

# Markov Chains under Nonlinear Expectation

Max Nendel

MAX.NENDEL@UNI-BIELEFELD.DE

Center for Mathematical Economics and CRC 1283, Bielefeld University, Germany

## 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 to a nonlinear setup, where the nonlinearity is due to parameter 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, continuous-time Markov chains under convex expectations and fully nonlinear ODEs. This extends a well-known characterization of  $Q$ -matrices. Moreover, the convex semigroup arising from a Markov process under a convex expectation is being described via the Fenchel-Legendre Transform of the generator.

**Keywords:** Nonlinear expectations, imprecise Markov chains, generators of nonlinear semigroups

## 1. Introduction and Main Result

We consider a finite non-empty 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  $S \rightarrow \mathbb{R}$  can therefore be identified by  $\mathbb{R}^d$ . A bounded random variable  $u$  thus will always be denoted as a vector of the form  $u = (u_1, \dots, u_d)^T \in \mathbb{R}^d$  identifying  $u(i) = u_i$  for  $i \in \{1, \dots, d\}$ . On  $\mathbb{R}^d$  we will always consider the norm

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

for a vector  $u \in \mathbb{R}^d$ . Moreover, for  $\alpha \in \mathbb{R}$  we denote by  $\alpha \in \mathbb{R}^d$  the constant vector  $u \in \mathbb{R}^d$  with  $u_i = \alpha$  for all  $i \in \{1, \dots, d\}$ . For a matrix  $a = (a_{ij})_{1 \leq i, j \leq d} \in \mathbb{R}^{d \times d}$ , we denote by  $\|a\|$  the operator norm of  $a: \mathbb{R}^d \rightarrow \mathbb{R}^d$  w.r.t. the norm  $\|\cdot\|_\infty$ , i.e.

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

Inequalities of vectors are always understood component-wise, i.e. for  $u, v \in \mathbb{R}^d$

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

All concepts in  $\mathbb{R}^d$  that include inequalities are to be understood w.r.t. the latter partial ordering. For example, a vector

field  $F: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is called convex if

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

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  $i \in \{1, \dots, d\}$ .

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

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

$$qu_0 = \lim_{h \searrow 0} \frac{\mathbb{E}(u_0(X_h)) - u_0}{h}$$

for all  $u_0 \in \mathbb{R}^d$ , where  $u_0(i)$  is the  $i$ -th component of  $u_0$  for  $i \in \{1, \dots, d\}$ ,  $\mathbb{P}_i$  stands for the probability measure under which the Markov chain  $(X_t)_{t \geq 0}$  satisfies  $\mathbb{P}_i(X_0 = i) = 1$  for  $i \in \{1, \dots, d\}$  and  $\mathbb{E}(Y) := (\mathbb{E}_{\mathbb{P}_1}(Y), \dots, \mathbb{E}_{\mathbb{P}_d}(Y))^T \in \mathbb{R}^d$  for any bounded random variable  $Y: \Omega \rightarrow \mathbb{R}$ .

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

$$\begin{aligned} u'(t) &= qu(t), \quad t \geq 0, \\ u(0) &= u_0, \end{aligned}$$

i.e.  $u(t) = e^{tq}u_0$  for all  $t \geq 0$ , where  $e^{tq}$  is the matrix exponential of  $tq$ . We refer to Norris [16] for a detailed illustration of this relation.

Here and throughout this paper, we say that a (possibly nonlinear) operator  $\mathcal{Q}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfies the *positive maximum principle* if for  $u = (u_1, \dots, u_d)^T \in \mathbb{R}^d$  and  $i \in \{1, \dots, d\}$  it holds  $(\mathcal{Q}u)_i \leq 0$  whenever  $u_i \geq u_j$  for all  $j \in \{1, \dots, d\}$ . This notion is motivated by the positive maximum principle for generators of Feller processes, see e.g. [12, 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. Indeed, property (iii) in the definition of a  $Q$ -matrix 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  $i \neq j$ . 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.

The notion of a nonlinear expectation was introduced by Peng [19]. Since then, nonlinear expectations have been widely used in order to describe model uncertainty in a probabilistic framework. Prominent examples of nonlinear expectations are the  $g$ -expectation, see Coquet et al. [2], and the  $G$ -expectation introduced by Peng [20],[21]. Given a measurable space  $(\Omega, \mathcal{F})$ , we consider the space  $\mathcal{L}^\infty(\Omega, \mathcal{F})$  of all bounded measurable functions  $\Omega \rightarrow \mathbb{R}$ . A *nonlinear expectation* is then a functional  $\mathcal{E}: \mathcal{L}^\infty(\Omega, \mathcal{F}) \rightarrow \mathbb{R}$  which satisfies  $\mathcal{E}(X) \leq \mathcal{E}(Y)$  whenever  $X(\omega) \leq Y(\omega)$  for all  $\omega \in \Omega$ , and  $\mathcal{E}(\alpha 1_\Omega) = \alpha$  for all  $\alpha \in \mathbb{R}$ . If  $\mathcal{E}$  is additionally convex, we say that  $\mathcal{E}$  is a *convex expectation*. It is well known, see e.g. [8] or [10], 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

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

for all  $X \in \mathcal{L}^\infty(\Omega, \mathcal{F})$ , where  $\mathbb{E}_{\mathbb{P}}$  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 possible models that are relevant for  $\mathcal{E}$ . 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,

i.e.  $\alpha_{\mathbb{P}}$  close or equal to 0, gives a lot of importance to the model  $\mathbb{P} \in \mathcal{P}$ .

If a nonlinear expectation  $\mathcal{E}$  is sublinear, then  $\rho(X) := \mathcal{E}(-X)$  defines a coherent monetary risk measure as introduced by Artzner et al. [1] and Delbaen [4],[5], see also Föllmer and Schied [10] for an overview of convex monetary risk measures. Moreover, if  $\mathcal{E}$  is a sublinear expectation, then  $\mathcal{E}$  is a coherent upper prevision cf. Walley [24] and vice versa. We would further like to mention that there is a one to one relation between the following three concepts: convex expectations, convex upper previsions cf. [17],[18] and convex risk measures cf. [10]. Another related concept are so-called (Choquet) capacities (see e.g. Dellacherie-Meyer [6]). However, in many applications the functional approach, using nonlinear expectations, has certain advantages, in particular regarding extension theorems or the existence of stochastic processes under model uncertainty, see e.g. Denk et al. [8].

In [19], Peng introduces a first notion of Markov chains under nonlinear expectations. 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 amongst others by Hartfiel [11], Škulj [22] and De Cooman et al. [3]. In [11], Hartfiel considers so-called Markov set-chains in discrete time, using matrix intervals in order to describe model uncertainty in the transition matrices. Later, Škulj [22] approached Markov chains under model uncertainty using Choquet capacities, which results in higher-dimensional matrices on the power set, while De Cooman et al. [3] considered imprecise Markov chains using an operator theoretic approach with upper and lower expectations. In [8, Example 5.3], model uncertainty in the transition matrix is being described by a transition operator, which 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 Škulj [23] or Krak et al. [13].

In this paper, we consider continuous-time Markov chains under convex expectations and extend the above relation between Markov chains,  $Q$ -matrices and ordinary differential equations to the convex case. This is done using convex duality, so-called Nisio semigroups (cf. Nisio [15]) and a convex version of Kolmogorov's extension theorem, see Denk et al. [8], which provides an extension to the whole path space. A similar approach has been used by Denk et al. [7] in order to construct Lévy processes under nonlinear expectations via solutions to fully nonlinear PDEs using Nisio semigroups. Restricting the time parame-

ter in the present work to the set of natural numbers leads to a discrete-time Markov chain, in the sense of [8, Example 5.3]. In order to state the main result, we need the following definitions:

**Definition 1** A (possibly nonlinear) map  $\mathcal{Q}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is called a  $Q$ -operator if the following conditions are satisfied:

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

**Definition 2** A convex Markov chain is a quadruple  $(\Omega, \mathcal{F}, \mathcal{E}, (X_t)_{t \geq 0})$ , where

- (i)  $(\Omega, \mathcal{F})$  is a measurable space.
- (ii)  $X_t: \Omega \rightarrow \{1, \dots, d\}$  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 make use of 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  $v_0 \in (\mathbb{R}^d)^{(n+1)}$  we have that

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

with  $Y := (X_{t_1}, \dots, X_{t_n})$  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  $\mathcal{E}: \mathcal{L}^\infty(\Omega, \mathcal{F}) \rightarrow \mathbb{R}^d$  is additively linear or sublinear, respectively.

The Markov property given in (iv) of the previous definition is the nonlinear analogon of the classical Markov property without using conditional expectations. Notice that due to the nonlinearity of the expectation, the definition and, in particular, the existence of a conditional (nonlinear) expectation is quite involved, which is why we avoid to introduce this concept.

**Definition 3** A family  $\mathcal{S} = (\mathcal{S}(t))_{t \geq 0}$  of (possibly nonlinear) operators  $\mathcal{S}(t): \mathbb{R}^d \rightarrow \mathbb{R}^d$  is called a semigroup if

- (i)  $\mathcal{S}(0) = I$ , where  $I = I_d$  is the  $d$ -dimensional identity matrix,

- (ii)  $\mathcal{S}(s+t) = \mathcal{S}(s)\mathcal{S}(t)$  for all  $s, t \geq 0$ .

- (iii) The mapping  $[0, \infty) \rightarrow \mathbb{R}^d$ ,  $t \mapsto \mathcal{S}(t)u_0$  is continuous for all  $u_0 \in \mathbb{R}^d$ .

Here and throughout, we make use of the notation  $\mathcal{S}(s)\mathcal{S}(t) := \mathcal{S}(s) \circ \mathcal{S}(t)$ . We call  $\mathcal{S}$  Markovian if

- (iv)  $\mathcal{S}(t)u_0 \leq \mathcal{S}(t)v_0$  for all  $t \geq 0$  and  $u_0, v_0 \in \mathbb{R}^d$  with  $u_0 \leq v_0$ ,

- (v)  $\mathcal{S}(t)\alpha = \alpha$  for all  $t \geq 0$  and  $\alpha \in \mathbb{R}$ , where we again use the notation  $\alpha := (\alpha, \dots, \alpha)^T \in \mathbb{R}^d$ .

We say that  $\mathcal{S}$  is linear, sublinear or convex if  $\mathcal{S}(t)$  is linear, sublinear or convex for all  $t \geq 0$ , respectively.

**Definition 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  $q_0 \in \mathcal{P}$ . We denote by

$$S_q(t)u_0 := e^{qt}u_0 + \int_0^t e^{qs}f_q ds = u_0 + \int_0^t e^{sq}(qu_0 + f_q) ds$$

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  $\mathcal{S}$  the Nisio semigroup or the semigroup envelope of  $(\mathcal{P}, f)$  if

- (i)  $\mathcal{S}(t)u_0 \geq S_q(t)u_0$  for all  $t \geq 0$ ,  $u_0 \in \mathbb{R}^d$  and  $q \in \mathcal{P}$ ,

- (ii) For any other semigroup  $\tilde{\mathcal{S}}$  satisfying (i) we have that  $\mathcal{S}(t)u_0 \leq \tilde{\mathcal{S}}(t)u_0$  for all  $t \geq 0$  and  $u_0 \in \mathbb{R}^d$ .

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

The following main theorem gives a full characterization of convex  $Q$ -operators.

**Theorem 5** Let  $\mathcal{Q}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a mapping. Then the following statements are equivalent:

- (i)  $\mathcal{Q}$  is a convex  $Q$ -operator.
- (ii)  $\mathcal{Q}$  is convex, satisfies the positive maximum principle and  $\mathcal{Q}\alpha = 0$  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  $\sup_{q \in \mathcal{P}} f_q = f_{q_0} = 0$  for some  $q_0 \in \mathcal{P}$  such that

$$\mathcal{Q}u_0 = \sup_{q \in \mathcal{P}} (qu_0 + f_q) \quad (1)$$

for all  $u_0 \in \mathbb{R}^d$ , where the suprema are to be understood componentwise.

- (iv) There exists a convex Markovian semigroup  $\mathcal{S}$  with

$$\mathcal{Q}u_0 = \lim_{h \searrow 0} \frac{\mathcal{S}(h)u_0 - u_0}{h}$$

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

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

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

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

$$u'(t) = \mathcal{Q}u(t) = \sup_{q \in \mathcal{P}} (qu(t) + f_q), \quad t \geq 0, \quad (2)$$

$$u(0) = u_0.$$

Moreover,  $u(t) = \mathcal{S}(t)u_0$  for all  $t \geq 0$  and  $\mathcal{S}$  is the Nisio semigroup w.r.t.  $(\mathcal{P}, f)$ .

**Remark 6** Consider the situation of Theorem 5.

- a) The dual representation in (iii) gives a model uncertainty interpretation to  $Q$ -operators. The set  $\mathcal{P}$  can be seen as the set of all plausible rate matrices, when considering the  $Q$ -operator  $\mathcal{Q}$ . For  $q \in \mathcal{P}$  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 has to exist at least one rate matrix  $q_0$  within the set of all plausible rate matrices  $\mathcal{P}$  to which we assign the maximal importance, that is minimal penalization.
- b) The Nisio semigroup  $\mathcal{S}$  w.r.t.  $(\mathcal{P}, f)$  can be constructed more explicitly. For details, we refer to Section 2. Moreover, notice that the Nisio semigroup  $\mathcal{S}$  can be constructed w.r.t. any dual representation  $(\mathcal{P}, f)$  as in (iii) and results in the unique classical solution of (2) independent of the choice of the representation  $(\mathcal{P}, f)$ .
- c) The same equivalence as in Theorem 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 5.
- d) Theorem 5 extends and includes the well-known relation between (linear) Markov chains,  $Q$ -matrices and ordinary differential equations.
- e) A consequence of Theorem 5 is that every convex Markovian semigroup, which is differentiable at 0, is the Nisio semigroup with respect to the Fenchel-Legendre transform (or any other dual representation as in (iii)) of its generator.

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

$$u'(t) = \mathcal{Q}u(t) \text{ for all } t \geq 0, \quad u(0) = u_0. \quad (3)$$

has a unique global solution.

g) Solutions to (3) remain bounded. Therefore, a Picard iteration or Runge-Kutta methods can be used for numerical computations and the convergence rate (depending on the size of the initial value  $u_0$ ) can be explicitly computed.

h) As in the linear case, by solving the differential equation (3) one can compute expressions of the form

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

under model uncertainty.

## 2. Proof of Theorem 5

Here, we only provide a proof of  $(v) \Rightarrow (ii) \Rightarrow (i) \Rightarrow (iii)$  and indicate the construction of the Nisio semigroup  $\mathcal{S}$ . For the remaining implications  $(iii) \Rightarrow (iv) \Rightarrow (v)$  and more details concerning the Nisio semigroup  $\mathcal{S}$ , we refer to [14] and [8].

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

$$\lambda p + (I - \lambda)q \in \mathcal{P},$$

where  $I = I_d$  is the  $d$ -dimensional identity matrix. Note that for all  $i \in \{1, \dots, d\}$  the  $i$ -th row of the matrix  $\lambda p + (I - \lambda)q$  is the convex combination of the  $i$ -th row of  $p$  and  $q$  with  $\lambda_i$ . Therefore, a set  $\mathcal{P}$  is row-convex if for all  $p, q \in \mathcal{P}$  the convex combination with different  $\lambda \in [0, 1]$  in every row is again an element of  $\mathcal{P}$ . Note that for example the set of all  $Q$ -matrices is row-convex.

**Remark 7** 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

$$\mathcal{P} := \{q \in \mathbb{R}^{d \times d} \mid \mathcal{Q}^*(q) < \infty\}$$

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  $\mathbb{R}^d$ . We refer to [9] or [10, Appendix A.1] for more details.

- a) The set  $\mathcal{P}$  is row-convex and the mapping  $\mathcal{P} \rightarrow \mathbb{R}$ ,  $q \mapsto f_q$  is continuous.



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

$$\mathcal{Q}_M: \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad u \mapsto \max_{q \in \mathcal{P}_M} (qu + f_q) \quad (4)$$

defines a convex operator which is Lipschitz continuous. Notice that the maximum in (4) is to be understood componentwise. However, for fixed  $u \in \mathbb{R}^d$  the maximum can be attained by a single element simultaneously in every component of  $\mathcal{P}_M$  since  $\mathcal{P}_M$  is row-convex, i.e. for all  $u \in \mathbb{R}^d$  there exists some  $q_0 \in \mathcal{P}_M$  with

$$\mathcal{Q}_M u = q_0 u + f_{q_0}.$$

c) Let  $R > 0$ . Then, there exists some  $M > 0$  such that

$$\mathcal{Q}u = \max_{q \in \mathcal{P}_M} (qu + f_q) = \mathcal{Q}_M u$$

for all  $u \in \mathbb{R}^d$  with  $\|u\|_\infty \leq R$ . In particular,  $\mathcal{Q}$  is locally Lipschitz continuous and admits a representation of the form

$$\mathcal{Q}u = \max_{q \in \mathcal{P}} (qu + f_q)$$

for all  $u \in \mathbb{R}^d$ , where for fixed  $u \in \mathbb{R}^d$  the maximum can be attained by a single element simultaneously in every component 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

(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  $v_0 := u_0 + \alpha$ . Then,

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

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

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

i.e.

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

which is a contradiction to

$$\|\mathcal{E}(v_0(X_h))\|_\infty \leq \|v_0\|_\infty.$$

This shows that  $\mathcal{Q}$  satisfies the positive maximum principle. (ii)  $\Rightarrow$  (i): This follows directly from the positive maximum principle, considering the vectors  $\lambda e_i$  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  $f = (f_q)_{q \in \mathcal{P}}$  as in Remark 7. Then, by Remark 7 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  $\alpha \in \mathbb{R}$  it holds

$$q\alpha = \frac{1}{\lambda} q(\lambda\alpha) \leq \frac{1}{\lambda} (\mathcal{Q}(\lambda\alpha) + \mathcal{Q}^*(q)) = \frac{1}{\lambda} \mathcal{Q}^*(q) \rightarrow 0,$$

as  $\lambda \rightarrow \infty$ , where we used Remark 7 c) in the second inequality. Therefore,  $q\alpha \leq 0$  for all  $\alpha \in \mathbb{R}$ . Since,  $q$  is linear, it follows  $q1 = 0$ . Now, let  $i \in \{1, \dots, d\}$ . Then, by Remark 7 c) and the definition of a  $Q$ -operator, we obtain that

$$q_{ii} \leq \frac{1}{\lambda} (\mathcal{Q}(\lambda e_i) + \mathcal{Q}^*(q))_i \leq \frac{1}{\lambda} (\mathcal{Q}^*(q))_i \rightarrow 0$$

as  $\lambda \rightarrow \infty$ , i.e.  $q_{ii} \leq 0$ . Now, let  $i, j \in \{1, \dots, d\}$  with  $i \neq j$ . Again by Remark 7 c) and the definition of a  $Q$ -operator, it follows that

$$-q_{ij} \leq \frac{1}{\lambda} (\mathcal{Q}(-\lambda e_i) + \mathcal{Q}^*(q))_j \leq \frac{1}{\lambda} (\mathcal{Q}^*(q))_j \rightarrow 0$$

as  $\lambda \rightarrow \infty$ , i.e.  $q_{ij} \geq 0$ . Therefore,  $q$  is a  $Q$ -matrix. ■

We conclude by constructing the Nisio semigroup  $\mathcal{S}$  w.r.t. the dual representation  $(\mathcal{P}, f)$  from the previous proof. For every  $q \in \mathcal{P}$ , we consider the linear ODE

$$u'(t) = qu(t) + f_q, \quad t \geq 0 \quad (5)$$

with  $u(0) = u_0 \in \mathbb{R}^d$ . Then, by variation of constant, the solution of (5) is given by

$$S_q(t)u_0 := e^{qt}u_0 + \int_0^t e^{qs}f_q ds = u_0 + \int_0^t e^{sq}(qu_0 + f_q) ds \quad (6)$$

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, i.e.

(i)  $S_q(0) = I$ , where  $I = I_d$  is the  $d$ -dimensional identity matrix,

(ii)  $S_q(s+t) = S_q(s)S_q(t)$  for all  $s, t \geq 0$ ,

(iii)  $\|S_q(t) - I\| \rightarrow 0$  as  $t \searrow 0$ .

### Remark 8

a) 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, i.e.

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

(ii)  $e^{tq}1 = 1$ .

In particular,  $e^{tq}u \leq e^{tq}v$  for all  $u, v \in \mathbb{R}^d$  with  $u \leq v$  and therefore, the semigroup  $S_q$  is monotone (see part b) below for a definition).

b) In line with [8, Definition 5.1], we say that a (possibly nonlinear) map  $\mathcal{E}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a kernel, if  $\mathcal{E}$  is monotone, i.e.  $\mathcal{E}(u) \leq \mathcal{E}(v)$  for all  $u, v \in \mathbb{R}^d$  with  $u \leq v$ , and  $\mathcal{E}$  preserves constants, i.e.  $\mathcal{E}(\alpha) = \alpha$  for all  $\alpha \in \mathbb{R}$ . By part a), it is clear that  $e^{tq} \in \mathbb{R}^{d \times d}$  is a linear kernel for all  $q \in \mathcal{P}$  and  $t \geq 0$ .

We consider the set of finite partitions

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

The set of partitions with end-point  $t$  will be denoted by  $P_t$ , i.e.  $P_t := \{ \pi \in P \mid \max \pi = t \}$ . Note that

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

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

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

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

$$S_q(h)u = e^{hq}u + \int_0^h e^{sq} f_q ds \leq e^{hq}u \leq \|u\|_\infty$$

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

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

since 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  $0 = t_0 < t_1 < \dots < t_m$ , 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  $\pi \in P$  being a concatenation of convex kernels.

**Proposition 9** The Nisio semigroup  $\mathcal{S} = (\mathcal{S}(t))_{t \geq 0}$  w.r.t.  $(\mathcal{P}, f)$  is given by

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

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

Note that  $\mathcal{S}(t): \mathbb{R}^d \rightarrow \mathbb{R}^d$  is well-defined and a convex kernel for all  $t \geq 0$  since  $\mathcal{E}_\pi$  is a convex kernel for all  $\pi \in P$ . For a proof of Proposition 9, we refer to [14].

## Acknowledgments

The author would like to thank Robert Denk and Michael Kupper for their valuable comments, support and suggestions related to this work. Financial support through the German Research Foundation via CRC 1283 ‘‘Taming Uncertainty’’ is gratefully acknowledged.

## References

- [1] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999. ISSN 0960-1627. doi:[10.1111/1467-9965.00068](https://doi.org/10.1111/1467-9965.00068).
- [2] Franois Coquet, Ying Hu, Jean Memin, and Shige Peng. Filtration-consistent nonlinear expectations and related  $g$ -expectations. *Probab. Theory Related Fields*, 123(1):1–27, 2002. ISSN 0178-8051. doi:[10.1007/s004400100172](https://doi.org/10.1007/s004400100172).
- [3] Gert De Cooman, Filip Hermans, and Erik Quaeghebeur. Imprecise Markov chains and their limit behavior. *Probab. Engrg. Inform. Sci.*, 23(4):597–635, 2009. ISSN 0269-9648. doi:[10.1017/S0269964809990039](https://doi.org/10.1017/S0269964809990039).
- [4] Freddy Delbaen. *Coherent risk measures*. Cattedra Galileiana. [Galileo Chair]. Scuola Normale Superiore, Classe di Scienze, Pisa, 2000.
- [5] Freddy Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, Berlin, 2002.
- [6] Claude Dellacherie and Paul-Andre Meyer. *Probabilities and potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam-New York; North-Holland Publishing Co., Amsterdam-New York, 1978. ISBN 0-7204-0701-X.
- [7] Robert Denk, Michael Kupper, and Max Nendel. A semigroup approach to nonlinear Levy processes. *Preprint*, 2017.
- [8] Robert Denk, Michael Kupper, and Max Nendel. Kolmogorov-type and general extension results for nonlinear expectations. *Banach J. Math. Anal.*, 12(3):515–540, 2018. ISSN 1735-8787. doi:[10.1215/17358787-2017-0024](https://doi.org/10.1215/17358787-2017-0024).
- [9] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. ISBN 0-89871-450-8. doi:[10.1137/1.9781611971088](https://doi.org/10.1137/1.9781611971088). Translated from the French.

- [10] Hans Föllmer and Alexander Schied. *Stochastic finance*. Walter de Gruyter & Co., Berlin, extended edition, 2011. ISBN 978-3-11-021804-6. doi:[10.1515/9783110218053](https://doi.org/10.1515/9783110218053). An introduction in discrete time.
- [11] Darald J. Hartfiel. *Markov set-chains*, volume 1695 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998. ISBN 3-540-64775-9. doi:[10.1007/BFb0094586](https://doi.org/10.1007/BFb0094586).
- [12] Niels Jacob. *Pseudo Differential Operators & Markov Processes: Markov Processes And Applications*, volume 3. Imperial College Press, 2001.
- [13] Thomas Krak, Jasper De Bock, and Arno Siebes. Imprecise continuous-time Markov chains. *Internat. J. Approx. Reason.*, 88:452–528, 2017. ISSN 0888-613X. doi:[10.1016/j.ijar.2017.06.012](https://doi.org/10.1016/j.ijar.2017.06.012).
- [14] Max Nendel. Markov chains under nonlinear expectation. *Preprint*, 2018.
- [15] Makiko Nisio. On a non-linear semi-group attached to stochastic optimal control. *Publ. Res. Inst. Math. Sci.*, 12(2):513–537, 1976/77. ISSN 0034-5318. doi:[10.2977/prims/1195190727](https://doi.org/10.2977/prims/1195190727).
- [16] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. ISBN 0-521-48181-3. Reprint of 1997 original.
- [17] Renato Pelessoni and Paolo Vicig. Convex imprecise previsions. *Reliable Computing*, 9(6):465–485, 2003.
- [18] Renato Pelessoni and Paolo Vicig. Uncertainty modelling and conditioning with convex imprecise previsions. *International Journal of Approximate Reasoning*, 39(2-3):297–319, 2005.
- [19] Shige Peng. Nonlinear expectations and nonlinear Markov chains. *Chinese Ann. Math. Ser. B*, 26(2):159–184, 2005. ISSN 0252-9599. doi:[10.1142/S0252959905000154](https://doi.org/10.1142/S0252959905000154).
- [20] Shige Peng.  $G$ -expectation,  $G$ -Brownian motion and related stochastic calculus of Itô type. In *Stochastic analysis and applications*, volume 2 of *Abel Symp.*, pages 541–567. Springer, Berlin, 2007. doi:[10.1007/978-3-540-70847-6\\_25](https://doi.org/10.1007/978-3-540-70847-6_25).
- [21] Shige Peng. Multi-dimensional  $G$ -Brownian motion and related stochastic calculus under  $G$ -expectation. *Stochastic Process. Appl.*, 118(12):2223–2253, 2008. ISSN 0304-4149. doi:[10.1016/j.spa.2007.10.015](https://doi.org/10.1016/j.spa.2007.10.015).
- [22] Damjan Škulj. Discrete time Markov chains with interval probabilities. *Internat. J. Approx. Reason.*, 50(8):1314–1329, 2009. ISSN 0888-613X. doi:[10.1016/j.ijar.2009.06.007](https://doi.org/10.1016/j.ijar.2009.06.007).
- [23] Damjan Škulj. Efficient computation of the bounds of continuous time imprecise Markov chains. *Appl. Math. Comput.*, 250:165–180, 2015. ISSN 0096-3003. doi:[10.1016/j.amc.2014.10.092](https://doi.org/10.1016/j.amc.2014.10.092).
- [24] Peter Walley. Statistical reasoning with imprecise probabilities. 1991.