FINANCE WITHOUT BROWNIAN MOTIONS: AN INTRODUCTION TO SIMPLIFIED STOCHASTIC CALCULUS

ALEŠ ČERNÝ AND JOHANNES RUF

Abstract. The paper introduces a simple way of recording and manipulating stochastic processes without explicit reference to a probability measure. In the new calculus, operations traditionally presented in a measure-specific way are instead captured by tracing the behaviour of jumps (also when no jumps are physically present). The new calculus is thus intuitive and compact. The calculus is also fail-safe in that, under minimal assumptions, all formal calculations are guaranteed to yield mathematically well-defined stochastic processes. Several illustrative examples of the new concept are given, among them a novel result on the Margrabe option to exchange one defaultable asset for another.

1. Introduction

"Because in mathematics we pile inferences upon inferences, it is a good thing whenever we can subsume as many of them as possible under one symbol. For once we have understood the true significance of an operation, just the sensible apprehension of its symbol will suffice to obviate the whole reasoning process that earlier we had to engage anew each time the operation was encountered."

— Carl Jacobi (1804–1851)

In this paper we introduce and formalize a way of thinking about stochastic calculus that simplifies some common calculations and achieves more with less, without sacrificing rigour. The paper is aimed at the research community whose members do not consider themselves to be experts in mathematics in general, or

Date: December 10, 2019.

2010 *Mathematics Subject Classification.* (Primary) 60H05, 60H10; 60G44, 60G48; (Secondary) 91B02, 91B25; 91G10.

Key words and phrases. Émery formula, semimartingale representation, simplified calculus, multiplicative compensator, Girsanov's theorem.

We thank Jan Kallsen, Jan-Frederik Mai, Lola Martinez-Miranda, and Johannes Muhle-Karbe for helpful comments.

stochastic analysis in particular, but who nevertheless use stochastic calculus as a modelling tool.

In this paper we assume that the reader is acquainted with Brownian motion, Lévy processes, the notion of quadratic variation, some version of Itô's formula, no matter how heuristic, and Girsanov's theorem. The examples in this paper are inspired by applications in Economics and Financial Mathematics but the broader lessons are clearly applicable to Science at large.

The paper is organized as follows. In the rest of this introduction, we trace how the novel concept of this paper, the semimartingale representation (1.17), arises from classical Itô calculus. Section 2 provides a thorough introduction to simplified calculus. It also explains how the proposed approach facilitates computation of drifts and expected values; in particular, it tackles the introductory example in the presence of jumps. Section 3 demonstrates the strength of the proposed approach on three additional examples. Section 4 amplifies this point by showcasing calculations that also require a change of measure. In particular, Example 4.3 contains a new result that makes use of a non-equivalent change of measure. Where practical, we contrast the new approach with the more involved classical notation. Section 5 highlights the robustness of the proposed approach whereby, for a given task, the same representation applies in both discrete and continuous models. Such unification is unattainable in standard calculus. Section 6 concludes.

1.1. **McKean calculus for Itô processes.**

For reasons of tractability, there is a preponderance of continuous-time stochastic models based on Brownian motion. Traditional stochastic calculus reflects this historical bias. As an example, consider a stochastic model for two economic variables, capital K and labour L,

$$\frac{\mathrm{d}K_t}{K_t} = \mu_K \mathrm{d}t + \sigma_K \mathrm{d}W_t, \tag{1.1}$$

$$\frac{\mathrm{d}L_t}{L_t} = \mu_L \mathrm{d}t + \sigma_L \left(\rho_{KL}\mathrm{d}W_t + \sqrt{1-\rho_{KL}^2}\,\mathrm{d}\widehat{W}_t\right). \tag{1.2}$$

Here W and $\widehat{W}$ are two independent Brownian motions. The inputs in this model are μ_K, μ_L, $\sigma_K > 0$, $\sigma_L > 0$, and $\rho_{KL} \in [-1, 1]$ describing the correlation between the changes in K and L. Informally, the 'drift part' $\mu_K \mathrm{d}t$ represents expected change, while the 'noise' $\sigma_K \mathrm{d}W_t$ is loosely interpreted as a shock with mean zero and variance $\sigma_K^2 \mathrm{d}t$.

The symbol $\mathrm{d}K_t$ represents an increase in capital over an infinitesimal time period $\mathrm{d}t$. The left-hand side of (1.1) signifies percentage change in capital over the same period. The percentage change *per unit of time* is not well defined because the derivative ${}^{\mathrm{d}W_t}/_{\mathrm{d}t}$ does not exist. However, the *expected* percentage change per unit of time is finite and equal to μ_K. This means the expected proportional increase in capital over a fixed time horizon T equals

$$\mathsf{E}\left[\frac{K_T}{K_0}\right] = \mathrm{e}^{\mu_K T}. \tag{1.3}$$

In the terminology of Samuelson (1965), K and L are *geometric Brownian motions* with drift rates μ_K and μ_L, respectively.

Suppose we are interested in the evolution of the capital-labour ratio, ${}^K/_L$. The standard Itô calculus (Itô 1951, Theorem 6) yields, after simplifications,

$$\begin{aligned}\mathrm{d}\left(\frac{K_t}{L_t}\right) =& \frac{K_t}{L_t}(\mu_K - \mu_L - \rho_{KL}\sigma_K\sigma_L + \sigma_L^2)\mathrm{d}t \\ &+ \frac{K_t}{L_t}(\sigma_K - \rho_{KL}\sigma_L)\mathrm{d}W_t - \frac{K_t}{L_t}\sigma_L\sqrt{1-\rho_{KL}^2}\,\mathrm{d}\widehat{W}_t.\end{aligned} \tag{1.4}$$

One can make two observations at this point:

(1) The formula (1.4) is not easy to decipher — we are doing some serious calculus here.

(2) The formula is 'misleading' — the processes K and L are already given, therefore the object on the left-hand side is defined path by path as the ratio ${}^{K_T(\omega)}/_{L_T(\omega)}$ and cannot depend on the reference probability measure. In contrast, some of the objects on the right-hand side are strongly measure-dependent: certainly if we change the reference probability measure there is no guarantee that W and $\widehat{W}$ will still be Brownian motions under the new measure.

McKean (1969, p. 33) addresses (1) and formally also (2) by rewriting (1.4) in the form

$$\mathrm{d}\left(\frac{K_t}{L_t}\right) = \frac{K_t}{L_t}\left(\frac{\mathrm{d}K_t}{K_t} - \frac{\mathrm{d}L_t}{L_t} - \frac{\mathrm{d}K_t}{K_t}\frac{\mathrm{d}L_t}{L_t} + \left(\frac{\mathrm{d}L_t}{L_t}\right)^2\right), \tag{1.5}$$

where $\mathrm{d}K_t\mathrm{d}L_t$ is understood to stand for $\mathrm{d}[K,L]_t$ and $[K,L]$ is the quadratic covariation of the processes K and L. In the present case we have

$$\frac{\mathrm{d}[K,L]_t}{K_tL_t} = \rho_{KL}\sigma_K\sigma_L\mathrm{d}t \qquad \text{and} \qquad \frac{\mathrm{d}[L,L]_t}{L_t^2} = \sigma_L^2\mathrm{d}t.$$

Let us make two further observations:

(3) Formula (1.5) is not only measure-independent, it is also model-free in the sense that it holds for *any* two continuous semimartingales K and L such that the integrals of ${}^{\mathrm{d}K_t}/_{K_t}$ and ${}^{\mathrm{d}L_t}/_{L_t}$ are well defined.

(4) McKean (1969) observes that one can obtain (1.5) much more directly without passing through (1.1, 1.2) and (1.4), simply by writing down a second-order Taylor expansion for $f(K,L) = {}^{K}/_{L}$ in the form

$$\mathrm{d}f(K_t,L_t) = \tfrac{\partial f}{\partial k}(K_t,L_t)\mathrm{d}K_t + \tfrac{\partial f}{\partial \ell}(K_t,L_t)\mathrm{d}L_t$$
$$+\frac{1}{2}\left(\tfrac{\partial^2 f}{\partial k^2}(K_t,L_t)(\mathrm{d}K_t)^2 + 2\tfrac{\partial^2 f}{\partial k\partial \ell}(K_t,L_t)\mathrm{d}K_t\mathrm{d}L_t + \tfrac{\partial^2 f}{\partial \ell^2}(K_t,L_t)(\mathrm{d}L_t)^2\right).$$

Suppose now we wish to evaluate the *expected* value of the capital–labour ratio. Here formula (1.4) is very helpful because it tells us immediately that ${}^{K}/_{L}$ is a geometric Brownian motion with drift rate

$$b = \mu_K - \mu_L - \rho_{KL}\sigma_K\sigma_L + \sigma_L^2. \tag{1.6}$$

With this coefficient in hand one swiftly concludes, in analogy to (1.3),

$$\mathsf{E}\left[\frac{K_T}{L_T}\right] = \frac{K_0}{L_0}\mathrm{e}^{bT}. \tag{1.7}$$

The need for equation (1.4) is only illusory, however. One can obtain formula (1.6) equally easily from the measure-independent McKean formula (1.5) by inserting the expected rate of change of K, L and their quadratic (co)variations on the right-hand side of (1.5) as implied by (1.1, 1.2),

$$\frac{\mathrm{d}({}^{K_t}/_{L_t})}{{}^{K_t}/_{L_t}} = \frac{\mathrm{d}K_t}{K_t} - \frac{\mathrm{d}L_t}{L_t} - \frac{\mathrm{d}[K,L]_t}{K_tL_t} + \frac{\mathrm{d}[L,L]_t}{L_t^2}, \tag{1.8}$$
$$\downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow$$
$$b = \mu_K - \mu_L - \rho_{KL}\sigma_K\sigma_L + \sigma_L^2. \tag{1.9}$$

Note that apart from the initial values K_0, L_0 and the time horizon T the calculation requires *five* characteristics of the capital and labour processes: μ_K, μ_L, σ_K, σ_L, and ρ_{KL}.

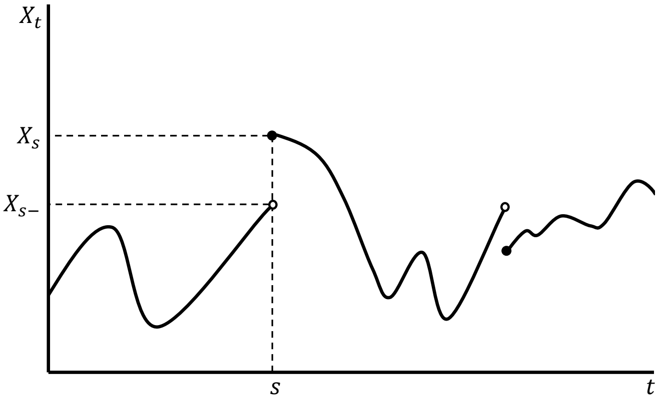

FIGURE 1.1. Illustration of a right-continuous path with left limits (process X).

1.2. **First steps.** Let us now modify (1.1, 1.2) by adding two jump components J^K and J^L that jointly form a Lévy process,

$$\frac{\mathrm{d}K_t}{K_{t-}} = \mu_K \mathrm{d}t + \sigma_K \mathrm{d}W_t + \mathrm{d}J_t^K; \tag{1.10}$$

$$\frac{\mathrm{d}L_t}{L_{t-}} = \mu_L \mathrm{d}t + \sigma_L \left(\rho_{KL} \mathrm{d}W_t + \sqrt{1-\rho_{KL}^2}\,\mathrm{d}\widehat{W}_t\right) + \mathrm{d}J_t^L; \tag{1.11}$$

$$\mathrm{d}J_t^K = \int_{|x|\leq 1} x_1 (N(\mathrm{d}t,\mathrm{d}x) - \Pi(\mathrm{d}x)\mathrm{d}t) + \int_{|x|>1} x_1 N(\mathrm{d}t,\mathrm{d}x); \tag{1.12}$$

$$\mathrm{d}J_t^L = \int_{|x|\leq 1} x_2 (N(\mathrm{d}t,\mathrm{d}x) - \Pi(\mathrm{d}x)\mathrm{d}t) + \int_{|x|>1} x_2 N(\mathrm{d}t,\mathrm{d}x). \tag{1.13}$$

Here N is a Poisson random measure and Π the corresponding Lévy measure. In particular, when (J^K, J^L) have finite activity, i.e., when $\Pi(\mathbb{R}^2) < \infty$, the quantity $\Pi(\mathbb{R}^2)$ is the arrival intensity of jumps and

$$x \mapsto \frac{\Pi((-\infty, x_1) \times (-\infty, x_2))}{\Pi(\mathbb{R}^2)}$$

is the bivariate cumulative distribution of jump sizes.

In the presence of jumps, by convention, the paths of K and L are assumed to be right-continuous with left limits. That means the value of capital *before* a jump is given by the left limit K_{t-} while K_t is the value *after* a jump. The jump is naturally defined as the difference between the two, $\Delta K_t = K_t - K_{t-}$ and likewise for L; see Figure 1.1.

Very little has changed between (1.1)–(1.2) and (1.10)–(1.11); we have replaced one process with time-homogeneous independent increments by another. But unlike in the Brownian case, it is now mathematically possible that the right-hand side of (1.11) — the percentage change in labour supply L — does not have a finite first moment, while the mean of ${}^{K}/_{L}$ is nevertheless finite.[1]

The main takeaway message is that decomposing a stochastic integral into 'signal' and 'noise' as suggested by (1.1, 1.2) is, in general, not straightforward. One possibility is to split J^L into two components containing the small and large jumps, respectively, and decompose only the small jump component into signal and noise as shown in (1.13). This makes (1.11) look more like (1.2) and largely represents the current practice in applications, see Kallsen (2000), Fujiwara and Miyahara (2003), Hubalek et al. (2006), Jeanblanc et al. (2007), Øksendal and Sulem (2007), Bender and Niethammer (2008), and Applebaum (2009).

We interpret the difficulty with signal–noise decomposition very differently, which is to say we refrain from using such a decomposition altogether and instead look for a measure-invariant representation of ${}^{K}/_{L}$. The sought expression must reduce to the McKean calculus (1.8) when K and L are continuous. At the same time, it must correctly account for changes due to jumps in K and L. Just such a formula, suitably reinterpreted, can be traced to Émery (1978, Section 3). Let us now describe what we mean, first informally and then rigorously.

In the present example we seek the representation of ${}^{K}/_{L}$. This will be written symbolically as

$$\mathrm{d}\left(\frac{K_t}{L_t}\right) = \frac{K_{t-} + \mathrm{d}K_t}{L_{t-} + \mathrm{d}L_t} - \frac{K_{t-}}{L_{t-}}, \tag{1.14}$$

which formally leads to an expression for percentage changes

$$\frac{\mathrm{d}\left({}^{K_t}/_{L_t}\right)}{{}^{K_{t-}}/_{L_{t-}}} = \frac{1 + {}^{\mathrm{d}K_t}/_{K_{t-}}}{1 + {}^{\mathrm{d}L_t}/_{L_{t-}}} - 1. \tag{1.15}$$

For Émery, the right-hand side of (1.15) represents a deterministic, time-constant function ξ, specifically

$$\xi(x) = \frac{1 + x_1}{1 + x_2} - 1, \tag{1.16}$$

[1]From a modelling point of view it is not realistic to believe that L has infinite mean. We are simply stating that the calculus must be general enough to entertain such possibility.

that acts on the increments ${}^{\mathrm{d}K_t}/_{K_{t-}}$ and ${}^{\mathrm{d}L_t}/_{L_{t-}}$. The real meaning of the expression $\xi(\mathrm{d}X_t)$ for a deterministic time-constant $\mathcal{C}^2$ function ξ and any semimartingale X is supplied by the Émery formula[2]

$$\xi(\mathrm{d}X_t) = D\xi(0)\mathrm{d}X_t + \frac{1}{2}\sum_{i,j=1}^{d} D^2_{ij}\xi(0)\,\mathrm{d}\big[X^{(i)}, X^{(j)}\big]^c_t + \left(\xi(\Delta X_t) - D\xi(0)\Delta X_t\right), \tag{1.17}$$

where the last term yields absolutely summable jumps of finite variation. Having assigned meaning to the right-hand side of (1.15) via the formula (1.17), equality (1.15) is no longer a formal expression but a theorem whose validity one needs to establish.

To accomplish this goal and build the simplified calculus, one can begin with the observation that (1.14), too, represents a function that acts on the increments of the underlying processes; in this case it acts on $\mathrm{d}K$ and $\mathrm{d}L$. One important difference is that the function in question is no longer deterministic, i.e., we formally have

$$\mathrm{d}\left(\frac{K_t}{L_t}\right) = \xi_t(\mathrm{d}K_t, \mathrm{d}L_t)$$

with

$$\xi_t(k,\ell) = \frac{K_{t-}+k}{L_{t-}+\ell} - \frac{K_{t-}}{L_{t-}}. \tag{1.18}$$

Another practical consideration is that the predictable function ξ in (1.18) is not finite-valued for $\ell = -L_-$.

To accommodate functions with restricted domain we define, for a given predictable function ξ, the set of semimartingales whose jumps are compatible with ξ, i.e.,

$$\mathsf{Dom}(\xi) = \{\text{semimartingale } X : \xi(\Delta X) \text{ is finite-valued}, \mathsf{P}\text{–a.s.}\}.$$

[2]The symbols $D\xi$ and $D^2\xi$ stand for first and second partial derivatives of ξ. In Émery (1978) the formula is given in a slightly different form

$$\xi(\mathrm{d}X_t) = D\xi(0)\mathrm{d}X_t + \frac{1}{2}\sum_{i,j=1}^{d} D^2_{ij}\xi(0)\,\mathrm{d}\big[X^{(i)}, X^{(j)}\big]_t + \Big(\xi(\Delta X_t) - D\xi(0)\Delta X_t - \frac{1}{2}\sum_{i,j=1}^{d} D^2_{ij}\xi(0)\,\Delta X^{(i)}_t \Delta X^{(j)}_t\Big).$$

The original formula considers X as a square-matrix-valued process, hence its indexation pattern is more involved than the one shown here.

It is formally straightforward to extend the formula (1.17) to a predictable function ξ by simply adding (or better just imagining) a subscript t in every occurence of ξ in (1.17). But having done this, it is *a priori* not clear that the extended formula will be well-defined for ξ such as (1.18), in particular it is not clear that one obtains

$$\sum_{0<t\leq\cdot} |\xi_t(\Delta X_t) - D\xi_t(0)\Delta X_t| < \infty. \tag{1.19}$$

Ideally, we would like to have a calculus that gives rise only to those predictable functions ξ for which the integral $\int_0^{\cdot} \xi_t(\mathrm{d}X_t)$ is always well-defined so we do not have to manually check admissibility every time a new ξ arises. The precise formulation of such a class, whose elements will be called *universally representing functions*, is given in Appendix A. We denote by $\mathfrak{I}^{d,n}_{0\mathbb{R}}$ the set of universally representing functions that map $\mathbb{R}^d$–valued processes to $\mathbb{R}^n$–dimensional processes. We also set $\mathfrak{I}_{0\mathbb{R}} = \bigcup_{d,n\in\mathbb{N}} \mathfrak{I}^{d,n}_{0\mathbb{R}}$. For computations involving characteristic functions, it is convenient to extend the Émery formula (1.17) to complex-differentiable functions by interpreting $D\xi$ and $D^2\xi$ as complex derivatives. The set of all universally representing complex-differentiable functions will be denoted by $\mathfrak{I}_{0\mathbb{C}}$.

Propositions A.4–A.6 in Appendix A show that the class $\mathfrak{I}_{0\mathbb{R}}$ is self-contained; if we use standard real-valued operations we are guaranteed to stay within $\mathfrak{I}_{0\mathbb{R}}$. The class $\mathfrak{I}_{0\mathbb{C}}$ is also self-contained provided the transformations we apply are complex-differentiable, as will be the case in all our examples involving complex numbers. These results guarantee that we will only ever encounter functions in $\mathfrak{I}_{0\mathbb{R}}$ (resp., $\mathfrak{I}_{0\mathbb{C}}$) and so will never have to check the conditions of Definition A.1 manually.

1.3. **Integral notation.** Just as the McKean calculus of Subsection 1.1, the simplified calculus is most intuitive when expressed in differential form, such as (1.14) and (1.15). When one wishes to speak of the integrated process whose increments are equal to $\xi_t(\mathrm{d}X_t)$, one typically just introduces a new label. For example, if we decided to call the new process Y, then in the differential form we would write

$$\mathrm{d}Y_t = \xi_t(\mathrm{d}X_t).$$

There is nothing wrong with the relabelling approach; it does deliver all the immediate benefits of the simplified calculus and helps to keep technicalities to a minimum.

Side-by-side with the intuitive differential approach, we want to offer the reader an alternative 'high-level' view of the calculus where the roles of ξ and X are acknowledged explicitly. Accordingly, the process with increments $\xi_t(\mathrm{d}X_t)$, starting at 0, will be denoted by $\xi \circ X$. The high-level notation may seem a little abstract at first, but it offers distinct benefits such as compactness and flexibility. For example, in the integral notation one can write

$$\begin{aligned}
[X,X] &= \int_0^{\cdot} (\mathrm{d}X_t)^2 = x^2 \circ X;\\
[[X,X],X] &= \int_0^{\cdot} (\mathrm{d}X_t)^3 = x^3 \circ X;\\
[[[X,X],X],X] = [[X,X],[X,X]] &= \int_0^{\cdot} (\mathrm{d}X_t)^4 = x^4 \circ X.
\end{aligned}$$

The notation $\xi \circ X$ also emphasizes the universality of the transformation $X \mapsto \xi \circ X$. In the same way that $[X, X]$ is well defined for *any* semimartingale X, the process $x^\alpha \circ X$ is well-defined for any semimartingale X and any $\alpha \geq 2$. This brings us to other universal transformations that are commonly used in the literature. For example, provided that X and X_- are different from zero, the literature defines $\mathcal{L}(X)$ as the process of cumulative percentage change in X, i.e.,

$$\mathrm{d}\mathcal{L}(X)_t = \frac{\mathrm{d}X_t}{X_{t-}}.$$

Thus, in the integral notation formula (1.15) reads

$$\mathcal{L}\left(\frac{K}{L}\right) = \left(\frac{1+x_1}{1+x_2} - 1\right) \circ (\mathcal{L}(K), \mathcal{L}(L)). \tag{1.20}$$

Provided that the cumulative percentage changes in K and L are well-defined, formula (1.20) holds for arbitrary semimartingales K and L.

2. Simplified calculus

2.1. **Composite rules.** The simplified calculus rests on a sequential application of Propositions A.4–A.6. In practice, one would not use these propositions directly but instead combine them into composite rules that transform a represented process $Y = Y_0 + \xi \circ X$ into another represented process Z. In differential notation the first two rules read as follows. We purposely do not spell out all the technicalities to highlight how natural the calculus is. Recall $\mathrm{d}Y_t = \xi_t(\mathrm{d}X_t)$.

(1) Stochastic integration with respect to Y. For a locally bounded process ζ the integral $\mathrm{d}Z_t = \zeta_t \mathrm{d}Y_t$ satisfies

$$\mathrm{d}Z_t = \zeta_t \xi_t(\mathrm{d}X_t).$$

(2) Itô formula for $Z = f(Y)$. For a suitably smooth function f such that Y and Y_- lie in the interior of the domain of f one has

$$\mathrm{d}Z_t = \mathrm{d}f(Y_t) = f(Y_{t-} + \mathrm{d}Y_t) - f(Y_{t-}) = f(Y_{t-} + \xi_t(\mathrm{d}X_t)) - f(Y_{t-}).$$

We now re-state the above fully rigorously in integral notation.

Corollary 2.1. *Let $\xi \in \mathfrak{I}^{d,n}_{0\mathbb{R}}$ (resp., $\mathfrak{I}^{d,n}_{0\mathbb{C}}$) with $X \in \mathsf{Dom}(\xi)$ and consider the n–dimensional process*

$$Y = Y_0 + \xi \circ X.$$

The following rules then apply.

- Stochastic integration: *For a locally bounded $\mathbb{R}^{m\times n}$–valued (resp., $\mathbb{C}^{m\times n}$–valued) predictable process ζ we have $\zeta\xi \in \mathfrak{I}^{d,m}_{0\mathbb{R}}$ (resp., $\mathfrak{I}^{d,m}_{0\mathbb{C}}$) and*

$$Z = Z_0 + \int_0^\cdot \zeta_u \mathrm{d}Y_u = Z_0 + \zeta\xi \circ X; \tag{2.1}$$

- Smooth transformation ('Itô's formula'): *Provided Y and Y_- remain in an open subset $\mathcal{U}$ of $\mathbb{R}^n$ (resp., $\mathbb{C}^n$) where the function $f : \mathcal{U} \to \mathbb{R}^m$ is twice continuously differentiable (resp. where $f : \mathcal{U} \to \mathbb{C}^m$ is analytic), we have $f(Y_- + \xi) - f(Y_-) \in \mathfrak{I}^{d,m}_{0\mathbb{R}}$ (resp., $\mathfrak{I}^{d,m}_{0\mathbb{C}}$) and*

$$Z = f(Y) = f(Y_0) + (f(Y_- + \xi) - f(Y_-)) \circ X. \tag{2.2}$$

Let us now outline a general scheme of how the two composite rules are applied in practice. At the outset, one will designate the primitive input to the problem at hand; this input process is thereafter labeled X. For example, in Subsections 1.1 and 1.2 it is natural to start from the bivariate process

$$X = (\mathcal{L}(K), \mathcal{L}(L)). \tag{2.3}$$

#	$\mathrm{d}Y$	operation $Y \to Z$	$\mathrm{d}Z$
1	$\mathrm{d}\begin{bmatrix}\mathcal{L}(K)\\ \mathcal{L}(L)\end{bmatrix}$	integration $\zeta = \begin{bmatrix}K_- & 0\\ 0 & L_-\end{bmatrix}$	$\mathrm{d}\begin{bmatrix}K\\ L\end{bmatrix} = \begin{bmatrix}K_- & 0\\ 0 & L_-\end{bmatrix}\mathrm{d}\begin{bmatrix}\mathcal{L}(K)\\ \mathcal{L}(L)\end{bmatrix}$
2	$\mathrm{d}\begin{bmatrix}K\\ L\end{bmatrix}$	smooth transformation $f(K,L) = \frac{K}{L}$	$\mathrm{d}\left(\frac{K}{L}\right) = \left(\frac{K_-+\mathrm{d}K}{L_-+\mathrm{d}L} - \frac{K_-}{L_-}\right) = \frac{K_-}{L_-}\left(\frac{1+\mathrm{d}\mathcal{L}(K)}{1+\mathrm{d}\mathcal{L}(L)} - 1\right)$
3	$\mathrm{d}\left(\frac{K}{L}\right)$	integration $\zeta = \frac{L_-}{K_-}$	$\mathrm{d}\mathcal{L}\left(\frac{K}{L}\right) = \frac{L_-}{K_-}\mathrm{d}\left(\frac{K}{L}\right) = \left(\frac{1+\mathrm{d}\mathcal{L}(K)}{1+\mathrm{d}\mathcal{L}(L)} - 1\right)$

TABLE 2.1. Schematic derivation of (1.20) by means of Corollary 2.1

Observe that X is always representable with respect to itself thanks to Proposition A.4 with ζ equal to the $d \times d$ identity matrix. Starting off from the trivial representation $X = X_0 + x \circ X$, i.e., taking $Y = X$ and $\xi(x) = x$ in Corollary 2.1, one applies smooth transformation or stochastic integration as required to obtain the first intermediate result Z. This intermediate result (relabeled Y) becomes the input to the next application of Corollary 2.1 producing the next intermediate output Z. The $Y \to Z$ pattern is repeated until one reaches the desired output Z; in our example the goal is

$$Z = \mathcal{L}\left(\frac{K}{L}\right) = \mathcal{L}\left(\frac{X^{(1)}}{X^{(2)}}\right). \tag{2.4}$$

Table 2.1 illustrates the steps required in the transition from (2.3) to (2.4).

The discussion above concerns formal calculations where the rules of the simplified calculus are applied mechanically. Many users will prefer a more intuitive approach whose main idea is apparent in the last column of Table 2.1. Here one observes that the calculus traces the behaviour of jumps and so one may effectively pretend that X, Y, and Z are finite-variation pure-jump processes. In this way, it is possible to arrive at the correct ξ even without applying formal rules. In the context of (1.20), for example, suppose K increases by 50% and L increases by 20%. The percentage change in ${}^{K}\!/_{L}$ is then precisely

$$\frac{1+0.5}{1+0.2} - 1 = 25\%.$$

Therefore, formula (1.20), among other things, describes jump transformations: every time $\mathcal{L}(K)$ jumps by x_1 (e.g. 0.5) and $\mathcal{L}(L)$ jumps by x_2 (e.g. 0.2) the process $\mathcal{L}(K/L)$ jumps by

$$\xi(x) = \frac{1+x_1}{1+x_2} - 1.$$

As a further example, let us see how the rules of the simplified calculus can be used to obtain the representation of the logarithmic return in terms of the rate of return.

Example 2.2 (Representation of the log return in terms of the rate of return)**.** Let $S > 0$ represent the value of an investment with $S_- > 0$ and let $X = \mathcal{L}(S) = \int_0^\cdot \mathrm{d}S_t/S_{t-}$ be the cumulative rate of return on this investment. On a purely intuitive level, thinking only of jump transformations, one can write

$$\mathrm{d}\ln S_t = \ln(S_{t-} + \mathrm{d}S_t) - \ln S_{t-} = \ln\left(1 + \frac{\mathrm{d}S_t}{S_{t-}}\right) = \ln\left(1 + \mathrm{d}\mathcal{L}(S)_t\right).$$

More formally, the integration rule yields $S = S_0 + S_- x \circ \mathcal{L}(S)$ and smooth transformation then gives

$$\ln S = \ln(S_0 + S_- x \circ \mathcal{L}(S)) = \ln S_0 + (\ln(S_- + S_- x) - \ln S_-) \circ \mathcal{L}(S).$$

Both approaches yield the representation

$$\ln S = \ln S_0 + \ln(1+x) \circ \mathcal{L}(S)$$

for any semimartingale S such that $S_- > 0$ and $S > 0$. □

As the final introductory example consider the representation of quadratic covariation.

Example 2.3 (Representation of quadratic covariation)**.** The quadratic covariation $[X, Y]$ satisfies (or, as in Meyer 1976, is defined by) the identity

$$XY = X_0 Y_0 + \int_0^\cdot X_{t-}\mathrm{d}Y_t + \int_0^\cdot Y_{t-}\mathrm{d}X_t + [X, Y].$$

This yields

$$\begin{aligned}
\mathrm{d}[X,Y]_t &= \mathrm{d}(X_t Y_t) - X_{t-}\mathrm{d}Y_t - Y_{t-}\mathrm{d}X_t \\
&= (X_{t-} + \mathrm{d}X_t)(Y_{t-} + \mathrm{d}Y_t) - X_{t-}Y_{t-} - X_{t-}\mathrm{d}Y_t - Y_{t-}\mathrm{d}X_t = \mathrm{d}X_t \mathrm{d}Y_t.
\end{aligned}$$

More formally, the integration and smooth transformation rules yield

$$[X,Y] = ((X_- + x)(Y_- + y) - X_-Y_- - X_-y - Y_-x) \circ (X,Y) = xy \circ (X,Y).$$

Thus, in the differential notation one can rigorously write $\mathrm{d}[X,X]_t = (\mathrm{d}X_t)^2$ for any univariate semimartingale X. □

Section 3 contains many more explicit representations that are useful in practice. Some of these are well known in the specialist literature, while others are new. Proposition A.6, in particular, is a powerful tool for obtaining new representations from old ones. We summarize it here in the form of a composition rule. Observe that in the differential notation the rule is completely natural; it asserts that $\mathrm{d}Z_t = \psi_t(\mathrm{d}Y_t)$ and $\mathrm{d}Y_t = \xi_t(\mathrm{d}X_t)$ imply $\mathrm{d}Z_t = \psi_t(\xi_t(\mathrm{d}X_t))$.

Corollary 2.4 (Composition of representations)**.** *Let $\xi \in \mathfrak{I}^{d,n}_{0\mathbb{R}}$ with $X \in \mathsf{Dom}(\xi)$ and consider the n–dimensional process $Y = Y_0 + \xi \circ X$. For $\psi \in \mathfrak{I}^{n,m}_{0\mathbb{R}}$ with $Y \in \mathsf{Dom}(\psi)$ one obtains $\psi(\xi) \in \mathfrak{I}^{d,m}_{0\mathbb{R}}$, $Y \in \mathsf{Dom}(\psi(\xi))$, and*

$$Z = Z_0 + \psi \circ Y = Z_0 + \psi(\xi) \circ X. \tag{2.5}$$

An analogous statement holds with $\mathfrak{I}_{0\mathbb{C}}$ in place of $\mathfrak{I}_{0\mathbb{R}}$.

The composition rule allows the user to store some common calculations and 'recycle' them later without having to revisit their detailed derivation. Suppose, for instance, that we are given the evolution of $(\ln K, \ln S)$ as the primitive input. Thanks to Corollary 2.4, there is no need to calculate everything afresh all the way from $(\ln K, \ln S)$ to $\mathcal{L}({}^K\!/\!{}_L)$. One only computes the passage from $(\ln K, \ln S)$ to $(\mathcal{L}(K), \mathcal{L}(S))$ which yields (see equation (3.4) below)

$$(\mathcal{L}(K), \mathcal{L}(S)) = (\mathrm{e}^{x_1} - 1, \mathrm{e}^{x_2} - 1) \circ (\ln K, \ln S),$$

while the passage from $(\mathcal{L}(K), \mathcal{L}(S))$ to $\mathcal{L}({}^K\!/\!{}_L)$ can be recycled from (1.20). The two results composed together give

$$\mathcal{L}\left(\frac{K}{L}\right) = \left(\mathrm{e}^{x_1 - x_2} - 1\right) \circ (\ln K, \ln L). \tag{2.6}$$

In differential notation,

$$\mathrm{d}K_t = \mathrm{d}\mathrm{e}^{\ln K_t} = \mathrm{e}^{\ln K_{t-} + \mathrm{d}\ln K_t} - \mathrm{e}^{\ln K_{t-}} = K_{t-}(\mathrm{e}^{\mathrm{d}\ln K_t} - 1)$$

substituted into (1.15) yields

$$\frac{\mathrm{d}\left({}^{K_t}/_{L_t}\right)}{{}^{K_{t-}}/_{L_{t-}}} = \frac{1+\mathrm{e}^{\mathrm{d}\ln K_t}-1}{1+\mathrm{e}^{\mathrm{d}\ln L_t}-1} - 1 = \mathrm{e}^{\mathrm{d}\ln K_t - \mathrm{d}\ln L_t} - 1,$$

which is the differential equivalent of formula (2.6).

2.2. **Émery formula and drift computation.** Having mastered the art of representing one process by means of another, we would like to obtain an analogon of (1.8)–(1.9), i.e., we want to be able to convert the characteristics of the representing process into the drift of the represented process.

To begin with, we collect the predictable characteristics of the input process $X = (\mathcal{L}(K), \mathcal{L}(L))$ in equations (1.10)–(1.13) in the more compact form

$$b^{X[1]} = \begin{bmatrix} \mu_K \\ \mu_L \end{bmatrix}; \qquad c^X = \begin{bmatrix} \sigma_K^2 & \rho_{KL}\sigma_K\sigma_L \\ \rho_{KL}\sigma_K\sigma_L & \sigma_L^2 \end{bmatrix}; \qquad F^X = \Pi. \tag{2.7}$$

Here $X[1]$ denotes the process X with jumps greater than 1 in absolute value removed,

$$X[1] = X_0 + x\mathbf{1}_{|x|\leq 1} \circ X. \tag{2.8}$$

Observe that this precisely matches the decomposition of jumps appearing in (1.12)–(1.13) and ensures that the drift of $X[1]$ is finite.[3] More generally, we will denote by $X[h]$ the process containing the small jumps of X as given by a specific truncation function h and observe that $X[1]$ corresponds to the choice

$$h(x) = x\mathbf{1}_{|x|\leq 1}. \tag{2.9}$$

The mechanics of truncation are described in Definition B.1.

The reader must be warned that $X[0]$, the continuous part of X, is not always well defined. For this reason, $X[0]$ cannot be universally represented, in contrast to $X[1]$ whose universal representation appears in (2.8). Nonetheless, the situations where $X[0]$ exists arise in practice, for example, in the Merton (1976) jump-diffusion model. In such models we may write

$$\mathrm{d}X_t = \mathrm{d}X[0]_t + \Delta X_t. \tag{2.10}$$

[3]In contrast, the drift of X need not be well-defined in general; see also footnote 1.

This, when substituted into (1.17), leads to a simplified expression

$$\xi_t(\mathrm{d}X_t) = D\xi(0)\mathrm{d}X[0] + \frac{1}{2}\sum_{i,j=1}^{d} D^2_{ij}\xi(0)\,\mathrm{d}\Big[X^{(i)},X^{(j)}\Big]^c_t + \xi_t(\Delta X_t) \tag{2.11}$$

that offers a valuable insight into the nature of the Émery formula. We observe that the first two terms of (2.11) correspond to the McKean calculus for the continuous part of X while the last term accounts for jumps in X. The two components do not interact and can be treated separately. In the most general situation where the decomposition (2.10) does *not* exist, one can make the result rigorous by adding small jumps to the first term and subtracting them in the last term to obtain

$$\xi(\mathrm{d}X_t) = D\xi(0)\mathrm{d}X[h]_t + \frac{1}{2}\sum_{i,j=1}^{d} D^2_{ij}\xi(0)\,\mathrm{d}\Big[X^{(i)},X^{(j)}\Big]^c_t + (\xi(\Delta X_t) - D\xi(0)h(\Delta X_t))\,. \tag{2.12}$$

The original Émery formula (1.17) corresponds simply to the case where *all* the jumps have been added to the first term and subtracted in the last term of (2.11).

We thus come to understand the Émery formula as a *spectrum* of equivalent expressions where one can dial the truncation function h all the way down to 0 or all the way up to $h(x) = x$. In this sense, the truncation is unimportant – we can always choose h to suit our needs. Thus one would pick $h(x) = 0$ if jumps of X have finite variation as in the Merton jump-diffusion model, failing that, $h(x) = x$ if the drift of X exists, and finally $h(x) = \mathbf{1}_{|x|\leq 1}$ in all remaining cases. With such choice of h, the drift of each contributing term in (2.12) is guaranteed to be finite.

We can now perform the feat previously achieved on a smaller scale in (1.8)–(1.9). By matching each term in (2.12) with its drift contribution, one obtains

$$b^{\xi\circ X} = D\xi(0)b^{X[h]} + \frac{1}{2}\sum_{i,j=1}^{d} D^2_{i,j}\xi(0)c^X_{ij} + \int(\xi(x) - D\xi(0)h(x))F^X(\mathrm{d}x). \tag{2.13}$$

Formula (2.13) is proved in Theorem B.6.

Specifically, with ξ given in (1.16) we obtain

$$D\xi(0) = [1 \quad -1], \qquad D^2\xi(0) = \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}.$$

For the specific input parameters in (2.7) and the corresponding truncation function in (2.9) the drift conversion formula (2.13) yields

$$b^{\mathcal{L}(K/L)} = \mu_K - \mu_L - \rho_{KL}\sigma_K\sigma_L + \sigma_L^2 + \int \left(\frac{1+x_1}{1+x_2} - 1 - (x_1 - x_2)\,\mathbf{1}_{|x|\leq 1}\right)\Pi(\mathrm{d}x), \tag{2.14}$$

which is the appropriate generalization of (1.9) provided the integral (2.14) converges.[4] Formula (1.7) continues to hold with this choice of b; see Theorems B.4 and B.6 below.

3. Further examples with drift computation

Many tasks where stochastic processes are concerned involve computation of the drift of some quantity. Hamilton–Jacobi–Bellman equations in optimal control, for example, express the fact that the optimal value function plus the integrated historical cost is a martingale and therefore has zero drift. Similarly, Feynman–Kac formulae reflect zero drift of an integral of costs discounted at a specified stochastic killing rate.

We will now showcase the strength of process representations such as (1.20) when it comes to computing drifts. We will do so side-by-side with the classical approach. Let us therefore start with an $\mathbb{R}$–valued Lévy process written in the classical notation,

$$X = X_0 + \int_0^\cdot \alpha \mathrm{d}s + \int_0^\cdot \sigma \mathrm{d}W_s + \int_0^\cdot \int_{|x|\leq 1} x\widehat{N}(\mathrm{d}s,\mathrm{d}x) + \int_0^\cdot \int_{|x|>1} xN(\mathrm{d}s,\mathrm{d}x), \tag{3.1}$$

where N is a Poisson jump measure, Π the corresponding Lévy measure,

$$\widehat{N}(\mathrm{d}t,\mathrm{d}x) = N(\mathrm{d}t,\mathrm{d}x) - \Pi(\mathrm{d}x)\mathrm{d}t$$

the compensated Poisson jump measure, and $\alpha, \sigma \in \mathbb{R}$.

[4] We might consider, for example, a model where the jumps in $\mathcal{L}(K)$ and $\mathcal{L}(L)$ are independent, in which case

$$\Pi(\mathrm{d}x_1,\mathrm{d}x_2) = \mathbf{1}_{x_2=0}\Pi^K(\mathrm{d}x_1) + \mathbf{1}_{x_1=0}\Pi^L(\mathrm{d}x_2),$$

meaning capital and labour do not jump simultaneously. We may take $\Pi^K(\mathrm{d}x_1)$ to be lognormal so that $\int_0^\infty x_1\Pi^K(\mathrm{d}x_1) < \infty$ and K has finite mean. The choice $\Pi^L(\mathrm{d}x_2) = x_2^{-2}\mathbf{1}_{x_2>0}\mathrm{d}x_2$ then provides an example where the mean of L is infinite while the mean of K/L remains finite.

In the simplified calculus we will never have to write out a full decomposition like (3.1). Instead we just note that X is an Itô semimartingale[5] with characteristics

$$\left(b^{X[1]} = \alpha, c^X = \sigma^2, F^X = \Pi\right). \tag{3.2}$$

The notation of (3.2) emphasizes the fact that some expressions below, such as (3.6), remain valid even if b^X, c^X, and F^X are stochastic.

In the next example we will find the representation for the cumulative percentage change in e^{vX} for fixed $v \in \mathbb{C}$ and use this to compute the moment generating function of X.

Example 3.1 (Drift of $\mathcal{L}(\mathrm{e}^{vX})$ for $v \in \mathbb{C}$)**.** By smooth transformation we have

$$\mathrm{e}^{vX} = \mathrm{e}^{vX_0} + \left(\mathrm{e}^{v(X_- + x)} - \mathrm{e}^{vX_-}\right) \circ X. \tag{3.3}$$

Consequently, by stochastic integration and the chain rule,

$$\mathcal{L}(\mathrm{e}^{vX}) = \int_0^{\cdot} \mathrm{e}^{-vX_{t-}} \mathrm{d}\mathrm{e}^{vX_t} = \mathrm{e}^{-vX_-}\left(\mathrm{e}^{v(X_- + x)} - \mathrm{e}^{vX_-}\right) \circ X = (\mathrm{e}^{vx} - 1) \circ X. \tag{3.4}$$

The representing function is $\xi(x) = \mathrm{e}^{vx} - 1$ with $\xi'(0) = v$ and $\xi''(0) = v^2$. The corresponding Émery formula (2.12) reads

$$\xi(\mathrm{d}X_t) = v\mathrm{d}X[h]_t \;+ \frac{1}{2}v^2 \mathrm{d}[X,X]^c + \left(\mathrm{e}^{v\Delta X_t} - 1 - vh(\Delta X_t)\right). \tag{3.5}$$

It is valid for *any* semimartingale X.

It is now straightforward to compute the drift in (3.5), provided it exists. Specifically, for an Itô semimartingale X, (3.5) yields a drift rate of

$$b^{\mathcal{L}(\mathrm{e}^{vX})} \;=\; vb^{X[h]} \quad + \frac{1}{2}v^2 c^X \qquad + \int_{\mathbb{R}} \left(\mathrm{e}^{vx} - 1 - vh(x)\right) F^X(\mathrm{d}x). \tag{3.6}$$

If, additionally, X is a Lévy process as in (3.1–3.2) we obtain

$$b^{\mathcal{L}(\mathrm{e}^{vX})} = \alpha v + \frac{1}{2}\sigma^2 v^2 + \int_{\mathbb{R}} \left(\mathrm{e}^{vx} - 1 - vx\mathbf{1}_{|x|\leq 1}\right) \Pi(\mathrm{d}x),$$

as long as the integral is finite and, in analogy to (1.3),

$$\mathsf{E}\left[\mathrm{e}^{v(X_T - X_0)}\right] = \exp\left(b^{\mathcal{L}(\mathrm{e}^{vX})} T\right). \tag{3.7}$$

[5]An Itô semimartingale is a generalization of (3.1) where α, σ^2, and Π are allowed to be stochastic; see Definition B.5 below.

In such case the function $\kappa^X(v) = b^{\mathcal{L}(e^{vX})}$ is known as the cumulant function of $X_1 - X_0$.[6] □

Remark 3.2. Let us now consider the same calculation using the form (3.1). Itô's formula (Applebaum 2009, Theorem 4.4.7) gives

$$\begin{aligned} e^{vX} - e^{vX_0} &= \int_0^\cdot v e^{vX_{s-}}(\alpha \mathrm{d}s + \sigma \mathrm{d}W_s) + \frac{1}{2}\int_0^t v^2 e^{vX_{s-}}\sigma^2 \mathrm{d}s \\ &\quad + \int_0^\cdot \int_{|x|\leq 1} \left(e^{v(X_{s-}+x)} - e^{vX_{s-}}\right) \widehat{N}(\mathrm{d}s, \mathrm{d}x) \\ &\quad + \int_0^\cdot \int_{|x|>1} \left(e^{v(X_{s-}+x)} - e^{vX_{s-}}\right) N(\mathrm{d}s, \mathrm{d}x) \\ &\quad + \int_0^\cdot \int_{|x|\leq 1} \left(e^{v(X_{s-}+x)} - e^{vX_{s-}} - v e^{vX_{s-}} x\right) \Pi(\mathrm{d}x)\mathrm{d}s. \end{aligned} \tag{3.8}$$

Integration yields (Applebaum 2009, Section 4.3.3)

$$\begin{aligned} \mathcal{L}(e^{vX}) &= \int_0^\cdot e^{-vX_{s-}} \mathrm{d}e^{vX_s} \\ &= \int_0^\cdot \left(\alpha v + \frac{1}{2}\sigma^2 v^2\right) \mathrm{d}s + \int_0^\cdot \sigma v \mathrm{d}W_s + \int_0^\cdot \int_{|x|\leq 1} (e^{vx} - 1)\widehat{N}(\mathrm{d}s, \mathrm{d}x) \\ &\quad + \int_0^\cdot \int_{|x|>1} (e^{vx} - 1) N(\mathrm{d}s, \mathrm{d}x) + \int_0^\cdot \int_{|x|\leq 1} (e^{vx} - 1 - vx)\Pi(\mathrm{d}x)\mathrm{d}s. \end{aligned} \tag{3.9}$$

Finally, the drift rate is evaluated by computing the drift of each contributing term in (3.9)

$$\begin{aligned} b^{\mathcal{L}(e^{vX})} &= \alpha v + \frac{1}{2}\sigma^2 v^2 + 0 + 0 \\ &\quad + \int_{|x|>1} (e^{vx} - 1)\Pi(\mathrm{d}x) + \int_{|x|\leq 1} (e^{vx} - 1 - vx)\,\Pi(\mathrm{d}x)\mathrm{d}s. \end{aligned} \tag{3.10}$$

Note the calculations (3.8–3.9) become much easier in the approach (3.3–3.5) because the rules of simplified calculus are more compact and easier to remember. □

The main advantage of the simplified calculus is that one does not have to keep track of the drift, volatility, and jump intensities through intermediate calculations. In the next example we will evaluate all three characteristics of the process $Y = \mathcal{L}(e^{vX})$ when X is an Itô semimartingale. Having all three characteristics is not necessary for our purposes; this example merely shows that the characteristics are easily recalled at any moment — if needed.

[6]Formula (3.7) holds thanks to Theorems B.4 and B.6. Note that (3.7) is in fact the Lévy-Khintchin formula (Sato 1999, Theorem 8.1).

Example 3.3 (Characteristics of a represented Itô semimartingale)**.** Consider $\xi \in \mathfrak{I}^{d,n}_{0\mathbb{R}} \cup \mathfrak{I}^{d,n}_{0\mathbb{C}}$ and an Itô semimartingale $X \in \mathsf{Dom}(\xi)$.

(1) For the 'volatility' of the represented process $Y = Y_0 + \xi \circ X$ we obtain

$$c^Y = D\xi(0)c^X D\xi(0)^\top.$$

(2) To compute the drift of $Y[g]$, observe that one naturally obtains

$$Y[g] = Y_0 + g \circ Y,$$

for any truncation function g equal to identity on a neighbourhood of zero (see Proposition B.2). The chain rule (2.5) now yields $Y[g] = Y_0 + g(\xi) \circ X$. As $Dg(0)$ is by assumption an identity matrix, we have

$$D(g \circ \xi)(0) = D\xi(0); \qquad D^2(g \circ \xi) = D^2\xi(0),$$

and the desired drift is

$$b^{Y[g]} = D\xi(0)b^{X[h]} + \frac{1}{2}\sum_{i,j=1}^{d} D^2_{ij}\xi(0)c^X_{ij} + \int_{\mathbb{R}} \left(g(\xi(x)) - D\xi(0)h(x)\right) F^X(\mathrm{d}x).$$

(3) Finally, let G be a closed n–dimensional set not containing zero. The process $\mathbf{1}_{y\in G} \circ Y$ counts the jumps of Y whose size is in G; its drift yields the jump arrival intensity $F^Y(G)$. The chain rule (2.5) gives

$$\mathbf{1}_{y\in G} \circ Y = \mathbf{1}_{\xi(x)\in G} \circ X.$$

The function $\psi = \mathbf{1}_{\xi(x)\in G}$ satisfies $D\psi(0) = D^2\psi(0) = 0$ which implies

$$F^Y(G) = b^{\psi\circ X} = 0 + 0 + \int_{\mathbb{R}} \mathbf{1}_{\xi(x)\in G} F^X(\mathrm{d}x),$$

whereby, for each (ω, t), we recognize F^Y as the image (a.k.a. push-forward) measure of F^X obtained via the mapping $x \mapsto \xi(x)$.

For concreteness, set $Y = \mathcal{L}(\mathrm{e}^{vX}) = (\mathrm{e}^{vx} - 1) \circ X$ for some fixed $v \in \mathbb{C}$ and take X to be the Lévy process defined by (3.1). As $D\xi(0) = v$ and $D^2\xi(0) = v^2$, we obtain

$$\begin{aligned}
b^{Y[1]} &= \alpha v + \frac{1}{2}\sigma^2 v^2 + \int_R \left((\mathrm{e}^{vx} - 1)\mathbf{1}_{|\mathrm{e}^{vx}-1|\leq 1} - vx\mathbf{1}_{|x|\leq 1}\right)\Pi(\mathrm{d}x);\\
c^Y &= \sigma^2 v^2;\\
F^Y(G) &= \int_{\mathbb{R}} \mathbf{1}_G(\mathrm{e}^{vx} - 1)\Pi(\mathrm{d}x).
\end{aligned}$$

We thus conclude that if X is a Lévy process then $Y = \mathcal{L}(\mathrm{e}^{vX})$ is again a Lévy process for all $v \in \mathbb{C}$. □

The next example illustrates the convenience of composing two representations without having to work with their predictable characteristics.

Example 3.4 (Maximization of exponential utility)**.** Fix a time horizon $T > 0$, and assume that X is a one-dimensional Lévy process given by (3.1–3.2). Consider an economy consisting of one bond with constant price one and of one risky asset with price process $S = \mathrm{e}^X$. Moreover, consider an agent with exponential utility function $u : w \mapsto -\mathrm{e}^{-w}$.

Since X is assumed to have stationary and independent increments and since we consider an exponential utility function, it is reasonable to conjecture that the optimal portfolio is a constant dollar amount $\lambda \in \mathbb{R}$ invested in the risky asset. Denote by $R = \mathcal{L}(\mathrm{e}^X)$ the cumulative yield on an 1\$ investment in the risky asset. Normalizing initial wealth to zero, the optimal wealth process equals λR and its expected utility is $\mathsf{E}[\mathrm{e}^{-\lambda R_T}]$. In analogy to (1.3) the expected utility can be obtained via the time rate of the expected percentage change of the quantity $\mathrm{e}^{-\lambda R}$. This is nothing other than the drift rate of the process $\mathcal{L}(\mathrm{e}^{-\lambda R})$. Provided this drift, commonly denoted by $\kappa^R(-\lambda)$, is finite, the expected utility will be equal to $\mathsf{E}[\mathrm{e}^{-\lambda R_T}] = \mathrm{e}^{\kappa^R(-\lambda)T}$, cf. (3.7).

Representation (3.4) and composition rule (2.5) give $R = \mathcal{L}(\mathrm{e}^X) = (\mathrm{e}^x - 1) \circ X$ and

$$\mathcal{L}\left(\mathrm{e}^{-\lambda R}\right) = \left(\mathrm{e}^{-\lambda y} - 1\right) \circ R = \left(\mathrm{e}^{-\lambda(\mathrm{e}^x-1)} - 1\right) \circ X. \tag{3.11}$$

For $\xi(x) = \mathrm{e}^{-\lambda(\mathrm{e}^x-1)} - 1$ one has $\xi'(0) = -\lambda$ and $\xi''(0) = \lambda^2 - \lambda$. Hence the desired drift reads

$$\kappa^R(-\lambda) = b^{\mathcal{L}(Y)} = -\alpha\lambda + \frac{\sigma^2}{2}\left(\lambda^2 - \lambda\right) + \int_{\mathbb{R}} \left(\mathrm{e}^{-\lambda(\mathrm{e}^x-1)} - 1 + \lambda x \mathbf{1}_{|x|\leq 1}\right) \Pi(\mathrm{d}x). \tag{3.12}$$

This expresses the cumulant function of R by means of the jump intensity of the process X.

Under the non-restrictive assumptions of Fujiwara and Miyahara (2003, Corollary 3.4) the expression (3.12) is finite for all $\lambda \in \mathbb{R}$ and $\lambda \mapsto \kappa^R(-\lambda)$ has a unique maximizer λ_* (Fujiwara and Miyahara 2003, Proposition 3.3). Under the same assumptions, R is locally bounded and it follows from the results in Biagini and Černý (2011) that ${}^{\lambda_*}/_{S_-}$ is the optimal strategy in a sufficiently wide class of admissible strategies for trading in S, therefore λ_* is the optimal dollar amount to be invested. □

Remark 3.5. In Fujiwara and Miyahara (2003) the previous calculation is performed in two steps: first the characteristics of the yield process $R = \mathcal{L}(\mathrm{e}^X)$ are computed and these are then plugged into the Lévy-Khintchin formula (3.6) of Example 3.1 to evaluate the cumulant function $\kappa^R(-\lambda)$, which after some cancellations and change of variables gives (3.12). The two-stage procedure is akin to using $\mathrm{d}t, \mathrm{d}W$ notation which, too, forces the user to keep track of the characteristics at every step of a multistage calculation, see (1.1, 1.2) and (1.4). Simplified calculus allows us to maintain a model-free formulation until the very end so that the drift calculation is performed only once, when the drift is finally needed. □

4. Drift under a change of measure

Next we will demonstrate that the calculus becomes very powerful when it comes to evaluating drifts under a different measure.

Example 4.1 (Minimal entropy martingale measure)**.** Let us continue in the economic setting of Example 3.4 with the stock price process $S = \mathrm{e}^X$, dollar yield process $R = \mathcal{L}(\mathrm{e}^X)$, exponential utility $u : w \mapsto \mathrm{e}^{-w}$, and optimal wealth process $\lambda_* R$. Under the assumptions of Fujiwara and Miyahara (2003, Corollary 3.4) the Radon-Nikodym derivative $Z_T = {}^{\mathrm{d}\mathsf{Q}}/_{\mathrm{d}\mathsf{P}}|_{\mathscr{F}_T}$ of the representative agent pricing measure is proportional to the marginal utility evaluated at the optimal wealth, that is, to $\mathrm{e}^{-\lambda_* R_T}$; see Fujiwara and Miyahara (2003, Theorem 3.1 and Corollary 4.4(3)). This Q is known in the literature as the minimal entropy martingale measure and the corresponding density process Z satisfies $Z_t = \mathrm{e}^{-\lambda_* R_t - \kappa^R(-\lambda_*)t}$ for all $t \in [0, T]$. The process Z is a true martingale thanks to Corollary C.3.

To value contingent claims on the stock $S = \mathrm{e}^X$ it is necessary to compute the characteristic function of X under Q. The required cumulant function $\kappa^X_{\mathsf{Q}}(v)$ is just the expected rate of change of $V = \mathrm{e}^{vX}$ under Q, i.e., the Q–drift rate of

$$\mathcal{L}\big(e^{vX}\big) = (\mathrm{e}^{vx} - 1) \circ X.$$

By Theorem C.2(ii) this Q–drift is the same as the P–drift of

$$\mathcal{L}\big(e^{vX}\big) + \big[\mathcal{L}\big(e^{vX}\big), \mathcal{L}\big(\mathrm{e}^{-\lambda_* R}\big)\big].$$

Recall from (3.11) that $\mathcal{L}(\mathrm{e}^{-\lambda_* R}) = (\mathrm{e}^{-\lambda_*(\mathrm{e}^x - 1)} - 1) \circ X$. In view of Example 2.3 and the composition rule (2.5) we have

$$\begin{aligned}
\big[\mathcal{L}\big(e^{vX}\big), \mathcal{L}\big(\mathrm{e}^{-\lambda_* R}\big)\big] &= (\mathrm{e}^{vx} - 1)\left(\mathrm{e}^{-\lambda_*(\mathrm{e}^x - 1)} - 1\right) \circ X;\\
\mathcal{L}\big(e^{vX}\big) + \big[\mathcal{L}\big(e^{vX}\big), \mathcal{L}\big(\mathrm{e}^{-\lambda_* R}\big)\big] &= (\mathrm{e}^{vx} - 1)\mathrm{e}^{-\lambda_*(\mathrm{e}^x - 1)} \circ X.
\end{aligned}$$

The function $\psi(x) = (\mathrm{e}^{vx} - 1)\mathrm{e}^{-\lambda_*(\mathrm{e}^x - 1)}$ satisfies

$$\psi'(0) = v \qquad \text{and} \qquad \psi''(0) = v^2 - 2\lambda_* v.$$

Consequently, if it exists, the P–drift of $\psi \circ X$ reads

$$b^{\psi \circ X} = v b^{X[h]} + \frac{c^X}{2}\left(v^2 - 2\lambda_* v\right) + \int_{\mathbb{R}} (\psi(x) - v h(x))\, F^X(\mathrm{d}x). \tag{4.1}$$

In the Lévy setting this yields

$$\begin{aligned}
\kappa^X_{\mathsf{Q}}(v) &= b^{\mathcal{L}(V)}_{\mathsf{Q}} = b^{\mathcal{L}(V) + [\mathcal{L}(V), \mathcal{L}(\mathrm{e}^{-\lambda_* R})]}\\
&= \alpha v + \frac{\sigma^2}{2}\left(v^2 - 2\lambda_* v\right) + \int_{\mathbb{R}} \left((e^{vx} - 1)\mathrm{e}^{-\lambda_*(\mathrm{e}^x - 1)} - v x \mathbf{1}_{|x| \leq 1}\right) \Pi(\mathrm{d}x),
\end{aligned} \tag{4.2}$$

whenever the integral on the right-hand side is finite. □

Remark 4.2. The standard calculus using the formulation (3.1) requires much more work. First, one must find an explicit expression for $\ln Z$, which after significant effort reads

$$\begin{aligned}
\ln Z = &- \int_0^{\cdot} \lambda_* \sigma \mathrm{d}W_s - \frac{1}{2}\int_0^{\cdot} \lambda_*^2 \sigma^2 \mathrm{d}s + \int_0^{\cdot}\int_{\mathbb{R}} -\lambda_*(\mathrm{e}^x - 1)\widehat{N}(\mathrm{d}s, \mathrm{d}x)\\
&+ \int_0^{\cdot}\int_{\mathbb{R}} \left(-\lambda_*(\mathrm{e}^x - 1) - \left(\mathrm{e}^{-\lambda_*(\mathrm{e}^x - 1)} - 1\right)\right) \Pi(\mathrm{d}x)\mathrm{d}s,
\end{aligned}$$

assuming $\ln Z$ has finite mean. Next, one constructs a new Brownian motion for the measure Q,

$$\mathrm{d}W^{\mathsf{Q}}_t = \mathrm{d}W_t + \lambda_* \sigma \mathrm{d}t,$$

and a new compensated Poisson jump measure

$$\widehat{N}^{\mathsf{Q}}(\mathrm{d}t,\mathrm{d}x)=\widehat{N}(\mathrm{d}t,\mathrm{d}x)+\left(1-\mathrm{e}^{-\lambda_*(\mathrm{e}^x-1)}\right)\Pi(\mathrm{d}x)\mathrm{d}t,$$

both using a custom-made formula, see Applebaum (2009, Theorem 5.2.12 and Exercise 5.2.14) and Øksendal and Sulem (2007, Theorem 1.32 and Lemma 1.33).

These quantities are then substituted into (3.9) to obtain

$$\begin{aligned}\mathcal{L}(\mathrm{e}^{vX})=&\int_0^\cdot\left(\alpha v+\frac{\sigma^2}{2}\left(v^2-2\lambda_*v\right)\right)\mathrm{d}s\\&+\int_0^\cdot\sigma v\mathrm{d}W_s^{\mathsf{Q}}+\int_0^\cdot\int_{\mathbb{R}}(\mathrm{e}^{vx}-1)\widehat{N}^{\mathsf{Q}}(\mathrm{d}x,\mathrm{d}s)\\&+\int_0^\cdot\int_{|x|\leq1}\left(\mathrm{e}^{-\lambda_*(\mathrm{e}^x-1)}(\mathrm{e}^{vx}-1)-vx\right)\Pi(\mathrm{d}x)\mathrm{d}s\\&+\int_0^\cdot\int_{|x|>1}\mathrm{e}^{-\lambda_*(\mathrm{e}^x-1)}(\mathrm{e}^{vx}-1)\Pi(\mathrm{d}x)\mathrm{d}s\end{aligned}\tag{4.3}$$

provided the Q–drift of $\mathcal{L}(\mathrm{e}^{vX})$ exists. The drift is now available by summing up the first, fourth, and fifth term in (4.3). In (4.2) the same result is available directly after plugging the specific form $h(x)=x\mathbf{1}_{|x|\leq1}$ and the characteristics (3.2) into the formula (4.1). The main difference between the two approaches is that (4.1) is more compact and arguably much easier to obtain than (4.3). □

We conclude this section with a bivariate example that makes use of a non-equivalent change of measure.

Example 4.3 (An option to exchange one defaultable asset for another)**.** Fix $d=2$ and let $X=(X^{(1)},X^{(2)})$ be an $\mathbb{R}^2$–valued Lévy martingale with the characteristic triplet

$$\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\sigma_1^2&\sigma_{12}\\\sigma_{12}&\sigma_2^2\end{bmatrix},\Pi\right)$$

relative to the truncation function $h(x)=x$.

Consider next two assets with price dynamics given by stochastic exponentials (see B.2) as

$$S^{(1)}=\mathscr{E}\big(X^{(1)}\big)=1+\int_0^\cdot S_{t-}^{(1)}\mathrm{d}X_t^{(1)};\qquad S^{(2)}=\mathscr{E}\big(X^{(2)}\big)=1+\int_0^\cdot S_{t-}^{(2)}\mathrm{d}X_t^{(2)}.$$

In financial economics one interprets $\mathscr{E}(X)$ as the value of a closed fund with initial investment of 1\$ following a trading strategy whose cumulative rate of return equals X. We assume that the Lévy measure Π is supported on $[-1,\infty)\times[-1,\infty)$ meaning both assets can default, perhaps simultaneously.

To value an option to exchange asset $S^{(1)}$ for asset $S^{(2)}$ on a specific date T one must compute the expectation

$$p = \mathsf{E}\left[\left(S_T^{(1)} - S_T^{(2)}\right)^+\right];$$

see Margrabe (1978). Let Q_k be the valuation measure with $S^{(k)}$ as a numéraire, i.e., ${}^{\mathrm{d}\mathsf{Q}_k}/_{\mathrm{d}\mathsf{P}} = S_T^{(k)}$ for each $k \in \{1,2\}$. Then we obtain an alternative expression for the price of the Margrabe option, namely

$$p = \mathsf{E}^{\mathsf{Q}_1}\left[\left(1 - \frac{S_T^{(2)}}{S_T^{(1)}}\right)^+\right]. \tag{4.4}$$

To evaluate (4.4) by integral transform methods one needs to compute the expectation

$$\mathsf{E}^{\mathsf{Q}_1}\left[\mathbf{1}_{\{S_T^2>0\}}\left(\frac{S_T^{(2)}}{S_T^{(1)}}\right)^v\right] \tag{4.5}$$

for certain values $v \in \mathbb{C}$. Let us fix such v. In the absence of default, the computation of (4.5) is achieved by evaluating the expected rate of change of $V = \mathbf{1}_{\{S_T^2>0\}}\left({}^{S^{(2)}}/_{S^{(1)}}\right)^v$ under the measure Q_1, in analogy to Example 4.1. This is an easy exercise in simplified calculus: for a semimartingale Y with $Y>0$ and $Y_- > 0$ one obtains

$$\frac{\mathrm{d}Y^v}{Y_-^v} = \frac{(Y_- + \mathrm{d}Y)^v - Y_-^v}{Y_-^v} = \left(1 + \frac{\mathrm{d}Y}{Y_-}\right)^v - 1$$

which yields

$$\mathcal{L}(Y^v) = ((1+y)^v - 1) \circ \mathcal{L}(Y).$$

Composition with $\mathcal{L}({}^{S^{(2)}}/_{S^{(1)}}) = \left({}^{(1+x_2)}/_{(1+x_1)} - 1\right) \circ (\mathcal{L}(S^{(1)}), \mathcal{L}(S^{(2)}))$ then gives

$$\mathcal{L}(V) = \left(\left(\frac{1+x_2}{1+x_1}\right)^v - 1\right) \circ X. \tag{4.6}$$

Representation (4.6) together with $\mathcal{L}\big(S^{(1)}\big) = X^{(1)} = x_1 \circ X$ yields

$$\mathcal{L}(V) + \big[\mathcal{L}(V), \mathcal{L}\big(S^{(1)}\big)\big] = (1+x_1)\left(\left(\frac{1+x_2}{1+x_1}\right)^v - 1\right) \circ X. \tag{4.7}$$

In conclusion, without default one obtains from Corollary C.3(ii)

$$\mathsf{E}^{\mathsf{Q}_1}\left[\mathbf{1}_{\{S_T^2>0\}}\left(\frac{S_T^{(2)}}{S_T^{(1)}}\right)^v\right] = \exp\left(b^{\psi \circ X} T\right) \tag{4.8}$$

with ψ given on the right-hand side of (4.7).

In the presence of default, V is no longer a P–semimartingale. However,

$$S^{(1)}_{\uparrow} = \mathscr{E}\big(\mathbf{1}_{x_1\neq -1}x_1 \circ X\big)$$

is Q_1–indistinguishable from $S^{(1)}$ with $S^{(1)}_{\uparrow} > 0$ and $S^{(1)}_{\uparrow -} > 0$, P–a.s. Therefore, the process

$$V_{\uparrow} = \mathbf{1}_{\{S^2>0\}}\left(\frac{S^{(2)}}{S^{(1)}_{\uparrow}}\right)^v = \mathscr{E}\left(-\mathbf{1}_{x_2=-1}\circ X\right)\left(\frac{\mathscr{E}\left(x_2\mathbf{1}_{x_2\neq -1}\circ X\right)}{\mathscr{E}\left(x_1\mathbf{1}_{x_1\neq -1}\circ X\right)}\right)^v$$
$$= \mathscr{E}\left(\left(\mathbf{1}_{x_2\neq -1}\left(\frac{1+\mathbf{1}_{x_2\neq -1}x_2}{1+\mathbf{1}_{x_1\neq -1}x_1}\right)^v - 1\right)\circ X\right)$$

is a P–semimartingale Q_1–indistinguishable from V. Corollary C.3(ii) shows that (4.8) goes through with a modified jump transformation function

$$\psi\left(x_1,x_2\right) = (1+x_1)\left(\mathbf{1}_{x_2\neq -1}\left(\frac{1+\mathbf{1}_{x_2\neq -1}x_2}{1+\mathbf{1}_{x_1\neq -1}x_1}\right)^v - 1\right). \tag{4.9}$$

We now proceed to compute the drift rate $b^{\psi\circ X}$ with ψ in (4.9). To this end, note that

$$D\psi(0) = v\begin{bmatrix} -1 & 1 \end{bmatrix}; \qquad D^2\psi(0) = v(v-1)\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Next, formula (2.13) with $h(x) = x$ for all $x \in \mathbb{R}$ yields

$$\begin{aligned} b^{\psi\circ X} =& \frac{1}{2}\left(\sigma_1^2 - 2\sigma_{12} + \sigma_2^2\right)v(v-1) \\ &+ \int_{\mathbb{R}^2}\left((1+x_1)\left(\left(\frac{1+x_2}{1+x_1}\right)^v\mathbf{1}_{x_2\neq -1} - 1\right) + vx_1 - vx_2\right)\Pi\left(\mathrm{d}x_1,\mathrm{d}x_2\right) \\ =& \frac{1}{2}\left(\sigma_1^2 - 2\sigma_{12} + \sigma_2^2\right)v(v-1) - \lambda_2^{\mathsf{Q}_1} + v\left(\lambda_2^{\mathsf{Q}_1} - \lambda_1^{\mathsf{Q}_2}\right) \\ &+ \int_{(-1,\infty)^2}\left((1+x_1)\left(\left(\frac{1+x_2}{1+x_1}\right)^v - 1\right) + vx_1 - vx_2\right)\Pi\left(\mathrm{d}x_1,\mathrm{d}x_2\right), \end{aligned}$$

as long as the expectation in (4.8) is finite. Here, the coefficient

$$\lambda_2^{\mathsf{Q}_1} = \int_{\mathbb{R}}\left(1+x_1\right)\mathbf{1}_{x_2=-1}\Pi\left(\mathrm{d}x_1,\mathrm{d}x_2\right)$$

signifies the arrival intensity of default of asset 2 under the probability measure Q_1 and $\lambda_1^{\mathsf{Q}_2}$ has the converse meaning.[7] Observe that without default $\kappa(v) = b^{\psi\circ X}$ can be interpreted as the cumulant function of $\ln S_1^{(2)}/S_1^{(1)}$ under Q_1.

[7] The coefficient $\lambda_2^{\mathsf{Q}_1}$ is the drift rate of the process $\mathbf{1}_{x_2=-1}\circ X$ under Q_1; see Theorem C.2(ii).

For concreteness let us now assume that, in the absence of default, our model follows a bivariate Merton (1976) jump-diffusion. In other words, on the open interval $(-1,\infty)\times(-1,\infty)$, the measure Π is a fixed multiple $\lambda \geq 0$ of a push-forward measure of a bivariate normal distribution with parameters

$$\left(\begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \begin{bmatrix} s_1^2 & s_{12} \\ s_{12} & s_2^2 \end{bmatrix}\right)$$

through the mapping $(z_1,z_2)\mapsto(\mathrm{e}^{z_1}-1,\mathrm{e}^{z_2}-1)$. Once the integrals have been evaluated one obtains

$$\begin{aligned}\kappa(v)=b^{\psi\circ X} &= \frac{1}{2}\left(\sigma_1^2-2\sigma_{12}+\sigma_2^2\right)v(v-1)-\lambda_2^{\mathsf{Q}_1}\\ &\quad+v\left(\lambda\left(\mathrm{e}^{m_1+\frac{1}{2}s_{11}}-\mathrm{e}^{m_2+\frac{1}{2}s_{22}}\right)+\lambda_2^{\mathsf{Q}_1}-\lambda_1^{\mathsf{Q}_2}\right)\\ &\quad+\lambda\mathrm{e}^{(1-v)m_1+vm_2+\frac{1}{2}(1-v)^2s_{11}+v(1-v)s_{12}+\frac{1}{2}v^2s_{22}}-\lambda\mathrm{e}^{m_1+\frac{1}{2}s_{11}}.\end{aligned}$$

Continuing now with the Fourier transform, Hubalek et al. (2006, Lemma 4.1) yields

$$(1-x)^+=\mathbf{1}_{x=0}+\mathbf{1}_{x>0}\int_{\beta+i\mathbb{R}}\psi(v)x^v\mathrm{d}v,\qquad x\geq 0,$$

where $\psi(v)=\frac{1}{2\pi i}\frac{1}{v(v-1)}$ and $\beta<0$. Consequently, using (4.4), the price of the Margrabe option is given as

$$\begin{aligned}p&=\mathsf{Q}_1\left[S_T^{(2)}=0\right]+\int_{\beta+i\mathbb{R}}\psi(v)\mathsf{E}^{\mathsf{Q}_1}\left[\mathbf{1}_{\left\{S_T^{(2)}>0\right\}}\left(\frac{S_T^{(2)}}{S_T^{(1)}}\right)^v\right]\mathrm{d}v\\ &=\mathrm{e}^{\kappa(0)T}+\int_{\beta+i\mathbb{R}}\frac{1}{2\pi i}\frac{1}{v(v-1)}\mathrm{e}^{\kappa(v)T}\mathrm{d}v,\end{aligned}$$

where $\kappa(0)=-\lambda_2^{\mathsf{Q}_1}$. The integrals are well defined and Fubini may be applied in the first equality because the function $v\mapsto|\psi(v)|\mathbf{1}_{\{S_T^{(2)}>0\}}(S_T^{(2)}/S_T^{(1)})^{\operatorname{Re}v}$ is product-integrable on $\beta+i\mathbb{R}$. □

5. Jumps at predictable times

The examples we have presented so far do not illustrate what happens in 'discrete-time' models. In general, 'discrete times' may occur at all rational times, meaning one is unable to enumerate the jump times in an increasing sequence. Furthermore, the 'discrete times' may be random. These features are handled in full generality in the companion papers Černý and Ruf (2019a,b).

The examples below preserve the independent increments feature of the Brownian and the Lévy-based examples in Subsection 1.1 and Sections 3 and 4. This forces the jumps to occur at fixed times. We illustrate only the 'truly discrete' case where the jump times have no finite accumulation point.

Example 5.1 (Maximization of expected utility)**.** Denote by $S > 0$ the value of a risky asset and assume the logarithmic price $X = \ln S$ is a discrete-time process (Definition B.5) with independent and identically distributed increments. Namely, for each $k \in \mathbb{N}$ we let the distribution of ΔX_k take three values, $\ln 1.1, 0,$ and $\ln 0.9$, with probabilities p_u, p_m, and p_d, respectively. With zero risk-free rate the value of a fund investing \$1 in the risky asset equals $R = \mathcal{L}(\mathrm{e}^X)$.

To evaluate the expected utility

$$\mathsf{E}\left[\mathrm{e}^{-\lambda R_t}\right], \qquad t \geq 0,$$

we recall from Example 3.4 the representation of the cumulative percentage change in $\mathrm{e}^{-\lambda R}$. Specifically, from (3.11) one obtains $\mathcal{L}(\mathrm{e}^{-\lambda R}) = \eta \circ X$ with

$$\eta(x) = \mathrm{e}^{-\lambda(\mathrm{e}^x - 1)} - 1. \tag{5.1}$$

Formulae (B.4) and (B.5) now yield

$$\begin{aligned}\mathsf{E}\left[\mathrm{e}^{-\lambda R_t}\right] &= \prod_{k=1}^{\lfloor t \rfloor} \mathsf{E}\left[1 + \eta(\Delta X_k)\right] \\ &= \prod_{k=1}^{\lfloor t \rfloor} \mathsf{E}\left[\mathrm{e}^{-\lambda(\mathrm{e}^{\Delta X_k} - 1)}\right] = \left(p_u \mathrm{e}^{-0.1\lambda} + p_m + p_d \mathrm{e}^{0.1\lambda}\right)^{\lfloor t \rfloor},\end{aligned}$$

for all $t \geq 0$. □

Example 5.2 (Minimal entropy martingale measure)**.** We now compute the minimal entropy martingale measure in the context of the previous example for some given time horizon $T > 0$. Optimizing (5.1) over λ, one obtains an explicit expression for the optimal dollar amount in the risky asset, namely

$$\lambda_* = \frac{\ln(p_u/p_d)}{0.2}.$$

Then the random variable $\mathrm{e}^{-\lambda_* R_T}/\mathsf{E}[\mathrm{e}^{-\lambda_* R_T}]$ gives the density $\mathrm{d}\mathsf{Q}/\mathrm{d}\mathsf{P}$ of the minimal entropy martingale measure Q. As in Example 4.1, we seek the expected value of e^{vX_t} under Q, for fixed $t \in [0, T]$ and $v \in \mathbb{C}$, provided the expectation is finite.

Because $\mathcal{L}(\mathrm{e}^{vX}) = \xi \circ X$ with

$$\xi(x) = \mathrm{e}^{vx} - 1,$$

the desired expectation is given by Corollary C.3(i) with η from (5.1) as follows,

$$\begin{aligned}
\mathsf{E}^{\mathsf{Q}}\left[\mathrm{e}^{vX_t}\right] &= \prod_{k=1}^{\lfloor t \rfloor} \frac{\mathsf{E}\left[(1+\eta_k(\Delta X_k))(1+\xi_k(\Delta X_k))\right]}{\mathsf{E}\left[1+\eta_k(\Delta X_k)\right]} \\
&= \prod_{k=1}^{\lfloor t \rfloor} \frac{\mathsf{E}\left[\mathrm{e}^{-\lambda_*(\mathrm{e}^{\Delta X_k}-1)}\mathrm{e}^{v\Delta X_k}\right]}{\mathsf{E}\left[\mathrm{e}^{-\lambda_*(\mathrm{e}^{\Delta X_k}-1)}\right]} = \left(\frac{(1.1^v + 0.9^v)\sqrt{p_u p_d} + p_m}{2\sqrt{p_u p_d} + p_m}\right)^{\lfloor t \rfloor}, \quad t \geq 0.
\end{aligned}$$

Here $\mathsf{E}^{\mathsf{Q}}[\mathrm{e}^{vX_t}]$ considered as a function of $v \in \mathbb{C}$ gives the moment generating function of $X_t = \ln S_t$ under Q for each $t \geq 0$ and can be therefore used to price contingent claims by integral transform methods. □

6. Concluding remarks

In this paper we have introduced the notion of 'X–representation' to describe a generic modelling situation where one starts from a (multivariate) process X whose predictable P–characteristics are given as the primitive input to the problem. The process X, which is trivially representable, is transformed by several applications of composite rules (2.1)–(2.2) to another process Y which is also X–representable. In many situations the required end product is the P–drift of Y. These examples include i) the construction of partial integro-differential equations from martingale criteria (e.g., Večeř and Xu 2004, Theorem 3.3); ii) the computation of exponential compensators (e.g., Duffie et al. 2003, Proposition 11.2); iii) the formulation of optimality conditions for various dynamic optimization problems (e.g., Øksendal and Sulem 2007, Theorem 3.1(v)).

Existing methods force us to keep track of the characteristics (drift, volatility, and jump intensities) throughout all intermediate calculations (e.g., Øksendal and Sulem 2007, Theorem 1.14). One of the drawbacks of describing processes via their characteristic triplets is that the drift and the jump intensities are measure-dependent and the drift additionally also depends on the truncation function h. The new calculus, in contrast, works with X–representations, which themselves do not depend on the characteristics in an overt way. This makes individual steps such as change of variables much simpler and the overall calculus more transparent and easy to use. An X–representation is converted into a drift only when the drift is really needed.

The proposed calculus emphasizes the universal nature of transformations such as stochastic integration or change of variables, which can typically be applied in the same way to any starting process X. For example, the conversion from the rate of return ${}^{\mathrm{d}X}/_{X_-}$ to the logarithmic return $\mathrm{d}\ln X$ always takes the form $\mathrm{d}\ln X = \ln(1 + {}^{\mathrm{d}X}/_{X_-})$. Robust results such as this are helpful in two ways. They offer an easy way to visualize fundamental relationships and separate what is fundamental from what is model-specific. Secondly, they open an avenue for studying richer models where, say, a Brownian motion is replaced with a more general process with independent increments. In the proposed calculus this is possible without additional overheads as long as the Markovian structure of the problem remains unchanged.

Further advantages of the new calculus become apparent when the drift of Y is to be computed under some new probability measure Q absolutely continuous with respect to P. The need to switch measures comes particularly from mathematical finance as illustrated in Examples 4.1, but it also arises in natural sciences as part of filtering theory (Särkkä and Sottinen 2008, and the references therein) and in Monte Carlo simulations (Grigoriu 2002, Section 5.4.2). In existing approaches a change of measure requires a custom-made formula that even depends on the form in which the density process M of ${}^{\mathrm{d}\mathsf{Q}}/_{\mathrm{d}\mathsf{P}}$ is supplied. If M is written as a stochastic exponential we need one formula, if it appears as an ordinary exponential we need another formula. These formulae convey little intuition and are consequently hard to memorize. In the new calculus there is no need to refer to a formula: we simply notice that by Girsanov's theorem the Q–drift of V equals the P–drift of $V + [V, \mathcal{L}(M)]$. Since it is typically very easy to write down the representation of $V + [V, \mathcal{L}(M)]$ the Girsanov computation comes at virtually no extra cost.

Somewhat surprisingly, the simplified calculus implies that one can perform classical Itô calculus on continuous processes by tracing the behaviour of a hypothetical pure-jump finite variation process. While this observation may seem paradoxical at first sight, we believe the emphasis on jumps makes the simplified calculus less intellectually taxing than classical approaches firmly rooted in Brownian motion.

References

Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus* (2nd ed.), Volume 116 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge.

Bender, C. and C. R. Niethammer (2008). On q-optimal martingale measures in exponential Lévy models. *Finance Stoch. 12*(3), 381–410.

Biagini, S. and A. Černý (2011). Admissible strategies in semimartingale portfolio selection. *SIAM J. Control Optim. 49*(1), 42–72.

Černý, A. and J. Ruf (2019a). Simplified calculus for semimartingales: Émery–McKean representations. In preparation.

Černý, A. and J. Ruf (2019b). Simplified calculus for semimartingales: Multiplicative compensators, PII-s, and change of measure. In preparation.

Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine processes and applications in finance. *Ann. Appl. Probab. 13*(3), 984–1053.

Émery, M. (1978). Stabilité des solutions des équations différentielles stochastiques application aux intégrales multiplicatives stochastiques. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 41*(3), 241–262.

Fujiwara, T. and Y. Miyahara (2003). The minimal entropy martingale measures for geometric Lévy processes. *Finance Stoch. 7*(4), 509–531.

Grigoriu, M. (2002). *Stochastic Calculus: Applications in Science and Engineering.* Springer, New York.

Hubalek, F., J. Kallsen, and L. Krawczyk (2006). Variance-optimal hedging for processes with stationary independent increments. *Ann. Appl. Probab. 16*(2), 853–885.

Itô, K. (1951). On a formula concerning stochastic differentials. *Nagoya Math. J. 3*, 55–65.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (2nd ed.), Volume 288 of *Comprehensive Studies in Mathematics.* Springer, Berlin.

Jeanblanc, M., S. Klöppel, and Y. Miyahara (2007). Minimal f^q-martingale measures of exponential Lévy processes. *Ann. Appl. Probab. 17*(5-6), 1615–1638.

Kallsen, J. (2000). Optimal portfolios for exponential Lévy processes. *Math. Methods Oper. Res. 51*(3), 357–374.

Larsson, M. and J. Ruf (2014). Convergence of local supermartingales and Novikov-Kazamaki type conditions for processes with jumps. Available from https://arxiv.org/abs/1411.6229.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance 33*(1), 177–186.

McKean, Jr., H. P. (1969). *Stochastic Integrals*. Probability and Mathematical Statistics, No. 5. Academic Press, New York.

Merton, R. C. (1976). Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics 3*, 125–144.

Meyer, P.-A. (1976). Un cours sur les intégrales stochastiques. In *Séminaire de Probabilités X, Strasbourg*, Volume 511 of *Lecture Notes in Math.*, pp. 245–400. Springer, Berlin.

Øksendal, B. and A. Sulem (2007). *Applied Stochastic Control of Jump Diffusions* (2nd ed.). Universitext. Springer, Berlin.

Samuelson, P. (1965). Rational theory of warrant pricing. *Ind. Manag. Rev. 6*, 13–32.

Särkkä, S. and T. Sottinen (2008). Application of Girsanov theorem to particle filtering of discretely observed continuous-time non-linear systems. *Bayesian Anal. 3*(3), 555–584.

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*, Volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge.

Večeř, J. and M. Xu (2004). Pricing Asian options in a semimartingale model. *Quant. Finance 4*(2), 170–175.

Appendix A. Notation and details about the representations

In this appendix, we provide the setup of this paper and the proofs of the statements in Section 1. Unless specified otherwise, d, m, and n are positive integers. The underlying filtered probability space is denoted by $(\Omega, \mathscr{F}, \mathfrak{F}, \mathsf{P})$. Complex integral of a locally bounded $\mathbb{C}^n$–valued process $\zeta = \zeta' + i\zeta''$ with respect to a $\mathbb{C}^n$–valued semimartingale $X = X' + iX''$ is the $\mathbb{C}$–valued semimartingale

$$\int_0^{\cdot} \zeta_t \mathrm{d}X_t = \int_0^{\cdot} \zeta_t' \mathrm{d}X_t' - \int_0^{\cdot} \zeta_t'' \mathrm{d}X_t'' + i \left(\int_0^{\cdot} \zeta_t'' \,\mathrm{d}X_t' + \int_0^{\cdot} \zeta_t' \mathrm{d}X_t'' \right).$$

We write $\overline{\mathbb{C}}^n = \mathbb{C}^n \cup \{\mathrm{NaN}\}$ for some 'non-number' $\mathrm{NaN} \notin \cup_{n \in \mathbb{N}} \mathbb{C}^n$ and $\overline{\Omega}_{\mathbb{C}}^n = \Omega \times [0, \infty) \times \overline{\mathbb{C}}^n$. The symbols $\overline{\mathbb{R}}$ and $\overline{\Omega}_{\mathbb{R}}^n$ have an analogous meaning. For a predictable function ξ we shall always assume that $\xi(\mathrm{NaN}) = \mathrm{NaN}$. If $\psi : \overline{\Omega}_{\mathbb{C}}^n \to \overline{\mathbb{C}}^m$, with $m \in \mathbb{N}$, denotes another predictable function we shall write $\psi \circ \xi$ or $\psi(\xi)$ to denote the predictable function $(\omega, t, x) \mapsto \psi(\omega, t, \xi(\omega, t, x))$ and likewise with $\mathbb{C}$ replace by $\mathbb{R}$.

Provided they exist, we write $D\xi$ and $D^2\xi$ for the complex derivatives of $\xi : \overline{\Omega}_{\mathbb{C}}^d \to \overline{\mathbb{C}}^n$, resp., the real derivatives of $\xi : \overline{\Omega}_{\mathbb{R}}^d \to \overline{\mathbb{R}}^n$. Note that $D\xi$ has dimension $n \times d$ and $D^2\xi$ has dimension $n \times d \times d$.

Definition A.1 (Two subclasses of universal representing functions)**.** Let $\mathfrak{I}_{0\mathbb{C}}^{d,n}$ denote the set of all predictable functions $\xi : \overline{\Omega}_{\mathbb{C}}^d \to \overline{\mathbb{C}}^n$ such that the following properties hold:

(1) $\xi(\omega, t, 0) = 0$ for all $(\omega, t) \in \Omega \times [0, \infty)$.
(2) There is a predictable process R locally bounded away from zero, i.e., with strictly positive running infimum, such that
   (a) $x \mapsto \xi(\omega, t, x)$ is analytic on $|x| \leq R(\omega, t)$, for all $(\omega, t) \in \Omega \times [0, \infty)$;
   (b) $\sup_{|x| \leq R} \left| D^2\xi(x) \right|$ is locally bounded.
(3) $D\xi(0)$ is locally bounded.

We write $\mathfrak{I}_{0\mathbb{C}} = \bigcup_{k,r \in \mathbb{N}} \mathfrak{I}_{0\mathbb{C}}^{k,r}$. The subclass $\mathfrak{I}_{0\mathbb{R}}$ of predictable functions $\xi : \overline{\Omega}_{\mathbb{R}}^d \to \overline{\mathbb{R}}^n$ is defined by replacing (a) with the requirement (a') $x \mapsto \xi(\omega, t, x)$ is twice differentiable for $|x| \leq R(\omega, t)$, for all $(\omega, t) \in \Omega \times [0, \infty)$. □

Let us provide some context to the previous definition. Most of the time, we are interested in 'real-valued' transformations ξ, which map an $\mathbb{R}^d$–valued semimartingale to an $\mathbb{R}^n$–valued process. The universal class $\mathfrak{I}_{0\mathbb{R}}$ is perfectly suited for this purpose. The Émery formula, as stated in (1.17), works also for complex-valued ξ if we interpret $D\xi(0)$ and $D^2\xi(0)$ as complex derivatives. Such extension from $\mathbb{R}$ to $\mathbb{C}$, albeit limited by forcing ξ to be analytic at 0, is helpful when computing characteristic functions, for example. This leads to the definition of $\mathfrak{I}_{0\mathbb{C}}$, which is now a proper subclass within a larger class of universally representing complex-valued functions $\mathfrak{I}_0$ that nests also $\mathfrak{I}_{0\mathbb{R}}$.

We do not define $\mathfrak{I}_0$ itself in this paper but it can be shown that $\mathfrak{I}_0^{d,n}$ has a one-to-one correspondence with $\mathfrak{I}_{0\mathbb{R}}^{2d,2n}$. A more general Émery formula is available for $\mathfrak{I}_0$ but will not be needed in this paper. We hence refer the interested reader to Černý and Ruf (2019a) for more details. All computations in this paper can be performed either within $\mathfrak{I}_{0\mathbb{R}}$ or within $\mathfrak{I}_{0\mathbb{C}}$ and the two classes are never used jointly. The arguments for the two cases are often identical but should be read and understood as two separate arguments because the meaning of $D\xi$ is different in the two cases. The proofs for each of the two classes are self-contained.

Let us now briefly show that all terms in (1.17) are well-defined.

Lemma A.2. *If $\xi \in \mathfrak{I}_{0\mathbb{R}} \cup \mathfrak{I}_{0\mathbb{C}}$ then the integrals $\int_0^\cdot D\xi_t(0)\mathrm{d}X_t$ and*

$$\int_0^\cdot \sum_{i,j=1}^d D^2_{i,j}\xi_t(0)\,\mathrm{d}\left[X^{(i)},X^{(j)}\right]^c_t$$

are well-defined. If, additionally, X belongs to $\mathsf{Dom}(\xi)$ then (1.19) holds.

Proof. Because $D\xi(0)$ and $D^2\xi(0)$ are locally bounded, the two integrals are well defined by Jacod and Shiryaev (2003, Theorem I.4.31). By assumption, $(\tau_n)_{n\in\mathbb{N}}$ given by $\tau_n = \inf\{t : R^*_t \leq {}^1\!/\!_n\}$ is a localizing sequence. Next, let $(\rho_n)_{n\in\mathbb{N}}$ be the localizing sequence from Definition A.1(2)(b). Then $(\tau_n \wedge \rho_n)_{n\in\mathbb{N}}$ is again a localizing sequence such that, after localization, $|\xi(x) - D\xi(0)x| \leq K|x|^2$ for all $|x| \leq \delta$ for some constants $K > 0$ and $\delta > 0$. This yields, after localization,

$$\begin{aligned}&\sum_{0<t\leq\cdot}|\xi_t(\Delta X_t) - D\xi_t(0)\Delta X_t|\\&\qquad=\sum_{\substack{0<t\leq\cdot\\|\Delta X_t|\leq\delta}}|\xi_t(\Delta X_t) - D\xi_t(0)\Delta X_t| + \sum_{\substack{0<t\leq\cdot\\|\Delta X_t|>\delta}}|\xi_t(\Delta X_t) - D\xi_t(0)\Delta X_t| < \infty\end{aligned}$$

as the last sum has only finitely many summands. □

Remark A.3. Thanks to Lemma A.2 the expression $\int_0^{\cdot} \xi(\mathrm{d}X_t) = \xi \circ X$ is now indeed well-defined by the Émery formula[8] in (1.17), provided $\xi \in \mathfrak{I}_{0\mathbb{R}}$ (resp., $\mathfrak{I}_{0\mathbb{C}}$) and $X \in \mathsf{Dom}(\xi)$. Indeed, it is easy to see that if $\xi \in \mathfrak{I}_{0\mathbb{R}} \cap \mathfrak{I}_{0\mathbb{C}}$ and $X \in \mathsf{Dom}(\xi)$ is real-valued then both the real and the complex interpretation of (1.17) yield the same result. □

We are now ready to state and prove the main properties of semimartingale representations.

Proposition A.4 (Representation of stochastic integrals)**.** *Let ζ be a locally bounded predictable $\mathbb{R}^{n\times d}$–valued (resp., $\mathbb{C}^{n\times d}$–valued) process. Then the predictable function $\xi : x \mapsto \zeta x$ belongs to $\mathfrak{I}^{d,n}_{0\mathbb{R}}$ (resp., $\mathfrak{I}^{d,n}_{0\mathbb{C}}$) and for any $\mathbb{R}^d$–valued (resp., $\mathbb{C}^d$–valued) semimartingale X one has*

$$\int_0^{\cdot} \zeta_t \mathrm{d}X_t = (\zeta x) \circ X. \tag{A.1}$$

Proof. We start with the complex-valued case. As $D\xi(x) = \zeta$ and $D^2\xi(x) = 0$ for all $x \in \mathbb{C}^d$, we have $\xi \in \mathfrak{I}_{0\mathbb{C}}$ with $X \in \mathsf{Dom}(\xi)$ for any $\mathbb{C}^d$–valued semimartingale X. The Émery formula (1.17) now yields (A.1). The real-valued proof proceeds analogously. □

Proposition A.5 (Representation of smooth transformations)**.** *Let $\mathcal{U} \subset \mathbb{R}^d$ (resp., $\mathcal{U} \subset \mathbb{C}^d$) be an open set such that $X_-, X \in \mathcal{U}$, let $f : \mathcal{U} \to \mathbb{R}^n$ be a twice continuously differentiable function (resp., let $f : \mathcal{U} \to \mathbb{C}^n$ be an analytic function), and let*

$$\xi^{f,X}(x) = \begin{cases} f(X_- + x) - f(X_-), & X_- + x \in \mathcal{U} \\ \mathrm{NaN}, & X_- + x \notin \mathcal{U} \end{cases}, \quad x \in \mathbb{R}^d \quad (\textit{resp., } x \in \mathbb{C}^d).$$

Then $\xi^{f,X} \in \mathfrak{I}^{d,n}_{0\mathbb{R}}$ (resp. $\xi^{f,X} \in \mathfrak{I}^{d,n}_{0\mathbb{C}}$), $X \in \mathsf{Dom}(\xi^{f,X})$, and

$$f(X) = f(X_0) + \xi^{f,X} \circ X.$$

Proof. The first part of the proof is identical for both cases. Note that $D\xi^{f,X}(0) = Df(X_-)$ and $D^2\xi^{f,X}(0) = D^2 f(X_-)$. As both $Df(X_-)$ and $D^2 f(X_-)$ are finite-valued predictable processes, they are locally bounded by Larsson and Ruf (2014,

[8] Émery (1978) considers only the case when ξ is deterministic and constant in time. We explicitly allow ξ to be predictable in order to develop a calculus that includes stochastic integration and Itô's formula.

Proposition 3.2). Next, denote by $R \in (0,1]$ the minimum of 1 and half of the distance from X_- to the boundary of $\mathcal{U}$ and by R^* its running infimum. The left-continuity of R now yields $R^* > 0$. Next, observe that

$$\tau_n = \inf\left\{t \geq 0 : R^*_t < \frac{1}{n}\right\} \wedge \inf\{t \geq 0 : |X_{t-}| > n\}, \qquad n \in \mathbb{N},$$

is a localizing sequence of stopping times that makes $\sup_{|x|\leq R} \left|D^2\xi(x)\right|$ locally bounded, yielding $\xi^{f,X} \in \mathfrak{I}_{0\mathbb{R}}$ (resp., $\mathfrak{I}_{0\mathbb{C}}$). As $\xi^{f,X}(\Delta X) = f(X) - f(X_-)$, we have $X \in \mathsf{Dom}(\xi^{f,X})$.

For $\xi \in \mathfrak{I}_{0\mathbb{R}}$, Lemma A.2 and the Émery formula (1.17) now yield that $f(X_0) + \xi^{f,X} \circ X$ is the Itô-Meyer change of variables formula (Jacod and Shiryaev 2003, I.4.57) and hence equal to $f(X)$. For $\xi \in \mathfrak{I}_{0\mathbb{C}}$ the result follows by identifying $\mathbb{C}^d$ with $\mathbb{R}^{2d}$ and using the real-valued statement we have just proved. □

Proposition A.6 (Composition of universally representing functions)**.** *The space $\mathfrak{I}_{0\mathbb{R}}$ is closed under dimensionally correct composition, i.e., if $\xi \in \mathfrak{I}^{d,n}_{0\mathbb{R}}$ and $\psi \in \mathfrak{I}^{n,m}_{0\mathbb{R}}$ then $\psi \circ \xi \in \mathfrak{I}^{d,m}_{0\mathbb{R}}$. An analogous statement holds for $\mathfrak{I}_{0\mathbb{C}}$.*

Proof. The proof is identical for both cases. By localization we may assume that $D\psi(0)$ is bounded and that there exists a constant $\delta_\psi > 0$ such that $\sup_{|y|\leq\delta_\psi} D^2\psi(y)$ and consequently also $\sup_{|y|\leq\delta_\psi} D\psi(y)$ are bounded. By the same construction, we may assume that there exists a constant $\delta_\xi > 0$ such that $\sup_{|x|\leq\delta_\xi} D^2\xi(x)$ and $\sup_{|x|\leq\delta_\xi} D\xi(x)$ are bounded. Moreover, there exists also $\delta_{\psi\circ\xi} \in (0, \delta_\xi)$ such that $\sup_{|x|\leq\delta_{\psi\circ\xi}} \xi(x) < \delta_\psi$.

By direct computation we now have for all $|x| \leq \delta_{\psi\circ\xi}$

$$D(\psi\circ\xi)(0) = \sum_{k=1}^{n} D_k\psi(0) D\xi^{(k)}(0);$$

$$D^2(\psi\circ\xi)(x) = \sum_{k,l=1}^{n} D^2_{k,l}\psi(\xi(x)) D\xi^{(k)}(x)^\top D\xi^{(l)}(x) + \sum_{k=1}^{n} D_k\psi(\xi(x)) D^2\xi^{(k)}(x).$$

This yields a positive non-increasing sequence $\left(\delta^{(n)}_{\psi\circ\xi}\right)_{n\in\mathbb{N}}$ and a localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times such that $D(\psi\circ\xi)(0)$ and

$$\sup_{|x|\leq\delta^{(n)}_{\psi\circ\xi}} D^2(\psi\circ\xi)(x)$$

are bounded on the stochastic interval $[\![\tau_{n-1}, \tau_n[\![$ for each $n \in \mathbb{N}$. The desired process $R_{\psi\circ\xi}$ is obtained by setting $\sum_{n\in\mathbb{N}} \delta^{(n)}_{\psi\circ\xi} \mathbf{1}_{[\![\tau_{n-1},\tau_n[\![}$. □

Appendix B. Truncation and predictable compensators

In this appendix we complement the observations in Subsection 2.2. We begin by formally introducing truncation functions.

Definition B.1 (Truncation function for X)**.** We say that a predictable function $h \in \mathfrak{I}^{d,d}_{0\mathbb{R}} \cup \mathfrak{I}^{d,d}_{0\mathbb{C}}$ is a truncation function for a semimartingale X if $X \in \mathsf{Dom}(h)$, $\sum_{0<t\leq \cdot} |\Delta X_t - h_t(\Delta X_t)| < \infty$, and if

$$X[h] = X - \sum_{0<t\leq \cdot} (\Delta X_t - h_t(\Delta X_t))$$

is a special semimartingale, i.e., if $X[h]$ can be decomposed into the sum of a local martingale and a predictable process of finite variation. □

Proposition B.2 (Universal truncation functions)**.** *If $h \in \mathfrak{I}^{d,d}_{0\mathbb{R}}$ is bounded and if $Dh(0)$ is an identity matrix then h is a truncation function for any d–dimensional semimartingale X and*

$$X[h] = X_0 + h \circ X.$$

The same statement extends to $h \in \mathfrak{I}^{d,d}_{0\mathbb{C}}$.

Proof. From Lemma A.2 with $\xi = h$ we obtain

$$\sum_{0<t\leq \cdot} |h_t(\Delta X_t) - \Delta X_t| < \infty.$$

Next, observe that $\Delta X[h]_t = h_t(\Delta X_t)$ is bounded, therefore $X[h]$ is special by Jacod and Shiryaev (2003, I.4.24). Finally, the Émery formula (1.17) yields

$$X_0 + h \circ X = X + \sum_{0<t\leq \cdot} (h_t(\Delta X_t) - \Delta X_t) = X[h],$$

which completes the proof. □

Proposition B.3 (Émery formula with truncation)**.** *Let $\xi \in \mathfrak{I}_{0\mathbb{R}}$ (resp., $\mathfrak{I}_{0\mathbb{C}}$), let $X \in \mathsf{Dom}(\xi)$, and let h be a truncation function for X. Then*

$$\sum_{0<t\leq \cdot} |\xi_t(\Delta X_t) - D\xi_t(0) h_t(\Delta X_t)| < \infty$$

and

$$\begin{aligned}\xi \circ X = \int_0^{\cdot} D\xi_t(0)\mathrm{d}X[h]_t + \frac{1}{2}\int_0^{\cdot} \sum_{i,j=1}^{d} D^2_{i,j}\xi_t(0)\,\mathrm{d}\big[X^{(i)}, X^{(j)}\big]^c_t \\ + \sum_{0<t\leq \cdot} (\xi_t(\Delta X_t) - D\xi_t(0) h_t(\Delta X_t))\,.\end{aligned} \tag{B.1}$$

Proof. First, the triangle inequality gives

$$\begin{aligned}\sum_{0<t\leq \cdot} |\xi_t(\Delta X_t) - D\xi_t(0)h_t(\Delta X_t)| &\leq \sum_{0<t\leq \cdot} |\xi_t(\Delta X_t) - D\xi_t(0)\Delta X_t| \\ &\quad + \sum_{0<t\leq \cdot} |D\xi_t(0)\Delta X_t - D\xi_t(0)h_t(\Delta X_t)| \\ &< \infty, \qquad t \geq 0.\end{aligned}$$

Here the second sum is finite thanks to Lemma A.2 and the third due to the local boundedness of $D\xi(0)$ and Definition B.1. The identity

$$\int_0^{\cdot} D\xi_t(0)\mathrm{d}X_t = \int_0^{\cdot} D\xi_t(0)\mathrm{d}X[h]_t + \sum_{0<t\leq \cdot} (D\xi_t(0)\Delta X_t - D\xi_t(0)h_t(\Delta X_t))$$

and the Émery formula (1.17) now yield the second part of the claim. □

We now introduce notation dealing with predictable compensators. If X is a special semimartingale we denote by B^X its predictable compensator, i.e., the unique predictable finite variation process starting at zero such that $X - B^X$ is a local P–martingale. If Q is another probability measure absolutely continuous with respect to P and X is Q–special we denote the corresponding Q–compensator by B^X_{Q}. We denote by C^X the continuous quadratic variation of X, i.e.,

$$C^X_{ij} = [X^{(i)}, X^{(j)}] - \sum_{0<t\leq \cdot} \Delta X^{(i)}_t \Delta X^{(j)}_t, \qquad i, j \in \{1, \dots, d\}.$$

Finally, we denote by ν^X the predictable P–compensator of the jumps of X, i.e. for any compact interval $J \subset \mathbb{R}^d$ (resp., $\mathbb{C}^d$) not containing the origin, $\nu^X([0, \cdot] \times J)$ is the predictable compensator of the finite variation process

$$\sum_{0<t\leq \cdot} \mathbf{1}_{\{\Delta X_t \in J\}}.$$

We shall say that a process is PII if it has independent increments. The following result for PII processes relates drifts to expected values, and hence shall be very useful. It is proved in Černý and Ruf (2019b).

At this point, we remind the reader that the stochastic exponential $\mathscr{E}(X)$ of a one-dimensional semimartingale X is given as the (unique) solution of the stochastic differential equation

$$\mathscr{E}(X) = 1 + \int_0^{\cdot} \mathscr{E}(X)_{t-}\mathrm{d}X_t. \tag{B.2}$$

Theorem B.4. *Let ξ be a predictable function in $\mathfrak{I}_{0\mathbb{R}} \cup \mathfrak{I}_{0\mathbb{C}}$ with $X \in \mathsf{Dom}(\xi)$. If ξ is deterministic and X is PII then $\xi \circ X$ is also PII and one has*

$$\mathsf{E}[(\xi \circ X)_t] = B_t^{\xi\circ X}, \qquad t \geq 0; \tag{B.3}$$

$$\mathsf{E}[\mathscr{E}(\xi \circ X)_t] = \mathscr{E}\big(B^{\xi\circ X}\big)_t, \qquad t \geq 0. \tag{B.4}$$

Below we evaluate the right-hand-side of (B.3) and (B.4) for two important classes of stochastic processes.

Definition B.5. We say that X is a discrete-time process if X is constant on $[k-1, k)$ for each $k \in \mathbb{N}$. We say that X is an Itô semimartingale if for all truncation functions h for X there exists a triplet $(b^{X[h]}, c^X, F^X)$ of predictable processes such that $B^{X[h]} = \int_0^\cdot b^{X[h]}\mathrm{d}t$, $C^X = \int_0^\cdot c^X \mathrm{d}t$, and ν can be written in disintegrated form as $\nu = \int_0^\cdot \int F^X(\mathrm{d}x)\mathrm{d}t$. □

Theorem B.6. *Let X be a semimartingale and let h be a truncation function for X. Let ξ be a predictable function in $\mathfrak{I}_{0\mathbb{R}} \cup \mathfrak{I}_{0\mathbb{C}}$ with $X \in \mathsf{Dom}(\xi)$ and assume $\xi \circ X$ is special. The following statements then hold.*

(i) *If X is a discrete-time process then $\xi \circ X$ is a discrete-time process and*

$$B_t^{\xi\circ X} = \sum_{k=1}^{\lfloor t\rfloor} \mathsf{E}_{k-}[\xi_k(\Delta X_k)], \qquad t \geq 0;$$

$$\mathscr{E}(B^{\xi\circ X})_t = \prod_{k=1}^{\lfloor t\rfloor} \mathsf{E}_{k-}\left[1 + \xi_k(\Delta X_k)\right], \qquad t \geq 0. \tag{B.5}$$

(ii) *If X is an Itô semimartingale then $\xi \circ X$ is an Itô semimartingale and*

$$b^{\xi\circ X} = D\xi(0) b^{X[h]} + \frac{1}{2}\sum_{i,j=1}^{d} D^2_{i,j}\xi(0) c^X_{i,j} + \int_{\mathbb{R}^d} \left(\xi(x) - D\xi(0)h(x)\right) F^X(\mathrm{d}x);$$

$$B^{\xi\circ X} = \int_0^\cdot b_t^{\xi\circ X}\mathrm{d}t; \qquad \mathscr{E}(B^{\xi\circ X}) = \exp\left(\int_0^\cdot b_t^{\xi\circ X}\mathrm{d}t\right).$$

Proof. By (B.1), we have

$$B^{\xi\circ X} = \int_0^\cdot D\xi_t(0)\mathrm{d}B_t^{X[h]} + \frac{1}{2}\int_0^\cdot \sum_{i,j=1}^{d} D^2_{i,j}\xi_t(0)\,\mathrm{d}\big[X^{(i)}, X^{(j)}\big]_t^c$$
$$+ \int_0^\cdot \int_{\mathbb{R}^d} \left(\xi_t(x) - D\xi_t(0)h_t(x)\right)\nu^X(\mathrm{d}t, \mathrm{d}x),$$

yielding the statement. □

Remark B.7. We treat discrete models and Itô semimartingales separately for pedagogical reasons. The two can be addressed jointly using that the decomposition of $X - X_0$ into the sum of an Itô semimartingale X' and a discrete-time process X'', both starting at 0, is unique if it exists. Then, for ξ in $\mathfrak{I}_0$ with $X \in \mathsf{Dom}(\xi)$ one obtains

$$\xi \circ (X' + X'') = \xi \circ X' + \xi \circ X'',$$

which is again the sum of an Itô semimartingale and a discrete-time process. Furthermore, $\xi \circ (X' + X'')$ is special if and only if both $\xi \circ X'$ and $\xi \circ X''$ are special. Then, if either of these conditions holds, one obtains

$$B^{\xi \circ (X'+X'')} = B^{\xi \circ X'} + B^{\xi \circ X''};$$
$$\mathscr{E}\left(B^{\xi \circ (X'+X'')}\right) = \mathscr{E}\left(B^{\xi \circ X'}\right) \mathscr{E}\left(B^{\xi \circ X''}\right).$$

Indeed, these formulas can be generalized even further; for this we refer to the companion papers. □

Appendix C. Change of measure

This appendix collects results on changes of measures.

Theorem C.1 (Girsanov's theorem for absolutely continuous probability measures)**.** *Let N be a P–semimartingale such that*

$$M = \mathscr{E}(N)$$

is a uniformly integrable P–martingale with $M \geq 0$. Define the probability measure Q by

$$\frac{\mathrm{d}\mathsf{Q}}{\mathrm{d}\mathsf{P}} = M_\infty.$$

For a Q–semimartingale V and a P–semimartingale $V_\uparrow$, Q–indistinguishable from V, the following are equivalent.

(1) *V is Q–special.*
(2) *$V_\uparrow + [V_\uparrow, N]$ is P–special.*

If either condition holds then the corresponding compensators are equal, i.e.,

$$B^V_{\mathsf{Q}} = B^{V_\uparrow + [V_\uparrow, N]}, \qquad \mathsf{Q}\text{–almost surely.}$$

The proof is quite classical and we do not reproduce it here. For details see Černý and Ruf (2019b).

Theorem C.2 (Girsanov's theorem – representations)**.** *Let $\eta, \xi \in \mathfrak{I}_{0\mathbb{R}}$ (resp., $\mathfrak{I}_{0\mathbb{C}}$) with $\eta \geq -1$ and X a semimartingale with $X \in \mathsf{Dom}(\eta) \cap \mathsf{Dom}(\xi)$ such that $\eta \circ X$ is special and $\Delta B^{\eta \circ X} > -1$. Assume further that $M = {}^{\mathscr{E}(\eta \circ X)}/_{\mathscr{E}\left(B^{\eta \circ X}\right)}$ is a uniformly integrable P–martingale and define the probability measure Q by*

$$\frac{\mathrm{d}\mathsf{Q}}{\mathrm{d}\mathsf{P}} = M_\infty.$$

Assume also that $\xi \circ X$ is Q–special. Then the following statements hold.

(i) *If X is a discrete-time process under P then $\xi \circ X$ is a discrete-time process under Q and*

$$B^{\xi \circ X}_{\mathsf{Q},t} = \sum_{k=1}^{\lfloor t \rfloor} \mathsf{E}_{k-}\left[\xi_k(\Delta X_k)\frac{1+\eta_k(\Delta X_k)}{1+\mathsf{E}_{k-}[\eta_k(\Delta X_k)]}\right], \qquad t \geq 0.$$

(ii) *If X is an Itô P–semimartingale then $\xi \circ X$ is an Itô Q–semimartingale and*

$$b^{\xi \circ X}_{\mathsf{Q}} = b^{(1+\eta)\xi \circ X} = D\xi(0) b^{X[h]} + \frac{1}{2}\sum_{i,j=1}^{d}\left(D^2_{i,j}\xi(0) + 2D_i\xi(0)D_j\eta(0)\right)c^X_{i,j}$$
$$+ \int_{\mathbb{R}^d}\left(\xi(x)(1+\eta(x)) - D\xi(0)h(x)\right)F^X(\mathrm{d}x).$$

Proof. In this proof all predictable functions appearing in representations are in $\mathfrak{I}_{0\mathbb{R}}$ (resp., $\mathfrak{I}_{0\mathbb{C}}$). By a standard calculation, the process $M = {}^{\mathscr{E}(\eta \circ X)}/_{\mathscr{E}\left(B^{\eta \circ X}\right)}$ satisfies $M = \mathscr{E}(N)$ with

$$N = \left(\frac{1+\eta(x)}{1+y} - 1\right) \circ (X, B^{\eta \circ X}).$$

From the representation of quadratic covariation in Example 2.3 we then obtain

$$V + [V,N] = \xi(x)\frac{1+\eta(x)}{1+y} \circ (X, B^{\eta \circ X}) = \xi(x)\frac{1+\eta(x)}{1+\Delta B^{\eta \circ X}} \circ X.$$

The rest follows from the general Girsanov theorem (Theorem C.1) and the drift formulae in Theorem B.6. □

Corollary C.3. *With the notation and assumptions as in Theorem C.2 above, if X is a PII under P stopped at a finite time and η is deterministic then $M =$*

$\mathscr{E}(\eta \circ X)/\mathscr{E}\left(B^{\eta\circ X}\right)$ *is a uniformly integrable martingale. Furthermore, if* ξ *too is deterministic then* $\xi \circ X$ *is PII under* Q *and the following statements hold for all* $t \geq 0$.

(i) *If* X *is a discrete-time process under* P *then*

$$\mathsf{E}^{\mathsf{Q}}[\mathscr{E}(\xi \circ X)_t] = \prod_{k=1}^{\lfloor t \rfloor} \mathsf{E}\left[(1 + \xi_k(\Delta X_k))\frac{1 + \eta_k(\Delta X_k)}{1 + \mathsf{E}[\eta_k(\Delta X_k)]}\right].$$

(ii) *If* X *is an Itô* P*–semimartingale then*

$$\mathsf{E}^{\mathsf{Q}}[\mathscr{E}(\xi \circ X)_t] = \exp\left(\int_0^t b_u^{(1+\eta)\xi\circ X} \mathrm{d}u\right).$$

Proof. First note that if η is deterministic and if X is PII, then Example 3.3 yields that $\eta \circ X$ is also PII. Next, Theorem B.4 yields the martingale property of M. The PII property of X under Q follows from Girsanov's theorem. The argument then follows from Theorems B.4, B.6, and C.2. □

Aleš Černý, Cass Business School, City, University of London
E-mail address: ales.cerny.1@city.ac.uk

Johannes Ruf, Department of Mathematics, London School of Economics and Political Science
E-mail address: j.ruf@lse.ac.uk