Received: 16 October 2019 | Accepted: 12 September 2020

DOI: 10.1111/mafi.12289

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics WILEY

ORIGINAL ARTICLE

Markov chains under nonlinear expectation

Max Nendel

Center for Mathematical Economics, Bielefeld University, Bielefeld, Germany

Correspondence
Max Nendel, Center for Mathematical Economics, Bielefeld University, 33615 Bielefeld, Germany.
Email: Max.Nendel@uni-bielefeld.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: CRC 1283

Abstract

In this paper, we consider continuous-time Markov chains with a finite state space under nonlinear expectations. We define so-called *Q-operators* as an extension of Q-matrices or rate matrices to a nonlinear setup, where the nonlinearity is due to model uncertainty. The main result gives a full characterization of convex *Q-operators* in terms of a positive maximum principle, a dual representation by means of Q-matrices, time-homogeneous Markov chains under convex expectations, and a class of nonlinear ordinary differential equations. This extends a classical characterization of generators of Markov chains to the case of model uncertainty in the generator. We further derive an explicit primal and dual representation of convex semigroups arising from Markov chains under convex expectations via the Fenchel–Legendre transformation of the generator. We illustrate the results with several numerical examples, where we compute price bounds for European contingent claims under model uncertainty in terms of the rate matrix.

KEYWORDS

generator of nonlinear semigroup, imprecise Markov chain, model uncertainty, nonlinear expectation, nonlinear ODE

1 | INTRODUCTION AND MAIN RESULT

In mathematical finance, model uncertainty or ambiguity is an almost omnipresent phenomenon, which, for example, appears due to incomplete information about certain aspects of an underlying asset or insufficient data in order to perform reliable statistical estimation methods for the parameters of a stochastic process. The latter typically leads to so-called parameter uncertainty in the generator of a stochastic process. Prominent examples for this type of uncertainty include a Black–Scholes model with uncertain volatility, the so-called uncertain volatility model, cf. Avellaneda, Levy, and Parás (1995), Avellaneda and Parás (1996), and Vorbrink (2014), and a Brownian motion under drift or volatility uncertainty leading to the g-framework, see, for example, Coquet, Hu, Mémin, and Peng (2002) or the G-framework by Peng (2007) and Peng (2008), respectively. Lately, these approaches have been generalized to Lévy processes with uncertainty in the Lévy triplet, cf. Denk, Kupper, and Nendel (2020), Hu and Peng (2009), and Neufeld and Nutz (2017), and uncertainty in the generator of Feller processes, cf. Nendel and Röckner (2019). While these works give sufficient conditions in order to guarantee the existence of stochastic processes under model uncertainty and to establish a connection to nonlinear partial differential equations, there is no necessary condition that determines the maximal degree of ambiguity that can be captured by an uncertain process.

In the present paper, we address this issue in a simplified setup, where we consider a finite state space. We provide sufficient and necessary conditions in terms of the generators of time-homogeneous continuous-time Markov chains that guarantee the existence of a continuous-time Markov chain under a convex expectation. We further establish a one-to-one relation between the transition operators of convex Markov chains and a class of nonlinear ordinary differential equations. In particular, we extend a classical relation between Markov chains, rate matrices, and ordinary differential equations to the case of model uncertainty. The ordinary differential equation related to a convex Markov chain is a spatially discretized version of a Hamilton–Jacobi–Bellman equation, and the nonlinear transition operators are related, via a dual representation, to a control problem where, roughly speaking, "nature" tries to control the system into the worst possible scenario (see Remark 4.18). The explicit description of the transition operators gives rise to a numerical scheme, different from Runge–Kutta methods, for the computation of price bounds for European contingent claims under model uncertainty. We illustrate this method and other numerical methods in several examples, where we consider an underlying Markov chain, which is a discrete version, more precisely, the generator is a finite difference discretization of the generator of a Brownian motion with uncertain drift, cf. Coquet et al. (2002), and uncertain volatility, cf. Peng (2007) and Peng (2008). The main tools, we use in our analysis, are convex duality, a semigroup-theoretic approach to control problems due to Nisio (1976/77), see also Denk et al. (2020) and Nendel and Röckner (2019), and a convex version of Kolmogorov's extension theorem due to Denk, Kupper, and Nendel (2018), which allows to extend the expectation to functionals that depend on the whole path. Restricting the time parameter, in the present work, to the set of natural numbers leads to a discrete-time Markov chain, in the sense of Denk et al. (2018, Example 5.3).

The concept we use to describe ambiguity is the notion of a nonlinear expectation introduced by Peng (2005). Nonlinear expectations closely relate to other concepts describing model uncertainty, backward stochastic differential equations (BSDEs), cf. Cohen (2012), and Coquet et al. (2002), and 2BSDEs, cf. Cheridito, Soner, Touzi, and Victoir (2007) and Denis, Hu, and Peng (2011). We refer to Pardoux and Peng (1992), Pardoux and Peng (1990), and El Karoui, Peng, and Quenez (1997) for a detailed study of BSDEs and their applications within the field of mathematical finance. If a nonlinear expectation $\mathcal{E}$ is sublinear, then $\rho(X) := \mathcal{E}(-X)$ defines a coherent monetary risk measure as introduced by Artzner, Delbaen, Eber, and Heath (1999), Delbaen (2000), and Delbaen (2002), see also Föllmer and Schied (2011) for an overview of monetary risk measures. Moreover, if $\mathcal{E}$ is a sublinear expectation, then $\mathcal{E}$ is a coherent upper prevision, cf. Walley (1991), and vice versa. There is a similar one-to-one relation between convex expectations, convex upper previsions, cf. Pelessoni and Vicig (2003) and Pelessoni and Vicig (2005), and convex risk measures, cf. Föllmer and Schied (2002) and Frittelli and Rosazza Gianin (2002). Further concepts, which are closely related to nonlinear expectations and describe model uncertainty, are Choquet capacities (see, e.g., Dellacherie & Meyer, 1978), game-theoretic probability by Vovk and Shafer (2014), and niveloids, see, for example, Cerreia-Vioglio, Maccheroni, Marinacci, and Rustichini (2014).

Our setup is inspired by Peng (2005), where Markov chains under nonlinear expectations are considered in an axiomatic way. However, the existence of stochastic processes under nonlinear expectations has only been considered in terms of finite-dimensional nonlinear marginal distributions, whereas completely path-dependent functionals could not be regarded. Markov chains under model uncertainty have been considered among others by Avellaneda and Buff (1999), De Cooman, Hermans, and Quaeghebeur (2009), Hartfiel (1998), and Škulj (2009). Avellaneda and Buff (1999) study a finite difference discretization of the uncertain volatility model leading to a Markov chain setting. Hartfiel (1998) considers so-called Markov set-chains in discrete time, using matrix intervals in order to describe model uncertainty in the transition matrices. Later, Škulj (2009) approached Markov chains under model uncertainty using Choquet capacities, which results in higher dimensional matrices on the power set, while De Cooman et al. (2009) considered imprecise Markov chains using an operator-theoretic approach with upper and lower expectations. In Denk et al. (2018, Example 5.3), Denk et al. describe model uncertainty in the transition matrix via a nonlinear transition operator, which, together with the results obtained in Denk et al. (2018), allows the construction of discrete-time Markov chains on the canonical path space. In continuous time, in particular, computational aspects of sublinear imprecise Markov chains have been studied amongst others by Krak, De Bock, and Siebes (2017) and Škulj (2015).

Another concept that is closely related to Markov chains under nonlinear expectations, as discussed in the present paper, are BSDEs on Markov chains by Cohen and Elliott (2008) and Cohen and Elliott (2010a), see also Cohen and Szpruch (2012), Cohen and Hu (2013), and Cohen and Elliott (2010b) for the discrete-time case. Here, a reference Markov chain $X = (X_t)_{t\geq 0}$ with generator $(q_t)_{t\geq 0}$ is fixed, and one considers BSDEs driven by X. This can be viewed as a discretization of the classical BSDE setup, where the state space is $\mathbb{R}$, the driving process is a Brownian Motion, and the generator is $\frac{1}{2}\partial_{xx}$. Cohen and Szpruch (2012) show that Markovian solutions to BSDEs on Markov chains are related via their driver to a system

$$u'(t) = f(t, u(t)) + A(t)u(t) \quad \text{for all } t \geq 0, \quad u(0) = u_0$$

of nonlinear ordinary differential equations with a nonlinear function f that is assumed to be globally Lipschitz in the variable u. In the present paper, $f(t, u) = Qu$ for a convex operator Q. The biggest difference between our approach and the theory of BSDEs on Markov chains lies in the fact that we do not consider a fixed reference Markov chain that drives the model. On the other hand, our approach is restricted to considering Markovian solutions to BSDEs on Markov chains. From a technical standpoint, further differences are that the theory of BSDEs allows for more generality in terms of nonlinearity of the driver, while we do not require global Lipschitz continuity of the generator allowing for a possibly unbounded convex conjugate. Additionally, we only focus on the time-homogeneous case. However, regarding the existence of Markov chains under convex expectations and their connection to nonlinear ordinary differential equations (ODEs), this restriction could easily be overcome with a slight modification of the construction of the transition operators.

Dentcheva and Ruszczyński (2018) consider Markov risk measures for a countable state space, see also Fan and Ruszczyński (2018a), Fan and Ruszczyński (2018b), and Ruszczyński (2010) for the discrete-time case. Here, the focus lies on time-consistent risk measurement related to a fixed reference continuous-time Markov chain $X = (X_t)_{t\geq 0}$. Using so-called semiderivatives in the direction of the generator A, the authors derive, in the case of a coherent risk measure, a sublinear ordinary differential equation related to the risk measure, where the dual representation of the nonlinear generator depends on the generator A of the baseline model X. Clearly, in the theory of Markov risk measures, the focus lies more on law-invariant risk measures such as the average value at risk, and is therefore not directly comparable with our approach, where we explicitly avoid to fix a baseline model but rather try to capture very general forms of uncertainty in the generator. However, on a technical level, our approach also allows to consider risk evaluations related to convex generators that do not depend on a fixed reference generator.

In view of the aforementioned existing literature on imprecise versions of Markov chains, the contribution of this paper can be summarized as follows (see Remark 2.6 for further details):

- We propose a framework describing Markov chains under model uncertainty in terms of the rate matrix. Our approach complements the existing literature on BSDEs on Markov chains and Markov risk measures covering a different range of examples and applications in a consistent way. The key difference between our framework and the aforementioned existing approaches lies in the fact that we do not consider a fixed reference Markov chain describing the dynamics of an underlying asset. Moreover, our approach relies on analytic rather than stochastic methods using distributional rather than pathwise properties, and thus leading to restrictions in certain directions but advantages in other directions.
- We show that, as in the linear case, Markov chains under convex expectations with certain regularity at time 0 are linked via a one-to-one relation to certain convex functions (their generator) and to solutions to convex differential equations, which can be solved, for example, by using an explicit Euler method or any other Runge–Kutta method. In particular, we prove the global existence of solutions to a class of convex differential equations with unbounded convex conjugate, that is, without a global Lipschitz condition on the generator.
- We show that the transition semigroup of a convex Markov chain can be explicitly constructed using any (!) dual representation of the generator. In particular, for numerical computations, a “minimal” dual representation in terms of certain “corner points” can be used to solve the nonlinear Kolmogorov equation. Based on the explicit construction of the semigroup, we propose a novel algorithm for the numerical computation of solutions to a class of nonlinear ODEs. Moreover, we show that every convex transition semigroup is the least upper bound (in the sense of semigroups) of a family of linear transition semigroups, and vice versa.
- The convex expectations we consider are defined on the whole path space without fixing any reference measure. We show that the nonlinear expectation, although possibly undominated, always admits a dual representation in terms of countably additive probability measures. Moreover, we derive an explicit dual representation in terms of an optimal control problem, where nature tries to control the system into the worst possible scenario, giving a control-theoretic interpretation to Markov chains under convex expectations.

1.1 | Structure of the paper

In Section 2, we fix the notation, introduce our setup and basic definitions, and state the main result (Theorem 2.5). In Section 3, we prove the first part of Theorem 2.5 (implications $(v) \Rightarrow (ii) \Rightarrow (i) \Rightarrow (iii)$). The main tool, we use in this part, is convex duality in $\mathbb{R}^d$. Moreover, we discuss how, in the sublinear case, computational efficiency can be improved by reducing compact and suitably convex sets of generator matrices to their "corner points." The effectiveness of this reduction is demonstrated in Section 5. In Section 4, we prove the remaining implications $(iii) \Rightarrow (iv) \Rightarrow (v)$ of Theorem 2.5. Here, we use a combination of so-called Nisio semigroups, as introduced in Nisio (1976/77), the theory of ordinary differential equations, and a Kolmogorov-type extension theorem for convex expectations derived in Denk et al. (2018). We conclude this section by showing that the semigroup envelope admits a dual representation as a cost functional related to an optimal control problem. In Section 5, we use and compare two different numerical methods, based on the results from Sections 3 and 4, in order to compute price bounds for European contingent claims, where the underlying is a discrete version of a Brownian motion with drift uncertainty (g-framework) and volatility uncertainty (G-framework).

2 | NOTATION, BASIC DEFINITIONS, AND MAIN RESULT

Given a measurable space $(\Omega, \mathcal{F})$, we denote the space of all bounded measurable functions $\Omega \to \mathbb{R}$ by $\mathcal{L}^\infty(\Omega, \mathcal{F})$. A *nonlinear expectation* is then a functional $\mathcal{E} : \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}$, which satisfies

- $\mathcal{E}(X) \le \mathcal{E}(Y)$ whenever $X(\omega) \le Y(\omega)$ for all $\omega \in \Omega$,
- $\mathcal{E}(\alpha 1_\Omega) = \alpha$ for all $\alpha \in \mathbb{R}$.

If $\mathcal{E}$ is additionally convex, that is, for all $X, Y \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ and $\theta \in [0,1]$,

$$\mathcal{E}(\theta X + (1-\theta)Y) \le \theta \mathcal{E}(X) + (1-\theta)\mathcal{E}(Y),$$

we say that $\mathcal{E}$ is a *convex expectation*. It is well known (see, e.g., Denk et al., 2018 or Föllmer & Schied, 2011) that every convex expectation $\mathcal{E}$ admits a dual representation in terms of finitely additive probability measures. If $\mathcal{E}$, however, even admits a dual representation in terms of (countably additive) probability measures, we say that $(\Omega, \mathcal{F}, \mathcal{E})$ is a *convex expectation* space. More precisely, we say that $(\Omega, \mathcal{F}, \mathcal{E})$ is a *convex expectation* space if there exists a set $\mathcal{P}$ of probability measures on $(\Omega, \mathcal{F})$ and a family $(\alpha_\mathbb{P})_{\mathbb{P} \in \mathcal{P}} \subset [0, \infty)$ with $\inf_{\mathbb{P} \in \mathcal{P}} \alpha_\mathbb{P} = 0$ such that

$$\mathcal{E}(X) = \sup_{\mathbb{P} \in \mathcal{P}} \left(\mathbb{E}_\mathbb{P}(X) - \alpha_\mathbb{P}\right)$$

for all $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$. Here, $\mathbb{E}_\mathbb{P}$ denotes the expectation w.r.t. a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$. If $\alpha_\mathbb{P} = 0$ for all $\mathbb{P} \in \mathcal{P}$, we say that $(\Omega, \mathcal{F}, \mathcal{E})$ is a *sublinear expectation space*. Here, the set $\mathcal{P}$ represents the set of all models that are relevant under the expectation $\mathcal{E}$. In the case of a sublinear expectation space, the functional $\mathcal{E}$ is the best case among all plausible models $\mathcal{P}$. In the case of a convex expectation space, the functional $\mathcal{E}$ is a weighted best case among all plausible models $\mathcal{P}$ with an additional penalization term $\alpha_\mathbb{P}$ for every $\mathbb{P} \in \mathcal{P}$. Intuitively, $\alpha_\mathbb{P}$ can be seen as a measure for how much importance we give to the prior $\mathbb{P} \in \mathcal{P}$ under the expectation $\mathcal{E}$. For example, a low penalization, that is, $\alpha_{\mathbb{P}}$ close or equal to 0, gives more importance to the model $\mathbb{P} \in \mathcal{P}$ than a high penalization.

Throughout, we consider a finite nonempty state space S with cardinality $d := |S| \in \mathbb{N}$. We endow S with the discrete topology 2^S and w.l.o.g. assume that $S = \{1, \dots, d\}$. The space of all bounded measurable functions $S \to \mathbb{R}$ can therefore be identified by $\mathbb{R}^d$ via

$$u = (u_1, \dots, u_d)^T \quad \text{with } u_i := u(i) \quad \text{for all } i \in \{1, \dots, d\}.$$

Therefore, we denote bounded measurable functions u as vectors of the form $u = (u_1, \dots, u_d)^T \in \mathbb{R}^d$, where u_i represents the value of u in the state $i \in \{1, \dots, d\}$. On $\mathbb{R}^d$, we consider the norm

$$\|u\|_\infty := \max_{i=1,\dots,d} |u_i| = \max_{i\in\{1,\dots,d\}} |u(i)|$$

for a vector $u \in \mathbb{R}^d$. Moreover, for $\alpha \in \mathbb{R}$, the vector $\alpha \in \mathbb{R}^d$ denotes the constant vector $u \in \mathbb{R}^d$ with $u_i = \alpha$ for all $i \in \{1, \dots, d\}$. For an arbitrary matrix $q = (q_{ij})_{1\le i,j\le d} \in \mathbb{R}^{d\times d}$, we denote by $\|q\|$ the operator norm of $q : \mathbb{R}^d \to \mathbb{R}^d$ w.r.t. the norm $\|\cdot\|_\infty$, that is,

$$\|q\| = \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\|qv\|_\infty}{\|v\|_\infty} = \max_{i=1,\dots,d} \left(\sum_{j=1}^{d} |q_{ij}| \right).$$

Inequalities of vectors are always understood componentwise, that is, for $u, v \in \mathbb{R}^d$,

$$u \le v \iff \forall i \in \{1, \dots, d\} : u_i \le v_i.$$

In the same way, all concepts in $\mathbb{R}^d$ that include inequalities are to be understood componentwise. For example, a vector field $F : \mathbb{R}^d \to \mathbb{R}^d$ is called *convex* if

$$F_i(\lambda u + (1-\lambda)v) \le \lambda F_i(u) + (1-\lambda)F_i(v)$$

for all $i \in \{1, \dots, d\}$, $u, v \in \mathbb{R}^d$ and $\lambda \in [0, 1]$. A vector field F is called *sublinear* if it is convex and positive homogeneous (of degree 1). Moreover, for a set $M \subset \mathbb{R}^d$ of vectors, we write $u = \sup M$ if $u_i = \sup_{v\in M} v_i$ for all $i \in \{1, \dots, d\}$ and $u = \max M$ if $u = \sup M$ and, for all $i \in \{1, \dots, d\}$, there exists some $v \in M$ with $u_i = v_i$.

In the following, we briefly recall the basic definitions and concepts from the theory of (time-homogeneous) Markov chains. A *(time-homogeneous) Markov chain* is a quadruple $(\Omega, \mathcal{F}, (\mathbb{P}_1, \dots, \mathbb{P}_d), (X_t)_{t\ge0})$, where:

(M1) $(\Omega, \mathcal{F})$ is a measurable space.

(M2) $X_t : \Omega \to \{1, \dots, d\}$ is $\mathcal{F}$-measurable for all $t \ge 0$.

(M3) $(\mathbb{P}_1, \dots, \mathbb{P}_d)$ is a collection of probability measures, where, for $i \in \{1, \dots, d\}$, $\mathbb{P}_i(X_0 = i) = 1$, that is, $\mathbb{P}_i$ denotes the probability distribution under which the Markov chain starts in the state i. Moreover, we use the notation

$$\mathbb{E}_i(Y) := \mathbb{E}_{\mathbb{P}_i}(Y) \quad \text{and} \quad \mathbb{E}(Y) := (\mathbb{E}_1(Y), \dots, \mathbb{E}_d(Y))^T$$

for $i \in \{1, \dots, d\}$ and all random variables $Y : \Omega \to \mathbb{R}$.

(M4) For all $s, t \geq 0$ and $i \in \{1, \dots, d\}$,

$$\mathbb{E}_i(u(X_{s+t})|\mathcal{F}_s) = \mathbb{E}_i(u(X_{t+s})|X_s) = \mathbb{E}_{X_s}(u(X_t)).$$

In particular, $\mathbb{E}_i(u(X_{t+s})|X_s = j) = \mathbb{E}_j(u(X_t)) \quad$ for all $i, j \in \{1, \dots, d\}$.

A matrix $q = (q_{ij})_{1\leq i,j\leq d} \in \mathbb{R}^{d\times d}$ is called a *Q-matrix* or *rate matrix* if it satisfies the following conditions:

(Q1) $q_{ii} \leq 0$ for all $i \in \{1, \dots, d\}$,
(Q2) $q_{ij} \geq 0$ for all $i, j \in \{1, \dots, d\}$ with $i \neq j$,
(Q3) $\sum_{j=1}^{d} q_{ij} = 0$ for all $i \in \{1, \dots, d\}$.

It is well known that every continuous-time Markov chain with certain regularity properties at time $t = 0$ can be related to a *Q-matrix* and vice versa. More precisely, for a matrix $q \in \mathbb{R}^{d\times d}$, the following statements are equivalent:

(i) q is a *Q-matrix*.
(ii) There is a Markov chain $(\Omega, \mathcal{F}, (\mathbb{P}_1, \dots, \mathbb{P}_d), (X_t)_{t\geq 0})$ such that

$$qu_0 = \lim_{h\searrow 0} \frac{\mathbb{E}(u_0(X_h)) - u_0}{h} \quad \text{for all } u_0 \in \mathbb{R}^d,$$

where $u_0(i)$ is the ith component of u_0 for $i \in \{1, \dots, d\}$.

In this case, for each vector $u_0 \in \mathbb{R}^d$, the function $u : [0, \infty) \to \mathbb{R}^d,\ t \mapsto \mathbb{E}(u_0(X_t))$ is the unique classical solution $u \in C^1([0, \infty); \mathbb{R}^d)$ to the initial value problem

$$u'(t) = qu(t), \quad t \geq 0,$$
$$u(0) = u_0,$$

that is, $u(t) = e^{tq}u_0$ for all $t \geq 0$, where e^{tq} is the matrix exponential of tq. We refer to Norris (1998) for a detailed illustration of this relation.

We say that a (possibly nonlinear) operator $Q : \mathbb{R}^d \to \mathbb{R}^d$ satisfies the *positive maximum principle* if, for every $u = (u_1, \dots, u_d)^T \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$,

$$(Qu)_i \leq 0 \quad \text{whenever } u_i \geq u_j \text{ for all } j \in \{1, \dots, d\}.$$

This notion is motivated by the positive maximum principle for generators of Feller processes, see, for example, Jacob (2001, Equation (0.8)). Notice that a matrix $q \in \mathbb{R}^{d\times d}$ is a *Q-matrix* if and only if it satisfies the positive maximum principle and $q1 = 0$, where $1 := (1, \dots, 1)^T \in \mathbb{R}^d$ denotes the constant 1 vector. In fact, Property (Q3) is just a reformulation of $q1 = 0$. Moreover, if q satisfies the positive maximum principle, then $q_{ii} = (qe_i)_i \leq 0$ for all $i \in \{1, \dots, d\}$ and $-q_{ij} = (q(-e_i))_j \leq 0$ for all $i, j \in \{1, \dots, d\}$ with $i \neq j$. That is, q fulfills (Q1) and (Q2). On the other hand, if q is a *Q-matrix*, $u = (u_1, \dots, u_d)^T \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$ with $u_i \geq u_j$ for all $j \in \{1, \dots, d\}$, then $(qu)_i = \sum_{j=1}^{d} q_{ij}u_j \leq u_i \sum_{j=1}^{d} q_{ij} = 0$, which shows that q satisfies the positive maximum principle.

To state the main result, we introduce the following definitions.

Definition 2.1. A (possibly nonlinear) map $Q : \mathbb{R}^d \to \mathbb{R}^d$ is called a *Q-operator* if the following conditions are satisfied:

(i) $(Q\lambda e_i)_i \leq 0$ for all $\lambda > 0$ and all $i \in \{1, \dots, d\}$,
(ii) $(Q(-\lambda e_j))_i \leq 0$ for all $\lambda > 0$ and all $i, j \in \{1, \dots, d\}$ with $i \neq j$,
(iii) $Q\alpha = 0$ for all $\alpha \in \mathbb{R}$, where we identify α with $(\alpha, \dots, \alpha)^T \in \mathbb{R}^d$.

Definition 2.2. A *convex Markov chain* is a quadruple $(\Omega, \mathcal{F}, \mathcal{E}, (X_t)_{t\geq 0})$ that satisfies the following conditions:

(i) $(\Omega, \mathcal{F})$ is a measurable space.
(ii) $X_t : \Omega \to \{1, \dots, d\}$ is $\mathcal{F}$-measurable for all $t \geq 0$.
(iii) $\mathcal{E} = (\mathcal{E}_1, \dots, \mathcal{E}_d)^T$, where $(\Omega, \mathcal{F}, \mathcal{E}_i)$ is a convex expectation space for all $i \in \{1, \dots, d\}$ and $\mathcal{E}(u_0(X_0)) = u_0$. Here and in the following, we use the notation

$$\mathcal{E}(Y) := (\mathcal{E}_1(Y), \dots, \mathcal{E}_d(Y))^T \in \mathbb{R}^d$$

for $Y \in \mathcal{L}^\infty(\Omega, \mathcal{F})$.
(iv) The following version of the Markov property is satisfied: For all $s, t \geq 0$, $n \in \mathbb{N}$, $0 \leq t_1 < \dots < t_n \leq s$, and $v_0 \in (\mathbb{R}^d)^{(n+1)}$,

$$\mathcal{E}(v_0(Y, X_{s+t})) = \mathcal{E}\left[\mathcal{E}_{X_s,t}(v_0(Y, \cdot\,))\right], \tag{1}$$

where $Y := (X_{t_1}, \dots, X_{t_n})$ and $\mathcal{E}_{i,t}(u_0) := \mathcal{E}_i(u_0(X_t))$ for all $u_0 \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$.

We say that the Markov chain $(\Omega, \mathcal{F}, \mathcal{E}, (X_t)_{t\geq 0})$ is *linear* or *sublinear* if the mapping $\mathcal{E} : \mathcal{L}^\infty(\Omega, \mathcal{F}) \to \mathbb{R}^d$ is, additionally, linear, or sublinear, respectively.

Notice that the properties (i)–(iii) in the previous definition are a one-to-one translation of (M1)–(M3) to a convex setup. The Markov property given in (iv) of the previous definition is the nonlinear analog of the classical Markov property (M4) without using conditional expectations. Due to the nonlinearity of the expectation, the definition and, in particular, the existence of a conditional (nonlinear) expectation are quite involved, which is why we avoid to introduce this concept. In order to get the idea behind the formulation in (iv), choose $v_0 = u(X_{s+t})1_B(Y)$ for a measurable function $u : \{1, \dots, d\} \to \mathbb{R}$ and arbitrary $B \subset \{1, \dots, d\}^n$. Then, if $\mathcal{E}$ is linear, Equation (1) reads as

$$\mathcal{E}(u(X_{s+t})1_B(Y)) = \mathcal{E}\left(\mathcal{E}_{X_s,t}(u)1_B(Y)\right),$$

which is equivalent to (M4). On the other hand, for every linear Markov chain, Property (M4) implies Property (iv). Hence, in the linear case, Definition 2.2 is consistent with the classical definition of a Markov chain.

In line with Denk et al. (2018, Definition 5.1), we say that a (possibly nonlinear) map $\mathcal{E} : \mathbb{R}^d \to \mathbb{R}^d$ is a *kernel*, if $\mathcal{E}$ is *monotone*, that is, $\mathcal{E}(u) \leq \mathcal{E}(v)$ for all $u, v \in \mathbb{R}^d$ with $u \leq v$, and $\mathcal{E}$ *preserves constants*, that is, $\mathcal{E}(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$.

Definition 2.3. A family $\mathscr{S} = (\mathscr{S}(t))_{t\geq 0}$ of (possibly nonlinear) operators $\mathscr{S}(t) : \mathbb{R}^d \to \mathbb{R}^d$ is called a *semigroup* if

(i) $\mathscr{S}(0) = I$, where $I = I_d$ is the d-dimensional identity matrix,
(ii) $\mathscr{S}(s+t) = \mathscr{S}(s)\mathscr{S}(t)$ for all $s, t \geq 0$.

Here and throughout, we make use of the notation $\mathscr{S}(s)\mathscr{S}(t) := \mathscr{S}(s) \circ \mathscr{S}(t)$. If, additionally, $\mathscr{S}(h) \to I$ uniformly on compact sets as $h \searrow 0$, we say that the semigroup $\mathscr{S}$ is *uniformly continuous*. We call $\mathscr{S}$ *Markovian* if $\mathscr{S}(t)$ is a *kernel* for all $t \geq 0$. We say that $\mathscr{S}$ is *linear*, *sublinear*, or *convex* if $\mathscr{S}(t)$ is linear, sublinear, or convex for all $t \geq 0$, respectively.

Definition 2.4. Let $\mathcal{P} \subset \mathbb{R}^{d\times d}$ be a set of Q-matrices and $f = (f_q)_{q\in\mathcal{P}}$ a family of vectors with $\sup_{q\in\mathcal{P}} f_q = f_{q_0} = 0$ for some $q_0 \in \mathcal{P}$, that is, $f_q \leq 0$ for all $q \in \mathcal{P}$ and there exists some $q_0 \in \mathcal{P}$ with $f_{q_0} = 0$. We denote by

$$S_q(t)u_0 := e^{qt}u_0 + \int_0^t e^{qs} f_q \,\mathrm{d}s = u_0 + \int_0^t e^{sq}\big(qu_0 + f_q\big)\,\mathrm{d}s$$

for $t \geq 0$, $u_0 \in \mathbb{R}^d$ and $q \in \mathcal{P}$. Then, $S_q = (S_q(t))_{t\geq 0}$ is an affine linear semigroup. We call a semigroup $\mathscr{S}$ the *(upper) semigroup envelope* (later also *Nisio semigroup*) of $(\mathcal{P}, f)$ if

(i) $\mathscr{S}(t)u_0 \geq S_q(t)u_0$ for all $t \geq 0$, $u_0 \in \mathbb{R}^d$ and $q \in \mathcal{P}$,
(ii) for any other semigroup $\mathscr{T}$ satisfying (i) we have that $\mathscr{S}(t)u_0 \leq \mathscr{T}(t)u_0$ for all $t \geq 0$ and $u_0 \in \mathbb{R}^d$.

That is, the semigroup envelope $\mathscr{S}$ is the smallest semigroup that dominates all semigroups $(S_q)_{q\in\mathcal{P}}$.

The following main theorem gives a full characterization of convex Q-*operator*s.

Theorem 2.5. *Let $Q : \mathbb{R}^d \to \mathbb{R}^d$ be a mapping. Then, the following statements are equivalent:*

(*i*) *Q is a convex* Q-operator.
(*ii*) *Q is convex, satisfies the positive maximum principle, and $Q\alpha = 0$ for all $\alpha \in \mathbb{R}$, where $\alpha := (\alpha, \dots, \alpha)^T \in \mathbb{R}^d$.*
(*iii*) *There exists a set $\mathcal{P} \subset \mathbb{R}^{d\times d}$ of* Q-*matrices and a family $f = (f_q)_{q\in\mathcal{P}} \subset \mathbb{R}^d$ of vectors with $f_q \leq 0$ for all $q \in \mathcal{P}$ and $f_{q_0} = 0$ for some $q_0 \in \mathcal{P}$, such that*

$$Qu_0 = \sup_{q\in\mathcal{P}} \big(qu_0 + f_q\big) \tag{2}$$

for all $u_0 \in \mathbb{R}^d$, where the supremum is to be understood componentwise.
(*iv*) *There exists a uniformly continuous convex Markovian semigroup $\mathscr{S}$ with*

$$Qu_0 = \lim_{h\searrow 0} \frac{\mathscr{S}(h)u_0 - u_0}{h}$$

for all $u_0 \in \mathbb{R}^d$.

(v) *There is a convex Markov chain* $(\Omega, \mathcal{F}, \mathcal{E}, (X_t)_{t\geq 0})$ *such that*

$$Qu_0 = \lim_{h \searrow 0} \frac{\mathcal{E}(u_0(X_h)) - u_0}{h}$$

for all $u_0 \in \mathbb{R}^d$.

In this case, for each initial value $u_0 \in \mathbb{R}^d$*, the function* $u : [0,\infty) \to \mathbb{R}^d$*,* $t \mapsto \mathcal{E}(u_0(X_t))$ *is the unique classical solution* $u \in C^1([0,\infty); \mathbb{R}^d)$ *to the initial value problem*

$$u'(t) = Qu(t) = \sup_{q\in\mathcal{P}} \big(qu(t) + f_q\big), \quad t \geq 0, \tag{3}$$

$$u(0) = u_0.$$

Moreover, the Markovian semigroup $\mathscr{S}$ *from (iv) is the (upper) semigroup envelope of* $(\mathcal{P}, f)$*, and* $u(t) = \mathscr{S}(t)u_0$ *for all* $t \geq 0$.

Remark 2.6. Consider the situation of Theorem 2.5.

(a) The dual representation in (iii) gives a model uncertainty interpretation to Q-*operator*s. The set $\mathcal{P}$ can be seen as the set of all plausible rate matrices, when considering the Q-*operator* Q. For every $q \in \mathcal{P}$, the vector $f_q \leq 0$ can be interpreted as a penalization, which measures how much importance we give to each rate matrix q. The requirement that there exists some $q_0 \in \mathcal{P}$ with $f_{q_0} = 0$ can be interpreted in the following way: There exists at least one rate matrix q_0 within the set of all plausible rate matrices $\mathcal{P}$ to which we assign the maximal importance, which is the minimal penalization.

(b) The semigroup envelope $\mathscr{S}$ of $(\mathcal{P}, f)$ can be constructed more explicitly, in particular, an explicit (in terms of $(\mathcal{P}, f)$) dual representation can be derived. For details, we refer to Section 4 (Definition 4.2 and Remark 4.18). Moreover, we would like to highlight that the semigroup envelope $\mathscr{S}$ can be constructed w.r.t. any dual representation $(\mathcal{P}, f)$ as in (iii) and results in the unique classical solution to (3) independent of the choice of the dual representation $(\mathcal{P}, f)$ of Q. This gives, in some cases, the opportunity to efficiently compute the semigroup envelope numerically via its primal/dual representation (see Remark 3.3 and Example 5.2).

(c) The same equivalence as in Theorem 2.5 holds if convexity is replaced by sublinearity in (i), (ii), (iv), and (v) and $f_q = 0$ for all $q \in \mathcal{P}$ in (iii). In this case, the set $\mathcal{P}$ in (iii) can be chosen to be compact as we will see in the proof of Theorem 2.5.

(d) Theorem 2.5 extends and includes the well-known relation between (linear) Markov chains, Q-matrices, and ordinary differential equations.

(e) A remarkable consequence of Theorem 2.5 is that every convex Markovian semigroup, which is differentiable at time $t = 0$, is the semigroup envelope with respect to the Fenchel–Legendre transformation (or any other dual representation as in (iii) of its generator, which is a convex Q-*operator*.

(f) Although $\mathcal{Q}$ has an unbounded convex conjugate, the convex initial value problem

$$u'(t) = \mathcal{Q}u(t) \text{ for all } t \geq 0, \quad u(0) = u_0, \tag{4}$$

has a unique global solution.
(g) Solutions to (4) remain bounded. Therefore, a Picard iteration or Runge–Kutta methods, such as the explicit Euler method, can be used for numerical computations, and the convergence rate (depending on the size of the initial value u_0) can be derived from the a priori estimate in Banach's fixed point theorem.
(h) As in the linear case, by solving the differential equation (4), one can (numerically) compute expressions of the form

$$u(t) = \mathcal{E}(u_0(X_t)).$$

We illustrate this computation procedure in Example 5.1.

3 PROOF OF $(v) \Rightarrow (ii) \Rightarrow (i) \Rightarrow (iii)$

We say that a set $\mathcal{P} \subset \mathbb{R}^{d\times d}$ of matrices is *row-convex* if, for any diagonal matrix $\theta \in \mathbb{R}^{d\times d}$ with $\theta_i := \lambda_{ii} \in [0,1]$ for all $i \in \{1, \dots, d\}$,

$$\theta p + (I - \theta)q \in \mathcal{P} \quad \text{for all } p, q \in \mathcal{P},$$

where $I = I_d \in \mathbb{R}^{d\times d}$ is the d-dimensional identity matrix. Notice that, for all $i \in \{1, \dots, d\}$, the ith row of the matrix $\theta p + (I - \theta)q$ is the convex combination of the ith row of p and q with θ_i. Notice that a set $\mathcal{P} \subset \mathbb{R}^{d\times d}$ is row-convex if and only if it is convex and, for arbitrary $p, q \in \mathcal{P}$, the matrix that results from replacing the ith row of p by the ith row of q is again an element of $\mathcal{P}$. For example, the set of all Q-matrices is row-convex.

Remark 3.1. Let $\mathcal{Q}$ be a convex Q-operator. For every matrix $q \in \mathbb{R}^{d\times d}$, let

$$\mathcal{Q}^*(q) := \sup_{u\in\mathbb{R}^d} (qu - \mathcal{Q}(u)) \in [0, \infty]^d$$

be the *conjugate function* of $\mathcal{Q}$. Notice that $0 \leq \mathcal{Q}^*(q)$ for all $q \in \mathbb{R}^{d\times d}$, since $\mathcal{Q}(0) = 0$. Let

$$\mathcal{P}^* := \left\{ q \in \mathbb{R}^{d\times d} \,\middle|\, \mathcal{Q}^*(q) \in [0, \infty)^d \right\}$$

and $f_q^* := -\mathcal{Q}^*(q)$ for all $q \in \mathcal{P}^*$. Then, the following facts are well-known results from convex duality theory in $\mathbb{R}^d$.

(a) The set $\mathcal{P}^*$ is row-convex and the mapping $\mathcal{P}^* \to \mathbb{R}^d,\ q \mapsto \mathcal{Q}^*(q)$ is lower semicontinuous.

(b) Let $M \geq 0$ and $\mathcal{P}_M^* := \{q \in \mathbb{R}^{d\times d} \,|\, \mathcal{Q}^*(q) \leq M\}$. Then, $\mathcal{P}_M^* \subset \mathbb{R}^{d\times d}$ is compact and row-convex. Therefore,

$$\mathcal{Q}_M : \mathbb{R}^d \to \mathbb{R}^d, \quad u \mapsto \max_{q\in\mathcal{P}_M^*} \left(qu + f_q^*\right) \tag{5}$$

defines a convex operator, which is Lipschitz continuous. Notice that the maximum in (5) is to be understood componentwise. However, for fixed $u_0 \in \mathbb{R}^d$, the maximum can be attained, simultaneously in every component, by a single element of $\mathcal{P}_M^*$, that is, for all $u_0 \in \mathbb{R}^d$, there exists some $q_0 \in \mathcal{P}_M^*$ with

$$\mathcal{Q}_M u_0 = q_0 u_0 + f_{q_0}^*.$$

This is due to the fact that $\mathcal{P}_M^*$ is row convex and that, for $q \in \mathcal{P}^*$, the ith component of the vector f_q^* only depends on the ith row of q.

(c) Let $R \geq 0$. Then, there exists some $M \geq 0$, such that

$$\mathcal{Q}u_0 = \max_{q\in\mathcal{P}_M^*} \left(qu_0 + f_q^*\right) = \mathcal{Q}_M u_0$$

for all $u_0 \in \mathbb{R}^d$ with $\|u_0\|_\infty \leq R$. In particular, $\mathcal{Q}$ is locally Lipschitz continuous and

$$\mathcal{Q}u_0 = \max_{q\in\mathcal{P}^*} \left(qu_0 + f_q^*\right) \quad \text{for all } u_0 \in \mathbb{R}^d,$$

where, for fixed $u_0 \in \mathbb{R}^d$, the maximum can be attained, simultaneously in every component, by a single element of $\mathcal{P}^*$. In particular, there exists some $q_0 \in \mathcal{P}^*$ with $f_{q_0}^* = \sup_{q\in\mathcal{P}^*} f_q^* = \mathcal{Q}(0) = 0$.

Proof of Theorem 2.5. $(v) \Rightarrow (ii)$: As $\mathcal{E}_i$ is a convex expectation for all $i \in \{1, \dots, d\}$, it follows that the operator $\mathcal{Q}$ is convex with $\mathcal{Q}\alpha = 0$ for all $\alpha \in \mathbb{R}$. Now, let $u_0 \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$ with $u_{0,i} \geq u_{0,j}$ for all $j \in \{1, \dots, d\}$. Let $\alpha > 0$ be such that

$$\|u_0 + \alpha\|_\infty = (u_0 + \alpha)_i = u_{0,i} + \alpha,$$

and define $v_0 := u_0 + \alpha$. Then,

$$\mathcal{Q}v_0 = \lim_{h\searrow 0} \frac{\mathcal{E}(u_0(X_h) + \alpha) - v_0}{h} = \lim_{h\searrow 0} \frac{\mathcal{E}(u_0(X_h)) - u_0}{h} = \mathcal{Q}u_0.$$

Assume that $(\mathcal{Q}u_0)_i > 0$. Then, there exists some $h > 0$ such that

$$\mathcal{E}_i(v_0(X_h)) - v_{0,i} > 0.$$

Hence,

$$\left\|\mathcal{E}(v_0(X_h))\right\|_\infty \geq \mathcal{E}_i(v_0(X_h)) > v_{0,i} = \|v_0\|_\infty,$$

which is a contradiction to

$$\left\|\mathcal{E}(v_0(X_h))\right\|_\infty \le \|v_0\|_\infty.$$

This shows that Q satisfies the positive maximum principle.

$(ii) \Rightarrow (i)$: This follows directly from the positive maximum principle, considering the vectors λe_i and $-\lambda e_i$ for all $\lambda > 0$ and $i \in \{1, \dots, d\}$.

$(i) \Rightarrow (iii)$: Let $\mathcal{Q}$ be a convex *Q-operator*. Moreover, let $\mathcal{P}^*$ and $f^* = (f_q^*)_{q\in\mathcal{P}^*}$ be as in Remark 5. Then, by Remark 5 (c), it only remains to show that every $q \in \mathcal{P}^*$ is a *Q-matrix*. To this end, fix an arbitrary $q \in \mathcal{P}^*$. Then, for all $\alpha \in \mathbb{R}$,

$$q\alpha = \frac{1}{\lambda}q(\lambda\alpha) \le \frac{1}{\lambda}(\mathcal{Q}(\lambda\alpha) + \mathcal{Q}^*(q)) = \frac{1}{\lambda}\mathcal{Q}^*(q) \to 0 \quad \text{as } \lambda \to \infty.$$

Therefore, $q\alpha \le 0$ for all $\alpha \in \mathbb{R}$. Since q is linear, it follows that $q1 = 0$. Now, let $i \in \{1, \dots, d\}$. Then, by definition of a *Q-operator*, we obtain that

$$q_{ii} \le \frac{1}{\lambda}(\mathcal{Q}(\lambda e_i) + \mathcal{Q}^*(q))_i \le \frac{1}{\lambda}(\mathcal{Q}^*(q))_i \to 0 \quad \text{as } \lambda \to \infty,$$

that is, $q_{ii} \le 0$. Now, let $i, j \in \{1, \dots, d\}$ with $i \ne j$. Then, again by definition of a *Q-operator*, it follows that

$$-q_{ij} \le \frac{1}{\lambda}(\mathcal{Q}(-\lambda e_i) + \mathcal{Q}^*(q))_j \le \frac{1}{\lambda}(\mathcal{Q}^*(q))_j \to 0 \quad \text{as } \lambda \to \infty,$$

that is, $q_{ij} \ge 0$. Therefore, q is a *Q-matrix.*

It remains to show the implications $(iii) \Rightarrow (iv) \Rightarrow (v)$, which is done in the entire next section. □

Before we start with the proof of the remaining implications $(iii) \Rightarrow (iv) \Rightarrow (v)$, we would like to point out how, in the sublinear case, the set $\mathcal{P}^*$ of Q-matrices from Remark 3.1 can be reduced to certain “corner points.” This can be done using the concept of row convexity, introduced at the beginning of this section, together with Minkowski's theorem on extremal points of convex sets in $\mathbb{R}^d$. Let $\mathcal{M} \subset \mathbb{R}^{d\times d}$ be a nonempty set of matrices. Then, we define the *row-convex hull* of $\mathcal{M}$ by

$$\operatorname{rch}(\mathcal{M}) := \left\{ \sum_{i=1}^n \theta^i q^i \,\middle|\, n \in \mathbb{N},\, \theta^1, \dots, \theta^n \in [0,\infty)^{d\times d},\, \sum_{i=1}^n \theta^i = I,\, q^1, \dots q^n \in \mathcal{M} \right\}.$$

For a convex set $C \subset \mathbb{R}^d$, we denote the set of all extreme points of C by $E(C)$. Recall that an extreme point of a convex set $C \subset \mathbb{R}^d$ is an element $x \in C$ such that $x = \lambda y + (1-\lambda)z$, for $\lambda \in (0,1)$ and $y, z \in C$, implies that $x = y = z$. For a matrix $q \in \mathbb{R}^{d\times d}$ and $i \in \{1, \dots, d\}$, we denote by

$$q_i := (q_{i1}, \dots, q_{id}) \in \mathbb{R}^d$$

the ith row of q. Let $\mathcal{P} \subset \mathbb{R}^{d\times d}$ be a nonempty compact row-convex set of matrices. Then, we say that a set $\mathcal{R} \subset \mathcal{P}$ is *$\mathcal{P}$-row-extreme* if

$$\{q_i \,|\, q \in \mathcal{R}\} = E(\{q_i \,|\, q \in \mathcal{P}\}) \quad \text{for all } i \in \{1, \dots, d\}.$$

That is, the set of all ith rows of $\mathcal{R}$ is the set of all extreme points of the ith rows of $\mathcal{P}$. We say that a set $\mathcal{M} \subset \mathcal{P}$ is *minimal $\mathcal{P}$-row-extreme*, if $\mathcal{M}$ is row-extreme for $\mathcal{P}$ and $\mathcal{R} \subset \mathcal{M}$ implies $\mathcal{R} = \mathcal{M}$ for any $\mathcal{P}$-row-extreme set $\mathcal{R} \subset \mathcal{P}$.

Proposition 3.2. *Let $\mathcal{P} \subset \mathbb{R}^{d\times d}$ be nonempty, compact, and row-convex. Then, there exists a minimal $\mathcal{P}$-row-extreme set $\mathcal{M} \subset \mathcal{P}$. Moreover, $\mathcal{P} = \mathrm{rch}(\mathcal{R})$ is the row-convex hull of any (minimal) $\mathcal{P}$-row-extreme set $\mathcal{R} \subset \mathcal{P}$ and*

$$\max_{q\in\mathcal{P}} qu_0 = \max_{q\in\mathcal{R}} qu_0 \quad \textit{for all } u_0 \in \mathbb{R}^d, \tag{6}$$

where the maxima are to be understood componentwise.

Proof. By Minkowski's theorem, the set of all $\mathcal{P}$-row-extreme sets is nonempty, and one readily verifies that the latter together with the partial order $\preceq$, given by $\mathcal{R}_1 \preceq \mathcal{R}_2$ if and only if $\mathcal{R}_1 \supset \mathcal{R}_2$, has the chain property. Hence, by Zorn's lemma, there exists a maximal element $\mathcal{M}$ within the set of all $\mathcal{P}$-row-extreme sets, which, by definition, is a minimal $\mathcal{P}$-row-extreme set. Now, let $\mathcal{R}$ be an arbitrary $\mathcal{P}$-row-extreme set and $u_0 \in \mathbb{R}^d$. Then,

$$\max_{q\in\mathcal{P}} (qu_0)_i = \max_{q\in\mathcal{P}} (q_i \cdot u_0) = \max_{q\in\mathcal{R}} (q_i \cdot u_0) = \max_{q\in\mathcal{R}} (qu_0)_i.$$

□

Remark 3.3. Let $\mathcal{Q} : \mathbb{R}^d \to \mathbb{R}^d$ be a sublinear *Q-operator*, and $\mathcal{P}^*$ as in Remark 3.1. Then,

$$\mathcal{P}^* = \left\{ q \in \mathbb{R}^d \,\middle|\, f_q^* = \mathcal{Q}^*(q) = 0 \right\}$$

is a nonempty, compact, and row-convex set. By the previous proposition, there exists a minimal $\mathcal{P}^*$-row-extreme set $\mathcal{M} \subset \mathcal{P}^*$, and, for all $u_0 \in \mathbb{R}^d$,

$$\mathcal{Q}u_0 = \max_{q\in\mathcal{M}} qu_0,$$

where the maximum is to be understood componentwise. Since $f^* = (f_q^*)_{q\in\mathcal{P}^*} = 0$, it follows that $(\mathcal{M}, 0)$ is a dual representation as in Theorem 2.5(iii). Notice that, in many cases, the cardinality of $\mathcal{M}$ is way smaller than the cardinality of $\mathcal{P}^*$. Therefore, concerning computational aspects, the dual representation $(\mathcal{M}, 0)$ is often way more tractable than the dual representation $(\mathcal{P}^*, 0)$, and, by Theorem 2.5, both representations result in the same semigroup envelope, and thus, the same solution to the ODE (3).

Example 3.4. Let $q_0, q \in \mathbb{R}^{d\times d}$ be two fixed Q-matrices and $\lambda_l, \lambda_h \in \mathbb{R}$ with $\lambda_l \leq \lambda_h$. We define the sublinear *Q-operator* $\mathcal{Q} : \mathbb{R}^d \to \mathbb{R}^d$ by

$$\mathcal{Q}u_0 := q_0u_0 + \max_{\lambda\in[\lambda_l,\lambda_h]} \lambda q u_0 \quad \text{for all } u_0 \in \mathbb{R}^d.$$

We consider the maximal row-convex set $\mathcal{P}^* \subset \mathbb{R}^{d\times d}$ representing $\mathcal{Q}$, defined as in Remark 3.1. Then,

$$\mathcal{P}^* = \left\{ p_0 + \lambda p \,\middle|\, \lambda \in \operatorname{diag}([\lambda_l, \lambda_h]) \right\},$$

where $\operatorname{diag}([\lambda_l, \lambda_h])$ denotes the set of all diagonal matrices $\lambda \in \mathbb{R}^{d\times d}$ with diagonal entries $\lambda_{ii} \in [\lambda_l, \lambda_h]$ for all $i \in \{1, \dots, d\}$. Now, let

$$\mathcal{M} := \{q_0 + \lambda_l q, q_0 + \lambda_h q\}.$$

Then, $\mathcal{M}$ is a minimal $\mathcal{P}^*$-row-extreme set, and thus, $\mathcal{P}^* = \operatorname{rch}(\mathcal{M})$. In particular, by the previous remark, the tuple

$$(\{q_0 + \lambda_l q, q_0 + \lambda_h q\}, (0, 0))$$

is a dual representation as in Theorem 2.5(iii), which is way more tractable than the dual representation $(\mathcal{P}^*, 0)$.

4 | PROOF OF $(iii) \Rightarrow (iv) \Rightarrow (v)$

Throughout, let $\mathcal{P} \subset \mathbb{R}^{d\times d}$ be a set of Q-matrices and $f = (f_q)_{q\in\mathcal{P}} \subset \mathbb{R}^d$ with $f_q \leq 0$ for all $q \in \mathcal{P}$ and $f_{q_0} = 0$ for some $q_0 \in \mathcal{P}$, such that the map

$$\mathcal{Q} : \mathbb{R}^d \to \mathbb{R}^d, \quad u \mapsto \sup_{q\in\mathcal{P}} \left(qu + f_q\right)$$

is well-defined. For every $q \in \mathcal{P}$, we consider the linear ODE

$$u'(t) = qu(t) + f_q, \quad \text{for } t \geq 0, \tag{7}$$

with $u(0) = u_0 \in \mathbb{R}^d$. Then, by a variation of constant, the solution to (7) is given by

$$u(t) = e^{qt}u_0 + \int_0^t e^{qs} f_q \,\mathrm{d}s = u_0 + \int_0^t e^{sq}\left(qu_0 + f_q\right) \mathrm{d}s =: S_q(t)u_0 \tag{8}$$

for $t \geq 0$, where $e^{tq} \in \mathbb{R}^{d\times d}$ is the matrix exponential of tq for all $t \geq 0$. Then, the family $S_q = (S_q(t))_{t\geq 0}$ defines a uniformly continuous semigroup of affine linear operators (see Definition 2.3).

Remark 4.1. Note that, for all $q \in \mathcal{P}$ and $t \geq 0$, the matrix exponential $e^{tq} \in \mathbb{R}^{d\times d}$ is a *stochastic matrix*, that is,

(i) $(e^{tq})_{ij} \geq 0$ for all $i, j \in \{1, \dots, d\}$,
(ii) $e^{tq}1 = 1$.

Therefore, $e^{tq} \in \mathbb{R}^{d\times d}$ is a linear kernel, that is, $e^{tq}u_0 \leq e^{tq}v_0$ for all $u_0, v_0 \in \mathbb{R}^d$ with $u_0 \leq v_0$ and $e^{tq}\alpha = \alpha$ for all $\alpha \in \mathbb{R}$, which implies that $S_q(t)$ is monotone for all $q \in \mathcal{P}$ and $t \geq 0$.

For the family $(S_q)_{q\in\mathcal{P}}$ or, more precisely, for $(\mathcal{P}, f)$, we will now construct the *Nisio semigroup*, and show that it gives rise to the unique classical solution to the nonlinear ODE (3). To this end, we consider the set of finite partitions

$$P := \left\{\pi \subset [0,\infty) \,\middle|\, 0 \in \pi, |\pi| < \infty\right\}.$$

The set of partitions with end point $t \geq 0$ will be denoted by P_t, that is, $P_t := \{\pi \in P \mid \max \pi = t\}$. Notice that

$$P = \bigcup_{t\geq 0} P_t.$$

For all $h \geq 0$ and $u_0 \in \mathbb{R}^d$, we define

$$\mathcal{E}_h u_0 := \sup_{q\in\mathcal{P}} S_q(h)u_0,$$

where the supremum is taken componentwise. Note that $\mathcal{E}_h$ is well-defined since

$$S_q(h)u_0 = e^{hq}u_0 + \int_0^h e^{sq} f_q \, \mathrm{d}s \leq e^{hq}u_0 \leq \|u_0\|_\infty$$

for all $q \in \mathcal{P}$, $h \geq 0$ and $u_0 \in \mathbb{R}^d$, where we used the fact that e^{hq} is a kernel. Moreover, $\mathcal{E}_h$ is a convex kernel, for all $h \geq 0$, as it is monotone and

$$\mathcal{E}_h\alpha = \alpha + \sup_{q\in\mathcal{P}} \int_0^h e^{sq} f_q \, \mathrm{d}s = \alpha$$

for all $\alpha \in \mathbb{R}$, where we used the fact that there is some $q_0 \in \mathcal{P}$ with $f_{q_0} = 0$. For a partition $\pi = \{t_0, t_1, \dots, t_m\} \in P$ with $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_m$, we set

$$\mathcal{E}_\pi := \mathcal{E}_{t_1 - t_0} \dots \mathcal{E}_{t_m - t_{m-1}}.$$

Moreover, we set $\mathcal{E}_{\{0\}} := \mathcal{E}_0$. Then, $\mathcal{E}_\pi$ is a convex kernel for all $\pi \in P$ since it is a concatenation of convex kernels.

Definition 4.2. The *Nisio semigroup* $\mathscr{S} = (\mathscr{S}(t))_{t\geq 0}$ of $(\mathcal{P}, f)$ is defined by

$$\mathscr{S}(t)u_0 := \sup_{\pi\in P_t} \mathcal{E}_\pi u_0$$

for all $u_0 \in \mathbb{R}^d$ and $t \geq 0$.

Notice that $\mathscr{S}(t) : \mathbb{R}^d \to \mathbb{R}^d$ is well-defined and a convex kernel for all $t \geq 0$ since $\mathcal{E}_\pi$ is a convex kernel for all $\pi \in P$. In many of the subsequent proofs, we will first concentrate on the case, where the family f is bounded and then use an approximation of the Nisio semigroup by means of other Nisio semigroups. This approximation procedure is specified in the following remark.

Remark 4.3. Let $M \geq 0$, $\mathcal{P}_M := \{q \in \mathcal{P} \mid \|f_q\|_\infty \leq M\}$ and $f_M := (f_q)_{q \in \mathcal{P}_M}$. Notice that, by assumption, there exists some $q_0 \in \mathcal{P}$ with $f_{q_0} = 0$, which implies that $q_0 \in \mathcal{P}_M$. Since $\mathcal{P}_M \subset \mathcal{P}$ (and by definition of f_M), the operator

$$\mathcal{Q}_M : \mathbb{R}^d \to \mathbb{R}^d, \quad v \mapsto \sup_{q \in \mathcal{P}_M} \left(qv + f_q\right)$$

is well-defined. Let $\mathscr{S}_M$ be the Nisio semigroup w.r.t. $(\mathcal{P}_M, f_M)$ for all $M \geq 0$. Since

$$\bigcup_{M \geq 0} \mathcal{P}_M = \mathcal{P},$$

it follows that $\mathcal{Q}_M \nearrow \mathcal{Q}$ and $\mathscr{S}_M(t) \nearrow \mathscr{S}(t)$, for all $t \geq 0$, as $M \to \infty$. Moreover, for all $q \in \mathcal{P}_M$, $u_0 \in \mathbb{R}^d$ with $\|u_0\|_\infty = 1$, and $i \in \{1, \dots, d\}$,

$$(qu_0)_i \leq \left(\mathcal{Q}u_0 - f_q\right)_i \leq \|\mathcal{Q}u_0\|_\infty + \|f_q\|_\infty \leq M + \max_{v \in \mathbb{S}^{d-1}} \|\mathcal{Q}v\|_\infty,$$

where $\mathbb{S}^{d-1} := \{v \in \mathbb{R}^d \mid \|v\|_\infty = 1\}$ and, in the last step, we used the fact that $\mathcal{Q} : \mathbb{R}^d \to \mathbb{R}^d$ is convex and therefore continuous. This implies that the set $\mathcal{P}_M$ is bounded in the sense that $\sup_{q \in \mathcal{P}_M} \|q\| < \infty$. In particular,

$$\sup_{q \in \mathcal{P}_M} \|qu_0 + f_q\|_\infty \leq \sup_{q \in \mathcal{P}_M} \left(\|q\|\|u_0\|_\infty + \|f_q\|_\infty\right) \leq M + \sup_{q \in \mathcal{P}_M} \|q\|\|u_0\|_\infty < \infty \tag{9}$$

for all $u_0 \in \mathbb{R}^d$.

Lemma 4.4. *Assume that the family f is bounded, that is, $(\mathcal{P}, f) = (\mathcal{P}_M, f_M)$ for some $M \geq 0$. Then, for all $u_0 \in \mathbb{R}^d$, the mapping $[0, \infty) \to \mathbb{R}^d,\ h \mapsto \mathcal{E}_h u_0$ is Lipschitz continuous.*

Proof. Let $u_0 \in \mathbb{R}^d$ and $0 \leq h_1 < h_2$. Then, by (8), for all $q \in \mathcal{P}$, we have that

$$\|S_q(h_2)u_0 - S_q(h_1)u_0\|_\infty \leq \int_{h_1}^{h_2} \left\|e^{qs}(qu_0 + f_q)\right\|_\infty \mathrm{d}s \leq (h_2 - h_1)\|qu_0 + f_q\|_\infty,$$

which implies that

$$\|\mathcal{E}_{h_2}u_0 - \mathcal{E}_{h_1}u_0\|_\infty \leq \sup_{q \in \mathcal{P}} \|S_q(h_2)u_0 - S_q(h_1)u_0\|_\infty \leq (h_2 - h_1)\left(\sup_{q \in \mathcal{P}} \|qu_0 + f_q\|_\infty\right). \tag{10}$$

Note that $\sup_{q \in \mathcal{P}} \|qu_0 + f_q\|_\infty < \infty$ by (9). □

Lemma 4.5. *Assume that the family f is bounded. Then,*

$$\|\mathscr{S}(t)u_0 - u_0\|_\infty \le t\left(\sup_{q\in\mathcal{P}} \|qu_0 + f_q\|_\infty\right)$$

for all $t \ge 0$ and $u_0 \in \mathbb{R}^d$. In particular, the map $[0,\infty) \to \mathbb{R}^d,\ t \mapsto \mathscr{S}(t)u_0$ is Lipschitz continuous for all $u_0 \in \mathbb{R}^d$.

Proof. Let $u_0 \in \mathbb{R}^d$. Then, for any partition $\pi \in P$ of the form $\pi = \{t_0, t_1, \dots, t_m\}$ with $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_m$, (10) together with the fact that $\mathcal{E}_h$ is a kernel, for all $h \ge 0$, implies that

$$\begin{aligned}\|\mathcal{E}_\pi u_0 - u_0\|_\infty &\le \sum_{k=1}^m \|\mathcal{E}_{h_k} u_0 - u_0\|_\infty \le \sum_{k=1}^m h_k\left(\sup_{q\in\mathcal{P}} \|qu_0 + f_q\|_\infty\right)\\ &= t_m\left(\sup_{q\in\mathcal{P}} \|qu_0 + f_q\|_\infty\right),\end{aligned}$$

where $h_k := t_k - t_{k-1}$ for all $k \in \{1, \dots, m\}$. By definition of $\mathscr{S}(t)$, for $t \ge 0$, it follows that

$$\|\mathscr{S}(t)u_0 - u_0\|_\infty \le \sup_{\pi\in P_t} \|\mathcal{E}_\pi u_0 - u_0\|_\infty \le t\left(\sup_{q\in\mathcal{P}} \|qu_0 + f_q\|_\infty\right).$$

Now, let $s, t \ge 0$. Then, since $\mathscr{S}(h)$ is a kernel for all $h \ge 0$, it follows that

$$\|\mathscr{S}(t)u_0 - \mathscr{S}(s)u_0\|_\infty \le \|\mathscr{S}(|t-s|)u_0 - u_0\|_\infty \le |t-s|\left(\sup_{q\in\mathcal{P}} \|qu_0 + f_q\|_\infty\right).$$

□

For a partition $\pi = \{t_0, t_1, \dots, t_m\} \in P$ with $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_m$, we define the *(maximal)* mesh size of π by

$$|\pi|_\infty := \max_{j=1,\dots,m} (t_j - t_{j-1}).$$

Moreover, we set $|\{0\}|_\infty := 0$. Let $u_0 \in \mathbb{R}^d$. In the following, we consider the limit of $\mathcal{E}_\pi u_0$ when the mesh size of the partition $\pi \in P$ tends to zero. For this, we first remark that, for $h_1, h_2 \ge 0$,

$$\begin{aligned}\mathcal{E}_{h_1+h_2} u_0 &= \sup_{q\in\mathcal{P}} S_\lambda(h_1+h_2)u_0 = \sup_{q\in\mathcal{P}} S_\lambda(h_1)S_\lambda(h_2)u_0\\ &\le \sup_{q\in\mathcal{P}} S_\lambda(h_1)\mathcal{E}_{h_2}u_0 = \mathcal{E}_{h_1}\mathcal{E}_{h_2}u_0,\end{aligned}$$

which implies the inequality

$$\mathcal{E}_{\pi_1} u_0 \le \mathcal{E}_{\pi_2} u_0 \tag{11}$$

for $\pi_1, \pi_2 \in P$ with $\pi_1 \subset \pi_2$. The following lemma now states that $\mathscr{S}(t)$, for $t \geq 0$, can be obtained by a pointwise monotone approximation with finite partitions letting the mesh size tend to zero.

Lemma 4.6. *Let $t \geq 0$ and $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with $\pi_n \subset \pi_{n+1}$ for all $n \in \mathbb{N}$ and $|\pi_n|_\infty \searrow 0$ as $n \to \infty$. Then, for all $u_0 \in \mathbb{R}^d$,*

$$\mathcal{E}_{\pi_n} u_0 \nearrow \mathscr{S}(t)u_0, \quad n \to \infty.$$

Proof. Let $u_0 \in \mathbb{R}^d$. For $t = 0$ the statement is trivial. Therefore, assume that $t > 0$, and let

$$u_\infty := \sup_{n\in\mathbb{N}} \mathcal{E}_{\pi_n} u_0. \tag{12}$$

As $\pi_n \subset \pi_{n+1}$ for all $n \in \mathbb{N}$, (11) implies that

$$\mathcal{E}_{\pi_n} u_0 \nearrow u_\infty, \quad n \to \infty.$$

Since $(\pi_n)_{n\in\mathbb{N}} \subset P_t$, we obtain that

$$u_\infty \leq \mathscr{S}(t)u_0.$$

Next, we assume that the family f is bounded. Let $\pi = \{t_0, t_1, \dots, t_m\} \in P_t$ with $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_m = t$. Since $|\pi_n|_\infty \searrow 0$ as $n \to \infty$, we may w.l.o.g. assume that $|\pi_n| \geq m+1$ for all $n \in \mathbb{N}$. Again, since $|\pi_n|_\infty \searrow 0$ as $n \to \infty$, there exist $0 = t_0^n < t_1^n < \cdots < t_m^n = t$ for all $n \in \mathbb{N}$ with $\pi_n' := \{t_0^n, t_1^n, \dots, t_m^n\} \subset \pi_n$ and $t_i^n \to t_i$ as $n \to \infty$ for all $i \in \{1, \dots, m\}$. Then, by Lemma 4.4, we have that

$$\|\mathcal{E}_\pi u_0 - \mathcal{E}_{\pi_n'} u_0\|_\infty \to 0, \quad n \to \infty,$$

and therefore,

$$u_\infty \geq \mathcal{E}_{\pi_n} u_0 \geq \mathcal{E}_{\pi_n'} u_0 \geq \mathcal{E}_\pi u_0 - \|\mathcal{E}_\pi u_0 - \mathcal{E}_{\pi_n'} u_0\|_\infty.$$

Letting $n \to \infty$, we obtain that $u_\infty \geq \mathcal{E}_\pi u_0$. Taking the supremum over all $\pi \in P_t$ yields the assertion for bounded f.

Now, let f again be (possibly) unbounded. It remains to show that $u_\infty \geq \mathscr{S}(t)u_0$. By the previous step, we have that $u_\infty \geq u_{\infty,M} = \mathscr{S}_M(t)$ for all $M \geq 0$, where $u_{\infty,M}$ is given by (12) but w.r.t. $(\mathcal{P}_M, f_M)$ instead of $(\mathcal{P}, f)$. Since $\mathscr{S}_M(t)u_0 \nearrow \mathscr{S}(t)u_0$ as $M \to \infty$, we obtain that $u_\infty \geq \mathscr{S}(t)u_0$, which ends the proof. □

Choosing, for example, $\pi_n = \{\frac{kt}{2^n} : k \in \{0, \dots, 2^n\}\}$ or $\pi_n = \left\{\frac{kt}{n!} : k \in \{0, \dots, n!\}\right\}$ in Lemma 4.6, we obtain the following corollaries.

Corollary 4.7. *For all $t \geq 0$, there exists a sequence $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with*

$$\mathcal{E}_{\pi_n} u_0 \nearrow \mathscr{S}(t)u_0$$

as $n \to \infty$ for all $u_0 \in \mathbb{R}^d$.

Corollary 4.8. *For all $t \geq 0$ and $u_0 \in \mathbb{R}^d$,*

$$\mathscr{S}(t)u_0 = \sup_{n\in\mathbb{N}} \mathcal{E}^n_{\frac{1}{n}} u_0 = \lim_{n\to\infty} \mathcal{E}^{2^n}_{2^{-n}} u_0.$$

Proposition 4.9. *The family $\mathscr{S} = (\mathscr{S}(t))_{t\geq 0}$ defines a semigroup of convex kernels from $\mathbb{R}^d$ to $\mathbb{R}^d$. In particular, for all $s, t \geq 0$, we have the dynamic programming principle*

$$\mathscr{S}(s+t) = \mathscr{S}(s)\mathscr{S}(t). \tag{13}$$

Moreover, the Nisio semigroup $\mathscr{S}$ of $(\mathcal{P}, f)$ coincides with the semigroup envelope of $(\mathcal{P}, f)$ (cf. Definition 2.4).

Proof. We have already shown that $\mathscr{S}(t)$ is a convex kernel for all $t \geq 0$, and, by definition, $\mathscr{S}(0) = I_d$. Let $u_0 \in \mathbb{R}^d$. If $s = 0$ or $t = 0$, the statement is trivial. Therefore, let $s, t > 0$, $\pi_0 \in P_{s+t}$ and $\pi := \pi_0 \cup \{s\}$. Then, $\pi \in P_{s+t}$ with $\pi_0 \subset \pi$. Hence, by (11), we obtain that

$$\mathcal{E}_{\pi_0} u_0 \leq \mathcal{E}_\pi u_0.$$

Let $m \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots t_m = s + t$ with $\pi = \{t_0, \dots, t_m\}$ and $i \in \{1, \dots, m\}$ with $t_i = s$. Then,

$$\pi_1 := \{t_0, \dots, t_i\} \in P_s \quad \text{and} \quad \pi_2 := \{t_i - s, \dots, t_m - s\} \in P_t$$

with

$$\mathcal{E}_{\pi_1} = \mathcal{E}_{t_1 - t_0} \dots \mathcal{E}_{t_i - t_{i-1}} \quad \text{and} \quad \mathcal{E}_{\pi_2} = \mathcal{E}_{t_{i+1} - t_i} \dots \mathcal{E}_{t_m - t_{m-1}}.$$

We thus see that

$$\begin{aligned}\mathcal{E}_{\pi_0} u_0 \leq \mathcal{E}_\pi u_0 &= \mathcal{E}_{t_1 - t_0} \dots \mathcal{E}_{t_m - t_{m-1}} u_0 = \big(\mathcal{E}_{t_1 - t_0} \dots \mathcal{E}_{t_i - t_{i-1}}\big)\big(\mathcal{E}_{t_{i+1} - t_i} \dots \mathcal{E}_{t_m - t_{m-1}} u_0\big) \\ &= \mathcal{E}_{\pi_1}\mathcal{E}_{\pi_2} u_0 \leq \mathcal{E}_{\pi_1}(\mathscr{S}(t)u_0) \leq \mathscr{S}(s)\mathscr{S}(t)u_0.\end{aligned}$$

Taking the supremum over all $\pi_0 \in P_{s+t}$, it follows that $\mathscr{S}(s+t)u_0 \leq \mathscr{S}(s)\mathscr{S}(t)u_0$.

Now, let $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with $\mathcal{E}_{\pi_n} u_0 \nearrow \mathscr{S}(t)u_0$ as $n \to \infty$ (see Corollary 4.7), and fix $\pi_0 \in P_s$. Then, for all $n \in \mathbb{N}$,

$$\pi_n' := \pi_0 \cup \{s + \tau : \tau \in \pi_n\} \in P_{s+t}$$

with $\mathcal{E}_{\pi_n'} = \mathcal{E}_{\pi_0}\mathcal{E}_{\pi_n}$. Therefore,

$$\mathcal{E}_{\pi_0}(\mathscr{S}(t)u_0) = \lim_{n\to\infty} \mathcal{E}_{\pi_0}\mathcal{E}_{\pi_n} u_0 = \lim_{n\to\infty} \mathcal{E}_{\pi_n'} u_0 \leq \mathscr{S}(s+t)u_0.$$

Taking the supremum over all $\pi_0 \in P_s$ yields that $\mathscr{S}(s)\mathscr{S}(t)u_0 \leq \mathscr{S}(s+t)u_0$. Therefore, $\mathscr{S}$ satisfies the semigroup property (13).

It remains to show that the family $\mathscr{S}$ is the semigroup envelope of $(\mathcal{P}, f)$. We have already shown that $\mathscr{S}$ is a semigroup and, by definition, $\mathscr{S}(t)u_0 \geq S_q(t)u_0$ for all $t \geq 0$, $u_0 \in \mathbb{R}^d$ and $q \in \mathcal{P}$.

Let $(\mathscr{T}(t))_{t\geq 0}$ be a semigroup with $\mathscr{T}(t)u_0 \geq S_q(t)u_0$ for all $t \geq 0$, $u_0 \in \mathbb{R}^d$, and $q \in \mathcal{P}$. Then,

$$\mathcal{E}_h u_0 \leq \mathscr{T}(h)u_0 \quad \text{for all } h \geq 0 \text{ and } u_0 \in \mathbb{R}^d.$$

Since $(\mathscr{T}(t))_{t\geq 0}$ is a semigroup and $\mathcal{E}_h$ is monotone for all $h \geq 0$, it follows that

$$\mathcal{E}_\pi u_0 \leq \mathscr{T}(t)u_0 \quad \text{for all } t \geq 0,\ \pi \in P_t \text{ and } u_0 \in \mathbb{R}^d.$$

Taking the supremum over all $\pi \in P_t$, it follows that $\mathscr{S}(t)u_0 \leq \mathscr{T}(t)u_0$ for all $t \geq 0$ and $u_0 \in \mathbb{R}^d$. □

To finish the proof of the implication $(iii) \Rightarrow (iv)$, it remains to show that the Nisio semigroup $\mathscr{S}$ is uniformly continuous and that it gives rise to the unique classical solution to the nonlinear ODE (3).

Remark 4.10. Assume that the set $\mathcal{P}$ is bounded, that is, $\sup_{q\in\mathcal{P}} \|q\| < \infty$.

(a) Since $\mathcal{P}$ is bounded, it follows that $\mathcal{Q}$ is Lipschitz continuous. Therefore, the Picard–Lindelöf Theorem implies that, for every $u_0 \in \mathbb{R}^d$, the initial value problem

$$\begin{aligned} u'(t) &= \mathcal{Q}u(t), \quad t \geq 0, \\ u(0) &= u_0, \end{aligned} \tag{14}$$

has a unique solution $u \in C^1([0,\infty);\mathbb{R}^d)$. We will show that this solution u is given by $u(t) = \mathscr{S}(t)u_0$ for all $t \geq 0$. That is, the unique solution of the ODE (14) is given by the Nisio semigroup.

(b) Since $\mathcal{P}$ is bounded, the mapping

$$\mathbf{q} : \mathbb{R}^d \to \mathbb{R}^d, \quad u \mapsto \sup_{q\in\mathcal{P}} qu$$

is well-defined.

The following key estimate and its proof are a straightforward adaption of the proof of (Nisio, 1976/77, Proposition 5) to our setup. Recall that, by Remark 4.3, the boundedness of the family f implies the boundedness of the set $\mathcal{P}$.

Lemma 4.11. *Assume that the family f is bounded. Then,*

$$\mathscr{S}(t)u_0 - u_0 \leq \int_0^t \Sigma(s)\mathcal{Q}u_0 \,\mathrm{d}s$$

for all $u_0 \in \mathbb{R}^d$ and $t \geq 0$. Here, $(\Sigma(t))_{t\geq 0}$ is the Nisio semigroup w.r.t. the sublinear Q-operator $\mathbf{q}$ *from the previous remark, or more precisely, the Nisio semigroup w.r.t. $(\mathcal{P}, f)$, where $f_q = 0$ for all $q \in \mathcal{P}$.*

Proof. Let $u_0 \in \mathbb{R}^d$ and $h > 0$. Then, by (8), we have that

$$S_q(h)u_0 - u_0 = \int_0^h e^{sq}\big(qu_0 + f_q\big)\,\mathrm{d}s \le \int_0^h \Sigma(s)Qu_0\,\mathrm{d}s.$$

Notice that, by Lemma 4.5, the mapping $[0,\infty) \to \mathbb{R}^d$, $t \mapsto \Sigma(t)v_0$ is continuous and therefore locally integrable for all $v_0 \in \mathbb{R}^d$. Hence, for all $\tau \ge 0$,

$$\mathcal{E}_h u_0 - u_0 \le \int_0^h \Sigma(s)Qu_0\,\mathrm{d}s = \int_\tau^{\tau+h} \Sigma(s-\tau)Qu_0\,\mathrm{d}s. \tag{15}$$

Next, we show that

$$\mathcal{E}_\pi u_0 - u_0 \le \int_0^{\max\pi} \Sigma(s)Qu_0\,\mathrm{d}s \tag{16}$$

for all $\pi \in P$ by an induction on $m = |\pi|$, where $|\pi|$ denotes the cardinality of π. If $m = 1$, that is, if $\pi = \{0\}$, the statement is trivial. Hence, assume that

$$\mathcal{E}_{\pi'} u_0 - u_0 \le \int_0^{\max\pi'} \Sigma(s)Qu_0\,\mathrm{d}s$$

for all $\pi' \in P$ with $|\pi'| = m$ for some $m \in \mathbb{N}$. Let $\pi = \{t_0, t_1, \dots, t_m\} \in P$ with $0 = t_0 < t_1 < \cdots < t_m$ and $\pi' := \pi \setminus \{t_m\}$. Then, we obtain that

$$\begin{aligned}\mathcal{E}_\pi u_0 - \mathcal{E}_{\pi'} u_0 &\le \Sigma(t_{m-1})\big(\mathcal{E}_{t_m - t_{m-1}} u_0 - u_0\big)\\ &\le \Sigma(t_{m-1})\left(\int_{t_{m-1}}^{t_m} \Sigma(s - t_{m-1})Qu_0\,\mathrm{d}s\right) \le \int_{t_{m-1}}^{t_m} \Sigma(s)Qu_0\,\mathrm{d}s,\end{aligned}$$

where, in the second inequality, we used (15) with $h = t_m - t_{m-1}$ and $\tau = t_{m-1}$, and, in the last inequality, we used the sublinearity of $\Sigma(t)$. Using the induction hypothesis, we thus see that

$$\begin{aligned}\mathcal{E}_\pi u_0 - u_0 &= (\mathcal{E}_\pi u_0 - \mathcal{E}_{\pi'} u_0) + (\mathcal{E}_{\pi'} u_0 - u_0)\\ &\le \int_{t_{m-1}}^{t_m} \Sigma(s)Qu_0\,\mathrm{d}s + \int_0^{t_{m-1}} \Sigma(s)Qu_0\,\mathrm{d}s = \int_0^{\max\pi} \Sigma(s)Qu_0\,\mathrm{d}s.\end{aligned}$$

By (16), it follows that

$$\mathcal{E}_\pi u_0 - u_0 \le \int_0^t \Sigma(s)Qu_0\,\mathrm{d}s$$

for all $\pi \in P_t$. Taking the supremum over all $\pi \in P_t$, we obtain the assertion. □

The following proposition states that the Nisio semigroup $(\mathscr{S}(t))_{t\ge 0}$ is differentiable at zero if the family f is bounded.

Proposition 4.12. *Assume that f is bounded. Then, for all $u_0 \in \mathbb{R}^d$,*

$$\left\| \frac{\mathscr{S}(h)u_0 - u_0}{h} - Qu_0 \right\|_\infty \to 0, \quad h \searrow 0.$$

Proof. Since f is bounded, it follows that $\mathcal{P}$ is bounded (see Remark 4.3). Let $\varepsilon > 0$ and $u_0 \in \mathbb{R}^d$. Using Lemma 4.5, the boundedness of $\mathcal{P}$ and (9), there exists some $h_0 > 0$ such that, for all $0 < h \leq h_0$,

$$\begin{aligned} \left\| e^{hq}\big(qu_0 + f_q\big) - \big(qu_0 + f_q\big) \right\|_\infty &\leq \|e^{hq} - I_d\| \cdot \|qu_0 + f_q\|_\infty \\ &\leq \big(e^{\|q\|h} - 1\big)\|qu_0 + f_q\|_\infty \leq \varepsilon, \end{aligned}$$

for all $q \in \mathcal{P}$, and

$$\Sigma(h)Qu_0 - Qu_0 \leq \varepsilon.$$

Let $0 < h \leq h_0$. Then,

$$\mathscr{S}(h)u_0 - u_0 \geq S_q(h)u_0 - u_0 = \int_0^h e^{tq}\big(qu_0 + f_q\big)\,\mathrm{d}s \geq \big(qu_0 + f_q - \varepsilon\big)h$$

for all $q \in \mathcal{P}$. Dividing by h and taking the supremum over all $q \in \mathcal{P}$, it follows that

$$\frac{\mathscr{S}(h)u_0 - u_0}{h} \geq Qu_0 - \varepsilon. \tag{17}$$

Moreover, by Lemma 4.11,

$$\mathscr{S}(h)u_0 - u_0 - hQu_0 \leq \int_0^h \Sigma(s)Qu_0\,\mathrm{d}s - hQu_0 = \int_0^h (\Sigma(s)Qu_0 - Qu_0)\,\mathrm{d}s \leq h\varepsilon.$$

Dividing again by $h > 0$ yields

$$\frac{\mathscr{S}(h)u_0 - u_0}{h} - Qu_0 \leq \varepsilon,$$

which, together with (17), implies that

$$\left\| \frac{\mathscr{S}(h)u_0 - u_0}{h} - Qu_0 \right\|_\infty \leq \varepsilon.$$

□

Corollary 4.13. *Let f be bounded, $u_0 \in \mathbb{R}^d$, and $u(t) := \mathscr{S}(t)u_0$ for all $t \geq 0$. Then, $u \in C^1([0,\infty);\mathbb{R}^d)$ is the unique classical solution to the ODE*

$$u'(t) = Qu(t), \quad t \geq 0$$

with $u(0) = u_0$.

Proof. Let $u_0 \in \mathbb{R}^d$ and $t \geq 0$. Then, by Proposition 4.12,

$$\lim_{h \searrow 0} \frac{\mathscr{S}(t+h)u_0 - \mathscr{S}(t)u_0}{h} \lim_{h \searrow 0} \frac{\mathscr{S}(h)\mathscr{S}(t)u_0 - \mathscr{S}(t)u_0}{h} = Q\mathscr{S}(t)u_0.$$

This shows that the map $u : [0, \infty) \to \mathbb{R}^d,\ t \mapsto \mathscr{S}(t)u_0$ is continuous (see Lemma 4.5) and right differentiable with continuous right derivative

$$[0, \infty) \to \mathbb{R}^d, \quad t \mapsto Q\mathscr{S}(t)u_0,$$

where we used that the fact that $Q : \mathbb{R}^d \to \mathbb{R}^d$ is convex and thus continuous. Therefore, u is continuously differentiable with $u'(t) = Qu(t)$, for all $t \geq 0$, and $u(0) = u_0$. The Picard–Lindelöf theorem together with the local Lipschitz continuity of the convex map $Q : \mathbb{R}^d \to \mathbb{R}^d$ implies the uniqueness of u. □

Corollary 4.14. *Let f be bounded. Then, there exists some constant $L > 0$ such that*

$$\|\mathscr{S}(t)u_0 - u_0\|_\infty \leq Lt\|u_0\|_\infty$$

for all $t \geq 0$ and $u_0 \in \mathbb{R}^d$.

Proof. Since f is bounded, we have that $\mathcal{P}$ is bounded, and therefore, Q is Lipschitz continuous with Lipschitz constant $L := \sup_{q \in \mathcal{P}} \|q\|$. For all $u_0 \in \mathbb{R}^d$, we thus obtain that

$$\|\mathscr{S}(t)u_0 - u_0\|_\infty \leq \int_0^t \|Q\mathscr{S}(s)u_0\|_\infty \,\mathrm{d}s \leq \int_0^t L\|\mathscr{S}(s)u_0\|_\infty \,\mathrm{d}s \leq Lt\|u_0\|_\infty.$$

□

In order to end the proof of Theorem 2.5, we have to extend Corollary 4.13 to the unbounded case. We start with the following remark, which is the key observation in order to finish the proof of Theorem 2.5.

Remark 4.15. Let $\mathcal{P}^* := \{q \in \mathbb{R}^{d\times d} \mid Q^*(q) < \infty\}$ and $f_q^* := -Q^*(q)$ for all $q \in \mathcal{P}^*$, where Q^* is the conjugate function of Q (cf. Remark 3.1). For all $M \geq 0$, let $\mathcal{P}_M^*$ and Q_M be as in Remark 4.3 with $\mathcal{P}$ being replaced by $\mathcal{P}^*$. Moreover, let $(\mathscr{S}_M^*(t))_{t\geq 0}$ be the Nisio semigroup w.r.t. $(\mathcal{P}_M^*, (f_q^*)_{q\in\mathcal{P}_M^*})$ for $M \geq 0$. As

$$\bigcup_{M\geq 0} \mathcal{P}_M^* = \mathcal{P}^*,$$

it follows that $\mathscr{S}_M^*(t) \nearrow \mathscr{S}^*(t)$ as $M \to \infty$ for all $t \geq 0$, where $(\mathscr{S}^*(t))_{t\geq 0}$ is the Nisio semigroup w.r.t. $(\mathcal{P}^*, f^*)$. Let $R > 0$ be fixed. Then, there exists some $M_0 \geq 0$ such that $Qu = Q_{M_0}u$ for all $u \in \mathbb{R}^d$ with $\|u\|_\infty \leq R$, by choice of $\mathcal{P}^*$ and f^*. Let $u_0 \in \mathbb{R}^d$ with $\|u_0\|_\infty \leq R$. Then, it follows that $\|\mathscr{S}_M^*(t)u_0\|_\infty \leq R$ for all $t \geq 0$ and $M \geq 0$, which implies that $\mathscr{S}_M^*(t)u_0 = \mathscr{S}_{M_0}^*(t)u_0$ for all

$t \geq 0$ and $M \geq M_0$ by the uniqueness obtained from the Picard–Lindelöf theorem. In particular, $\mathscr{S}^*(t)u_0 = \mathscr{S}^*_{M_0}(t)u_0$ for all $t \geq 0$, which shows that the nonlinear ODE (3) has a unique classical solution $u^* \in C^1([0,\infty);\mathbb{R}^d)$ with $u^*(0) = u_0$. This solution is given by $u^*(t) = \mathscr{S}^*(t)u_0$ for all $t \geq 0$. By Corollary 4.14, we thus get that $\mathscr{S}^*(t) \to I$ as $t \searrow 0$ uniformly on compact sets.

We are now able to finish the proof of Theorem 2.5. The following proposition summarizes the results from this section, and proves the implication $(iii) \Rightarrow (iv)$.

Proposition 4.16. *The Nisio semigroup $(\mathscr{S}(t))_{t\geq 0}$ is a uniformly continuous convex Markovian semigroup and, for all $u_0 \in \mathbb{R}^d$, the function $u : [0,\infty) \to \mathbb{R}^d$, $t \mapsto \mathscr{S}(t)u_0$ is the unique classical solution $u \in C^1([0,\infty);\mathbb{R}^d)$ to the initial value problem*

$$u'(t) = \mathcal{Q}u(t), \quad t \geq 0,$$
$$u(0) = u_0.$$

Moreover, the Nisio semigroup $\mathscr{S}$ of $(\mathcal{P}, f)$ coincides with the semigroup envelope of $(\mathcal{P}, f)$.

Proof. In view of Proposition 4.9, it remains to show that the Nisio semigroup is uniformly continuous and that u is the unique solution to the ODE $u' = \mathcal{Q}u$ with initial value $u(0) = u_0$. By Remark 4.15, the initial value problem

$$u'(t) = \mathcal{Q}u(t), \quad t \geq 0,$$
$$u(0) = u_0,$$

has a unique classical solution $u^* \in C^1([0,\infty);\mathbb{R}^d)$, which is given by

$$u^*(t) := \mathscr{S}^*(t)u_0 \quad \text{for all } t \geq 0.$$

We show that $u^*(t) = \mathscr{S}(t)u_0$, for all $t \geq 0$. Let $R := \|u_0\|_\infty$. For all $M \geq 0$, let $(\mathcal{P}_M, f_M)$, $\mathcal{Q}_M$ and $\mathscr{S}_M = (\mathscr{S}_M(t))_{t\geq 0}$ be as in Remark 4.3. Let $\varepsilon > 0$. Since $\mathcal{Q} : \mathbb{R}^d \to \mathbb{R}^d$ is convex, it is locally Lipschitz. Hence, by Dini's lemma, there exists some $M_0 \geq 0$ such that

$$\|\mathcal{Q}v_0 - \mathcal{Q}_{M_0}v_0\|_\infty \leq \varepsilon$$

for all $v_0 \in \mathbb{R}^d$ with $\|v\|_\infty \leq R$. Further, there exists some constant $L > 0$ such that

$$\|\mathcal{Q}v_1 - \mathcal{Q}v_2\|_\infty \leq L\|v_1 - v_2\|_\infty$$

for all $v_1, v_2 \in \mathbb{R}^d$ with $\|v_1\|_\infty \leq R$ and $\|v_2\|_\infty \leq R$. Since $\|u^*(t)\|_\infty \leq R$ and $\|\mathscr{S}_M(t)u_0\|_\infty \leq R$ for all $M \geq 0$ and $t \geq 0$, we obtain that

$$\|\mathscr{S}_M(t)u_0 - u^*(t)\|_\infty = \left\| \int_0^t \mathcal{Q}_M\mathscr{S}_M(s)u_0 - \mathcal{Q}u^*(s)\,\mathrm{d}s \right\|_\infty$$

$$\leq \int_0^t \|\mathcal{Q}_M \mathscr{S}_M(s)u_0 - \mathcal{Q}u^*(s)\|_\infty \,\mathrm{d}s$$

$$\leq \int_0^t (\|\mathcal{Q}\mathscr{S}_M(s)u_0 - \mathcal{Q}u^*(s)\|_\infty + \varepsilon)\,\mathrm{d}s$$

$$\leq \int_0^t L\|\mathscr{S}_M(s)u_0 - u^*(s)\|_\infty + \varepsilon \,\mathrm{d}s$$

for all $t \geq 0$ and $M \geq M_0$. By Gronwall's lemma, we thus get that

$$\|\mathscr{S}_M(t)u_0 - u^*(t)\|_\infty \leq \varepsilon t e^{Lt}$$

for all $t \geq 0$ and $M \geq M_0$, showing that $\mathscr{S}_M(t)u_0 \to u^*(t)$ as $M \to \infty$ for all $t \geq 0$. However, since $\mathscr{S}_M(t)u_0 \nearrow \mathscr{S}(t)u_0$ as $M \to \infty$ for all $t \geq 0$, we obtain that $u^*(t) = \mathscr{S}(t)u_0$. This shows that $\mathscr{S}(t) = \mathscr{S}^*(t)$ for all $t \geq 0$, which, together with Remark 4.15, implies that $\mathscr{S}(t) = \mathscr{S}^*(t) \to I$ uniformly on compact sets as $h \searrow 0$. This ends the proof of this proposition and also the proof of the implication $(iii) \Rightarrow (iv)$ in Theorem 2.5. □

The remaining implication $(iv) \Rightarrow (v)$ is a direct consequence of (Denk et al., 2018, Theorem 5.6), which we summarize in the following proposition.

Proposition 4.17. *Let $\mathscr{S}$ be a uniformly continuous convex Markovian semigroup. Then, there exists a convex Markov chain $(\Omega, \mathcal{F}, \mathcal{E}, (X_t)_{t\geq 0})$ such that*

$$(\mathscr{S}(t)u_0)_i = \mathcal{E}_i(u_0(X_t))$$

for all $u_0 \in \mathbb{R}^d$, $t \geq 0$ and $i \in \{1, \dots, d\}$.

Restricting the time parameter of this process to $\mathbb{N}_0$ leads to a discrete-time Markov chain with transition operator $\mathscr{S}(1)$ (cf. Denk et al., 2018, Example 5.3). We conclude this section with the following remark, where we derive a dual representation of the semigroup envelope.

Remark 4.18. We will now derive a dual representation of the semigroup envelope by viewing the semigroup envelope as the cost functional of an optimal control problem, where, roughly speaking, "nature" tries to control the system into the worst possible scenario (using controls within the set $\mathcal{P}$). For $q = (q^1, \dots q^d) \in \mathcal{P}^d$ and $t \geq 0$, let $S_q(t) \in \mathbb{R}^{d\times d}$ be given by

$$\left(S_q(t)u_0\right)_i := \left(S_{q^i}(t)u_0\right)_i \tag{18}$$

for all $u_0 \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$. That is, $S_q(t)$ is the matrix whose ith row is the ith row of $S_{q^i}(t)$ for all $i \in \{1, \dots, d\}$. Here, the interpretation is that, in every state $i \in \{1, \dots, d\}$, "nature" is allowed to choose a different model $q \in \mathcal{P}$. We now add a dynamic component, and define

$$Q_t := \left\{ (q_k, h_k)_{k=1,\dots,m} \in \left(\mathcal{P}^d \times [0,t]\right)^m \,\middle|\, m \in \mathbb{N},\ \sum_{k=1}^m h_k = t \right\}.$$

Roughly speaking, Q_t corresponds to the set of all (space-time discrete) admissible controls for the control set $\mathcal{P}$. For an admissible control $\theta = (q_k, h_k)_{k=1,\ldots,m} \in Q_t$ with $m \in \mathbb{N}$ and $u_0 \in \mathbb{R}^d$, we then define

$$S_\theta u_0 := S_{q_1}(h_1) \cdots S_{q_m}(h_m) u_0,$$

where $S_{q_k}(h_k)$ is defined as in (18) for $k = 1, \ldots, m$. Then, for all $u_0 \in \mathbb{R}^d$,

$$\mathscr{S}(t)u_0 = \sup_{\pi \in P_t} \mathcal{E}_\pi u_0 = \sup_{\theta \in Q_t} S_\theta u_0. \tag{19}$$

In fact, by definition of Q_t, it follows that $S_q(t)u_0 \leq \sup_{\theta \in Q_t} S_\theta u_0 \leq \mathscr{S}(t)u_0$ for all $q \in \mathcal{P}$, $t \geq 0$ and $u_0 \in \mathbb{R}^d$. On the other hand, one readily verifies that $\mathscr{T}(t)u_0 := \sup_{\theta \in Q_t} S_\theta u_0$, for $t \geq 0$ and $u_0 \in \mathbb{R}^d$, gives rise to a semigroup $(\mathscr{T}(t))_{t \geq 0}$. Since $(\mathscr{S}(t))_{t \geq 0}$ is the semigroup envelope of $(\mathcal{P}, f)$, it follows that $\mathscr{T}(t) = \mathscr{S}(t)$ for all $t \geq 0$.

5 | COMPUTATION OF PRICE BOUNDS UNDER MODEL UNCERTAINTY

In this section, we demonstrate how price bounds for European contingent claims under uncertainty can be computed numerically in certain scenarios, first, via the explicit primal/dual description (19) of the semigroup envelope and, second, by solving the pricing ODE (3). Throughout, we consider two Q-matrices $q_0 \in \mathbb{R}^{d \times d}$ and $q \in \mathbb{R}^{d \times d}$ and, for $\lambda_l, \lambda_h \in \mathbb{R}$ with $\lambda_l \leq \lambda_h$, the interval $[\lambda_l, \lambda_h]$. Then, we consider the *Q-operator* $\mathcal{Q} : \mathbb{R}^d \to \mathbb{R}^d$ given by

$$\mathcal{Q}u_0 := q_0 u_0 + \max_{\lambda \in [\lambda_l, \lambda_h]} \lambda q u_0 \quad \text{for all } u_0 \in \mathbb{R}^d.$$

Then, by Example 3.4, $\mathcal{Q}$ is sublinear and has the (minimal) dual representation $(\{q_0 + \lambda_l q, q_0 + \lambda_h q\}, (0, 0))$. Choosing the latter as a dual representation as in Theorem 2.5 (iii), we may compute $\mathcal{Q}$ and $\mathcal{E}_h$, for $h \geq 0$, via

$$\mathcal{Q}u_0 = \max_{\lambda = \lambda_l, \lambda_h} \lambda q_0 u_0 + q u_0 \quad \text{for all } u_0 \in \mathbb{R}^d, \tag{20}$$

and

$$\mathcal{E}_h u_0 = \max_{\lambda = \lambda_l, \lambda_h} e^{h(q_0 + \lambda q)} u_0 \quad \text{for all } u_0 \in \mathbb{R}^d. \tag{21}$$

In the sequel, we use (20) and (21) in order to compute upper bounds for prices of European contingent claims under uncertainty. Replacing the maximum by a minimum in (20) and (21), we obtain lower bounds for the prices. In the examples, we consider, for suitable $\delta > 0$, the rate

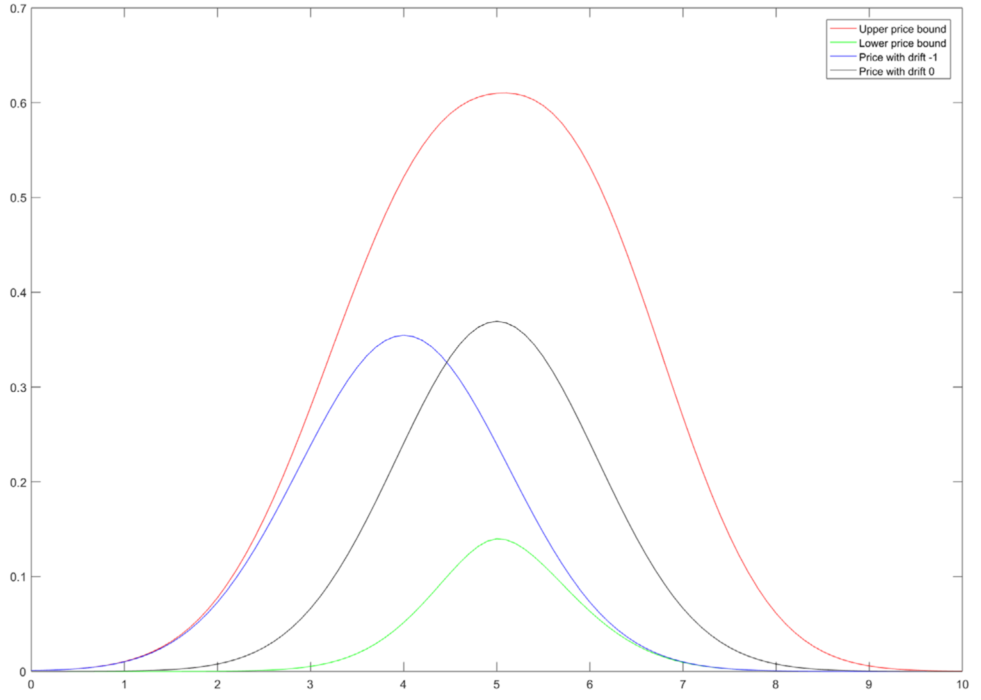

FIGURE 1 Upper and lower price bounds for a butterfly spread (24) with $K = 4$ and $L = 5$ under drift uncertainty depending on the current price in red and green, respectively. In blue and black, we see the value of the butterfly in the Bachelier model with drift -1 and 0, respectively [Color figure can be viewed at wileyonlinelibrary.com]

matrix

$$a := \frac{1}{\delta^2}\begin{pmatrix} -1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -2 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 & -2 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 1 & -1 \end{pmatrix}, \tag{22}$$

which is a discretization of the second space derivative with Neumann boundary conditions, and the rate matrix

$$b := \frac{1}{\delta}\begin{pmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 & 0 \\ 0 & \cdots & 0 & 0 & -1 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix} \tag{23}$$

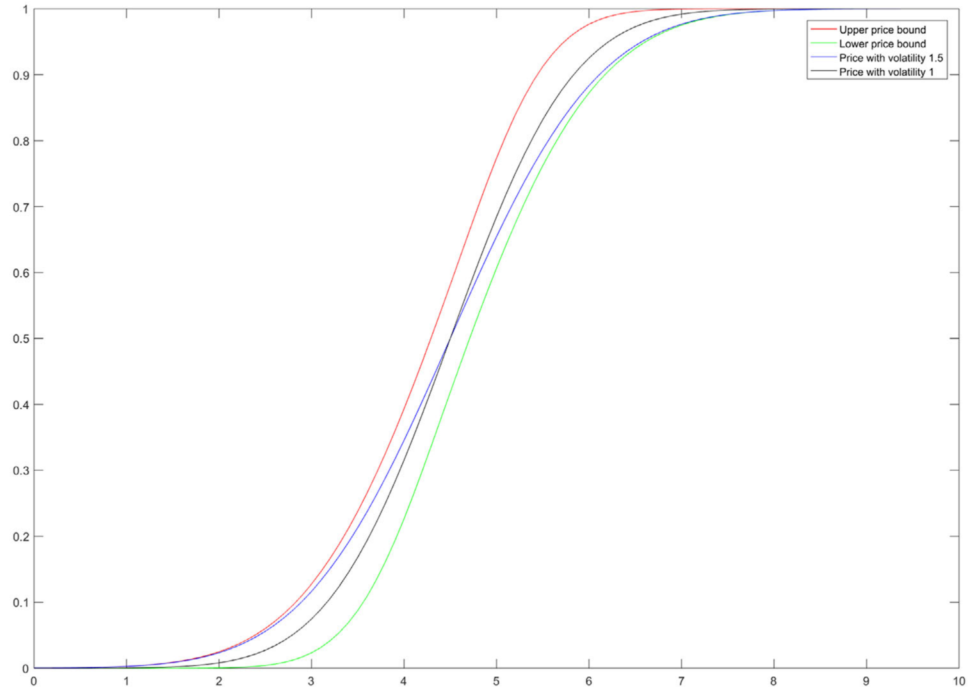

FIGURE 2 Upper and lower price bounds for a bull spread (25) with $K = 4$ and $L = 5$ under volatility uncertainty depending on the current price in red and green, respectively. In black and blue, we see the value of the butterfly in the Bachelier model with drift 1 and 1.5, respectively [Color figure can be viewed at wileyonlinelibrary.com]

as a discretization of the first space derivative. Then, the rate matrix

$$\frac{\sigma^2}{2}a + \mu b, \quad \text{for } \sigma > 0 \text{ and } \mu \in \mathbb{R},$$

is a finite-difference discretization of $\frac{\sigma^2}{2}\partial_{xx} + \mu\partial_x$, which is the generator of a Brownian motion with volatility σ and drift μ.

We start with an example, where we demonstrate how the semigroup envelope can be computed by solving the nonlinear pricing ODE (3). In the following example, we compute the upper and lower semigroup envelope for a discretized version of a Brownian motion (Bachelier model) with drift or volatility uncertainty. The solutions resemble the price bounds resulting from the parameter uncertainty of the underlying asset (the discretized version of a Brownian Motion) for a particular European contingent claim with fixed maturity as a function of the current price of the underlying asset.

Example 5.1. In this example, we compute the semigroup envelope $(\mathscr{S}(t))_{t\geq 0}$ by solving the ODE $u' = \mathcal{Q}u\, u(0) = u_0 \in \mathbb{R}^d$ with the explicit Euler method. The latter could be replaced by any other Runge–Kutta method. We consider the case, where $d = 101$, $\delta = \frac{1}{10}$. The state space is $S = \{i\delta \mid i \in \{0, \dots, 100\}\}$, which as a discretization of the interval [0,10], the maturity is $t = 1$, and we choose 1,000 time steps in the explicit Euler method. We consider the following two examples.

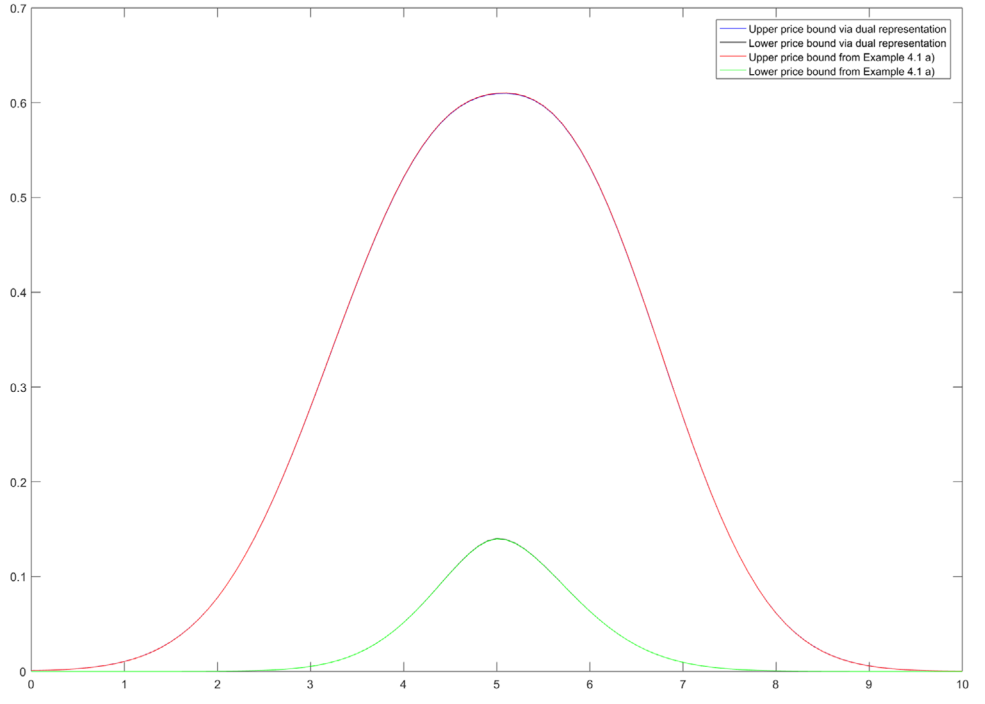

FIGURE 3 Upper and lower price bounds for a butterfly spread (24) with $K = 4$ and $L = 5$ under drift uncertainty from Example 5.1(a) in red and green, respectively. In blue and black, the upper and lower price bounds, computed via (26), respectively [Color figure can be viewed at wileyonlinelibrary.com]

(a) Let Q be given by (20) with $q_0 := a$, $q := b$, $\lambda_l := -1$, and $\lambda_h := 1$, that is, we consider the case of an uncertain drift parameter in the interval $[-1, 1]$. We price a butterfly spread, which is given by

$$u_0(x) = (L - K - |x - L|)^+, \quad \text{for } x = i\delta \text{ and } i \in \{1, \dots, 100\}, \tag{24}$$

with $K = 4$ and $L = 5$. In Figure 1, we depict the upper and lower price bounds as well as the prices corresponding to the Bachelier model with drift -1 and 0 in blue and black, respectively.

(b) Now, let $q_0 := 0$, $q := a$, $\lambda_l := 0.5$, and $\lambda_h := 1.5$ in (20). That is, we consider the case of an uncertain volatility in the interval [0.5,1.5]. We price a bull spread

$$u_0(x) = \min\left\{(x - K)^+, L - K\right\}, \quad \text{for } x = i\delta \text{ and } i \in \{1, \dots, 100\}, \tag{25}$$

with $K = 4$ and $L = 5$. In Figure 2, we see the upper and lower price bounds as well as the prices corresponding to the Bachelier model with volatilities 1 and 1.5 in black and blue, respectively.

The following example presents a second algorithm, using the primal/dual representation of the semigroup envelope, for the computation of price bounds for European contingent claims

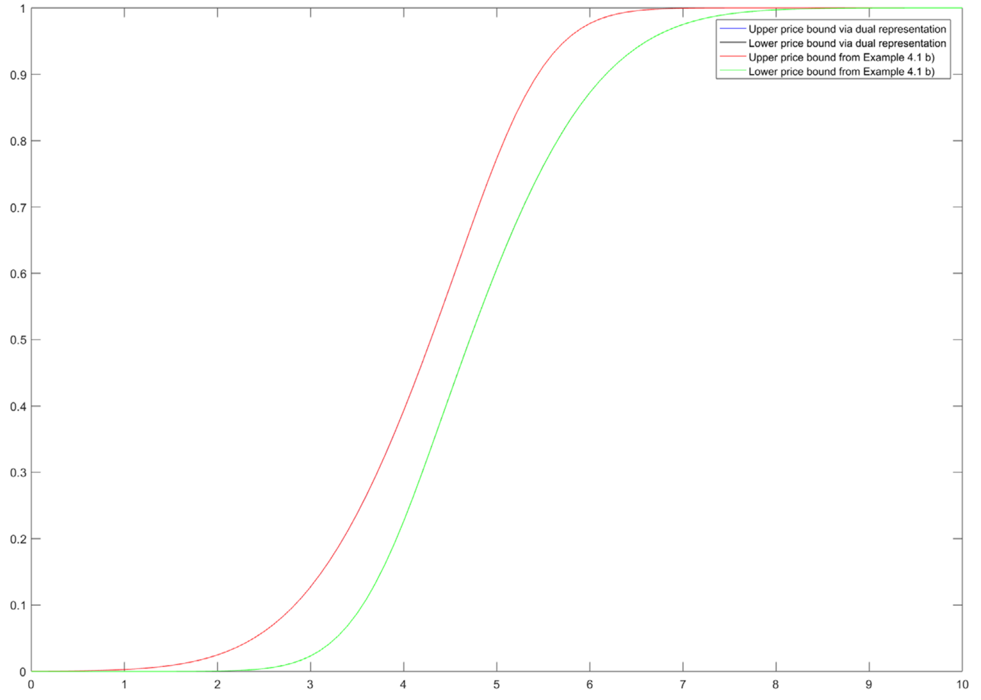

FIGURE 4 Upper and lower price bounds for a bull spread (25) with $K = 4$ and $L = 5$ under volatility uncertainty from Example 5.1(b) in red and green, respectively. In blue and black, the upper and lower price bounds, computed via (26), respectively [Color figure can be viewed at wileyonlinelibrary.com]

under model uncertainty. We compare the results with the ones from the previous example, which were obtained using Euler's method.

Example 5.2. For a fixed maturity $t \geq 0$, we consider the partitions

$$\pi_n := \{k2^{-n}t \mid k = 0, \dots, 2^n\}, \quad \text{for } n \in \mathbb{N}_0,$$

of the time interval $[0, t]$. We are then able to approximate the upper bound for prices of European contingent claims under uncertainty with maturity $t = 1$ by computing, for $n \in \mathbb{N}_0$ sufficiently large,

$$\underbrace{\mathcal{E}_{2^{-n}t} \cdots \mathcal{E}_{2^{-n}t}}_{2^n\text{-times}} u \tag{26}$$

with $\mathcal{E}_h$ given by (21) for $h \geq 0$. The fundamental system $e^{h(q_0+\lambda q)}$, for $\lambda = \lambda_h, \lambda_l$, appearing in (21) can either be computed via the Jordan decomposition of $q_0 + \lambda q$, by the approximation

$$\left(I + \frac{h}{k}(q_0 + \lambda q)\right)^k \tag{27}$$

with $k \in \mathbb{N}_0$ sufficiently large or by numerically solving the matrix-valued ODE

$$U' = (q_0 + \lambda q)U \quad \text{with} \quad U(0) = I,$$

where $I = I_d$ is the $d \times d$-identity matrix. We illustrate the approximation of the semigroup envelope via (26) in the following two examples, where a and b are given by (22) and (23). Again, we consider the case, where, $d = 101$, $\delta = \frac{1}{10}$ and the maturity is $t = 1$. In both examples, we choose $n = 10$, that is, we consider the partition π_{10} with $t = 1$, and use (27) with $k = 10$ for the computation of $e^{h(q_0+\lambda q)}$ for $\lambda = \lambda_h, \lambda_l$.

(a) As in Example 5.1(a), let $q_0 := a, q := b, \lambda_l := -1$ and $\lambda_h := 1$. Again, we compute the price of a butterfly spread, which is given by (24) with $K = 4$ and $L = 5$. In Figure 3, we see the upper and lower price curves from the previous example as well as the price bounds computed in this example. We observe that the price bounds match very well.
(b) We consider the case of an uncertain volatility parameter from Example 5.1(b), that is, let $q_0 := 0, q := a, \lambda_l := 0.5$, and $\lambda_h := 1.5$. As in Example 5.1(b), we price a bull spread given by (25) with $K = 4$ and $L = 5$. In Figure 4, we again depict the upper and lower price bounds from the previous example and this example. As in part (a), we observe that the price bounds perfectly match. □

ACKNOWLEDGMENTS

The author would like to thank Robert Denk and Michael Kupper for their valuable comments, support, and suggestions related to this work. Moreover, the author expresses his gratitude toward three anonymous referees for their helpful suggestions and remarks. Financial support through the German Research Foundation via CRC 1283 is gratefully acknowledged.

Open access funding enabled and organized by Projekt DEAL.

ORCID

Max Nendel https://orcid.org/0000-0002-9253-9518

REFERENCES

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203–228.

Avellaneda, M., & Buff, R. (1999). Combinatorial implications of nonlinear uncertain volatility models: The case of barrier options. *Applied Mathematical Finance*, *6*(1), 1–18.

Avellaneda, M., Levy, A., & Parás, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, *2*(2), 73–88.

Avellaneda, M., & Parás, A. (1996). Managing the volatility risk of portfolios of derivative securities: The lagrangian uncertain volatility model. *Applied Mathematical Finance*, *3*(1), 21–52.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Rustichini, A. (2014). Niveloids and their extensions: Risk measures on small domains. *Journal of Mathematical Analysis and Applications*, *413*(1), 343–360.

Cheridito, P., Soner, H. M., Touzi, N., & Victoir, N. (2007). Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Communications on Pure and Applied Mathematics*, *60*(7), 1081–1110.

Cohen, S. N. (2012). Representing filtration consistent nonlinear expectations as g-expectations in general probability spaces. *Stochastic Processes and their Applications*, *122*(4), 1601–1626.

Cohen, S. N., & Elliott, R. J. (2008). Solutions of backward stochastic differential equations on Markov chains. *Communications on Stochastic Analysis*, *2*(2), 251–262.

Cohen, S. N., & Elliott, R. J. (2010a). Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. *Annals of Applied Probability*, *20*(1), 267–311.

Cohen, S. N., & Elliott, R. J. (2010b). A general theory of finite state backward stochastic difference equations. *Stochastic Processes and their Applications*, *120*(4), 442–466.

Cohen, S. N., & Hu, Y. (2013). Ergodic BSDEs driven by Markov chains. *SIAM Journal on Control and Optimization*, *51*(5), 4138–4168.

Cohen, S. N., & Szpruch, L. (2012). On Markovian solutions to Markov chain BSDEs. *Numerical Algebra, Control and Optimization*, *2*(2), 257–269.

Coquet, F., Hu, Y., Mémin, J., & Peng, S. (2002). Filtration-consistent nonlinear expectations and related *g*-expectations. *Probability Theory and Related Fields*, *123*(1), 1–27.

de Cooman, G., Hermans, F., & Quaeghebeur, E. (2009). Imprecise Markov chains and their limit behavior. *Probability in the Engineering and Informational Sciences*, *23*(4), 597–635.

Delbaen, F. (2000). *Coherent risk measures*. Cattedra Galileiana. [Galileo Chair]. Pisa: Scuola Normale Superiore, Classe di Scienze.

Delbaen, F. (2002). Coherent risk measures on general probability spaces. In K. Sandmann & P. J. Schönbucher (Eds.), *Advances in finance and stochastics* (pp. 1–37). Berlin: Springer.

Dellacherie, C., & Meyer, P.-A. (1978). *Probabilities and potential*, Volume 29 of *North-Holland Mathematics Studies*. Amsterdam/New York: North-Holland Publishing Co.

Denis, L., Hu, M., & Peng, S. (2011). Function spaces and capacity related to a sublinear expectation: Application to G-Brownian motion paths. *Potential Analysis*, *34*(2), 139–161.

Denk, R., Kupper, M., & Nendel, M. (2018). Kolmogorov-type and general extension results for nonlinear expectations. *Banach Journal of Mathematical Analysis*, *12*(3), 515–540.

Denk, R., Kupper, M., & Nendel, M. (2020). A semigroup approach to nonlinear Lévy processes. *Stochastic Processes and their Applications*, *130*(3), 1616–1642.

Dentcheva, D., & Ruszczyński, A. (2018). Time-coherent risk measures for continuous-time Markov chains. *SIAM Journal on Financial Mathematics*, *9*(2), 690–715.

El Karoui, N., Peng, S., & Quenez, M. C. (1997). Backward stochastic differential equations in finance. *Mathematical Finance*, *7*(1), 1–71.

Fan, J., & Ruszczyński, A. (2018a). Process-based risk measures and risk-averse control of discrete-time systems. *Mathematical Programming*, 1–28. https://doi.org/10.1007/s10107-018-1349-2

Fan, J., & Ruszczyński, A. (2018b). Risk measurement and risk-averse control of partially observable discrete-time Markov systems. *Mathematical Methods of Operations Research*, *88*(2), 161–184.

Föllmer, H., & Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, *6*(4), 429–447.

Föllmer, H., & Schied, A. (2011). *Stochastic finance: An introduction in discrete time (extended ed.)*. Berlin: Walter de Gruyter & Co.

Frittelli, M., & Rosazza Gianin, E. (2002). Putting order in risk measures. *Journal of Banking & Finance*, *26*(7), 1473–1486.

Hartfiel, D. J. (1998). *Markov set-chains*, Volume 1695 of Lecture Notes in Mathematics. Berlin: Springer-Verlag.

Hu, M., & Peng, S. (2009). G-Lévy processes *under sublinear expectations*. Preprint arXiv:0911.3533.

Jacob, N. (2001). Fourier analysis and semigroups. In N. Jacob (Ed.), *Pseudo differential operators and Markov processes. Vol. I*. London: Imperial College Press.

Krak, T., De Bock, J., & Siebes, A. (2017). Imprecise continuous-time Markov chains. *International Journal of Approximate Reasoning*, *88*, 452–528.

Nendel, M., & Röckner, M. (2019). *Upper envelopes of families of Feller semigroups and viscosity solutions to a class of nonlinear Cauchy problems*. Preprint arXiv:1906.04430.

Neufeld, A., & Nutz, M. (2017). Nonlinear Lévy processes and their characteristics. *Transactions of the American Mathematical Society*, *369*(1), 69–95.

Nisio, M. (1976/77). On a non-linear semi-group attached to stochastic optimal control. *Publications of the Research Institute for Mathematical Sciences*, *12*(2), 513–537.

Norris, J. R. (1998). *Markov chains, Volume 2 of Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge University Press. Reprint of 1997 original.

Pardoux, E., & Peng, S. G. (1990). Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, *14*(1), 55–61.

Pardoux, E., & Peng, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. In B. L. Rozovskii & R. B. Sowers (Eds.), *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of Lecture Notes Control of Information Science (pp. 200–217). Berlin: Springer.

Pelessoni, R., & Vicig, P. (2003). Convex imprecise previsions. *Reliable Computing: An International Journal Devoted to Reliable Mathematical Computations based on Finite Representations and Guaranteed Accuracy*, *9*, 465–485. Special issue on dependable reasoning about uncertainty.

Pelessoni, R., & Vicig, P. (2005). Uncertainty modelling and conditioning with convex imprecise previsions. *International Journal of Approximate Reasoning*, *39*(2–3), 297–319.

Peng, S. (2005). Nonlinear expectations and nonlinear Markov chains. *Chinese Annals of Mathematics, Series B*, *26*(2), 159–184.

Peng, S. (2007). G-expectation, G-Brownian motion and related stochastic calculus of Itô type. In A. Truman & D. Williams (Eds.), *Stochastic analysis and applications*, volume 2 of Abel Symposium (pp. 541–567). Berlin: Springer.

Peng, S. (2008). Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Processes and their Applications*, *118*(12), 2223–2253.

Ruszczyński, A. (2010). Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming*, *125*(2, Ser. B), 235–261.

Škulj, D. (2009). Discrete time Markov chains with interval probabilities. *International Journal of Approximate Reasoning*, *50*(8), 1314–1329.

Škulj, D. (2015). Efficient computation of the bounds of continuous time imprecise Markov chains. *Applied Mathematics and Computation*, *250*, 165–180.

Vorbrink, J. (2014). Financial markets with volatility uncertainty. *Journal of Mathematical Economics*, *53*, 64–78.

Vovk, V., & Shafer, G. (2014). Game-theoretic probability. In T. Augustin, F. P. A. Coolen, G. de Cooman, & M. C. M. Troffaes (Eds.), *Introduction to imprecise probabilities*, Wiley Series in Probability and Statistics (pp. 114–134). Chichester: Wiley.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*, Volume 42 of Monographs on Statistics and Applied Probability. London: Chapman and Hall, Ltd.

How to cite this article: Nendel M. Markov chains under nonlinear expectation. *Mathematical Finance*. 2021;31:474–507. https://doi.org/10.1111/mafi.12289