal{G}1 = 0$, and (A.1) and (A.2) yield

$$\begin{aligned} \mathcal{G}(p(\cdot - x))(x) &= c_0 \mathcal{G}1(x) + \sum_{i=1}^d c_i \mathcal{G}(e_i^\top(\cdot - x))(x) \\ &\quad + \sum_{i,j=1}^d \mathcal{G}(p_{ij}(\cdot - x)e_i^\top(\cdot - x)e_j^\top(\cdot - x))(x) \\ &= \sum_{i=1}^d c_i b_i(x) + \sum_{i,j=1}^d \left(\int q_{ij}(\xi) \xi_i \xi_j \nu(x, d\xi) + q_{ij}(0) a_{ij}(x) \right) \\ &= \frac{1}{2} \text{Tr}(a(x) \nabla^2 q(0)) + b(x)^\top \nabla q(0) \\ &\quad + \int (q(\xi) - q(0) - \xi^\top \nabla q(0)) \nu(x, d\xi). \end{aligned}$$

Thus, with $p(\xi) = f(x + \xi)$ for a polynomial $f \in \text{Pol}(E)$, we obtain the desired form (2.3), where the right-hand side is computed using a representative of f , the choice of which is arbitrary.

It remains to verify that the a , b , and ν satisfy the additional stated properties. First, $a(x)$ is positive semidefinite since $u^\top a(x)u = \sum_{i,j=1}^d u_i u_j \nu_{ij}(x, \{0\}) = \nu_u(x, \{0\}) \geq 0$, and ν clearly satisfies the support conditions $\nu(x, \{0\}) = 0$ and $\nu(x, (E - x)^c) = 0$. Next, since \mathcal{G} maps polynomials to continuous functions, it is clear from (A.1) that b is bounded and measurable. Similarly, $x \mapsto \int p(\xi) \nu_u(x, d\xi) = \mathcal{G}(p(\cdot - x)(u^\top(\cdot - x))^2)(x)$ is continuous, hence bounded and measurable, for every $p \in \text{Pol}(E)$, and so by the monotone class theorem $\nu_u(\cdot, A)$ is measurable for every Borel set $A \subseteq E - x$. Thus $\nu_u(x, d\xi)$ is a kernel, from which it follows that a is measurable and $\nu(x, d\xi)$ is a kernel. Finally, continuity in x of

$$\text{Tr}(a(x)) + \int |\xi|^2 \nu(x, d\xi) = \nu_{e_1}(x, E - x) + \cdots + \nu_{e_d}(x, E - x) = \mathcal{G}(|\cdot - x|^2)(x)$$

implies that a and $\int |\xi|^2 \nu(\cdot, d\xi)$ are bounded on E . \square

B Proof of Theorem 2.10 and Lemma 2.11

Proof of Theorem 2.10. Let $\alpha, \beta \in \mathbb{R}_+$ and $\mathcal{G}_1, \mathcal{G}_2 \in \mathcal{K}$, and fix $f \in \text{Pol}_k(E)$ for some $k \in \mathbb{N}$. Then, since $\mathcal{G}_1 f, \mathcal{G}_2 f \in \text{Pol}_k(E)$, we have that

$$\mathcal{G}f := \alpha \mathcal{G}_1 f + \beta \mathcal{G}_2 f \in \text{Pol}_k(E)$$

as well, proving that it is polynomial. By Theorem 2.3, the well-posedness of the martingale problem for \mathcal{G} follows directly by the well-posedness of the martingale problems of \mathcal{G}_1 and \mathcal{G}_2 . For the second part, set $(\mathcal{G}_n)_{n \in \mathbb{N}}$ as in the statement of the theorem and recall that $\text{Pol}_k(E)$ is closed under pointwise convergence for each $k \in \mathbb{N}_0$. Fixing $f \in \text{Pol}_k(E)$, since $\mathcal{G}_n f \in \text{Pol}_k(E)$ by the polynomial property of

\mathcal{G}_n , we can conclude that $\mathcal{G}f \in \text{Pol}_k(E)$ as well. Again, existence and uniqueness of a solution to the martingale problem is guaranteed by Theorem 2.3, since $\mathcal{G}1 = 0$ and the positive maximum principle is preserved in the limit. \square

Proof of Lemma 2.11. In order to prove the first part of the lemma, it is enough to observe that for all $n \in \mathbb{N}$ and $|\mathbf{k}| \geq 1$,

$$\mathcal{G}_n((\cdot - x)^{\mathbf{k}})(x) = \begin{cases} b_i^n(x) & \text{if } \mathbf{k} = e_i, \\ a_{ij}^n(x) + \int \xi_i \xi_j \nu_n(x, d\xi) & \text{if } \mathbf{k} = e_i + e_j, \\ \int \xi^{\mathbf{k}} \nu_n(x, d\xi) & \text{if } |\mathbf{k}| \geq 3. \end{cases} \quad (\text{B.1})$$

For the second part of the lemma, if \mathcal{G} is well-defined we know from Theorem 2.10 that it has a Lévy-Khintchin representation, for some coefficients a and b , and a kernel $\nu(x, d\xi)$. As a result, the analog of (B.1) holds true for \mathcal{G} and by definition of the limit we thus obtain

$$b_i(x) = \mathcal{G}((\cdot - x)^{\mathbf{k}})(x) = \lim_{n \rightarrow \infty} \mathcal{G}_n((\cdot - x)^{\mathbf{k}})(x) = \lim_{n \rightarrow \infty} b_i^n(x),$$

and similarly

$$a_{ij}(x) \mathbf{1}_{\{\mathbf{k}=e_i+e_j\}} + \int \xi^{\mathbf{k}} \nu(x, d\xi) = \lim_{n \rightarrow \infty} \left(a_{ij}^n(x) \mathbf{1}_{\{\mathbf{k}=e_i+e_j\}} + \int \xi^{\mathbf{k}} \nu_n(x, d\xi) \right), \quad (\text{B.2})$$

for all $|\mathbf{k}| \geq 2$. Since $\nu(x, d\xi)$ does not have mass in 0,

$$\nu(x, d\xi) = |\xi|^{-4} (|\xi|^4 \nu(x, d\xi)).$$

Moreover, using that moments completely determine compactly supported finite distribution, the kernel $|\xi|^4 \nu(x, d\xi)$, and thus $\nu(x, d\xi)$, is uniquely determined by (B.2). \square

C Proof of Theorem 3.3

Throughout the proof we assume without loss of generality that E has nonempty interior. Suppose that $\nu(x, d\xi)$ is the zero measure for all $x \in E$. Setting $\lambda = 0$, the form of $\gamma(x, y)$ and the measure μ are irrelevant and we are thus free to choose $K \leq 1$. We may therefore suppose that $\nu(x, \cdot)$ is nonzero for some $x \in E$, and thus in particular

$$\mu(\gamma(x, \cdot) \neq 0) > 0 \quad (\text{C.1})$$

for at least one $x \in E$. As in Remark 3.2, we can assume without loss of generality that μ is compactly supported and hence all its moments of order at least two are finite. Set then

$$p_n := \int \gamma^n(\cdot, y) \mu(dy) \quad \text{and} \quad r_n := \int \xi^n \nu(\cdot, d\xi) = \lambda p_n$$

and note that, by the integrability conditions on μ and condition (3.3) respectively, $p_{\mathbf{n}}$ and $r_{\mathbf{n}}$ are polynomials on E for all $|\mathbf{n}| \geq 3$. In particular, p_{4e_i} is a nonzero polynomial for at least one $i \in \{1, \dots, d\}$ by (C.1), and thus

$$\lambda(x) = \frac{r_{4e_i}(x)}{p_{4e_i}(x)} \quad (\text{C.2})$$

for all $x \in E \setminus \{p_{4e_i} = 0\}$.

Since E has nonempty interior by assumption, each polynomial $p \in \text{Pol}(E)$ has a unique representative $\bar{p} \in \text{Pol}(\mathbb{R}^d)$ such that $\bar{p}|_E = p$. In particular, the degree of a polynomial on E always coincides with the maximal degree of its monomials. Assume now for contradiction that K cannot be chosen less than or equal one. Let

$$\mathbf{n}_j := \sup\{\mathbf{k} : \mu(y_{\mathbf{k}}^j) \neq 0\} \neq 0$$

be the multi-index corresponding to the leading monomial of $\gamma_j(x, y)$, with respect to some graded lexicographic order. Choose $j \in \{1, \dots, d\}$ such that $|\mathbf{n}_j| \geq 2$ and note that by the maximality of \mathbf{n}_j and since $\int (y_{\mathbf{n}_j}^j)^{10} \mu(dy) > 0$ we have that

$$\deg(p_{10e_j}) = \deg\left(\int (y_{\mathbf{n}_j}^j)^{10} \mu(dy) x^{10\mathbf{n}_j}\right) = 10|\mathbf{n}_j|.$$

Analogously, $\deg(p_{4e_j}) = 4|\mathbf{n}_j|$ and thus (C.2) holds true for $i = j$. Since $p_{10e_j}(x)r_{4e_j}(x) = p_{4e_j}(x)r_{10e_j}(x)$, using that $|\mathbf{n}_j| \geq 2$ we can compute

$$\deg(p_{10e_j}r_{4e_j}) \geq 10|\mathbf{n}_j| > 4|\mathbf{n}_j| + 10 \geq \deg(p_{4e_j}r_{10e_j}),$$

and obtain the desired contradiction. As a result, K can always be chosen smaller than or equal one. \square

D The unit interval: Proof of Lemma 4.1 and Theorem 4.3

Proof of Lemma 4.1. Assume \mathcal{G} is a polynomial operator and its martingale problem is well-posed. Theorem 2.3 and Theorem 2.8 imply that \mathcal{G} is of Lévy type for some triplet (a, b, ν) , so that in particular (i) holds. Condition (iii) follows from Lemma 2.9. To verify (ii), let f_n be polynomials on $[0, 1]$ with $0 \leq f_n \leq 1$, $f_n(0) = 1$, $xf_n(x) \leq 1$, and $f_n(x) \downarrow 0$ for $x \in (0, 1]$. For example, one can choose $f_n(x) := \frac{n-1}{n}(1-x)^n + \frac{1}{n}$. Let $g_n(x) := \frac{x}{n} - x^2 f_n(x)$. Then g_n has a minimum at $x = 0$, so by the positive maximum principle,

$$0 \leq \mathcal{G}g_n(0) = -a(0) + \frac{1}{n}b(0) - \int f_n(\xi)\xi^2\nu(0, d\xi) \rightarrow -a(0), \quad n \rightarrow \infty,$$

where the dominated convergence theorem was used to pass to the limit. Similarly, $h_n(x) := xf_n(x)$ is nonnegative on $[0, 1]$ with a minimum at $x = 0$, so the monotone convergence theorem yields

$$0 \leq \mathcal{G}h_n(0) = b(0) - \int \xi(1 - f_n(\xi))\nu(0, d\xi) \rightarrow b(0) - \int \xi\nu(0, d\xi), \quad n \rightarrow \infty.$$

We have thus shown (ii) for the boundary point $x = 0$. The case $x = 1$ is similar.

We now prove the converse. Lemma 2.9 and (iii) imply that \mathcal{G} is polynomial. Next, clearly $\mathcal{G}1 = 0$. Thus, by Theorem 2.3 it only remains to verify the positive maximum principle in order to deduce that the martingale problem for \mathcal{G} is well-posed. To this end, let $f \in \text{Pol}(E)$ be an arbitrary polynomial having a maximum over E on some $x \in E$. If $x \in \text{int}(E)$ it follows that $f'(x) = 0$, $f''(x) \leq 0$, and $f(x) \geq f(x + \xi)$ for all $\xi \in E - x$. Hence, using that $a \geq 0$ on E , we conclude that $\mathcal{G}f(x) \leq 0$. On the other hand, if $x \in \partial E = \{0, 1\}$ we use that $a(x) = 0$ and the integrability of ξ with respect to $\nu(x, \cdot)$ to write

$$\mathcal{G}f(x) = \left(b(x) - \int \xi \nu(x, d\xi) \right) f'(x) + \int (f(x + \xi) - f(x)) \nu(x, d\xi).$$

The classical Karush-Kuhn-Tucker conditions (see e.g. Proposition 3.3.1 in Bertsekas (1995)) imply that $f'(x) \leq 0$ if $x = 0$ and $f'(x) \geq 0$ if $x = 1$, and thus the first summand is nonnegative by (ii). Using as before that $f(x) \geq f(x + \xi)$ for all $\xi \in E - x$, we conclude that $\mathcal{G}f(x) \leq 0$. \square

Proof of Theorem 4.3. By assumption \mathcal{G} is polynomial and its martingale problem is well-posed. Hence, conditions (i)-(iii) of Lemma 4.1 are satisfied. As in Remark 3.2, we can assume without loss of generality that μ is compactly supported. In particular, all its moments of order at least two are finite. For all $n \geq 2$ set then

$$p_n := \int \gamma^n(\cdot, y) \mu(dy) \quad \text{and} \quad r_n := \int \xi^n \nu(\cdot, d\xi) = \lambda p_n. \quad (\text{D.1})$$

Note that $p_n \in \text{Pol}_n(E)$ for all $n \geq 2$ by the integrability conditions on μ , and $r_n \in \text{Pol}_n(E)$ for all $n \geq 3$ by condition (iii) of Lemma 4.1. By Remark 3.4 we know that

$$\lambda(x) = \frac{r_4(x)}{p_4(x)} \mathbf{1}_{\{p_4(x) \neq 0\}}$$

and hence the condition $\nu(x, (E - x)^c) = 0$ of (2.4) implies that μ can be chosen to be supported on $[0, 1]^2$ and such that $\gamma(x, y) = y_1(-x) + y_2(1 - x)$ μ -a.s. By Lemma 4.1(iii) we also know that $b \in \text{Pol}_1(E)$ and, by Lemma 4.1(ii), that the boundary conditions

$$b(0) \geq \lambda(0) \int \gamma(0, y) \mu(dy) \quad \text{and} \quad b(1) \leq \lambda(1) \int \gamma(1, y) \mu(dy) \quad (\text{D.2})$$

hold. We consider now five complementary assumptions, which will lead to Types 0 to 4.

Assume that $\nu(x, d\xi) = 0$. Then Lemma 4.1 implies that $a(x) = Ax(1 - x)$ for some $A \in \mathbb{R}_+$. This proves that \mathcal{G} is an operator of Type 0.

Assume now that $\nu(x, d\xi) \neq 0$ and λ can be chosen to be constant. We can then without loss of generality set $\lambda = 1$. Moreover, since in this case $r_2 \in \text{Pol}_2(E)$, we can conclude as before that $a(x) = Ax(1 - x)$ for some $A \in \mathbb{R}_+$. This proves that \mathcal{G} is an operator of Type 1.

Assume that $\nu(x, d\xi) \neq 0$, λ cannot be chosen to be constant, and $p_4(x^*) = 0$ for some $x^* \in \mathbb{R}$. By definition of p_4 this automatically implies that $\gamma(x^*, y) = 0$, and in particular $x^* = y_2(y_1 + y_2)^{-1}$, for μ -a.e. $y \in [0, 1]^2$. As a result, x^* lies in E , and setting $\bar{y} := y_1 + y_2$ we obtain $\gamma(x, y) = -\bar{y}(x - x^*)$. Moreover, since $\bar{y} = \frac{y_2}{x^*} = \frac{y_1}{1-x^*}$ μ -a.s., we can conclude that it is square-integrable and takes values in the set $[0, (x^* \vee (1 - x^*))^{-1}]$ μ -a.s. By (D.1) we can then write

$$\lambda(x) = \frac{r_3(x)}{p_3(x)} = \frac{\bar{r}_3(x)}{(x - x^*)^3}$$

for some $\bar{r}_3 \in \text{Pol}_3(E)$, for all $x \in E \setminus \{x^*\}$. Since in this case $\nu(x^*, \cdot) = 0$ we are free to choose $\lambda(x^*) = 0$. By Lemma 4.1 we also know that r_2 is bounded on E . Therefore, noting that

$$r_2(x) = \frac{\bar{r}_3(x)}{(x - x^*)} \int (-\bar{y})^2 \mu(d\bar{y}) \quad \text{for all } x \in E \setminus \{x^*\},$$

it follows that $\bar{r}_3(x^*) = 0$, and thus $\lambda(x) = \frac{q_2(x)}{(x - x^*)^2}$ for some $q_2 \in \text{Pol}_2(E)$ and all in $x \in E \setminus \{x^*\}$. This in particular implies that $r_2 \in \text{Pol}_2(E \setminus \{x^*\})$ and hence $a \in \text{Pol}_2(E \setminus \{x^*\})$. Knowing that $a + r_2$ has to be continuous by condition (iii) of Lemma 4.1, we can finally deduce that

$$a(x^*) = \lim_{x \rightarrow x^*} a(x) + q_2(x^*) \int \bar{y}^2 \mu(d\bar{y}).$$

Suppose now $x^* \in \{0, 1\}$. Then, since $a \geq 0$ on E and $a(0) = a(1) = 0$, we conclude that $q_2(x^*) = 0$, $a(x) = Ax(1 - x)$ for some $A \in \mathbb{R}_+$, and thus $\lambda(x) = \frac{q_1(x)}{(x - x^*)} \mathbf{1}_{\{x \neq x^*\}}$ for some $q_1 \in \text{Pol}_1(E)$. For $x^* = 0$, respectively $x^* = 1$, the nonnegativity of λ implies that $q_1(x) = 1 + qx$, respectively $q_1(x) = 1 + q(1 - x)$, for some $q \in [-1, \infty)$. As a result, \mathcal{G} is an operator of Type 2.

On the other hand, if $x^* \in (0, 1)$, using again that $a \geq 0$ on E and $a(0) = a(1) = 0$, we conclude that

$$a(x) = Ax(1 - x) + \left(q_2(x^*) \int \bar{y}^2 \mu(d\bar{y}) \right) \mathbf{1}_{\{x = x^*\}},$$

for some $A \in \mathbb{R}_+$, proving that \mathcal{G} is an operator of Type 3.

Assume now that $\nu(x, d\xi) \neq 0$, λ cannot be chosen to be constant, and $p_4(x) \neq 0$ for all $x \in \mathbb{R}$. We must argue that \mathcal{G} is then necessarily of Type 4. By (D.1) we have $\lambda(x) = \frac{r_4(x)}{p_4(x)}$ on E , and thus on \mathbb{R} . Consequently, λ is locally bounded, nonnegative, and non-constant. Moreover, (D.1) yields the expression $\lambda(x) = \frac{r_3(x)}{p_3(x)}$ for all $x \in E$ with $p_3(x) \neq 0$. These facts combined with the fundamental theorem of algebra imply that

$$\lambda(x) = \frac{q_2(x)}{(x - \alpha)(x - \bar{\alpha})} \tag{D.3}$$

for some positive $q_2 \in \text{Pol}_2(E)$ and $\alpha \in \mathbb{C} \setminus \mathbb{R}$. Without losing generality we choose to satisfy $\text{Im}(\alpha) < 0$. Note that no further cancellation of polynomial

factors is possible in (D.3) since λ is non-constant. Furthermore, condition (iii) of Lemma 4.1 yields $r_n \in \text{Pol}_n(E)$ for all $n \geq 3$. Therefore, since $p_n(x)q_2(x) = (x - \alpha)(x - \bar{\alpha})r_n(x)$ due to (D.1) and (D.3), it follows that $p_n(x) = (x - \alpha)(x - \bar{\alpha})R_{n-2}(x)$ for all $x \in E$ and $n \geq 3$, for some $R_{n-2} \in \text{Pol}_{n-2}(E)$. This already implies (4.2) for $n \geq 3$, i.e.,

$$p_n(\alpha) = \int \gamma^n(\alpha, y)\mu(dy) = \int (y_1(-\alpha) + y_2(1 - \alpha))^n \mu(dy) = 0, \quad n \geq 3. \quad (\text{D.4})$$

Next, we will establish (4.2) also for $n = 2$. In preparation for an application of Lemma D.1 later on, choose a constant C_α such that

$$\left| 1 + \frac{i\gamma(\alpha, y)}{C_\alpha} \right| < 1 \quad \text{and} \quad \frac{2|\gamma(\alpha, y)|}{C_\alpha} < \tan^{-1} \left| \frac{\text{Im}(1 - \alpha)}{\text{Re}(1 - \alpha)} \right| \quad (\text{D.5})$$

for all $y \in [0, 1]^2 \setminus \{0\}$. Define

$$f_k(z) := 1 - \left(1 + \frac{iz}{C_\alpha} \right)^k,$$

and note that $|f_k(\gamma(\alpha, y))| \leq 2$ and $\lim_{k \rightarrow \infty} f_k(\gamma(\alpha, y)) = 1$ for all $y \in [0, 1]^2 \setminus \{0\}$. By dominated convergence we then obtain

$$\int \gamma^2(\alpha, y)\mu(dy) = \lim_{k \rightarrow \infty} \int \gamma^2(\alpha, y)f_k(\gamma(\alpha, y))\mu(dy) = 0, \quad (\text{D.6})$$

where the last equality follows since, for each $k \geq 3$, the integral on the right-hand side is a linear combination of $\int \gamma^n(\alpha, y)\mu(dy)$ with $3 \leq n \leq k$, and therefore vanishes due to (D.4). Hence, (4.2) holds also for $n = 2$.

We now derive some consequences. First, $r_2 \in \text{Pol}_2(E)$ and hence $a(x) = Ax(1 - x)$ for some $A \in \mathbb{R}_+$. Moreover, since

$$\text{Arg}(\gamma^2(\alpha, \cdot)) \subseteq [2\text{Arg}(1 - \alpha), 2\text{Arg}(-\alpha)]$$

holds μ -a.s., (D.6) implies that $\text{Arg}(-\alpha) - \text{Arg}(1 - \alpha) \geq \pi/2$, or equivalently, $|2\alpha - 1| \leq 1$. In the description of Type 4 it is claimed in addition that the inequality is strict, i.e.

$$|2\alpha - 1| < 1. \quad (\text{D.7})$$

To see this, observe that in case of equality, (D.6) would imply that $\text{Arg}(\gamma(\alpha, y)) \subseteq \{\text{Arg}(-\alpha), \text{Arg}(1 - \alpha)\}$ for μ -a.e. $y \in [0, 1]^2$, which is clearly incompatible with having $\int \gamma^3(\alpha, y)\mu(dy) = 0$ for some nontrivial measure μ .

Next, we claim that $\int y_1\mu(dy) = \int y_2\mu(dy) = \infty$. We prove this by excluding the complementary possibilities. First, assume for contradiction that $\int y_1\mu(dy) < \infty$ and $\int y_2\mu(dy) < \infty$. Then $\int |\gamma(\alpha, y)|\mu(dy) < \infty$. Proceeding as with (D.6) we then deduce

$$\int \gamma(\alpha, y)\mu(dy) = \lim_{k \rightarrow \infty} \int \gamma(\alpha, y)f_k(\gamma(\alpha, y))\mu(dy) = 0,$$

which is clearly not possible since $\text{Im}(\gamma(\alpha, y)) > 0$ μ -a.s. This is the desired contradiction.

Suppose instead $\int y_1 \mu(dy) < \infty$ and $\int y_2 \mu(dy) = \infty$. Define

$$g_k(y) := \text{Re}(\gamma(\alpha, y) f_k(\gamma(\alpha, y))). \quad (\text{D.8})$$

Set $C := \text{Im}(1 - \alpha)(\text{Re}(1 - \alpha))^{-1}$ and observe that $C > 0$ due to the fact that $\text{Re}(1 - \alpha) > 0$ in view of (D.7). Then define the set

$$A := \left\{ y \in [0, 1]^2 : \frac{\text{Im}(\gamma(\alpha, y)/C_\alpha)}{\text{Re}(\gamma(\alpha, y)/C_\alpha)} \in [C, 2C], \right. \\ \left. \frac{2|\gamma(\alpha, y)|}{C_\alpha} \leq \tan^{-1}(C), \text{Im}\left(\frac{\gamma(\alpha, y)}{C_\alpha}\right) \geq 0 \right\}.$$

Choosing $\varepsilon > 0$ small enough such that $\{y_1 < \varepsilon y_2\} \cap [0, 1]^2 \subseteq A$, by Lemma D.1 we have that

$$\{y_1 < \varepsilon y_2\} \cap [0, 1]^2 \subseteq A \subseteq \left\{ h_k\left(\frac{\gamma(\alpha, y)}{C_\alpha}\right) \geq 0 \right\} = \left\{ \frac{g_k(y)}{C_\alpha} \geq 0 \right\} = \{g_k(y) \geq 0\}.$$

We can then compute

$$g_k(y) \geq -2|\gamma(\alpha, y)| \mathbb{1}_{\{g_k(y) < 0\}} \geq -2|\gamma(\alpha, y)| \mathbb{1}_{\{y_1 \geq \varepsilon y_2\}} \geq -2(y_1|\alpha| + \varepsilon^{-1}y_1|1 - \alpha|),$$

for all $k \in \mathbb{N}$ and μ -a.e. $y \in [0, 1]^2$. Fatou's lemma then yields

$$0 = \lim_{k \rightarrow \infty} \int g_k(y) \mu(dy) \geq -\text{Re}(\alpha) \int y_1 \mu(dy) + \text{Re}(1 - \alpha) \int y_2 \mu(dy) = \infty,$$

using in the last step that $\text{Re}(1 - \alpha) > 0$. Again we arrive at a contradiction.

Finally, suppose $\int y_1 \mu(dy) = \infty$ and $\int y_2 \mu(dy) < \infty$. We may then repeat the arguments from the first case, using the function $-g_k$ instead of g_k to obtain the required contradiction. In summary, we have shown that $\int y_1 \mu(dy) = \int y_2 \mu(dy) = \infty$, as claimed.

Finally, the boundary conditions (D.2) now forces $\lambda(0) = \lambda(1) = 0$, which in view of (D.3) yields $q_2(0) = q_2(1) = 0$. Therefore $q_2(x) = Lx(1 - x)$ for some constant $L > 0$, as claimed. As a result, \mathcal{G} is an operator of Type 4, and the proof of Theorem 4.3 is complete. \square

Lemma D.1. Fix $C > 0$ and set $h_k(z) := \text{Re}(z(1 - (1 + iz)^k))$ for all $k \in \mathbb{N}$. Then there is some $K \in \mathbb{N}$ such that

$$\left\{ \text{Im}(z)/\text{Re}(z) \in [C, 2C], 2|z| \leq \tan^{-1}(C), \text{Im}(z) \geq 0 \right\} \subseteq \{h_k(z) \geq 0\}$$

for all $k \geq K$.

Proof. Fix $c \in [C, 2C]$ and let $z \in \mathbb{C}$ such that $\text{Im}(z) = c\text{Re}(z)$ and $\text{Im}(z) \geq 0$. Define then $w := 1 + iz$ and compute

$$h_k(z) = h_k(i - iw) = \text{Im}(w)(1 - \text{Re}(w^k) + c\text{Im}(w^k)).$$

Let $x := \text{Arg}(w)$, note that $x = \text{Arg}(1 + iz) \in [0, 2|z|]$ and moreover

$$\text{Re}(w^k) - c\text{Im}(w^k) = \frac{\cos(kx) - c\sin(kx)}{(c\sin(x) + \cos(x))^k} = \frac{\sqrt{1+c^2}\cos(kx + \tan^{-1}(c))}{(\sqrt{1+c^2}\cos(x - \tan^{-1}(c)))^k}. \quad (\text{D.9})$$

Since $\text{Im}(w) = \text{Re}(z) \geq 0$, it is then enough to show that for k big enough this expression is smaller than or equal to 1 for all $x \in [0, \tan^{-1}(C)]$ and $c \in [C, 2C]$.

Let $x_k^c := (\pi - \tan^{-1}(c))/k$ be the first minimum of the numerator. Observe that for $x = x_k^c$ the denominator converges to $\exp(c(\pi - \tan^{-1}(c))) > \sqrt{1+c^2}$ uniformly on compact sets. As a result, for k big enough, we have that

$$\sup_{c \in [C, 2C]} \frac{\cos(kx) - c\sin(kx)}{(c\sin(x) + \cos(x))^k} \leq \sup_{c \in [C, 2C]} \frac{\cos(kx_k^c) - c\sin(kx_k^c)}{(c\sin(x_k^c) + \cos(x_k^c))^k} \leq 1,$$

for all $x \in [x_k^c, \tan^{-1}(C)]$. Since expression (D.9) takes value 1 for $x = 0$ and is decreasing in x on $[0, x_k^c]$, we conclude the proof. \square

E The unit simplex: Proof of Lemma 6.1 and Theorem 6.3

Proof of Lemma 6.1. Assume \mathcal{G} is a polynomial operator and its martingale problem is well-posed. Theorem 2.3 and Theorem 2.8 imply that \mathcal{G} is of Lévy type for some triplet (a, b, ν) , so that in particular (i) holds. Condition (iv) follows from Lemma 2.9. To prove (ii), fix $x \in E \cap \{x_i = 0\}$. Let $g_n^i(x) := g_n(x_i)$ and $h_n^i(x) := h_n(x_i)$, where g_n and h_n are the functions on $[0, 1]$ described in the proof of Lemma 4.1. Then by the positive maximum principle we conclude that

$$0 \leq \mathcal{G}g_n^i(x) \rightarrow -a_{ii}(x) \quad \text{and} \quad 0 \leq \mathcal{G}h_n^i(x) \rightarrow b_i(x) - \int \xi_i \nu(x, d\xi).$$

The positive semidefiniteness of $a(x)$ then implies that $a_{ij}(x) = 0$ for all $j \in \{1, \dots, d\}$. In order to verify (iii), note that setting $f^\Delta(x) := x^\top \mathbf{1} - 1$, by the positive maximum principle we have

$$0 = \mathcal{G}(f^\Delta)(x) = b(x)^\top \mathbf{1} \quad \text{and} \quad 0 = \mathcal{G}((\cdot - x)^{e_j} f^\Delta)(x) = a_j(x)^\top \mathbf{1},$$

where $a_j(x)$ denotes the j -th column of $a(x)$.

Conversely, Lemma 2.9 and (iv) imply that \mathcal{G} is polynomial. Thus by Theorem 2.3, the martingale problem for \mathcal{G} is well-posed, provided that $\mathcal{G}\mathbf{1} = 0$ and \mathcal{G} satisfies the positive maximum principle. The first condition is clearly satisfied. For the second one, let $g^i(x) := x_i$ and $f \in \text{Pol}(E)$ be an arbitrary polynomial having a maximum over E at some $x \in E$. Observe that

$$E = \{x \in \mathbb{R}^d : f^\Delta = 0 \text{ and } g^i \geq 0 \text{ for all } i \in \{1, \dots, d\}\},$$

and let $I(x)$ be the set of all active inequality constraints at point x , that is, $I(x)$ is the set of all $i \in \{1, \dots, d\}$ such that $x_i = 0$. By the necessity of the Karush-Kuhn-Tucker conditions (see e.g. Proposition 3.3.1 in Bertsekas (1995)), there exist multipliers $\mu \in \mathbb{R}_+^d$ and $\lambda \in \mathbb{R}$ such that $\mu_i = 0$ for all $i \in \{1, \dots, d\} \setminus I(x)$,

$$\nabla f(x) = - \sum_{i \in I(x)} \mu_i \nabla g^i(x) + \lambda \nabla f^\Delta(x) = - \sum_{i \in I(x)} \mu_i e_i + \lambda \mathbf{1},$$

and $v^\top \nabla^2 f(x) v \leq 0$ for all $v \in \mathbb{R}^d$ such that $v^\top \mathbf{1} = 0$ and $v_i = 0$ for all $i \in I(x)$. Since $\xi^\top \mathbf{1} = 0$ for $\nu(x, \cdot)$ -a.e. ξ , $b(x)^\top \mathbf{1} = 0$ by (iii), and $\int |\xi_i| \nu(x, d\xi) < \infty$ for all $i \in I(x)$ by (ii), we can thus write

$$\begin{aligned} \mathcal{G}f(x) &= \frac{1}{2} \text{Tr} (a(x) \nabla^2 f(x)) + \sum_{i \in I(x)} -\mu_i \left(b_i(x) - \int \xi_i \nu(x, d\xi) \right) \\ &\quad + \int (f(x + \xi) - f(x)) \nu(x, d\xi). \end{aligned}$$

We must argue that $\mathcal{G}f(x) \leq 0$. The second term on the right-hand side is nonpositive by (ii). The last term is also nonpositive since f is maximal over E at x . It remains to show that the first term is nonpositive. To this end, let $\sqrt{a(x)}$ denote the symmetric and positive semidefinite square root of $a(x)$. Condition (iii) yields $a(x)\mathbf{1} = 0$ and thus $\sqrt{a(x)}\mathbf{1} = 0$. By symmetry of $\sqrt{a(x)}$ we deduce

$$(\sqrt{a(x)}v)^\top \mathbf{1} = v^\top \sqrt{a(x)}\mathbf{1} = 0 \quad \text{for all } v \in \mathbb{R}^d.$$

Moreover, by (ii) we also have that $a(x)e_i = 0$, and hence $\sqrt{a(x)}e_i = 0$, for all $i \in I(x)$. This implies that $(\sqrt{a(x)})_{ij} = 0$ and thus $(\sqrt{a(x)}v)_i = 0$, for all $i \in I(x)$ and $v \in \mathbb{R}^d$. As a result,

$$v^\top (\sqrt{a(x)} \nabla^2 f(x) \sqrt{a(x)}) v = (\sqrt{a(x)}v)^\top \nabla^2 f(x) (\sqrt{a(x)}v) \leq 0,$$

which implies that $\sqrt{a(x)} \nabla^2 f(x) \sqrt{a(x)}$ is negative semidefinite. This gives the desired inequality

$$\text{Tr}(a(x) \nabla^2 f(x)) = \text{Tr}(\sqrt{a(x)} \nabla^2 f(x) \sqrt{a(x)}) \leq 0,$$

showing that $\mathcal{G}f(x) \leq 0$. This completes the proof of the lemma. \square

Before starting the proof of Theorem 6.3, we prove three auxiliary lemmas.

Lemma E.1. *Consider a polynomial $p \in \text{Pol}_n(E)$.*

(a) *If p vanishes on $E \cap \{x_i = x_j = 0\}$, it can be written as*

$$p(x) = x_i p_{n-1}^i(x) + x_j p_{n-1}^j(x) \quad \text{for some } p_{n-1}^i, p_{n-1}^j \in \text{Pol}_{n-1}(E). \quad (\text{E.1})$$

(b) *If p vanishes on $E \cap (\{x_i = 0\} \cup \{x_j = 0\})$ for some $i \neq j$, it can be written as*

$$p(x) = x_i x_j p_{n-2}(x) \quad \text{for some } p_{n-2} \in \text{Pol}_{n-2}(E). \quad (\text{E.2})$$

(c) If p vanishes on $E \cap \{cx_i = x_j\}$ for some $c \geq 0$ and $i \neq j$, it can be written as

$$p(x) = (-cx_i + x_j)p_{n-1}(x) \quad \text{for some } p_{n-1} \in \text{Pol}_{n-1}(E). \quad (\text{E.3})$$

Proof. Since every affine function on E can be written as a linear one, there is a real collection $(p_{\mathbf{n}})_{|\mathbf{n}|=n}$ such that $p(x) = \sum_{|\mathbf{n}|=n} p_{\mathbf{n}}x^{\mathbf{n}}$, for all $x \in E$. Observe that for all $x \in E \cap \{x_i = x_j = 0\}$ we have that

$$0 = p(x) = \sum_{\substack{|\mathbf{n}|=n \\ n_i=n_j=0}} p_{\mathbf{n}}x^{\mathbf{n}}.$$

Assume without loss of generality that $i = d$ and $j = d - 1$ (resp. $i = j = d$ if $i = j$) and note that, the polynomial $q \in \text{Pol}(\mathbb{R}^{j-1})$ given by

$$q(x) := \sum_{\substack{|\mathbf{n}|=n \\ n_i=n_j=0}} p_{\mathbf{n}}x^{\bar{\mathbf{n}}},$$

where $\bar{\mathbf{n}} = \sum_{k=1}^{j-1} n_k e_k$, is a homogeneous polynomial vanishing on the simplex. This directly implies that $q = 0$ and hence $p_{\mathbf{n}} = 0$ for all $|\mathbf{n}| = n$ such that $n_i = n_j = 0$. We can thus conclude that p satisfies (E.1).

Proceeding as before for the second part, we obtain that $p_{\mathbf{n}} = 0$ for all $|\mathbf{n}| = n$ such that $n_i = 0$ or $n_j = 0$ and can thus conclude that p satisfies (E.2).

Finally, for the third part it is enough to note that the polynomial $\bar{p} \in \text{Pol}_n(E)$ given by

$$\bar{p}(x) := p\left(x + \frac{cx_i}{1+c}(e_j - e_i)\right)$$

vanishes on $E \cap \{x_j = 0\}$. By (a) this gives us that

$$p(x) = \bar{p}(x + cx_i(e_i - e_j)) = (-cx_i + x_j)\bar{p}_{n-1}^j(x)$$

on $E \cap \{x_j \geq cx_i\}$ and thus on E , proving that p satisfies (E.3). \square

Lemma E.2. Let μ be a nonzero measure on $(\Delta^d)^d \setminus \{e_1, \dots, e_d\}$. The function $\gamma : E \times (\Delta^d)^d \rightarrow \mathbb{R}$ given by $\gamma(x, y) = \sum_{i=1}^d (y^i - e_i)x_i$ can be represented as

$$\gamma(x, y) = H(y)P_1(x) \quad \mu\text{-a.s.}, \quad (\text{E.4})$$

for a measurable function $H : (\Delta^d)^d \rightarrow \mathbb{R}^d$ and $P_1 \in \text{Pol}_1(E)$, if and only if one of the following cases holds true

(a) $\gamma(x, y) = (y^i - e_i)x_i$ μ -a.s. for some $i \in \{1, \dots, d\}$.

(b) $\gamma(x, y) = y_i^j(-cx_i + x_j)(e_i - e_j)$ μ -a.s. for some $i \neq j$ and $c > 0$. In this case $y_i^j \in (0, \frac{1}{c} \wedge 1]$ μ -a.s.

Proof. First assume that (E.4) holds true. Since $P_1 \in \text{Pol}_1(E)$, and as every affine function on E has a linear representation, we can write $P_1(x) = C^\top x$, for some $C \in \mathbb{R}^d$. If $C = 0$, the support of μ has to be contained in $\{e_1, \dots, e_d\}$, which is not possible by assumption.

If $C_i \neq 0$ for some $i \in \{1, \dots, d\}$, item (a) follows if $C_j = 0$ for all $j \neq i$.

If C_i and C_j are nonzero for some $i \neq j$, item (b) follows if $C_\ell = 0$ for all $\ell \notin \{i, j\}$. Indeed, by assumption we have $(y^k - e_k) = C_k H(y)$ for $k \in \{i, j\}$ and thus

$$y^i - e_i = \frac{C_i}{C_j}(y^j - e_j).$$

Since $y^i, y^j \in \Delta^d$ μ -a.s., we can conclude that $y_\ell^i = y_\ell^j = 0$ for all $\ell \notin \{i, j\}$ and hence

$$\frac{y_j^i}{y_i^j} = \frac{1 - y_i^i}{y_i^j} = -\frac{C_i}{C_j} =: c \quad (\text{E.5})$$

proving that the conditions of item (b) hold true. In this case $y_i^j \in (0, \frac{1}{c} \wedge 1]$ μ -a.s. by (E.5).

Finally, if $C_i \neq 0$ for at least three different values of i , the same reasoning as for case (b) implies $y_\ell^i = 0$ for all $\ell \neq i$ and thus $H = 0$ μ -a.s., which is not possible by assumption.

The converse direction is clear. \square

Lemma E.3. *The following assertions are equivalent.*

- (i) *The matrix $a(x) \in \mathbb{S}_+^d$ satisfies $a\mathbf{1} = 0$, $a_{ij} \in \text{Pol}_2(E)$, and $a_{ii} = 0$ on $E \cap \{x_i = 0\}$.*
- (ii) *The matrix $a(x)$ satisfies*

$$a_{ii}(x) = \sum_{i \neq j} \alpha_{ij} x_i x_j \quad \text{and} \quad a_{ij}(x) = -\alpha_{ij} x_i x_j \quad \text{for all } i \neq j,$$

for some $\alpha_{ij} = \alpha_{ji} \in \mathbb{R}_+$.

Proof. We start by proving (i) \Rightarrow (ii): By Lemma E.1 we already know that for all $i \neq j$ we have $a_{ij} = -\alpha_{ij} x_i x_j$ for some $\alpha_{ij} \in \mathbb{R}$. Moreover, as $a\mathbf{1} = 0$ on E , we also have that

$$a_{ii}(x) = -\sum_{j \neq i} a_{ij}(x) = \sum_{i \neq j} \alpha_{ij} x_i x_j$$

for all $x \in E$. Since $a \in \mathbb{S}_+^d$ on E and $\alpha_{ij} = 4a_{ii}((e_i + e_j)/2)$ for all $i \neq j$, it follows that $\alpha_{ij} \in \mathbb{R}_+$, which finishes the proof of the first direction. Concerning (ii) \Rightarrow (i), the only condition which is not obvious is positive semidefiniteness of a on E , which follows exactly as in the proof of Proposition 6.6 in Filipović and Larsson (2016). \square

Proof of Theorem 6.3. As \mathcal{G} is polynomial and its martingale problem is well-posed, the conditions of Lemma 6.1 are satisfied. As in Remark 3.2, we can assume without loss of generality that μ is compactly supported and all its moments of order greater or equal two are thus finite. Analogously to (D.1) we set then

$$p_{\mathbf{n}} := \int \gamma^{\mathbf{n}}(\cdot, y) \mu(dy) \quad \text{and} \quad r_{\mathbf{n}} := \int \xi^{\mathbf{n}} \nu(\cdot, d\xi) = \lambda p_{\mathbf{n}} \quad (\text{E.6})$$

for all $|\mathbf{n}| \geq 2$. Note that $p_{\mathbf{n}} \in \text{Pol}_{|\mathbf{n}|}(E)$ for all $|\mathbf{n}| \geq 2$ by the integrability conditions on μ . By condition (iv) of Lemma 6.1 we also have that $r_{\mathbf{n}} \in \text{Pol}_{|\mathbf{n}|}(E)$ for all $|\mathbf{n}| \geq 3$. By Remark 3.4 we know that

$$\lambda(x) = \frac{\int |\xi|^4 \nu(x, d\xi)}{\int |\gamma(x, y)|^4 \mu(dy)} \mathbb{1}_{\{\int |\gamma(x, y)|^4 \mu(dy) \neq 0\}},$$

and hence condition $\nu(x, (E - x)^c) = 0$ implies that μ can be chosen supported on $(\Delta^d)^d$ and such that

$$\gamma(x, y) = \sum_{i=1}^d (y^i - e_i) x_i \quad \mu\text{-a.s.}$$

By definition of affine jump sizes, the measure μ has to be square-integrable.

Concerning the statement on the drift this is a consequence of Lemma 6.1. Indeed (iv) yields the affine (and thus linear) form of the drift, (ii) leads to

$$\sum_{j \neq i} \left(B_{ij} x_j - \lambda(x) \int y_i^j x_j \mu(dy) \right) \geq 0, \quad x \in E \cap \{x_i = 0\} \quad (\text{E.7})$$

and finally $B^\top \mathbf{1} = 0$ is a consequence of (iii), namely $b^\top \mathbf{1} = 0$ for all $x \in E$. Since condition (E.7) yields $\sum_{j \neq i} B_{ij} x_j \geq 0$, choosing $x = e_j$ we get $B_{ij} \geq 0$ for $j \neq i$ and $B_{ii} = -\sum_{j \neq i} B_{ji}$ for all $i \in \{1, \dots, d\}$. We will now consider four complementary assumptions, which will lead to Type 0 to 3.

Assume that $\nu(x, d\xi) = 0$. Then by Lemma 6.1 we can apply Lemma E.3 to conclude that a satisfies (6.2). This proves that in this case \mathcal{G} is an operator of Type 0.

Assume now that $\nu(x, d\xi) \neq 0$ and λ can be chosen to be constant. We can then without loss of generality set $\lambda = 1$. Moreover, since in this case $r_{e_i + e_j} \in \text{Pol}_2(E)$ for all $i, j \in \{1, \dots, d\}$, condition (iv) of Lemma 6.1 implies that the entries diffusion matrix $a_{ij} \in \text{Pol}_2(E)$. We can thus conclude as before that a can be chosen to be of form (6.2). Finally, condition (E.7) can be rewritten as $\sum_{j \neq i} (B_{ij} - \int y_i^j \mu(dy)) x_j \geq 0$ for all $x \in E \cap \{x_i = 0\}$, which yields $B_{ij} - \int y_i^j \mu(dy)$ for all $i \neq j$. As a result, \mathcal{G} is an operator of Type 1.

Assume now that $\nu(x, d\xi) \neq 0$ and λ cannot be chosen to be constant. We already know that $\lambda = \frac{p(x)}{q(x)}$ for some $p, q \in \text{Pol}(E)$. Supposing without loss of generality that $p(x)$ and $q(x)$ are coprime polynomials, we necessarily have due to Assumption A that $q(x)$ is a divisor of $\gamma_i(x, y)^3$ for each $i \in \{1, \dots, d\}$

and μ -a.e. $y \in (\Delta^d)^d$. Since $\gamma_i(\cdot, y) \in \text{Pol}_1(E)$ μ -a.s., this in turn implies that $\gamma(x, y) = H(y)P_1(x)$ μ -a.s. with a measurable function $H : (\Delta^d)^d \rightarrow \mathbb{R}^d$ and $P_1 \in \text{Pol}_1(E)$.

Choose now $i \in \{1, \dots, d\}$ such that $\mu(H_i(y) \neq 0) > 0$. By equation (E.6) we have that

$$\lambda(x) = \frac{r_{3e_i}(x)}{p_{3e_i}(x)} = \frac{\bar{r}_{3e_i}(x)}{(P_1(x))^3}$$

for some $\bar{r}_{3e_i} \in \text{Pol}_3(E)$, for all $x \in E \setminus \{P_1 = 0\}$. Since in this case $\nu(x, d\xi) = 0$ for all $x \in E \cap \{P_1 = 0\}$, we are free to choose $\lambda(x) = 0$ on this set. By Lemma 6.1 we also know that r_{2e_i} has to be a bounded function on E . Noting that for all $x \in E \setminus \{P_1 = 0\}$

$$r_{2e_i} = \frac{\bar{r}_{3e_i}(x)}{P_1(x)} \int H_i^2(y) \mu(dy),$$

we see that this condition holds true if and only if $\bar{r}_{3e_i}(x) = 0$ for all $x \in \{P_1 = 0\}$. Since we know by Lemma E.2 that $P_1(x) = -cx_i + x_j$ for some $c \geq 0$, by Lemma E.1 we thus have that

$$\lambda(x) = \frac{q_2(x)}{(P_1(x))^2} \mathbb{1}_{\{P_1 \neq 0\}}$$

for some $q_2 \in \text{Pol}_2(E)$. This in particular implies that $r_{e_k+e_\ell} \in \text{Pol}_2(E \setminus \{P_1 = 0\})$ and hence, by condition (iv) of Lemma 6.1, $a_{k\ell} \in \text{Pol}_2(E \setminus \{P_1 = 0\})$ for all $k, \ell \in \{1, \dots, d\}$. By the same condition we also have that for all $x \in E \cap \{P_1 = 0\}$

$$a_{k\ell}(x) = a_{k\ell}(x) + r_{e_k+e_\ell}(x) = \lim_{z \rightarrow x} a_{k\ell}(z) + q_2(x) \int H_k(y) H_\ell(y) \mu(dy), \quad (\text{E.8})$$

and thus in particular, by positive semidefiniteness of $a(x)$ and condition (ii) of Lemma 6.1,

$$a_{ii}(x) = \lim_{z \rightarrow x} a_{ii}(z) = q_2(x) = 0 \quad (\text{E.9})$$

for all $x \in E \cap \{P_1 = 0\} \cap \{x_i = 0\}$. Setting $a_{k\ell}^c \in \text{Pol}_2(E)$ be such that $a_{k\ell}^c|_{E \setminus \{P_1=0\}} = a_{k\ell}|_{E \setminus \{P_1=0\}}$, we obtain that $a^c := (a_{k\ell}^c)_{k\ell}$ satisfies the conditions of Lemma E.3 and thus a^c can be chosen to be of the form (6.2). By (E.8) we can then conclude that

$$a(x) = a^c(x) + q_2(x) \mathbb{1}_{\{P_1=0\}} \int H(y) H(y)^\top \mu(dy).$$

By Lemma E.2, we know that there are only two complementary choices of H and P_1 .

The first choice is $H(y) = (y^i - e_i)$ and $P_1(x) = x_i$ for some fixed $i \in \{1, \dots, d\}$. Then by (E.9) and Lemma E.1 we have $q_2(x) = q_1(x)x_i$ for some $q_1 \in \text{Pol}_1(E)$. Moreover, using that $q_1(x) = \sum_{j=1}^d q_1(e_j)x_j$, condition (E.7) can be rewritten as

$$\sum_{j \neq k} \left(B_{kj} - q_1(e_j) \mathbb{1}_{\{x_i \neq 0\}} \int y_k^i \mu(dy) \right) x_j \geq 0, \quad x \in E \cap \{x_k = 0\}$$

for all $k \in \{1, \dots, d\}$, which yields (6.5) for all $k \neq i$ and $j \neq k$. As a result, \mathcal{G} is an operator of Type 2.

The second choice of H and P_1 is $H(y) = y_i^j(e_i - e_j)$ and $P_1(x) = -cx_i + x_j$ for some $i, j \in \{1, \dots, d\}$ such that $i \neq j$, where $y_i^j \in (0, \frac{1}{c} \wedge 1]$. Then by (E.9) and Lemma E.1 we have $q_2(x) = \sum_{k=1}^d q_{ik}x_ix_k + q_{jk}x_jx_k$ for some $q_{kl} \in \mathbb{R}$. Since condition (E.7) coincides with condition (6.6), we can conclude that \mathcal{G} is an operator of Type 3. \square

Chapter III

Probability measure-valued jump-diffusions

1 Introduction

Probability measure-valued processes have been studied for many years in many different contexts. Among the most famous ones, one can cite Fleming and Viot ([Fleming and Viot, 1979](#)) who introduced the well-known homonym process, or Dawson and Watanabe ([Watanabe, 1968](#); [Dawson, 1977, 1978](#)): the significance of their works is reflected in the choice to give their names to the corresponding superprocesses. After them, a wide theory grew and the field was explored in many different directions. Among many others, there is a very remarkable collection of lecture notes of the famous summer school in St. Flour by Sznitman, Dawson, and Perkins ([Sznitman, 1991](#); [Dawson, 1993](#); [Perkins, 2002](#)), a deep and rigorous study done by Ethier and Kurtz ([Ethier and Kurtz, 1987, 1993, 2005](#)), but also the works of many other prominent authors like Etheridge, Hochberg, Vaillancourt, or Xiong.

As explained in the general introduction, the applications in this respect are rich and range from population genetics, interacting particle systems, stochastic partial differential equations, and statistical physics. From a mathematical finance perspective, one can for instance be interested in modeling high dimensional financial markets involving many assets in a tractable and robust way. By the notion of high dimensional financial markets we mean for instance term structures in fixed income, commodities or electricity markets involving potentially an uncountably infinite number of assets. But also joint stochastic modeling of a large finite number of stocks or market capitalizations constituting the major indices (e.g., 500 in the case of S&P 500) falls in this category. This is an essential task, in particular in the area of stochastic portfolio theory ([Fernholz, 2002](#); [Fernholz and Karatzas, 2009](#)). Measure-valued processes can serve as universal models to describe the time evolution of a large number or a continuum of assets, market weights, zero-coupon bonds, etc. In this way, high dimensionality can be captured using *one* measure valued process.

Let us be more precise on this construction: let E be a locally compact

Polish space and denote by \mathcal{C} some finite or infinite subset of bounded continuous functions. Consider a process X taking values in $M_+(E)$, the set of finite, positive measures on E . Then asset prices (or market capitalizations) can be defined via

$$S_t^g := \int_E g(x) X_t(dx), \quad g \in \mathcal{C}. \quad (1.1)$$

Similarly, in the context of stochastic portfolio theory the market weights obtained from market capitalizations by normalizing with the total market capitalization can be modeled by

$$\mu_t^g := \int_E g(x) X_t(dx), \quad (1.2)$$

where now X takes values in $M_1(E)$, the set of probability measures on E , and g is, for instance, an indicator function $\mathbb{1}_{A^k}$, where $(A^k)_{k=1}^d$ is a partition of E .

The latter example already indicates that a particular relevant subclass are *probability measure-valued processes*. Apart from modeling market weights, such processes can be used to describe the empirical measure of the capitalizations,

$$\frac{1}{|\mathcal{C}|} \sum_{g \in \mathcal{C}} \delta_{S_t^g}. \quad (1.3)$$

The empirical measure is important for the study of capital distribution curves (Fernholz, 2002; Shkolnikov, 2013). Analyzing (1.3) as the number $|\mathcal{C}|$ of assets tends to infinity allows to draw conclusions on the shape and fluctuations of the implied capital distribution curves for large markets.

Another area of mathematical finance which is closely related to probability measure-valued jump-diffusions is given by mean field games. This field plays an important and constantly growing role in a wide range of applications. In particular, examples include models of population growth, computational methods in probability and stochastic processes (including simulation), genetics and other stochastic models in biology and the life sciences, stochastic control, stochastic models in stochastic optimization, stochastic models in the physical sciences (see e.g. Carmona and Delarue (2017) for theory and applications).

As explained in the introduction, our approach is based on the martingale problem and consists in directly characterizing the positive maximum principle of the corresponding generators and in providing some general optimality conditions. Chapter II provides the guidelines needed for doing this in practice. The present chapter is organized as follows. Section 2 summarizes some notation and provide the mathematical background needed throughout this chapter and Chapter IV. Section 3 treats the martingale problem, its connection with the positive maximum principle in the context of existence of solutions to the martingale problem. In that section, we also illustrate how some properties of a jump-diffusion can be read directly from the corresponding generator. Section 4 is concerned with Lévy type operators and Section 5 with optimality conditions for functions of polynomial argument. The proof of Theorem 4.2 and those of some technical lemmas are gathered in appendices.

2 Preliminaries

The aim of this section is to introduce the mathematical background needed to deal with probability measure-valued jump-diffusions. As explained in the introduction, we approach this field following the line suggested by Chapter II and we thus need to understand which notion of polynomial, and more generally of derivative, is convenient to use. We address this task in Sections 2.2 and 2.3. In Section 2.4 we study functions of measure argument, whose derivative are particularly well-behaved.

2.1 Notation

Throughout this chapter E is a locally compact Polish space endowed with its Borel σ -algebra. We denote by $M_+(E)$ the space of all finite measures on E with the topology of weak convergence. The subset $M_1(E) \subseteq M_+(E)$ denotes the set of probability measures, and $M(E) = M_+(E) - M_+(E)$ the space of signed measures of bounded variation (i.e., they are of the form $\nu = \nu_+ - \nu_-$ with $\nu_+, \nu_- \in M_+(E)$), again topologized by weak convergence. For $\mu, \nu \in M(E)$ we write $\mu \leq \nu$ if $\nu - \mu \in M_+(E)$ and $|\nu|$ for $\nu_+ + \nu_-$. Moreover, given two Polish spaces E_1 and E_2 , both endowed with their Borel σ -algebra, some $\mu \in M(E_1)$, and a measurable map $f : E_1 \rightarrow E_2$, we denote by $f_*\mu \in M(E_2)$ the pushforward of μ with respect to f .

We let $C(E)$ denote the space of all continuous functions $f : E \rightarrow \mathbb{R}$. The subscript b (respectively c) indicates that they are also bounded (respectively have compact support). Both spaces $C_b(E)$ and $C_c(E)$ are given with the topology of uniform convergence and we denote by $\|\cdot\|$ the supremum norm. If E is noncompact, we let $E^\Delta = E \cup \{\Delta\}$ be the one-point compactification (or Alexandroff compactification) of E . Note that since E is a locally compact Polish space, E^Δ is a compact Polish space. In order to simplify the notation we write $E^\Delta = E$ when E is compact. We let $C_\Delta(E^k)$ be the closed subspace of $C_b(E^k)$ given by

$$C_\Delta(E^k) := \left\{ f|_{E^k} : f \in C_b((E^\Delta)^k) \right\}$$

and let $C_0(E)$ denote the subspace of $C_\Delta(E)$ of all continuous functions vanishing at infinity. Concerning notation, we will switch between $C_\Delta(E)$ and $C(E^\Delta)$ every time it is needed, which is possible due to their one to one correspondence. For E compact, we have $C(E) = C_b(E) = C_\Delta(E) = C_0(E)$ and we therefore simply write $C(E)$. Finally, with a small abuse of notation, for $p : M(E) \rightarrow \mathbb{R}$ we write $p \in C(M_1(E))$ if $p|_{M_1(E)}$ is continuous with respect to the topology of weak convergence and $p \in C(M_1(E^\Delta))$ if $p|_{M_1(E)} = \bar{p}|_{M_1(E)}$ for some continuous map $\bar{p} : M_1(E^\Delta) \rightarrow \mathbb{R}$.

Remark 2.1. *For each locally compact Polish space E , the spaces $M_+(E)$ and $M_1(E)$ are Polish. This in particular implies that both spaces are metrizable and thus sequential spaces. If E is also compact, then $M_+(E)$ is locally compact and*

$M_1(E)$ is a compact. However, if E is noncompact, $M_1(E)$ is not even locally compact.

We let $\widehat{C}_\Delta(E^k)$ denote the closed subspace of $C_\Delta(E^k)$ consisting of functions f that are symmetric, i.e., $f(x_1, \dots, x_k) = f(x_{\sigma(1)}, \dots, x_{\sigma(k)})$ for all $\sigma \in \Sigma_k$, where Σ_k denotes the permutation group of k elements. $\widehat{C}_0(E^k)$ and $\widehat{C}(E^k)$ are defined similarly. For any $g \in \widehat{C}_\Delta(E^k), h \in \widehat{C}_\Delta(E^\ell)$ we denote by $g \otimes h$ the function in $\widehat{C}_\Delta(E^{k+\ell})$ given by

$$\begin{aligned} (g \otimes h)(x_1, \dots, x_{k+\ell}) \\ = \frac{1}{(k+\ell)!} \sum_{\sigma \in \Sigma_{k+\ell}} g(x_{\sigma(1)}, \dots, x_{\sigma(k)}) h(x_{\sigma(k+1)}, \dots, x_{\sigma(k+\ell)}). \end{aligned} \quad (2.1)$$

Moreover, for a linear subspace $D \subseteq C_\Delta(E)$ we set $D \otimes D := \text{span}\{g \otimes h : g, h \in D\}$.

For any sufficiently differentiable function $g : \mathbb{R}^d \rightarrow \mathbb{R}$ we write ∇g for the gradient of g and $\nabla^2 g$ for the Hessian of g . If $d = 1$, we simplify the notation by setting $g' := \nabla g$ and $g'' := \nabla^2 g$. By convention we set $\nabla g(\Delta) = 0$ and $\nabla^2 g(\Delta) = 0$. Finally, given $E \subseteq \mathbb{R}^d$ and $n \in \mathbb{N} \cup \{\infty\}$, we denote by $C_\Delta^n(E)$ the space of all $f \in C_\Delta(E)$ such that $f = g|_E$ for some $g : \mathbb{R}^d \rightarrow \mathbb{R}$ n -times continuously differentiable. $C_c^n(E)$ and $C^n(E)$ are defined analogously.

2.2 Differentiation for functions of measure argument

Following the classical literature, we will use a notion of derivative which is well-known since the work of [Fleming and Viot \(1979\)](#).

Fix $x \in E$. If $p : M(E) \rightarrow \mathbb{R}$ is any function, we say that p is differentiable at μ in direction δ_x if

$$\partial_x p(\mu) := \lim_{\varepsilon \rightarrow 0} \frac{p(\mu + \varepsilon \delta_x) - p(\mu)}{\varepsilon}$$

exists. We then write $\partial p(\mu)$ for the map $x \mapsto \partial_x p(\mu)$ and we use the notation

$$\partial_{x_1, \dots, x_k}^k p(\mu) := \partial_{x_1} \cdots \partial_{x_k} p(\mu)$$

for iterated derivatives, writing $\partial^k p(\mu)$ for the corresponding map. Observe that for $p(\mu) = \int g(x) \mu(dx)$ for some $g \in C_\Delta(E)$ we get

$$\partial_x p(\mu) = \lim_{\varepsilon \rightarrow 0} \frac{\int g(y) \varepsilon \delta_x(dy)}{\varepsilon} = g(x).$$

2.3 Monomials and polynomials of measure argument

The compactness explained in [Remark 2.1](#) justify the use of a class of polynomial of measure arguments. As in the finite dimensional setting, we can use the Stone–Weierstrass’s theorem to prove the density of this class in the space of continuous function (see [Lemma 2.8\(iii\)](#) later for more details).

A *monomial* on $M(E)$ is an expression of the form

$$\langle g, \nu^k \rangle := \int_{E^k} g(x_1, \dots, x_k) \nu(dx_1) \cdots \nu(dx_k)$$

for some $k \in \mathbb{N}_0$, where $g \in \widehat{C}_\Delta(E^k)$ is referred to as the *coefficient* of the monomial. We identify $\widehat{C}_\Delta(E^0)$ with \mathbb{R} , so that for $k = 0$ we have $\langle g, \nu^0 \rangle = g \in \mathbb{R}$. If g is not the zero function, then k is called the degree of $\langle g, \nu^k \rangle$. If g is the zero function, the degree is defined to be $-\infty$. It is clear that the map $\nu \mapsto \langle g, \nu^k \rangle$ is homogeneous of degree k , and that $g \mapsto \langle g, \nu^k \rangle$ is linear. Furthermore, one has the identity $\langle g, \nu^k \rangle \langle h, \nu^\ell \rangle = \langle g \otimes h, \nu^{k+\ell} \rangle$, where the symmetric tensor product $g \otimes h$ is defined in (2.1). A polynomial on $M(E)$ is now defined as a (finite) linear combination of monomials,

$$p(\nu) = \sum_{k=0}^m \langle g_k, \nu^k \rangle, \quad (2.2)$$

with coefficients $g_k \in \widehat{C}_\Delta(E^k)$. The degree of the polynomial $p(\nu)$, denoted by $\deg(p)$, is the largest k such that g_k is not the zero function. The representation (2.2) is unique; see Corollary 2.5 below.

Example 2.2. Let $E = \{1, \dots, d\}$ be a finite set. Then every element $\nu \in M(E)$ is of the form

$$\nu = z_1 \delta_1 + \cdots + z_d \delta_d, \quad (z_1, \dots, z_d) \in \mathbb{R}^d,$$

where δ_i is the Dirac mass concentrated at $\{i\}$. Monomials take the form

$$\langle g, \nu^k \rangle = \sum_{i_1, \dots, i_k} g(i_1, \dots, i_k) z_{i_1} \cdots z_{i_k},$$

where the summation ranges over $E^k = \{1, \dots, d\}^k$. Therefore, as g ranges over all symmetric functions on E^k , we recover all homogeneous polynomials of total degree k in the d variables z_1, \dots, z_d . In particular, in view of Corollary 2.6 later, this relation provides a one to one correspondence between polynomials on the unit simplex Δ^d , namely the set

$$\Delta^d := \left\{ z \in \mathbb{R}^d : \sum_{i=1}^d z_i = 1, z_i \geq 0 \right\}$$

defined in Section II.6, and polynomials on $M_1(E)$.

We will usually think of a polynomial on $M(E)$ as the map it induces. We are then led to consider the following function spaces.

Definition 2.3. The set

$$P := \{ \nu \mapsto p(\nu) : p \text{ is a polynomial on } M(E) \}$$

denotes the algebra of all polynomials on $M(E)$ regarded as real-valued maps, equipped with pointwise addition and multiplication.

Just like ordinary polynomials, the elements of P are smooth. This is made precise in the next lemma.

Lemma 2.4. (i) *Each $p \in P$ is sequentially continuous (on $M(E)$) and continuous on $M_+(E)$. In particular, every polynomial is continuous on $M_1(E)$. Moreover, each $p \in P$ can be uniquely extended to a polynomial on $M(E^\Delta)$.*

(ii) *Let $p \in P$ be a monomial of the form $p(\nu) = \langle g, \nu^k \rangle$. Then, for every $x \in E$ and $\nu \in M(E)$,*

$$\partial_x p(\nu) = k \langle g(\cdot, x), \nu^{k-1} \rangle,$$

where $g(\cdot, x) \in \widehat{C}_\Delta(E^{k-1})$ is the map $(x_1, \dots, x_{k-1}) \mapsto g(x_1, \dots, x_{k-1}, x)$. If $k = 0$, the right-hand side should be read as zero.

(iii) *For each $p \in P$ and $x \in E$ the map $\partial_x p: \nu \mapsto \partial_x p(\nu)$ lies in P .*

(iv) *For each $p \in P$ and $\nu \in M(E)$, the map $\partial p(\nu): x \mapsto \partial_x p(\nu)$ lies in $C_\Delta(E)$.*

(v) *The identity*

$$\partial_x(pq)(\nu) = p(\nu)\partial_x q(\nu) + q(\nu)\partial_x p(\nu)$$

holds for all $p, q \in P$, $x \in E$, $\nu \in M(E)$.

(vi) *The Taylor representation*

$$p(\nu + \mu) = \sum_{\ell=0}^k \frac{1}{\ell!} \langle \partial^\ell p(\nu), \mu^\ell \rangle,$$

holds for all $p \in P$ and $\nu, \mu \in M(E)$, where k denotes the degree of p .

Proof. (i): For $h \in C_\Delta(E)^{\otimes k}$ we can write $h = \sum_{\ell=1}^L \lambda_\ell h_\ell^{\otimes k}$ for some $h_\ell \in C_\Delta(E)$ and $\lambda_\ell \in \mathbb{R}$. Since $\langle h_\ell, \nu \rangle$ is continuous by definition of weak convergence,

$$\langle h, \nu^k \rangle = \sum_{\ell=1}^L \lambda_\ell \langle h_\ell^{\otimes k}, \nu^k \rangle = \sum_{\ell=1}^L \lambda_\ell \langle h_\ell, \nu \rangle^k$$

is continuous as well. Note then that by linearity in (2.2) it is enough to prove the result for $p(\nu) = \langle g, \nu^k \rangle$ and $g \in \widehat{C}_\Delta(E^k)$. Choose $h \in C_\Delta(E)^{\otimes k}$ such that $\|g - h\| \leq \varepsilon$ and let $\nu_n \in M(E)$ converge weakly to $\nu \in M(E)$. Observe that, by Banach–Steinhaus theorem, $\sup_n |\nu_n|(E) < \infty$. Then

$$|\langle g, \nu_n^k \rangle - \langle g, \nu^k \rangle| \leq |\langle h, \nu_n^k \rangle - \langle h, \nu^k \rangle| + \varepsilon (\sup_n |\nu_n|(E)^k + |\nu|(E)^k) \xrightarrow{n \rightarrow \infty} C\varepsilon$$

for some $C \geq 0$. Since ε is arbitrary, this proves sequential continuity of p on $M(E)$. Finally, since by Remark 2.1 $M_+(E)$ is a Polish space and thus a sequential space, we can conclude that p is continuous on $M_+(E)$. The last part follows from the observation that every function in $C_\Delta(E^k)$ can be uniquely extended to a function in $C((E^\Delta)^k)$.

(ii): A direct calculation yields

$$\begin{aligned} p(\nu + \varepsilon\delta_x) - p(\nu) &= \int g(x_1, \dots, x_k) \left(\prod_{i=1}^k (\nu + \varepsilon\delta_x)(dx_i) - \prod_{i=1}^k \nu(dx_i) \right) \\ &= \varepsilon \sum_{i=1}^k \int g(x_1, \dots, x_k) \delta_x(dx_i) \prod_{j \neq i} \nu(dx_j) + O(\varepsilon^2) \\ &= \varepsilon k \int g(x_1, \dots, x_{k-1}, x) \prod_{j=1}^{k-1} \nu(dx_j) + o(\varepsilon), \end{aligned}$$

using the symmetry of g in the last equality. The expression for $\partial_x p(\nu)$ follows.

For the remaining part of the proof it suffices to consider monomials $p(\nu) = \langle g, \nu^k \rangle$ for $g \in \widehat{C}_\Delta(E^k)$ due to the linearity in (2.2).

(iii): Fix $x \in E$ and note that $kg(\cdot, x) \in \widehat{C}_\Delta(E^{k-1})$. The claim follows by (ii).

(iv): For $p(\nu) = \langle g, \nu^k \rangle$ we have

$$|\partial_x p(\nu)| = |\langle kg(\cdot, x), \nu^{k-1} \rangle| \leq k \|g\| |\nu|(E)^{k-1} < \infty.$$

Continuity of $x \mapsto \partial_x p(\nu)$ follows from the dominated convergence theorem and the fact that E is Polish, and thus a sequential space.

(v): For monomials $p(\nu) = \langle g, \nu^k \rangle$ and $q(\nu) = \langle h, \nu^\ell \rangle$, we have $pq(\nu) = \langle g \otimes h, \nu^{k+\ell} \rangle$. We can then compute

$$\begin{aligned} \partial_x pq(\nu) &= k \langle g(\cdot, x), \nu^{k-1} \rangle \langle h, \nu^\ell \rangle + \ell \langle g, \nu^k \rangle \langle h(\cdot, x), \nu^{\ell-1} \rangle \\ &= p(\nu) \partial_x q(\nu) + q(\nu) \partial_x p(\nu) \end{aligned}$$

for all $x \in E$ and $\nu \in M(E)$.

(vi): Observing that for $p(\nu) = \langle g, \nu^k \rangle$ it holds

$$p(\nu + \mu) = \sum_{\ell=0}^k \binom{k}{\ell} \int g(x_1, \dots, x_k) \prod_{i=\ell+1}^k \nu(dx_i) \prod_{i=1}^{\ell} \mu(dx_i)$$

the result follows by (ii). \square

From Lemma 2.4(ii) one can deduce the uniqueness of the representation (2.2).

Corollary 2.5. *Suppose $p(\nu) = \sum_{k=0}^m \langle g_k, \nu^k \rangle$ equals zero for all $\nu \in M(E)$. Then $g_k = 0$ for all k .*

Proof. Let $x_1, \dots, x_m \in E$ be arbitrary and note that differentiating m times using Lemma 2.4(ii) we get $m! g_m(x_1, \dots, x_m) = \partial_{x_1 x_2 \dots x_m} p(\nu) = 0$. Thus $g_m = 0$. Now repeat this successively for $g_{m-1}, g_{m-2}, \dots, g_0$. \square

The next property is particularly useful in the context of probability measure-valued polynomial jump-diffusions (see Remark IV.2.6 later). In the finite-dimensional setting, the result states that every polynomial on the unit simplex has a homogeneous representative.

Corollary 2.6. *Every polynomial on $M(E)$ has a unique homogeneous representative on $M_1(E)$. That is, for every $p \in P$ with $\deg(p) \leq k$ there is a unique $g \in \widehat{C}_\Delta(E^k)$ such that*

$$p(\nu) = \langle g, \nu^k \rangle \quad \text{for all } \nu \in M_1(E).$$

Proof. Corollary 2.5 yields a unique set of coefficients g_0, \dots, g_k with $g_\ell \in \widehat{C}_\Delta(E^\ell)$ and $p(\nu) = \sum_{\ell=0}^k \langle g_\ell, \nu^\ell \rangle$. The result follows by setting $g := \sum_{\ell=0}^k g_\ell \otimes 1^{\otimes(k-\ell)}$. \square

Remark 2.7. *If we choose to work with coefficients in $\widehat{C}_0(E^k)$ instead of $\widehat{C}_\Delta(E^k)$ we would obtain the same class of polynomials on $M_1(E)$. This is because every $g \in \widehat{C}_\Delta(E^k)$ equals $\sum_{i=0}^k g_i \otimes 1^{\otimes(k-i)}$ for some $g_i \in \widehat{C}_0(E^i)$, and therefore $\langle g, \nu^k \rangle = \sum_{i=0}^k \langle g_i, \nu^i \rangle$ for all $\nu \in M_1(E)$. Indeed, the g_i are given iteratively by*

$$g_0 := g(\Delta, \dots, \Delta) \quad \text{and} \quad g_i := \binom{k}{i} \left(g(\Delta, \dots, \Delta, \cdot) - \sum_{j=0}^{i-1} g_j \otimes 1^{\otimes(i-j)} \right).$$

However, not every such polynomial admits a homogenous representative on $M_1(E)$ in the sense of Corollary 2.6, unless E is compact. An example is $1 + \langle g, \nu \rangle$ with $g \in C_0(E)$ nonzero. The existence of homogeneous representatives leads to significant notational simplifications in the context of probability measure-valued polynomial diffusions. This is the main reason for working with the space $C_\Delta(E)$ instead of $C_0(E)$.

2.4 Cylindrical functions and cylindrical polynomials

As we saw in Lemma 2.4(iv), derivatives of polynomials inherit continuity from their coefficients, but additional regularity is sometimes needed. This leads us to consider subspaces of polynomials whose coefficients are sufficiently nice.

More precisely, let $D \subseteq C_\Delta(E)$ be a dense linear subspace containing the constant function 1 and for all $n \in \mathbb{N}$ let $C^{n,D}$ denote the space of functions on $M(E)$ defined by

$$C^{n,D} := \left\{ \mu \mapsto \phi(\langle g_1, \mu \rangle, \dots, \langle g_k, \mu \rangle) : k \in \mathbb{N}, g_1, \dots, g_k \in D, \phi \in C^n(\mathbb{R}^k) \right\}.$$

We call this space the space of *cylindrical functions*. The space of *cylindrical polynomials* with coefficients in D is then the subspace of $C^{n,D}$ given by

$$P^D := \text{span} \left\{ 1, \mu \mapsto \langle g, \mu \rangle^k : k \in \mathbb{N}, g \in D \right\}. \quad (2.3)$$

Thus P^D is the subalgebra of P consisting of all (finite) linear combinations of the constant polynomial and “rank-one” monomials $\langle g \otimes \dots \otimes g, \nu^k \rangle = \langle g, \nu \rangle^k$ with $g \in D$. Equivalently, P^D consists of all polynomials $p(\nu) = \phi(\langle g_1, \nu \rangle, \dots, \langle g_k, \nu \rangle)$ with $k \in \mathbb{N}$, $g_1, \dots, g_k \in D$, and ϕ a polynomial on \mathbb{R}^k .

The next lemma refines the properties given in Lemma 2.4. In order to simplify the notation, for $\phi \in C^n(\mathbb{R}^k)$ we write ϕ_{i_1, \dots, i_n} for $\frac{d^n}{dx_{i_1} \dots dx_{i_n}} \phi$.

Lemma 2.8. (i) Fix $\phi \in C^n(\mathbb{R}^k)$ and $g_1, \dots, g_k \in D$. Then $p \in C^{n,D}$ given by

$$p(\mu) = \phi(\langle g_1, \mu \rangle, \dots, \langle g_k, \mu \rangle) \quad (2.4)$$

is n -times differentiable at μ in direction δ_x for each $\mu \in M(E)$ and $x \in E$. Moreover, for all $x_1, \dots, x_n \in E$

$$\partial_{x_1, \dots, x_n}^n p(\mu) = \sum_{i_1, \dots, i_n=1}^k g_{i_1}(x_1) \cdots g_{i_n}(x_n) \phi_{i_1, \dots, i_n}(\langle g_1, \mu \rangle, \dots, \langle g_k, \mu \rangle).$$

This in particular yields

$$\partial_x p(\mu) = \nabla \phi(\langle g_1, \mu \rangle, \dots, \langle g_k, \mu \rangle)(g_1(x), \dots, g_k(x))^\top$$

for all $x \in E$ and, whenever $n \geq 2$,

$$\partial_{xy}^2 p(\mu) = (g_1(y), \dots, g_k(y)) \nabla^2 \phi(\langle g_1, \mu \rangle, \dots, \langle g_k, \mu \rangle)(g_1(x), \dots, g_k(x))^\top,$$

for all $x, y \in E$.

- (ii) For $p \in C^{n,D}$ and $\mu \in M(E)$, we have $\partial^k p(\mu) \in D^{\otimes k}$ for all $k \in \{1, \dots, n\}$. In particular, each $p \in P^D$ is infinitely differentiable and $\partial^k p(\mu) \in D^{\otimes k}$.
- (iii) P^D , adequately extended and restricted, is a dense subset of the space of continuous functions on $M_1(E^\Delta)$.
- (iv) For all $\nu, \mu \in M(E)$, the Taylor approximation

$$p(\mu + \varepsilon\nu) = \sum_{\ell=0}^n \frac{\varepsilon^\ell}{\ell!} \langle \partial^\ell p(\mu), \nu^\ell \rangle + o(\varepsilon^n)$$

holds for all $p \in C^{n,D}$.

- (v) Let p be as in (2.4) for some differentiable function $\phi : \mathbb{R}^k \rightarrow \mathbb{R}$. Then

$$\partial_\nu p(\mu) := \lim_{\varepsilon \rightarrow 0} \frac{p(\mu + \varepsilon\nu) - p(\mu)}{\varepsilon} = \langle \partial p(\mu), \nu \rangle.$$

For those functions, the map $\nu \mapsto \partial_\nu p(\mu)$ thus coincides with the Fréchet derivative of p at μ . This is in particular the case for all $p \in C^{1,D}$.

Proof. (i) For each $\nu \in M(E)$ set $v_\nu := (\langle g_1, \nu \rangle, \dots, \langle g_k, \nu \rangle)^\top$. Noting that for each $i \in \{1, \dots, k\}$ it holds $\langle g_i, \mu + \varepsilon\delta_x \rangle = \langle g_i, \mu \rangle + \varepsilon\langle g_i, \delta_x \rangle$, differentiability of ϕ yields

$$\begin{aligned} \partial_x p(\mu) &= \lim_{\varepsilon \rightarrow 0} \frac{\phi(v_\mu + \varepsilon v_{\delta_x}) - \phi(v_\mu)}{\varepsilon} = \nabla \phi(v_\mu) v_{\delta_x}^\top \\ &= \nabla \phi(\langle g_1, \mu \rangle, \dots, \langle g_k, \mu \rangle)(g_1(x), \dots, g_k(x))^\top. \end{aligned}$$

The conclusion for higher order derivatives follows similarly.

(ii) Follows by (i) and the regularity of the coefficients g_1, \dots, g_k .

(iii) Continuity of polynomials follows by Lemma 2.4(i). Stone–Weierstrass and the fact that D is densely contained in $C(E^\Delta)$ yield the density.

(iv) Similarly to (i), the result follows by noting that $p(\mu + \varepsilon\nu) = \phi(v_\mu + \varepsilon v_\nu)$ and applying the classical Taylor approximation theorem.

(v) Follows by (iv). □

3 The martingale problem

Again, let E be a locally compact Polish space and $D \subseteq C_\Delta(E)$ be a dense linear subspace. Since we are not interested in reducing the generality at this point, we consider a linear operator

$$L : P^D \rightarrow F(M(E))$$

where $F(M(E))$ denotes the set of all real-valued functions on $M(E)$. Fix $\mathcal{S} \subseteq M(E)$ and let \mathcal{S}^τ be the topological space of \mathcal{S} endowed with a topology τ . We assume that $Lp|_{\mathcal{S}}$ are continuous (with respect to τ) for all $p \in P^D$.

Definition 3.1. (i) Let X be an \mathcal{S} -valued process defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and suppose that X has càdlàg paths with respect to τ . Then X is called conservative \mathcal{S}^τ -solution (or simply \mathcal{S}^τ -solution) to the martingale problem for L with initial condition $\mu \in \mathcal{S}$ if $X_0 = \mu$ \mathbb{P} -a.s. and

$$N_t^p = p(X_t) - p(X_0) - \int_0^t Lp(X_s) ds \quad (3.1)$$

defines a local martingale for every $p \in P^D$. Uniqueness of solutions of the martingale problem is always understood in the sense of law. The martingale problem for L is well-posed if for every $\mu \in \mathcal{S}$ there exists a unique \mathcal{S}^τ -solution to the martingale problem for L with initial condition μ .

(ii) More generally, let μ^\dagger denote an isolated point playing the role of a cemetery state. Define $(L^\dagger p)|_{\mathcal{S}} := L((p - p(\mu^\dagger))|_{\mathcal{S}})$ and $L^\dagger p(\mu^\dagger) = 0$ for all $p : \mathcal{S} \cup \{\mu^\dagger\} \rightarrow \mathbb{R}$ such that $p|_{\mathcal{S}} \in P^D$. Then, any $(\mathcal{S} \cup \{\mu^\dagger\})^\tau$ -solution to the martingale problem for L^\dagger with initial condition $\mu \in \mathcal{S}$ is called possibly killed \mathcal{S}^τ -solution to the martingale problem for L with initial condition μ .¹

Remark 3.2. Observe that since $L^\dagger p(\mu^\dagger) = 0$ for all $p : \mathcal{S} \cup \{\mu^\dagger\} \rightarrow \mathbb{R}$, once the corresponding possibly killed solution X reaches μ^\dagger it stays there forever. In this case, we say that the solution is killed. If on the contrary X is \mathcal{S} -valued, one can easily verify that it is also a conservative \mathcal{S}^τ -solution.

It is interesting to note that in most of the applications $Lp|_{\mathcal{S}}$ is bounded for every $p \in P^D$. This implies that for any solution X to the martingale problem for L and any $p \in P^D$, the process N^p is a bounded local martingale on $[0, T]$, and thus a true martingale. This in particular yields

$$\mathbb{E}[p(X_t)] = \mathbb{E}\left[p(X_0) + \int_0^t Lp(X_s) ds\right].$$

The notion of $M_1(E)$ -solution permits us to introduce the object that the title of the thesis refers to.

¹With $(\mathcal{S} \cup \{\mu^\dagger\})^\tau$ we indicate the topological space of $\mathcal{S} \cup \{\mu^\dagger\}$ endowed with the topology given by $\{U \subseteq \mathcal{S} \cup \{\mu^\dagger\} : U \setminus \{\mu^\dagger\} \in \tau\}$.

Definition 3.3. An $M_1(E)$ -solution X to the martingale problem for a linear operator $L : P^D \rightarrow C(M_1(E))$ is called probability measure-valued jump-diffusion.

If $\langle g, X \rangle$ is a real valued martingale for all $g \in D$, then X is called probability measure-valued martingale.

Probability measure-valued martingales are particularly interesting for applications in mathematical finance, see for instance [Beiglböck et al. \(2017\)](#).

3.1 The positive maximum principle

We introduce now the main tool of this chapter. The power behind this concept is given by [Lemma 3.5](#) and [Lemma 3.6](#).

Definition 3.4. Fix $\mu \in M(E)$. Then L is said to satisfy the positive maximum principle on \mathcal{S} at μ if

$$p \in P^D, \quad \sup_{\mathcal{S}} p = p(\mu) \geq 0 \quad \Rightarrow \quad Lp(\mu) \leq 0.$$

If this holds for all $\mu \in \mathcal{S}$, then L is said to satisfy the positive maximum principle on \mathcal{S} .

The next two lemmas illustrate how the positive maximum principle is essentially equivalent to the existence of a (possibly killed) solution to the martingale problem for L for all initial conditions. The proof of [Lemma 3.5](#) is standard and we thus omit it. See for instance the proof of [Lemma 2.3](#) in [Filipović and Larsson \(2016\)](#).

Lemma 3.5. Let $L : P^D \rightarrow C(M_1(E))$ be a linear operator, fix $\mu \in M_1(E)$ and suppose that there exists a (possibly killed) $M_1(E)$ -solution X to the martingale problem for L with initial condition μ . Then L satisfies the positive maximum principle on $M_1(E)$ at μ . If X is conservative, then $L1 = 0$.

The next lemma is an adaptation of a classical result from [Ethier and Kurtz \(2005\)](#). For the application of this result it is crucial for L to be an operator on the space of continuous functions on a locally compact, separable, metrizable space. Since this is not the case of $M_1(E)$ for E noncompact, we need to go around this problem. The principle is simple. Instead of working with $M_1(E)$ we first accept the possibility that the total mass of the solution to the corresponding martingale problem is not preserved. Mathematically, this can be achieved by working with $M_1(E^\Delta)$ -valued processes, where E^Δ denotes the one-point compactification of E . Since E is a locally compact Polish space, E^Δ is a compact Polish space and $M_1(E^\Delta)$ is a compact Polish space with respect to the corresponding topology of weak convergence (see [Remark 2.1](#)). The result of [Ethier and Kurtz \(2005\)](#) can then be applied and we just have to check that if the initial condition of an $M_1(E^\Delta)$ -solution X assigns mass 1 to E , then $X_t(E) = 1$ almost surely for each $t \geq 0$.

From a technical point of view, we need now to guarantee that L can be seen as an operator acting on continuous functions on $M(E^\Delta)$. Observe that by

Lemma 2.4(i) every polynomial on $M(E)$ can be uniquely extended to a polynomial on $M(E^\Delta)$. If we also have that Lp has a continuous extension to $M_1(E^\Delta)$, we can then view the operator L as an operator acting on continuous functions on $M_1(E^\Delta)$. This is in particular the case if $Lp \in P$ for all $p \in P^D$. Recall that for $p \in F(M(E))$ we write $p \in C(M_1(E^\Delta))$ if

$$p|_{M_1(E)} = \bar{p}|_{M_1(E)}$$

for some continuous map $\bar{p} : M_1(E^\Delta) \rightarrow \mathbb{R}$.

Lemma 3.6. *Let $L : P^D \rightarrow C(M_1(E^\Delta))$ be a linear operator satisfying the positive maximum principle on $M_1(E^\Delta)$. Then, there exists a (possibly killed) $M_1(E^\Delta)$ -solution to the martingale problem for L for each initial condition $\mu \in M_1(E)$. If $L1 = 0$, then every possibly killed $M_1(E^\Delta)$ -solution is conservative.*

Suppose in addition that there exists a sequence $(g_n)_{n \in \mathbb{N}} \subseteq D \cap C_0(E)$ such that setting $p_n(\mu) = \langle g_n, \mu \rangle$ it holds

$$\lim_{n \rightarrow \infty} g_n = 1, \quad \text{and} \quad \lim_{n \rightarrow \infty} (Lp_n)^- = 0 \quad (3.2)$$

bounded pointwise on E and $M_1(E^\Delta)$, respectively. Then, every $M_1(E^\Delta)$ -solution with initial condition $\mu \in M_1(E)$ is also an $M_1(E)$ -solution.²

Proof. The first part of the result is a consequence of Theorem 4.5.4 in Ethier and Kurtz (2005) and the successive Remark 4.5.5. We now explain how the necessary conditions hold true. By Remark 2.1 and Lemma 2.8(iii) we know that $M_1(E^\Delta)$ is a compact separable metrizable space and

$$P^D(M_1(E^\Delta)) := \{p|_{M_1(E^\Delta)} : p \in P^D\}$$

is a dense subset of the space of continuous functions on $M_1(E^\Delta)$. Moreover, the positive maximum principle implies that $Lp|_{M_1(E^\Delta)} = Lq|_{M_1(E^\Delta)}$ for all $p, q \in P^D$ such that $p|_{M_1(E^\Delta)} = q|_{M_1(E^\Delta)}$. We may thus regard L as an operator on the space of continuous functions on $M_1(E^\Delta)$ with domain $P^D(M_1(E^\Delta))$.

For the second part, observe that by the dominated convergence theorem, (3.1), and Fatou's lemma we can compute

$$\mathbb{E}[X_t(E)] = \lim_{n \rightarrow \infty} \mathbb{E}[\langle g_n, X_t \rangle] = \lim_{n \rightarrow \infty} \left(\langle g_n, \mu \rangle + \mathbb{E} \left[\int_0^t Lp_n(X_s) ds \right] \right) \geq \mu(E) = 1.$$

Finally, note that a càdlàg process X on $M_1(E^\Delta)$ such that $X_t(E) = 1$ almost sure is càdlàg also with respect to the topology of weak convergence on $M_1(E)$. \square

Clearly, it would be nice if the conditions of Lemma 3.6 could be expressed in terms of $M_1(E)$ instead of $M_1(E^\Delta)$. Even if this is not possible in general, we provide now a result connecting the positive maximum principle on $M_1(E)$ with the positive maximum principle on $M_1(E^\Delta)$.

²More precisely, every $M_1(E^\Delta)$ -solution X with initial condition $\mu \in M_1(E^\Delta)$ such that $\mu(E) = 1$, satisfies $X_t(E) = 1$ almost surely for all $t \geq 0$. Moreover, the paths of the solution are càdlàg with respect to the topology of weak convergence on $M_1(E)$.

Lemma 3.7. *Suppose that E is noncompact.*

- (i) *Fix $\mu \in M_1(E)$. Then L satisfies the positive maximum principle on $M_1(E)$ at μ if and only if it does so on $M_1(E^\Delta)$ at μ .*
- (ii) *Suppose that $D \subseteq \mathbb{R} + C_c(E)$ and $Lp \in C(M_1(E^\Delta))$ for all $p \in D$. Then L satisfies the positive maximum principle on $M_1(E)$ if and only if it does so on $M_1(E^\Delta)$.*

Proof. (i) Note that the closure of $M_1(E)$ with respect to the topology of weak convergence on $M_1(E^\Delta)$ is given by $M_1(E^\Delta)$. Combined with Lemma 2.8(iii) this implies that $\sup_{M_1(E)} p = \max_{M_1(E^\Delta)} p$ for all $p \in P^D$ and the result follows.

(ii) Let L satisfy the positive maximum principle on $M_1(E)$. Fix $\mu \in M_1(E^\Delta)$ and $p \in P^D$ such that $p(\mu) = \max_{M_1(E^\Delta)} p \geq 0$. Note that for $|x|$ big enough

$$p\left(\mu + \mu(\{\Delta\})(\delta_x - \delta_\Delta)\right) = p(\mu) = \max_{M_1(E^\Delta)} p \geq 0$$

and hence, by the first part and the positive maximum principle on $M_1(E)$, $Lp(\mu + \mu(\{\Delta\})(\delta_x - \delta_\Delta)) \leq 0$. Letting $|x|$ go to infinity, the continuity of Lp yields $Lp(\mu) \leq 0$. The converse implication follows by (i). \square

3.2 Properties of $M_1(E)$ -solution

The same principle behind condition (3.2) can be used to investigate other properties of a solution to the martingale problem. As illustrative examples, we present now sufficient conditions for the so-called support containment condition (item (i)), for guaranteeing that a given solution is not absolutely continuous (item (ii)), and for the invariance of the Wasserstein space (items (iii) and (iv)). It will be clear to the reader that the same method can be used for the investigation of many other properties of a solution.

Lemma 3.8. *Let $L : P^D \rightarrow C(M_1(E^\Delta))$ be a linear operator satisfying the positive maximum principle on $M_1(E^\Delta)$ and let X be an $M_1(E)$ -solution on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ to the corresponding martingale problem with initial condition $\mu \in M_1(E)$.*

- (i) *Fix an arbitrary subset $A \subseteq E$ and suppose that there exists a sequence $(g_n)_{n \in \mathbb{N}} \subseteq D$ such that setting $p_n(\nu) := \langle g_n, \nu \rangle$ it holds*

$$\lim_{n \rightarrow \infty} g_n = \mathbf{1}_A, \quad \text{and} \quad \lim_{n \rightarrow \infty} (Lp_n)^- = 0$$

bounded pointwise on E and $M_1(E)$, respectively. Then, $\mu(A) = 1$ yields $X_t(A) = 1$ a.s. for all $t \geq 0$.

- (ii) *Suppose that there exists a sequence $(g_n)_{n \in \mathbb{N}} \subseteq D \otimes D$ such that setting $p_n(\nu) := \langle g_n, \nu^2 \rangle$ it holds*

$$\lim_{n \rightarrow \infty} g_n(x, y) = \mathbf{1}_{\{x=y\}}, \quad \text{and} \quad \lim_{n \rightarrow \infty} (Lp_n)^- = 0$$

bounded pointwise on $E \times E$ and on $M_1(E)$, respectively. Then,

$$\lim_{n \rightarrow \infty} (L(p_n)(\mu))^- > 0$$

yields X_t not absolutely continuous with positive probability, for all $t > 0$. More precisely, in this case for each $t > 0$ there is a positive probability that X_t has at least one atom.

- (iii) Suppose that $E \subseteq \mathbb{R}^d$ and let $D \subseteq \mathbb{R} + C_c(E)$. Consider an increasing function $\psi \in C^\infty(\mathbb{R}_+)$ such that

$$0 \leq \psi(x) \leq C(1+x)$$

for some $C \in \mathbb{R}_+$ and $\lim_{x \rightarrow \infty} \psi(x) = \infty$. Fix also $m_p \in C^\infty(\mathbb{R}^d)$ such that $m_p(x) = |x|^p$ for $|x|$ big enough. Suppose that there exists a nonnegative increasing sequence $(g_n)_{n \in \mathbb{N}} \subseteq D$ converging to m_p on E and such that for $p_n(\nu) := \psi(\langle g_n, \nu \rangle)$ it holds

$$\sup_n \sup_{M_1(E)} (Lp_n)^+ < \infty.$$

Then $\langle |\cdot|^p, \mu \rangle < \infty$ yields $\langle |\cdot|^p, X_t \rangle < \infty$ a.s. for all $t \geq 0$.

- (iv) Let E and D be as in (iii), and suppose that $(\mathcal{F}_t)_{t \geq 0}$ is right continuous and complete. Then if (iii) is true for some ψ being also strictly increasing and

$$\sup_n \sup_{M_1(E)} |Lp_n| < \infty,$$

the process X has a càdlàg modification with respect to the p -Wasserstein distance.

Proof. (i) By the dominated convergence theorem, (3.1), and Fatou's lemma we compute

$$\mathbb{E}[X_t(A)] = \lim_{n \rightarrow \infty} \mathbb{E}[\langle g_n, X_t \rangle] = \lim_{n \rightarrow \infty} \left(\langle g_n, \mu \rangle + \mathbb{E} \left[\int_0^t Lp_n(X_s) ds \right] \right) \geq \mu(A) = 1.$$

(ii) First observe that for each measure $\nu \in M_1(E)$ it holds $\sum_{x \in E} \nu(\{x\})^2 = \langle \mathbf{1}_{\{x=y\}}, \nu^2 \rangle$. By the dominated convergence theorem, (3.1), and Fatou's lemma we then compute

$$\mathbb{E} \left[\sum_{x \in E} X_t(\{x\})^2 \right] = \lim_{n \rightarrow \infty} \mathbb{E}[\langle g_n, X_t^2 \rangle] = \lim_{n \rightarrow \infty} \left(\langle g_n, \mu^2 \rangle + \mathbb{E} \left[\int_0^t Lp_n(X_s) ds \right] \right) > 0.$$

(iii) By the monotone convergence theorem, (3.1), and Fatou's lemma we compute

$$\begin{aligned} 0 \leq \mathbb{E}[\psi(\langle m_p, X_t \rangle)] &= \lim_{n \rightarrow \infty} \mathbb{E}[p_n(X_t)] = \lim_{n \rightarrow \infty} \left(p_n(\mu) + \mathbb{E} \left[\int_0^t Lp_n(X_s) ds \right] \right) \\ &\leq \psi(\langle m_p, \mu \rangle) + t \sup_n \|(Lp_n)^+\| < \infty, \end{aligned}$$

proving that $\psi(\langle m_p, X_t \rangle)$, and thus $\langle |\cdot|^p, X_t \rangle$, is finite \mathbb{P} -a.s.

(iv) Since we already know that X has càdlàg paths with respect to the topology of weak convergence, it suffices to show that $(\langle m_p, X_t \rangle)_{t \geq 0}$ has càdlàg paths. Since ψ is strictly increasing, this can be proved by showing that $(\psi(\langle m_p, X_t \rangle))_{t \geq 0}$ has càdlàg paths. First note that setting $L(\psi(\langle m_p, \cdot \rangle))(\nu) := \limsup_{n \rightarrow \infty} Lp_n(\nu)$ and

$$M_t := \psi(\langle m_p, X_t \rangle) - \int_0^t L(\psi(\langle m_p, \cdot \rangle))(X_s) ds,$$

monotone convergence theorem and dominated convergence theorem yield

$$\mathbb{E}[M_t - M_s | \mathcal{F}_s] = \lim_{n \rightarrow \infty} \mathbb{E} \left[\psi(\langle g_n, X_t \rangle) - \psi(\langle g_n, X_s \rangle) - \int_s^t Lp_n(X_u) du \mid \mathcal{F}_s \right] = 0.$$

Since, as a by-product of the proof of (ii), we already know that M_t is integrable, this proves that M is a martingale. By Theorem 2.2.9 in Revuz and Yor (1999) we can then conclude that M_t , and thus $\langle m_p, X_t \rangle$, has a càdlàg modification. \square

The Fleming–Viot diffusion was introduced by Fleming and Viot (1979) and subsequently studied by several other authors. It can be defined as the $M_1(\mathbb{R})$ -solution to the martingale problem for $L : P^D \rightarrow P$ given by

$$Lp(\nu) = \int_E B(\partial p(\nu))(x) \nu(dx) + \frac{1}{2} \int_{E^2} \partial_{xy}^2 p(\nu) \nu(dx) (\delta_x(dy) - \nu(dy)), \quad \nu \in M_1(E),$$

where $D = C_{\Delta}^2(\mathbb{R})$ and $Bg(x) = \frac{1}{2} \sigma^2 g''(x)$ for some $\sigma > 0$. It is well-known, see for instance Konno and Shiga (1988), that it takes values among absolutely continuous probability measures. In the next example, we show that choosing $\sigma = 0$ this property is not satisfied anymore.

Example 3.9 (Fleming–Viot martingale). *Fix $\sigma = 0$ such that L can be rewritten as $Lp(\nu) = \langle Q(\partial^2 p(\nu)), \nu^2 \rangle$ for $Qg(x, y) = \frac{1}{2}(g(x, x) + g(y, y) - 2g(x, y))$. Choose an arbitrary sequence $(g_n)_{n \in \mathbb{N}} \subseteq C_c^\infty(\mathbb{R}) \otimes C_c^\infty(\mathbb{R})$ converging bounded pointwise on \mathbb{R}^2 to $\mathbf{1}_{\{x=y\}}$. Setting $p_n(\nu) := \langle g_n, \nu^2 \rangle$, the boundedness of Q yields*

$$\lim_{n \rightarrow \infty} Lp_n(\nu) = \langle \mathbf{1}_{\{x \neq y\}}, \nu^2 \rangle,$$

bounded pointwise on $M_1(\mathbb{R})$. By Lemma 3.8(ii) we can conclude that any $M_1(\mathbb{R})$ -solution X of the corresponding martingale problem has at least an atom with positive probability for each positive time. More precisely, noting that

$$\mathbb{E}[\langle \mathbf{1}_{\{x=y\}}, X_t^2 \rangle] = \mathbb{E}[\langle \mathbf{1}_{\{x=y\}}, X_0^2 \rangle] + \int_0^t 1 - \mathbb{E}[\langle \mathbf{1}_{\{x=y\}}, X_s^2 \rangle] ds,$$

we can conclude that if the initial condition is free of atoms, then

$$\mathbb{E} \left[\sum_{x \in \mathbb{R}} X_t(\{x\})^2 \right] = \mathbb{E}[\langle \mathbf{1}_{\{x=y\}}, X_t^2 \rangle] = 1 - e^{-t}.$$

This in particular shows that for t going to infinity X_t converges to a Dirac measure.

3.3 Continuity of $M_1(E)$ -solutions

The operator L also contains information about path continuity of corresponding $M_1(E)$ -solutions to the martingale problem. As in [Bakry and Émery \(1985\)](#), we now introduce the two central notions to formulate this more precisely.

Definition 3.10. *The carré-du-champ operator of L is the symmetric bilinear map $\Gamma: P^D \times P^D \rightarrow F(M(E))$ defined by*

$$\Gamma(p, q) = L(pq) - pLq - qLp. \quad (3.3)$$

For $\mathcal{S} \subseteq M(E)$, we say that a symmetric bilinear map $\Gamma: P^D \times P^D \rightarrow F(M(E))$ is an \mathcal{S} -derivation if

$$\Gamma(pq, r) = p\Gamma(q, r) + q\Gamma(p, r) \quad \text{on } \mathcal{S}$$

for all $p, q, r \in P^D$.

The next lemma states, roughly speaking, that path continuity for a jump-diffusion corresponding to L holds precisely when the carré-du-champ operator of L is a derivation.

Lemma 3.11. *Let $L: P^D \rightarrow C(M_1(E))$ be a linear operator. If the carré-du-champ operator Γ of L is an $M_1(E)$ -derivation, then any $M_1(E)$ -solution to the martingale problem for L has continuous paths. Conversely, if for every initial condition $\nu \in M_1(E)$ there is an $M_1(E)$ -solution to the martingale problem for L with continuous paths, then the carré-du-champ operator Γ associated to L is an $M_1(E)$ -derivation.*

Proof. Let X be an $M_1(E)$ -solution to the martingale problem for L . By Proposition 2 in [Bakry and Émery \(1985\)](#), the real-valued process $p(X)$ is continuous for every $p \in P^D$, in particular for every linear monomial $p(\nu) = \langle h, \nu \rangle$ with $h \in D$. Choose any $g \in C_\Delta(E)$ and $\varepsilon > 0$. Since D is dense in $C_\Delta(E)$ there is $h \in D$ such that $\|g - h\| \leq \varepsilon$. Therefore,

$$\lim_{s \rightarrow t} |\langle g, X_s \rangle - \langle g, X_t \rangle| \leq \varepsilon + \lim_{s \rightarrow t} |\langle h, X_s \rangle - \langle h, X_t \rangle| = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $\lim_{s \rightarrow t} \langle g, X_s \rangle = \langle g, X_t \rangle$. In particular $X_t(E) = 1 = \lim_{s \rightarrow t} X_s(E)$, and we deduce that X is continuous with respect to the topology of weak convergence on $M_1(E)$.

Conversely, if X is an $M_1(E)$ -solution to the martingale problem for L with continuous paths, then, by Lemma 2.4(i), the map $t \mapsto p(X_t)$ is continuous for all $p \in P^D$. The result now follows by Proposition 1 in [Bakry and Émery \(1985\)](#). \square

4 Lévy type operators

Let $E \subseteq \mathbb{R}^d$ be a closed set. As explained in Remark 2.7, the class of polynomials on $M_1(E)$ with coefficients in $\widehat{C}_\Delta(E^k)$ coincides with the class of polynomials

on $M_1(E)$ with coefficients in $\widehat{C}_0(E^k)$. Till now, we always worked with $C_\Delta(E)$ since this space is more suitable for the study of polynomial jump-diffusions. However, for the content of this section, choosing $C_0(E)$ as space of coefficients is more convenient. The main advantage in this framework is that a polynomial extension from $M_1(E)$ to $M_1(E^\Delta)$ coincides with its extension from $M_1(E)$ to the set of subprobabilities on E , denoted $M_{\leq 1}(E)$. More precisely, for $g \in C_0(E)$ and $\mu \in M_{\leq 1}(E)$ we have that

$$\begin{aligned} \langle g, \mu + (1 - \mu(E))\delta_\Delta \rangle &:= \lim_{|x| \rightarrow \infty} \langle g, \mu + (1 - \mu(E))\delta_x \rangle \\ &= \langle g, \mu \rangle + (1 - \mu(E)) \lim_{|x| \rightarrow \infty} g(x) = \langle g, \mu \rangle, \end{aligned}$$

and thus similarly for $D \subseteq C_0(E)$ and $p \in P^D$ we get

$$p(\mu + (1 - \mu(E))\delta_\Delta) = p(\mu).$$

This in particular implies that a linear operator $L : P^D \rightarrow F(M(E))$ for $D \subseteq C_0(E)$ satisfies the positive maximum principle on $M_1(E^\Delta)$ is and only if it does so on $M_{\leq 1}(E)$.

Because of those considerations and also considering Lemma 3.7(ii), in this section we decided to work with

$$P_c^\infty(\mathbb{R}^d) := \text{span}\{\nu \mapsto \langle g, \nu \rangle^k : g \in C_c^\infty(\mathbb{R}^d), k \in \mathbb{N}_0\}.$$

Observe that the name of the set is mnemonic for the space $C_c^\infty(\mathbb{R}^d)$ of coefficients. In this framework, we denote by $M_{\leq 1}^v(E)$ the space of subprobabilities on E endowed with the topology of vague convergence.³ Note that, as $M_1(E^\Delta)$, $M_{\leq 1}^v(E)$ is a locally compact Polish space.

Consider a linear operator $L : P^D \rightarrow F(M(E))$. As in the finite dimensional case, the positive maximum principle imposes some structure to the operator L (see e.g. Courrège (1965), Section 2.2 in Hoh (1998), but also Theorem II.2.8 for operators on functions of E -arguments, where E is a compact subset of \mathbb{R}^d).

Definition 4.1. Fix $\mu \in M_{\leq 1}(E)$. The operator L is said to be of Lévy type at μ if the operator mapping $P_c^\infty(\mathbb{R}^d)$ to \mathbb{R} given by $p \mapsto Lp(\mu)$ can be represented as

$$\begin{aligned} Lp(\mu) &= -\Gamma(\mu)p(\mu) + B(\partial p(\mu), \mu) \\ &\quad + \frac{1}{2}Q(\partial^2 p(\mu), \mu) + \int p(\xi) - p(\mu) - \langle \partial p(\mu), \xi - \mu \rangle N(\mu, d\xi), \end{aligned} \quad (4.1)$$

where the quadruplet $(\Gamma(\mu), B(\cdot, \mu), Q(\cdot, \mu), N(\mu, \cdot))$ consists of some constant $\Gamma(\mu) \in \mathbb{R}_+$, some linear operators

$$B(\cdot, \mu) : C_c^\infty(\mathbb{R}^d) \rightarrow \mathbb{R}, \quad \text{and} \quad Q(\cdot, \mu) : C_c^\infty(\mathbb{R}^d) \otimes C_c^\infty(\mathbb{R}^d) \rightarrow \mathbb{R},$$

³Recall that a sequence $(\mu_n)_{n \in \mathbb{N}} \subseteq M(E)$ is said to converge vaguely to $\mu \in M(E)$ if $\langle g, \mu_n \rangle$ converges to $\langle g, \mu \rangle$ for all $g \in C_c(\mathbb{R}^d)$.

and some (positive) measure $N(\mu, \cdot)$ on $M_{\leq 1}(E) \setminus \{\mu\}$ satisfying

$$Q(g \otimes g, \mu) \geq 0 \quad \text{and} \quad \int \langle g, \xi - \mu \rangle^2 N(\mu, d\xi) < \infty,$$

for all $g \in C_c^\infty(\mathbb{R}^d)$. If L is a Lévy type operator at μ for all $\mu \in M_{\leq 1}(E)$, then it is said to be of Lévy type.

Observe that because of Lemma 2.8(ii) the operators B and Q need not to be defined on the whole spaces $C_0(\mathbb{R}^d)$ and $\widehat{C}_0((\mathbb{R}^d)^2)$, respectively. This will in particular allow them to be differential operators.

The proof of the next theorem follows the proof of its finite dimensional analogue given by Theorem II.2.8. The only technical difference is that in the finite dimensional setting the square of the norm is a polynomial (and thus a test function) whose value in x is 0 if and only if $x = 0$. Such a test function does not exist in the present setting and we thus need to work with a separating sequence, i.e. a sequence $(g_n)_{n \in \mathbb{N}} \subseteq C_c^\infty(\mathbb{R}^d)$ such that for $\mu \in M_{\leq 1}(E)$ it holds $\langle g_n, \mu \rangle = 0$ for all $n \in \mathbb{N}$ if and only if $\mu = 0$. The proof can be found in Section A.

Theorem 4.2. *Fix $\mu \in M_{\leq 1}(E)$. If L satisfies the positive maximum principle on $M_{\leq 1}(E)$ (or equivalently on $M_1(E)$) at μ , then L is of Lévy type at μ .*

For $E \subseteq \mathbb{R}^d$ being also compact the space $M_1(E)$ is closed with respect to the vague (or equivalently weak) topology. This yields the following refinement of Theorem 4.2.

Lemma 4.3. *Consider the setting of Theorem 4.2 and suppose that $E \subseteq \mathbb{R}^d$ is also compact. Then, the measure $N(\mu, d\xi)$ appearing in (4.1) is a measure on $M_1(E) \setminus \{\mu\}$.*

Proof. Choose $g \in C_c^\infty(\mathbb{R}^d)$ such that $g|_E = 1$ and note that $p(\nu) = (1 - \langle g, \nu \rangle)^4$ satisfies $p(\nu) = 0$ for all $\nu \in M_1(E)$. The positive maximum principle then yields $0 = Lp(\mu) = \int p(\xi)N(\mu, d\xi)$ proving that $\xi(E) = \langle g, \xi \rangle = 1$ for $N(\mu, \cdot)$ -a.e. $\xi \in M_{\leq 1}(E)$. \square

Observe that by Theorem 4.2, if L satisfies the positive maximum principle on $M_1(E)$ we can define some operators $\Gamma : M_1(E) \rightarrow \mathbb{R}_+$,

$$B : C_c^\infty(\mathbb{R}^d) \times M_1(E) \rightarrow \mathbb{R}, \quad \text{and} \quad Q : C_c^\infty(\mathbb{R}^d) \otimes C_c^\infty(\mathbb{R}^d) \times M_1(E) \rightarrow \mathbb{R},$$

and some kernel $N(\cdot, d\xi)$ from $M_1(E)$ to $M_{\leq 1}(E)$ such that L is a Lévy type operator at μ where the corresponding quadruplet is given by

$$(\Gamma(\mu), B(\cdot, \mu), Q(\cdot, \mu), N(\mu, \cdot))$$

for each $\mu \in M_1(E)$. By Lemma 3.5 we know that this is always true if there exists a (possibly killed) $M_1(E)$ -solution to the martingale problem for L for all initial conditions $\mu \in M_1(E)$.

If we additionally have that $Lp|_{M_{\leq 1}(E)} \in C(M_{\leq 1}^v(E))$ for all $p \in P_c^\infty(\mathbb{R}^d)$, by Lemma 3.7(ii) we get that L satisfies the positive maximum principle on $M_{\leq 1}(E)$. In this case, the form of L explained before extends to all $\mu \in M_{\leq 1}(E)$.

The previous considerations illustrate how all the relevant linear operators L corresponding to a jump-diffusion need in fact to be of Lévy type. This conclusion strongly supports the choice to work with the notion of derivative introduced in Section 2.2. Secondly, it decomposes the complexity of a probability measure-valued jump-diffusion into four (deterministic) parameters Γ, B, Q, N , each of them with a specific, classical, interpretation. In Lemma 4.5, Lemma 4.6, and by means of numerous examples (see Section IV.5.2) we will deepen this aspect. Finally, but not less important, by Lemma 3.6 we know that the positive maximum principle on $M_1(E)$ is the first step in order to guarantee the existence of an $M_1(E)$ -solution to the martingale problem for L . From a practical point of view, checking the positive maximum principle always reduces to finding (necessary) optimality conditions for the functions in its domain. Knowing a priori that L is a Lévy type operator gives us an intuition about the form of the optimality conditions we would like to have, and guarantees that the conditions we will present in Section 5 are of effective utility.

In fact, more properties can be deduced using the positive maximum principle, as we illustrate in the next lemma.

Lemma 4.4. *Let $\mu \in M_1(E)$, L be a Lévy type operator at μ , and suppose that it satisfies the positive maximum principle on $M_1(E)$ at μ . Then*

$$\langle g, \mu \rangle = \sup_E g \quad \Rightarrow \quad Q(g \otimes g, \mu) = 0 \quad \text{and} \quad B(g, \mu) - \int \langle g, \xi - \mu \rangle N(\mu, d\xi) \leq 0,$$

and

$$\langle g^2, \mu \rangle = 0 \quad \Rightarrow \quad Q(g \otimes g, \mu) + \int \langle g, \xi - \mu \rangle^2 N(\mu, d\xi) \leq B(g^2, \mu).$$

Proof. Without loss of generality suppose that $g \geq 0$ and $\sup_E g = 1$. Let then $(F_n)_{n \in \mathbb{N}}$ be polynomials on $[0, 1]$ with $0 \leq F_n \leq 1$, $F_n(0) = 1$, $xnF_n(x) \leq 1$, and $F_n(x) \downarrow 0$ for $x \in (0, 1]$. For example, one can choose $F_n(x) := \frac{n-1}{n}x^n + \frac{1}{n}$. Let $\phi_n(x) := \frac{1-x}{n} - (1-x)^2 F_n(x)$. Then ϕ_n has a minimum at $x = 1$, so by the positive maximum principle,

$$0 \leq L(\phi_n(\langle g, \cdot \rangle))(\mu) = -\frac{1}{n}B(g, \mu) - Q(g \otimes g, \mu) - \int F_n(\langle g, \xi \rangle)(1 - \langle g, \xi \rangle)^2 N(\mu, d\xi)$$

and hence, by the dominated convergence theorem, $Q(g \otimes g, \mu) \leq 0$. Similarly, $\psi_n(x) := (1-x)F_n(x)$ is nonnegative on $[0, 1]$ with a minimum at $x = 1$, so

$$0 \leq L(\psi_n(\langle g, \cdot \rangle))(\mu) = -B(g, \mu) + \int \langle g, \xi - \mu \rangle (1 - F_n(\langle g, \xi \rangle)) N(\mu, d\xi)$$

and the monotone convergence theorem yields $B(g, \mu) - \int \langle g, \xi - \mu \rangle N(\mu, d\xi) \leq 0$. The last bound follows by noting that $p(\nu) = \langle g^2, \nu \rangle - \langle g, \nu \rangle^2$ has a minimum in

μ and hence

$$0 \leq Lp(\mu) = B(g^2, \mu) - Q(g \otimes g, \mu) - \int \langle g, \xi - \mu \rangle^2 N(\mu, d\xi).$$

□

We now provide two results, illustrating how some interesting information about the corresponding $M_1(E)$ -solution to the martingale problem can be read from the parameters Γ , B , Q , and N . As for $C(M_1(E))$ and $C(M_{\leq 1}(E))$, with a small abuse of notation we will write $p \in C(M_{\leq 1}(E))$ if the map $p|_{M_{\leq 1}(E)} : M_{\leq 1}^v(E) \rightarrow \mathbb{R}$ is continuous.

As in Section 3.3, the result concerning continuity of $M_1(E)$ -solutions is based on Bakry and Émery (1985).

Lemma 4.5. *Let $L : P_c^\infty(\mathbb{R}^d) \rightarrow C(M_{\leq 1}^v(E))$ be a linear operator.*

- (i) *If $\Gamma = 0$, then any possibly killed $M_1(E)$ -solution to the martingale problem for L is conservative. Conversely, if for every initial condition $\nu \in M_1(E)$ there is a conservative $M_1(E)$ -solution to the martingale problem for L , then $\Gamma = 0$.*
- (ii) *If $\Gamma = 0$ and $N = 0$, then any possibly killed $M_1(E)$ -solution to the martingale problem for L is conservative and has continuous paths. Conversely, if for every initial condition $\nu \in M_1(E)$ there is a conservative $M_1(E)$ -solution to the martingale problem for L with continuous paths, then $\Gamma = 0$ and $N = 0$.*

Proof. (i) Since $L1 = -\Gamma$, the conclusion follows by Lemma 3.5 and Lemma 3.6.

(ii) From Lemma 3.11, it is sufficient to show that $\Gamma = 0$ and $N = 0$ if and only if the carré-du-champ operator of L is an $M_1(E)$ -derivation, i.e.

$$L(p_1 p_2 p_3)(\mu) = \sum_{i=1}^3 \sum_{j \neq i} \sum_{k \neq i, j} p_i(\mu) L(p_j p_k)(\mu) - p_i(\mu) p_j(\mu) L p_k(\mu), \quad (4.2)$$

for all $p_1, p_2, p_3 \in P_c^\infty(\mathbb{R}^d)$ and $\mu \in M_1(E)$. Noting that by Lemma 2.8(i) we know that $\partial(p_1 p_2) = p_1 \partial p_2 + p_2 \partial p_1$, a direct computation shows that if $\Gamma = 0$ and $N = 0$ condition (4.2) holds true. Conversely, applying condition (4.2) to $p_1 = p_2 = p_3 = 1$ directly yields $\Gamma = 0$. Fix then $\mu \in M_1(E)$. Applying condition (4.2) to $p_1(\nu) = p_2(\nu) = p_3(\nu) = (\langle g, \nu \rangle - \langle g, \mu \rangle)^2$ then yields

$$\int (\langle g, \nu \rangle - \langle g, \mu \rangle)^6 N(\mu, d\xi) = 0, \quad g \in C_c^\infty(\mathbb{R}^d),$$

proving that $N(\mu, \cdot) = 0$. □

The next lemma provides an expression for the predictable quadratic variation and covariation of cylindrical polynomials evaluated in probability measure-valued jump-diffusions. Due to the conflict of notation with the monomials, we indicate those two quantities with $\text{PQV}(\cdot)$ and $\text{PQC}(\cdot, \cdot)$.

Since the result holds true also for cylindrical functions (in sense of Section 2.4), we state the result in this more general framework.

Lemma 4.6. Fix $p, q : M(E) \rightarrow \mathbb{R}$ such that $p, q \in C^2, C_c^\infty(\mathbb{R}^d)$. Let L be a Lévy type operator and X be an $M_1(E)$ -solution to the martingale problem for L . The predictable quadratic variation of $p(X)$ is then given by

$$\text{PQV}(p(X)) = \int Q(\partial p(X_s) \otimes \partial p(X_s), X_s) ds.$$

Similarly, the predictable quadratic covariation of the pair $(p(X), q(X))$ reads

$$\text{PQC}(p(X), q(X)) = \int Q(\partial p(X_s) \otimes \partial q(X_s), X_s) ds.$$

Proof. Fix $g_1, g_2 \in C_c^\infty(\mathbb{R}^d)$ and observe that an application of the Itô formula yields

$$\begin{aligned} \langle g_1, X_t \rangle \langle g_2, X_t \rangle &= \langle g_1, X_0 \rangle \langle g_2, X_0 \rangle + \sum_{i \neq j} \int_0^t \langle g_i, X_s \rangle d\langle g_j, X_s \rangle \\ &\quad + \text{PQC}(\langle g_1, X \rangle, \langle g_2, X \rangle) + \sum_{s \leq t} \Delta \langle g_1, X_s \rangle \Delta \langle g_2, X_s \rangle. \end{aligned}$$

(3.1) and the form of L then yield

$$\begin{aligned} \langle g_i, X_t \rangle &= \langle g_i, X_0 \rangle + \int_0^t B(g_i, X_s) ds + (\text{martingale}), \quad i \in \{1, 2\}, \\ \langle g_1, X_t \rangle \langle g_2, X_t \rangle &= \langle g_1, X_0 \rangle \langle g_2, X_0 \rangle + \sum_{i \neq j} \int_0^t \langle g_i, X_s \rangle B(g_j, X_s) ds \\ &\quad + \int_0^t Q(g_1 \otimes g_2, X_s) ds + \sum_{s \leq t} \Delta \langle g_1, X_s \rangle \Delta \langle g_2, X_s \rangle + (\text{martingale}), \end{aligned}$$

proving that $\text{PQC}(\langle g_1, X \rangle, \langle g_2, X \rangle) = \int Q(g_1 \otimes g_2, X_s) ds$. The result for the general case follows by the linearity of Q and Lemma 2.8(i). \square

We conclude this section giving a second look at Lemma 3.8, under the assumption that L is a Lévy type operator.

Remark 4.7. Observe that if $\Gamma = 0$ and $B = 0$, condition (3.2) in Lemma 3.6 is automatically satisfied. This choice of the parameters automatically implies that every possibly killed $M_{\leq 1}^v(E)$ -solution X to the martingale problem for L is a conservative $M_1(E)$ -solution and also a martingale (in the sense of Definition 3.3). By Lemma 3.8(i) we also get that in this case for all $A \subseteq E$ and $t \geq s$,

$$\mathbb{P}(X_t(A) = 1 | X_s(A) = 1) = 1.$$

It is important to note that this condition does not imply that if X_0 is nonatomic the same is true for X_t . Example 3.9 shows how this can in fact happen.

Lemma 4.8. *Consider a Lévy type operator L for $\Gamma = 0$, and let X be the corresponding $M_1(E)$ -solution to the martingale problem with initial condition μ . Let m_p be as in Lemma 3.8(iii) and suppose also that there exists a nonnegative sequence $(g_n)_{n \in \mathbb{N}} \subseteq C_c^\infty(\mathbb{R}^d)$ such that g_n increases to m_p and*

$$\sup_n \sup_{M_1(E)} \left(\frac{B(g_n, \cdot)^+}{1 + \langle g_n, \cdot \rangle} \right) < \infty.$$

Then $\langle |\cdot|^p, \mu \rangle < \infty$ yields $\langle |\cdot|^p, X_t \rangle < \infty$ a.s. for all $t \geq 0$.

Proof. Set p_n as in Lemma 3.8(iii) for $\psi(x) = \log(x+1)$ and note that

$$\partial p_n(\mu) = (\langle g_n, \mu \rangle + 1)^{-1} g_n \quad \text{and} \quad \partial^2 p_n(\mu) = -(\langle g_n, \mu \rangle + 1)^{-2} g_n \otimes g_n.$$

Since $\Gamma(\mu) \geq 0$, $Q(g_n \otimes g_n, \mu) \geq 0$, and $\log(x+1) - x \leq 0$ for all $\mu \in M_1(E)$, $n \in \mathbb{N}$, and $x > -1$, we get that $(Lp_n(\mu))^+ \leq B(g_n, \mu)^+(1 + \langle g_n, \mu \rangle)^{-1}$. The result follows by Lemma 3.8 (iii). \square

5 Optimality conditions

Let E be a locally compact Polish space. We now develop optimality conditions for polynomials of measure arguments, which are instrumental when working with the positive maximum principle on $M_1(E^\Delta)$. Our first result, Theorem 5.1, extends the classical first and second order Karush–Kuhn–Tucker conditions for functions on the finite dimensional unit simplex (see e.g. Bertsekas (1995)). It is derived by slightly changing an optimizer $\nu_* \in M_1(E^\Delta)$ via shifting small amounts of mass to some arbitrary point on E^Δ . Our second result, Theorem 5.3, is obtained by deforming the optimizer ν_* using a group of isometries of $C_\Delta(E)$. The resulting condition is genuinely infinite dimensional, as can be seen from Lemma 5.5.

5.1 Generalization of the classical Karush–Kuhn–Tucker conditions

We use the operator Ψ , which maps any function $g: E \times E \rightarrow \mathbb{R}$ to the function $\Psi(g): E \times E \rightarrow \mathbb{R}$ given by

$$\Psi(g)(x, y) = \frac{1}{2} (g(x, x) + g(y, y) - 2g(x, y)). \quad (5.1)$$

Theorem 5.1. *Let $p \in P$ and $\nu_* \in M_1(E^\Delta)$ satisfy $p(\nu_*) = \max_{M_1(E^\Delta)} p$. Then the following first and second order optimality conditions hold:*

- (i) $\langle \partial p(\nu_*), \mu \rangle = \sup_E \partial p(\nu_*)$, for all $\mu \in M_1(E^\Delta)$ such that $\text{supp}(\mu) \subseteq \text{supp}(\nu_*)$.
In particular,

$$\partial_x p(\nu_*) = \sup_E \partial p(\nu_*) \quad \text{for all } x \in \text{supp}(\nu_*). \quad (5.2)$$

(ii) $\langle \partial^2 p(\nu_*), \mu^2 \rangle \leq 0$ for all signed measures $\mu \in M(E^\Delta)$ such that $\langle 1, \mu \rangle = 0$ and $\text{supp}(|\mu|) \subseteq \text{supp}(\nu_*)$. In particular,

$$\Psi(\partial^2 p(\nu_*))(x, y) \leq 0 \quad \text{for all } x, y \in \text{supp}(\nu_*). \quad (5.3)$$

Proof. Pick any $x \in \text{supp}(\nu_*)$ and $y \in E^\Delta$. For each $n \in \mathbb{N}$, let A_n be the ball of radius $1/n$ centered at x , intersected with $\text{supp}(\nu_*)$. Then $\nu_*(A_n) > 0$, and the probability measures $\mu_n := \nu_*(\cdot \cap A_n)/\nu_*(A_n)$ converge weakly to δ_x as $n \rightarrow \infty$. Choose $\varepsilon_n \in (0, \nu_*(A_n))$. Then $\nu_* \geq \varepsilon_n \mu_n$ since for all $B \in \mathcal{B}(E)$

$$\nu_*(B) - \varepsilon_n \frac{\nu_*(B \cap A_n)}{\nu_*(A_n)} \geq \frac{\nu_*(B \cap A_n)}{\nu_*(A_n)} (\nu_*(A_n) - \varepsilon_n) \geq 0.$$

Hence $\nu_n := \nu_* + \varepsilon_n(\delta_y - \mu_n)$ is a probability measure. Maximality of ν_* and Lemma 2.4(vi) now give

$$0 \geq p(\nu_n) - p(\nu_*) = \varepsilon_n \langle \partial p(\nu_*), \delta_y - \mu_n \rangle + o(\varepsilon_n).$$

Dividing by ε_n , sending n to infinity, and using that $x \mapsto \partial_x p(\nu_*)$ is bounded and continuous, we obtain $\partial_x p(\nu_*) \geq \partial_y p(\nu_*)$. We deduce (5.2), which immediately implies (i).

Next, in addition to the above, suppose y is in $\text{supp}(\nu_*)$. Since we also have that $\text{supp}(\mu_n) \subseteq \text{supp}(\nu_*)$, we get $\langle \partial p(\nu_*), \delta_y - \mu_n \rangle = 0$ due to (i). Maximality of ν_* and Lemma 2.4(vi) then give

$$0 \geq p(\nu_n) - p(\nu_*) = \frac{1}{2} \varepsilon_n^2 \langle \partial^2 p(\nu_*), (\delta_y - \mu_n)^2 \rangle + o(\varepsilon_n^2),$$

and therefore $\langle \partial^2 p(\nu_*), (\delta_y - \delta_x)^2 \rangle \leq 0$. More generally, consider measures of the form

$$\nu_n := \nu_* + \varepsilon_n \left(\sum_{i=1}^m \lambda_i \delta_{y_i} - \sum_{i=1}^m \gamma_i \mu_{i,n} \right)$$

for some points $y_i \in \text{supp}(\nu_*)$, convex weights $\lambda_1, \dots, \lambda_m$ and $\gamma_1, \dots, \gamma_m$, and $\mu_{i,n}$ constructed as μ_n above with x replaced by $x_i \in \text{supp}(\nu_*)$. Letting ε_n decrease to zero sufficiently rapidly, the above argument gives $\langle \partial^2 p(\nu_*), \mu^2 \rangle \leq 0$ for the signed measure

$$\mu = \sum_{i=1}^m \lambda_i \delta_{y_i} - \sum_{i=1}^m \gamma_i \delta_{x_i}.$$

Passing to the weak closure yields (ii) with the additional restriction that the positive and negative parts of μ are probability measures. The general case is obtained by scaling. Finally, since $\langle \partial^2 p(\nu_*), (\delta_y - \delta_x)^2 \rangle = 2\Psi(\partial^2 p(\nu_*))(x, y)$ we obtain (5.3). \square

It is very interesting to note the similarity between the conditions of Theorem 5.1 and the classical Karush–Kuhn–Tucker conditions on the finite dimensional simplex Δ^d , that we report here. Let $f \in C^2(\mathbb{R}^d)$ and $x^* \in \Delta^d$ satisfy $f(x^*) = \max_{\Delta^d} f$. Then the first and second order (necessary) Karush–Kuhn–Tucker conditions on Δ^d hold:

(i) For each $v \in \Delta^d$ such that $v_i = 0$ whenever $x_i^* = 0$,

$$\nabla f(x^*)^\top v = \max_{j \in \{1, \dots, d\}} \frac{\partial f}{\partial x_j}(x^*).$$

(ii) For all $v \in \mathbb{R}^d$ such that $\mathbf{1}^\top v = 0$ and $v_i = 0$ whenever $x_i^* = 0$,

$$v^\top \nabla^2 f(x^*) v \leq 0,$$

where $\mathbf{1} := (1, \dots, 1)^\top$.

Remark 5.2. Taking $E = \{1, \dots, d\}$ as example, the appearance of Ψ in (5.3) can be understood as follows. Suppose $z \in \Delta^d$ maximizes a function $f \in C^2(\mathbb{R}^d)$ over Δ^d . Then for every i, j such that $z_i > 0$ and $z_j > 0$, we must have $(e_i - e_j)^\top \nabla^2 f(z)(e_i - e_j) \leq 0$, where e_i is the i -th canonical unit vector. Indeed, otherwise $z \pm \varepsilon(e_i - e_j)$ would lie in Δ^d and give a higher function value for small $\varepsilon > 0$. More explicitly, we must have

$$\partial_{ii}^2 f(z) + \partial_{jj}^2 f(z) - 2\partial_{ij}^2 f(z) \leq 0,$$

where the left hand side is equal to $2\Psi(\partial^2 f(z))(i, j)$ on $E = \{1, \dots, d\}$.

5.2 An optimality condition beyond the finite dimensional setting

In the remaining part of the section, $D \subseteq C_\Delta(E)$ is a linear subspace, and P^D is defined by (2.3).

Our next optimality condition is more subtle, in that it becomes trivial in the finite-dimensional case; see Lemma 5.5. In order to understand the interpretation of this result, let us first focus on its application, Corollary 5.7. This condition is obtained by slightly changing an optimizer ν^* via shifting its support in a way that guarantees that the resulting measure is still supported on E . The next theorem generalizes this result. The operator A described there generates a group of isometries $\{T_t\}_{t \in \mathbb{R}}$ of $C_\Delta(E)$. For any $\mu \in M_+(E^\Delta)$, the group induces a flow of measures $\mu_t \in M_+(E^\Delta)$ via the formula $\langle g, \mu_t \rangle = \langle T_t g, \mu \rangle$ for every $g \in C_\Delta(E)$. The value of a polynomial in its maximizer ν_* cannot be less than its value in $\nu_* - \mu + \mu_t$, for any t , and this leads to an optimality condition in terms of the group generator A .

The tensor notation $A \otimes A$ is used to denote the linear operator from $D \otimes D$ to $\widehat{C}_\Delta(E^2)$ determined by

$$(A \otimes A)(g \otimes g) := (Ag) \otimes (Ag)$$

for a given linear operator $A: D \rightarrow C_\Delta(E)$.

Theorem 5.3. Let $p \in P^D$ and $\nu_* \in M_1(E^\Delta)$ satisfy $p(\nu_*) = \max_{M_1(E^\Delta)} p$. Let A be the generator of a strongly continuous group of positive isometries of $C_\Delta(E)$, and assume the domain of A contains both D and $A(D)$. Then

$$\langle A^2(\partial p(\nu_*)), \mu \rangle + \langle (A \otimes A)(\partial^2 p(\nu_*)), \mu^2 \rangle \leq 0$$

for every $\mu \in M_+(E^\Delta)$ with $\mu \leq \nu_*$.

Proof. Let $\{T_t\}_{t \in \mathbb{R}}$ be the group generated by A . For any $\mu \in M_+(E^\Delta)$, the group induces a flow of measures $\mu_t \in M(E^\Delta)$ via the formula

$$\langle g, \mu_t \rangle = \langle T_t g, \mu \rangle \quad \text{for} \quad g \in C_\Delta(E).$$

The positivity and isometry property of T_t implies that μ_t is nonnegative and has constant total mass $\mu_t(E^\Delta) = \mu(E^\Delta)$. Therefore, assuming henceforth that $\mu \leq \nu_*$, it follows that $\nu_* + \mu_t - \mu$ is a probability measure. Since $\|T_t g - g\| = O(t)$ for every $g \in D$, we have $\langle g, (\mu_t - \mu)^k \rangle = O(t^k)$ for every $g \in D^{\otimes k}$. Maximality of ν_* and Lemma 2.4(vi) then give

$$\begin{aligned} 0 &\geq p(\nu_* + \mu_t - \mu) - p(\nu_*) \\ &= \langle \partial p(\nu_*), \mu_t - \mu \rangle + \frac{1}{2} \langle \partial^2 p(\nu_*), (\mu_t - \mu)^2 \rangle + o(t^2) \\ &= \langle (T_t - \text{id}) \partial p(\nu_*), \mu \rangle + \frac{1}{2} \langle (T_t \otimes T_t - 2T_t \otimes \text{id} + \text{id} \otimes \text{id}) \partial^2 p(\nu_*), \mu^2 \rangle + o(t^2). \end{aligned} \quad (5.4)$$

We claim that both A and $-A$ satisfy the positive maximum principle on E^Δ . Indeed, for $f \in D$ and $x \in E^\Delta$ with $f(x) = \max_{E^\Delta} f \geq 0$, the positivity and isometry property give

$$T_t f(x) \leq T_t f^+(x) \leq \|T_t f^+\| = \|f^+\| = f(x). \quad (5.5)$$

Thus

$$A f(x) = \lim_{t \downarrow 0} (T_t f(x) - f(x))/t \leq 0 \quad \text{and} \quad -A f(x) = \lim_{t \downarrow 0} (T_{-t} f(x) - f(x))/t \leq 0,$$

proving the claim. Since $\partial_x p(\nu_*) = \sup_E \partial p(\nu_*)$ for all $x \in \text{supp}(\nu_*)$ due to Theorem 5.1, it follows that $A(\partial p(\nu_*))(x) = 0$ for all such x . As a result, using that $\text{supp}(\mu) \subseteq \text{supp}(\nu_*)$ and that the domain of A contains $A(D)$, we get

$$\langle (T_t - \text{id}) \partial p(\nu_*), \mu \rangle = \langle (T_t - \text{id} - tA) \partial p(\nu_*), \mu \rangle = \frac{1}{2} t^2 \langle A^2(\partial p(\nu_*)), \mu \rangle + o(t^2). \quad (5.6)$$

Furthermore, using that

$$(T_t \otimes T_t - 2T_t \otimes \text{id} + \text{id} \otimes \text{id})(g \otimes g) = (T_t g - g) \otimes (T_t g - g)$$

for all $g \in D$, we deduce that

$$\langle (T_t \otimes T_t - 2T_t \otimes \text{id} + \text{id} \otimes \text{id})g, \mu^2 \rangle = t^2 \langle (A \otimes A)g, \mu^2 \rangle + o(t^2) \quad (5.7)$$

for all $g \in D \otimes D$. Inserting (5.6) and (5.7) into (5.4), dividing by t^2 , and sending t to zero yields

$$0 \geq \frac{1}{2} \langle A^2(\partial p(\nu_*)), \mu \rangle + \frac{1}{2} \langle (A \otimes A) \partial^2 p(\nu_*), \mu^2 \rangle.$$

This completes the proof. \square

Remark 5.4. We claim that for A as in Theorem 5.3, the operator A^2 satisfies the positive maximum principle on E^Δ . Indeed, let $f \in D$ and $x \in E^\Delta$ with $f(x) = \max_{E^\Delta} f \geq 0$. Then, as in (5.5) and with the same notation, we have $T_t f(x) \leq f(x)$, and $Af(x) = 0$ since both A and $-A$ satisfy the positive maximum principle on E^Δ . Hence $A^2 f(x) = \lim_{t \downarrow 0} (T_t f(x) - f(x) - Af(x))/t \leq 0$, which proves the claim.

The next lemma illustrates the pure infinite dimensional nature of the condition provided in Theorem 5.3.

Lemma 5.5. Let A be the generator of a strongly continuous group of positive isometries of $C_\Delta(E)$. If the domain of A is all of $C_\Delta(E)$, then $A = 0$. This is in particular the case if A is bounded or E consists of finitely many points.

Proof. Note that A and $-A$ satisfy the positive maximum principle on E , and $A1 = 0$. Therefore Lemma B.1 implies that A and $-A$ are both of the form (B.1) with $B = \pm A$. As a result,

$$0 = Ag(x) - Ag(x) = \int (g(\xi) - g(x))(\nu_A + \nu_{-A})(x, d\xi)$$

for all $x \in E$ and $g \in C(E^\Delta)$. This implies that both $\mathbb{1}_{\{x\}^c}(\xi)\nu_A(x, d\xi)$ and $\mathbb{1}_{\{x\}^c}(\xi)\nu_{-A}(x, d\xi)$ are zero for all $x \in E$ and hence that $A = 0$. Since each linear operator on a finite dimensional vector space is bounded, and the domain of a bounded operator on $C_\Delta(E)$ can be extended to all of $C_\Delta(E)$, the second part follows. \square

5.3 An application of Theorem 5.3

Since Theorem 5.3 is quite abstract, we provide a typical example of its application. We will see that this is particularly important in the context of particle system. Throughout this section $E \subseteq \mathbb{R}^d$ is a closed subset and $D \subseteq C_\Delta^2(\mathbb{R}^d)$. For $M \in \mathbb{R}^{d \times d}$ we write M_j for the j -th column of M and M_{ij} for its ij -th entry. Finally, $\Sigma^d(E)$ denotes the set of possible diffusion matrices for a process on E , namely

$$\Sigma^d(E) := \left\{ \tau : E \rightarrow \mathbb{R}^{d \times d} : \tau_{ij} \in C_\Delta^1(E), \quad \text{and} \right. \quad (5.8)$$

$$\left. \tau(x)^\top \nabla f(x) = 0 \quad \text{for } x \in E, f \in C_\Delta^2(E) : f(x) = \max_E f \right\}.$$

Lemma 5.6. For $\tau \in \Sigma^d(E)$, the operator on $C_\Delta(E)$ given by

$$A_j g := \tau_j^\top \nabla g \quad \text{for all } g \in D$$

satisfies the conditions of Theorem 5.3, for $j \in \{1, \dots, d\}$.

The resulting statement is the following. Since it just consists in a reformulation of Theorem 5.3, we omit the proof.

Corollary 5.7. Fix $\tau \in \Sigma^d(E)$. Let $p \in P^D$ and $\nu_* \in M_1(E^\Delta)$ satisfy $p(\nu_*) = \max_{M_1(E^\Delta)} p$ and for all $g \in D$ set

$$B_\tau g := \sum_{j=1}^d (\tau_j^\top \nabla)^2 g(x) = \sum_{j=1}^d \tau_j(x)^\top D\tau_j(x) \nabla g(x) + \text{Tr}(\tau(x) \tau(x)^\top \nabla^2 g(x)),$$

$$Q_\tau(g \otimes g) := \text{Tr}((\tau^\top \nabla g) \otimes (\tau^\top \nabla g)^\top), \quad (5.9)$$

where for $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$, Df denotes the Jacobian of f .⁴ Then

$$\langle B_\tau(\partial p(\nu_*)), \nu_* \rangle + \langle Q_\tau(\partial^2 p(\nu_*)), \nu_*^2 \rangle \leq 0.$$

As part of the proof we will see that the involved operators B_τ and Q_τ are well-defined, which means, that $(\tau_j^\top \nabla)^2 g$ and $\text{Tr}((\tau^\top \nabla g) \otimes (\tau^\top \nabla g)^\top)$ only depend on g through its values on E . This permits us to use the expressions given in (5.9) without specifying with respect to which representative of τ_j and g the appearing derivatives are computed.

Proof of Lemma 5.6. We first prove the well-definedness. Fix $\tau, \sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ enough differentiable such that $\tau|_E = \sigma|_E \in \Sigma^d(E)$. Set then $A_j g := \tau_j^\top \nabla g = \sigma_j^\top \nabla g$ for all $g \in C_\Delta^2(\mathbb{R}^d)$ and note that condition (5.8) then yields that $A_j, -A_j$, and thus A_j^2 satisfy the positive maximum principle on E^Δ . Then, for all $g_1, g_2 \in C_\Delta^2(\mathbb{R}^d)$ such that $g_1|_E = g_2|_E$ we get that $(\tau_j^\top \nabla)^2 g_1 = (\sigma_j^\top \nabla)^2 g_2$ and

$$\text{Tr}((\tau^\top \nabla g_1) \otimes (\tau^\top \nabla g_1)^\top) = \text{Tr}((\sigma^\top \nabla g_2) \otimes (\sigma^\top \nabla g_2)^\top)$$

on E and E^2 respectively.

For the second part, note that by Proposition 2.5 in Da Prato and Frankowska (2004) there exist maps $(t, x) \mapsto \phi_j(t, x)$ from $\mathbb{R} \times E$ to E such that

$$\frac{\partial}{\partial t} \phi_j(t, x) = \tau_j(\phi_j(t, x)), \quad \phi_j(0, x) = x,$$

where τ_j defines the j -th column of τ . The strongly continuous group of isometries of $C_\Delta(E)$ corresponding to A_j is then given by $T_t g(x) = g(\phi_j(t, x))$, where the continuity of $T_t g$ is guaranteed by the local Lipschitz continuity of τ . Finally, $\tau_{ij} \in C_\Delta^1(E)$ yields that the domain of A_j contains both D and $A_j(D)$, proving that the linear operator A_j satisfies the conditions of Theorem 5.3. \square

Choosing $E = \mathbb{R}^d$, the state space has no boundaries anymore and the conditions on τ can be relaxed. The resulting condition does not fall in the framework of Theorem 5.3, but can be proved directly.

Lemma 5.8. Fix $E = \mathbb{R}^d$ and $\tau : E \rightarrow \mathbb{R}^{d \times d}$ such that $\tau_{ij} \in C_\Delta(E)$ and $\tau(x)$ is positive semidefinite for all $x \in E$. Let again $p \in P^D$ and $\nu_* \in M_1(E^\Delta)$ satisfy $p(\nu_*) = \max_{M_1(E^\Delta)} p$ and for all $g \in D$ set

$$B_\tau g := \text{Tr}(\tau(x) \tau(x)^\top \nabla^2 g(x)) \quad \text{and} \quad Q_\tau(g \otimes g) := \text{Tr}((\tau^\top \nabla g) \otimes (\tau^\top \nabla g)^\top).$$

Then $\langle B_\tau(\partial p(\nu_*)), \nu_* \rangle + \langle Q_\tau(\partial^2 p(\nu_*)), \nu_*^2 \rangle \leq 0$.

⁴Observe that for $d = 1$ these expressions reduce to $B_\tau g = \tau \tau' g' + \tau g''$ and $Q_\tau(g \otimes g) = (\tau g') \otimes (\tau g')$.

Proof. Define $\nu_\varepsilon := f_*^\varepsilon \nu_*$ where $f^\varepsilon(x) := x + \tau_j(x)\varepsilon$ for $x \in \mathbb{R}^d$. Clearly $\nu_\varepsilon \in M_{\leq 1}(\mathbb{R}^d)$ and thus by the maximality of ν_* and Lemma 2.8(iv)

$$\begin{aligned} 0 &\geq p(\nu_\varepsilon) - p(\nu_*) = \langle \partial_{f^\varepsilon} p(\nu_*) - \partial p(\nu_*), \nu_* \rangle \\ &\quad + \frac{1}{2} \langle \partial^2 p(\nu_*) - 2\partial_{(\cdot)f^\varepsilon}^2 p(\nu_*) + \partial_{f^\varepsilon f^\varepsilon}^2 p(\nu_*), \nu_*^2 \rangle + o(\varepsilon^2). \end{aligned}$$

Since Theorem 5.1(i) yields $\nabla \partial_y p(\nu_*) = 0$ for all $y \in \text{supp}(\nu_*)$, dividing the above expression by ε^2 , letting ε go to 0, and summing over $1 \leq j \leq d$ we get the result. \square

A Proof of Theorem 4.2

Consider the function $F_g : M_{\leq 1}(E) \rightarrow \mathbb{R}$ as $F_g(\nu) := \langle g, \nu - \mu \rangle$. Define $\Gamma(\mu) \in \mathbb{R}$ and $B(\cdot, \mu) : C_c^\infty(\mathbb{R}^d) \rightarrow \mathbb{R}$ as

$$\Gamma(\mu) := -L1(\mu) \quad \text{and} \quad B(g, \mu) := LF_g(\mu),$$

and note that $B(\cdot, \mu)$ heritates the linearity of L . By the positive maximum principle, we also know that $\Gamma(\mu)$ is nonnegative on $M_1(E)$.

Consider the operator $L_g : P_c^\infty(\mathbb{R}^d) \rightarrow \mathbb{R}$ given by $L_g p := L(F_g^2 p)(\mu)$. Observe that by the positive maximum principle $L_g p_1 = L_g p_2$ whenever $p_1|_{M_{\leq 1}(E)} = p_2|_{M_{\leq 1}(E)}$, $L_g p$ is nonnegative whenever $p|_{M_{\leq 1}(E)}$ is nonnegative, and

$$\sup_{\|p\|=1} |L_g(p)| = |L_g(1)| < \infty.$$

This implies that L_g is a well-defined bounded positive operator on $\{p|_{M_{\leq 1}(E)} : p \in P_c^\infty(\mathbb{R}^d)\}$. By Lemma 2.8(iii), it can thus be extended to a bounded positive (linear) operator on $C(M_{\leq 1}^v(E))$. Since $M_{\leq 1}^v(E)$ is a compact Polish space, the Riesz-Markov representation theorem applies and yields

$$L(F_g^2 p)(\mu) = L_g(p) = \int p(\xi) N(g, \mu, d\xi) + p(\mu) G(g, \mu), \quad (\text{A.1})$$

for some $G(g, \mu) \geq 0$ and some (positive) measure $N(g, \mu, \cdot)$ on $M_{\leq 1}(E) \setminus \{\mu\}$. Our goal is now to construct a measure $N(\mu, \cdot)$ on $M_{\leq 1}(E)$ such that

$$F_g(\xi)^2 N(\mu, d\xi) = \mathbf{1}_{\{F_g \neq 0\}}(\xi) N(g, \mu, d\xi) \quad (\text{A.2})$$

for all $g \in C_c^\infty(\mathbb{R}^d)$. Fix a separating sequence, i.e. a sequence $(g_n)_n \subseteq C_c^\infty(\mathbb{R}^d)$ such that $\nu = \mu$ whenever $F_{g_n}(\nu) = 0$ for all $n \in \mathbb{N}$. Define

$$F_K(\nu) := \sum_{n=1}^K F_{g_n}(\nu)^2 \quad \text{and} \quad A_K := \{\nu \in M_{\leq 1}(E) : F_K(\nu) \neq 0\}.$$

Observe that the sequence $(A_K)_{K \in \mathbb{N}}$ is increasing and converges to $M_{\leq 1}(E) \setminus \{\mu\}$. For all set $A \subseteq A_K$ we can then define

$$N(\mu, A) := \int_A F_K(\xi)^{-1} \mathbb{1}_{A_K}(\xi) \sum_{n=1}^K N(g_n, \mu, d\xi).$$

Observe that for any $p \in P_c^\infty(\mathbb{R}^d)$ such that $\text{supp}(p) \cap M_{\leq 1}(E) \subseteq A_K$ we can compute

$$\begin{aligned} & \int \frac{F_K(\xi) F_{K+1}(\xi) p(\xi)}{F_K(\xi)} \sum_{n=1}^K N(g_n, \mu, d\xi) \\ &= \sum_{n=1}^K L_{g_n}(F_{K+1} p) = \sum_{n=1}^K L \left(F_{g_n}^2 \sum_{m=1}^{K+1} F_{g_m}^2 p \right) (\mu) = \sum_{m=1}^{K+1} L_{g_m}(F_K p) \\ &= \int \frac{F_K(\xi) F_{K+1}(\xi) p(\xi)}{F_{K+1}(\xi)} \sum_{m=1}^{K+1} N(g_m, \mu, d\xi), \end{aligned}$$

proving that

$$F_K(\xi)^{-1} \mathbb{1}_{A_K}(\xi) \sum_{n=1}^K N(g_n, \mu, d\xi) = F_{K+1}(\xi)^{-1} \mathbb{1}_{A_K}(\xi) \sum_{n=1}^{K+1} N(g_n, \mu, d\xi)$$

and hence that $N(\mu, \cdot)$ is well-defined on A_K . This permits to define the measure $N(\mu, \cdot)$ on $M_{\leq 1}(E)$ as

$$N(\mu, A) = \lim_{K \rightarrow \infty} N(\mu, A \cap A_K), \quad A \in \mathcal{B}(M_{\leq 1}^v(E)),$$

where $\mathcal{B}(M_{\leq 1}^v(E))$ denotes the Borel σ -algebra on $M_{\leq 1}^v(E)$. In order to prove (A.2), fix now $g \in C_c^\infty(\mathbb{R}^d)$. Observe that for each $p_K \in P_c^\infty(\mathbb{R}^d)$ such that $\text{supp}(p_K) \cap M_{\leq 1}(E) \subseteq A_K \cap \text{supp}(F_g)$ we can compute

$$\int p_K(\xi) F_K(\xi) F_g(\xi)^2 N(\mu, d\xi) = L(p_K F_g^2 F_K)(\mu) = \int p_K(\xi) F_K(\xi) N(g, \mu, d\xi)$$

proving that

$$F_g(\xi)^2 \mathbb{1}_{A_K}(\xi) N(\mu, d\xi) = \mathbb{1}_{A_K}(\xi) \mathbb{1}_{\{F_g \neq 0\}}(\xi) N(g, \mu, d\xi)$$

and hence condition (A.2).

Finally, set $Q(\cdot, \mu) : C_c^\infty(\mathbb{R}^d) \otimes C_c^\infty(\mathbb{R}^d) \rightarrow \mathbb{R}$ as

$$Q \left(\sum_{i=1}^n \lambda_i g_i \otimes g_i, \mu \right) := L \left(\sum_{i=1}^n \lambda_i F_{g_i}^2 \right) (\mu) - \sum_{i=1}^n \lambda_i N(g_i, \mu, M_{\leq 1}(E))$$

and note that it is well-defined by (A.2) and linear in the first argument by linearity of L . Equation (A.1) for $p \equiv 1$ then yields $Q(g \otimes g, \mu) = G(g, \mu) \geq 0$.

Fix $p(\nu) := \langle g, \nu \rangle^k$ and note that by Lemma 2.4(ii) and (vi)

$$p(\nu) = p(\mu) + \langle \partial p(\mu), \nu - \mu \rangle + \langle g, \nu - \mu \rangle^2 f_g(\nu), \quad \nu \in M_{\leq 1}(E) \quad (\text{A.3})$$

where $f_g(\nu) := \sum_{\ell=2}^k \binom{k}{\ell} \langle g, \nu - \mu \rangle^{\ell-2} \langle g, \mu \rangle^{k-\ell}$. Linearity of L , and (A.1) then yields

$$\begin{aligned} Lp(\mu) &= -\Gamma(\mu)p(\mu) + B(\partial p(\mu), \mu) + f_g(\mu)G(g, \mu) + \int f_g(\xi)N(g, \mu, d\xi) \\ &= -\Gamma(\mu)p(\mu) + B(\partial p(\mu), \mu) + \frac{1}{2}Q(\partial^2 p(\mu), \mu) \\ &\quad + \int p(\xi) - p(\mu) - \langle \partial p(\mu), \xi - \mu \rangle N(\mu, d\xi), \end{aligned}$$

where in the second equality we use (A.2) and again (A.3). By linearity, we can conclude that the result holds true for all $p \in P_c^\infty(\mathbb{R}^d)$ and $\mu \in M_1(E)$. \square

B Auxiliary Lemmas

Lemma B.1. *Let $B: C(E^\Delta) \rightarrow C(E^\Delta)$ be a linear operator. Then $B1 = 0$ and B satisfies the positive maximum principle on E if and only if there is a (nonnegative, finite) kernel ν_B from E to E^Δ such that*

$$Bg(x) = \int (g(\xi) - g(x))\nu_B(x, d\xi) \quad (\text{B.1})$$

for all $x \in E$ and $g \in C(E^\Delta)$. In this case, B is bounded and satisfies the positive maximum principle on E^Δ , and $\{e^{tB}\}_{t \geq 0}$ is a strongly continuous contraction semigroup. Moreover, there is some nonnegative (finite) measure $\nu_B(\Delta, \cdot)$ such that (B.1) holds also for $x = \Delta$.

Proof. Assume there is a (nonnegative, finite) kernel ν_B from E to E^Δ such that (B.1) holds for all $x \in E$ and $g \in C(E^\Delta)$. Then clearly $B1 = 0$. Suppose $g \in C(E^\Delta)$, $x \in E$, and $g(x) = \max_E g \geq 0$. Then $g(x) = \max_{E^\Delta} g$, so that $g(\xi) - g(x) \leq 0$ for all $\xi \in E^\Delta$ and hence $Bg(x) \leq 0$. Thus B satisfies the positive maximum principle on E , which proves sufficiency.

To prove necessity, assume $B1 = 0$ and B satisfies the positive maximum principle on E . By Lemmas 4.2.1 and 1.2.11 in Ethier and Kurtz (2005), the restriction $B|_{C_0(E)}$ is dissipative, hence closable, and even closed since it is globally defined on $C_0(E)$. By the closed graph theorem $B|_{C_0(E)}$ is bounded, and then so is B since $B1 = 0$. Pick any $g \in C(E^\Delta)$ with $g(\Delta) = \max_{E^\Delta} g \geq 0$. Then $g - g(\Delta) \leq 0$, so there exist functions $h_n \in C_c(E)$ with $h_n \leq 0$ and $h_n \rightarrow g - g(\Delta)$ uniformly. Then $Bh_n \rightarrow B(g - g(\Delta)) = Bg$ uniformly as well. Taking x_n such that $h_n(x_n) = 0$ and $x_n \rightarrow \Delta$, we obtain $Bg(\Delta) = \lim_{n \rightarrow \infty} Bh_n(x_n) \leq 0$. We have thus proved that B is bounded and satisfies the positive maximum principle on E^Δ . As a result, Lemma 4.2.1 and Theorem 1.7.1 in Ethier and Kurtz (2005) yield that $\{e^{tB}\}_{t \geq 0}$ is a strongly continuous contraction semigroup.

It remains to exhibit a kernel ν_B from E^Δ to E^Δ such that (B.1) holds for all $x \in E^\Delta$ and $g \in C(E^\Delta)$. To this end, fix $x \in E^\Delta$ and define $h \in C(E^\Delta)$ by $h(y) := d(x, y)$, where $d(\cdot, \cdot)$ is a compatible metric for the Polish space E^Δ . Since B satisfies the positive maximum principle on E^Δ , the map

$$C(E^\Delta) \rightarrow \mathbb{R}, \quad g \mapsto B(gh)(x)$$

is a positive linear functional. By the Riesz–Markov representation theorem, there is a measure $\mu(x, \cdot) \in M_+(E^\Delta)$ such that $B(gh)(x) = \int_{E^\Delta} g(\xi)\mu(x, d\xi)$ for all $g \in C(E^\Delta)$. We define

$$\nu_B(x, d\xi) := \mathbf{1}_{E^\Delta \setminus \{x\}}(\xi) \frac{1}{h(\xi)} \mu(x, d\xi),$$

which is permissible since $h(y) > 0$ for all $y \neq x$. For every $g \in C_c(E^\Delta \setminus \{x\})$ we have $g/h \in C(E^\Delta)$, and therefore

$$Bg(x) = B\left(\frac{g}{h}\right)(x) = \int_{E^\Delta} \frac{g(\xi)}{h(\xi)} \mu(x, d\xi) = \int_{E^\Delta} g(\xi) \nu_B(x, d\xi).$$

Since B is bounded, the identity $Bg(x) = \int_{E^\Delta} g(\xi) \nu_B(x, d\xi)$ extends by continuity to all $g \in C(E^\Delta)$ with $g(x) = 0$. Thus, using also that $B1 = 0$,

$$Bg(x) = B(g - g(x))(x) = \int_{E^\Delta} (g(\xi) - g(x)) \nu_B(x, d\xi).$$

Repeating this for every $x \in E^\Delta$ yields that ν_B satisfies (B.1) for all $x \in E^\Delta$ and $g \in C(E^\Delta)$. To see that $\nu_B(x, E^\Delta) < \infty$, just note that $\int_{E^\Delta} g(\xi) \nu_B(x, d\xi) \leq \|B\|$ whenever $g \in C(E^\Delta)$ satisfies $0 \leq g \leq 1$. Measurability of $\nu_B(\cdot, A)$ for every Borel set $A \subseteq E^\Delta$ follows from a monotone class argument, so that ν_B is indeed a kernel from E^Δ to E^Δ . \square

Chapter IV

Probability measure-valued polynomial diffusions and other applications

1 Introduction

Inspired by the study of jump-diffusions taking value on the unit simplex provided in Chapter II, our first goal here is to *systematically* characterize the class of probability-valued polynomial diffusions, i.e. the class of probability-valued continuous jump-diffusions whose generator maps a polynomial of measure argument (in sense of Section III.2.3) to a polynomial of measure argument of equal or lower degree. As explained in the introduction, this leads to the class of probability measure-valued *polynomial* diffusions whose k -th moments can be computed by solving a k -dimensional linear partial (integro) differential equation.

More precisely, to each polynomial operator L we associate some dual operators L_1, L_2, \dots each of them being the generator of a jump-diffusion taking values in E, E^2, \dots , respectively. We then provide a moment formula stating that if $u: \mathbb{R}_+ \times E^k \rightarrow \mathbb{R}$ is a (regular enough) solution of

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x) &= L_k u(t, \cdot)(x), & (t, x) \in \mathbb{R}_+ \times E^k, \\ u(0, x) &= g(x), & x \in E^k, \end{aligned} \tag{1.1}$$

then each jump-diffusion X corresponding to L satisfies

$$\mathbb{E}[\langle g, X_T^k \rangle \mid \mathcal{F}_t] = \langle u(T-t, \cdot), X_t^k \rangle$$

for all $T \geq t$. Moreover, if the moment formula holds for a sufficiently large set of coefficients g , then the law of X is uniquely determined by L and X_0 .

As explained in Chapter II, the class of finite dimensional polynomial jump-diffusions is considerably tractable. However, unfortunately, it is not immune to the curse of dimensionality: on the unit simplex Δ^d , the number of ODEs to solve in order to characterize moments up to degree k is given by $\binom{k+d-1}{k}$, which grows

with d as d^k . The probability measure-valued approach is a tractable alternative for the cases where high dimensionality breaks the tractability of the finite dimensional setting. For illustrating the numerical and computational advantages of that framework, in particular in comparison with finite dimensional polynomial jump-diffusions, let us come back to the linear factor model introduced in (III.1.1). Consider a d -dimensional process S defined via

$$S_t^i := \langle g_i, X_t \rangle, \quad i \in \{1, \dots, d\}, \quad (1.2)$$

where $g_i, i \in \{1, \dots, d\}$ are continuous (or, if needed, even more regular) functions and X is a probability measure-valued diffusion with underlying state space $E = \mathbb{R}$. We here think of truly high dimensional situations with $d \sim 10^3$. Using the moment formula for computing moments of degree k of S , i.e.

$$\mathbb{E}[S_t^k] = \mathbb{E}[\langle g_1, X_t \rangle^{k_1} \cdots \langle g_d, X_t \rangle^{k_d}] = \mathbb{E}[\langle g_1^{\otimes k_1} \otimes g_2^{\otimes k_2} \cdots \otimes g_d^{\otimes k_d}, X_t^k \rangle]$$

amounts

- (i) to solving a k -dimensional PIDE as given in (1.1) with initial condition

$$g := g_1^{\otimes k_1} \otimes g_2^{\otimes k_2} \cdots \otimes g_d^{\otimes k_d},$$

- (ii) and computing a k -fold integral with respect to the initial measure X_0 .

In order to implement these two steps numerically, the most basic approach is to consider a space discretization into n points such that the PIDE becomes an ODE in \mathbb{R}^{n^k} and the integral a sum with n^k summands.

For comparison consider now a classical d -dimensional polynomial diffusion for S . Then computing $\mathbb{E}[S_t^k]$ means

- (i) solving the following linear ODE

$$\partial u_t = Gu,$$

with initial condition corresponding to the nonzero coefficients in the polynomial S_t^k . Here, G is the $N \times N$ matrix representation of the polynomial operator associated with S with $N = \binom{k+d-1}{k} \sim d^k$.

- (ii) computing a scalar product with two vectors of length N (i.e. a sum with N summands).

From this comparison it is now obvious that the competition lies in *space discretization* versus *dimension* of the underlying process. While the first can be chosen according to the accuracy we want to achieve when computing moments, the dimension of the underlying objects that we are interested in cannot be altered. Clearly the second approach gives *exact* results, however at the expense of a curse of dimensionality. At the loss of exact results, this curse of dimensionality can be broken within the first approach when fixing the space discretization n and increasing the dimension d . The key additional structure that is exploited here is the regularity in space, coming from the coefficients g_i .

Finally, linear factor models based on probability measure-valued diffusions are more flexible not only in view of computational aspects, but also when it comes to modeling. Indeed, as projections of an infinite dimensional process they constitute a much richer class than polynomial models on subsets of \mathbb{R}^d .

The second aim of this chapter is to illustrate how the technology presented in Chapter III can be applied in other fields of research in mathematical finance. We will in particular focus on finite and infinite particle systems with mean fields interaction and common noise (continuous and purely discontinuous), on the corresponding (weighted) empirical measure, and on their connection with McKean–Vlasov equations. In the last part we also take a quick look at empirical measures of branching processes, providing some intuition about the asymptotic situation. During this exploration, the effects of the different parameters of a polynomial operator, or of a Lévy type operator, on the corresponding (polynomial) jump-diffusion will be clarified.

Recall that Section III.2 summarizes some notation and provide the mathematical background needed throughout this chapter. The chapter is then organized as follows. Section 2 concerns polynomial operators and their dual operators. In Section 3 we introduce the moment formula, its link with uniqueness of solutions to the martingale problem, and finally we explore some conditions under which existence for each initial condition and well-posedness holds. Section 4 focuses on the analysis of the results of the previous section in the context of finite or real underlying state spaces. Section 5 collects many different examples whose goal is two-fold: explain how the different parameters influence the corresponding probability measure-valued jump-diffusion (Sections 5.1–5.2), illustrate how the proposed machinery is very flexible and can be used in other fields (Sections 5.3–5.4). Some of the proofs are gathered in appendices.

2 Polynomial operators

Let E be a locally compact Polish space. We now define polynomial operators, which constitute a class of possibly unbounded linear operators acting on polynomials. They are not defined on all of P in general, but only on the subspace P^D for some dense subspace $D \subseteq C_\Delta(E)$; see Section III.2.4.

Definition 2.1. Fix $\mathcal{S} \subseteq M(E)$. A linear operator $L: P^D \rightarrow P$ is called \mathcal{S} -polynomial if for every $p \in P^D$ there is some $q \in P$ such that $q|_{\mathcal{S}} = Lp|_{\mathcal{S}}$ and

$$\deg(q) \leq \deg(p).$$

For a finite dimensional diffusion it is known that its generator is polynomial if and only if the drift and diffusion coefficients are polynomial of first and second degree, respectively; see Cuchiero et al. (2012) and Filipović and Larsson (2016). Theorem 2.2 is the generalization of this fact to the probability measure-valued setting. This connection becomes clear from the perspective of Section III.3.3. In particular, recall that Lemma III.3.11 relates path continuity of solutions to

the martingale problem to the carré-du-champ operator being a derivation. The proof is given in Section A.

Theorem 2.2. *Let $L: P^D \rightarrow P$ be a linear operator. Then L is $M_1(E)$ -polynomial and its carré-du-champ operator Γ is an $M_1(E)$ -derivation if and only if*

$$\begin{aligned} Lp(\nu) &= \langle B(\partial p(\nu)), \nu \rangle + \frac{1}{2} \langle Q(\partial^2 p(\nu)), \nu^2 \rangle, \quad \nu \in M_1(E), \\ &\text{for some linear operators } B: D \rightarrow C_\Delta(E) \text{ and} \\ &Q: D \otimes D \rightarrow \widehat{C}_\Delta(E^2). \end{aligned} \tag{2.1}$$

In this case, B and Q are uniquely determined by L .

An analogue of Theorem 2.2 holds for L being \mathcal{S} -polynomial, where \mathcal{S} is an arbitrary subset of $M(E)$; see Theorem A.1.

Remark 2.3. *It is important to note that the operators B and Q described in Theorem 2.2 do not coincide with B and Q given in Section III.4. With the notation of that section, $B(g, \nu)$ is given by $\langle Bg, \nu \rangle$ and $Q(g, \nu)$ is given by $\langle Qg, \nu^2 \rangle$. Since, because of Theorem 2.2, in the framework of polynomial diffusions we are always be interested in operators L satisfying (2.1), we decided to accept this change in order to lighten the notation and emphasize the polynomial structure of the operators.*

As in the previous chapter, we will use the operator Ψ , which maps any function $g: E \times E \rightarrow \mathbb{R}^k$ to the function $\Psi(g): E \times E \rightarrow \mathbb{R}^k$ given by

$$\Psi(g)(x, y) = \frac{1}{2} (g(x, x) + g(y, y) - 2g(x, y)). \tag{2.2}$$

Example 2.4 (The Fleming–Viot process). *Let $E = \mathbb{R}$ and $D = C_\Delta^2(\mathbb{R})$. We already quickly presented the Fleming–Viot process in the discussion before Example III.3.9. This process takes values in $M_1(\mathbb{R})$, and its generator L acts on polynomials $p \in P^D$ by*

$$Lp(\nu) = \int_E B(\partial p(\nu))(x) \nu(dx) + \frac{1}{2} \int_{E^2} \partial_{xy}^2 p(\nu) \nu(dx) (\delta_x(dy) - \nu(dy)), \quad \nu \in M_1(E),$$

where $Bg(x) := \frac{1}{2} \sigma^2 g''(x)$ for some $\sigma > 0$. This is an $M_1(\mathbb{R})$ -polynomial operator of the form (2.1), where $Q = \Psi$. For more details, see Chapter 10.4 of [Ethier and Kurtz \(2005\)](#).

Because of Corollary III.2.6 we know that any polynomial on $M_1(E)$ has a unique homogeneous representative of the same degree. This implies that an operator L satisfying (2.1) maps in fact any monomial $\nu \mapsto \langle g, \nu^k \rangle$ to a unique monomial $\nu \mapsto \langle h, \nu^k \rangle$ on $M_1(E)$, for each $k \in \mathbb{N}$. This induces an operator L_k acting on the corresponding coefficients as $L_k g := h$. The operators L_1, L_2, \dots are the key objects for the computation of conditional moments of jump-diffusions corresponding to L .

Definition 2.5. Let $L: P^D \rightarrow P$ satisfy (2.1). The k -th dual operator of L is defined as the unique linear operator $L_k: D^{\otimes k} \rightarrow \widehat{C}_\Delta(E^k)$ determined by

$$Lp(\nu) = \langle L_k g, \nu^k \rangle, \quad \nu \in M_1(E), \quad (2.3)$$

for every $p(\nu) = \langle g, \nu^k \rangle$ with $g \in D^{\otimes k}$.

Because of (2.1), the k -th dual operator L_k can be written as

$$L_k = kB \otimes \text{id}^{\otimes(k-1)} + \frac{k(k-1)}{2} Q \otimes \text{id}^{\otimes(k-2)}, \quad (2.4)$$

where the tensor notation $B_1 \otimes \dots \otimes B_N$ is used to denote the linear operator from $D^{\otimes k}$ to $\widehat{C}_\Delta(E^k)$ determined by

$$(B_1 \otimes \dots \otimes B_N)(g^{\otimes k}) := B_1(g^{\otimes n_1}) \otimes \dots \otimes B_N(g^{\otimes n_N})$$

for given linear operators $B_i: D^{\otimes n_i} \rightarrow \widehat{C}_\Delta(E^{n_i})$ with $n_1 + \dots + n_N = k$. More explicitly, we have

$$L_k = B_k + Q_k$$

where B_k and Q_k are defined by

$$B_k g := \sum_{i=1}^k B^{(i)} g \quad \text{and} \quad Q_k g := \frac{1}{2} \sum_{i,j=1}^k Q^{(ij)} g \quad (2.5)$$

for $B^{(i)} g(x) := Bg(\dots, x_{i-1}, \cdot, x_{i+1}, \dots)(x_i)$ and

$$Q^{(ij)} g(x) := Q(g(\dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_{j-1}, \cdot, x_{j+1}, \dots))(x_i, x_j).$$

Remark 2.6. Observe that without the existence of a homogeneous representative (guaranteed by Corollary III.2.6), expression (2.3) would read

$$Lp(\nu) = \langle L_k^k g, \nu^k \rangle + \langle L_k^{k-1} g, \nu^{k-1} \rangle + \dots + L_k^0 g, \quad \nu \in M_1(E),$$

and the k -th dual operator would thus consist in a $(k+1)$ -tuple of operators L_k^k, \dots, L_k^0 . In the context of the moment formula later, this would translate in replacing (3.1) with a system of $(k+1)$ PIDE. The implied growth in the amount of needed notation and the consequent lost in clarity is evident.

If one is interested in studying jump-diffusions taking value in other subspaces of $M(E)$, as e.g. $M_+(E)$, a homogeneous representative can no longer be found.

3 Existence and uniqueness of polynomial diffusions on $M_1(E)$

Let E be a locally compact Polish space, D a dense linear subspace of $C_\Delta(E)$ containing the constant function 1, and $L: P^D \rightarrow P$ a linear operator. In this section we study existence and uniqueness of $M_1(E)$ -valued polynomial diffusions, and derive the moment formula.

Definition 3.1. Let L be $M_1(E)$ -polynomial. Any continuous $M_1(E)$ -solution to the martingale problem for L is called a probability measure-valued polynomial diffusion.

3.1 Moment formula and uniqueness in law

Polynomial diffusions are of interest in applications because they generally satisfy a *moment formula*, which allows moments of the process to be computed tractably. If E is a finite set, the moment formula always holds, but technical conditions are needed in general.

Theorem 3.2. *Suppose L satisfies (2.1) and fix $k \in \mathbb{N}$. Assume the k -th dual operator L_k is closable, and let g be in the domain $\mathcal{D}(\bar{L}_k)$ of its closure \bar{L}_k .¹ Suppose that there is a solution $u: \mathbb{R}_+ \times E^k \rightarrow \mathbb{R}$ of*

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x) &= \bar{L}_k u(t, \cdot)(x), & (t, x) \in \mathbb{R}_+ \times E^k, \\ u(0, x) &= g(x), & x \in E^k, \end{aligned} \quad (3.1)$$

and suppose that $\sup_{t \in [0, T]} \|\bar{L}_k u(t, \cdot)\| < \infty$ for all $T \in \mathbb{R}_+$. In particular, $u(t, \cdot)$ is assumed to be in the domain of \bar{L}_k for all $t \geq 0$. Then for any continuous $M_1(E)$ -solution X to the martingale problem for L , one has the moment formula

$$\mathbb{E}[\langle g, X_T^k \rangle \mid \mathcal{F}_t] = \langle u(T-t, \cdot), X_t^k \rangle. \quad (3.2)$$

Proof. We will follow the proof of Theorem 4.4.11 in Ethier and Kurtz (2005) in order to obtain a slightly more general result. Fix $T \in \mathbb{R}_+$, $t \in [0, T]$, $A \in \mathcal{F}_t$, and for all $(s_1, s_2) \in [0, T-t] \times [0, T-t]$ set $f(s_1, s_2) := \mathbb{E}[\langle u(s_1, \cdot), X_{t+s_2}^k \rangle \mathbb{1}_A]$. Fix $s_2 \in [0, T-t]$. (3.1) and the fundamental theorem of calculus then yield

$$\begin{aligned} f(s_1, s_2) - f(0, s_2) &= \mathbb{E}[\langle u(s_1, \cdot) - u(0, \cdot), X_{t+s_2}^k \rangle \mathbb{1}_A] \\ &= \int_0^{s_1} \mathbb{E}[\langle \bar{L}_k u(s, \cdot), X_{t+s_2}^k \rangle \mathbb{1}_A] ds. \end{aligned}$$

Fix then $s_1 \in [0, T-t]$. Since $u(t, \cdot) \in \mathcal{D}(\bar{L}_k)$ for all $t \in \mathbb{R}_+$, (III.3.1) yields

$$\begin{aligned} f(s_1, s_2) - f(s_1, 0) &= \mathbb{E}[\mathbb{E}[\langle u(s_1, \cdot), X_{t+s_2}^k \rangle - \langle u(s_1, \cdot), X_t^k \rangle \mid \mathcal{F}_t] \mathbb{1}_A] \\ &= \int_0^{s_2} \mathbb{E}[\langle \bar{L}_k u(s_1, \cdot), X_{t+s}^k \rangle \mathbb{1}_A] ds. \end{aligned}$$

Since $\sup_{s_1, s_2 \in [0, T-t]} |\mathbb{E}[\langle \bar{L}_k u(s_1, \cdot), X_{t+s_2}^k \rangle \mathbb{1}_A]| \leq \sup_{s_1 \in [0, T-t]} \|\bar{L}_k u(s_1, \cdot)\| < \infty$, we can then conclude that both $f(\cdot, s_2)$ and $f(s_1, \cdot)$ are absolutely continuous with bounded derivative. Lemma 4.4.10 in Ethier and Kurtz (2005) then yields $f(T-t, 0) - f(0, T-t) = 0$ and the result follows. \square

In order to avoid confusion, for the rest of the section we denote by u_g the solution of (3.1) with initial condition $u_g(0, \cdot) = g$.

In most cases of interest (see Remark 3.8(iii) later) the operator L_k satisfies the positive maximum principle on E^k , for each $k \in \mathbb{N}$. If this is the case,

¹ For more information about concepts related to the field of operator semigroups see e.g. Ethier and Kurtz (2005).

the existence of a solution u_g of (3.1) satisfying the conditions of Theorem 3.2 for sufficiently many g , is essentially equivalent for \bar{L}_k to generate a strongly continuous contraction semigroup on $\widehat{C}_\Delta(E^k)$. This is then in turn essentially equivalent for \bar{L}_k to be the generator of a Markov process. We state this more precisely in the next remark.

Remark 3.3. *Let L satisfy (2.1) and X denote an $M_1(E)$ -solution of the corresponding martingale problem with initial condition $\nu \in M_1(E)$. Assume that the corresponding k -th dual operator L_k satisfies the positive maximum principle on $(E^\Delta)^k$ (which in particular implies that L_k is closable), for each $k \in \mathbb{N}$.*

Let $D_0 \subseteq \mathcal{D}(\bar{L}_k)$ be a dense subset, and suppose that the conditions of Theorem 3.2 hold true for all $g \in D_0$. By Proposition 1.3.4 of Ethier and Kurtz (2005), if we additionally have that $t \mapsto L_k u_g(t)$ is continuous, then \bar{L}_k is the generator of a strongly continuous contraction semigroup $\{Y_t^k\}_{t \geq 0}$ on $\widehat{C}_\Delta(E^k)$ and $Y_t^k g = u_g(t, \cdot)$.

Conversely, if \bar{L}_k is the generator of a strongly continuous contraction semigroup $\{Y_t^k\}_{t \geq 0}$ on $\widehat{C}_\Delta(E^k)$, then for all $g \in \mathcal{D}(\bar{L}_k)$ the map $u_g(t, x) := Y_t^k g(x)$ satisfies the conditions of Theorem 3.2. By Hille–Yosida theorem, this is for instance the case if the range of $\lambda - L_k$ is dense in $\widehat{C}_\Delta(E^k)$ for some $\lambda > 0$. In this case, Corollary 4.2.8 in Ethier and Kurtz (2005) yields a solution $Z^{(k)}$ (without loss of generality defined on the same probability space as X) to the martingale problem for L_k with values in $(E^\Delta)^k$ and satisfying $Y_t^k g(x) = \mathbb{E}[g(Z_t^{(k)}) | Z_0^{(k)} = x]$. The moment formula then yields

$$\mathbb{E}[g(Z_t^{(k)}) | Z_0^{(k)} \sim \nu^k] = \mathbb{E}[\langle g, X_t^k \rangle]. \quad (3.3)$$

As in the finite dimensional case, the moment formula yields well-posedness of the martingale problem.

Corollary 3.4. *Suppose L satisfies (2.1), and let X be an $M_1(E)$ -solution to the martingale problem for L with initial condition $\nu \in M_1(E)$. If the moment formula (3.2) holds for all $g \in D^{\otimes k}$ and $k \in \mathbb{N}$, then the law of X is uniquely determined by L and ν .*

Proof. By the moment formula (3.2) we have $\mathbb{E}[\langle g, X_T^k \rangle] = \langle u_g(T, \cdot), \nu^k \rangle$ for all $k \in \mathbb{N}$ and $g \in D^{\otimes k}$. Since $g \mapsto u_g$ is determined by L , Lemma III.2.8(iii) yields that the one dimensional distributions of X are uniquely determined by L and ν . The conclusion follows by Theorem 4.4.2 in Ethier and Kurtz (2005). \square

Remark 3.5. *Suppose that L_k is a polynomial operator in sense of Filipović and Larsson (2016) for all $k \in \mathbb{N}$ and observe that the corresponding semigroup $\{Y_t^k\}_{t \geq 0}$ can then be explicitly computed by means of a matrix exponential. In this particular case, the moment formula given above provides explicit representations of the conditional moments of $M_1(E)$ -valued polynomial diffusion.*

As the dual operators L_1, L_2, \dots play a central role for determining uniqueness of the solution to the martingale problem for L and in particular for the expressing

the moment formula, we expect that the same is true in the framework of finite dimensional polynomial diffusions. In the next example, we illustrate that this is in fact the case. For a more precise description of the connection with the finite dimensional case, see Section 4.1.

Example 3.6 (Dual operators L_k in the finite dimensional setting). *Let $A \subseteq \mathbb{R}^d$ be closed set, fix $k \in \mathbb{N}$, and let $H : A \rightarrow \mathbb{R}^N$ be a basis of $\text{Pol}_k(A)$, where*

$$\text{Pol}_k(A) := \{p : A \rightarrow \mathbb{R} \mid p \text{ is a polynomial on } A \text{ of degree at most } k\}$$

and $N := \dim(\text{Pol}_k(A))$. Consider then a polynomial operator \mathcal{A} on $\text{Pol}(A) := \bigcup_k \text{Pol}_k(A)$ in sense of Filipović and Larsson (2016). How are then the corresponding dual operators L_k given by? In that setting, see e.g. the proof of Theorem II.2.3, one usually defines $G \in \mathbb{R}^{N \times N}$ as the unique matrix such that

$$\mathcal{A}p(x) = H(x)^\top G \vec{p}$$

for all $p \in \text{Pol}_k(A)$, where $\vec{p} \in \mathbb{R}^N$ represents the coefficients vector of p and is thus uniquely determined by $p = \vec{p}^\top H$. In this case $L_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is given by $L_k \vec{p} := G \vec{p}$.

3.2 Existence and well-posedness

The first main result of this section gives abstract sufficient conditions for existence of $M_1(E)$ -solutions to the martingale problem. Applications of this result are discussed in Section 4. Recall that E is throughout a locally compact Polish space.

For B as in (2.1) we say that B is E -conservative if there exist functions $g_n \in D \cap C_0(E)$ such that $\lim_{n \rightarrow \infty} g_n = 1$ and $\lim_{n \rightarrow \infty} (B g_n)^- = 0$ bounded pointwise on E and E^Δ , respectively.

Theorem 3.7. *Let $D \subseteq C_\Delta(E)$ be a dense linear subspace containing the constant function 1. Let $L : P^D \rightarrow P$ be a linear operator satisfying (2.1), where*

- (i) B is E -conservative and satisfies $B1 = 0$,
- (ii) Q is given by

$$Qg = \alpha \Psi(g) + \sum_{i=1}^n (A_i \otimes A_i)(g), \quad g \in D \otimes D,$$

where $\alpha : E^2 \rightarrow \mathbb{R}$ is a nonnegative symmetric function and, for $i = 1, \dots, n$, A_i is the generator of a strongly continuous group of positive isometries of $C_\Delta(E)$, and the domain of A_i contains both D and $A_i(D)$,

- (iii) $B - \frac{1}{2} \sum_{i=1}^n A_i^2$ satisfies the positive maximum principle on E^Δ .

Then L is $M_1(E)$ -polynomial and its martingale problem has an $M_1(E)$ -solution with continuous paths for every initial condition $\nu \in M_1(E)$. If in addition the moment formula (3.2) holds for all $g \in D^{\otimes k}$ and $k \in \mathbb{N}$, then the martingale problem for L is well-posed.

Note that (2.1) imposes the implicit condition on α that $\alpha\Psi(g)$ must lie in $\widehat{C}_\Delta(E^2)$ for every $g \in D \otimes D$. If $D = C_\Delta(E)$, then α is necessarily bounded, as is seen from Theorem 3.10 below. However, this does not hold for general $D \subseteq C_\Delta(E)$. For an example in this sense see Example 4.3.

Proof. Theorem 2.2 shows that L is $M_1(E)$ -polynomial. By Lemma III.2.4(i) we know that $Lp \in C(M_1(E^\Delta))$ for all $p \in P^D$. Moreover, $L1 = 0$ and the E -conservativeness of B implies that condition (III.3.2) holds true. Lemma III.3.6 then yields existence of an $M_1(E)$ -solution to the martingale problem for any initial condition (necessarily with continuous paths due to Lemma III.3.11) once we check that L satisfies the positive maximum principle on $M_1(E^\Delta)$. Let therefore $\nu_* \in M_1(E^\Delta)$ be a maximizer of $p \in P^D$ over $M_1(E^\Delta)$. The optimality conditions in Theorem III.5.1 yield

$$\partial_x p(\nu_*) = \sup_E \partial p(\nu_*) \quad \text{and} \quad \Psi(\partial^2 p(\nu_*))(x, y) \leq 0, \quad x, y \in \text{supp}(\nu_*).$$

Therefore, since $B - \frac{1}{2} \sum_{i=1}^n A_i^2$ satisfies the positive maximum principle and α is nonnegative, we get

$$Lp(\nu_*) \leq \frac{1}{2} \sum_{i=1}^n \left(\langle A_i^2(\partial p(\nu_*)), \nu_* \rangle + \langle (A_i \otimes A_i)(\partial^2 p(\nu_*)), \nu_*^2 \rangle \right).$$

The optimality condition in Theorem III.5.3 now yields $Lp(\nu_*) \leq 0$. This proves the positive maximum principle and thus the existence statement. The assertions regarding the moment formula and well-posedness follow from Theorem 3.2 and Corollary 3.4. \square

Remark 3.8. (i) With regard to item (iii) in Theorem 3.7, note that a linear operator $\mathcal{G} : D \rightarrow C_\Delta(E)$ satisfies the positive maximum principle on E^Δ if and only if \mathcal{G} satisfies the positive maximum principle on E and $\mathcal{G}g(\Delta) \geq 0$ for every nonnegative $g \in C_0(E) \cap D$. In many cases of interest, for instance $E \subseteq \mathbb{R}^d$ and $D \subseteq \mathbb{R} + C_c(E)$ (see Lemma III.3.7(ii)), the positive maximum principle on E implies the positive maximum principle on E^Δ .

(ii) Let us also remark, that the k -th dual operator \mathcal{G}_k associated to $\langle \mathcal{G}(\partial p(\nu)), \nu \rangle$ satisfies the positive maximum principle on $(E^\Delta)^k$ if it holds for \mathcal{G} on E^Δ . Indeed, if $x^* \in (E^\Delta)^k$ is a maximum of g , then x_i^* is a maximum of $g(\dots, x_{i-1}^*, \cdot, x_{i+1}^*, \dots)$. Hence \mathcal{G}_k given by

$$\mathcal{G}_k g = k\mathcal{G} \otimes \text{id}^{\otimes(k-1)} g = \sum_{j=1}^k \mathcal{G}^{(j)} g,$$

where we use the same notation as in (2.5), clearly satisfies the positive maximum principle on $(E^\Delta)^k$.

(iii) Consider the setting and the assumptions of Theorem 3.7 and define

$$G_k := k \left(B - \frac{1}{2} \sum_{i=1}^n A_i^2 \right) \otimes \text{id}^{\otimes(k-1)}, \quad C_k := \frac{k(k-1)}{2} (\alpha \Psi) \otimes \text{id}^{\otimes(k-2)},$$

$$T_k := k \left(\frac{1}{2} \sum_{i=1}^n A_i^2 \right) \otimes \text{id}^{\otimes(k-1)} + \frac{k(k-1)}{2} \left(\sum_{i=1}^n (A_i \otimes A_i) \right) \otimes \text{id}^{\otimes(k-2)}.$$

Note that by (2.4) we have $L_k = G_k + C_k + T_k$. We claim that G_k, C_k, T_k , and hence L_k , satisfy the positive maximum principle on $(E^\Delta)^k$.

By item (iii) in Theorem 3.7, $B - \frac{1}{2} \sum_{i=1}^n A_i^2$ satisfies the positive maximum principle on E^Δ , whence by (ii) it holds also for G_k on $(E^\Delta)^k$. The form of Ψ and the nonnegativity of α guarantee that this is also the case for C_k . Finally, since $T_k = \sum_{i=1}^n \frac{1}{2} (\sum_{j=1}^n A_i^{(j)})^2$ where

$$A_i^{(j)} g(x) = A_i g(\dots, x_{j-1}, \cdot, x_{j+1}, \dots)(x_j),$$

Remark III.5.4 yields the positive maximum principle on $(E^\Delta)^k$ also for T_k and thus all together for L_k .

The following result gives a useful condition for uniqueness when all the operators A_i are zero. Due to Lemma III.5.5 this happens, for instance, if $D = C_\Delta(E)$. An example where uniqueness holds when those operators are not all zero is given in Example 4.4.

Lemma 3.9. Consider setting and assumptions of Theorem 3.7, and assume that $A_i = 0$ for all i . Assume additionally that α is bounded and B is closable and its closure is the generator of a strongly continuous contraction semigroup on $C_\Delta(E)$. Then the moment formula (3.2) holds for all $g \in \widehat{C}_\Delta(E^k)$ and $k \in \mathbb{N}$.

Since B satisfies the positive maximum principle on E^Δ due to Theorem 3.7(iii), the Hille–Yosida theorem guarantees that the conditions of the lemma are satisfied whenever $\lambda - B$ has dense range in $C_\Delta(E)$ for some $\lambda > 0$.

Proof. Let $\{Y_t^1\}_{t \geq 0}$ be the semigroup corresponding to \overline{B} . Fix any $k \in \mathbb{N}$ and let B_k and Q_k be as in (2.5). It is straightforward to check that B_k is the restriction to $D^{\otimes k}$ of the generator of the strongly continuous contraction semigroup $\{(\overline{Y_t^1})^{\otimes k}\}_{t \geq 0}$ on $\widehat{C}_\Delta(E^k)$. Moreover, one has the estimate

$$\|Q_k g\| \leq k(k-1) \|\alpha\| \|g\|, \quad g \in \widehat{C}_\Delta(E^k),$$

whence Q_k is a bounded operator. It follows as in Theorem 1.7.1 and Corollary 1.7.2 in Ethier and Kurtz (2005) that $L_k = B_k + Q_k$ is closable and its closure is the generator of a strongly continuous contraction semigroup on $\widehat{C}_\Delta(E^k)$. By Remark 3.3 and Theorem 3.2 the result follows. \square

As in the finite dimensional case (see Filipović and Larsson (2016)), a polynomial operator generating a diffusion whose domain consists of all polynomials is fully characterized. While Theorem 3.7 only gives sufficient conditions for existence, the result is sharp. More precisely, the next result states that if D can be chosen to be $C_\Delta(E)$, no other polynomial specification exists. Note that this condition is always satisfied if E is a finite set. The proof is given in Section B.

Theorem 3.10. *Let $D = C_\Delta(E)$ and let $L: P^D \rightarrow P$ be a linear operator. Then L is $M_1(E)$ -polynomial, its martingale problem is well-posed, and all $M_1(E)$ -solutions have continuous paths, if and only if L satisfies (2.1) with*

$$Bg = \int (g(\xi) - g(\cdot)) \nu_B(\cdot, d\xi) \quad \text{and} \quad Qg = \alpha\Psi(g), \quad (3.4)$$

where ν_B is a (nonnegative), finite kernel from E to E , and $\alpha: (E^\Delta)^2 \rightarrow \mathbb{R}$ is nonnegative, symmetric, bounded, and continuous on $(E^\Delta)^2 \setminus \{x = y\}$. In this case, the moment formula (3.2) holds for all $g \in \widehat{C}_\Delta(E^k)$ and $k \in \mathbb{N}$. Moreover, B and Q , and hence each L_k , are bounded operators.

As in Theorem 3.7, condition (2.1) imposes implicit conditions on the different parameters. This is the case for the measure ν_B , which in particular needs to satisfy $\int g(\xi) - g(\cdot) \nu_B(\cdot, d\xi) \in C_\Delta(E)$ for all $g \in C_\Delta(E)$. This condition is clearly satisfied if the map from E to $M_+(E)$ given by $x \mapsto \nu_B(x, \cdot)$ is continuous. However the converse fails to be true as one can see by considering the following kernel

$$\nu_B(x, d\xi) = \delta_{\phi(x)} \mathbb{1}_{\{\phi(x) \neq x\}},$$

for some continuous $\phi: E \rightarrow E$ such that $\phi \neq \text{id}$.

Corollary 3.11. *Let $D \subseteq C_\Delta(E)$ be a dense linear subspace containing the constant function 1 and let L satisfy (2.1) with B and Q as in Theorem 3.10. Then L is $M_1(E)$ -polynomial, its martingale problem is well-posed, and all $M_1(E)$ -solutions have continuous paths. Moreover, the moment formula (3.2) holds for all $g \in D^{\otimes k}$ and $k \in \mathbb{N}$.*

Proof. Since by Theorem 3.10 each L_k is bounded, the operator L can be uniquely extended to $P^{C_\Delta(E)}$. The result then follows by the same theorem. \square

The next technical lemma exposes the key properties of a sequence of polynomials. This sequence turns out to be useful in many different places. We already used it for the proofs of Theorem II.4.1 and Theorem II.6.1.

Lemma 3.12. *Define $F_n(z) := \frac{n-1}{n}(1-z)^n + \frac{1}{n}$ for all $z \in [0, 1]$. Then*

$$F_n(z) \in [0, 1], \quad F_n(z)zn \leq 1, \quad \text{and} \quad F_n(z)\sqrt{zn} \leq 1,$$

for all $z \in [0, 1]$.

The last main result of this section characterizes probability measure-valued polynomial martingales, in sense of Definition III.3.3. Note that, unlike Theorem 3.7, the conditions are both necessary and sufficient, regardless of the choice of D .

Theorem 3.13. *Let $D \subseteq C_\Delta(E)$ be a dense linear subspace containing the constant function 1. Let $L: P^D \rightarrow P$ be a linear operator. Then L is $M_1(E)$ -polynomial, its martingale problem has an $M_1(E)$ -solution for any initial condition, and every solution is a martingale with continuous paths, if and only if L satisfies (2.1) with*

$$B = 0 \quad \text{and} \quad Q = \alpha\Psi$$

for some nonnegative symmetric function $\alpha: E^2 \rightarrow \mathbb{R}$. In this case, if in addition α is also bounded, the martingale problem is well-posed.

Proof. To prove the forward implication, first note that Lemma III.3.11 and Theorem 2.2 imply that L satisfies (2.1). To see that $B = 0$, pick any $g \in D$ and $x \in E$, and let X be an $M_1(E)$ -solution to the martingale problem with initial condition δ_x . Since $\langle g, X \rangle$ is a martingale, we have $\langle Bg, X \rangle = 0$ and hence $Bg(x) = \langle Bg, X_0 \rangle = 0$. The form of Q will follow from Lemma C.2. To verify its hypotheses, fix $g \in D$ and $\nu \in M_1(E)$, and define $p \in P^D$ by $p(\mu) := -(\langle g, \nu \rangle - \langle g, \mu \rangle)^2$. Then $\partial^2 p(\nu) = -2g \otimes g$, $p \leq 0$, and $p(\nu) = 0$, so the positive maximum principle yields

$$-\langle Q(g \otimes g), \nu^2 \rangle = Lp(\nu) \leq 0.$$

Next, fix $g \in D$ and $\nu \in M_1(E)$ such that g is constant on the support of ν . Define $p \in P^D$ by $p(\mu) := \langle g, \mu \rangle^2 - \langle g^2, \mu \rangle$. Then, again, $\partial^2 p(\nu) = 2g \otimes g$, and Jensen's inequality yields $p \leq 0$ and $p(\nu) = 0$. Consequently,

$$\langle Q(g \otimes g), \nu^2 \rangle = Lp(\nu) \leq 0.$$

The form of Q thus follows from Lemma C.2.

To prove the reverse implication, observe that existence of $M_1(E)$ -solutions of the martingale problem, along with path continuity, follows from Corollary 3.11, as does well-posedness if in addition α is bounded. Since $B = 0$, it is clear that $\langle g, X \rangle$ is a martingale for every $g \in D$ and every $M_1(E)$ -solution X of the martingale problem. This implies that X is a martingale. \square

4 Applications of probability measure-valued polynomial diffusions

4.1 The unit simplex

Set $E = \{1, \dots, d\}$ and $D \subseteq C(E)$ be a dense linear subspace containing the constant function 1. Note that since $C(E)$ is finite dimensional, the only possible choice for D is given by $C(E)$.

As we already saw many times, the similarity between the structure on the unit simplex Δ^d and that on $M_1(E)$ is very high. Example III.2.2 and Remark III.5.2 already formulate some aspects of this resemblance explicitly. In this section we analyze some other aspects of this correspondence, showing in particular that the

classical moment formula for polynomial jump-diffusions on the unit simplex (see e.g. Filipović and Larsson (2016)) can be deduced from Theorem 3.10.

Let ν_B be a finite kernel from E to E , $\alpha \in \widehat{C}(E^2)$ be nonnegative, and let L satisfy (2.1) with

$$Bg = \int (g(\xi) - g(\cdot)) \nu_B(\cdot, d\xi) \quad \text{and} \quad Qg = \alpha \Psi(g).$$

Recall that by Theorem 3.10 the corresponding martingale problem is well-posed. Consider the linear operator on $\text{Pol}(\Delta^d)$ given by

$$\begin{aligned} \mathcal{A}g(z) = & \sum_{i,j=1}^d \nu_B(i, \{j\}) z_i \left(\frac{d}{dz_j} g(z) - \frac{d}{dz_i} g(z) \right) \\ & + \frac{1}{2} \sum_{i,j=1}^d \alpha(i, j) z_i z_j \left(\frac{d^2}{dz_i^2} g(z) + \frac{d^2}{dz_j^2} g(z) - 2 \frac{d^2}{dz_i dz_j} g(z) \right). \end{aligned} \quad (4.1)$$

Observe that it can be alternatively be written as $\mathcal{A}g = b^\top \nabla g + \frac{1}{2} \text{Tr}(a \nabla^2 g)$ where $b_k(z) := \sum_{i=1}^d (\nu_B(i, \{k\}) z_i - \nu_B(k, \{i\}) z_k)$ and

$$a_{kk}(z) := \sum_{\ell \neq k} \frac{1}{2} \alpha(k, \ell) z_k z_\ell, \quad a_{k\ell}(z) := -\frac{1}{2} \alpha(k, \ell) z_k z_\ell, \quad \text{for all } k \neq \ell.$$

Since this operator is of Type 0 (see Section II.6 for more details), the corresponding martingale problem is well-posed. We can thus define $Z := (Z^1, \dots, Z^d)$ as the unique Δ^d -valued solution to the martingale problem for \mathcal{A} with initial value $z^0 \in \Delta^d$. We claim that

$$X_t := \sum_{i=1}^d Z_t^i \delta_i, \quad \text{for all } t \geq 0,$$

is the unique $M_1(E)$ -solution to the martingale problem for L with initial condition $z_1^0 \delta_1 + \dots + z_d^0 \delta_d$.

Set $p(\nu) := \langle g, \nu \rangle^k$ and $p_d(z) := p(z_1 \delta_1 + \dots + z_d \delta_d)$ for $z \in \Delta^d$. For N^p as in (III.3.1) we can then compute

$$\begin{aligned} N_t^p = & \left(\sum_{i=1}^d Z_t^i g(i) \right)^k - \left(\sum_{i=1}^d z_i^0 g(i) \right)^k \\ & - k \int_0^t \sum_{i=1}^d \left(\sum_{j=1}^d (g(j) - g(i)) \left(\sum_{\ell=1}^d Z_s^\ell g(\ell) \right)^{k-1} \nu_B(i, \{j\}) \right) Z_s^i ds \\ & - \frac{k(k-1)}{2} \int_0^t \sum_{i,j=1}^d \alpha(i, j) \left(g(i)^2 + g(j)^2 - 2g(i)g(j) \right) \\ & \quad \left(\sum_{\ell=1}^d Z_s^\ell g(\ell) \right)^{k-2} Z_s^i Z_s^j ds \\ = & p_d(Z_t) - p_d(z^0) - \int_0^t \mathcal{A}p_d(Z_s) ds. \end{aligned}$$

Using that Z is a solution to the martingale problem for \mathcal{A} we can thus conclude that N^p is a martingale and the claim follows.

4.1.1 The classical moment formula

Fix $k \in \mathbb{N}$. Let then $\overline{H} : E^k \rightarrow \mathbb{R}^N$ be a basis of $C(E)^{\otimes k} = \widehat{C}(E^k)$, which is now a finite dimensional vector space, and let $H : \Delta^d \rightarrow \mathbb{R}^N$ be the basis of $\text{Pol}_k(\Delta^d)$ given by

$$H(z) = \langle \overline{H}, (z_1\delta_1 + \dots + z_d\delta_d)^k \rangle \quad \text{for } z \in \Delta^d,$$

where $N = \dim(\widehat{C}(E^k)) = \dim(\text{Pol}_k(\Delta^d))$. From this prospective, the dual operator L_k is a linear transformation of \mathbb{R}^N and can thus be seen as an element of $\mathbb{R}^{N \times N}$. The corresponding semigroup (in sense of Remark 3.3) is given by $Y_t^k g = \overline{H}^\top e^{tL_k} \vec{g}$, where $\vec{g} \in \mathbb{R}^N$ denotes the coefficients vector of $g \in \widehat{C}(E^k)$ and is thus uniquely determined by $g = \overline{H}^\top \vec{g}$. By Theorem 3.10 the moment formula holds and we can then conclude that for any $p \in \text{Pol}_k(\Delta^d)$ and $\vec{p} \in \mathbb{R}^N$ such that $p = H^\top \vec{p}$ we have

$$\begin{aligned} \mathbb{E}[p(Z_{t+s}) | \mathcal{F}_s] &= \mathbb{E}[H(Z_{t+s})^\top \vec{p} | \mathcal{F}_s] = \mathbb{E}[\langle \overline{H}^\top \vec{p}, X_{t+s}^k \rangle | \mathcal{F}_s] = \langle Y_t^k(\overline{H}^\top \vec{p}), X_s^k \rangle \\ &= \langle \overline{H}^\top e^{tL_k} \vec{p}, X_s^k \rangle = H(Z_s)^\top e^{tL_k} \vec{p}, \end{aligned}$$

which coincides with the classical moment formula (II.2.2) for polynomial diffusions on the unit simplex Δ^d .

4.1.2 Approximation by diffusions on unit simplexes of increasing dimension

Let E be a locally compact Polish space and $D \subseteq C_\Delta(E)$ be a dense linear subspace containing the constant function 1. Suppose that $L : P^D \rightarrow P$ satisfies (2.1) for B and Q as in Theorem 3.10. Our goal is to approximate L by a sequence of polynomial operators generating diffusions on the unit simplexes of increasing dimension.

Fix a partition $(A_k^d)_{k=1}^d$ of E , some points $y_k^d \in A_k^d$ and let $E_d := \{y_1^d, \dots, y_d^d\}$. The idea is to approximate polynomials on $M_1(E)$ with polynomials on $M_1(E_d)$ and then, following the discussion at the beginning of Section 4.1, to identify those polynomials with polynomials on the unit simplex. Consider the operator $L^d : P^D \rightarrow P$ (where now P^D and P are subspaces of the space of polynomials on $M(E_d)$) satisfying (2.1) with

$$Bg = \int (g(\xi) - g(\cdot)) \nu_B^d(\cdot, d\xi) \quad \text{and} \quad Qg = \alpha^d \Psi(g),$$

for $\nu_B^d(y_i^d, \{y_j^d\}) := \nu_B(y_i^d, A_k^d)$ and $\alpha^d := \alpha|_{E_d \times E_d}$. Note that L^d corresponds to

the following generator of a polynomial diffusion on the unit simplex

$$\begin{aligned} \mathcal{A}^d g(z) &= \sum_{i,j=1}^d \nu_B(y_i^d, \{A_j^d\}) z_i \left(\frac{d}{dz_j} g(z) - \frac{d}{dz_i} g(z) \right) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \alpha(y_i^d, y_j^d) z_i z_j \left(\frac{d^2}{dz_i^2} g(z) + \frac{d^2}{dz_j^2} g(z) - 2 \frac{d^2}{dz_i dz_j} g(z) \right). \end{aligned} \quad (4.2)$$

As explained before, the corresponding martingale problem is well-posed. Moreover, it is now easy to see that $(\mathcal{A}^d)_{d \in \mathbb{N}}$ approximates L in the following sense

$$\begin{aligned} Lp(\nu) &= \lim_{d \rightarrow \infty} L^d(p|_{M_1(E^d)}) (\nu(A_1^d) \delta_{y_1^d} + \cdots + \nu(A_d^d) \delta_{y_d^d}) \\ &= \lim_{d \rightarrow \infty} \mathcal{A}^d(p_d) (\nu(A_1^d), \dots, \nu(A_d^d)), \end{aligned}$$

where p_d is again the polynomial on Δ^d corresponding to $p|_{M_1(E^d)}$, i.e. $p_d(z) := p(z_1 \delta_{y_1^d} + \cdots + z_d \delta_{y_d^d})$ for $z \in \Delta^d$.

Lemma 1.6.1 of [Ethier and Kurtz \(2005\)](#) then allows us to deduce that any $M_1(E)$ -solution to the martingale problem for L can be approximated by polynomial diffusions on unit simplexes of increasing dimension. More precisely, setting $y_\nu^d := (\nu(A_1^d), \dots, \nu(A_d^d))$ and letting (Z_1^d, \dots, Z_d^d) denote the unique solution to the martingale problem for \mathcal{A}^d we get

$$\mathbb{E}_{y_\nu^d} \left[p \left(\sum_{k=1}^d Z_{k,t}^d \delta_{y_k^d} \right) \right] \xrightarrow{d \rightarrow \infty} \mathbb{E}_\nu [p(X_t)], \quad p \in P^D.$$

4.2 The underlying space $E \subseteq \mathbb{R}^d$

The goal of this section is to analyze [Theorem 3.7](#) for $E \subseteq \mathbb{R}^d$ being a closed subset and $D \subseteq C_\Delta^2(E)$ dense linear subspace containing the constant function 1. Recall that Ψ is given by [\(2.2\)](#) and $\Sigma^d(E)$ denotes the set of suitable diffusion matrices for a process on E , precisely defined by [\(III.5.8\)](#).

Theorem 4.1. *Let $L: P^D \rightarrow P$ be a linear operator satisfying [\(2.1\)](#), where*

- (i) B is E -conservative and $B1 = 0$,
- (ii) Q is given by

$$Q(g \otimes g) = \alpha \Psi(g \otimes g) + \text{Tr} \left((\tau^\top \nabla g) \otimes (\tau^\top \nabla g)^\top \right) \quad g \in D,$$

for $\alpha: E^2 \rightarrow \mathbb{R}$ being nonnegative and symmetric and $\tau \in \Sigma^d(E)$

- (iii) $B - \sum_{j=1}^d (\tau_j^\top \nabla)^2$ satisfies the positive maximum principle on E^Δ .

Then conditions [\(i\)](#)-[\(iii\)](#) of [Theorem 3.7](#) hold true.

Proof. The result follows from [Lemma III.5.6](#). □

As explained after Corollary III.5.7, the operators defined in the theorem are well-defined, which means that $(\tau_j^\top \nabla)^2 g$ and $\text{Tr}((\tau^\top \nabla g) \otimes (\tau^\top \nabla g)^\top)$ only depends on g through its values on E . This consideration permits us to use the expressions given in (ii) and (iii) without specifying with respect to which representative of τ_j and g the appearing derivatives are computed.

The rest of the section is dedicated to the case $d = 1$ and $E = \mathbb{R}$. For simplicity, we also assume that $D \subseteq \mathbb{R} + C_c^2(\mathbb{R})$. Because of the third condition of Theorem 4.1 and Remark III.5.4 we already know to be interested in drift operators $B : D \rightarrow C_\Delta(\mathbb{R})$ satisfying the positive maximum principle on \mathbb{R}^Δ . As in Lemma III.3.7(ii), one can show that the choice D guarantees that this condition is implied by the positive maximum principle on \mathbb{R} . It is well-known, see e.g. Courrège (1965) or Hoh (1998), that under this condition B is a Lévy type operator, i.e.

$$Bg = bg' + \frac{1}{2}ag'' + \int g(\cdot + \xi) - g - \chi(\xi)g' F(\cdot, d\xi), \quad g \in D \quad (4.3)$$

for some continuous functions $b : E \rightarrow \mathbb{R}$, $a : E \rightarrow \mathbb{R}_+$, a truncation function χ , and a jump kernel $F(\cdot, d\xi)$ from E to E such that $\|\int |\xi|^2 \wedge 1 F(\cdot, d\xi)\| < \infty$. Observe that every operator of this form satisfies $B1 = 0$ and the positive maximum principle on \mathbb{R} .

The first result is a reformulation of Theorem 4.1 in this setting.

Corollary 4.2. *Let $L : P^D \rightarrow P$ be a linear operator satisfying (2.1), where B is as in (4.3) for $a := \sigma^2 + \tau^2$ and*

$$Q(g \otimes g)(x, y) = \frac{1}{2}\alpha(x, y)(g(x) - g(y))^2 + \tau(x)\tau(y)g'(x)g'(y), \quad g \in D,$$

where $\alpha \in \widehat{C}_\Delta(\mathbb{R}^2)$ is nonnegative, $\sigma \in C(\mathbb{R})$, and $\tau \in C_\Delta^1(\mathbb{R})$. Assume also that B is \mathbb{R} -conservative. Then conditions (i)-(iii) of Theorem 3.7 hold true.

In the next example we illustrate a possible interplay between α and τ .

Example 4.3. *Noting that for all $g \in D \otimes D$*

$$\lim_{y \rightarrow x} \frac{(g(x) - g(y))^2}{|x - y|^2} = g'(x)^2 \quad \text{for all } x \in \mathbb{R}$$

one can see that an interplay between τ and α leads to an extension of the class of specifications listed above, in the sense that α is not necessarily bounded. Indeed, fix $\alpha \in C((\mathbb{R}^\Delta)^2 \setminus \{x = y\})$ and $\tau_\alpha \in C_\Delta^1(\mathbb{R})$ such that

$$\alpha(x, x + y) = \tau_\alpha(x)^2|y|^{-2} + o(|y|^{-2}) \quad (4.4)$$

and observe that setting

$$Q(g \otimes g)(x, y) = \alpha(x, y)(g(x) - g(y))^2 + \tau_\alpha(x)\tau_\alpha(y)g'(x)g'(y)\mathbb{1}_{\{x=y\}}$$

the condition $Q(g \otimes g) \in \widehat{C}_\Delta(\mathbb{R}^2)$ is guaranteed for all $g \in D$. Setting then again B is as in (4.3) for $a := \sigma^2 + \tau_\alpha^2$ and assuming that it is \mathbb{R} -conservative, one can prove that the results of Corollary 4.2 still hold true. From a technical point of view, the only change in the proof is due to the application of Theorem III.5.3 for $\mu = \nu_*(y)\delta_y$ for every atom $y \in E^\Delta$ of ν_* .

Finally, it is interesting to note that those observations allow for a relaxation the conditions on α even in the case where $\tau = 0$, i.e. $Q = \alpha\Psi$. Indeed, whenever

$$\alpha \in C((\mathbb{R}^\Delta)^2 \setminus \{x = y\}) \quad \text{and} \quad \alpha(x, y) = o(|x - y|^{-2}),$$

condition (4.4) is satisfied for $\tau_\alpha = 0$.

In the following example L_k , G_k , C_k , and T_k are as in Remark 3.8(iii).

Example 4.4 (Well-posedness for $\tau \neq 0$). Consider the setting of Corollary 4.2 for $D = \mathbb{R} + C_c^\infty(\mathbb{R})$. Let then B be as in (4.3) for $F = 0$, $a := \sigma^2 + \tau^2$ for $\tau \in C_\Delta^1(\mathbb{R})$ and σ^2 being bounded away from 0. Assume that the parameters b and σ^2 are Lipschitz continuous and bounded. Then, by Theorem 8.1.6 of Ethier and Kurtz (2005), B is \mathbb{R} -conservative and the closure of $G_k + T_k$ generates a strongly continuous semigroup on $\widehat{C}_\Delta(E^k)$ for each $k \in \mathbb{N}$. Since C_k is bounded, \bar{L}_k generates a strongly continuous contraction semigroup on $\widehat{C}_\Delta(E^k)$ as well (for more details see e.g. Theorem 1.7.1 in Ethier and Kurtz (2005)). Since by Remark 3.8(iii) L_k satisfies the positive maximum principle, Remark 3.3 and Theorem 3.2 then yield the moment formula for all $g \in D^{\otimes k}$ and well-posedness follows by Theorem 3.7.

Suppose now that $\mathbb{R} + C_c^\infty(\mathbb{R}) \subseteq D$. The next result can easily be extended to general closed statespace $E \subseteq \mathbb{R}^d$.

Lemma 4.5. Let $L: P^D \rightarrow P$ be a linear operator satisfying (2.1) for B is as in (4.3). Suppose that L satisfies the positive maximum principle on \mathbb{R} . Then, for all $\lambda \in [0, 1]$, $g \in D$, and $x, y \in \mathbb{R}$ such that $g(x) = g(y)$ we have that

$$\langle Q(g \otimes g), \nu_\lambda^2 \rangle \leq \langle (ag')^2, \nu_\lambda \rangle, \quad \nu_\lambda = \lambda\delta_x + (1 - \lambda)\delta_y. \quad (4.5)$$

On the one hand, this lemma illustrates that the form of Q given in Corollary 4.2 is very general. Indeed, if in that setting we set $\sigma = 0$ (and thus $a = \tau$) and $\alpha = 0$ we get

$$Q(g \otimes g)(x, y) = \tau(x)\tau(y)g'(x)g'(y)$$

showing that condition (4.5) is tight for $\lambda \in \{0, 1\}$. To the other hand, it provides a useful tool in terms of interpretation, that we will present in the next corollary.

Proof. Fix $g \in D$ such that $g(x) = g(y)$. Since, by Lemma C.1, $B1 = 0$ and $Q(g \otimes 1) = 0$ it is enough to consider the case $g(x) = g(y) = 1$. The result will follow from Lemma C.3. Indeed, if we let $(p_n)_{n \in \mathbb{N}}$ and $(f_n)_{n \in \mathbb{N}}$ be the sequences described there, by the positive maximum principle of L on \mathbb{R} we get

$$0 \geq Lp_n(\nu_\lambda) = \langle Bf_n, \nu_\lambda \rangle + \frac{1}{2} \langle Q(g \otimes g), \nu_\lambda^2 \rangle$$

and letting n go to ∞ we can conclude the proof.

To verify the hypotheses of Lemma C.3, observe that Lemma C.1 yields

$$\langle Q(g \otimes g), \nu_\lambda^2 \rangle \geq 0 \quad \text{for all } \lambda \in [0, 1].$$

Fix some $g \in D$ and $x, y \in \mathbb{R}$ such that $g(z) = g'(z) = 0$ for $z \in \{x, y\}$, and suppose that $\|g\| = 1$. Let $F_n : [0, 1] \rightarrow \mathbb{R}$ be the function defined in Lemma 3.12. Consider then the sequence of polynomials given by

$$p_n(\nu) = \langle g, \nu \rangle^2 F_n(\langle H, \nu \rangle) - \frac{1}{n} \langle H, \nu \rangle,$$

where, for some compactly supported function $\rho \in C_\Delta^\infty(\mathbb{R})$ such that $\rho = 1$ on some neighborhood of x and y and $\rho(\mathbb{R}) \subseteq [0, 1]$,

$$H(z) = C|z - x|^2|z - y|^2\rho(z) + (1 - \rho(z)).$$

Observe that the conditions on g guarantee that for C big enough $|g| \leq H$ and thus $|\langle g, \nu \rangle| \leq \langle H, \nu \rangle$ for all $\nu \in M_1(\mathbb{R})$. For $\text{supp}(\rho)$ small enough we also have that $\|H\| \leq 1$. Lemma 3.12 then yields $\langle g, \nu \rangle^2 F_n(\langle H, \nu \rangle) \leq \frac{1}{n} \langle H, \nu \rangle$ for all $\nu \in M_1(\mathbb{R})$, and therefore $p_n \leq 0$ on $M_1(\mathbb{R})$. This automatically implies that p_n has a maximum at ν_λ for all $\lambda \in [0, 1]$. Proceeding as in the proof of Theorem 3.10 we then obtain that $\langle Q(g \otimes g), \nu_\lambda^2 \rangle = 0$ for any $g \in D$ such that $g(x) = g(y) = 1$ and $g'(x) = g'(y) = 0$. Choosing $\lambda = 0, 1, 1/2$ we get the result. \square

Corollary 4.6. *Let B admit representation (4.3). If $a = 0$, i.e. if B does not have a diffusion coefficient, then $Q = \alpha\Psi$ for some nonnegative $\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that $\alpha\Psi g \in \widehat{C}_\Delta(\mathbb{R}^2)$ for all $g \in D$.*

Proof. Lemma C.1 and condition (4.5) for $\lambda = 0, 1, 1/2$ yield the conditions of Lemma C.2(i) and the result follows. \square

This corollary illustrates that if the process generated by B is not diffusive, then the diffusion operator Q cannot include a term involving derivatives of its arguments. By Section 5.1.4, this in particular implies that the corresponding measure valued diffusion cannot present a spatial-type diffusive behavior.

5 Examples

The goal of this section is two-fold. On the one hand, we explain which effects the parameters used for describing both polynomial operators and Lévy type operators have on the $M_1(E)$ -solutions to the corresponding martingale problem. To the other hand, we illustrate how the approach exposed in this thesis is very flexible, and its technology can be used in many different contexts. It is particularly interesting to observe that empirical measures of particle systems fall very naturally within the framework described in this thesis. This constitutes an important motivation for our results and, as for the results about the Lévy type representation of operators presented in Section III.4, it strongly supports the choice to work with the notion of derivative introduced in Section III.2.2.

5.1 The role of the parameters of a polynomial operators

For the whole section, we let E be a locally compact Polish space, D be a dense linear subspace containing the constant function 1, and $L : P^D \rightarrow P$ be a linear operator satisfying the conditions of Theorem 3.7.

5.1.1 The role of B

Suppose that for each $k \in \mathbb{N}$ the closure of the dual operator \bar{L}_k is the generator of a strongly continuous contraction semigroup $\{Y_t^k\}_{t \geq 0}$ on $\widehat{C}_\Delta(E^k)$. Then, by Remark 3.3 and Corollary 3.4 the martingale problem for L is well-posed and we can denote by X be the corresponding unique $M_1(E)$ -solution with initial condition $\mu \in M_1(E)$ defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Following Remark 3.3, denote by $Z^{(1)}$ the solution to the martingale problem for $L_1 = B$ with initial condition $Z_0^{(1)} \sim \mu$. Without loss of generalities assume that it is defined on the same filtered probability space. Note that E -conservativity of B guarantees that $\mathbb{P}(Z_t^{(1)} \in E) = 1$. The moment formula then yields

$$\mathbb{E}[\langle g, X_t \rangle] = \langle Y_t^1 g, \mu \rangle = \mathbb{E}[g(Z_t^{(1)}) | Z_0^{(1)} \sim \mu],$$

showing that the “expected solution” $\mathbb{E}[X_t(\cdot)]$ coincides with the distribution $\mathbb{P}(Z_t^{(1)} \in \cdot | Z_0^{(1)} \sim \mu)$ of the spatial motion $Z_t^{(1)}$, for each $t \geq 0$. From this we also infer that

$$X_t^{(1)} := \mathbb{P}(Z_t^{(1)} \in \cdot | Z_0^{(1)} \sim \mu), \quad t \geq 0,$$

is the unique $M_1(E)$ -solution to the martingale problem for $L : P^D \rightarrow P$ given by $Lp(\nu) = \langle B(\partial p(\nu)), \nu \rangle$. Observe that, as any pure drift process, $X^{(1)}$ is deterministic.

5.1.2 The role of the jumps term in B

As we saw in Section 5.1.1, the operator B can often be seen as the generator of a jump-diffusion on E . This in particular does not exclude the possibility that the corresponding E -valued jump-diffusion presents some discontinuities. At a first glance this observation can be surprising, in particular since by Theorem 2.2 we know that any $M_1(E)$ -solution to the martingale problem for L has continuous paths. The next example illustrates the role of the jumps in the drift term B , and how they do not compromise the continuity of the corresponding $M_1(E)$ -solutions. In order to enrich the example, we do not assume that the coefficient α is trivial.

Set $E := \{0\} \cup A$ where $A = [\frac{1}{2}, 1]$ and suppose that $A_i = 0$ for each i , $Bg(x) = g(0) - g(x)$, and fix $\alpha \in \widehat{C}_\Delta(E^2)$ such that $\alpha(\cdot, 0) = \alpha(0, \cdot) = 0$. Observe that B is the generator of a jump-diffusion which is constant for an $\exp(1)$ -distributed random time and then performs a jump to 0.

Since the martingale problem for L is well-posed by Corollary 3.11, we can denote by X the corresponding unique $M_1(E)$ -solution with initial condition ν

for some $\nu \in M_1(E)$ such that $\text{supp}(\nu) \subseteq A$. Choosing $g \in C_\Delta(E)$ such that $g|_A = 1$ and $g(0) = 0$, we can then compute

$$\begin{aligned} \mathbb{E}[X_t(A)^k] &= \mathbb{E}[\langle g, X_t \rangle^k] = \mathbb{E}[\langle g, \nu \rangle^k] + \int_0^t \mathbb{E}[\langle -kg, X_s \rangle \langle g, X_s \rangle^{k-1}] ds \\ &= 1 + \int_0^t -k \mathbb{E}[X_s(A)^k] ds. \end{aligned}$$

As a result $\mathbb{E}[X_t(A)^k] = e^{-kt}$ for all $k \in \mathbb{N}$ and thus

$$X_t(\{0\}) = 1 - X_t(A) = 1 - e^{-t} \quad \mathbb{P}\text{-a.s.}$$

This example consists thus in a probability measure-valued polynomial diffusion whose mass is progressively (more precisely, weakly continuously) moved from the set A , where a diffusion is taking place, to the isolated point $\{0\}$.

5.1.3 The role of α

Fix now $B = 0$, $A_i = 0$ for each i , and $\alpha \in \widehat{C}_\Delta(E^2)$. Moreover, for some fixed points $z_1, \dots, z_d \in E$ let $Y := (Y^1, \dots, Y^d)$ be the unique Δ^d -valued solution to the martingale problem for

$$Ag(z) = \frac{1}{2} \sum_{i,j=1}^d \alpha(z_i, z_j) \left(\frac{d^2}{dy_i^2} g(y) + \frac{d^2}{dy_j^2} g(y) - 2 \frac{d^2}{dy_i dy_j} g(y) \right), \quad (5.1)$$

and initial value $y^0 \in \Delta^d$. As explained in Section 4.1, the unique $M_1(E)$ -solution X of the martingale problem for L with initial condition $y_1^0 \delta_{z_1} + \dots + y_d^0 \delta_{z_d}$ is given by $X_t := \sum_{i=1}^d Y_t^i \delta_{z_i}$.

For instance, for $d = 2$, X is a probability measure-valued martingale, whose mass oscillates between two atoms following a Jacobi diffusion until one of the two reaches mass 0. Also, more generally, the value $\alpha(x, y)$ can be interpreted as kind of volatility coefficient driving the mass exchange between point x and y .

5.1.4 The role of τ

What we report here is the most prominent example of a jump-diffusion whose generator involves a τ -term. This in particular ensures the importance of including the A_i 's terms in the description of the linear operator L considered in Theorem 3.7.

Let $E = \mathbb{R}$, $D = C_\Delta^2(\mathbb{R})$, and W denote a brownian motion on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$. In the setting of Corollary 4.2 set $\alpha = 0$, $Bg = \frac{1}{2}g''$, and $\tau = 1$. The resulting operator Q is then given by $Q(g \otimes g) = g' \otimes g'$, respectively. Then

$$X := \delta_W$$

is an $M_1(\mathbb{R})$ -solution to the martingale for L with initial condition δ_0 . One can prove it by checking condition (III.3.1) with a simple application of the Itô formula to $\langle g, X_t \rangle^k = g(W_t)^k$ for all $g \in D$ and $k \in \mathbb{N}_0$.

As explained in Section III.5.2, the optimality condition concerning τ is obtained by slightly changing an optimizer ν^* via shifting its support. This example enforces the intuition that τ has an impact on the dynamics of support and not on weights, as was the case for α . Indeed, one can easily see that X_t coincide with the initial condition up to a support's shift. This concept will be explained with more examples in Section 5.2.3.

Finally, it is interesting to observe that by Section 5.1.1 we know that setting $\tau = 0$ (and thus $Q = 0$) the corresponding $M_1(E)$ -solution to the martingale problem would be

$$X_t^{(1)} := \mathbb{P}(W_t^{(1)} \in \cdot) = \mathcal{N}(0, t),$$

where $\mathcal{N}(0, t)$ denotes the centered normal distribution with variance t . This in particular implies that $\text{supp}(X_t^{(1)}) = \mathbb{R}$ for all $t > 0$. Comparing $X^{(1)}$ with X we can conclude that the operator Q given by $Q(g \otimes g) = g' \otimes g'$ provides the exact amount of diffusion needed to maintain the support of X_t concentrate in one single point \mathbb{P} -a.s. for all $t \geq 0$.

5.1.5 When α and τ appear together

The goal of this example is to explicitly construct an $M_1(E)$ -solution to the martingale problem for an operator L as in Corollary 4.2, where both the coefficients α and τ are assumed to be nonzero. It will have the form of a weighted empirical measure, where the weights follow a diffusion on the unit simplex, and the particles a diffusion on \mathbb{R} with common noise.

Let $E = \mathbb{R}$ and $D \subseteq C_\Delta^2(\mathbb{R})$. Then set B is as in (4.3) for $b \in C_c^1(\mathbb{R})$, $a := \tau^2$ for $\tau \in C_c^1(\mathbb{R})$, and $F = 0$, and

$$Q(g \otimes g)(x, y) = \alpha(x, y)\Psi(g \otimes g)(x, y) + \tau(x)\tau(y)g'(x)g'(y), \quad g \in D,$$

where $\alpha \in \widehat{C}_\Delta(\mathbb{R}^2)$ is nonnegative. The conditions on b and τ already guarantee that B is \mathbb{R} -conservative. Let then $(Z^i)_{i \in \mathbb{N}}$ be a solution of the SDE

$$dZ_t^i = b(Z_t^i)dt + \tau(Z_t^i)dW_t^0$$

with initial value $Z_0^i = z_0^i \in \mathbb{R}$, where W^0 is a brownian motion. Observe that the generator of each Z^i is given by B . Let then Y^d be the polynomial diffusion on the unit simplex given by

$$dY_t^d = a^d(Y_t^d, Z_t^1, \dots, Z_t^d)^{1/2}dW_t$$

with initial value $Y_0^d = y_0^d \in \Delta^d$, where W denotes a d dimensional brownian motion independent of W^0 and

$$a_{kk}^d(y, z_1, \dots, z_d) := \sum_{\ell \neq k} \frac{1}{2} \alpha(z_k, z_\ell) y_k y_\ell, \quad a_{k\ell}^d(y, z_1, \dots, z_d) := -\frac{1}{2} \alpha(z_k, z_\ell) y_k y_\ell.$$

The choice of a^d becomes clear if one compares it with the construction provided in the initial discussion of Section 4.1.

Lemma 5.1. *The process $X^d := \sum_{i=1}^d Y_i^d \delta_{z_i}$ is an $M_1(\mathbb{R})$ -solution to the martingale problem for L with initial condition $\sum_{i=1}^d y_0^d \delta_{z_i}$.*

Proof. Setting $p(\nu) = \langle g, \nu \rangle^k$, by Itô formula we can compute

$$\begin{aligned} p(X_t^d) - p(X_0^d) &= (\text{martingale}) + \int_0^t k \langle g, X_s^d \rangle^{k-1} \langle Bg, X_s^d \rangle \\ &\quad + \frac{k(k-1)}{2} \langle g, X_s^d \rangle^{k-2} \left(\langle \tau g', X_s^d \rangle^2 + \langle \alpha \Psi(g \otimes g), (X_s^d)^2 \rangle \right) ds \\ &= (\text{martingale}) + \int_0^t \langle B(\partial p(X_s^d)), X_s^d \rangle + \frac{1}{2} \langle Q(\partial^2 p(X_s^d)), (X_s^d)^2 \rangle ds, \end{aligned}$$

and the result follows. \square

5.2 The role of the parameters Q and N of Lévy type operators

We now turn to the setting of Lévy type operators, in order to understand the role of the parameters appearing there. The role of B and of Γ have already been explained in Section 5.1.1 and in Lemma III.4.5(i), respectively. Lemma III.4.6 gives an intuition about the interpretation of Q .

In the first section, we focus our attention to the parameter Q , which without the assumption for L to be polynomial can have a much simpler form. We then use the intuition behind Δ^d -valued polynomial jump-diffusions of Type 1 to construct an $M_1(E)$ -valued polynomial jump-diffusion. Finally, we investigate a possible interaction between Q and N , and B , in the spirit of Section 5.1.4.

5.2.1 Probability measure-valued martingale

The example presented in this section has been suggested to us by Sigrid Källblad.

Fix $h \in C_c(\mathbb{R})$ and let $L : P_c^\infty(\mathbb{R}) \rightarrow C(M_{\leq 1}^v(\mathbb{R}))$ be a Lévy type operator for $\Gamma = 0$, $B = 0$, $N = 0$, and

$$Q(g \otimes g, \mu) = \left(\langle gh, \mu \rangle - \langle g, \mu \rangle \langle h, \mu \rangle \right)^2 = \langle g \otimes g, \mu_h^2 \rangle, \quad \mu \in M_{\leq 1}(\mathbb{R}), \quad g \in C_c^\infty(\mathbb{R}),$$

for $\mu_h(dx) = h(x)\mu(dx) - \langle h, \mu \rangle \mu(dx)$. Observe that L maps $P_c^\infty(\mathbb{R})$ to $C(M_{\leq 1}^v(\mathbb{R}))$ and by Theorem III.5.1(ii) it satisfies the positive maximum principle on $M_1(\mathbb{R})$. By Remark III.4.7, Lemma III.3.6 guarantees the existence of an $M_1(\mathbb{R})$ -solution to the martingale problem for L . In order to have a better intuition about such a solution, observe that a simple application of the Itô formula proves that every $M_1(\mathbb{R})$ -valued process X with càdlàg paths satisfying

$$d\langle g, X_t \rangle = (\langle gh, X_t \rangle - \langle h, X_t \rangle \langle g, X_t \rangle) dW_t, \quad \langle g, X_0 \rangle = \langle g, \mu \rangle, \quad \forall g \in C_c^\infty(\mathbb{R}),$$

where W denotes a brownian motion, is an $M_1(\mathbb{R})$ -solution with initial condition μ . More explicitly, one can for instance let $X_t = \sum_{i=1}^n Y_t^i \delta_{x_i}$ for some $x_1, \dots, x_n \in$

\mathbb{R} , where Y solves

$$dY_t^i = \left(h(x_i) - \sum_{j=1}^n h(x_j) Y_t^j \right) Y_t^i dW_t, \quad Y_0^i = y^i,$$

and $\mu = \sum_{i=1}^n y_i \delta_{x_i} \in M_1(\mathbb{R})$.

Heuristically, this martingale consists of a flow of probability measures, whose mass performs an oscillation between the region where the value of h is over its mean with respect to the measure, and the the region where it is below the mean. Choosing for instance $h(x) = x$, we get an oscillation between the right and the left tails of the distribution.

This construction does not depend on the choice of \mathbb{R} as state space, indeed L satisfies the positive maximum principle on $M_1(E)$ for every closed $E \subseteq \mathbb{R}$.

5.2.2 Polynomial jump-diffusions

Let L be a Lévy type operator for $\Gamma = 0$, $Q = 0$,

$$N(\mu, d\xi) = \gamma(\mu, y) * F(dy), \quad \text{and} \quad B(g, \mu) = \int \langle g, \xi - \mu \rangle N(\mu, d\xi) \quad (5.2)$$

for all $g \in C_c^\infty(\mathbb{R})$ and $\mu \in M_{\leq 1}(\mathbb{R})$, where $\gamma(\mu, y) = \langle \nu(\cdot, y), \mu \rangle$ for some continuous map $\nu(\cdot, y) : \mathbb{R} \rightarrow M_1(\mathbb{R})$, and F is a finite measure on \mathbb{R} . More explicitly, L is given by

$$Lp(\mu) = \int p(\gamma(\mu, y)) - p(\mu) F(dy), \quad \mu \in M_{\leq 1}(\mathbb{R}).$$

With this specification, if a jump occurs the moments of the process would jump from $\langle g, \mu \rangle$ to $\langle g, \gamma(\mu, y) \rangle = \int \langle g, \nu(x, y) \rangle \mu(dx)$, where y is F -distributed.

Observe that L clearly satisfies the positive maximum principle on $M_1(\mathbb{R})$ and condition (III.3.2). However, in order to apply Lemma III.3.6, one needs to check case by case if L maps $P_c^\infty(\mathbb{R})$ to $C(M_{\leq 1}^v(\mathbb{R}))$. We now propose some basic examples satisfying this condition. It is straightforward to show that all of the are $M_1(\mathbb{R})$ -polynomial in sense of Definition 2.1.

- (i) Let $\nu(x, y) = \mathcal{N}(y, 1)$. If a jump occurs, the probability measure-valued jump-diffusion jumps to $\mathcal{N}(y, 1)$ where y is F -distributed. The resulting operator is given by

$$Lp(\mu) = \int p(\mathcal{N}(y, 1)) - p(\mu) F(dy).$$

- (ii) Let $F = \frac{1}{2}(\delta_0 + \delta_1)$ and

$$\nu(x, 0) = \delta_{-x} \quad \text{and} \quad \nu(x, 1) = \delta_0.$$

If a jump occurs, with probability 1/2 the support of μ is reflected with respect to 0, and with probability 1/2 the process jumps to 0. The resulting operator is given by $L(\langle g, \cdot \rangle^k)(\mu) = \frac{1}{2}(\langle \bar{g}, \mu \rangle^k + g(0)^k) - \langle g, \mu \rangle^k$, where $\bar{g}(x) = g(-x)$.

- (iii) Let $\nu(x, y) = \mathcal{U}([x, x + y])$. If a jump occurs, the measure valued process jumps from μ to $\gamma(\mu, y)(dz) = \frac{1}{y}\mu([z - y, z])dz$ where y is F -distributed. The resulting operator is given by

$$L(\langle g, \cdot \rangle^k)(\mu) = \int \left\langle \frac{1}{y} \int g(z) \mathbb{1}_{[\cdot, \cdot + y]}(z) dz - g, \mu \right\rangle^k F(dy).$$

It is important to note that with this specification B cannot be chosen to be 0, or equivalently, the operator L cannot be the generator of an $M_1(E)$ -valued martingale. Indeed, by Lemma III.4.4 we know that if L satisfies the positive maximum principle on $M_1(\mathbb{R})$, the operator B needs to be greater or equal the quantity specified in (5.2) for all $g \in C_c^\infty(\mathbb{R})$ such that $\langle g, \mu \rangle = \sup_{\mathbb{R}} g$. For those g , the expression $\int \langle g, \gamma(\mu, y) - \mu \rangle F(dy)$ is nonpositive with equality only if $g(x) = \sup_{\mathbb{R}} g$ for all $x \in \text{supp}(\gamma(\mu, y))$. But this automatically implies that $\text{supp}(\gamma(\delta_x, y)) \subseteq \{x\}$ for each x and the form of γ implies then that $\gamma(\mu, y) = \mu$ for F -a.e. $y \in \mathbb{R}$, which is not possible.

5.2.3 Spatial jump-diffusions

In this sections we illustrate a possible interplay between B , and Q and H , respectively. The corresponding solution to the martingale problem will be a flow of probability which are all identical to their initial condition, up to a support's shift. This class of jump-diffusions inspire the third condition for optimality of Theorem III.5.1, and in particular of Corollary III.5.7. The corresponding generators satisfy the second condition of Lemma III.4.4 in a nontrivial way, meaning that the expression $Q(g \otimes g, \mu) + \int \langle g, \xi - \mu \rangle^2 N(\mu, d\xi)$ does not vanish for all test functions g and probabilities μ such that $\langle g^2, \mu \rangle = 0$, as it is the case for generators of diffusions of Fleming–Viot type.

The most illustrative example is probably given by δ_{W_t} where W represents a brownian motion and has been already explained in details in Section 5.1.4.

For any $x \in \mathbb{R}$, let Z^x be a strong solution to the SDE

$$dZ_t^x = b(Z_t^x)dt + \tau(Z_t^x)dW_t^0 + \int \ell(Z_{t-}^x, y)(\mathcal{P}^0(dt, dy) - F(dy)dt), \quad Z_0^x = x$$

where $b, \tau \in C_\Delta^\infty(\mathbb{R})$, $\ell \in C_\Delta^\infty(\mathbb{R}^2)$, W^0 is a brownian motion, and \mathcal{P}^0 is a poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $F(dy)dt$. Define L as a Lévy type operator for $\Gamma = 0$,

$$B(g, \mu) = \langle \mathcal{G}g, \mu \rangle, \quad Q(g \otimes g, \mu) = \langle \tau g', \mu \rangle^2, \quad \text{and} \quad N(\mu, d\xi) = \gamma(\mu, y)_* F(dy),$$

for all $g \in C_c^\infty(\mathbb{R})$ and $\mu \in M_{\leq 1}(\mathbb{R})$, where

$$\mathcal{G}g = bg' + \frac{1}{2}\tau^2 g'' + \int g(\cdot + \ell(\cdot, y)) - g - g'\ell(\cdot, y)F(dy)$$

denotes the generator of each Z^x and $\gamma(\mu, y) = \langle \delta_{\cdot + \ell(\cdot, y)}, \mu \rangle = (\cdot + \ell(\cdot, y))_* \mu$. Observe that L maps $P_c^\infty(\mathbb{R})$ to $C(M_{\leq 1}^v(\mathbb{R}))$ and by Theorem III.5.1(i) and Theorem III.5.3 it satisfies the positive maximum principle on $M_{\leq 1}(\mathbb{R})$. Intuitively,

we expect the corresponding $M_1(\mathbb{R})$ -solution to be driven by W^0 and \mathcal{P}^0 . The next lemma illustrates how this is in fact the case.

Lemma 5.2. $X_t = (Z_t)_* \mu$ is an $M_1(\mathbb{R})$ -solution to the martingale problem for L with initial condition μ . For example, choosing $\mu = \mathcal{U}(a, b)$ we get $X_t = \mathcal{U}(Z_t^a, Z_t^b)$.

Proof. Let X_t and Z^\cdot be as in the lemma and fix $\mu \in M_1(\mathbb{R})$. In order to simplify the notation set also $\mu(dx) = \mu(dx_1) \cdots \mu(dx_k)$ and $Z_t^x = (Z_t^{x_1}, \dots, Z_t^{x_k})$ for each $x \in \mathbb{R}^k$. Then by Itô formula we can compute

$$\langle g, X_t \rangle^k = \int \prod_{i=1}^k g(Z_t^{x_i}) \mu(dx) = \langle g, \mu \rangle^k + \int \int_0^t L^k g(Z_s^x) ds \mu(dx) + (\text{martingale}) \quad (5.3)$$

where

$$\begin{aligned} L^k g(Z_s^x) &= \sum_{i=1}^k \mathcal{G}g(Z_s^{x_i}) \prod_{j \neq i} g(Z_s^{x_j}) + \sum_{j \neq i} \tau(Z_s^{x_i}) \tau(Z_s^{x_j}) g'(Z_s^{x_i}) g'(Z_s^{x_j}) \prod_{\ell \neq i, j} g(Z_s^{x_\ell}) \\ &\quad + \int \prod_{i=1}^k g(Z_s^{x_i} + \ell(Z_s^{x_i}, y)) - \prod_{i=1}^k g(Z_s^{x_i}) \\ &\quad - \sum_{i=1}^k \left(g(Z_s^{x_i} + \ell(Z_s^{x_i}, y)) - g(Z_s^{x_i}) \right) \prod_{j \neq i} g(Z_s^{x_j}) F(dy). \end{aligned}$$

Since $\int L^k g(Z_s^x) \mu(dx) = L(\langle g, \cdot \rangle^k)(X_s)$, this concludes the proof. \square

Let us now analyze this result for some particular choices of the parameters. Setting $b = \tau = 0$, $\ell(\cdot, y) = y$, and $F = \delta_1$ we get $Q = 0$,

$$B(g, \mu) = \langle g(\cdot + 1) - g - g', \mu \rangle, \quad \text{and} \quad N(\mu, d\xi) = \delta_{(\cdot + 1)_* \mu}.$$

The processes Z^x is then given by $Z_t^x = x + N_t - t$, where N is a homogeneous Poisson process with intensity 1. The $M_1(\mathbb{R})$ -solution to the corresponding martingale problem is then given by

$$X_t = (\cdot + N_t - t)_* \mu.$$

For example, choosing $\mu = \mathcal{U}(a, b)$ we get $X_t = \mathcal{U}(a + N_t - t, b + N_t - t)$.

This behavior is similar to the case where $b = 0$, $\tau = 1$, and $F = 0$, and hence

$$N = 0, \quad B(g, \mu) = \left\langle \frac{1}{2} g'', \mu \right\rangle, \quad \text{and} \quad Q(g \otimes g, \mu) = \langle g', \mu \rangle^2,$$

and the $M_1(\mathbb{R})$ -solution to the corresponding martingale problem is given by

$$X_t = (\cdot + W_t)_* \mu.$$

Again, choosing $\mu = \mathcal{U}(a, b)$ we get $X_t = \mathcal{U}(a + W_t, b + W_t)$.

5.3 Particle systems with mean fields interaction

Starting by the work of McKean (1967) limits of empirical measure processes for systems of interacting particles have been studied by several authors (among many others, Dawson, Ethier, Kotolenez, Kurtz, Sznitman, Xiong, ...). Those processes fit very nicely also in the framework presented in this thesis. In particular, they can be described using the technology of the martingale problem and their properties can be understood by studying the corresponding generator.

In the setting of Lévy type operators, consider three continuous maps $b, \sigma, \tau : M_{\leq 1}^v(\mathbb{R}) \rightarrow C(\mathbb{R})$ given by $\mu \mapsto b_\mu$, $\mu \mapsto \sigma_\mu$, and $\mu \mapsto \tau_\mu$. This in particular means that b_{μ_n} converges to b_μ uniformly on compact sets whenever μ_n goes to μ vaguely, and the same is true for σ and τ . For $g \in C_c^\infty(\mathbb{R})$ set then

$$B_\mu g = b_\mu g' + \frac{1}{2}(\sigma_\mu^2 + \tau_\mu^2)g'', \quad \Sigma_\mu g = \tau_\mu g',$$

for all $\mu \in M_{\leq 1}^v(\mathbb{R})$. Let L be a Lévy type operator for $\Gamma = 0$, $N = 0$,

$$B(g, \mu) = \langle B_\mu g, \mu \rangle, \text{ and } Q(g \otimes g, \mu) = \langle \Sigma_\mu g, \mu \rangle^2, \quad \mu \in M_{\leq 1}(\mathbb{R}), \quad g \in C_c^\infty(\mathbb{R}).$$

By Theorem III.5.1(i) and Theorem III.5.3 L satisfies the positive maximum principle on $M_{\leq 1}(\mathbb{R})$.

By means of the next two lemmas, we will now verify that all the conditions of Lemma III.3.6 are satisfied. This will let us conclude that the martingale problem for L has an $M_1(\mathbb{R})$ -solution for each initial condition $\mu \in M_1(\mathbb{R})$.

Lemma 5.3. *L maps $P_c^\infty(\mathbb{R})$ to $C(M_{\leq 1}^v(\mathbb{R}))$.*

Proof. Fix $p(\mu) := \langle g, \mu \rangle^k$ for some $g \in C_c^\infty(\mathbb{R})$ and $k \in \mathbb{N}$ and note that

$$\partial p(\mu) = k \langle g, \mu \rangle^{k-1} g \quad \text{and} \quad \partial^2 p(\mu) = k(k-1) \langle g, \mu \rangle^{k-1} g \otimes g.$$

Fix now $(\mu_n)_n \subseteq M_{\leq 1}(\mathbb{R})$ and $\mu \in M_{\leq 1}(\mathbb{R})$ such that μ_n converges vaguely to μ . Then, the properties of b, σ and τ guarantee that

$$\varepsilon_n := \|B_{\mu_n} \partial p(\mu_n) - B_\mu \partial p(\mu)\| = k |\langle g, \mu \rangle|^{k-1} \|B_{\mu_n} g - B_\mu g\|$$

converges to 0 for n going to infinity. Since the same properties guarantee that $B_\mu g \in C_c(\mathbb{R})$, we get

$$|B(g, \mu_n) - B(g, \mu)| \leq \varepsilon_n + k |\langle g, \mu \rangle|^{k-1} |\langle B_\mu g, \mu_n - \mu \rangle| \xrightarrow{n \rightarrow \infty} 0.$$

Proceeding similarly for Q and using that $C(M_{\leq 1}^v(\mathbb{R}))$ is a Polish space and thus a sequential space, we can conclude that L maps $P_c^\infty(\mathbb{R})$ to $C(M_{\leq 1}^v(\mathbb{R}))$. \square

Lemma 5.4. *Assume that for some $K \geq 0$*

$$x b_\mu(x), \sigma_\mu(x)^2 + \tau_\mu(x)^2 \leq K(1 + x^2) \tag{5.4}$$

for all $x \in \mathbb{R}$ and $\mu \in M_{\leq 1}(\mathbb{R})$. Then condition (III.3.2) holds true.

Proof. Fix $h \in C_c^\infty(\mathbb{R})$ such that $\mathbf{1}_{[-1,1]} \leq h \leq \mathbf{1}_{[-2,2]}$ and $h'(x)x \leq 0$ for all $x \in \mathbb{R}$. Then set $g_n(x) := h(x/n)$ and note that $\sup_n \|g_n\| = 1$ and $g_n(x) \rightarrow 1$ for all $x \in \mathbb{R}$. Moreover

$$\begin{aligned} B_\mu g_n(x) &= \frac{1}{n^2} \left(-b_\mu(x)x \frac{|h'(x/n)|}{|x|/n} + \frac{1}{2} (\sigma_\mu(x)^2 + \tau_\mu(x)^2) h''(x/n) \right) \\ &\geq -\frac{K(1+(2n)^2)}{n^2} \sup_{x \in \mathbb{R}} \left(\frac{|h'(x)|}{|x|} + \frac{1}{2} (h''(x))^- \right) \geq -C \end{aligned}$$

for some $C \in \mathbb{R}_+$, proving that $L(\langle g_n, \cdot \rangle)(\mu)$ is bounded from below with respect to μ and n . Finally, for all $\mu \in M_{\leq 1}(\mathbb{R})$ by Fatou lemma we can compute

$$\liminf_{n \rightarrow \infty} L(\langle g_n, \cdot \rangle)(\mu) = \liminf_{n \rightarrow \infty} \langle B_\mu g_n, \mu \rangle \geq \langle \lim_{n \rightarrow \infty} B_\mu g_n, \mu \rangle = 0.$$

□

5.3.1 Dynamics of the first order moments

Let now X be an $M_1(\mathbb{R})$ -solution to the martingale problem corresponding to L . By (III.3.1) and Lemma III.4.6, we know that for all $g, h \in C_c^\infty(\mathbb{R})$

$$\begin{aligned} \langle g, X_t \rangle - \langle g, X_0 \rangle - \int_0^t \langle B_{X_s} g, X_s \rangle ds \quad \text{is a martingale} \\ \text{PQC}(\langle g, X \rangle, \langle h, X \rangle) = \int Q(g \otimes h, X_s) ds = \int \langle \Sigma_{X_s} g, X_s \rangle \langle \Sigma_{X_s} h, X_s \rangle ds. \end{aligned}$$

Moreover, since $\Gamma = 0$ and $N = 0$ we also know that X is conservative and continuous (see Lemma III.4.5(ii)). Combining those observations we can see that the dynamics of X are compatible with the following system of SDEs

$$d\langle g, X_t \rangle = \langle B_{X_t} g, X_t \rangle dt + \langle \Sigma_{X_t} g, X_t \rangle dW_t^0 \quad \forall g \in C_c^\infty(\mathbb{R}),$$

where W^0 denotes a brownian motion.

5.3.2 Particle systems and McKean-Vlasov equations

Consider a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and let $(X, (Z^i)_{i \in \mathbb{N}})$ be a solution of the system

$$X_t = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{Z_t^i} \quad \text{and} \quad dZ_t^i = b_{X_t}(Z_t^i) dt + \sigma_{X_t}(Z_t^i) dW_t^i + \tau_{X_t}(Z_t^i) dW_t^0, \quad (5.5)$$

with initial value $Z_0^i = x \in \mathbb{R}$, where W^0, W^1, \dots are independent brownian motions. Assume that Z^1, Z^2, \dots are exchangeable. By De Finetti theorem (see e.g. Theorem 4.1 in Kotelenetz and Kurtz (2008) or, for a general overview, also Section 12.3 in Klenke (2013)) we get that $(Z_t^i)_{i \in \mathbb{N}}$ are conditionally i.i.d. with respect to the invariant σ -algebra $\mathcal{F}_t^\infty = \sigma(X_s, s \leq t)$ and $X_t = \mathbb{P}(Z_t^1 \in \cdot | \mathcal{F}_t^\infty)$.

Lemma 5.5. *X is an $M_1(\mathbb{R})$ -solution to the martingale problem for L with initial condition δ_x .*

Proof. Set $Z := Z^1$ and note that for all $g_1, \dots, g_k \in C_c^\infty(\mathbb{R})$

$$\mathbb{E}[g_1(Z_t^1) \cdots g_k(Z_t^k) | \mathcal{F}_t^\infty] = \mathbb{E}[g_1(Z_t) | \mathcal{F}_t^\infty] \cdots \mathbb{E}[g_k(Z_t) | \mathcal{F}_t^\infty].$$

For all $g \in C_c^\infty(\mathbb{R})$ setting $\bar{Z} := (Z^1, \dots, Z^k)$ we then get that

$$N_t^{g,k} := g(Z_t^1) \cdots g(Z_t^k) - g(x)^k - \int_0^t L^k g(\bar{Z}_s, X_s) ds$$

is a bounded $(\mathcal{F}_t)_{t \geq 0}$ -martingale, where

$$\begin{aligned} L^k g(\bar{Z}_s, X_s) &= \sum_{i=1}^k \left(b_{X_s}(Z_s^i) g'(Z_s^i) + \frac{1}{2} (\sigma_{X_s}(Z_s^i) + \tau_{X_s}(Z_s^i)) g''(Z_s^i) \right) \prod_{j \neq i} g(Z_s^j) \\ &\quad + \sum_{i=1}^k \sum_{j \neq i} \tau_{X_s}(Z_s^i) \tau_{X_s}(Z_s^j) g'(Z_s^i) g'(Z_s^j) \prod_{\ell \neq i, j} g(Z_s^\ell). \end{aligned}$$

Since $\mathcal{F}_t^\infty \subseteq \mathcal{F}_t$ this implies that $\mathbb{E}[N_t^{g,k} | \mathcal{F}_t^\infty]$ is an $(\mathcal{F}_t^\infty)_{t \geq 0}$ -martingale and hence

$$\begin{aligned} \mathbb{E}[\langle g, X_t \rangle^k | \mathcal{F}_s^\infty] - \langle g, X_s \rangle^k &= \mathbb{E} \left[\int_s^t L^k g(\bar{Z}_u, X_u) du \middle| \mathcal{F}_s^\infty \right] \\ &= \mathbb{E} \left[\int_s^t \mathbb{E}[L^k g(\bar{Z}_u, X_u) | \mathcal{F}_u^\infty] du \middle| \mathcal{F}_s^\infty \right] \\ &= \mathbb{E} \left[\int_s^t L(\langle g, \cdot \rangle^k)(X_u) du \middle| \mathcal{F}_s^\infty \right] \end{aligned}$$

proving that X is an $M_1(\mathbb{R})$ -solution to the martingale problem for L . \square

We now propose an interesting remark based on Theorem 2.3 in [Kurtz and Xiong \(1999\)](#) and its proof.

Remark 5.6. *Under the additional assumption of pathwise uniqueness for the solution of (5.5), we get that*

$$X_t = \mathbb{P}(Z_t^1 \in \cdot | \mathcal{F}_t^0), \quad \text{where} \quad \mathcal{F}_t^0 = \sigma(W_s^0, s \leq t).$$

In fact, as a characterization of X , the particle system (5.5) is essentially equivalent to the following McKean–Vlasov equation

$$X_t = \mathbb{P}(Z_t \in \cdot | \mathcal{F}_t^0) \quad \text{and} \quad dZ_t = b_{X_t}(Z_t) dt + \sigma_{X_t}(Z_t) dW_t^1 + \tau_{X_t}(Z_t) dW_t^0, \quad (5.6)$$

where we require Z to be compatible with (W^1, W^0) in the sense that, for each $t \geq 0$, the increments of W^1 and W^0 after t are independent of $\sigma(Z_s, W_s, W_s^0, s \leq t)$.

More precisely, the existence of a solution (X, Z, W^1, W^0) of (5.6) yields the existence of a solution

$$(\tilde{X}, (\tilde{Z}^i)_{i \in \mathbb{N}}, (\tilde{W}^i)_{i \in \mathbb{N}}, \tilde{W}^0)$$

of (5.5) such that $(\tilde{X}, \tilde{Z}^1, \tilde{W}^1, \tilde{W}^0)$ has the same distribution as (X, Z, W^1, W^0) . Conversely, if $(X, (Z^i)_{i \in \mathbb{N}}, (W^i)_{i \in \mathbb{N}}, W^0)$ is a pathwise unique solution of (5.5), then (X, Z^1, W^1, W^0) is a solution of (5.6).

5.3.3 A jump-diffusion with mean fields interactions

Consider now a probability measure F on \mathbb{R} , and a continuous map $\ell : M_{\leq 1}(\mathbb{R}) \rightarrow C_c(\mathbb{R} \times \mathbb{R})$ given by $\mu \mapsto \ell_\mu$. For $g \in C_c^\infty(\mathbb{R})$ set then

$$B_\mu g = b_\mu g' + \frac{1}{2}(\sigma_\mu^2 + \tau_\mu^2)g'' + \int g(\cdot + \ell_\mu(\cdot, y)) - g - g' \ell_\mu(\cdot, y) F(dy), \quad \Sigma_\mu g = \tau_\mu g',$$

for all $\mu \in M_{\leq 1}^v(\mathbb{R})$. Let then L be a Lévy type operator for $\Gamma = 0$,

$$B(g, \mu) = \langle B_\mu g, \mu \rangle, \quad Q(g \otimes g, \mu) = \langle \Sigma_\mu g, \mu \rangle^2 \quad \text{and} \quad N(\mu, d\xi) = \gamma(\mu, y)_* F(dy),$$

where $\gamma(\mu, y) = (\cdot + \ell_\mu(\cdot, y))_* \mu$, for all $\mu \in M_{\leq 1}(\mathbb{R})$, $g \in C_c^\infty(\mathbb{R})$. Note that L is explicitly given by

$$\begin{aligned} Lp(\mu) &= \langle B_\mu \partial p(\mu), \mu \rangle + \frac{1}{2} \langle \Sigma_\mu \otimes \Sigma_\mu \partial^2 p(\mu), \mu^2 \rangle \\ &\quad + \int p(\gamma(\mu, y)) - p(\mu) - \langle \partial p(\mu), \gamma(\mu, y) - \mu \rangle F(dy). \end{aligned} \quad (5.7)$$

Consider then a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and consider a solution $(X, (Z^i)_{i \in \mathbb{N}})$ of the system given by $X_t = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{Z_t^i}$ and

$$dZ_t^i = b_{X_t}(Z_t^i)dt + \sigma_{X_t}(Z_t^i)dW_t^i + \tau_{X_t}(Z_t^i)dW_t^0 + \int \ell_{X_t}(Z_t^i, y)(\mathcal{P}^0(dt, dy) - F(dy)dt)$$

with initial value $Z_0^i = x \in \mathbb{R}$, where W^0, W^1, \dots are independent brownian motions and \mathcal{P}^0 is a poisson random measure with compensator $F(dy)dt$. Again, assume that Z^1, Z^2, \dots are exchangeable.

Lemma 5.7. *X is an $M_1(\mathbb{R})$ -solution to the martingale problem for L with initial condition δ_x .*

Proof. By De Finetti theorem we get that $(Z_t^i)_{i \in \mathbb{N}}$ are conditionally i.i.d. with respect to the invariant σ -algebra $\mathcal{F}_t^\infty = \sigma(X_s, s \leq t)$ and $X_t = \mathbb{P}(Z_t^1 \in \cdot | \mathcal{F}_t^\infty)$. The proof follows the proof of Lemma 5.5. \square

5.3.4 Interactions through weighted empirical measures

[Kurtz and Xiong \(1999\)](#) provides a representation of the unique solution of a class of nonlinear stochastic differential equations by means of a wighted empirical measure. Since we are interested in probability measure-valued processes, we restrict our attention to those cases where sum of the weights is equal to 1.

We approach this framework following the cited article. Consider a continuous map $v : M_{\leq 1}(\mathbb{R}) \rightarrow C_c(\mathbb{R})$ given by $\mu \mapsto v_\mu$. Consider then a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and consider a solution $(X, (Z^i)_{i \in \mathbb{N}}, (Y^i)_{i \in \mathbb{N}})$ of the system given by

$$\begin{aligned} X_t &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_t^i \delta_{Z_t^i} \\ dZ_t^i &= b_{X_t}(Z_t^i)dt + \sigma_{X_t}(Z_t^i)dW_t^i + \tau_{X_t}(Z_t^i)dW_t^0 \quad \text{and} \quad dY_t^i = Y_t^i v_{X_t}(Z_t^i)dW_t^i, \end{aligned} \quad (5.8)$$

with initial value $Z_0^i = x \in \mathbb{R}$, $Y_0^i = 1$, where W^0, W^1, \dots are independent brownian motions. Assume also that $(Z^1, Y^1), (Z^2, Y^2), \dots$ are exchangeable. Observe that, differently from Section 5.1.5, each weight $\frac{1}{n}Y^i$ is correlated with the corresponding particle Z^i but it is uncorrelated to the other weights $\frac{1}{n}Y^j$, for $j \neq i$.

Proceeding as in the proof of Proposition 2.1 of Kurtz and Xiong (1999) we can show that $\mathbb{E}[\sup_{0 \leq s \leq t} (Y_s^i)^2] < \infty$ and following the proof of Theorem 3.1 of the same paper, we can then show that by exchangeability of $(Y^i)_{i \in \mathbb{N}}$ and boundedness of $(\mu, z) \mapsto v_\mu(z)$

$$X_t(\mathbb{R}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_t^i = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left(1 + \int_0^t Y_s^i v_{X_s}(Z_s^i) dW_s^i \right) = 1.$$

We can thus conclude that $X_t \in M_1(\mathbb{R})$, for all $t \geq 0$. It is also interesting to note that by the De Finetti theorem and integrability of $\sup_{0 \leq s \leq t} (Y_s^1)^2$ we know that

$$X_t = \mathbb{E}[Y_t^1 \mathbb{1}_{\{Z_t^1 \in \cdot\}} | \mathcal{F}_t^\infty],$$

where $\mathcal{F}_t^\infty = \sigma(\bar{X}_s, s \leq t)$ for $\bar{X}_t = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{(Z_t^i, Y_t^i)}$.

For $g \in C_c^\infty(\mathbb{R})$ and $\mu \in M_{\leq 1}(\mathbb{R})$ set then

$$B_\mu g = (b_\mu + \sigma_\mu v_\mu)g' + \frac{1}{2}(\sigma_\mu^2 + \tau_\mu^2)g'', \quad \Sigma_\mu g = \tau_\mu g',$$

and define L as a Lévy type operator for $\Gamma = 0$, $N = 0$

$$B(g, \mu) = \langle B_\mu g, \mu \rangle \quad \text{and} \quad Q(g \otimes g, \mu) = \langle \Sigma_\mu g, \mu \rangle^2.$$

Lemma 5.8. *X is an $M_1(\mathbb{R})$ -solution to the martingale problem for L with initial condition δ_x .*

Proof. Set $Z := Z^1$ and $Y := Y^1$. By De Finetti theorem we get that $(Z_t^i, Y_t^i)_{i \in \mathbb{N}}$ are conditionally i.i.d. with respect to $(\mathcal{F}_t^\infty)_{t \geq 0}$. For all $g \in C_c^\infty(\mathbb{R})$, setting $\bar{Z} := (Z^1, \dots, Z^k)$ and $\bar{Y} := (Y^1, \dots, Y^k)$ we then we have that

$$N_t^{g,k} := Y_t^1 g(Z_t^1) \cdots Y_t^k g(Z_t^k) - g(x)^k - \int_0^t L_k g(\bar{Z}_s, \bar{Y}_s, X_s) ds$$

is a bounded $(\mathcal{F}_t)_{t \geq 0}$ -martingale, where

$$\begin{aligned} L_k g(z, y, \mu) &= \sum_{i=1}^k y_i B_\mu g(z_i) \prod_{j \neq i} y_j g(z_j) \\ &\quad + \sum_{j \neq i} \tau_\mu(z_i) \tau_\mu(z_j) y_i g'(z_i) y_j g'(z_j) \prod_{\ell \neq i, j} y_\ell g(z_\ell). \end{aligned}$$

Since $\mathcal{F}_t^\infty \subseteq \mathcal{F}_t$ this implies that $\mathbb{E}[N_t^{g,k} | \mathcal{F}_t^\infty]$ is an $(\mathcal{F}_t^\infty)_{t \geq 0}$ -martingale and hence proceeding as in the proof of Lemma 5.5 we can conclude that X is an $M_1(\mathbb{R})$ -solution to the martingale problem for L . \square

It is interesting to note that as in Remark 5.6 if we additionally assume path-wise uniqueness of the solution of the particle system (5.8) we get that

$$X_t = \mathbb{E}[Y_t^1 \mathbf{1}_{\{Z_t^1 \in \cdot\}} | \mathcal{F}_t^0],$$

for $\mathcal{F}_t^0 = \sigma(W_s^0, s \leq t)$.

Observe that by Section 5.3.2 we know that L is the generator of

$$\tilde{X}_t := \mathbb{P}(\tilde{Z}_t \in \cdot | \mathcal{F}_t^0),$$

where the dynamics of \tilde{Z} coincides with those of Z up to a correction term $\sigma_\mu \nu_\mu$ in the drift.

5.4 $M_1(E)$ -solutions taking values in a subset of $M_1(E)$

Consider again the setting of Lévy type operators. There are cases, where we are interested in $M_1(E)$ -solutions taking values in some subset M of $M_1(E)$ being closed in a suitable sense.² One can show that in order to prove the existence of such a solution one does not need to check the positive maximum principle on M instead of $M_1(E)$. This conclusion is reasonable, since it is intuitively clear that the behavior of a potential solution outside M should not affect the behavior of a solution starting in M and living in M .

We provide here two examples in this sense. In the first one, M is given by $M_n := \{\nu \in M_1(\mathbb{R}) : \nu = \frac{1}{n} \sum_{i=1}^n \delta_{z_i}, z_i \in \mathbb{R}\}$ and in the second one

$$M_{\mathbb{N}} := \left\{ \nu \in M_1(\mathbb{N} \times \mathbb{R}) : \nu = \delta_n \times \frac{1}{n} \sum_{i=1}^n \delta_{z_i}, z_i \in \mathbb{R}, n \in \mathbb{N} \right\}.$$

Observe that, in the second example the killing parameter Γ will not assumed to be 0.

5.4.1 Systems of finitely many particles

Fix $n \in \mathbb{N}$ and consider $b, \sigma, \tau \in C(\mathbb{R})$. For $g \in C_c^\infty(\mathbb{R})$ set then

$$\mathcal{G}g = bg' + \frac{1}{2}(\sigma^2 + \tau^2)g'' \quad \text{and} \quad \Sigma^\tau g = \tau g',$$

and define L as a Lévy type operator for $\Gamma = 0, N = 0$,

$$B(g, \nu) = \langle \mathcal{G}g, \nu \rangle \quad \text{and} \quad Q(g \otimes g, \nu) = \frac{1}{n} \langle (\Sigma^\sigma g)^2, \nu \rangle + \langle \Sigma^\tau g, \nu \rangle^2.$$

Consider a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and a solution $Z := (Z^1, \dots, Z^n)$ of

$$dZ_t^i = b(Z_t^i)dt + \sigma(Z_t^i)dW_t^i + \tau(Z_t^i)dW_t^0,$$

with initial value $Z_0^i = x \in \mathbb{R}$, where W^0, W^1, \dots, W^n are $(n + 1)$ independent brownian motions.

² Since the procedure for checking existence always start with an embedding in $M_{\leq 1}(E)$, we need to guarantee that $M = \{\nu \in \bar{M} : \nu(E) = 1\}$ for some closed subset \bar{M} of $M_{\leq 1}^v(E)$.

Lemma 5.9. $X = \frac{1}{n} \sum_{i=1}^n \delta_{Z_i}$ is an $M_1(\mathbb{R})$ -solution to the martingale problem for L with initial condition δ_x .

Proof. The result follows by an application of the Itô lemma to $\langle g, X_t \rangle^k = \frac{1}{n^k} (\sum_{i=1}^n \delta_{Z_i})^k$ for all $g \in C_c^\infty(\mathbb{R})$. \square

Even if one cannot guarantee that the positive maximum principle on $M_{\leq 1}(\mathbb{R})$ holds true, it holds on M_n . Observe also that for $n \rightarrow \infty$ the operator L converges to the operator described at the beginning of Section 5.3 (without mean fields interactions).

5.4.2 Empirical measures of branching processes

This last example belongs to the frameworks of mean field games with branching. It is inspired by the work of Claisse, Ren, and Tan presented during the workshop *Advances in Stochastic Analysis for Risk Modeling*, taking place from 13 to 17 November 2017 at the CIRM in Luminy. Also in this case, the technology of the martingale problem can be used to understand some properties of the process, for instance when the population is assumed to be asymptotically large.

Consider a population whose branching dynamics are given by the following characteristics. Each individual Z^i follows a diffusion which is independent to the other individuals and whose generator is given by

$$\mathcal{G}f(z) = -\lambda(z)f(z) + b(z)f'(z) + \frac{1}{2}f''(z).$$

This in particular implies that the deathtime τ_i of the i -th individual is given by

$$\mathbb{P}(\tau_i \geq t | Z_s^i, s \leq t) = \exp\left(-\int_0^t \lambda(Z_s^i) ds\right).$$

While the i -th individual vanishes, she gives birth to new individuals. We denote by F the distribution on \mathbb{N}_0 of the number of offsprings and we assume for simplicity that the offsprings inherit the position of their parent. Define then $V_t := \{i : Z^i \text{ is alive at time } t\}$ and consider the $M_1(\mathbb{N} \times \mathbb{R})$ -valued process

$$X_t = \delta_{|V_t|} \times \frac{1}{|V_t|} \sum_{i \in V_t} \delta_{Z_t^i},$$

where $|V_t|$ denote the number of elements in V_t . Our goal is now to provide a Lévy type operator L such that X is a possibly killed $M_1(\mathbb{R})$ -solution to the martingale problem for L . Before to start with the construction, a small comment about the choice of the state space is due. In fact, we are just interested in the empirical measure $\frac{1}{|V_t|} \sum_{i \in V_t} \delta_{Z_t^i}$ but this process unfortunately not Markov and our theory can thus not be applied. This observation became clear if one notes that the distribution of X_t given $X_s = \delta_0$ depends on $|V_s|$ as well and $|V_s|$ is $\sigma(X_u, u \leq s)$ -measurable.

Observe that if a branch occurs, the process X performs a jump. More precisely, suppose that at time t an individual dies in the position $z \in \mathbb{R}$ and

gives birth to $y \in \mathbb{N}_0$ new individuals. Then if $X_{t-} = \mu_1 \times \mu_2$ we get $X_t = \bar{\mu}_1(y, z) \times \bar{\mu}_2(y, z) =: \bar{\mu}(y, z)$ for

$$\bar{\mu}_1(y, z) := \delta_{\langle w, \mu_1 \rangle - 1 + y} \quad \text{and} \quad \bar{\mu}_2(y, z) := \frac{1}{\langle w, \mu_1 \rangle - 1 + y} \left(\langle w, \mu_1 \rangle \mu_2 + (-1 + y) \delta_z \right)$$

where $w(v) := v$. If $\langle w, \mu_1 \rangle - 1 + y = 0$, the dying individual is the last living one ($\langle w, \mu_1 \rangle = 1$) and she does not give birth to any offspring ($y = 0$). In this case extinction occurs and we write $X_t = \mu^\dagger$, where μ^\dagger denotes the cemetery state.

This reasoning provides the intuition behind the following choice of the killing and jump parameters.

$$\begin{aligned} \Gamma(\mu) &:= - \int \int \mathbf{1}_{\{y=0\}} F(dy) \lambda(z) \mu_2(dz) \langle h_1, \mu_1 \rangle = -F(0) \langle \lambda, \mu_2 \rangle \langle h_1, \mu_1 \rangle \\ N(\mu, d\xi) &:= \bar{\mu}(y, z)_* \left(F_\mu(dy) \times \lambda(z) \langle w, \mu_1 \rangle \mu_2(dz) \right) \end{aligned}$$

where for $h_1 \in C_c^\infty(\mathbb{N})$ such that $h_1(1) = 1$ and $h_1(v) = 0$ for all $v \in \mathbb{N} \setminus \{1\}$, $F_\mu = F$ for $\mu_1 \neq \delta_1$, and $F_\mu(dy) = \mathbf{1}_{\{y>0\}} F(dy)$ if $\mu_1 = \delta_1$. Since between two jumps the process X is just an empirical measure, the intuition behind the choice of B and Q follows from Section 5.4.1:

$$\begin{aligned} B((g_1, g_2), \mu) &:= \int \langle (g_1, g_2), \xi - \mu \rangle N(\mu, d\xi) + \langle g_1, \mu_1 \rangle \langle \bar{\mathcal{G}} g_2, \mu_2 \rangle, \\ Q((g_1, g_2) \otimes (g_1, g_2), \mu) &:= \frac{1}{\langle w, \mu_1 \rangle} \langle g_1, \mu_1 \rangle^2 \langle (g_2')^2, \mu_2 \rangle. \end{aligned}$$

where $\bar{\mathcal{G}}f(z) = b(z)f'(z) + \frac{1}{2}f''(z)$ is the generator of an individual's diffusion, assuming that no killing can occur.

We now (heuristically) analyze the situation where the number of particles is infinite. Fix $g_1 \equiv 1$ and set $g_2 = g$. For $\bar{\mu}_v = \delta_v \times \mu_v$ such that $\mu_v = \frac{1}{v} \sum_{i=1}^v \delta_{z_i}$ and $\mu_v \xrightarrow{v \rightarrow \infty} \mu$ we get that

$$\begin{aligned} L\left(\langle (1, g), \cdot \rangle\right)(\bar{\mu}_v) &= -F(0) \langle \lambda, \mu_v \rangle \mathbf{1}_{\{v=1\}} + \langle \bar{\mathcal{G}}g, \mu_v \rangle \\ &\quad + \int \frac{v(y-1)}{v+y-1} F_{\mu_v}(dy) \left(\langle g\lambda, \mu_v \rangle - \langle g, \mu_v \rangle \langle \lambda, \mu_v \rangle \right) \end{aligned}$$

which for v going to ∞ converges to

$$L^\infty\left(\langle (1, g), \cdot \rangle\right)(\mu) := F_1\left(\langle g\lambda, \mu \rangle - \langle g, \mu \rangle \langle \lambda, \mu \rangle\right) + \langle \bar{\mathcal{G}}g, \mu \rangle,$$

for $F_1 = \int yF(dy) - 1$ being the expected change of size of the population after a death.

The interpretation of the (possibly killed) $M_1(E)$ -solution of the corresponding martingale problem depends on the sign of F_1 . If it is positive, i.e. if the population tends to increase after a death of an individual, we can note that

$$L^\infty\left(\langle (1, g), \cdot \rangle\right)(\mu) = \left\langle \int (g(\xi) - g) F_\mu^+(d\xi) + \bar{\mathcal{G}}g, \mu \right\rangle,$$

where $F_\mu^+(d\xi) = F_1\lambda(\xi)\mu(d\xi)$. Following the intuition of Section 5.1.1 we can see that the possibly killed $M_1(E)$ -solution to the corresponding martingale problem is given by $\mathbb{P}(Z_t^+ \in \cdot)$ where the dynamics of Z^+ are generated by

$$\mathcal{G}^+g := \int (g(\xi) - g)F_\mu^+(d\xi) + \bar{\mathcal{G}}g.$$

An individual whose dynamics are given by Z^+ moves like the original ones before their death till when a jump occurs. Its jump rate is given by $F_1\langle\lambda, \mu\rangle$, which in particular does not depend on the position of the individual. The jumps distribution is then given by F^+ , which in particular is more concentrated to regions with a high mortality rate (λ large) and where more particles are present (i.e. where μ is more concentrated).

If F_1 is negative, i.e. if the population tends to decrease after a death, we can note that

$$L^\infty\left(\langle(1, g), \cdot\rangle\right)(\mu) = \left\langle \int (g(\xi) - g)F_\mu^-(\cdot, d\xi) + \bar{\mathcal{G}}g, \mu \right\rangle,$$

where $F_\mu^-(x, d\xi) = -F_1\lambda(x)\mu(d\xi)$. Again, the (possibly killed) $M_1(E)$ -solution to the corresponding martingale problem is given by $\mathbb{P}(Z_t^- \in \cdot)$ where the dynamics of Z^- are generated by

$$\mathcal{G}^-g := \int (g(\xi) - g)F_\mu^-(d\xi) + \bar{\mathcal{G}}g.$$

Also in this case, an individual whose dynamics are given by Z^- moves like the original ones before their dead. Its jump rate is given by $-F_1\lambda$, which in particular depends on the position of the individual. Finally, the jumps distribution is then given by F_μ^- , which in particular is more concentrated to regions where more particles are present (i.e. where μ is more concentrated).

A Proof of Theorem 2.2 and a generalization

We first prove Theorem 2.2. Assume that L is of the stated form. Then for monomials $p(\nu) = \langle g, \nu \rangle^k$ with $g \in D$, $k \in \mathbb{N}$ and $\nu \in M_1(E)$ one has

$$\begin{aligned} Lp(\nu) &= \langle B(\partial p(\nu)), \nu \rangle + \frac{1}{2} \langle Q(\partial^2 p(\nu)), \nu^2 \rangle \\ &= k \langle g, \nu \rangle^{k-1} \langle B(g), \nu \rangle + \frac{1}{2} k(k-1) \langle g, \nu \rangle^{k-2} \langle Q(g \otimes g), \nu^2 \rangle, \end{aligned}$$

which is a polynomial in ν of degree at most k . Moreover, $L1 = 0$. By linearity, this shows that L is $M_1(E)$ -polynomial. Next, a direct calculation yields

$$\Gamma(p, q)(\nu) = \langle Q(\partial p(\nu) \otimes \partial q(\nu)), \nu^2 \rangle \quad \text{for all } \nu \in M_1(E),$$

which is easily seen to be an $M_1(E)$ -derivation due to the product rule given in Lemma III.2.4(v) for differentiating polynomials.

Conversely, assume L is $M_1(E)$ -polynomial and Γ is an $M_1(E)$ -derivation. Consider arbitrary first degree monomials $q(\nu) = \langle g, \nu \rangle$ and $r(\nu) = \langle h, \nu \rangle$, $g, h \in D$. The $M_1(E)$ -polynomial property and Corollary III.2.6 yield

$$Lq(\nu) = \langle Bg, \nu \rangle \quad \text{for all } \nu \in M_1(E),$$

for some map $B : D \rightarrow C_\Delta(E)$ that are easily seen to be linear due to the linearity of L . Furthermore, the $M_1(E)$ -polynomial property, definition (III.3.3) of Γ , and Corollary III.2.6 imply that

$$\Gamma(q, r)(\nu) = \langle Q(g \otimes h), \nu^2 \rangle \quad \text{for all } \nu \in M_1(E),$$

where Q inherits symmetry and linearity from Γ and take values in $\widehat{C}_\Delta(E^2)$. Thus, by taking linear combinations, we can and do extend them to operators on $D \otimes D$.

Explicit calculation now shows that Lp is of the form (2.1) for $p = q$ and $p = q^2$. Furthermore, since Γ is an $M_1(E)$ -derivation we have $\Gamma(1, 1) = 2\Gamma(1, 1)$, hence $\Gamma(1, 1) = 0$, and therefore $L1 = L(1^2) = 0 + 2L1$. Thus $L1 = 0$, so that (2.1) holds also for $p = 1$.

We now make more substantial use of the fact that Γ is an $M_1(E)$ -derivation in order to extend (2.1) to higher degree monomials. We proceed by induction on k , and assume Lp is of the form (2.1) for all $p = q^l$, $l \leq k$. So far we have proved this for $k = 2$. The definition (III.3.3) of Γ and the fact that it is an $M_1(E)$ -derivation give the identity on $M_1(E)$

$$L(q^{k+1}) = 2qL(q^k) - q^2L(q^{k-1}) + q^{k-1}\Gamma(q, q)$$

for $k \geq 2$. Due to the induction assumption, the right-hand side can be computed explicitly using (2.1). The result is

$$(k+1)q(\nu)^k \langle Bg, \nu \rangle + \frac{1}{2}(k+1)kq(\nu)^{k-1} \langle Q(g \otimes g), \nu^2 \rangle,$$

which is equal to $\langle B(\partial p(\nu)), \nu \rangle + \frac{1}{2} \langle Q(\partial^2 p(\nu)), \nu^2 \rangle$ with $p = q^{k+1}$, for all $\nu \in M_1(E)$. This concludes the induction step. It follows by induction that (2.1) holds for all monomials $\langle g, \nu \rangle^k$, and by linearity for all $p \in P^D$. Finally, the uniqueness assertion is immediate from the way B and Q were obtained above. This completes the proof of Theorem 2.2.

We now state a generalization of Theorem 2.2, where $M_1(E)$ is replaced by a general state space. We let E be a locally compact Polish space, $D \subseteq C_\Delta(E)$ be a dense linear subspace, and fix $\mathcal{S} \subseteq M(E)$.

Theorem A.1. *Let $L : P^D \rightarrow P$ be a linear operator. Then L is \mathcal{S} -polynomial and its carré-du-champs operator Γ is an $M_1(E)$ -derivation if and only if L admits a representation*

$$\begin{aligned} Lp(\nu) = & B_0(\partial p(\nu)) + \langle B_1(\partial p(\nu)), \nu \rangle \\ & + \frac{1}{2} \left(Q_0(\partial^2 p(\nu)) + \langle Q_1(\partial^2 p(\nu)), \nu \rangle + \langle Q_2(\partial^2 p(\nu)), \nu^2 \rangle \right), \quad \nu \in \mathcal{S} \end{aligned}$$

for some linear operators $B_0 : D \rightarrow \mathbb{R}$, $B_1 : D \rightarrow C_\Delta(E)$, $Q_0 : D \otimes D \rightarrow \mathbb{R}$, $Q_1 : D \otimes D \rightarrow C_\Delta(E)$, $Q_2 : D \otimes D \rightarrow \widehat{C}_\Delta(E^2)$. If \mathcal{S} contains an open subset of $M(E)$, these operators are uniquely determined by L .

Proof. The proof of this result follows the proof of Theorem 2.2. \square

B Proof of Theorem 3.10

Assume L satisfies (2.1) with B and Q as in (3.4), where ν_B is a (nonnegative, finite) kernel from E to E , and $\alpha : (E^\Delta)^2 \rightarrow \mathbb{R}$ is nonnegative, symmetric, bounded, and continuous on $(E^\Delta)^2 \setminus \{x = y\}$. Clearly Q is bounded with operator norm $2\|\alpha\|$. Identifying $C_\Delta(E)$ and $C(E^\Delta)$, we infer from Lemma III.B.1 that B is bounded, satisfies $B1 = 0$ as well as the positive maximum principle on E^Δ , and that $\{e^{tB}\}_{t \geq 0}$ is a strongly continuous contraction semigroup. By considering any sequence of functions $g_n \in C_0(E)$ with $0 \leq g_n(x) \uparrow 1$ for all $x \in E$, and using that $\nu_B(x, \{\Delta\}) = 0$ for all $x \in E$, one sees that B is E -conservative. Theorem 3.7 then yields that L is $M_1(E)$ -polynomial and its martingale problem has an $M_1(E)$ -solution with continuous paths for every initial condition $\nu \in M_1(E)$. Well-posedness follows by Theorem 3.9.

We now prove the opposite implication. Assume L is $M_1(E)$ -polynomial, its martingale problem is well-posed, and all $M_1(E)$ -solutions have continuous paths. Theorem 2.2 and Lemma III.3.11 imply that L satisfies (2.1), and then also the positive maximum principle on $M_1(E)$ due to Lemma III.3.5.

By Lemma C.1, the operator B satisfies the positive maximum principle on E and Lemma III.B.1 thus shows that B has the form in (3.4) for some (nonnegative, finite) kernel ν_B from E^Δ to E^Δ . Additionally, B is bounded, satisfies the positive maximum principle on E^Δ , and is the generator of the strongly continuous contraction semigroup $\{e^{tB}\}_{t \geq 0}$. We must prove that $\nu_B(x, \{\Delta\}) = 0$ for all $x \in E$; this will allow us to view ν_B as a kernel from E to E . Fix a sequence of functions $h_n \in C_0(E)$ with $0 \leq h_n \uparrow 1$. By Theorem 1.2.1 in Ethier and Kurtz (2005), the inverse $(\text{id} - B)^{-1}$ exists and is a bounded operator, whence there are functions $g_n \in C(E^\Delta)$ with $g_n - Bg_n = h_n$ and $\sup_n \|g_n\| < \infty$. Fix any $x \in E$ and let X be an $M_1(E)$ -solution to the martingale problem for L with initial condition δ_x . Then $\langle g_n, X_t \rangle - \int_0^t \langle Bg_n, X_s \rangle ds$ defines a martingale for each n . By Lemma 4.3.2 in Ethier and Kurtz (2005), so does $e^{-t} \langle g_n, X_t \rangle + \int_0^t e^{-s} \langle g_n - Bg_n, X_s \rangle ds$. Taking expectation and sending t to infinity yields

$$g_n(x) = \langle g_n, X_0 \rangle = \mathbb{E} \left[\int_0^\infty e^{-s} \langle g_n - Bg_n, X_s \rangle ds \right] = \mathbb{E} \left[\int_0^\infty e^{-s} \langle h_n, X_s \rangle ds \right].$$

Sending n to infinity and using that $X_t(E) = 1$ for all $t \geq 0$, we infer that $g_n(x) \rightarrow 1$ for all $x \in E$. After passing to subsequence we also have $g_n(\Delta) \rightarrow c$ for some $c \in \mathbb{R}$. We may therefore send n to infinity in the identity

$$\int_{E^\Delta} (g_n(\xi) - g_n(x)) \nu_B(x, d\xi) = Bg_n(x) = g_n(x) - h_n(x)$$

to get $(c - 1)\nu_B(x, \{\Delta\}) = 0$. On the other hand,

$$g_n(\Delta) = g_n(\Delta) - h_n(\Delta) = Bg_n(\Delta) = \int_E (g_n(\xi) - g_n(\Delta))\nu_B(\Delta, d\xi),$$

which in the limit yields $c = (1 - c)\nu_B(\Delta, E)$, forcing $c \neq 1$. It follows that $\nu_B(x, \{\Delta\}) = 0$ as required. This proves that B is of the stated form.

The form of Q will follow from Lemma C.2. To verify its hypotheses, note that by Lemma C.1 $\langle Q(g \otimes g), \nu^2 \rangle \geq 0$. Next, fix some $g \in D$ and $\nu \in M_1(E)$ such that $g = 0$ on the support of ν , and suppose that $\|g\| = 1$. For each $n \in \mathbb{N}$, define the polynomial

$$p_n(\mu) = \langle g, \mu \rangle^2 F_n(\langle |g|, \mu \rangle) - \frac{1}{n} \langle |g|, \mu \rangle,$$

where F_n is as in Lemma 3.12. Since $D = C_\Delta(E)$, we have $p_n \in P^D$. Moreover, since $F_n(z)zn \leq 1$ for all $z \in [0, 1]$, we get

$$\langle g, \mu \rangle^2 F_n(\langle |g|, \mu \rangle) \leq \frac{1}{n} \langle |g|, \mu \rangle, \quad \mu \in M_1(E),$$

and therefore $p_n \leq 0$ on $M_1(E)$. Since $g = 0$ on the support of ν , $p_n(\nu) = 0$. Applying the positive maximum principle and using the form (2.1) of L , as well as $\langle g, \nu \rangle = \langle |g|, \nu \rangle = 0$ and $F_n(0) = 1$ we obtain

$$0 \geq Lp_n(\nu) = -\frac{1}{n} \langle B(|g|), \nu \rangle + \langle Q(g \otimes g), \nu^2 \rangle$$

for all n , whence $\langle Q(g \otimes g), \nu^2 \rangle \leq 0$. By scaling, this actually holds for any $g \in D$ and $\nu \in M_1(E)$ such that $g = 0$ on the support of ν . If g equals some other constant $c \in \mathbb{R}$ on the support of ν , we still get

$$\langle Q(g \otimes g), \nu^2 \rangle = \langle Q((g - c) \otimes (g - c)), \nu^2 \rangle \leq 0$$

using that $Q(g \otimes 1) = 0$ by Lemma C.1. Thus Lemma C.2(ii) holds, and we conclude that $Q = \alpha\Psi$ for some nonnegative symmetric function $\alpha: E^2 \rightarrow \mathbb{R}$. It remains to use that $\alpha\Psi(g) \in \widehat{C}_\Delta(E^2)$ to show that this function can be extended to a bounded continuous function on $(E^\Delta)^2 \setminus \{x = y\}$.

Continuity is clear. For proving boundedness, choose a sequence of pairs $(x_n, y_n) \in (E^\Delta)^2 \setminus \{x = y\}$ such that $\alpha(x_n, y_n) \xrightarrow{n \rightarrow \infty} \infty$. Since we can assume without loss of generalities that $\alpha(x_i, y_i) > 0$, $x_i \neq x_j$, $x_i \neq y_j$, and $y_i \neq y_j$ for all $i, j \in \mathbb{N}$, we can construct $g \in C_\Delta(E)$ such that

$$(g(x_n) - g(y_n))^4 = \alpha(x_n, y_n)^{-1}.$$

This yields $\alpha(x_n, y_n)\Psi(g \otimes g)(x_n, y_n) = \alpha(x_n, y_n)^{1/2}$ proving that $\alpha\Psi(g \otimes g)$ is unbounded and providing the necessary contradiction. \square

C Auxiliary Lemmas

Let E be a locally compact Polish space.

Lemma C.1. *Let $D \subseteq C_\Delta(E)$ be a dense linear subspace containing the constant function 1, and let $L: P^D \rightarrow P$ be a linear operator satisfying (2.1) and the positive maximum principle on $M_1(E)$. Then B satisfies the positive maximum principle on E , $B1 = 0$, $\langle Q(g \otimes g), \nu^2 \rangle \geq 0$, and $Q(g \otimes 1) = 0$ for all $g \in D$ and $\nu \in M_1(E)$.*

Proof. Note that $L1 = 0$ since L satisfies (2.1). For any $g \in D$ and $x \in E$ such that $g(x) = \max_E g \geq 0$, the polynomial $p(\nu) = \langle g, \nu \rangle$ lies in P^D and satisfies $p(\delta_x) = \max_{M_1(E)} p \geq 0$. Thus $Bg(x) = Lp(\delta_x) \leq 0$. Furthermore, taking $p(\nu) = \langle 1, \nu \rangle \equiv 1$ we get $B1(x) = Lp(\delta_x) = 0$ for all $x \in E$. Fix then g and ν as in the lemma and define $p \in P^D$ by $p(\mu) = -(\langle g, \nu \rangle - \langle g, \mu \rangle)^2$. Then $p \leq 0$, $p(\nu) = 0$, $\partial p(\nu) = 0$, and $\partial^2 p(\nu) = -2g \otimes g$, so the positive maximum principle yields $-\langle Q(g \otimes g), \nu^2 \rangle = Lp(\nu) \leq 0$. Furthermore, taking $p(\nu) = \langle g \otimes 1, \nu \rangle - \langle g, \nu \rangle \equiv 0$ we get $0 = \langle g, \nu \rangle \langle B1, \nu \rangle + \frac{1}{2} \langle Q(g \otimes 1), \nu^2 \rangle = \frac{1}{2} \langle Q(g \otimes 1), \nu^2 \rangle$ for all $\nu \in M_1(E)$, proving the claim. \square

Lemma C.2. *Let $D \subseteq C_\Delta(E)$ be a dense linear subspace containing the constant function 1, and let $Q: D \otimes D \rightarrow \widehat{C}_\Delta(E^2)$ be a linear operator. The following conditions are equivalent:*

- (i) $Q(g \otimes g)(x, y) \geq 0$ for all $g \in D$ and $x, y \in E$, with equality if $g(x) = g(y)$.
- (ii) $\langle Q(g \otimes g), \nu^2 \rangle \geq 0$ for all $g \in D$ and $\nu \in M_1(E)$, with equality if g is constant on the support of ν .

If either condition is satisfied, then Q is of the form $Q = \alpha\Psi$ for some nonnegative symmetric function $\alpha: E^2 \rightarrow \mathbb{R}$.

Proof. It is clear that (i) implies (ii). For the converse, first note that for any $x \in E$ and $g \in D$, trivially g is constant on the support of δ_x . Thus $Q(g \otimes g)(x, x) = \langle Q(g \otimes g), \delta_x^2 \rangle = 0$. Taking $\nu = \frac{1}{2}(\delta_x + \delta_y)$ for any $x, y \in E$ then yields $Q(g \otimes g)(x, y) = \langle Q(g \otimes g), \nu^2 \rangle \geq 0$, with equality if $g(x) = g(y)$ since g is then constant on the support of ν . This proves that (ii) implies (i).

It remains to obtain the stated form of Q under the assumption that (i) holds. If E is a singleton then $Q = 0$, so we may assume that E contains at least two points. Fix $x, y \in E$ with $x \neq y$. Due to (i), the map $(g, h) \mapsto Q(g \otimes h)(x, y)$ is bilinear and positive semidefinite, and therefore satisfies the Cauchy–Schwarz inequality

$$|Q(g \otimes h)(x, y)| \leq \sqrt{Q(g \otimes g)(x, y)} \sqrt{Q(h \otimes h)(x, y)}.$$

Along with (i) this implies that $Q(g \otimes h)(x, y)$ depends on g and h only through their values at x and y . Moreover, since D is dense in $C_\Delta(E)$, for every $a \in \mathbb{R}^2$

there exists $g \in D$ such that $a = (g(x), g(y))$. Thus there is a unique map $T: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ such that

$$Q(g \otimes h)(x, y) = T(a, b) \quad \text{where} \quad a = \begin{pmatrix} g(x) \\ g(y) \end{pmatrix}, \quad b = \begin{pmatrix} h(x) \\ h(y) \end{pmatrix}.$$

The map T inherits bilinearity and positive semidefiniteness from Q . Since $Q(g \otimes 1)(x, y) = 0$ due to the Cauchy–Schwarz inequality and (i), we also have $T(a, b) = 0$ for $b = (1, 1)$. This implies that $T(a, b) = \frac{1}{2}\alpha(x, y)(a_1 - a_2)(b_1 - b_2)$ for some $\alpha(x, y) \in \mathbb{R}_+$. Thus,

$$Q(g \otimes h)(x, y) = \frac{1}{2}\alpha(x, y)(g(x) - g(y))(h(x) - h(y)) = \alpha(x, y)\Psi(g \otimes h)(x, y).$$

Defining $\alpha(x, x)$ arbitrarily, we obtain the map $\alpha: E^2 \rightarrow \mathbb{R}$, which is symmetric due to the symmetry of $Q(g \otimes h)$. □

For the next lemma we consider the setting of Lemma 4.5, i.e. $E = \mathbb{R}$ and

$$\mathbb{R} + C_c^\infty(\mathbb{R}) \subseteq D \subseteq \mathbb{R} + C_c^2(\mathbb{R}).$$

Lemma C.3. *Consider two operators $B: D \rightarrow C_\Delta(\mathbb{R})$ and $Q: D \otimes D \rightarrow \widehat{C}_\Delta(\mathbb{R}^2)$ such that B is as in (4.3) and Q satisfies*

$$Q(h \otimes h)(x, y) \geq 0 \text{ for all } h \in D, \\ \text{with equality if } h(x) = h(y) \text{ and } h'(x) = h'(y) = 0.$$

Then, for each $g \in D$ and $x, y \in \mathbb{R}$ such that $g(x) = g(y) = 1$ there exists a sequence $(p_n)_{n \in \mathbb{N}} \subseteq P^D$ such that

$$p_n(\nu_\lambda) = \max_{M_1(\mathbb{R})} p_n, \quad \partial p_n(\nu_\lambda) = f_n, \quad \text{and} \quad \langle Q(\partial^2 p_n(\nu_\lambda)), \nu_\lambda^2 \rangle = \langle Q(g \otimes g), \nu_\lambda^2 \rangle$$

for all $n \in \mathbb{N}$ and $\lambda \in [0, 1]$, where $\nu_\lambda = \lambda\delta_x + (1 - \lambda)\delta_y$ and $(f_n)_{n \in \mathbb{N}}$ satisfies

$$\lim_{n \rightarrow \infty} -2Bf_n(z) = (a(z)g'(z))^2,$$

for all $z \in \{x, y\}$.

Proof. Fix $g \in D$ such that $g(x) = g(y) = 1$. Let $F_n: [0, 1] \rightarrow \mathbb{R}$ as in Lemma 3.12 and fix a compactly supported function $\rho \in C_c^\infty(\mathbb{R})$ such that $\rho = 1$ on some neighborhoods of x and y and $\rho(\mathbb{R}) \subseteq [0, 1]$. Set then

$$\begin{aligned} \bar{g}_n(z) &= 1 + g'(x)(z - x) \frac{(z - y)^2}{(x - y)^2} F_{n^4} \left(\frac{|z - x|^2}{C_x} \right) \\ &\quad + g'(y)(z - y) \frac{(z - x)^2}{(x - y)^2} F_{n^4} \left(\frac{|z - y|^2}{C_y} \right), \end{aligned}$$

where $C_x = 2 \sup_{z \in \text{supp}(\rho)} (z - x)^2$. Setting $g_n = 1 + (\bar{g}_n - 1)\rho$ we get $g_n \in \mathbb{R} + C_c^\infty(\mathbb{R}) \subseteq D$. For n even, define now the polynomial

$$p_n(\nu) = \frac{1}{n(n-1)} (\langle g_n, \nu \rangle^n - \langle g_n^n, \nu \rangle).$$

Since $p_n(\nu_\lambda) = 0$ and by Jensen inequality $p_n \leq 0$, we can conclude that ν_λ maximizes p_n for all n even and $\lambda \in [0, 1]$. Observe that

$$\partial p_n(\nu_\lambda) = \frac{1}{n-1} \left(g_n - \frac{1}{n} g_n^n \right) =: f_n \quad \text{and} \quad \partial^2 p_n(\nu_\lambda) = g_n \otimes g_n.$$

Proceeding as in the proof of Lemma C.2, we can use the assumptions on Q to prove that $Q(g \otimes h)(x, y)$ depends on g and h only through their values and the values of their derivatives at x and y . Since $g_n(z) = g(z) = 1$ and $g'_n(z) = g'(z)$ for all n even and $z \in \{x, y\}$, this implies that $\langle Q(g_n \otimes g_n), \nu_\lambda^2 \rangle = \langle Q(g \otimes g), \nu_\lambda^2 \rangle$. Finally, the representation of B given by (4.3) yields

$$-2Bf_n(z) = (a(z)g'(z))^2 - 2 \int \frac{1}{n-1} \left(g_n(z+\xi) - \frac{1}{n} g_n(z+\xi)^n \right) - \frac{1}{n} F(z, d\xi),$$

for all $z \in \{x, y\}$. Since by the dominated convergence theorem the integral term converges to 0 for n going to ∞ , this concludes the proof. \square

Bibliography

- D. Akerer and D. Filipović. Linear Credit Risk Models. *ArXiv e-prints*, 2016. URL <https://arxiv.org/abs/1605.07419>.
- D. Akerer, D. Filipović, and S. Pulido. The Jacobi Stochastic Volatility Model. *Finance and Stochastics*, Forthcoming, 2016.
- D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*, pages 177–206. Springer, 1985.
- M. Beiglböck, A. M. G. Cox, M. Huesmann, and S. Källblad. Measure-valued martingales and optimality of Bass-type solutions to the Skorokhod Embedding Problem. *ArXiv e-prints*, 2017. URL <https://arxiv.org/abs/1708.07071>.
- D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1995.
- S. Biagini and Y. Zhang. Polynomial diffusion models for life insurance liabilities. *Insurance: Mathematics and Economics*, 71:114 – 129, 2016.
- B. Böttcher, R. Schilling, and J. Wang. Lévy matters. III. *Lévy-type Processes: Construction, Approximation and Sample Path Properties.*, 2013.
- P. Cardaliaguet. Notes from P. L. Lions’ lectures at the Collège de France. Technical report, 2012. URL <https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf>.
- R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2017.
- P. Courrège. Sur la forme intégrô-différentielle des opérateurs de C_k^∞ dans C satisfaisant au principe du maximum. *Séminaire Brelot-Choquet-Deny (Théorie du Potentiel)*, 10(2), 1965.
- C. Cuchiero. Polynomial processes in stochastic portfolio theory. *ArXiv e-prints*, 2017. URL <https://arxiv.org/abs/1705.03647>.
- C. Cuchiero, M. Keller-Ressel, and J. Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16:711–740, 2012.

- C. Cuchiero, K. Gellert, M. Guiricich, A. Platts, S. Sookdeo, and J. Teichmann. Polynomial models for market weights at work in stochastic portfolio theory: theory, tractable estimation, calibration and implementation. *Working paper*, 2016.
- C. Cuchiero, M. Larsson, and S. Svaluto-Ferro. Polynomial jump-diffusions on the unit simplex. *Annals of Applied Probability*, 28(4):2451–2500, 2018a.
- C. Cuchiero, M. Larsson, and S. Svaluto-Ferro. Probability measure-valued polynomial diffusions. *ArXiv e-prints*, 2018b. URL <https://arxiv.org/abs/1807.03229>.
- G. Da Prato and H. Frankowska. Invariance of stochastic control systems with deterministic arguments. *Journal of Differential Equations*, 200(1):18 – 52, 2004.
- D. Dawson. The critical measure diffusion process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 40(2):125–145, 1977.
- D. Dawson. Geostochastic calculus. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 6(2):143–168, 1978.
- D. Dawson. Measure-valued Markov processes. In *École d’Été de Probabilités de Saint-Flour XXI–1991*, Lecture Notes in Mathematics, pages 1–260. Springer, 1993.
- D. Dawson and K. Hochberg. Wandering Random Measures in the Fleming-Viot Model. *Annals of Probability*, 10(3):554–580, 1982.
- D. Dawson and Z. Li. Skew convolution semigroups and affine Markov processes. *Annals of Probability*, 34(3):1103–1142, 2006.
- F. Delbaen and H. Shirakawa. An interest rate model with upper and lower bounds. *Asia-Pacific Financial Markets*, 9:191–209, 2002.
- D. Duffie. *Credit risk modeling with affine processes*. Publications of the Scuola Normale Superiore. Scuola Normale Superiore, 2004.
- C. L. Epstein and R. Mazzeo. *Degenerate diffusion operators arising in population biology*. Number 185 in Annals of Mathematics Studies. Princeton University Press, 2013.
- A. Etheridge. *Some Mathematical Models from Population Genetics: École D’Été de Probabilités de Saint-Flour XXXIX-2009*. Lecture Notes in Mathematics. Springer, 2011.
- S. N. Ethier and T. G. Kurtz. *The Infinitely-Many-Alleles Model with Selection as a Measure-Valued Diffusion*, pages 72–86. Springer Berlin Heidelberg, Berlin, Heidelberg, 1987.

-
- S. N. Ethier and T. G. Kurtz. Fleming–viot processes in population genetics. *SIAM Journal on Control and Optimization*, 31(2):345–386, 1993.
- S. N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 2 edition, 2005.
- R. Fernholz. *Stochastic Portfolio Theory*. Applications of Mathematics. Springer-Verlag, New York, 2002.
- R. Fernholz and I. Karatzas. Relative arbitrage in volatility-stabilized markets. *Annals of Finance*, 1(2):149–177, 2005.
- R. Fernholz and I. Karatzas. Stochastic portfolio theory: an overview. *Handbook of numerical analysis*, 15:89–167, 2009.
- D. Filipović and M. Larsson. Polynomial diffusions and applications in finance. *Finance and Stochastics*, 20(4):931–972, 2016.
- D. Filipović and M. Larsson. Polynomial Jump-Diffusion Models. *ArXiv e-prints*, 2017. URL <https://arxiv.org/abs/1711.08043>.
- D. Filipović and A. B. Trolle. The term structure of interbank risk. *Journal of Financial Economics*, 109(3):707 – 733, 2013.
- D. Filipović and S. Willems. A term-structure model for dividends and interest rates. *ArXiv e-prints*, 2017. URL <https://arxiv.org/abs/1803.02249>.
- D. Filipović, E. Gourier, and L. Mancini. Quadratic variance swap models. *Journal of Financial Economics*, 119(1):44–68, 2016.
- D. Filipović, M. Larsson, and A. Trolle. Linear-rational term structure models. *The Journal of Finance*, 72(2):655–704, 2017.
- D. Filipović, M. Larsson, and T. Ware. Polynomial processes for power prices. *ArXiv e-prints*, 2017. URL <https://arxiv.org/abs/1710.10293>.
- W. H. Fleming and M. Viot. Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.*, 28(5):817–843, 1979.
- L. Gallardo and M. Yor. *Some remarkable properties of the Dunkl martingales*, pages 337–356. Springer, 2006.
- C. Gourieroux and J. Jasiak. Multivariate Jacobi process with application to smooth transitions. *Journal of Econometrics*, 131(1):475–505, 2006.
- E. K. Haviland. On the momentum problem for distribution functions in more than one dimension. *American Journal of Mathematics*, 57(3):562–568, 1935.
- E. K. Haviland. On the Momentum Problem for Distribution Functions in More Than One Dimension. II. *American Journal of Mathematics*, 58(1):164–168, 1936.

- P. Henry-Labordère, X. Tan, and N. Touzi. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Processes and their Applications*, 124(2):1112–1140, 2014.
- W. Hoh. Pseudo differential operators generating Markov processes. *Habilitationschrift, Universität Bielefeld*, 1998.
- J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, second edition, 2003.
- R. Jarrow and S. Turnbull. A unified approach for pricing contingent claims on multiple term structures. *Review of Quantitative Finance and Accounting*, 10(1):5–19, 1998.
- M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer Finance. Springer London, 2009.
- M. Kimura. Diffusion models in population genetics. *Journal of Applied Probability*, 1(2):177–232, 1964.
- A. Klenke. *Probability Theory: A Comprehensive Course*. Universitext. Springer London, 2 edition, 2013.
- N. Konno and T. Shiga. Stochastic partial differential equations for some measure-valued diffusions. *Probability Theory and Related Fields*, 79(2):201–225, 1988.
- P. Kotelenetz and T. Kurtz. Macroscopic limits for stochastic partial differential equations of McKean–Vlasov type. *Probability Theory and Related Fields*, 146(1):189, 2008.
- T. Krabichler and J. Teichmann. The Jarrow & Turnbull Setting Revisited. *Working paper*, 2017.
- P. Krühner and M. Larsson. Affine processes with compact state space. *Electronic Journal of Probability*, 23, 2018.
- T. Kurtz and J. Xiong. Particle representations for a class of nonlinear SPDEs. *Stochastic Processes and their Applications*, 83, 1999.
- K. Larsen and M. Sørensen. Diffusion models for exchange rates in a target zone. *Mathematical Finance*, 17(2):285–306, 2007. ISSN 1467-9965.
- M. Larsson and S. Svaluto-Ferro. An alternative approach to probability measure-valued jump-diffusions. *Working paper*, 2018.
- T. Liggett. *Interacting particle systems*, volume 276. Springer Science & Business Media, 2012.
- P. L. Lions. *Théorie des jeux à champs moyen et applications*. Lectures at the Collège de France, 2006–2007. 2007-2008. URL [//www.college-de-france.fr/site/pierre-louis-lions/course-2006-2007.htm](http://www.college-de-france.fr/site/pierre-louis-lions/course-2006-2007.htm).

-
- M. Marshall. *Positive Polynomials and Sums of Squares*. Mathematical surveys and monographs. American Mathematical Society, 2008.
- H. McKean. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57, 1967.
- E. Perkins. Dawson-Watanabe Superprocesses and Measure-valued Diffusions. In *École d'Été de Probabilités de Saint-Flour XXIX–1999*, Lecture Notes in Mathematics, pages 125–329. Springer, 2002.
- D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. A series of comprehensive studies in mathematics. Springer, 1999.
- M. Shkolnikov. Large volatility-stabilized markets. *Stochastic Processes and their Applications*, 123(1):212 – 228, 2013.
- A. S. Sznitman. Topics in Propagation of Chaos. In *École d'Été de Probabilités de Saint-Flour XIX–1989*, Lecture Notes in Mathematics, pages 165–251. Springer, 1991.
- S. Watanabe. A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.*, 8(1):141–167, 1968.
- H. Zheng. Interaction of credit and liquidity risks: Modelling and valuation. *Journal of Banking & Finance*, 30:391–407, 2006.