Chapter 2
Dynamic Modeling of Dependence in Finance via Copulae Between Stochastic Processes

Tomasz R. Bielecki, Jacek Jakubowski and Mariusz Niewęgłowski

Abstract Modeling of stochastic dependence is crucial to pricing and hedging of basket derivatives, as well as to pricing and hedging of some other financial products, such as rating-triggered corporate step-up bonds. The classical approach to modeling of dependence in finance via static copulae (and Sklar's theorem) is inadequate for consistent valuation and hedging in time. In this survey we present recent developments in the area of modeling of dependence between stochastic processes with given marginal laws. Some of these results have already been successfully applied in finance in connection with the portfolio credit risk.

2.1 Introduction

Dynamic modeling of dependence between financial risks is crucial to achieving consistent calibration through time to market data, as well as to dynamic hedging of these risks.

The classical approaches to modeling dependence in finance were typically rooted in the static copula theory (see e.g. [7]). A standard example is the Gaussian copula model introduced by David Li [25], which was widely used by practitioners.

Tomasz R. Bielecki
Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL, USA
e-mail: bielecki@iit.edu

Jacek Jakubowski
Institute of Mathematics, University of Warsaw, Warszawa, Poland; Faculty of Mathematics and Information Science, Warsaw University of Technology, Warszawa, Poland
e-mail: jakub@mimuw.edu.pl

Mariusz Niewęgłowski
Faculty of Mathematics and Information Science, Warsaw University of Technology, Warszawa, Poland
e-mail: M.Nieweglowski@mini.pw.edu.pl

P. Jaworski et al. (eds.), *Copula Theory and Its Applications*, Lecture Notes in Statistics 198,
DOI 10.1007/978-3-642-12465-5_2,

But the static copulae models were not capable of effectively dealing with the dynamic aspects of dependence between financial risks. Some people went to the extreme of blaming the recent financial crisis on the use of static models (cf. e.g. [28]), accusations that definitely went too far (for a critique of these accusations see e.g. [14] or [33]).

Recently, an effort has been started to model dependence between financial risks in a dynamic way. The so called Lévy copulae were studied in Kallsen and Tankov [21]. Markov copulae were introduced in Bielecki, Jakubowski, Vidozzi and Vidozzi [4], and subsequently studied in Bielecki, Vidozzi and Vidozzi [6] and Bielecki, Jakubowski and Niewęgłowski [3]. Motivated by the results in [4] the so called semimartingale copulae were formally defined and studied in L. Vidozzi [32]. Stochastic dependence between the components of a multivariate Markov process in terms of its infinitesimal operator was investigated in A. Vidozzi [31]. A related, but different line of research devoted to modeling dynamic dependence between stopping times and applications to credit risk was originated in El Karoui, Jeanblanc and Jiao [12], where the conditional density approach is used.

In this chapter we shall describe an approach to dynamic modeling of dependence in finance, based on modeling of dependence between stochastic processes, and using the results from Bielecki et al. [4, 6], Vidozzi [32] and Bielecki et al. [3], Cont and Tankov [8] and Kallsen and Tankov [21]. In particular, we shall describe various ways of modeling dependence between stochastic processes so that the laws of individual components of a multivariate process agree with some prescribed laws. Therefore, with an abuse of terminology, we shall refer to relevant constructions as *semimartingale copulae* and *Markov copulae*. It needs to be stressed, though, that the term "copula" is used here for convenience and for its historical connotation only.

The objective of the methodology outlined in this article is different from that in Lageras [22], in which results of Darsow et al. [11] are extended. Those two papers aim at relating the classical concept of copula and the concept of Markov property. In this context they investigate dependence along the time line in the case of a one-dimensional Markov process, and characterize the Markov property in terms of copulae. Next, Ibragimov [17] generalized results of Darsow et al. [11] to higher order Markov processes. The problem that we present here is also different from [12], since their interest is not in building models with prescribed marginal laws.

Sections 2.3 and 2.4.2 are, for the most part, taken from [31, 32], respectively. The proofs are skipped though in this survey article. We refer the interested reader to [4, 31, 32] for a comprehensive treatment of the relevant topics. Analogous remarks apply to Sect. 2.4.3 based on [3], and Sect. 2.5 based on [5].

The paper is organized as follows. Section 2.2 describes Lévy copulae. Semimartingale copulae are defined and investigated in Sect. 2.3. We consider semimartingales that are uniquely characterized, in the sense of their probability laws, by their characteristics. We construct a process X whose i-th univariate law, i.e., the law of the i-th component X^i, is the same as the law of a given process Y^i, $i = 1, \ldots, n$. Section 2.4 is devoted to Markov copulae. We present two different approaches: generator based and symbolic. The first is based on infinitesimal generators, the second applies pseudo-differential operators. The last section presents an application to finance of copulae so defined.

2.2 Lévy Copulae

As is well known, the law of a multivariate Lévy process is entirely determined by any of its one-dimensional distributions. Thus, creating dependence between univariate components of a multivariate Lévy process essentially amounts to creating dependence between finite dimensional random variables. The problem is that if one wants to do this things in terms of the Lévy characteristics of the process, then one needs, among other, to create dependence between the marginal Lévy measures, which, in general, are not finite measures. This leads to certain technical difficulties that however were successfully dealt with in the papers by Tankov [30] and Kallsen and Tankov [21].

Tankov [30] introduced Lévy copulae to characterize dependence between components of a multidimensional Lévy process. His construction is for Lévy processes with positive jumps in every component. Later, Kallsen and Tankov [21] generalized this concept to arbitrary Lévy processes. A Lévy copula is a counterpart of the notion of copula for multivariate distributions. Copulae give a characterization of possible dependence structures of a random vector, given the margins, and allow one to construct a multidimensional distribution with specified dependence from a collection of one-dimensional distributions. Similarly, the aim of Lévy copulae is to provide a way to construct multivariate Lévy processes with given marginals.

Since the dependence structure of the Brownian motion part of a Lévy process X is characterized entirely by its covariance matrix, and since the Brownian motion part of X is independent of the jump part, it remains to describe the dependence structure of the purely discontinuous part of X, and this is done by means of Lévy copulae.

Now, we present formal definitions.

Definition 2.2.1. Let $\overline{\mathbb{R}} := (-\infty, +\infty]$. A function $F : \overline{\mathbb{R}}^d \to \mathbb{R}$ is d-increasing if $F(u_1, \ldots, u_d) \neq \infty$ for $(u_1, \ldots, u_d) \neq (\infty, \ldots, \infty)$ and

$$\sum_{c \in \{a_1, b_1\} \times \ldots \times \{a_d, b_d\}} (-1)^{N(c)} F(c) \geq 0$$

for any $-\infty < a_i \leq b_i \leq \infty$ and $N(c) := \#\{k : c_k = a_k\}$.

For a d-increasing function we can define margins in a similar way as for a probability distribution function. To do this, we set

$$\operatorname{sgn} x = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases}$$

Definition 2.2.2. Let F be a d-increasing function. For any nonempty index set $I \subset \{1,\dots,d\}$ the I-margin of F is the function $F^I : \overline{\mathbb{R}}^I \to \overline{\mathbb{R}}$ defined by

$$F^I((u_i)_{i\in I}) := \lim_{k\to\infty} \sum_{(u_j)_{j\in I^c}\in\{-k,\infty\}^{I^c}} F(u_1,\dots,u_d) \prod_{j\in I^c} \operatorname{sgn} u_j,$$

where $I^c := \{1,\dots,d\} \setminus I$.

In particular, for $I = \{i\}$, $F^{\{i\}}$, the i-th margin of F, is given by

$$F^{\{i\}}(x) := \lim_{c\to-\infty} \left(F(+\infty,\dots,+\infty,x,+\infty,\dots,+\infty) - F(c,\dots,c,x,c,\dots,c)\right).$$

Definition 2.2.3. A function $F : \overline{\mathbb{R}}^d \to \overline{\mathbb{R}}$ is called a *Lévy copula* if

1. $F(u_1,\dots,u_d) = 0$ if $u_i = 0$ for at least one $i \in \{1,\dots,d\}$,
2. F is d-increasing,
3. $F^{\{i\}}(u) = u$ for any $i \in \{1,\dots,d\}$, $u \in \mathbb{R}$.

It is worth noting that Lévy copulae have properties similar to ordinary copulae; in particular, they are Lipschitz continuous. In order to use Lévy copulae to investigate dependence between components of a general multivariate Lévy process we have to define a tail integral.

Definition 2.2.4. Let X be an $\mathbb{R}^d$-valued Lévy process with Lévy measure ν. The tail integral of X is the function $U : (\mathbb{R}\setminus\{0\})^d \to \mathbb{R}$ defined by

$$U(x_1,\dots,x_d) := \prod_{i=1}^{d} \operatorname{sgn}(x_i)\nu\left(\prod_{j=1}^{d} \mathfrak{I}(x_j)\right),$$

where, for $x \in \mathbb{R}$, we denote

$$\mathfrak{I}(x) := \begin{cases} (x,\infty), & x \geq 0, \\ (-\infty,x], & x < 0. \end{cases}$$

The tail integral does not determine the Lévy measure uniquely, in general, since it does not give any information about mass on coordinate axes.

This motivates introducing the *I-marginal tail integral* U^I, for nonempty set $I \subset \{1,\dots,d\}$. It is the tail integral of the Lévy process $(X^i)_{i\in I}$, or equivalently the tail integral of the I-marginal of the Lévy measure ν, that is, the measure ν^I defined by

$$\nu^I(A) := \nu(x \in \mathbb{R}^d : (x_i)_{i\in I} \in A \setminus \{0\}) \quad \text{for } A \in \mathscr{B}(\mathbb{R}^{|I|}).$$

It turns out that to determine the Lévy measure uniquely we have to know all marginal tail integrals, i.e., we have to know U^I for all $I \subset \{1,\dots,d\}$ (see [21, Lemma 3.5]). In fact, there is one-to-one correspondence between the Lévy measure and the set of all marginal tail integrals.

The most important feature of Lévy copulae is that they allow separating the margins and the dependence structure of Lévy measures. This is clear from the following counterpart of Sklar's theorem, proved by Kallsen and Tankov [21].

Theorem 2.2.1. *1) Let $X = (X_1,\dots,X_d)$ be an $\mathbb{R}^d$-valued Lévy process. Then there exists a Lévy copula F such that the tail integrals of X satisfy*

$$U^I\left((x_i)_{i\in I}\right) = F^I\left((U_i(x_i))_{i\in I}\right) \tag{2.1}$$

for any nonempty $I \subset \{1,\dots,d\}$ and any $(x_i)_{i\in I} \in (\mathbb{R}\setminus\{0\})^I$. The Lévy copula F is unique on $\prod_{i=1}^d \overline{\operatorname{Ran} U_i}$.
2) Let F be a d-dimensional Lévy copula and U_i, $i = 1,\dots,d$, be the tail integrals of real valued Lévy processes. Then there exists an $\mathbb{R}^d$-valued Lévy process whose components have tail integrals $U_1,\dots,U_d$ and whose marginal tail integral satisfies condition (2.1) *for any nonempty $I \subset \{1,\dots,d\}$ and any $x \in (\mathbb{R}\setminus\{0\})^I$. The Lévy measure ν of X is uniquely determined by F and U_i, $i = 1,\dots,d$.*

We now proceed with a few examples of Lévy copulae.

Example 2.2.1. In Kallsen and Tankov [21] it is shown that a pure jump Lévy process has independent coordinates if and only if its Lévy copula is given by the following formula:

$$F_\perp(x_1,\dots,x_d) := \sum_{i=1}^d x_i \prod_{j\neq i} 1_{\{\infty\}}(x_j).$$

Example 2.2.2. Kallsen and Tankov [21] introduce an *Archimedean Lévy copula*, analogously to an ordinary Archimedean copula, by setting

$$F(x_1,\dots,x_d) := \varphi\left(\prod_{i=1}^d \tilde{\varphi}(u_i)\right),$$

where $\varphi : [-1,1] \to [-\infty,\infty]$ is a strictly increasing continuous function with $\varphi(-1) = -\infty$, $\varphi(0) = 0$, $\varphi(1) = \infty$, and having derivatives up to order d on the intervals $(-1,0)$ and $(0,1)$, satisfying

$$\frac{\partial^d \varphi(e^x)}{\partial x^d} \geq 0, \quad \frac{\partial^d \varphi(-e^x)}{\partial x^d} \leq 0, \quad x \in (-\infty,0),$$

and where $\tilde{\varphi}$ is defined by

$$\tilde{\varphi}(u) := 2^{d-2}(\varphi(u) - \varphi(-u)).$$

Example 2.2.3. Bauerle et al. [2] observed that in the case $d > 2$ the family of Archimedean Lévy copulae fails to generate positively dependent Lévy processes.[1] In or-

[1] Three concepts of dependence are introduced in [2]: (positive) association, positive orthant dependence (POD), and positive supermodular dependence (PSMD). According to Corollary 3.10 in [2], all three concepts are equivalent in the case of multivariate Lévy processes. Thus, in the context of Lévy processes we give the same name to all three concepts: *positive dependence.*

der to overcome this problem they proposed to generalize this family. In the first step they noted that for $\phi : (0,\infty) \to (0,\infty)$ a strictly decreasing function with alternating signs of derivatives up to order d, and with $\lim_{t\downarrow 0}\phi(t) = \infty$ and $\lim_{t\downarrow\infty}\phi(t) = 0$ the function defined by

$$F_\phi(u_1,\dots,u_d) := \phi\left(\sum_{i=1}^{d}\phi^{-1}(u_i)\right), \quad u_1,\dots,u_d > 0, \tag{2.2}$$

is a Lévy copula on $(0,\infty)^d$. Then, the main idea of [2] was to spread these positive Lévy copulae on all orthants with additional weighting functions. This construction is a generalization of an ordinary Archimedean copula that uses the additive generator rather than the multiplicative one. To make the above idea precise, given functions $F_{\phi_{\mathbf{i}}}$ defined by (2.2) for $\mathbf{i} \in I := \{-1,1\}^d$, let

$$F(x_1,\dots,x_d) := \begin{cases} \sum_{\mathbf{i}\in I}\left(\eta(\mathbf{i})F_{\phi_{\mathbf{i}}}(|u_1|,\dots,|u_d|)1_{\{u\in O_{\mathbf{i}}\}}\prod_{i=1}^{d}\operatorname{sgn}(u_i)\right) & \text{if } |u_j| > 0, \\ & j = 1,\dots,d, \\ 0 & \text{otherwise,} \end{cases}$$

where $O_{\mathbf{i}}$ denotes the orthant with signs in $\mathbf{i}$, i.e.,

$$O_{\mathbf{i}} := \left\{x \in \mathbb{R}^d : sgn(x_j) = i_j, j = 1,\dots,d\right\},$$

and $\eta : I \to [0,1]$ is a weight function having the property

$$\sum_{\mathbf{i}:i_k=-1}\eta(\mathbf{i}) = \sum_{\mathbf{i}:i_k=1}\eta(\mathbf{i}) = 1.$$

The above function F defines a Lévy copula on $\mathbb{R}^d$ which generates positively dependent Lévy processes if and only if $\eta(1,1,...,1) = \eta(-1,-1,...,-1) = 1$.

Example 2.2.4. One can obtain a Clayton type Lévy copula by choosing $\varphi(u) = u^{-1/\theta}$ and $\eta(-1,-1) = \eta(1,1) = 1$. Then

$$F_\theta(u_1,u_2) = \left(u_1^{-\theta} + u_2^{-\theta}\right)^{-1/\theta} 1_{\{u\in\mathbb{R}^2_{++}\}} + \left((-u_1)^{-\theta} + (-u_2)^{-\theta}\right)^{-1/\theta} 1_{\{u\in\mathbb{R}^2_{--}\}}$$

with $\mathbb{R}^2_{++} := \mathbb{R}_+ \times \mathbb{R}_+$ and $\mathbb{R}^2_{--} := \mathbb{R}_- \times \mathbb{R}_-$, is a Lévy copula for positively dependent Lévy processes.

Remark 2.2.1. In various applications we often need an appropriate algorithm for Monte Carlo simulation of *dependent* Lévy processes. Tankov [29] discusses the issue of generation of sample paths of Lévy processes with given Lévy copulae by using series representations of Lévy processes.

2.3 Semimartingale Copulae

In this section, which is based on [4, 32], we study certain aspects of stochastic dependence between some classes of finite dimensional semimartingale processes in terms of their infinitesimal characteristics. We shall only consider the semimartingales that are uniquely characterized, in the sense of their probability laws, by their characteristics.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be some underlying probability space with $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, and let $X = (X^1, X^2, \ldots, X^n)$ be an $\mathbb{R}^n$-valued semimartingale with respect to some filtration, defined on this probability space. Let also $Y^1, Y^2, \ldots, Y^n$ be a collection of semimartingales on $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i)$, with respect to some filtrations. All filtrations in what follows are assumed to satisfy the usual conditions.

For the most part of this section, for simplicity of presentation, we shall only consider the bivariate case, that is, $n = 2$, although the results presented can be generalized to higher dimensions in a straightforward manner.

Our study is motivated by the question which arises naturally in various applications, such as valuation and hedging of financial derivatives written on baskets of underlying securities: What conditions on the local characteristics of a process X are sufficient for the law of X^i to be the same as the law of Y^i, $i = 1, 2, \ldots, n$, where the Y^i are given processes. So our aim is to construct a process X so that its i-th univariate law, i.e., the law of the i-th component X^i, is the same as the law of a given process Y^i, $i = 1, 2, \ldots, n$.

In this context, the question is reminiscent of the concept of copula functions, and the celebrated Sklar theorem (see [26]). Unfortunately, the complex structure of the cylindrical sigma algebras on canonical spaces does not allow a direct extension of Sklar's results to random variables on function spaces. However, infinitesimal characteristics of a stochastic process are often available, so we study dependence between processes in terms of those infinitesimal characteristics. Consequently, for historical reasons, we somewhat abuse terminology when using the term "copula": various "copulae" that we define below are not really copula functions. Nevertheless, we find this terminology useful and convenient.

Our approach was in part inspired by Tankov [30] and Kallsen and Tankov [21] (cf. Sect. 2.2). Although very appealing, their approach cannot be extended to construct more general processes, as its validity relies on the fact that the jump characteristic of a Lévy process is a measure on a finite dimensional space.

The key role in this section will be played by the canonical characteristics of a semimartingale, that is, the characteristics expressed as functions of the trajectory of the process.

2.3.1 Copulae for Special Semimartingales

Assume that we are given two real valued semimartingales Y^1, Y^2, whose finite dimensional distributions are uniquely determined by the corresponding infinitesimal characteristics. The processes are possibly defined on different (canonical) probability spaces, say $(\Omega_i, \mathscr{F}^{Y^i}, \mathbb{P}^i)$, $i = 1,2$, endowed with the canonical filtrations $\mathbb{F}^{Y^i}$. We would like to construct a probability space $(\Omega, \mathscr{F}^X, \mathbb{P})$, with $\Omega = \Omega_1 \times \Omega_2$, such that the finite dimensional distributions of the components of the canonical process $X = (X^1, X^2)$ on that space are identical to those of Y^1, Y^2. In what follows we use the notation:

$$\mathscr{F}_t^X = \bigcap_{s>t} \sigma(X_r, r \le s), \quad \mathscr{F}^X := \mathscr{F}_\infty^X := \bigvee_{t>0} \mathscr{F}_t^X, \quad \mathbb{F}^X := \{\mathscr{F}_t^X\}_{t\ge 0},$$

for a semimartingale X.

Consider a bivariate semimartingale, $X = (X^1, X^2)$, defined on a stochastic basis $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$, where the components X^1, X^2 are real valued semimartingales. We assume that the finite dimensional distributions of the vector process X are uniquely determined by its $\mathbb{F}^X$ characteristic triple.[2]

We first examine the problem of finding the $\mathbb{F}^{X^i}$ characteristic triple of $X^i, i = 1,2$ (i.e., the canonical characteristic triple of the coordinate processes) knowing the $\mathbb{F}^X$ characteristic triple of the components $X^i, i = 1,2$. We provide a characterization of the canonical characteristics of X^i in terms of projections of their $\mathbb{F}^X$ characteristic triple. Next we illustrate the theory on some examples, for which we compute explicitly the $\mathbb{F}^{X^i}$ characteristics of the coordinate processes. Finally, we explore how to extend these to determine the distribution of $X = (X^1, X^2)$ in some cases of interest.

Characteristics of the coordinate processes. Assume that X is a semimartingale $X = (X^1, X^2)$ taking values in $\mathbb{R}^2$ and defined on the stochastic basis $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$. Moreover we are given the $\mathbb{F}^X$-characteristics of X, say, (B, C, ν), where $B = (B^i)$ and $C = [C^{ij}]$ with $i, j = 1,2$ are predictable processes taking values in $\mathbb{R}^2$ and $\mathbb{R}^{2\times 2}$ respectively, and ν is a predictable random measure on $\mathscr{B}(\mathbb{R}^2) \otimes \mathscr{B}(\mathbb{R}_+)$ (the dual predictable projection of the integer valued, optional random measure μ counting the jumps of X).

We introduce the following notation:

- μ^i is the integer valued, optional random measure on $\mathscr{B}(\mathbb{R}) \otimes \mathscr{B}(\mathbb{R}_+)$ counting the jumps of the process X^i;
- ν^i is the compensator of μ^i in the filtration $\mathbb{F}^X$;
- $\widetilde{\nu}^i$ is the compensator of μ^i with respect to $\mathbb{F}^{X^i}$;
- ${}^{o_i}(Z)$ (or ${}^{p_i}(Z)$) is the optional (resp. predictable) projection of the process Z on $\mathbb{F}^{X^i}$;
- $(Z)^{p_i}$ (or $(\mu)^{p_i}$) is the dual predictable projection of the process Z (resp. of the random measure μ) on $\mathbb{F}^{X^i}$.

[2] The precise meaning of this statement is the following: If $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$ and $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{Q})$ are two probability spaces, where $(\Omega, \mathscr{F}, \mathbb{F})$ is a canonical space endowed with canonical filtration such that the canonical process X is a semimartingale on both stochastic bases with the same characteristics, then $\mathbb{P} = \mathbb{Q}$. This implies that two semimartingales, defined on $(\Omega, \mathscr{F}, \mathbb{F})$, that have the same characteristics, also have the same law. This uniqueness property can be verified in terms of uniqueness of the so called *martingale problem* (cf. [32] for details).

Clearly $\mu^2(dx,dt) = \mu(\mathbb{R},dx,dt)$, $\mu^1(dx,dt) = \mu(dx,\mathbb{R},dt)$ and likewise for ν^1,ν^2. We make the following standing assumptions:
A1. For $i = 1,2$ and for all $\mathbb{F}^X$ local martingales under consideration, there exists a fundamental sequence of $\mathbb{F}^{X^i}$ stopping times.
A2. The process X is a special semimartingale.

Assumption A1 ensures that we do not destroy the local martingale property by taking projections (if a local martingale is a genuine martingale this assumption is trivially satisfied: consider the sequence $T_n = n$). Assumption A2 ensures that X has a unique canonical semimartingale decomposition. The following results yield the $\mathbb{F}^{X^i}$ characteristics of the processes X^i.

Proposition 2.3.1 (see [32]). *Let $X^i = X_0^i + M^i + B^i$ denote the canonical decomposition of the semimartingale X^i in the filtration $\mathbb{F}^X$ and let B^{i+} and B^{i-} denote the processes in the Jordan decomposition of B^i. Then X^i admits the following canonical decomposition in the filtration $\mathbb{F}^{X^i}$:*

$$X^i = X_0^i + \widetilde{M}^i + \widetilde{B}^i,$$

where

$$\widetilde{M}^i = {}^{o_i}(M^i) + L^{i+} - L^{i-},$$

with L^{i+} and L^{i-} the local martingale parts of the Doob Meyer decomposition of ${}^{o_i}(B^{i+})$ and ${}^{o_i}(B^{i-})$ respectively, and

$$\widetilde{B}^i = ({}^{o_i}(B^{i+}))^{p_i} - ({}^{o_i}(B^{i-}))^{p_i}. \tag{2.3}$$

The above proposition yields the first two $\mathbb{F}^{X^i}$ characteristics of the process X^i, i.e., $\widetilde{B}^i$ is given by (2.3), and $\widetilde{C}^{ii} = \langle (\widetilde{M}^i)^c, (\widetilde{M}^i)^c \rangle$. Using similar arguments, we compute the jump characteristic of X^i in the filtration $\mathbb{F}^{X^i}$. To this end, we shall need the following results:

Proposition 2.3.2 (see [32]). *Fix $i = 1,2$, $s \geq 0$, $A \in \mathscr{B}(\mathbb{R})$ and $B \in \mathscr{F}_s^{X^i}$. Then the process*

$$L_t^i = \mathbb{1}_B \left(\mu^i((s,t],A) - ({}^{o_i}(\nu^i((s,t],A)))^{p_i} \right), \quad t \geq s,$$

is an $\mathbb{F}^{X^i}$ local martingale.

From now on we assume that the jumps of the processes X^1 and X^2 take values in a finite set,[3] say $E = \{x_1, \ldots, x_M\} \subset \mathbb{R}$. For every $x \in E$ and every interval $(s,t]$ we put

$$\widetilde{\nu}^i((s,t],x) := ({}^{o_i}(\nu^i((s,t],x)))^{p_i}.$$

[3] The set E can be interpreted as the mark space.

One can uniquely extend $\widetilde{\nu}^i$ to a measure on $\mathscr{B}(\mathbb{R}_+) \otimes 2^E$. The next proposition shows that this unique extension, denoted by $\widetilde{\nu}^i(dt,dx)$, is indeed the $\mathbb{F}^{X^i}$ compensator of $\mu_i(dt,dx)$.

Proposition 2.3.3. *The measure $\widetilde{\nu}^i(dt,dx)$ is the $\mathbb{F}^{X^i}$ dual predictable projection of the counting measure $\mu^i(dt,dx)$.*

If the compensator is absolutely continuous, then we can compute the projections $({}^{o_i}(\nu^i([s,t],A_n)))^{p_i}$ in a simple way:

Lemma 2.3.1 (see [32]). *Assume that $\nu^i((s,t],A)$ is (locally) integrable for every set A in $\mathscr{B}(\mathbb{R})$. In addition, assume that $\nu^i((s,t],A)$ is absolutely continuous, i.e.,*

$$\nu^i((s,t],A)(\omega) = \int_s^t \int_A K^i(u,\omega,dx)\,du$$

for some $\mathscr{F}^X \otimes \mathscr{B}(\mathbb{R})$ measurable kernel K^i. Then, for any $s < t < \infty$,

$$({}^{o_i}(\nu^i((s,t],A)))^{p_i} = \int_s^t {}^{o_i}(K^i(u,A))\,du.$$

We can obtain a similar result for the process $\widetilde{B}^i$, that is, for the finite variation part of the $\mathbb{F}^{X^i}$ semimartingale decomposition of X^i.

Lemma 2.3.2 (see [32]). *Assume that, for $i = 1,2$, B^i is (locally) integrable and absolutely continuous, i.e.,*

$$B^i_t = \int_0^t b^i_s\,ds$$

for some progressively measurable process b^i. Then, for any $t < \infty$,

$$\widetilde{B}^i_t = ({}^{o_i}(B^{i+}))^{p_i}_t - ({}^{o_i}(B^{i-}))^{p_i}_t = \int_0^t {}^{o_i}(b^i_s)\,ds.$$

Now, we compute explicitly the $\mathbb{F}^{X^i}$ characteristics of the vector semimartingale $X = (X^1,X^2)$ in some special cases.

Example 2.3.1. Consider the stochastic basis $(\Omega,\mathscr{F},\mathbb{F},\mathbb{P})$, where $\mathbb{F} := \mathbb{F}^W \vee \mathbb{F}^{X^2}$, W is an SBM and X^2 is a Markov chain that takes values in $\{e_1,\dots,e_N\}$, where $(e_i)_{i=1}^N$ is the standard basis in $\mathbb{R}^N$. We assume that X^2 admits a constant generator matrix A. Consider the vector process $X = (X^1,X^2)$, where

$$dX^1_t = \langle b, X^2_t\rangle\,dt + \sigma dW_t, \quad X^1_0 = x^1 \in \mathbb{R},$$

with $\sigma \in \mathbb{R}^+$ and $b \in \mathbb{R}^N$. Note that $\mathbb{F}^X = \mathbb{F}$ and the first $\mathbb{F}^X$ characteristic of X^1 is given by the process $B^1_t = \int_0^t \langle b, X^2_s\rangle\,ds$. In view of Lemma 2.3.2, the first $\mathbb{F}^{X^1}$ characteristic of X^1 is given by the process $\widetilde{B}^i_t = \int_0^t \langle b, p_s\rangle\,ds$, where $p_s := {}^{o_1}(X^2_s)$.

It is a well known result in filtering theory (the so called Wonham filter, see Elliot [13]) that the process p satisfies the following vector SDE:

$$p_t = p_0 + \int_0^t A^T p_s\,ds + \frac{1}{\sigma^2}\int_0^t \mathrm{diag}(p_s)(b - \langle b, p_s\rangle \mathbb{1})(dX_s^1 - \langle b, p_s\rangle\,ds),$$

where p_0 is the initial distribution of the chain X^2, and $\mathbb{1} := (1,\dots,1)^\top$.

Example 2.3.2. Consider a stochastic basis $(\Omega, \mathscr{F}^N, \mathbb{F}^N, \mathbb{P})$, where $\Omega = \Omega_1 \times \Omega_2$ is the canonical space of a bivariate one-point process, and a canonical process, say $N = (N^1, N^2)$, on this space. Observe that N can be identified with a pair of positive random variables $T_1 : \Omega_1 \to \mathbb{R}_+$ and $T_2 : \Omega_2 \to \mathbb{R}_+$ given by $T_1 := \inf\{t > 0 : \Delta N_t^1 \neq 0\}$ and $T_2 := \inf\{t > 0 : \Delta N_t^2 \neq 0\}$. We assume that, under $\mathbb{P}$, the joint probability of (T_1, T_2) admits a density function $f(u,v)$. We first compute the $\mathbb{F}^N$ compensator of N^i. By a straightforward application of the Fubini theorem, the $\mathbb{F}^N$ jump characteristic of N^1 is

$$\nu^1(ds,dx) = \delta_1(dx)\left(\frac{\int_s^\infty f(s,v)\,dv}{\int_s^\infty\int_s^\infty f(u,v)\,du\,dv}\mathbb{1}_{\{s\le T_1\wedge T_2\}} + \frac{f(s,T_2)}{\int_s^\infty f(u,T_2)\,du}\mathbb{1}_{\{T_2<s\le T_1\}}\right)ds.$$

To compute the canonical jump characteristic of the coordinate processes N^1, say $\widetilde{\nu}^1([0,t),\{1\})$, we use Propositions 2.3.2 and Lemma 2.3.1:

$$\begin{aligned}\widetilde{\nu}^1(ds,\{1\}) &= \mathbb{E}\left(\frac{\int_s^\infty f(s,v)\,dv}{\int_s^\infty\int_s^\infty f(u,v)\,du\,dv}\mathbb{1}_{\{s\le T_1\wedge T_2\}}\big|\mathscr{F}_s^{N^1}\right)ds\\ &+ \mathbb{E}\left(\frac{f(s,T_2)}{\int_s^\infty f(u,T_2)\,du}\mathbb{1}_{\{T_2<s\le T_1\}}\big|\mathscr{F}_s^{N^1}\right)ds.\end{aligned}$$

Since the process $\mathbb{1}_{\{T_1\ge t\}}$ is predictable, it is adapted to $(\mathscr{F}_{t-}^{N^1}, t\ge 0)$. Therefore, for any $s<\infty$, we have

$$\begin{aligned}\widetilde{\nu}^1(ds,\{1\}) &= \frac{\int_s^\infty f(s,v)\,dv}{\int_s^\infty\int_s^\infty f(u,v)\,du\,dv}\mathbb{P}(\{T_2\ge s\}\big|\{T_1\ge s\})\mathbb{1}_{\{T_1\ge s\}}\,ds\\ &+\mathbb{E}\left(\frac{f(s,T_2)}{\int_s^\infty f(u,T_2)\,du}\mathbb{1}_{\{T_2<s\le T_1\}}\big|\{T_1\ge s\}\right)\mathbb{1}_{\{T_1\ge s\}}\,ds\\ &= \frac{\int_s^\infty f(s,v)\,dv}{\int_s^\infty\int_s^\infty f(u,v)\,du\,dv}\,\frac{\int_s^\infty\int_s^\infty f(u,v)\,du\,dv}{\int_s^\infty\int_0^\infty f(u,v)\,du\,dv}\mathbb{1}_{\{T_1\ge s\}}\,ds\\ &+\int_0^s \frac{f(s,v)}{\int_s^\infty f(u,v)\,du}\,\frac{\int_s^\infty f(u,v)\,du}{\int_s^\infty\int_0^\infty f(u,v)\,du\,dv}\,dv\mathbb{1}_{\{T_1\ge s\}}\,ds\\ &= \frac{\int_0^\infty f(s,v)\,dv}{\int_s^\infty\int_0^\infty f(u,v)\,du\,dv}\mathbb{1}_{\{T_1\ge s\}}\,ds.\end{aligned}$$

Remark 2.3.1. Let $X = (X^1, X^2)$ be a two dimensional semimartingale defined on the stochastic basis $(\Omega, \mathscr{F}^X, \mathbb{F}^X, \mathbb{P})$ and such that its $\mathbb{F}^X$ characteristic triple uniquely determines its finite dimensional distributions. We know, at least in some special cases, how to compute the $\mathbb{F}^{X^i}$ characteristics of the components $X^i, i = 1,2$, from the corresponding $\mathbb{F}^X$ characteristic triple. We still have to establish whether the $\mathbb{F}^{X^i}$ characteristic triple uniquely determines the finite dimensional distributions of the components X^i, say $(\widetilde{B}^i, \widetilde{C}^i, \widetilde{\nu}^i)$. We can give a positive answer under the assumption that there exists a unique probability measure $\mathbb{P}^i$ on $\mathscr{F}^{X^i}$ such that X^i is a semimartingale with $\mathbb{F}^{X^i}$ characteristic triple $(\widetilde{B}^i, \widetilde{C}^i, \widetilde{\nu}^i)$. Then necessarily the restriction $\mathbb{P}|_{\mathscr{F}^{X^i}}$ must coincide with $\mathbb{P}^i$. This implies that, at least in this case, the $\mathbb{F}^{X^i}$ characteristics indeed determine the finite dimensional distributions of X^i.

Semimartingale copulae. We are now ready to proceed with the presentation of semimartingale copulae. In fact, the discussion earlier in this section, and, in particular, the discussion in Remark 2.3.1 indicate a recipe for constructing bivariate semimartingales with given margins. We construct $(\Omega, \mathscr{F}^X, \mathbb{F}^X, \mathbb{P})$ in such a way that the $\mathbb{F}^{X^i}$ characteristics of the coordinate processes X^i, $i = 1,2$, are identical (as functions of trajectories) to the $\mathbb{F}^{Y^i}$ characteristics of Y^i, $i = 1,2$. This implies that for $i = 1,2$, X^i and Y^i are equal in law.

Indeed, X^i and Y^i live in the same canonical space Ω_i (this means that the canonical σ-algebras $\mathscr{F}^{Y^i}$ and $\mathscr{F}^{X^i}$ contain the same events), and the $\mathbb{F}^{X^i}$ characteristics of X^i and the $\mathbb{F}^{Y^i}$ characteristics of Y^i coincide (as functions of trajectories), so uniqueness implies that X^i and Y^i have the same finite dimensional distributions.

We now proceed to define the concept of semimartingale copula for two dimensional semimartingales (this definition can be readily extended to higher dimensional processes).

Let Y^1, Y^2 be two $\mathbb{R}$-valued semimartingales defined on possibly different (canonical) filtered probability spaces $(\Omega_i, \mathscr{F}^{Y^i}, \mathbb{F}^{Y^i}, \mathbb{P}^i)$. Let $(\widehat{B}^i, \widehat{C}^i, \widehat{\nu}^i)$ denote the characteristics of Y^i, $i = 1,2$, and assume that the finite dimensional distributions of Y^i are uniquely determined by its characteristic triple. Let X denote the vector valued, canonical process on the filtered canonical stochastic basis $(\Omega, \mathscr{F}^X, \mathbb{F}^X)$, where $\Omega = \Omega_1 \times \Omega_2$.

Definition 2.3.1. We say that a triple (B, C, ν) defined on the basis $(\Omega, \mathscr{F}^X, \mathbb{F}^X)$ is a semimartingale copula for $Y^i, i = 1,2$, if the following conditions hold:

i) there is a unique probability measure $\mathbb{P}$ on $\mathscr{F}^X$ such that the canonical process on the stochastic basis $(\Omega, \mathscr{F}^X, \mathbb{F}^X)$ is a semimartingale with characteristic triple (B, C, ν);
ii) under $\mathbb{P}$, the $\mathbb{F}^{X^i}$ characteristics of X^i, say $(\widetilde{B}^i, \widetilde{C}^i, \widetilde{\nu}^i)$, are equal (as functions of trajectories) to $(\widehat{B}^i, \widehat{C}^i, \widehat{\nu}^i)$ for $i = 1,2$.

Now we introduce a suitable measure of dependence between components of the process X. Let $\mathscr{J} = 2^{\{1,\dots,d\}}$, $\mathscr{J}_i = \{S \in \mathscr{J} : S \text{ contains at least } i \text{ elements}\}$ and let $\mathrm{card}(S)$ denote the cardinality of the set S.

Definition 2.3.2. Let $X = (X^1, X^2, \dots, X^d)$ be an $E = \times_{i=1}^d \mathbb{E}_i \subset \mathbb{R}^d$ valued locally square integrable semimartingale. Let $T < \infty$. The *d-volume* is defined as

$$\mathrm{Dvol}(X_T) = \mathbb{E}\left(\int_0^T \frac{1}{2}\sum_{i,j=1,i\neq j}^{m} |d\langle X^i, X^j\rangle_s^c| + \sum_{S\in\mathscr{J}_2} \mathrm{card}(S)\nu^S(T)\right), \tag{2.4}$$

where $\langle X^i, X^j\rangle^c$ is the process compensating $[X^i, X^j]^c$ in the filtration $\mathbb{F}^X$, ν_t^S is the $\mathbb{F}^X$-dual predictable projection of the process

$$J_t^S := \int_0^t \int_{E_S\times\mathbf{0}_{S^c}} \mu(dx, ds),$$

$\mu(dx, ds)$ is an integer valued random measure counting the jumps of the process X, and $E_S \times \mathbf{0}_{S^c}$ is a set in $\mathbb{R}^d$ defined as $\mathsf{X}_{i\in S}\mathbb{E}_i \times \mathsf{X}_{i\in S^c}\{0\}$.

Remark 2.3.2. The d-volume is a measure of “dynamic” dependence between components of a multivariate semimartingale X. The dependence between processes is also related to the dependence structure intrinsic to the initial state X_0, but this "static" dependence is not accounted for in the d-volume.

To motivate this choice, we give the following simple example.

Example 2.3.3. Let $X = (X^1, X^2)$ be a nondegenerate jump diffusion with

$$\begin{aligned} dX^1 &= \mu_1(X^1)dt + \sigma_{11}(X^1)dW^1(t) + \sigma_{12}(X^1)dW^2(t) + dN^1(t) + dN^2(t),\\ dX^2 &= \mu_2(X^2)dt + \sigma_{21}(X^2)dW^1(t) + \sigma_{22}(X^2)dW^2(t) + dN^2(t) + dN^3(t), \end{aligned}$$

where W^1, W^2 are independent Brownian motions and N^1, N^2, N^3 are independent Poisson processes with intensities $\lambda_1, \lambda_2, \lambda_3$ respectively. In this case the covariance between the continuous components of X^1 and X^2 is measured by

$$dc_{ij} = \left(\sigma_{12}(X^1)\sigma_{21}(X^2) + \sigma_{12}(X^1)\sigma_{22}(X^2)\right) dt,$$

while the tendency of the processes to jump together is measured by

$$\nu(dt) = \lambda_2\, dt.$$

Proposition 2.3.4. *Let $X = (X^1, X^2, \ldots, X^m)$ be an E-valued semimartingale. Then*

$$0 \le \mathrm{Dvol}(X_T) \le \mathbb{E}\left(\frac{1}{2}\sum_{i,j=1,i\neq j}^{m} (\langle X^i, X^i\rangle_T^c)^{\frac{1}{2}} (\langle X^j, X^j\rangle_T^c)^{\frac{1}{2}} + \sum_{S\in\mathscr{J}_1} \mathrm{card(S)}\nu^S(T)\right).$$

Proof. This is an immediate consequence of the Kunita-Watanabe inequality (see Protter [27, Chap. II, Theorem 25]) and the observation that $\sum_{S\in\mathscr{J}_2}\mathrm{card}(S)\nu_T^S \le \sum_{S\in\mathscr{J}_1}\mathrm{card}(S)\nu_T^S$. □

Example 2.3.4. We construct now a semimartingale copula for a vector one-point process. Suppose we are given two one-point processes Y^i defined on the basis $(\Omega_i, \mathscr{F}^{Y^i}, \mathbb{F}^{Y^i}, \mathbb{P}^i)$, where Ω_i is the canonical space of one-point processes on $\mathbb{R}$. Let $\widetilde{T}_1$ and $\widetilde{T}_2$ denote the jump times of Y^1 and Y^2, respectively, and assume that, under $\mathbb{P}^i$, $\widetilde{T}^i$ is exponentially distributed with parameter λ_i, $i = 1,2$. Let F_i denote the corresponding distribution function. It can be easily verified that Y^i admits jump characteristic $\widetilde{\nu}^i(dt,dx) = \delta_1(dx)\lambda_i \mathbb{1}_{\{\widetilde{T}_i \geq t\}}\, dt$. Let $N = (N^1, N^2)$ denote the canonical point process on the stochastic basis $(\Omega, \mathscr{F}^N, \mathbb{F}^N)$, where $\Omega = \Omega_1 \times \Omega_2$. Next define two positive random variables T_1, T_2 as follows: $T_1 = \inf\{t \geq 0 : \Delta N_t^1 = 1\}$, $T_2 = \inf\{t \geq 0 : \Delta N_t^2 = 1\}$. Let $C(\cdot,\cdot)$ be an arbitrary two dimensional absolutely continuous copula function, and $c(\cdot,\cdot)$ the density of the distribution function $C(F_1(\cdot), F_2(\cdot))$. In addition, define the following random measure on $\mathscr{B}(\mathbb{R}_+) \otimes 2^E$ (here $E = \{0,1\}^2$):

$$
\begin{aligned}
\nu(ds,\{(0,0)\}) &= 0,\\
\nu(ds,\{(1,0)\}) &= \left(\frac{\int_s^\infty c(s,v)dv}{\int_s^\infty \int_s^\infty c(u,v)dudv} I_{\{s \leq T_1 \wedge T_2\}} + \frac{c(s,T_2)}{\int_s^\infty c(u,T_2)du} I_{\{T_2 < s \leq T_1\}} \right) ds,\\
\nu(ds,\{(0,1)\}) &= \left(\frac{\int_s^\infty c(u,s)du}{\int_s^\infty \int_s^\infty c(u,v)dudv} I_{\{s \leq T_1 \wedge T_2\}} + \frac{c(T_1,s)}{\int_s^\infty c(T_1,v)dv} I_{\{T_1 < s \leq T_2\}} \right) ds,\\
\nu(ds,\{(1,1)\}) &= 0.
\end{aligned}
$$

First we prove that the measure $\nu(dt,dx)$ is $\mathbb{F}^N$ predictable. To this end consider the simple random function $W = \mathbb{1}_A(\mathbb{1}_{C_0}\mathbb{1}_{\{t \leq T_1 \wedge T_2\}} + \mathbb{1}_{C_1}\mathbb{1}_{\{T_1 \wedge T_2 < t \leq T_1 \vee T_2\}} + \mathbb{1}_{C_2}\mathbb{1}_{\{t > T_1 \vee T_2\}})$, where A is a set in E, $C_0 \in \mathscr{F}_0^N$, $C_1 \in \mathscr{F}^N_{T_1 \wedge T_2}$ and $C_2 \in \mathscr{F}^N_{T_1 \vee T_2}$. Since $W * \nu$ is of the form $D_0 \mathbb{1}_{\{t \leq T_1 \wedge T_2\}} + D_1 \mathbb{1}_{\{T_1 \wedge T_2 < t \leq T_1 \vee T_2\}} + D_2 \mathbb{1}_{\{t > T_1 \vee T_2\}}$, where $D_0 \in \mathscr{F}_0^N$, $D_1 \in \mathscr{F}^N_{T_1 \wedge T_2}$ and $D_2 \in \mathscr{F}^N_{T_1 \vee T_2}$ we see that $W * \nu$ is an $\mathbb{F}^N$ predictable process (Lemma III 1.29 in [20]). Hence ν is an $\mathbb{F}^N$ predictable random measure by a monotone class argument. By direct verification we see that the probability measure $\mathbb{P}$ defined by the distribution function

$$
\mathbb{P}(T_1 \leq t_1, T_2 \leq t_2) = C(F_1(t_1), F_2(t_2))
$$

is a solution to the martingale problem for ν. In fact, if $\mu(dt,dx)$ denote the optional counting measure associated to N, then $\mu((0,t],(1,0)) = N_t^1$ and $\mu((0,t],(0,1)) = N_t^2$, and the martingale property of $\mu((0,t],A) - \nu((0,t],A)$, for all $A \in 2^E$, follows from a straightforward application of the Fubini theorem. Next, we deduce that, for all predictable simple random functions W, the process $W * (\mu - \nu)$ is a martingale, implying that $\mathbb{P}$ is a solution to the martingale problem for ν. It follows from known results (e.g. [24, Chap. 4, Theorem 5]) that $\mathbb{P}$ is the unique probability measure on $\mathscr{F}^N$ such that the canonical process N is a bivariate one-point process with compensator ν.

Moreover, by arguments analogous to those used in Example 2.3.2, the $\mathbb{F}^{N^1}$ dual predictable projection of N^1 is given by

$$
\begin{aligned}
\widetilde{\nu}^1(ds,dx) &= \delta_1(dx) \frac{\int_0^\infty c(s,s_2)ds_2}{\int_s^\infty \int_0^\infty c(s_1,s_2)ds_2\, ds_2} \mathbb{1}_{\{s \leq T_1\}}\, ds\\
&= \delta_1(dx) \frac{\lambda_1 \exp(-\lambda_1 s)}{\exp(-\lambda_1 s)} \mathbb{1}_{\{s \leq T_1\}}\, ds = \delta_1(dx)\lambda_1 \mathbb{1}_{\{s \leq T_1\}}\, ds.
\end{aligned}
$$

We conclude that the random measure $\nu(dt,dx)$ is a semimartingale copula for Y^1,Y^2.

The above example relied on the fact that we are able to compute the projections of the $\mathbb{F}^N$ jump characteristic of $N^i, i=1,2$. In the general case, if we had to compute projections in order to construct a semimartingale copula, the practical usefulness of the theory presented so far would be rather restricted. So we try to construct nontrivial semimartingale copulae without computing projections. We start from a construction of a vector one-point process $N=(N^1,N^2)$ different from that given in Example 2.3.4.

Example 2.3.5. Let the setting be as in Example 2.3.4. Again, we would like to construct a probability measure on $(\Omega,\mathscr{F}^N,\mathbb{F}^N)$ such that the distributions of the jump time of the components N^i, $i=1,2$, are exponential with intensity λ_i, i.e., N^i are equal in law to Y^i for $i=1,2$. To do this, let λ_{12} be a positive real number satisfying the condition $\lambda_{12}\le\lambda_1\wedge\lambda_2$ and define the following random measure on $\mathscr{B}(\mathbb{R}_+)\otimes 2^E$ ($E=\{0,1\}^2$):

$$\begin{aligned}
\nu(ds,\{(0,0)\}) &= 0,\\
\nu(ds,\{(1,0)\}) &= \left(\lambda_1\mathbb{1}_{\{s\le T_1\}}-\lambda_{12}\mathbb{1}_{\{s\le T_1\wedge T_2\}}\right)ds,\\
\nu(ds,\{(0,1)\}) &= \left(\lambda_2\mathbb{1}_{\{s\le T_2\}}-\lambda_{12}\mathbb{1}_{\{s\le T_1\wedge T_2\}}\right)ds,\\
\nu(ds,\{(1,1)\}) &= \left(\lambda_{12}\mathbb{1}_{\{s\le T_1\wedge T_2\}}\right)ds.
\end{aligned}$$

By arguments analogous to those used in Example 2.3.4 we infer that the nonnegative measure $\nu(dt,dx)$ is $\mathbb{F}^N$ predictable. By a direct verification, a solution to the martingale problem for ν is given by the probability measure $\mathbb{P}$ defined by the distribution function

$$\mathbb{P}(T_1\le t_1,T_2\le t_2)=1-e^{-\lambda_1 t_1}-e^{-\lambda_2 t_2}+e^{-(\lambda_1-\lambda_{12})t_1-(\lambda_2-\lambda_{12})t_2-\lambda_{12}t_1\vee t_2}.$$

To see this, we prove that $M_t=\mu((0,t],A)-\nu([0,t],A)$, where $\mu(dt,dx)$ is the counting measure associated to N, is an $\mathbb{F}^N$ martingale for every set A in 2^E. Let, for example, $A=\{(1,1)\}$. Then $\mu((0,t],(1,1))=\mathbb{1}_{\{T_1\wedge T_2\le t,T_1=T_2\}}$ and by direct computation

$$\begin{aligned}
\mathbb{P}(T_1\le t,T_2\le t,T_1=T_2|\{T_1\wedge T_2>s\}) &= \int_s^t\lambda_{12}e^{-(\lambda_1+\lambda_2-\lambda_{12})(u-s)}\,du,\\
\mathbb{P}(T_1\wedge T_2\ge t|\{T_1\wedge T_2>s\}) &= e^{-(\lambda_1+\lambda_2-\lambda_{12})(t-s)}.
\end{aligned}$$

Therefore, M_t satisfies

$$\begin{aligned}
&\mathbb{E}(M_t|\mathscr{F}_s^N)=M_s+\mathbb{P}(\{s<T_1\wedge T_2\le t,T_1=T_2\}|\mathscr{F}_s^N)-\lambda_{12}\int_s^t\mathbb{P}(T_1\wedge T_2\ge u|\mathscr{F}_s^N)\,du\\
&=M_s+\int_s^t e^{-(\lambda_1+\lambda_2-\lambda_{12})(u-s)}\,du\,\mathbb{1}_{\{T_1\wedge T_2>s\}}-\int_s^t e^{-(\lambda_1+\lambda_2-\lambda_{12})(u-s)}\,du\,\mathbb{1}_{\{T_1\wedge T_2>s\}}=M_s.
\end{aligned}$$

$\mathbb{P}$ is the unique probability measure on $\mathscr{F}^N$ such that the canonical process N is a bivariate one-point process with compensator ν. Moreover, under $\mathbb{P}$, the $\mathbb{F}^N$ compensator and the $\mathbb{F}^{N^i}$ compensator of the coordinate process N^i, $i = 1,2$, coincide, i.e., $\widetilde{\nu}^i = \nu^i$. In fact, the $\mathbb{F}^N$ compensator of N^1 is given by

$$\begin{aligned}\nu^1(ds,\{1\}) &= \nu(ds,\{(1,0)\}\cup\{(1,1)\}) \\ &= \lambda_1 \mathbb{1}_{\{s\le T_1\}} - \lambda_{12}\mathbb{1}_{\{s\le T_1\wedge T_2\}} + \lambda_{12}\mathbb{1}_{\{s\le T_1\wedge T_2\}} = \lambda_1 \mathbb{1}_{\{s\le T_1\}},\end{aligned}$$

and $\nu^1(ds,\{x\}) = 0$ if $x \neq 1$, i.e., $\nu^1(ds,dx) = \delta_1(dx)\lambda_1\mathbb{1}_{\{s\le T_1\}}\,ds$, and similarly for $\nu^2(ds,dx)$. Therefore, for $i = 1,2$, $\nu^i(ds,\{1\})$ is $\mathbb{F}^{N^i}$ predictable and $\nu^i(ds,dx)$ is as well the compensator of N^i in the filtration $\mathbb{F}^{N^i}$. It is then obvious that the $\mathbb{F}^{N^i}$ characteristic of the component process is the same (as a function of trajectories) as the $\mathbb{F}^{Y^i}$ characteristic of Y^i, i.e., the random measure $\nu(dt,dx)$ is a semimartingale copula.

Remark 2.3.3. The probability measure $\mathbb{P}$ solving the martingale problem for $\nu(dt,dx)$ given in Example 2.3.5 can be constructed via the Marshall-Olkin copula between two exponential random variables with intensity λ_1 and λ_2 (see [26, Sect. 3.1.1]). By direct computation, if $t_1,t_2 > 0$, then $\mathbb{P}(T_1 \le t_1, T_2 \le \infty) = 1 - e^{-\lambda_1 t_1}$ and $\mathbb{P}(T_1 \le \infty, T_2 \le t_2) = 1 - e^{-\lambda_2 t_2}$.

Remark 2.3.4. Note that the nature of dependence between the coordinate processes exhibited in Example 2.3.5 is very different from that seen in Example 2.3.4. In Example 2.3.4 the $\mathbb{F}^N$ characteristics of each component may also depend functionally on the trajectories of the other components. On the other hand, in Example 2.3.5, the dependence between components is only given by the possibility of common jumps of the processes (the jump measure of the set $\{(1,1)\}$ is positive).

2.3.2 Consistent Semimartingale Copulae

From the above examples we can draw some useful conclusions. Given the $\mathbb{F}^X$ characteristic triple of a multivariate process X, the computation of the $\mathbb{F}^{X^i}$ characteristics of the components X^i can be made much simpler if we construct the multivariate process in such a way that the $\mathbb{F}^X$ characteristic triple of X^i is $\mathbb{F}^{X^i}$ predictable. In this case the computation of projections can be avoided. As we have seen in the examples, projecting the $\mathbb{F}^X$ characteristics in the filtration $\mathbb{F}^{X^i}$ is rather difficult, and in fact this is possible only in very simple and special cases. These observations suggest the following definition:

Definition 2.3.3. We say that a two dimensional semimartingale $X = (X^1,X^2)$ defined on a stochastic basis $(\Omega,\mathscr{F}^X,\mathbb{F}^X,\mathbb{P})$ is consistent with respect to $\mathbb{F}^{X^i}$ if the $\mathbb{F}^X$ characteristic triple (B^i,C^i,ν^i) and the $\mathbb{F}^{X^i}$ characteristic triple $(\widetilde{B}^i,\widetilde{C}^i,\widetilde{\nu}^i)$ of the component process X^i coincide (as functions of trajectories).

Let Y^1, Y^2 be two $\mathbb{R}$-valued semimartingales defined on possibly different (canonical) filtered probability spaces $(\Omega_i, \mathscr{F}^{Y^i}, \mathbb{F}^{Y^i}, \mathbb{P}^i)$, $i = 1,2$. Moreover, let $(\widehat{B}^i, \widehat{C}^i, \widehat{\nu}^i)$ denote the characteristics of Y^i, $i = 1,2$, and assume that the finite dimensional distributions of Y^i are uniquely determined by its characteristic triple. Let $X = (X^1, X^2)$ denote the vector valued, canonical process on the filtered canonical stochastic basis $(\Omega, \mathscr{F}^X, \mathbb{F}^X)$ where $\Omega = \Omega_1 \times \Omega_2$.

Definition 2.3.4. We say that a triple (B, C, ν) defined on the basis $(\Omega, \mathscr{F}^X, \mathbb{F}^X)$ is a consistent semimartingale copula for $Y^i, i = 1,2$, if the following conditions hold:

i) there is a unique probability measure $\mathbb{P}$ on $\mathscr{F}^X$ such that the canonical process on the stochastic basis $(\Omega, \mathscr{F}^X, \mathbb{F}^X)$ is a semimartingale with characteristic triple (B, C, ν);

ii) under $\mathbb{P}$, X is consistent with respect to $\mathbb{F}^{X^i}$ and (B^i, C^i, ν^i) are equal (as functions of trajectories) to $(\widehat{B}^i, \widehat{C}^i, \widehat{\nu}^i)$ for $i = 1,2$.

Note that the difference in the definitions of semimartingale copula and consistent semimartingale copula lies in the requirement of *consistency* imposed in the latter case; as we explained earlier, the consistency property allows one to avoid computations of projections of the characteristics on smaller filtrations. We devote the rest of this section to construct examples of consistent semimartingale copulae for some important classes of semimartingales.[4]

Copulae between pure jump Lévy processes. We shall now provide an elementary example of a semimartingale copula that is also a Lévy copula. There is a one-to-one correspondence between a homogeneous Poisson process with values in $\mathbb{R}^2$ and a homogeneous Poisson measure on $E = \{0,1\}^2 \setminus \{(0,0)\}$. We let ν denote the $\mathbb{F}$ dual predictable projection of a Poisson measure μ. The measure ν is a measure on a finite set, so it is uniquely determined by its values on the atoms in E. Therefore a Poisson process X in $\mathbb{R}^2$ is uniquely determined by

$$\nu(dt, \{1,0\}) = \lambda_{10} dt, \quad \nu(dt, \{0,1\}) = \lambda_{01} dt, \quad \nu(dt, \{1,1\}) = \lambda_{11} dt \tag{2.5}$$

for some positive constants λ_{10}, λ_{01} and λ_{11}.

Example 2.3.6. Let us consider two Poisson processes X^1 and X^2 with values in $\mathbb{R}^1$, with intensities λ_1 and λ_2 respectively. We will show that if real numbers $\lambda_{10}, \lambda_{01}, \lambda_{11}$ satisfy

$$\begin{aligned} \lambda_1 &= \lambda_{10} + \lambda_{11}, \\ \lambda_2 &= \lambda_{01} + \lambda_{11}, \\ \lambda_{11} &\in [0, \lambda_1 \wedge \lambda_2], \end{aligned} \tag{2.6}$$

then the measure ν in (2.5) is a semimartingale copula for X^1, X^2. First, (2.6) implies that ν is positive. Moreover, ν defines uniquely the probability law of a Poisson random measure on $\{0,1\}^2$. The vector Poisson process corresponding to ν can, in fact,

[4] Since all semimartingale copulae constructed below are consistent we often omit the qualifier consistent and we only talk about semimartingale copulae.

be easily constructed from a vector of three independent unit Poisson processes, say (N^1,N^2,N^3), by using time-changing. If $Y^1_t = N^1_{\lambda_{10}t} + N^2_{\lambda_{11}t}$ and $Y^2_t = N^3_{\lambda_{01}t} + N^2_{\lambda_{11}t}$, then it is straightforward to verify that the probability law of $Y = (Y^1,Y^2)$ is a solution of the martingale problem for $\nu(dt,dx)$. Uniqueness of the martingale problem for the triple $(0,0,\nu)$ follows from [24, Chap. 4, Theorem 5]. Finally, since

$$\nu^1(dt,\{1\}) = \nu(dt,\{(1,0)\}) + \nu(dt,\{(1,1)\}) = \lambda_{10}dt + \lambda_{11}dt = \lambda_1 dt$$

and $\nu^1(dt,\{x\}) = 0$, $\forall x \neq 1$ (and similarly for $\nu^2(dt,\{1\})$) we conclude that $\nu(dt,dx)$ is a semimartingale copula.[5]

Copulae between diffusion processes. Let us consider two $\mathbb{R}$-valued diffusion processes X_1 and X_2 defined on the spaces $(\Omega'_1,\mathscr{G}^1,\mathbb{P}^1)$ and $(\Omega'_2,\mathscr{G}^2,\mathbb{P}^2)$, where $(\Omega'_i,\mathscr{G}^i,\mathbb{P}^i)$ supports the Standard Brownian Motion (SBM) W^i, $i=1,2$, and the filtration $\widetilde{\mathbb{G}}^i$ is generated by the SBM W^i. Assume that the diffusions are driven by the following SDEs:

$$dX^i(t) = \mu_i(X^i(t))dt + \sigma_i(X^i(t))dW^i(t), \quad X^i(0) = x_i, \quad i = 1,2. \tag{2.7}$$

For the moment, we suppose that X^i, $i = 1,2$, are strong solutions. We shall relax this assumption later on. It is well known that the $\mathbb{F}^{X^i}$ characteristics of X^i are determined by μ_i and σ_i.

Now, let $(\Omega',\mathscr{G},\mathbb{G})$ be a filtered probability space supporting a two dimensional Brownian motion, W, where $\mathbb{G}$ is the filtration generated by W and $\Omega' = \Omega'_1 \times \Omega'_2$. The problem of constructing a semimartingale copula for X^i is equivalent to finding functions $m = [m_1,m_2]^T : \mathbb{R}^2 \to \mathbb{R}^2$ and $\Sigma = [\sigma_{ij}] : \mathbb{R}^2 \to L(\mathbb{R}^2,\mathbb{R}^2)$ such that the $\mathbb{F}^Y$ characteristic triple of the diffusion process $Y = (Y^1,Y^2)$, solving the SDE

$$dY(t) = m(Y(t))dt + \Sigma(Y(t))dW(t), \quad Y^i(0) = x^i, \tag{2.8}$$

satisfies Definition 2.3.1.

Remark 2.3.5. Note that, in the diffusion case, the filtered stochastic basis $(\Omega',\mathscr{G},\mathbb{G})$ is not constructed according to the canonical setting. In fact, in this case, the filtration $\mathbb{G}$ may be strictly larger (or smaller) than $\mathbb{F}^Y$. However, in view of Jacod and Shiryaev [20, Sect. 2.26, Theorem III 2.26], if a unique solution-measure to (2.8) exists, it is the unique probability measure on the canonical space $(\Omega,\mathscr{F}^Y,\mathbb{F}^Y)$ such that the process Y has characteristic triple $(m(Y),\sigma(Y),0)$. We can therefore construct a semimartingale copula on the stochastic basis $(\Omega',\mathscr{G},\mathbb{G})$.

Proposition 2.3.5. *Suppose that the function Σ is chosen so that a strong solution of* (2.8) *exists and*

$$\sigma^2_{11}(x,y) + \sigma^2_{12}(x,y) = \sigma^2_1(x), \quad \sigma^2_{21}(x,y) + \sigma^2_{22}(x,y) = \sigma^2_2(y), \tag{2.9}$$

[5] The first requirement in condition (iii) of Definition 2.3.1, namely that Y is consistent with respect to $\mathbb{F}^{Y^i}$, $i = 1,2$, is trivially satisfied since the characteristics are deterministic.

with

$$\sup_y \sigma_{12}^2(x,y) \le \sigma_1^2(x) \quad \textit{and} \quad \sup_x \sigma_{21}^2(x,y) \le \sigma_2^2(y). \tag{2.10}$$

In addition, suppose that the function m satisfies

$$m_1(x,y) = \mu_1(x), \quad m_2(x,y) = \mu_2(y). \tag{2.11}$$

Then the processes Y^i are diffusion processes, and $m = [m_1, m_2]^T$ and $\Sigma = [\sigma_{ij}]$ define a semimartingale copula for X^1, X^2, which we term a diffusion copula.

The next proposition is the counterpart of Proposition 2.3.5 when the system of SDEs (2.8) does not necessarily admit a unique strong solution. We now only assume that the coefficients of Eq. (2.7) satisfy all conditions needed for existence of a weak solution.

Proposition 2.3.6 (see [4, 32]). *Let Y be a (weak) solution of the SDE* (2.8). *Suppose that the function Σ is measurable and satisfies the conditions (2.9) with*

$$\sup_y \sigma_{12}^2(x,y) < \sigma_1^2(x) \quad \textit{and} \quad \sup_x \sigma_{21}^2(x,y) < \sigma_2^2(y). \tag{2.12}$$

In addition, suppose that the function m satisfies (2.11). *Then the processes Y^i are diffusion processes, and $m = [m_1, m_2]^T$ and $\Sigma = [\sigma_{ij}]$ define a semimartingale copula for X^1, X^2, which we term a diffusion copula.*

Note that in view of (2.11) we have no freedom in the choice of the function m. However, we do have freedom in the choice of Σ. Dependence between components of Y is then fully described in terms of the functions σ_{12} and σ_{21}. It is easy to verify that in the diffusion case

$$\mathrm{Dvol}(Y_T) = \mathbb{E}\left(\int_0^T \left(\sigma_{11}(Y_s^1, Y_s^2)\sigma_{21}(Y_s^1, Y_s^2) + \sigma_{12}(Y_s^1, Y_s^2)\sigma_{22}(Y_s^1, Y_s^2)\right) ds\right).$$

We immediately have the following bounds for the d-volume associated to Y:

$$0 < \mathrm{Dvol}(Y_T) \le \mathbb{E}\left(\int_0^T \left(\sigma_1(Y_s^1)^2 + \sigma_2(Y_s^2)^2\right) ds\right) \tag{2.13}$$

as this condition is necessary for the diffusion matrix to be nonnegative definite (see [32] for a more detailed discussion).

Copulae between finite Markov chains. Since, in general, when dealing with Markov chains we are given their generator matrix, rather than their characteristic triple, we find it convenient to work with a counting measure which can be more directly related to the infinitesimal generator of the chain.

Finite Markov chains and related random measures. As before, let $(\Omega, \mathscr{F}^X, \mathbb{P})$ be the underlying probability space. We consider on this space a stochastic process $X = (X_t)_{t \ge 0}$ with values in a finite set $\mathscr{X} = \{1, 2, \dots, N\} \subset \mathbb{N}$. As usual, by $\mathbb{F}^X$

we shall denote the natural filtration generated by X. As anticipated above, rather than looking at the jump characteristic of the chain, we are going to introduce a counting measure that can be more easily associated to its infinitesimal generator. To this end, for any two states $i,j \in \mathscr{X}$ such that $i \neq j$, we define the following $\mathbb{F}^X$-optional random measure on $[0,\infty)$:

$$N^{ij}((0,t]) = \sum_{0<s\leq t} \mathbb{1}_{\{X_{s-}=i, X_s=j\}}. \tag{2.14}$$

We shall simply write $N^{ij}(t)$ in place of $N^{ij}((0,t])$. Obviously, $N^{ij}(t)$ represents the number of jumps from state i to state j that X executes over the time interval $(0,t]$. Let us denote by ν^{ij} the dual predictable projection, with respect to $\mathbb{F}^X$, of the random measure N^{ij}. We are now going to relate the collection of the random measures ν^{ij}, $i,j \in \mathscr{X}$, $i \neq j$, to the infinitesimal generator of the chain X. To this end, let us define a matrix valued function A on $[0,\infty)$ by

$$A(t) = [\lambda_{i,j}(t)]_{i,j\in\mathscr{X}}, \tag{2.15}$$

where $\lambda_{i,j}$ are real valued, locally integrable functions on $[0,\infty)$ such that for $t \geq 0$ and $i,j \in \mathscr{X}$, $i \neq j$, we have

$$\lambda_{i,j}(t) \geq 0 \quad \text{and} \quad \lambda_{i,i}(t) = -\sum_{j\neq i} \lambda_{i,j}(t).$$

Note that $\lambda_{i,j}(t)$ is the time-t intensity of jump from state i to state j. The following proposition establishes the connection between the random measures ν^{ij}, $i,j \in \mathscr{X}$, $i \neq j$, and the infinitesimal generator of X.

Proposition 2.3.7 (see [4]). *The process X is a Markov chain (with respect to $\mathbb{F}^X$) with infinitesimal generator $A(t) = [\lambda_{i,j}(t)]$ iff the dual predictable projections with respect to $\mathbb{F}^X$ of the counting measures $N^{ij}(dt)$, $i,j \in \mathscr{X}$, are of the form*

$$\nu^{ij}(dt) = \mathbb{1}_{\{X_{t-}=i\}} \lambda_{i,j}(t)dt. \tag{2.16}$$

Copulae between Markov chains. As usual, we shall only consider the case of bivariate Markov chains. The general multivariate case can be treated similarly. In the rest of this section we denote by $\mathscr{S}$ and $\mathscr{O}$ two finite sets. Let $X = (X^1, X^2)$ denote a two dimensional Markov chain on $\mathscr{X} = \mathscr{S} \times \mathscr{O}$, with generator function $A(t) = [\lambda^X_{ih,jk}(t)]_{i,j\in\mathscr{S},k,h\in\mathscr{O}}$. Assume that the following conditions hold:

$$\sum_{k\in\mathscr{O}} \lambda^X_{ih,jk}(t) = \sum_{k\in\mathscr{O}} \lambda^X_{ih',jk}(t), \quad \forall h,h' \in \mathscr{O},\ \forall i,j, \in \mathscr{S}, i \neq j, \tag{2.17a}$$

$$\sum_{j\in\mathscr{S}} \lambda^X_{ih,jk}(t) = \sum_{j\in\mathscr{S}} \lambda^X_{i'h,jk}(t), \quad \forall i,i' \in \mathscr{S},\ \forall k,h \in \mathscr{O}\ h \neq k. \tag{2.17b}$$

Intuitively, conditions (2.17a) and (2.17b) requires that the jump intensity of the component X^1 does not depend on the state of X^2 and vice versa. As shown in the following proposition, conditions (2.17a) and (2.17b) are sufficient to yield Markovianity of the components X^i, $i = 1,2$, in the filtration $\mathbb{F}^X$. Moreover, we obtain an explicit characterization of the infinitesimal generator matrix of the components X^i in terms of $A(t)$.

Proposition 2.3.8 (see [4]). *Suppose that conditions (2.17a) and (2.17b) hold and define*

$$f_{i,j}(t) := \sum_{k\in\mathscr{O}} \lambda^X_{ih,jk}(t), \quad i,j \in \mathscr{S}, i \neq j, \quad f_{i,i}(t) := -\sum_{j\in\mathscr{S}, j\neq i} f_{i,j}(t), \quad \forall i \in \mathscr{S}, \tag{2.18}$$

and

$$g_{h,k}(t) := \sum_{j\in\mathscr{S}} \lambda^X_{ih,jk}(t), \quad k,h \in \mathscr{O}, h \neq k, \quad g_{h,h}(t) := -\sum_{k\in\mathscr{O}, k\neq h} g_{h,k}(t), \quad \forall h \in \mathscr{O}. \tag{2.19}$$

Then the components X^1 and X^2 of the Markov chain X are Markov chains with respect to their natural filtrations with generator functions $A^1(t) = [f_{i,j}(t)]_{i,j\in\mathscr{S}}$ and $A^2(t) = [g_{h,k}(t)]_{k,h\in\mathscr{O}}$, respectively.

In view of Proposition 2.3.8, it is clear how to construct the generator of a bivariate Markov chain whose components have prescribed infinitesimal generators.

Corollary 2.3.1. *Consider two Markov chains Y_1 and Y_2, with respect to their own filtrations, $\mathbb{F}^{Y^1}$ and $\mathbb{F}^{Y^2}$, with values in $\mathscr{S}$ and $\mathscr{O}$, respectively. Suppose that their generators are $A^1(t) = [\lambda^{Y^1}_{i,j}(t)]_{i,j\in\mathscr{S}}$ and $A^2(t) = [\lambda^{Y^2}_{h,k}(t)]_{h,k\in\mathscr{O}}$. Next, consider the system of equations in the unknowns $\lambda^X_{ih,jk}(t)$, where $i,j \in \mathscr{S}$, $h,k \in \mathscr{O}$ and $(i,h) \neq (j,k)$:*

$$\begin{cases} \sum_{k\in\mathscr{O}} \lambda^X_{ih,jk}(t) = \lambda^{Y^1}_{i,j}(t), & \forall h \in \mathscr{O}, \forall i,j \in \mathscr{S}, i \neq j \\ \sum_{j\in\mathscr{S}} \lambda^X_{ih,jk}(t) = \lambda^{Y^2}_{h,k}(t), & \forall i \in \mathscr{S}, \forall h,k \in \mathscr{O}, h \neq k. \end{cases} \tag{2.20}$$

Suppose that the above system admits a solution such that the matrix function $A(t) = [\lambda^X_{ih,jk}(t)]_{i,j\in\mathscr{S},k,h\in\mathscr{O}}$, with

$$\lambda^X_{ih,ih}(t) = -\sum_{(j,k)\in\mathscr{S}\times\mathscr{O},(j,k)\neq(i,h)} \lambda^X_{ih,jk}(t), \tag{2.21}$$

properly defines an infinitesimal generator function of a Markov chain with values in $\mathscr{S}\times\mathscr{O}$. Consider a bivariate Markov chain $X := (X_1, X_2)$ on $\mathscr{S}\times\mathscr{O}$ with generator function $A(t)$. Then the components X_1 and X_2 are Markov chains with respect to their own filtrations, with generators $A^1(t)$ and $A^2(t)$.

Note that, typically, system (2.20) contains many more unknowns than equations. In fact, given that the cardinalities of $\mathscr{S}$ and $\mathscr{O}$ are $K_{\mathscr{S}}$ and $K_{\mathscr{O}}$, respectively, the system consists of $K_{\mathscr{S}}(K_{\mathscr{S}}-1)+K_{\mathscr{O}}(K_{\mathscr{O}}-1)$ equations in $K_{\mathscr{S}}K_{\mathscr{O}}(K_{\mathscr{S}}K_{\mathscr{O}}-1)$ unknowns. Thus, in principle, one can create several bivariate Markov chains X with given margins.

Now we embed the above results in the framework of semimartingale copulae. First note that the elements of the mark space of X^1, say $\mathscr{J}^1$, can be uniquely identified with the integers $l=j-i$ with $i,j\in\mathscr{S}$. Similarly, the elements of the mark space of X^2, say $\mathscr{J}^2$, can be uniquely identified with the integers $r=k-h$ with $h,k\in\mathscr{O}$. Thus, we can uniquely construct the jump characteristic[6] of X^1 from the collection of the random measures ν^{ij} in the following way:

$$\nu^1([0,t),l)=\sum_{i,j:j-i=l}\nu_t^{ij},$$

where, as usual, $\nu^{ij}(dt)=\mathbb{1}_{\{X_{t-}=i\}}\lambda_{i,j}(t)dt$. Now let ν be a random measure on $\mathscr{J}^1\times\mathscr{J}^2$ given by

$$\begin{aligned}\nu(dt,(0,0))&=0,\\ \nu(dt,(l,r))&=\sum_{i,j:j-i=l}\sum_{h,k:k-h=r}\nu^{ih,jk}(dt),\end{aligned}\tag{2.22}$$

where, as usual, $\nu^{ih,jk}(dt)=\mathbb{1}_{\{X^1_{t-}=i,X^2_{t-}=h\}}\lambda^X_{ih,jk}(t)dt$.

Proposition 2.3.9 (see [4]). *Let ν be the random measure in (2.22). If, for $i,j\in\mathscr{S}$, $h,k\in\mathscr{O}$ and $(i,h)\neq(j,k)$, $\lambda^X_{ih,jk}$ is a positive solution of the system (2.20), then the triplet $(0,0,\nu)$ is a semimartingale copula for Y^1 and Y^2.*

It needs to be stressed that in the case of semimartingales that are also Markov processes, we can either apply the semimartingale copula methodology described above, or the Markov copula methodology, described in the following section, to model dependence between univariate components of a multivariate process with preservation of margins. This remark applies, in particular, to the Markov chain case of Proposition 2.3.9, as well as to the diffusion case of Proposition 2.3.6. In Sect. 2.5 we shall use semimartingale copulae for Markov chains in order to price ratings triggered step-up bonds.

2.4 Markov Copulae

In this section, which is based on [6, 31], we tackle the problem of defining and constructing "Markov copulae" using infinitesimal generators. First we introduce the class of *consistent* Markov processes for which we next define and construct Markov copulae.

---

[6] Analogous arguments hold for the jump characteristic of X^2 as well.

2.4.1 Consistent Markov Processes

Let $E = \times_{i=1}^{d} E_i$, where E_i are locally compact separable spaces. We recall the notation:

1. For any index set $I \subset \{1,\dots,d\}$, we denote by I^c its complementary set, and we write $E_I = \times_{i\in I} E_i$. For $x \in E$ we use the notation $x_I = (x_i,\, i \in I)$.
2. $B(E)$ is the space of bounded functions on E endowed with the supremum norm. Likewise, $B(E_I)$ is the space of bounded functions on E_I.
3. For a linear operator $A \subset B(E) \times B(E)$, we denote its domain by $\mathscr{D}(A)$. For a suitably large set $D(E) \subseteq \mathscr{D}(A)$, we let $L(E) := \overline{D(E)}$[7].

To ensure regularity of the sample paths of the Markov processes under consideration, we shall assume that $D(E) \subseteq C_0(E)$ (and therefore $L(E) \subseteq C_0(E)$ as well). Additionally, we assume that $D(E)$ is the closure of a tensor product space, i.e.,

$$D(E) = D_I(E_I) \widehat{\otimes} D_{I^c}(E_{I^c}),$$

for suitable spaces $D_I(E_I) \subseteq C_0(E_I)$ and $D_{I^c}(E_{I^c}) \subseteq C_0(E_{I^c})$.[8] In addition we assume the space $L(E)$ is a separating subspace of $B(E)$. This condition as well will be satisfied in all the cases considered in this survey.

Definition 2.4.1. We define the following subspaces:

1. $B_I(E) := \{f \otimes \mathbb{1}_{E_{I^c}} : f \in B(E_I)\}$,
2. $D_I(E) := \{h \in B_I(E) : (h,g) \in$ the b.p. closure of A for $g \in B(E)$ and $h = f \otimes \mathbb{1}_{E_{I^c}}$ for some $f \in D_I(E_I)\}$[9],
3. $L_I(E_I) := \overline{D_I(E_I)}$, where we assume that $L_I(E_I)$ is a separating subset of the space $B(E_I)$.

Remark 2.4.1. Note that if $(f_n^1, Af_n^1) \overset{b.p.}{\rightarrow} (f,h^1)$ and $(f_n^2, Af_n^2) \overset{b.p.}{\rightarrow} (f,h^2)$, then

$$h^1(x) = h^2(x) = \lim_{t\downarrow 0} \frac{T(t)f(x) - f(x)}{t}, \quad \forall x \in E.$$

This implies that if (f,h^1) and (f,h^2) both belong to the bounded pointwise closure of A, then $h^1 = h^2$.

Let now $\mathbb{P}$ be a probability measure on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{F})$ and let $X = (X^1,\dots,X^d)$ be an $\mathbb{F}$-Markov process under $\mathbb{P}$, taking values in E. In general, the components of a vector $\mathbb{F}$-Markov process are not $\mathbb{F}$-Markovian themselves.

[7] For an operator $A \subset B(E) \times B(E)$ or a subspace X of $B(E)$, the notations $\overline{A}$ and $\overline{X}$ signify that the closure is taken in the $\|\cdot\|_\infty$ norm.

[8] It is shown in [31] that this assumption will be satisfied in all the cases considered in this section.

[9] We implicitly assume that all functions of the form $f \otimes \mathbb{1}_{E_{I^c}}$, where $f \in D_I(E_I)$ belong to the bounded pointwise closure of $D(E)$ (for definition of the b.p. closure see [15]). This is the case in all our applications.

Example 2.4.1. Consider the vector process $Y_t := (W_t, M_t)$, where W is a SBM,

$$M_t := N\left(\int_0^t \exp(\sigma W_s)\, ds\right)$$

and N is a standard Poisson process. It is well known that M is a Cox process, Y is Markov in its natural filtration $\mathbb{F}^Y$, but the component M is not $\mathbb{F}^Y$-Markov. In fact, since the filtration generated by $\int_0^t \exp(\sigma W_s)\, ds$ is contained in $\mathscr{F}_t^W$, W is still an $\mathbb{F}^Y$ Brownian motion and

$$\mathbb{E}(M_{t+s}|\mathscr{F}_t^Y) = M_t + \mathbb{E}\left(\int_t^{t+s} \exp(\sigma W_u)\, du|\mathscr{F}_t^Y\right) = M_t + \exp(\sigma W_t)\frac{2\exp(\frac{\sigma^2}{2}s) - 1}{\sigma^2}.$$

It turns out that the converse statement is also false, namely we construct a vector process which is not Markov itself with all components Markovian (in their natural filtration).

Example 2.4.2. Consider a pair of random times τ_1 and τ_2 with exponential distribution and intensities λ_1 and λ_2, and denote by $\mathbb{H} = \mathbb{H}^1 \vee \mathbb{H}^2$ the minimal filtration making τ_1 and τ_2 stopping times. It can be checked that the indicator processes $H_t^i := \mathbb{1}_{\{\tau_i \le t\}}$, $i = 1, 2$, are Markovian in the respective natural filtrations $\mathbb{H}^i$, $i = 1, 2$. In fact, for $t \ge s$,

$$\mathbb{P}\left(H_t^i = 0 \mid \mathscr{H}_s^i\right) = \mathbb{P}\left(\tau_i \ge t \mid \mathscr{H}_s^i\right) = (1 - H_s^i)\exp(-\lambda_i(t-s)).$$

Assume that the joint distribution of (τ_1, τ_2) is given by a Gaussian copula, i.e.,

$$\mathbb{P}(\tau_1 \le t_1, \tau_2 \le t_2) := C(F_1(t_1), F_2(t_2)),$$

with $C(\cdot,\cdot) := \Phi_2(\Phi^{-1}(\cdot), \Phi^{-1}(\cdot))$, where Φ_2 is the CDF of a bivariate Gaussian random variable with mean vector $(0,0)$ and covariance matrix Σ, Φ is the CDF of a standard Gaussian random variable, and $F_i(t_i) = 1 - \exp(\lambda_i t_i)$, $i = 1, 2$, are the marginal CDFs.
Consider now the bivariate process (H^1, H^2). It is not Markovian in its natural filtration $\mathbb{H}$. To see this, it is sufficient to notice that[10]

$$\begin{aligned}\mathbb{P}(H_t^1 = 1 \mid \mathscr{H}_s) = &\left(\frac{1 - \partial_2 C(F_1(t), F_2(\tau_2))}{1 - \partial_2 C(F_1(s), F_2(\tau_2))}\right)\mathbb{1}_{\{\tau_1 > s, \tau_2 \le s\}} \\ &+ \left(\frac{1 - F_1(t)}{1 - C(F_1(s), F_2(s))}\right)\mathbb{1}_{\{\tau_1 > s, \tau_2 > s\}},\end{aligned}$$

which is clearly nonmeasurable in $\sigma(H_s^1) \vee \sigma(H_s^2)$. In this example, starting from Markovian "marginal processes" H^1 and H^2, we constructed a non-Markovian vector process with given marginal laws.

[10] By $\partial_2 C(u_1, u_2)$ we denote the partial derivative of C with respect to its second variable, evaluated at u_2.

Remark 2.4.2. The above examples show that requiring that the $\mathbb{F}$-Markov process X has $\mathbb{F}$-Markov component X^I is a stringent requirement. However, if the components of a multivariate Markov process X are themselves Markovian, then one can apply the rich analytical apparatus of Markov processes to the analysis of both X and its components. This observation motivates the following definition.

Definition 2.4.2. We say that a Markov process X has the *Markovian consistency property* for X^I (or briefly consistency property if X^I is predetermined) if

$$\mathbb{E}(f(X^I_{t+s}) \mid \mathscr{F}_t) = \mathbb{E}(f(X^I_{t+s}) \mid X^I_t), \quad \forall f \in B(E_I), \tag{2.23}$$

where $X^I = (X^i, i \in I)$.

If, in addition, the law of X^I agrees with the law of a given Markov process Y taking values in E_I, and defined on some probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$, i.e., for any positive integer n, any $t_1, t_2, \ldots, t_n \geq 0$, and any measurable subsets of E_I, $A_1, A_2, \ldots, A_n$,

$$\mathbb{P}(X^I_{t_i} \in A_i, i = 1, 2, \ldots, n) = \tilde{\mathbb{P}}(Y_{t_i} \in A_i, i = 1, 2, \ldots, n), \tag{2.24}$$

then we say that X has the *Markovian consistency property* for (X^I, Y).

Remark 2.4.3. Let $\mathbb{F}^I$ be the natural filtration of the process X^I and let $\mathbb{G}$ be any filtration satisfying $\mathbb{F}^I \subseteq \mathbb{G} \subseteq \mathbb{F}$. It is an immediate consequence of (2.23) and the chain rule for conditional expectation that X^I remains a Markov process with respect to $\mathbb{G}$. In other words, Markovian consistency also implies Markovianity of the component in its own filtration.

2.4.2 Markov Copulae: Generator Approach

In this section, which is based on [6, 31], we tackle the problem of defining and constructing "Markov copulae" using infinitesimal generators. In what follows, we provide conditions on the infinitesimal generator of X that ensure that the Markov consistency property for X^I holds. We can and will assume that the paths of X are $\mathbb{P}$-a.s. in $\mathbb{D}_E[0, \infty)$. By $T(t)$ we denote the semigroup of operators on $B(E)$ defined by the transition function corresponding to X, and by A its infinitesimal generator. We fix an index set $I \subset \{1, \ldots, d\}$. Proposition 2.4.1 below yields a necessary condition for Markovian consistency to hold.

Remark 2.4.4. For $f \in \mathscr{D}(A)$, $Af(x)$ determines the expected infinitesimal evolution of the process $f(X_t)$, given the initial state $X_t = x$. Intuitively, for X^I to have the Markov property in the filtration $\mathbb{F}$, its infinitesimal probabilistic behavior should not depend on the state of the components X_{I^c}. In terms of the infinitesimal generator, this means that for a function f which is a constant function of x_{I^c}, $Af(x)$ should only depend on the variables x_I. This intuition is formalized in the following proposition.

Proposition 2.4.1 (see [6, 31]). *Assume that X is conservative and that its component X^I is $\mathbb{F}$-Markov. If $f \in D_I(E)$ and $D(E) \times A(D(E)) \ni (f_n, Af_n) \stackrel{b.p.}{\to} (f, h)$, then h belongs to $B_I(E)$.*

The next proposition gives sufficient conditions on the infinitesimal generator of a Markov process that ensure the Markovian consistency for components X^I. In addition, it provides an explicit characterization of the infinitesimal generator of X^I, which will be very helpful in the actual construction of Markov copulae.

Proposition 2.4.2. *Let A be the infinitesimal generator of an E-valued Markov process X, and assume that X is conservative. In addition, assume that*

$$\begin{aligned} &\forall g \in D_I(E_I), \text{ there is a sequence } f_n \in D(E), \text{and } h^g \in C_0(E_I) \text{ such that} \\ &(f_n, Af_n) \stackrel{b.p.}{\to} (g \otimes \mathbb{1}_{E^{I^c}}, h^g \otimes \mathbb{1}_{E^{I^c}}). \end{aligned} \tag{2.25}$$

Then:
(i) we can define an operator $(A^I, D_I(E_I))$ by

$$A^I g = h^g, \text{ for all } g \in D_I(E_I). \tag{2.26}$$

Assume, in addition, that $R(\lambda \cdot \mathrm{Id} - A^I)$ is dense in $L_I(E_I)$ for some $\lambda > 0$. Then:
(ii) $\overline{A^I}$ generates a strongly continuous contraction semigroup $T^I(t)$ on $L_I(E_I)$, and
(iii) X^I is the unique Markov process corresponding to $T^I(t)$.

Later in this section we shall be concerned with constructing operators A satisfying Markovian consistency conditions, starting from the infinitesimal generators of the component processes X^I. The following corollary will be useful to this end:

Corollary 2.4.1. *Let A, X and $\mathbb{F}$ be as in Proposition 2.4.2. Assume that condition (2.25) holds and that the operator $\overline{A^I}$, as defined in Proposition 2.4.2, generates a strongly continuous contraction semigroup on $L_I(E_I)$. Then X^I is an $\mathbb{F}$-Markov process and A^I coincides with the infinitesimal generator of X^I on $D_I(E_I)$.*

Finally, we state sufficient conditions on the multivariate generator A such that the component X^I is Markovian with given finite dimensional distributions.

Proposition 2.4.3. *Let X be a Markov process on E with generator A, and let Y be an E_I-valued Markov process, with infinitesimal generator A^Y. Suppose that the conditions of Proposition 2.4.2 are satisfied and define A^I by (2.26). Moreover, suppose that $A^I = A^Y$ on $D_I(E_I)$. Then X satisfies the Markovian consistency conditions for (X^I, Y).*

In Proposition 2.4.2 we considered a vector-valued $\mathbb{F}$-Markov process and provided conditions on its generator ensuring that a given component is also Markov. Now we consider the problem from the opposite perspective. Given a collection of Markov processes, say (Y^i), where Y^i is E_i-valued, $i = 1, \ldots, d$, we want to construct a vector process $X = (X^1, X^2, \ldots, X^d)$ with values in $E = E_1 \times E_2 \times \cdots \times E_d$, that is Markov with respect to its natural filtration, say $\mathbb{F}$, and has the Markovian consistency property for (X^i, Y^i), $i = 1, \ldots, d$.

Suppose that we are given a collection of operators, say $\mathscr{A} = \{(A^i, \mathscr{D}(A^i)) : i = 1, \ldots, d\}$, such that the closure of A^i restricted to $D_i(E_i) \subset \mathscr{D}(A^i)$ generates a strongly continuous, positive contraction semigroup on $L_i(E_i) := \overline{D_i(E_i)} \subseteq C_0(E_i)$. We denote by Y^i, $i = 1, \ldots, d$, the corresponding E_i-valued Markov processes.

Definition 2.4.3. Assume $D(E) \subseteq \hat{\otimes}_{i=1}^d D_i(E_i)$. Let $\mathscr{C}^{\mathscr{A}} \subset \mathscr{L}(B(E), B(E))$ be a set of linear operators satisfying[11]:
i) For every element A in $\mathscr{C}^{\mathscr{A}}$, the operator $\overline{A|_{D(E)}}$ generates a strongly continuous positive contraction semigroup on $L(E) := \overline{D(E)}$,
ii) For each $i = 1, 2, \ldots, d$, and for every $g \in D_i(E_i)$ there exists $f_n \in D(E)$ such that $(f_n, Af_n) \overset{b.p.}{\to} (g \otimes \mathbb{1}_{E_{i^c}}, A^i g \otimes \mathbb{1}_{E_{i^c}})$.

If $\mathscr{A}$ is not empty, then we call an element in $\mathscr{C}^{\mathscr{A}}$ a Markov copula for $\mathscr{A}$ with respect to $D(E)$.

Remark 2.4.5. The question whether $\mathscr{A}$ is nonempty is not easy. In some special cases it is shown by construction that $\mathscr{A}$ is not empty. In general, this question requires analysis of existence of appropriate solutions to an operator equation. In Sect. 2.4.3 we discuss symbolic Markov copulae. The question of existence of such copulae corresponds to the question of existence of solutions to certain functional equations.

Let us fix i. Then in view of Corollary 2.4.1, the process X^i is Markov with respect to the natural filtration of X and admits a generator A^i. Now by applying Proposition 2.4.3 we have

Proposition 2.4.4. *Let A be an element of $\mathscr{C}^{\mathscr{A}}$. Then the canonical Markov process $X = (X^1, \ldots, X^d)$ corresponding to the semigroup generated by A has the Markovian consistency property for (X^i, Y^i), $i = 1, \ldots, d$.*

2.4.2.1 Examples

In this subsection, we consider some important classes of Markov processes and provide a constructive answer to the problem introduced at the beginning of this section. Given a collection of Markov processes Y^i, how do we construct a multivariate process $X = (X^1, \ldots, X^d)$ that is Markov with respect to its natural filtration $\mathbb{F}$, and satisfies the Markovian consistency conditions for (X^i, Y^i)? We shall construct elements of the set $\mathscr{C}^{\mathscr{A}}$ in the following cases, which we consider important for applications:

1. The marginal processes Y^i, generated by A^i, are $\mathbb{R}$-valued diffusion processes;
2. The marginal processes Y^i, generated by A^i, are $\mathbb{R}$-valued pure jump Markov processes;

[11] We use the notation introduced in Definition 2.4.1.

3. The marginal processes Y^i, generated by A^i, are $\mathbb{R}$-valued diffusion modulated jump processes.

Diffusion Processes. We consider a collection of d operators,

$$A^i f(x_i) = b_i(x_i)\partial_{x_i} f(x_i) + \frac{1}{2}\sigma_i(x_i)^2 \partial_{x_i}\partial_{x_i} f(x_i), \tag{2.27}$$

on $D_i(E_i) := C_c^\infty(E_i)$, where the coefficients $b_i(x_i)$ and $\sigma_i(x_i)$ are given functions in $C_b^2(E_i)$. We know that A^i is a core of the infinitesimal generator of a Markov diffusion Y^i, taking values in $E_i = \mathbb{R}$ (see [15, Chap. 8, Theorem 2.1]). In this section we assume that $E = \mathbb{R}^d$.

In what follows we use the shorthand notation $I^{\{i\}^c} \hat{\otimes} A^i$ for $I^1 \hat{\otimes} \ldots \hat{\otimes} A^i \hat{\otimes} \ldots \hat{\otimes} I^d$, where I^m is the identity operator on the space $B(E_m)$, for $m = 1, \ldots, d$.

Proposition 2.4.5. *Let A^i be as in (2.27) and define a linear operator A on $D(E) := C_0^\infty(E)$ as*

$$Af(x) := \sum_{i=1}^{d} I^{\{i\}^c} \hat{\otimes} A^i f(x) + \sum_{i,j=1, i\neq j}^{d} \frac{1}{2} a_{ij}(x_i, x_j)\partial_{x_i}\partial_{x_j} f(x) \tag{2.28}$$

where $a_{ij}(x_i, x_j)$ are such that $a_{ii}(x_i) = \sigma_i^2(x_i)$ and the (diffusion) matrix $\Sigma(x) = [a_{ij}(x_i, x_j)]$ is symmetric nonnegative definite and admits a square root $[\sigma_{ij}] := \Sigma^{\frac{1}{2}} \in C_b^2(E)$. Then the operator A is a Markov copula for $\{(A^i, \mathscr{D}(A_i)) : i = 1, \ldots, d\}$.

Remark 2.4.6. In view of Proposition 2.4.4, the (canonical) Markov process X, corresponding to the semigroup generated by A, has the Markovian consistency property for (X^i, Y^i).

Remark 2.4.7. Note that dependence between the components X^i is entirely characterized by the functions $a_{ij}(\cdot,\cdot)$, $i \neq j$. Therefore, every diffusion copula can be associated to a particular choice of the functions $a_{ij}(\cdot,\cdot)$.

Markov jump processes: General case. In this section, we assume that $E_i \subset \mathbb{R}$ are compact sets for all i. For $i = 1, \ldots, d$, we consider a family of operators on $D(E_i) := C(E_i)$ given by

$$A^i f(x_i) = \eta^i(x_i) \int_{E_i} (f(z_i) - f(x_i))\, \nu^i(x_i, dz_i), \tag{2.29}$$

where $\eta^i(x_i)$ are continuous functions, $\nu^i(x_i, dz_i) \in \mathscr{P}(E_i)$, and the mapping $x_i \to \nu^i(x_i, B)$ is continuous for all i and $B \subset E_i$. In view of the discussion in [15, Chap. 8, Sect. 3], A^i are the generators of pure jump Feller processes taking values in E_i, $i = 1, \ldots, n$.

Proposition 2.4.6 (see [32]). *Let A^i be as in (2.29), and define an operator A on $C(E)$ as*

$$Af(x) := \sum_{i=1}^{d} I^{\{i\}^c} \hat{\otimes} A^i f(x) + \sum_{S \in \mathscr{J}_2} \lambda^S(x) \int_E (f(z) - f(x))\, \nu^S(x,dz) \tag{2.30}$$
$$- \sum_{i=1}^{d} \sum_{S \in \mathscr{J}_2 : i \in S} \lambda^S(x) \int_E (f(z) - f(x))\, \nu^{\{i\}}(x,dz),$$

where $\mathscr{J}_n = \{S \in 2^{\{1,\dots,d\}} : \mathrm{card}(S) \geq n\}$ *and:*
i) $\nu^S(x,dz) \in \mathscr{P}(E)$ *is defined for* $S \in 2^{\{1,\dots,d\}} \setminus \emptyset$ *as*

$$\nu^S(x,dz) := \otimes_{i \in S} \nu^i(x_i, dz_i) \otimes_{j \in S^c} \delta_{x_j}(dz_j),$$

ii) for any $S \in \mathscr{J}_2$*, the functions* λ^S *are nonnegative, continuous, and*

$$\sum_{S \in \mathscr{J}_2 : i \in S} \lambda^S(x) \leq \eta^i(x_i), \quad \forall x \in E, \forall i \in \{1,\dots,d\}. \tag{2.31}$$

Let $D(E) = C(E)$*. Then the operator* A *is a Markov copula for* $\{(A^i, \mathscr{D}(A_i)) : i = 1,\dots,d\}$*.*

Remark 2.4.8. a) Notice that we can rewrite (2.30) in the form

$$Af(x) := \sum_{S \in \mathscr{J}_1} \lambda^S(x) \int_E (f(z) - f(x))\, \nu^S(x,dz), \tag{2.32}$$

with $\lambda^{\{i\}}(x) = \eta^i(x_i) - \sum_{S \in \mathscr{J}_2 : i \in S} \lambda^S(x)$ for all i. This, in particular, implies that $\sum_{S \in \mathscr{J}_1 : i \in S} \lambda^S(x) = \eta^i(x_i)$.
b) In view of Proposition 2.4.4, the Markov process X, corresponding to the semigroup generated by A, has the Markovian consistency property for (X^i, Y^i), $i = 1,\dots,d$.

Markov jump processes: Case of space homogeneous jump size distribution. In this subsection, we consider special pure jump processes with generators $\{A^i, C(E_i)\}$ defined by

$$A^i f(x_i) = \eta^i(x_i) \int_{E_i} (f(z_i) - f(x_i))\, \nu^i(dz_i), \tag{2.33}$$

where $\eta^i(x_i)$ are continuous and $\nu^i(dz_i) \in \mathscr{P}(E_i)$. Let $D_i(E_i) = C(E_i)$. The jump distribution $\nu^i(dz_i)$ is space homogeneous (does not depend on x). It turns out that it is possible to construct multivariate Markov jump processes with an arbitrary jump distribution.

Proposition 2.4.7 (see [31]). *Let* A^i *be as in (2.33) and let*

$$Af(x) := \sum_{i=1}^{d} I^{\{i\}^c} \hat{\otimes} A^i f(x) + \sum_{S \in \mathscr{J}_2} \lambda^S(x) \int_E (f(z) - f(x))\, \nu^S(dz) \qquad (2.34)$$
$$- \sum_{i=1}^{d} \sum_{S \in \mathscr{J}_2 : i \in S} \lambda^S(x) \int_E (f(z) - f(x))\, \nu^{\{i\}}(dz)$$

be an operator on $C(E)$, where
i) $\nu^S(dz) \in \mathscr{P}(E)$ is defined as

$$\nu^S(dz) := C^S(\nu^i(dz_i), i \in S) \otimes_{j \in S^c} \delta_{x_j}(z_j),$$

for some copula function $C^S : [0,1]^S \to [0,1]$,
ii) the nonnegative continuous functions λ^S are such that:

$$\sum_{S \in \mathscr{J}_2 : i \in S} \lambda^S(x) \leq \eta^i(x_i), \quad \forall x \in E, \forall i \in \{1, \dots, d\}.$$

Let $D(E) := C(E)$. Then the operator A is a Markov copula for $\{(A^i, \mathscr{D}(A_i)) : i = 1, \dots, d\}$.

Diffusion modulated Markov jump processes. Let Y be a diffusion process in $\mathbb{R}^n$, with infinitesimal generator $\mathscr{L}$ given by

$$\mathscr{L} f(y) = b(y)^T \nabla f(y) + \mathrm{trace}(a(y) \nabla \nabla^T) f(y),$$

where $b(\cdot)$ and $a(\cdot)$ are regular enough to ensure that $\overline{\mathscr{L}}|_{C_c^\infty(E)}$ generates a strongly continuous contraction semigroup on $C_0(\mathbb{R}^n)$. Using $\mathscr{L}$ we define a collection of d operators on $C_c^{\infty,0}(\mathbb{R}^n \times E_i)$, $i = 1, \dots, d$, by

$$A^i f(y, x_i) = \mathscr{L} \hat{\otimes} I^i f(y, x_i) + \tilde{A}^i f(y, x_i), \qquad (2.35)$$

where I^i is the identity operator on E_i (a compact subset of $\mathbb{R}$) and

$$\tilde{A}^i f(y, x_i) = \eta^i(y, x_i) \int_{E_i} (f(y, z_i) - f(y, x_i))\, \nu^i(y, dz_i), \qquad (2.36)$$

$\eta^i(\cdot, \cdot)$ is a continuous and bounded function of both arguments, and $\nu^i(y, dz_i)$ is a probability measure for every y, such that for every measurable set B in E_i the map $y \to \nu^i(y, B)$ is continuous and bounded for $i = 1, 2, \dots, d$.

Using our assumptions about $\mathscr{L}$, i.e., boundedness of the operators $\tilde{A}^i$, we see by [15, Chap. 1, Corollary 7.2] that for each i the operator A^i generates a strongly continuous semigroup of operators on $C_0(E_i \times \mathbb{R}^n)$. The functions η^i, $i = 1, \dots, d$, are nonnegative, so $\tilde{A}^i$ satisfies the positive maximum principle. Hence the semigroup generated by A^i is positive and contractive in view of the Ethier and Kurtz theorem [15, Chap. 1, Theorem 7.1]. Hence (2.35) is indeed the generator of a Markov process on $\mathbb{R}$.

Proposition 2.4.8. *Define an operator A on $C_c^{\infty,0}(\mathbb{R}^n \times E)$ by*

$$Af(x,y) := \mathscr{L}\hat{\otimes}If(y,x) + \sum_{i=1}^{d} I^{\{i\}^c}\hat{\otimes}\tilde{A}^i f(y,x) \tag{2.37}$$
$$+ \sum_{S\in\mathscr{J}_2} \lambda^S(y,x)\int_E (f(y,z)-f(y,x))\,\nu^S(y,dz)$$
$$- \sum_{i=1}^{d}\sum_{S\in\mathscr{J}_2:i\in S} \lambda^S(y,x)\int_E (f(y,z)-f(y,x))\,\nu^{\{i\}}(y,dz),$$

where

i) $\nu^S(y,dz) \in \mathscr{P}(E)$ *is defined as*

$$\nu^S(y,dz) := C^S(\nu^i(y,dz_i),\, i\in S)\otimes_{j\in S^c}\delta_{x_j}(dz_j)$$

for some copula function $C^S : [0,1]^S \to [0,1]$,
ii) the nonnegative continuous functions $\lambda^S(y,x)$ *are such that*

$$\sum_{S\in\mathscr{J}_2:i\in S} \lambda^S(y,x) \le \eta^i(y,x_i), \quad \forall (x,y)\in E\times\mathbb{R}^n, \forall i\in\{1,\dots,d\}.$$

Let $D(E) := C_c^{\infty,0}(\mathbb{R}^n \times E)$. *Then* A *is a Markov copula for* $\{(A^i,\mathscr{D}(A_i)) : i = 1,\dots,d\}$.

It is now easy to measure the dependence between the Markovian components (Y,X^i). The d-volume is, in this case[12]:

$$\mathrm{Dvol}((Y,X_T)) = \mathbb{E}\int_0^T \left(\frac{1}{2}\sum_{i,j=1}^{d}|a(Y_s)| + \sum_{S\in\mathscr{J}_2}\mathrm{card(S)}\lambda(Y_s,X_s)\right)ds.$$

2.4.3 Markov Copulae: Symbolic Approach

In the previous section we have presented a construction of a copula between Markov processes in terms of infinitesimal generators. Here, based on [3], we present a *symbolic approach*, which makes use of pseudo-differential operators (PDO). This approach is more transparent and gives relatively simple conditions guaranteeing that a multivariate Markov process have Markovian components with respect to their own filtration. It also allows one to construct a Markov process with prescribed marginal laws. In this approach, to construct the symbol corresponding to a Markov copula, one just has to construct nonnegative definite functions satisfying appropriate conditions, whereas in the approach in [4] one has to construct an operator acting on functions. Thus, the symbolic approach allows one to avoid

[12] For an element S of $\mathscr{S}$ (space of symmetric matrices), the norm $|S|$ is understood to be given by $\sqrt{trace(SS)}$.

using tensor products of infinitesimal generators and investigation of b.p. closure of operators, but it has some limitations which follow from Hoh's theorem.

In Sect. 2.4.3.1 we investigate the connection of Markovian consistency properties with the corresponding PDOs, in particular we study the question of constructing a multivariate Feller process with given marginal laws in terms of symbols of some related PDOs. Examples are provided in Sect. 2.4.3.2. In what follows, we shall only consider time-homogeneous Markov processes.

2.4.3.1 Dependence and Symbols

Consider $X = (X^j, j = 1,\dots,n)$, a time-homogeneous Markov process, defined on an underlying probability space $(\Omega,\mathscr{F},\mathbb{P})$, taking values in $\mathbb{R}^n$.

As before, we are interested in the Markovian consistency properties (see Definition 2.4.2) of a Feller Markov process X. For simplicity of exposition we limit ourselves to one-dimensional margins, i.e., we consider $I = \{j\}$, but all results can be extended to the case of an arbitrary subset of components of the process X.

Therefore our first goal is to provide necessary and sufficient conditions which guarantee that the components of X are Markov processes with respect to their natural filtrations. The second goal is to provide necessary and sufficient conditions which guarantee that the Markovian consistency condition holds for (X^j,Y^j) for a given one-dimensional Markov process Y^j defined on $(\widetilde{\Omega},\widetilde{\mathscr{F}},\widetilde{\mathbb{P}})$.

The Markov consistency properties, which are properties of transition probabilities, can be formulated in terms of the conditional characteristic functions of X_t and X_t^j defined as

$$\lambda_t(x,\xi) := \mathbb{E}\left(e^{-i(X_t-x,\xi)}\Big|X_0 = x\right), \quad \lambda_t^j(x_j,\xi_j) := \mathbb{E}\left(e^{-i(X_t^j-x_j)\xi_j}\Big|X_0^j = x_j\right),$$

$$\psi_t^j(y_j,\xi_j) := \widetilde{\mathbb{E}}\left(e^{-i\xi_j(Y_t^j-y_j)}\Big|Y_0^j = y_j\right).$$

Since there are no explicit formulae for the transition probabilities or conditional characteristic functions of general Markov processes, solving our problems in terms of the entire families λ_t, λ_t^j and ψ_t^j is quite inconvenient. The general form of conditional characteristic functions is well known only in the case of Lévy processes.[13]

So we have to solve our problems using a different language. The key observation is that the study of our problems in terms of the families of functions λ_t, λ_t^j and ψ_t^j turns out to be equivalent to the study of these problems in terms of the Markov semigroups corresponding to X, X^j and Y^j. In Jacob [18] it is shown that for a family of functions λ_t we can compute the semigroup $(T_t)_{t\geq 0}$ corresponding to X in the following way:

$$T_t u(x) := \mathbb{E}^x u(X_t) = (2\pi)^{n/2}\int_{\mathbb{R}^n} e^{i(x,\xi)}\lambda_t(x,\xi)\hat{u}(\xi)d\xi,$$

[13] Note that in the case of Lévy processes these conditions can be significantly simplified in the sense that they can be reduced to considering $t = 1$ only.

where $\hat{u}$ denotes the Fourier transform of the function $u : \mathbb{R}^n \to \mathbb{R}$, that is,

$$\hat{u}(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i(x,\xi)} u(x)dx.$$

Analogous properties hold for the families λ_t^j and ψ_t^j and the corresponding semi-groups, say T_t^j and S_t^j.

Moreover, in view of the results of Courrège [9], the generator A of X, acting on $u \in C_0^\infty(\mathbb{R}^n)$, the space of infinitely differentiable functions with compact support, has a representation

$$Au(x) = -q(x,D)u(x) := -(2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i(x,\xi)} q(x,\xi)\widehat{u}(\xi)d\xi, \tag{2.38}$$

where $q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is a measurable, continuous function in ξ and for every x the function $q(x,\cdot)$ is negative definite. In this context, the function $q(x,\xi)$ is called the symbol of the pseudo-differential operator $q(x,D)$ (cf. [19]), and it has the following form:

$$q(x,\xi) = i(b(x),\xi) + (\xi, a(x)\xi) + \int_{\mathbb{R}^n\setminus\{0\}} \left(1 - e^{i(y,\xi)} + \frac{i(y,\xi)}{1+|y|^2}\right) \nu(x,dy) \tag{2.39}$$

where a, b are Borel measurable functions, $b(x) \in \mathbb{R}^n$, $a(x)$ is a symmetric nonnegative definite matrix, and $\nu(x,dy)$ is a Lévy kernel. Moreover, if $q(x,\xi)$ is continuous (in all variables) then $q(x,D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C(\mathbb{R}^n)$ [19, Vol. 1, Theorem 4.5.7, p. 337]. Analogous results hold for q^j and ρ^j. In particular, in the case of Y^j, the infinitesimal generator B^j, acting on $w \in C_0^\infty(\mathbb{R})$, satisfies

$$B^j w(x_j) = -(2\pi)^{-1/2} \int_{\mathbb{R}} e^{ix_j\xi_j} \widehat{w}(\xi_j)\rho_j(x_j,\xi_j)d\xi_j,$$

where

$$\rho_j(x_j,\xi_j) = ib_j(x_j)\xi_j + c_j(x_j)\xi_j^2 + \int_{\mathbb{R}\setminus\{0\}} \left(1 - e^{iz_j\xi_j} + \frac{iz_j\xi_j}{1+|z_j|^2}\right) \nu_j(x_j,dz_j). \tag{2.40}$$

So we pursue the study of our problems in terms of symbols of pseudo-differential operators. We shall adopt the following convention: We suppose that $f : \mathbb{R} \to \mathbb{R}$ and that $f_j : \mathbb{R}^n \to \mathbb{R}$ is defined by $f_j(x) = f(x_j)$. Note, however, that even though f may be of compact support, f_j will not be a function of compact support. Recall that by e_j we denote the standard unit vector in $\mathbb{R}^n$ with 1 in the j-th position.

Proposition 2.4.9. *Let X be a Feller process with symbol q and the corresponding generator A such that $C_0^\infty(\mathbb{R}^n) \subseteq D(A)$.*
a) For every $w \in C_0^\infty(\mathbb{R})$,

$$Aw_j(x) = -(2\pi)^{-1/2} \int_{\mathbb{R}} e^{ix_j\xi_j} \widehat{w}(\xi_j)q(x,Q_j(\xi))d\xi_j, \tag{2.41}$$

where $Q_j : \mathbb{R}^n \to \mathbb{R}^n$ is the projection $Q_j(\xi) := \mathbf{e}_j \xi_j$.
b) Assume that X^j is a one-dimensional Feller process with symbol q_j and that

$$q(x, e_j \xi_j) = q_j(x_j, \xi_j) \quad \text{for all } x \in \mathbb{R}^n \text{ and } \xi_j \in \mathbb{R}. \tag{2.42}$$

Then for every $w \in C_0^\infty(\mathbb{R})$,

$$A w_j(x) = A_j w(x_j), \ \ \forall x \in \mathbb{R}^n, \tag{2.43}$$

where A is the generator of X, and A_j is the generator of X^j.

The following proposition gives a necessary condition for Markovian consistency for X^j.

Proposition 2.4.10. *Let X be a Feller process with symbol q. If X^j is a* 1*-dimensional Feller processes with symbol q_j, then* (2.42) *holds.*

Before we consider the question of sufficiency of condition (2.42), we introduce conditions, due to Hoh [16], guaranteeing that the pseudo-differential operator $-q(x,D)$ has a unique extension which generates a Feller semigroup. Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite function such that for some positive constants r and c we have $\psi(\xi) \geq c|\xi|^r$ for $|\xi| \geq 1$. We define

$$\lambda(\xi) := (1 + \psi(\xi))^{1/2}$$

and let M be the smallest integer such that $M > (\frac{n}{r} \vee 2) + n$, and set $k = 2M + 1 - n$. For a continuous negative definite symbol $q : \mathbb{R}^n \times \mathbb{R} \to \mathbb{C}$ we make the following assumptions:
C0: the function q is continuous in both variables;
C1: the map $x \mapsto q(x,\xi)$ is k times continuously differentiable and

$$\left| \partial_x^\beta q(x,\xi) \right| \leq c\lambda^2(\xi), \quad \beta \in \mathbb{N}_0^n, |\beta| \leq k;$$

C2: for some strictly positive function $\gamma : \mathbb{R}^n \to \mathbb{R}$,

$$q(x,\xi) \geq \gamma(x)\lambda^2(\xi) \quad \text{for} \quad |\xi| \geq 1, x \in \mathbb{R}^n;$$

C3:

$$\sup_{x \in \mathbb{R}^n} |q(x,\xi)| \underset{\xi \to 0}{\longrightarrow} 0.$$

It is proved in Hoh [16, Theorem 5.24, p. 82] that under C0–C3 the pseudo-differential operator $-q(x,D) : C_0^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ has an extension which generates a Feller semigroup given by $P_t f(x) = \mathbb{E}^x f(X_t)$, where $\mathbb{E}^x$ is expectation with respect to the solution of the associated well-posed martingale problem starting at x.

Theorem 2.4.1. *Let X be a Feller process with symbol q. Assume that*
C4: *the function $q(x, e_j \xi_j)$ as a function of x depends only on x_j*
and denote

$$\tilde{q}_j(x_j,\xi_j) := q(x,e_j\xi_j). \tag{2.44}$$

If $\tilde{q}_j$ satisfies conditions C0–C3, then the component X^j of X is a Feller process with generator given by the symbol $\tilde{q}_j$.

The above theorem demonstrates that (2.42) (with $q_j = \tilde{q}_j$) is a sufficient condition for Markov consistency to hold, provided that $\tilde{q}_j$ satisfies conditions C0–C3.

Corollary 2.4.2. *Let X be a Feller process with symbol q satisfying conditions C0–C4. Then the component X^j of X is a Feller process with generator given by the symbol $\tilde{q}_j$.*

Now we formulate an answer to the problem of Markovian consistency conditions for (X^i,Y^i) which combines the results of Proposition 2.4.10 and Theorem 2.4.1.

Theorem 2.4.2. *Let $X = (X^1,\dots,X^n)$ be a Feller process with symbol q, and $Y^1,\dots,Y^n$ be an n-tuple of one-dimensional Feller processes with symbols $\rho_1,\dots,\rho_j$. Assume that $\rho_1,\dots,\rho_j$ satisfy assumptions C0–C3. The marginal distribution of the j-th coordinate of X is equal to the distribution of Y^j given by the symbol ρ_j if and only if*

$$q(x,e_j\xi_j) = \rho_j(x_j,\xi_j) \quad \textit{for all } x\in\mathbb{R}^n \textit{ and } \xi_j\in\mathbb{R}. \tag{2.45}$$

Therefore the j-th coordinate of X is a Feller process with respect to its natural filtration.

All these considerations allow us to provide an algorithm for construction of an n-dimensional Feller process with given marginal distributions, and such that its components are also Feller processes. In view of Theorem 2.4.2 we can introduce

Definition 2.4.4 (Symbolic Markov copulae). We say that a symbol q is a Markov symbolic copula for symbols $q_1,\dots,q_n$ if for all $j=1,\dots,n$,

$$q(x,e_j\xi_j) = q_j(x_j,\xi_j). \tag{2.46}$$

Now, our aim is to give a recipe for constructing a symbol q, starting from given one dimensional symbols $q_1,\dots,q_n$, such that q satisfies condition (2.46). Taking into account (2.39) and (2.40) we are looking for a vector function b such that

$$b_j(x) = d_j(x_j), \tag{2.47}$$

a symmetric nonnegative definite matrix function a such that

$$a_{jj}(x) = c_j(x_j), \tag{2.48}$$

and a Lévy measure $\nu(x,dy)$ on $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n\setminus\{0\}}\left(1-e^{iy_j\xi_j}+\frac{iy_j\xi_j}{1+|y|^2}\right)\nu(x,dy) = \int_{\mathbb{R}\setminus\{0\}}\left(1-e^{iy_j\xi_j}+\frac{iy_j\xi_j}{1+|y_j|^2}\right)\nu_j(x_j,dy_j). \tag{2.49}$$

The triple (d_j,c_j,ν_j) will be called the characteristic triple for the copula q_j.

2.4.3.2 Examples

Now we present some examples to illustrate how this idea works.

Example 2.4.3 (Product copula). A copula for symbols $q_1,\dots,q_n$ with characteristic triples $(d_j,c_j,\nu_j)_{j=1}^n$ is a *product copula* if its characteristic triple (b,a,ν) is defined by

$$b_j(x) := d_j(x_j), \quad a_{ij}(x) := c_j(x_j)\mathbf{1}_{\{i=j\}},$$

$$\nu(x,dy) := \sum_{j=1}^n \otimes_{k\neq j}\delta_{\{0\}}(dy_k)\otimes \nu_j(x_j,dy_j).$$

It is easy to see that the symbol q corresponding to (b,a,ν) defined above is a copula for given symbols $q_1,\dots,q_n$. It is a copula that corresponds to independent Feller processes.

Example 2.4.4 (Diffusion copula). Consider n one-dimensional diffusion processes with symbols given by $q_j(x_j,\xi_j) = id_j(x_j)\xi_j + c_j(x_j)\xi_j^2$, where d_i, c_i are functions such that q_i satisfies C0–C3. We define q by

$$q(\xi,x) = i(b(x),\xi) + (\xi,a(x)\xi),$$

where the functions $b:\mathbb{R}^n\to\mathbb{R}^n$ and $a:\mathbb{R}^n\to\mathbb{R}^n_n$ satisfy

$$b_j(x) = d_j(x_j), \quad a_{jj}(x) = c_j(x_j) \quad \forall j = 1,\dots,n, \tag{2.50}$$

and moreover b, a are chosen in such a way that C0–C3 hold. Then q is a symbolic copula for $q_1,\dots,q_n$.

Example 2.4.5 (Lévy copulae). Consider n one-dimensional Lévy processes $Z_1,\dots,Z_n$ with Lévy measures $\nu_1,\dots,\nu_n$, and a Lévy copula F. A Lévy measure ν, which is determined by the set of tail integrals defined by formula (2.1), gives a symbol that satisfies (2.49). Therefore the construction of Lévy copulae due to Kallsen and Tankov [21] provides also a construction of symbolic copulae.

Example 2.4.6 (Poisson copula). Consider two one-dimensional Poisson processes. Their symbols are given by $q_i(\xi_i) = (1-e^{i\xi_i})\eta_i$, where η_i are nonnegative constants for $i=1,2$. A symbol q given by

$$q(\xi_1,\xi_2) = (1-e^{i\xi_2})\lambda_{(0,1)} + (1-e^{i\xi_1})\lambda_{(1,0)} + (1-e^{i(\xi_1+\xi_2)})\lambda_{(1,1)},$$

where $\lambda_{(0,1)},\lambda_{(1,0)},\lambda_{(1,1)}$ are nonnegative constants, defines a Markov copula iff $\lambda_{(0,1)},\lambda_{(1,0)},\lambda_{(1,1)}$ satisfy the following system of linear equations:

$$\begin{cases}\lambda_{(0,1)} + \lambda_{(1,1)} = \eta_2,\\ \lambda_{(1,0)} + \lambda_{(1,1)} = \eta_1.\end{cases}$$

The above system has infinitely many solutions which can be parameterized by $\lambda_{(1,1)}$. Since we are interested in positive solutions, we restrict $\lambda_{(1,1)}$ to the interval $[0,\lambda_1\wedge\lambda_2]$. Generalization to the n-dimensional case is immediate.

We can generalize Example 2.4.6 by allowing λ to depend on x. In this case we have to deal with generalized n-dimensional Markov point processes. Moreover, we can construct copulae for Markov jump processes and Markov jump processes with the possibility of common jumps with space homogeneous jump size distribution. All these examples are presented in [3].

2.5 Applications in Finance

We present an application to finance of the results discussed in the previous sections. Towards this end we shall borrow some results from [5]. These results illustrate an application of Markov copulae (and semimartingale copulae) to valuation of so called *rating-triggered, step-up bonds*. We refer to [5] as well as to [1, 10] for some other financial applications of Markov copulae.

Rating-triggered step-up bonds were issued by some European telecom companies in recent 10 years. These products are of interest because they offer protection against credit events other than defaults. In particular, rating-triggered corporate step-up bonds (step-up bonds for short) are corporate coupon issues for which the coupon payment depends on the issuer's credit quality: in principle, the coupon payment increases when the credit quality of the issuer declines. In practice, credit quality is reflected in credit ratings assigned to the issuer by at least one credit rating agency (Moody's-KMV or Standard & Poor's). The provisions linking the cash flows of the step-up bonds to the credit rating of the issuer have different step amounts and different rating event triggers. In some cases, a step-up of the coupon requires a downgrade to the trigger level by both rating agencies. In other cases, there are step-up triggers for actions of each rating agency. Here, a downgrade by one agency will trigger an increase in the coupon regardless of the rating from the other agency. Provisions also vary with respect to step-down features which, as the name suggests, trigger a lowering of the coupon if the company regains its original rating after a downgrade. In general, there is no step-down below the initial coupon for ratings exceeding the initial rating.

Let R_t stand for some indicator of credit quality at time t (note that in this case, the process R may be composed of two, or more, distinct rating processes). Assume that t_i, $i = 1, 2, \ldots, n$, are coupon payment dates. Here we adopt the convention that coupon paid at date t_n depends only on the rating history through date t_{n-1}, that is: $c_n = c(R_t, t \le t_{n-1})$ are the coupon payments. In other words, we assume that no accrual convention is in force.

Assuming that the bond's notional amount is 1, the cumulative discounted cash flow of the step-up bond is (as usual we assume that the current time is 0)

$$(1 - H_T)\beta_T + \int_{(0,T]} (1 - H_u)\beta_u \, dC_u + \beta_\tau Z_\tau H_T, \tag{2.51}$$

where $C_t = \sum_{t_i \le t} c_i$, τ is the bond's default time, $H_t = \mathbb{1}_{\tau \le t}$, and Z_t is a (predictable) recovery process.

2.5.1 Pricing Rating-Triggered Step-Up Bonds via Simulation

Here, using our results on Markov copulae, we shall apply a simulation approach to pricing rating-triggered step-up bonds.

Let us consider a rating-triggered step-up bond issued by an obligor XYZ. Recall that, typically, cash flows associated with a step-up bond depend on ratings assigned to XYZ by both Moody‘s Investors Service (Moody’s in what follows) and Standard & Poor’s (S&P in what follows). Thus, a straightforward way to model joint credit migrations would be to consider a credit migration process R such that $R_t = (M_t, SP_t)$, where M_t and SP_t denote the time t credit rating assigned to XYZ by Moody’s and SP_t, respectively. We assume that the process M is a time-homogeneous Markov chain with respect to its natural filtration, under the statistical probability $\mathbb{P}$, and that its state space is $\mathscr{K}_1 = \{1, 2, \ldots, K_1\}$. Likewise, we assume that SP is a time-homogeneous Markov chain with respect to its natural filtration, under the statistical probability $\mathbb{P}$, and that its state space is $\mathscr{K}_2 = \{1, 2, \ldots, K_2\}$.

Typically, we are only provided with individual statistical characteristics of each of the processes M and SP. Thus, in a sense, we know the marginal distributions of the joint process R under the *statistical* measure $\mathbb{P}$ (where M and SP are considered as “univariate” margins). The crucial issue is thus appropriate modeling of dependence between M and SP. In particular, we want to model dependence, under $\mathbb{P}$, between M and SP so that the joint process R is a time-homogeneous Markov chain, and so that the components M and SP are time-homogeneous Markov chains with given $\mathbb{P}$-generators, say A^M and A^{SP}, respectively. Thus, essentially, we need to model a $\mathbb{P}$-generator matrix, say A^R, so that R is a time-homogeneous Markov chain with $\mathbb{P}$-generator A^R and that M and SP are time-homogeneous Markov chains with $\mathbb{P}$-generators A^M and A^{SP}. We can of course deal with this problem using the theory of Markov copulae.

Towards this end, we fix an underlying probability space $(\Omega, \mathscr{F}, \mathbb{P})$. On this space we consider two univariate Markov chains M and SP, with given infinitesimal $\mathbb{P}$-generators $A^M = [a^M_{ij}]$ and $A^{SP} = [a^{SP}_{hk}]$, respectively. Next, we consider the system of equations in variables $a^R_{ih,jk}$

$$\begin{cases} \sum_{k \in \mathscr{K}_2} a^R_{ih,jk} = a^M_{ij}, \ \forall i, j \in \mathscr{K}_1, i \neq j, \ \forall h \in \mathscr{K}_2, \\ \sum_{j \in \mathscr{K}_1} a^R_{ih,jk} = a^{SP}_{hk}, \ \forall h, k \in \mathscr{K}_2, h \neq k, \ \forall i \in \mathscr{K}_1. \end{cases} \tag{2.52}$$

Now, provided that the system (2.52) has a positive solution, it follows from Corollary 2.3.1 in Sect. 2.3.2 that the resulting matrix[14] $A^R = [a^R_{ih,jk}]_{i,j \in \mathscr{K}_1, h,k \in \mathscr{K}_2}$ satisfies conditions for a $\mathbb{P}$-generator matrix of a bivariate time-homogenous Markov chain, whose components take values in finite state spaces $\mathscr{K}_1$ and $\mathscr{K}_2$ with cardinalities K_1 and K_2, respectively, and, more importantly, they are Markov chains with the

[14] System (2.52) does not involve diagonal elements of A^R. These elements are obtained as $a^R_{ih,ih} = -\sum_{(j,k) \in \mathscr{K}_1 \times \mathscr{K}_2 \setminus \{(i,h)\}} a^R_{ih,jk}$.

same distributions as M and SP under $\mathbb{P}$. Thus, indeed, the system (2.52) essentially serves as a Markov copula between the Markovian margins M, SP and the bivariate Markov chain R.

Note that the system (2.52) contains many more variables than equations. Thus, one can create several bivariate Markov chains R with the given margins M and SP. In financial applications this feature leaves a lot of flexibility for various modeling options and for calibration of the model. For example, as observed by Lando and Mortensen [23], although the ratings assigned by S&P and Moody's to the same company do not necessarily coincide, split ratings are rare and are usually only observed in short time intervals. This feature can easily be modelled using the Markovian copula system (2.52) via imposing side constraints for the unknowns $a^R_{ih,jk}$. In order to model such observed behavior of the joint rating process, we thus impose additional constraints on the variables in the system (2.52). Specifically, we postulate that

$$a^R_{ih,jk} = \begin{cases} 0 & \text{if } i \neq j \text{ and } h \neq k \text{ and } j \neq k, \\ \alpha \min(a^M_{ij}, a^{SP}_{hk}) & \text{if } i \neq j \text{ and } h \neq k \text{ and } j = k, \end{cases} \tag{2.53}$$

where $\alpha \in [0,1]$ is a modeling parameter. Using constraints (2.53) we can easily solve system (2.52) (in this case the system actually becomes fully decoupled) and we can obtain the generator of the joint process. The interpretation of constraints (2.53) is the following: The components M and SP of the process R migrate according to their marginal laws, but they tend to join, that is, they tend to both take the same value. The strength of that tendency is measured by the parameter α. When $\alpha = 0$ then, in fact, the two components are independent processes; when $\alpha = 1$ the intensity of both components migrating simultaneously to the same rating category is maximum (given the specified functional form for the intensities of common jumps).

For pricing purposes the statistical probability measure is changed to the *EMM*. Typically, the Radon-Nikodym density is chosen in such a way that the resulting (risk-neutral) default probabilities are consistent with the term structure of CDS spreads. In addition, we require that the process R, which is Markovian under the statistical measure, is also Markovian under the pricing measure.

We recall that $A^R = [a^R_{ih,jk}]$ is the generator of R under the statistical measure $\mathbb{P}$. In view of Corollary 4.1 in [5], given a vector $h = [h_{11}, \cdots, h_{K_1K_2}] \in \mathbb{R}^{K_1K_2}$, we can change the statistical measure $\mathbb{P}$ to an equivalent "risk-neutral" measure $\mathbb{Q}$ in such a way that R is a time-homogeneous Markov chain under $\mathbb{Q}$, and its $\mathbb{Q}$-infinitesimal generator is given by

$$\widetilde{A}^R = [\widetilde{a}_{ih,jk}],$$

where

$$\widetilde{a}_{ih,jk} = a_{ih,jk}\frac{h_{jk}}{h_{ih}}, \quad \forall i,j \in \mathscr{K}_1, \quad \forall h,k \in \mathscr{K}_2, \quad (i,h) \neq (j,k)$$

and

$$\widetilde{a}_{ih,ih} = -\sum_{(j,k)\in\mathscr{K}_1\times\mathscr{K}_2\setminus\{(i,h)\}} a_{ih,jk}\frac{h_{jk}}{h_{ih}}.$$

Remark 2.5.1. Note that, although the change of measure preserves the Markov property of the joint process R, its components may not be Markov (in their natural filtration) under the new probability measure. This however is not an issue for us, as all we need to conduct computations is the Markov property of the joint process R under the new measure.

An arbitrary choice of the vector h may lead to a heavy parametrization of the pricing model. We suggest that the vector h_{ij} be chosen as follows:

$$h_{ij} = \exp(\alpha_1 i + \alpha_2 j), \quad \forall i \in \mathscr{K}_1, \forall j \in \mathscr{K}_2,$$

where α_1 and α_2 are parameters to be calibrated. It turns out, as the calibration results provided below indicate, that this is a good choice.

2.5.2 Model Calibration and Pricing

The model is fully specified by three parameters, namely α, α_1, α_2, which are calibrated to market data.

Let us consider a vanilla bond, which is equivalent[15] to the given step-up bond. One would presume that the price of a step-up bond is equal to the price of the equivalent vanilla bond plus the (positive) value of the step-up provision. In general, equivalent vanilla bonds are not traded on the market. However, their price can be synthesized by applying a standard bootstrapping-interpolation procedure to the market prices of traded vanilla bonds. Surprisingly, the value of the step-up provision is often negligible or even negative. This was already noted by some recent empirical literature (cf. e.g. [23]), which provides strong evidence that the market typically "underprices" step-up bonds. These findings suggest that step-up bond investors are more risk averse than vanilla bond investors. In particular, on the theoretical level, this means that the pricing kernel implied by step-up bonds prices should be different from that implied by vanilla bonds. For calibration purposes, this implies that the model parameters, or at least those relative to credit migrations, should not be calibrated to vanilla bond prices. Nevertheless, such data provides useful information. In particular, under the assumptions given below, vanilla bond prices can be used to compute a term structure of firm-specific, liquidity-adjusted, discount factors (risk-free rate + liquidity spread).

---

[15] By equivalent, we mean a coupon bearing bond, backed by the same company, whose all provisions, other than the step-up provision, are identical to those of the given step-up bond. That is, maturity and coupon dates are the same, and the coupons of the equivalent bond are equal to the fixed coupons of the step-up bond. In addition, credit risk is the same and liquidity risk is comparable. The term vanilla means that the step-up provision is not present.

Our first assumption is that the vanilla bond market assesses likelihood of the default event in the same way as the CDS (Credit Default Swap) market.[16] Our second assumption is that liquidity risk is priced identically by the step-up and vanilla bond markets.

Given the above, we can apply a standard bootstrapping-interpolation procedure to a pool of reference bonds[17] to obtain a term-structure of firm specific, liquidity adjusted, zero-coupons. The straightforward procedure is briefly described below. We are given a set of J reference bonds with associated cash flows $CF^j_{t^j_i}$, $j = 1, \ldots, J$, and coupon dates $t^j_0 = 0, \ldots, t^j_N = T^j$ such that $T^1 < T^2 < \cdots < T^J$. The cash flows are then adjusted by the default probability implied by the CDS spreads. Let τ denote the default time of the relevant obligor. Then the default adjusted cash flows are $\widetilde{CF}^j_{t^j_i} = CF^j_{t^j_i}\mathbb{Q}(\tau > t^j_i)$. The interpolation-bootstrapping procedure is now applied to the reference bonds with default-risk adjusted cash flows, so that the resulting discount factors account only for the firm specific liquidity spread.[18] At this point, the price of an arbitrary step-up bond can be computed by simulating the evolution of the joint rating process and the relative discounted cash flows.[19] The model parameters, α, α_1, α_2 are calibrated to step-up bond prices.

2.5.2.1 Calibration Results

We shall present now some calibration results. The bond data, obtained from Bloomberg's Corporate Bonds section, is relative to mid market quotes on April 5, 2006.

We calibrated the model parameters to a DT (Deutsche Telecom) step-up issue described in Table 2.1:

Table 2.1 DT step-up issue on April 5, 2006.

ISIN	XS0132407957
Maturity	07/11/11
Coupon	$6\frac{5}{8}$ Annual
Step provision	$\begin{cases} +50 \text{ bps}, & \text{if both downgraded below single Aaa3/A-;} \\ -50 \text{ bps}, & \text{if both subsequently upgraded above Baa1/BBB+} \end{cases}$

[16] This is not necessary since default risk can be inferred from yield spreads in the bond market, but the higher liquidity of the CDS market makes it a preferable choice.

[17] We adopt here terminology from [23] to denote vanilla bonds of several maturities which have comparable liquidity and are issued by the same company as the relevant step-up bond.

[18] Plus market risk spreads other than credit spread.

[19] Simulation seems to be the only feasible computation technique, because of certain path dependencies in the payoff structure, induced by the step-down provision present in most step-up issues. Such path dependency is well explained in [23].

Given the default probability implied by the 5-y CDS spread of DT (46 bps), the liquidity adjusted discount rates are obtained using the above mentioned bootstrapping-interpolation procedure from the following pool of reference bonds (Table 2.2):

Table 2.2 Reference bonds pool on April 5, 2006.

ISIN	Maturity	Coupon	Mid-price
XS0141544691	01/22/07	$5\frac{1}{4}$	1.015698
DE0002317807	05/20/08	$5\frac{1}{4}$	1.031821
XS0242840345	02/02/09	3	0.979798
XS0217817112	04/22/09	3	0.978352
XS0210319090	01/19/10	$3\frac{1}{4}$	0.976716
XS0210318795	01/19/15	4	0.960349

The calibration results are given in Table 2.3:

Table 2.3 Calibration results.

	Model price	Market price
Bond price	1.11705	1.11705
Step-up provision	0.00574	–

We remark that, since our calibration problem is overdetermined (three parameters are calibrated to one piece of data), the value of the step-up provision is not uniquely defined. This can be easily overcome by calibrating the model to more step-up issues of different maturities and/or provisions.

2.5.2.2 Valuation of Step-Up Bonds

Using the calibrated model, we price selected issues of DT step-up bonds; we refer to Tables 2.4 and 2.5 for the description of the bonds.

Table 2.4 DT step-up issue $XS0113709264$ on April 5, 2006.

ISIN	$XS0113709264$
Maturity	07/06/10
Coupon	$6\frac{5}{8}$ Annual
Step provision	$\begin{cases} +50 \text{ bps}, & \text{if both downgraded below single Aaa3/A-;} \\ -50 \text{ bps}, & \text{if both subsequently upgraded above Baa1/BBB+} \end{cases}$

Table 2.5 DT step-up issue $XS0155788150$ on April 5, 2006.

ISIN	$XS0155788150$
Maturity	10/07/09
Coupon	$6\frac{1}{2}$ Annual
Step provision	$\begin{cases} +50 \text{ bps, if both downgraded below Baa1/BBB+;} \\ -50 \text{ bps, if both subsequently upgraded above Baa2/BBB} \end{cases}$

Table 2.6 presents the pricing results as well as the corresponding market quotes. The results are very satisfactory, indicating that the model is robust and prices consistently across maturities and step-up provisions.

Table 2.6 Pricing results using calibrated model.

ISIN	Mkt price/Model price	
	XS0113709264	XS0155788150
Bond price	1.10105/1.103546	1.08435/1.08685
Step-up provision	– /0.003752	– /0.00215

References

1. Assefa, S., Bielecki, T.R., Crépey, S., Jeanblanc, M.: CVA computation for counter party risk assessment and hedging in credit portfolios. Working Paper (2009)
2. Bauerle, N., Blatter, A., Muller, A.: Dependence properties and comparison results for Lévy processes. Math. Methods Operat. Res. **67**(1), 161–186 (2008)
3. Bielecki, T.R., Jakubowski, J., Niewęgłowski, M.: Study of dependence for some stochastic processes part II: Symbolic Markov copulae. Submitted (2009)
4. Bielecki, T.R., Jakubowski, J., Vidozzi, A., Vidozzi, L.: Study of dependence for some stochastic processes. Stoch. Anal. Appl. **26**, 903–924 (2008)
5. Bielecki, T.R., Vidozzi, A., Vidozzi, L.: A Markov copulae approach to pricing and hedging of credit index derivatives and ratings triggered step-up bonds. J. Credit Risk **4**(1) (2008)
6. Bielecki, T.R., Vidozzi, A., Vidozzi, L.: Multivariate Markov processes with given Markovian margins, and Markov copulae. Preprint, Illinois Institute of Technology, Chicago, IL (2008)
7. Cherubini, U., Luciano, E., Vecchiato, W.: Copula Methods in Finance. Wiley Finance Series. Wiley, Chichester (2004)
8. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman & Hall/CRC Financial Mathematics Series. Chapman and Hall/CRC, Boca Raton, FL (2004)
9. Courrège, P.: Sur la forme intégro-différentielle des opérateurs de C_K^∞ dans C satisfaisant au principe du maximum. Sém. Théorie du Potentiel **2** (1965/1966), 3-01–3-48
10. Crépey, S., Jeanblanc, M., Zargari, B.: CDS with counterparty risk in a Markov chain copula model with joint defaults. Working Paper (2009)
11. Darsow, W.F., Nguyen, B., Olsen, E.T.: Copulas and Markov processes. Ill. J. Math. **36**(4), 600–642 (1992)
12. El Karoui, N., Jeanblanc, M., Jiao, Y.: Successive default times. Working Paper (2009)

13. Elliott, R.J.: Finite dimensional filters related to Markov chains. In: Duncan, T.E. et al. (eds.) Stochastic Theory and Adaptive Control, Proceedings of a Workshop, Lawrence, KS, 26–28 Sept 1991. Lecture Notes in Control Information Sciences, vol. 184, pp. 142–160. Springer, Berlin (1992)
14. Embrechts, P.: Did a mathematical formula really blow up Wall Street? Presentation Slides (2009)
15. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New York, Chichester, Brisbane, Toronto, and Singapore (1986)
16. Hoh, W.: Pseudo Differential Operators Generating Markov Processes. Habilitationsschrift, Bielefeld (1998)
17. Ibragimov, R.: Copula-based characterizations for higher order Markov processes. Econ. Theory **25**(3), 819–846 (2009)
18. Jacob, N.: Characteristic Functions and Symbols in the Theory of Lévy Processes. Potential Anal. **8**, 61–68 (1998)
19. Jacob, N.: Pseudo Differential Operators, Markov Processes. Imperial College Press, Cambridge (2005)
20. Jacod, J., Shiryaev, A.N.: Limit theorems for stochastic processes. Grundlehren der Mathematischen Wissenschaften, vol. 288. Springer, Berlin (1987)
21. Kallsen, J., Tankov, P.: Characterization of dependence of multidimensional Lévy processes using Lévy copulas. J. Multivar. Anal. **97** (2006)
22. Lagerås, A.N.: Copulas for Markovian dependence. Bernoulli **16**(2), 331–342 (2010)
23. Lando, D., Mortensen, A.: On the pricing of step-up bonds in the European telecom sector. J. Credit Risk **1**(1), 71–110 (Winter 2004/05)
24. Last, G., Brandt, A.: Marked point processes on the real line. The dynamic approach. Probability and Its Applications. Springer, New York, NY (1995)
25. Li, D.X.: On default correlation: a copula function approach. J. Fixed Income **9**, 43–54 (2000)
26. Nelsen, R.B.: An introduction to copulas. Lecture Notes in Statistics, vol. 139. Springer, New York, NY (1999)
27. Protter, P.E.: Stochastic integration and differential equations. Application of Mathematics. Springer, Berlin, Heidelberg, New York (2004)
28. Salmon, F.: Recipe for disaster: the formula that killed Wall Street. Wired Mag **17**(3) (2009)
29. Tankov, P.: Simulation and option pricing in Lévy copula model. Working paper available from http://people.math.jussieu.fr/tankov/
30. Tankov, P.: Dependence structure of spectrally positive multidimensional Lévy processes (2003). Working paper available from http://people.math.jussieu.fr/tankov/
31. Vidozzi, A.: Two essays in mathematical finance. Doctoral Dissertation, Illinois Institute of Technology, Chicago, IL (2009)
32. Vidozzi, L.: Two essays on multivariate stochastic processes and applications to credit risk modeling. Doctoral Dissertation, Illinois Institute of Technology, Chicago, IL (2009)
33. Whitehouse, M.: How a formula ignited market that burned some big investors. Wall St. J. **12**(18) (2005), 1

http://www.springer.com/978-3-642-12464-8

Copula Theory and Its Applications
Proceedings of the Workshop Held in Warsaw, 25-26 September 2009
(Eds.)P. Jaworski; F. Durante; W.K. Härdle; T. Rychlik
2010, XVIII, 327 p. 25 illus., Softcover
ISBN: 978-3-642-12464-8