



## City Research Online

### City, University of London Institutional Repository

---

**Citation:** Černý, A. & Ruf, J. (2020). Simplified stochastics calculus with applications in Economics and Finance. *European Journal of Operational Research*, 293(2), pp. 547-560. doi: 10.1016/j.ejor.2020.12.037

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/25430/>

**Link to published version:** <https://doi.org/10.1016/j.ejor.2020.12.037>

**Copyright:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

---

---

---

City Research Online:

<http://openaccess.city.ac.uk/>

[publications@city.ac.uk](mailto:publications@city.ac.uk)

---

# SIMPLIFIED STOCHASTIC CALCULUS WITH APPLICATIONS IN ECONOMICS AND FINANCE

ALEŠ ČERNÝ AND JOHANNES RUF

**ABSTRACT.** The paper introduces a simple way of recording and manipulating general 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 calculus is fail-safe in that, under minimal assumptions, all informal calculations yield mathematically well-defined stochastic processes. The calculus is also intuitive as it allows the user to pretend all jumps are of compound Poisson type. The new calculus is very effective when it comes to computing drifts and expected values that possibly involve a change of measure. Such drift calculations yield, for example, partial integro–differential equations, Hamilton–Jacobi–Bellman equations, Feynman–Kac formulae, or exponential moments needed in numerous applications. We provide several illustrations of the new technique, among them a novel result on the Margrabe option to exchange one defaultable asset for another.

*Keywords.* finance; drift; Émery formula; Girsanov’s theorem; simplified stochastic calculus

## 1. INTRODUCTION

Anyone who has attempted stochastic modelling with jumps will be aware of the sudden increase in mathematical complexity in models that are not of compound Poisson type. The difficulty is such that experienced researchers readily forgo generality in order to reduce the technical burden placed on their readers; see, for example, [Feng and Linetsky \(2008\)](#), [Cai and Kou \(2012\)](#), [Hong and Jin \(2018\)](#), and [Aït-Sahalia and Matthys \(2019\)](#).

In this paper we introduce an intuitive calculus that works for general processes but retains the simplicity of compound Poisson calculations. To achieve this, a change of paradigm is required. Classical Itô calculus is based on decomposing the increments of every process into signal (drift, expected change) and noise (Brownian motion, zero-mean shock). This is at once convenient and mathematically expedient. The convenience of knowing the drift is immediate. 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

---

BUSINESS SCHOOL (FORMERLY CASS), CITY, UNIVERSITY OF LONDON, 106 BUNHILL ROW, LONDON, EC1Y 8TZ, UK.

DEPARTMENT OF MATHEMATICS, LSE, HOUGHTON STREET, LONDON, WC2A 2AE, UK.

*E-mail addresses:* ales.cerny.1@city.ac.uk, j.ruf@lse.ac.uk.

*Date:* 19 December, 2020.

*MSC [2010].* (Primary) 60H05, 60H10; 60G44, 60G48; (Secondary) 91B02, 91B25; 91G10.

Previous versions of this paper were circulated under the title “Finance without Brownian motions: An introduction to simplified stochastic calculus.” We thank Jan Kallsen, Jan-Frederik Mai, Lola Martinez Miranda, Johannes Muhle-Karbe, and two anonymous referees for helpful comments.

drift. Similarly, Feynman–Kac formulae reflect zero drift of an integral of costs discounted at a specified stochastic killing rate. Closer to home, the Black–Scholes partial differential equation can be obtained by setting the drift of the discounted option price process to zero under the risk-neutral measure.

The expediency of the signal–noise decomposition comes from the early construction of the Itô integral where the drift is integrated path-by-path but the Brownian motion integral is performed, loosely speaking, by summing up uncorrelated square-integrable random variables with zero mean. The paradigm shift is applied here: we separate how a process is recorded from the drift calculation. In other words, we do not carry the drift with us at all times but only evaluate it when the drift is really needed. This feels a little uncomfortable at first but there are ample rewards for the small intellectual effort required.

The consequences of the subtle change in perspective are far-reaching. By recording processes in a measure-invariant manner the technicalities of stochastic integration fall away, the importance of Brownian motions and Poisson processes recedes, and one begins to see deeper into the fundamental relationships among the modelled variables, which now take center stage. Measure change, too, becomes an easy application of the simplified calculus, for as long as the new measure is directly driven by the variables being studied, which is overwhelmingly the case in practice.

We have prepared the paper with two audiences in mind. First and foremost, the paper is intended for the research community whose members do not consider themselves experts in mathematics in general, or stochastic analysis in particular, but who nevertheless use stochastic calculus as a modelling tool. To this readership we want to demonstrate that the new calculus is easy to understand and apply in practice. Second, but no less important, we address colleagues specializing in stochastic analysis whom we wish to convince that all our arguments are mathematically rigorous.

The stated goal is not without its challenges. Where practical, we contrast the new approach with the more involved classical notation. In order to perform such comparison, one has to introduce some advanced concepts, such as the Poisson random measure, which are needed in classical stochastic calculus. Plainly, the lay readers will not be acquainted with some of the advanced concepts, nor do they need to be. Familiarity with Brownian motion, compound Poisson processes, some version of Itô’s formula, and perhaps Girsanov’s theorem ought to be enough to sufficiently appreciate the backdrop against which the new calculus is constructed. The new calculus itself only needs a grasp of drift, volatility, and jump arrival intensity, plus three basic rules that are self-evident on an informal level.

The paper is organized as follows. In the rest of this introduction, we trace how the novel concept of this paper, the semimartingale representation (1.18), arises from classical Itô calculus. Section 2 provides a thorough introduction to the simplified stochastic 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. 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.

The examples in the paper are inspired by applications in Economics and Financial Mathematics but the broader lessons are clearly applicable to Science at large. We explore wider repercussions of the proposed methodology and briefly mention applications to Statistics and Engineering in the concluding Section 6.

**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  $L$ ,

$$\frac{dK_t}{K_t} = \mu_K dt + \sigma_K dW_t, \quad (1.1)$$

$$\frac{dL_t}{L_t} = \mu_L dt + \sigma_L \left( \rho_{KL} dW_t + \sqrt{1 - \rho_{KL}^2} d\widehat{W}_t \right). \quad (1.2)$$

Here  $W$  and  $\widehat{W}$  are two independent Brownian motions. The inputs in this model are  $\mu_K$ ,  $\mu_L$ ,  $\sigma_K > 0$ ,  $\sigma_L > 0$ , and  $\rho_{KL} \in [-1, 1]$  describing the correlation between the changes in  $K$  and  $L$ . Informally, the ‘drift part’  $\mu_K dt$  represents expected change, while the ‘noise’  $\sigma_K dW_t$  is loosely interpreted as a shock with mean zero and variance  $\sigma_K^2 dt$ .

The symbol  $dK_t$  represents an increase in capital over an infinitesimal time period  $dt$ . 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  $dW_t/dt$  does not exist. However, the *expected* percentage change per unit of time is finite and equal to  $\mu_K$ . This means the expected proportional increase in capital over a fixed time horizon  $T$  equals

$$\mathbb{E} \left[ \frac{K_T}{K_0} \right] = e^{\mu_K T}. \quad (1.3)$$

In the terminology of Samuelson (1965),  $K$  and  $L$  are *geometric Brownian motions* with drift rates  $\mu_K$  and  $\mu_L$ , respectively.

Suppose we are interested in the evolution of the capital-labour ratio,  $K/L$ . The standard Itô calculus (Itô 1951, Theorem 6; Karatzas and Shreve 1991, Theorem 3.6; Duffie 2001, p. 95; Björk 2009, Theorem 4.16) yields, after simplifications,

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

One can make two observations at this point:

- (1) The formula (1.4) is not easy to decipher — the calculations are involved.
- (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) addresses (1) and informally also (2) by rewriting (1.4) in the form

$$d\left(\frac{K_t}{L_t}\right) = \frac{K_t}{L_t} \left( \frac{dK_t}{K_t} - \frac{dL_t}{L_t} - \frac{dK_t}{K_t} \frac{dL_t}{L_t} + \left(\frac{dL_t}{L_t}\right)^2 \right), \quad (1.5)$$

where  $dK_t dL_t$  is understood to stand for  $d[K, L]_t$  and  $[K, L]$  is the quadratic covariation of the processes  $K$  and  $L$ . In the present case we have

$$\frac{d[K, L]_t}{K_t L_t} = \rho_{KL} \sigma_K \sigma_L dt \quad \text{and} \quad \frac{d[L, L]_t}{L_t^2} = \sigma_L^2 dt.$$

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  $dK_t/K_t$  and  $dL_t/L_t$  are well-defined.
- (4) McKean (1969, p. 33) 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

$$df(K_t, L_t) = \frac{\partial f}{\partial k}(K_t, L_t) dK_t + \frac{\partial f}{\partial \ell}(K_t, L_t) dL_t + \frac{1}{2} \left( \frac{\partial^2 f}{\partial k^2}(K_t, L_t) (dK_t)^2 + 2 \frac{\partial^2 f}{\partial k \partial \ell}(K_t, L_t) dK_t dL_t + \frac{\partial^2 f}{\partial \ell^2}(K_t, L_t) (dL_t)^2 \right). \quad (1.6)$$

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

$$b = \mu_K - \mu_L - \rho_{KL} \sigma_K \sigma_L + \sigma_L^2. \quad (1.7)$$

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

$$\mathbb{E} \left[ \frac{K_T}{L_T} \right] = \frac{K_0}{L_0} e^{bT}. \quad (1.8)$$

The need for equation (1.4) is only illusory, however. One can obtain formula (1.7) equally easily from the measure-independent 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),

$$\frac{d(K_t/L_t)}{K_t/L_t} = \frac{dK_t}{K_t} - \frac{dL_t}{L_t} - \frac{d[K, L]_t}{K_t L_t} + \frac{d[L, L]_t}{L_t^2}, \quad (1.9)$$

$$\begin{array}{ccccccccc} \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ b & = & \mu_K & - & \mu_L & - & \rho_{KL} \sigma_K \sigma_L & + & \sigma_L^2. \end{array} \quad (1.10)$$

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:  $\mu_K, \mu_L, \sigma_K, \sigma_L$ , and  $\rho_{KL}$ .

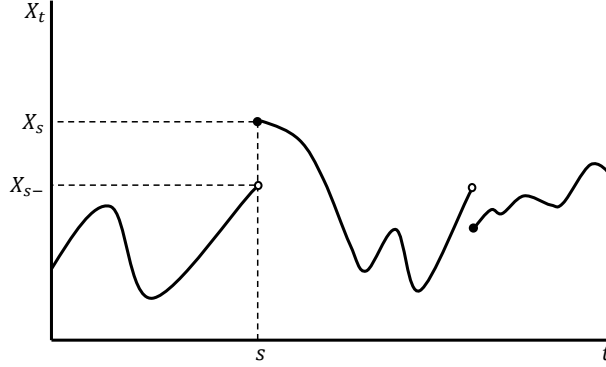


FIGURE 1.1. Illustration of a right-continuous path with left limits (process  $X$ ).

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,

$$\frac{dK_t}{K_{t-}} = \mu_K dt + \sigma_K dW_t + dJ_t^K; \quad (1.11)$$

$$\frac{dL_t}{L_{t-}} = \mu_L dt + \sigma_L \left( \rho_{KL} dW_t + \sqrt{1 - \rho_{KL}^2} d\widehat{W}_t \right) + dJ_t^L; \quad (1.12)$$

$$dJ_t^K = \int_{|x_1| \leq 1} x_1 (N(dt, d(x_1, x_2)) - \Pi(d(x_1, x_2))dt) + \int_{|x_1| > 1} x_1 N(dt, d(x_1, x_2)); \quad (1.13)$$

$$dJ_t^L = \int_{|x_2| \leq 1} x_2 (N(dt, d(x_1, x_2)) - \Pi(d(x_1, x_2))dt) + \int_{|x_2| > 1} x_2 N(dt, d(x_1, x_2)). \quad (1.14)$$

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

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

is the bivariate cumulative distribution of jump sizes. The complicated expression in (1.14) is required to accommodate the models where the small jumps of  $L$  have infinite variation while, at the same time, the large jumps of  $L$  have infinite mean.

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 (1.11)–(1.12); 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 (1.12) — the percentage change in labour supply  $L$  — does not have a finite first moment, while the mean of  $K/L$  is nevertheless finite.<sup>1</sup>

<sup>1</sup>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. There are other circumstances where infinite mean arises more naturally, for example stock price  $S$  will typically have finite mean but  $\ln S$  may plausibly have mean equal to  $-\infty$  in a model with jumps if returns close to -100% occur frequently enough.

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 (1.14). This makes (1.12) look more like (1.2) and largely represents the current practice in applications where jumps of the most general type are considered, 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).

In this paper, we do handle arbitrary jumps but 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 (1.6) 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

$$d\left(\frac{K_t}{L_t}\right) = \frac{K_{t-} + dK_t}{L_{t-} + dL_t} - \frac{K_{t-}}{L_{t-}}, \quad (1.15)$$

which informally leads to an expression for percentage changes

$$\frac{d(K_t/L_t)}{K_{t-}/L_{t-}} = \frac{1 + dK_t/K_{t-}}{1 + dL_t/L_{t-}} - 1. \quad (1.16)$$

For Émery, the right-hand side of (1.16) represents a deterministic, time-constant function

$$\eta(x_1, x_2) = \frac{1 + x_1}{1 + x_2} - 1 \quad (1.17)$$

that acts on the increments  $dK_t/K_{t-}$  and  $dL_t/L_{t-}$ . The real meaning of the expression  $\xi(dX_t)$  for a generic deterministic time-constant  $\mathcal{C}^2$  function  $\xi$  with  $\xi(0) = 0$  and any semimartingale  $X$  is supplied by the Émery formula<sup>2</sup>

$$\xi_t(dX_t) = D\xi_t(0)dX_t + \frac{1}{2} \sum_{i,j=1}^d D_{ij}^2 \xi_t(0) d[X^{(i)}, X^{(j)}]_t^c + (\xi_t(\Delta X_t) - D\xi_t(0)\Delta X_t), \quad (1.18)$$

where the last term yields absolutely summable jumps of finite variation and  $[\cdot, \cdot]^c$  stands for the continuous part of quadratic covariation. An easy way to memorize the Émery formula is suggested in equation (2.12) and the subsequent paragraph.

Having assigned meaning to the right-hand side of (1.16) via the formula (1.18) with  $\xi = \eta$  from (1.17), equality (1.16) is no longer an informal 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 (1.15), too, represents a function that acts on the increments of the underlying processes; in this case it acts on  $dK_t$  and  $dL_t$ . One important difference is that the

<sup>2</sup>The symbols  $D\xi$  and  $D^2\xi$  stand for first and second partial derivatives of  $\xi$ . Émery (1978) considers only the case when  $\xi$  is deterministic and constant in time. We explicitly allow  $\xi$  to be predictable in order to develop a calculus that includes stochastic integration and Itô’s formula.



function in question is no longer deterministic, i.e., we have

$$d\left(\frac{K_t}{L_t}\right) = \tilde{\eta}_t(dK_t, dL_t)$$

with

$$\tilde{\eta}_t(x_1, x_2) = \frac{K_{t-} + x_1}{L_{t-} + x_2} - \frac{K_{t-}}{L_{t-}}. \quad (1.19)$$

Another practical consideration is that the predictable function  $\tilde{\eta}$  in (1.19) is not finite-valued at the point where  $x_2 = -L_-$ .

To accommodate functions with restricted domain, we say that

a predictable function  $\xi$  is *compatible* with  $X$  if  $\xi(\Delta X)$  is finite-valued,  $\mathbb{P}$ -a.s.

It is formally straightforward to apply the formula (1.18) to a predictable function  $\xi$ . However, it is *a priori* not clear that the formula will be well-defined; in particular, for  $\xi = \tilde{\eta}$  in (1.19) it is not clear that one obtains

$$\sum_{0 < t \leq \cdot} |\xi_t(\Delta X_t) - D\xi_t(0)\Delta X_t| < \infty. \quad (1.20)$$

Ideally, we would like to have a calculus that gives rise only to those predictable functions  $\xi$  for which the integral  $\int_0^\cdot \xi_t(dX_t)$  is always well-defined so we do not have to manually check admissibility every time a new  $\xi$  arises. The precise formulation of *two* such subclasses, whose elements will be called *universal representing functions*, is given in Appendix A. We denote by  $\mathfrak{J}_{0\mathbb{R}}^{d,n}$  the subset of universal representing functions that map  $\mathbb{R}^d$ -valued processes to  $\mathbb{R}^n$ -dimensional processes. We also set  $\mathfrak{J}_{0\mathbb{R}} = \bigcup_{d,n \in \mathbb{N}} \mathfrak{J}_{0\mathbb{R}}^{d,n}$ . For computations involving characteristic functions, it is convenient to generalize the Émery formula (1.18) to a subset of complex predictable functions by interpreting  $D\xi$  and  $D^2\xi$  as complex derivatives. The subset of all universal representing functions that are complex-differentiable near the origin will be denoted by  $\mathfrak{J}_{0\mathbb{C}}$ .

Propositions A.3–A.5 in Appendix A show that the class  $\mathfrak{J}_{0\mathbb{R}}$  is self-contained; if we use standard real-valued operations we are guaranteed to stay within  $\mathfrak{J}_{0\mathbb{R}}$ . The class  $\mathfrak{J}_{0\mathbb{C}}$  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  $\mathfrak{J}_{0\mathbb{R}}$  (resp.,  $\mathfrak{J}_{0\mathbb{C}}$ ) 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 stochastic calculus is most intuitive when expressed in differential form, such as (1.15) and (1.16). When one wishes to speak of the integrated process whose increments are equal to  $\xi_t(dX_t)$ , one typically just introduces a new label, say  $Y$ , writing  $dY_t = \xi_t(dX_t)$ . 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  $\xi$  and  $X$  are acknowledged explicitly. Accordingly, the process with increments  $\xi_t(dX_t)$ , starting at 0, will be denoted by  $\xi \circ X$ , i.e.,

$$\xi \circ X = \int_0^\cdot \xi_t(dX_t).$$

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

$$\begin{aligned} [X, X] &= \int_0^\cdot (dX_t)^2 = \text{id}^2 \circ X; \\ [[X, X], X] &= \int_0^\cdot (dX_t)^3 = \text{id}^3 \circ X; \\ [[[X, X], X], X] &= [[X, X], [X, X]] = \int_0^\cdot (dX_t)^4 = \text{id}^4 \circ X. \end{aligned}$$

Here we let  $\text{id}$  denote the identity function;

$$\text{id}(x) = x.$$

Below we use  $\text{id}_1(x) = x_1$  for the first component,  $\text{id}_2$  for the second, etc., where required.

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  $|\text{id}|^\alpha \circ X$  is well-defined for any semimartingale  $X$  and any real  $\alpha \geq 2$ . 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.,

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

Thus, in the integral notation formula (1.16) reads

$$\mathcal{L}\left(\frac{K}{L}\right) = \left(\frac{1 + \text{id}_1}{1 + \text{id}_2} - 1\right) \circ (\mathcal{L}(K), \mathcal{L}(L)). \quad (1.21)$$

Provided that the cumulative percentage changes in  $K$  and  $L$  are well-defined, formula (1.21) holds for arbitrary semimartingales  $K$  and  $L$ ; it is new in this generality as far as we know.

## 2. SIMPLIFIED STOCHASTIC CALCULUS

**2.1. Composite rules.** The simplified stochastic calculus rests on a sequential application of Propositions A.3–A.5. 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, we have  $dY_t = \xi_t(dX_t)$  and the first two rules read as follows.

**Stochastic integration with respect to  $Y$ :** For a locally bounded process  $\zeta$  and the integral  $Z = \int_0^\cdot \zeta_t dY_t$  one has

$$dZ_t = \zeta_t \xi_t(dX_t).$$

**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

$$dZ_t = df(Y_t) = f(Y_{t-} + dY_t) - f(Y_{t-}) = f(Y_{t-} + \xi_t(dX_t)) - f(Y_{t-}).$$

We now restate the above fully rigorously in integral notation.

**Corollary 2.1.** *Assume  $\xi \in \mathfrak{J}_{0\mathbb{R}}^{d,n}$  (resp.,  $\mathfrak{J}_{0\mathbb{C}}^{d,n}$ ) is compatible with  $X$  and consider the  $n$ -dimensional process*

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

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

TABLE 2.1. Schematic derivation of (1.21) by means of Corollary 2.1

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  $\zeta$  we have  $\zeta \xi \in \mathfrak{J}_{0\mathbb{R}}^{d,m}$  (resp.,  $\mathfrak{J}_{0\mathbb{C}}^{d,m}$ ) and

$$Z = Z_0 + \int_0^\cdot \zeta_u dY_u = Z_0 + \zeta \xi \circ X; \quad (2.1)$$

- Smooth transformation ('Itô's formula'): Assume  $Y$  and  $Y_-$  remain in an open subset  $\mathcal{U}$  of  $\mathbb{R}^n$  (resp.,  $\mathbb{C}^n$ ) where the function  $f : \mathcal{U} \rightarrow \mathbb{R}^m$  is twice continuously differentiable (resp. where  $f : \mathcal{U} \rightarrow \mathbb{C}^m$  is analytic). Then  $f(Y_- + \xi) - f(Y_-)$  is in  $\mathfrak{J}_{0\mathbb{R}}^{d,m}$  (resp.,  $\mathfrak{J}_{0\mathbb{C}}^{d,m}$ ) and it is compatible with  $X$ . Furthermore,

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

Let us 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

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

Observe that  $X$  is always representable with respect to itself thanks to Proposition A.3 with  $\zeta$  equal to the  $d \times d$  identity matrix. Starting off from the trivial representation  $X = X_0 + \text{id} \circ X$ , 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 (relabelled  $Y$ ) becomes the input to the next application of Corollary 2.1 producing the next intermediate output  $Z$ . The  $Y \rightarrow Z$  pattern is repeated until one reaches the desired output  $Z$ ; in our example the goal is

$$Z = \mathcal{L}\left(\frac{K}{L}\right). \quad (2.4)$$

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  $\xi$  even without applying formal rules. In the context of (1.21), 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 (1.21), 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  $S > 0$  stand for the value of an investment with  $S_- > 0$  and let  $X = \mathcal{L}(S) = \int_0^\cdot dS_t/S_{t-}$  be the cumulative rate of return on this investment. On a purely intuitive level, thinking only of jump transformations, one can write

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

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

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

Both approaches yield the representation

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

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

$$XY = X_0Y_0 + \int_0^\cdot X_{t-}dY_t + \int_0^\cdot Y_{t-}dX_t + [X, Y].$$

This yields

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

More formally, the integration and smooth transformation rules yield

$$[X, Y] = ((X_- + \text{id}_1)(Y_- + \text{id}_2) - X_-Y_- - X_- \text{id}_2 - Y_- \text{id}_1) \circ (X, Y) = \text{id}_1 \text{id}_2 \circ (X, Y).$$

Thus, in the differential notation one can rigorously write  $d[X, X]_t = (dX_t)^2$  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.5, 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

$$dY_t = \xi_t(dX_t) \quad \text{and} \quad dZ_t = \psi_t(dY_t) \quad \text{yield} \quad dZ_t = \psi_t(\xi_t(dX_t)). \quad (2.5)$$

**Corollary 2.4** (Composition of representations). *Assume  $\xi \in \mathfrak{J}_{0\mathbb{R}}^{d,n}$  is compatible with  $X$  and consider the  $n$ -dimensional process  $Y = Y_0 + \xi \circ X$ . For  $\psi \in \mathfrak{J}_{0\mathbb{R}}^{n,m}$  compatible with  $Y$  one obtains that  $\psi(\xi) \in \mathfrak{J}_{0\mathbb{R}}^{d,m}$  is compatible with  $X$  and*

$$Z = Z_0 + \psi \circ Y = Z_0 + \psi(\xi) \circ X. \quad (2.6)$$

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

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  $(\ln K, \ln L)$  as the primitive input. Thanks to Corollary 2.4, there is no need to calculate everything afresh all the way from  $(\ln K, \ln L)$  to  $\mathcal{L}(K/L)$ . One only computes the passage from  $(\ln K, \ln L)$  to  $(\mathcal{L}(K), \mathcal{L}(L))$  which yields (see equation (3.4) below)

$$(\mathcal{L}(K), \mathcal{L}(L)) = (e^{\text{id}_1} - 1, e^{\text{id}_2} - 1) \circ (\ln K, \ln L),$$

while the passage from  $(\mathcal{L}(K), \mathcal{L}(L))$  to  $\mathcal{L}(K/L)$  can be recycled from (1.21). The two results composed together give

$$\mathcal{L}\left(\frac{K}{L}\right) = (e^{\text{id}_1 - \text{id}_2} - 1) \circ (\ln K, \ln L). \quad (2.7)$$

In differential notation,

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

substituted into (1.16) yields

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

which is the differential equivalent of formula (2.7).

**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 (1.9)–(1.10). Our task is to express the drift of the represented process with the help of the characteristics of the representing process.

To begin with, we collect the predictable characteristics (i.e., the drift rate, the covariance matrix of the associated Brownian motion, and the Lévy measure) of the input process  $X = (\mathcal{L}(K), \mathcal{L}(L))$  in equations (1.11)–(1.14) in the more compact form

$$b^{X[h^1]} = \begin{bmatrix} \mu_K \\ \mu_L \end{bmatrix}; \quad c^X = \begin{bmatrix} \sigma_K^2 & \rho_{KL}\sigma_K\sigma_L \\ \rho_{KL}\sigma_K\sigma_L & \sigma_L^2 \end{bmatrix}; \quad F^X = \Pi. \quad (2.8)$$

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

$$X[h^1] = X_0 + (\text{id}_1 \mathbf{1}_{|\text{id}_1| \leq 1}, \text{id}_2 \mathbf{1}_{|\text{id}_2| \leq 1}) \circ X. \quad (2.9)$$

Observe that this precisely matches the decomposition of jumps appearing in (1.13)–(1.14) and ensures that the drift of  $X[h^1]$  is finite.<sup>3</sup> 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  $X[h^1]$  corresponds to the choice

$$h^1 = (\text{id}_1 \mathbf{1}_{|\text{id}_1| \leq 1}, \dots, \text{id}_d \mathbf{1}_{|\text{id}_d| \leq 1}), \quad (2.10)$$

where  $d$  is the dimension of  $X$ . 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  $X[h^1]$ , whose universal representation appears in (2.9). Nonetheless, the situations where  $X[0]$  exists do arise in practice, for example, in the Merton (1976) jump-diffusion model. In such models we may write

$$dX_t = dX[0]_t + \Delta X_t. \quad (2.11)$$

This, when substituted into (1.18), leads to the simplified expression

$$\xi_t(dX_t) = D\xi_t(0)dX[0]_t + \frac{1}{2} \sum_{i,j=1}^d D_{ij}^2 \xi_t(0) d[X^{(i)}, X^{(j)}]_t^c + \xi_t(\Delta X_t), \quad (2.12)$$

which offers a valuable insight into the nature of the Émery formula. We observe that the first two terms of (2.12) correspond to the McKean calculus for the continuous part of  $X$  while the last term accounts for the jumps in  $X$ . The two components do not interact and can be treated separately. In the most general situation where the decomposition (2.11) does *not* exist, one can make (2.12) rigorous by adding small jumps to the first term and subtracting them in the last term to obtain

$$\xi_t(dX_t) = D\xi_t(0)dX[h]_t + \frac{1}{2} \sum_{i,j=1}^d D_{ij}^2 \xi_t(0) d[X^{(i)}, X^{(j)}]_t^c + (\xi_t(\Delta X_t) - D\xi_t(0)h(\Delta X_t)). \quad (2.13)$$

The original Émery formula (1.18) corresponds simply to the case where *all* the jumps of  $X$  have been added to the first term and subtracted in the last term of (2.12).

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. In a univariate case, one would thus pick  $h = 0$  if jumps of  $X$  have finite variation as in the Merton jump-diffusion model, failing that,  $h = \text{id}$  if the drift of  $X$  exists, and finally  $h = h^1 = |\text{id}| \mathbf{1}_{|\text{id}| \leq 1}$  in all remaining cases. In a multivariate case this choice can be performed component-wise. With such choice of  $h$ , the drift of each contributing term in (2.13) is guaranteed to be finite.

We can now perform the feat previously achieved on a smaller scale in (1.9)–(1.10). By matching each term in (2.13) with its drift contribution, one obtains

$$b^{\xi \circ X} = D\xi(0)b^{X[h]} + \frac{1}{2} \sum_{i,j=1}^d D_{ij}^2 \xi(0) c_{ij}^X + \int_{\mathbb{R}^d} (\xi(x) - D\xi(0)h(x)) F^X(dx). \quad (2.14)$$

Formula (2.14) is proved in Theorem B.6.

<sup>3</sup>In contrast, the drift of  $X$  need not be well-defined in general; see also Footnote 1.

Specifically, with  $\xi$  given in (1.17) we obtain

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

For the specific input parameters in (2.8) and the corresponding truncation function in (2.10) the drift conversion formula (2.14) yields

$$\begin{aligned} b^{\mathcal{L}(K/L)} &= \mu_K - \mu_L - \rho_{KL}\sigma_K\sigma_L + \sigma_L^2 \\ &+ \int_{\mathbb{R}^2} \left( \frac{1+x_1}{1+x_2} - 1 - (x_1\mathbf{1}_{|x_1|\leq 1} - x_2\mathbf{1}_{|x_2|\leq 1}) \right) \Pi(d(x_1, x_2)), \end{aligned} \quad (2.15)$$

which is the appropriate generalization of (1.10) provided the integral in (2.15) converges.<sup>4</sup> Formula (1.8) continues to hold with this choice of  $b$ ; see Theorems B.4 and B.6 below.

### 3. FURTHER EXAMPLES WITH DRIFT COMPUTATION

We will now showcase the strength of process representations such as (1.21) 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,

$$X = X_0 + \int_0^\cdot \alpha ds + \int_0^\cdot \sigma dW_s + \int_0^\cdot \int_{|x|\leq 1} x \widehat{N}(ds, dx) + \int_0^\cdot \int_{|x|>1} x N(ds, dx), \quad (3.1)$$

where  $N$  is a Poisson jump measure,  $\Pi$  the corresponding Lévy measure,

$$\widehat{N}(dt, dx) = N(dt, dx) - \Pi(dx)dt$$

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

In the simplified stochastic calculus we will never have to write out the decomposition (3.1) in full. Instead we just note that  $X$  is an Itô semimartingale<sup>5</sup> with characteristics

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

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  $\mathcal{L}(e^{vX})$  for  $v \in \mathbb{C}$ ). By integration and smooth transformation we have

$$d\mathcal{L}(e^{vX})_t = \frac{de^{vX_t}}{e^{vX_{t-}}} = \frac{e^{v(X_{t-}+dX_t)} - e^{vX_{t-}}}{e^{vX_{t-}}} = e^{v dX_t} - 1, \quad (3.3)$$

or equivalently,

$$\mathcal{L}(e^{vX}) = (e^{v \text{id}} - 1) \circ X. \quad (3.4)$$

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

$$\Pi(dx_1, dx_2) = \mathbf{1}_{x_2=0}\Pi^K(dx_1) + \mathbf{1}_{x_1=0}\Pi^L(dx_2),$$

meaning capital and labour do not jump simultaneously. We may take  $\Pi^K(dx_1)$  to be lognormal so that  $\int_0^\infty x_1 \Pi^K(dx_1) < \infty$  and  $K$  has finite mean. The choice  $\Pi^L(dx_2) = x_2^{-2} \mathbf{1}_{x_2>0} dx_2$  then provides an example where  $L$  has infinite variation (even with  $\sigma^L = 0$ ) and infinite mean while the mean of  $K/L$  remains finite (see also Theorem B.4).

<sup>5</sup>A generalization of (3.1) where  $\alpha$ ,  $\sigma^2$ , and  $\Pi$  are allowed to be stochastic; see Definition B.5 below.

The representing function is  $\xi = e^{v\text{id}} - 1$  with  $\xi'(0) = v$  and  $\xi''(0) = v^2$ . The corresponding Émery formula (2.13) reads<sup>6</sup>

$$\xi_t(dX_t) = v dX[h]_t + \frac{1}{2}v^2 d[X, X]_t^c + \left( e^{v\Delta X_t} - 1 - v h(\Delta X_t) \right). \quad (3.5)$$

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 the drift rate of

$$b^{\mathcal{L}(e^{vX})} = vb^{X[h]} + \frac{1}{2}v^2 c^X + \int_{\mathbb{R}} (e^{vx} - 1 - v h(x)) F^X(dx). \quad (3.6)$$

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

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

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

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

The drift rate  $\kappa^X(v) = b^{\mathcal{L}(e^{vX})}$  is known as the cumulant function of the r.v.  $X_1 - X_0$ .<sup>7</sup>  $\square$

*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

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

Integration yields (Applebaum 2009, Section 4.3.3)

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

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

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

The calculations (3.8)–(3.9) become much easier in the approach (3.3)–(3.5) because the rules of the simplified stochastic calculus are more compact and easier to remember.  $\square$

<sup>6</sup>A helpful mnemonic device for the Émery formula is shown in equation (2.12) and the subsequent paragraph.

<sup>7</sup>Formula (3.7) holds thanks to Theorems B.4 and B.6. Note that (3.7) is in fact the Lévy-Khintchin formula applied to the Lévy process  $X$  (Sato 1999, Theorem 8.1).



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.

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

- (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 (2.6) now yields  $Y[g] = Y_0 + g(\xi) \circ X$ . As  $Dg(0)$  is by assumption an identity matrix, we have  $D(g \circ \xi)(0) = D\xi(0)$ ,  $D^2(g \circ \xi) = D^2\xi(0)$ , and the desired drift is

$$b^{Y[g]} = D\xi(0)b^{X[h]} + \frac{1}{2} \sum_{i,j=1}^d D_{ij}^2 \xi(0) c_{ij}^X + \int_{\mathbb{R}} (g(\xi(x)) - D\xi(0)h(x)) F^X(dx).$$

- (3) Finally, let  $G$  be a closed  $n$ -dimensional set not containing zero. The process  $\mathbf{1}_{\text{id} \in G} \circ Y$  counts the jumps of  $Y$  whose size is in  $G$ ; its drift yields the jump arrival intensity  $F^Y(G)$ . The chain rule (2.6) gives  $\mathbf{1}_{\text{id} \in G} \circ Y = \mathbf{1}_{\xi \in G} \circ X$ . The function  $\psi = \mathbf{1}_{\xi \in G}$  satisfies  $D\psi(0) = D^2\psi(0) = 0$  which yields

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

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

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

$$b^{Y[h^1]} = \alpha v + \frac{1}{2} \sigma^2 v^2 + \int_{\mathbb{R}} ((e^{vx} - 1) \mathbf{1}_{|e^{vx} - 1| \leq 1} - vx \mathbf{1}_{|x| \leq 1}) \Pi(dx);$$

$$c^Y = \sigma^2 v^2;$$

$$F^Y(G) = \int_{\mathbb{R}} \mathbf{1}_G(e^{vx} - 1) \Pi(dx).$$

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

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 1 and of one risky asset with price process  $S = e^X$ . Moreover, consider an agent with exponential utility function  $u : w \mapsto -e^{-w}$ .

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  $R = \mathcal{L}(e^X)$  the cumulative yield on an 1\$ investment in the risky asset. Normalizing initial wealth to zero, the optimal wealth process equals  $\lambda R$  and its expected utility is  $\mathbb{E}[e^{-\lambda R_T}]$ . In analogy to (1.3) the expected utility can be obtained via the time rate of the expected percentage change of the quantity  $e^{-\lambda R}$ . This is nothing other than the drift rate of the process  $\mathcal{L}(e^{-\lambda R})$ . Provided this drift, commonly denoted by  $\kappa^R(-\lambda)$ , is finite, the expected utility will be equal to  $\mathbb{E}[e^{-\lambda R_T}] = e^{\kappa^R(-\lambda)T}$ , cf. (3.7).

Formula (3.3) and the composition rule (2.5) give  $dR_t = de^{X_t}/e^{X_{t-}} = e^{dX_t} - 1$  and

$$d\mathcal{L}(e^{-\lambda R})_t = \frac{de^{-\lambda R_t}}{e^{-\lambda R_{t-}}} = e^{-\lambda dR_t} - 1 = e^{-\lambda(e^{dX_t} - 1)} - 1. \quad (3.10)$$

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

$$\kappa^R(-\lambda) = b^{\mathcal{L}(e^{-\lambda R})} = -\alpha\lambda + \frac{\sigma^2}{2}(\lambda^2 - \lambda) + \int_{\mathbb{R}} \left( e^{-\lambda(e^x - 1)} - 1 + \lambda x \mathbf{1}_{|x| \leq 1} \right) \Pi(dx). \quad (3.11)$$

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 (3.11) is finite for all  $\lambda \in \mathbb{R}$  and  $\lambda \mapsto \kappa^R(-\lambda)$  has a unique maximizer  $\lambda_*$  (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  $\lambda_*$  is the optimal dollar amount to be invested.  $\square$

*Remark 3.5.* In Fujiwara and Miyahara (2003) the previous calculation is performed in two steps: first the characteristics of the yield process  $R = \mathcal{L}(e^X)$  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 (3.11). The two-stage procedure is akin to using  $dt, dW$  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). The simplified stochastic 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.  $\square$

The online appendix Černý and Ruf (2020c) discusses affine models and the derivation of Riccati equations as an additional example.

#### 4. DRIFT UNDER A CHANGE OF MEASURE

Next we will demonstrate that the simplified 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  $S = e^X$ , dollar yield process  $R = \mathcal{L}(e^X)$ , exponential utility  $u : w \mapsto e^{-w}$ , and optimal wealth process  $\lambda_* R$ . Under the assumptions of Fujiwara and Miyahara (2003, Corollary 3.4) the Radon-Nikodym derivative  $Z_T = d\mathbb{Q}/d\mathbb{P}|_{\mathcal{F}_T}$  of the representative agent pricing measure is proportional to the marginal utility evaluated at the optimal wealth, that is, to  $e^{-\lambda_* R_T}$ ; see Fujiwara and Miyahara (2003, Theorem 3.1 and Corollary 4.4(3)). This  $\mathbb{Q}$  is known in the literature as the minimal entropy martingale measure and the corresponding density process  $Z$  satisfies  $Z_t = e^{-\lambda_* R_t - \kappa^R(-\lambda_*)t}$  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  $S = e^X$ , it is necessary to compute the characteristic function of  $X$  under  $\mathbb{Q}$ . The required cumulant function  $\kappa_{\mathbb{Q}}^X(v)$  is just the expected rate of change of  $e^{vX}$  under  $\mathbb{Q}$ , i.e., the  $\mathbb{Q}$ -drift rate of  $\mathcal{L}(e^{vX})_t$ . By Theorem C.2(ii) and Example 2.3 this  $\mathbb{Q}$ -drift is the same as the  $\mathbb{P}$ -drift of

$$\begin{aligned} d\mathcal{L}(e^{vX})_t + d[\mathcal{L}(e^{vX}), \mathcal{L}(e^{-\lambda_* R})]_t &= d\mathcal{L}(e^{vX})_t + d\mathcal{L}(e^{vX})_t d\mathcal{L}(e^{-\lambda_* R})_t \\ &= (e^{v dX_t} - 1)e^{-\lambda_*(e^{dX_t} - 1)}, \end{aligned}$$

where the second equality combines the representations of  $\mathcal{L}(e^{vX})$  and  $\mathcal{L}(e^{-\lambda_* R})$  obtained earlier in (3.3) and (3.10).

The function  $\psi(x) = (e^{vx} - 1)e^{-\lambda_*(e^x - 1)}$  satisfies  $\psi'(0) = v$  and  $\psi''(0) = v^2 - 2\lambda_*v$ . Consequently, if the drift exists, the  $\mathbb{P}$ -drift rate of  $\psi \circ X$  reads

$$b^{\psi \circ X} = vb^{X[h]} + \frac{c^X}{2} (v^2 - 2\lambda_*v) + \int_{\mathbb{R}} (\psi(x) - vxh(x)) F^X(dx). \quad (4.1)$$

In the Lévy setting this yields

$$\begin{aligned} \kappa_{\mathbb{Q}}^X(v) &= b_{\mathbb{Q}}^{\mathcal{L}(e^{vX})} = b^{\mathcal{L}(e^{vX}) + [\mathcal{L}(e^{vX}), \mathcal{L}(e^{-\lambda_* R})]} \\ &= \alpha v + \frac{\sigma^2}{2} (v^2 - 2\lambda_*v) + \int_{\mathbb{R}} \left( (e^{vx} - 1)e^{-\lambda_*(e^x - 1)} - vx\mathbf{1}_{|x| \leq 1} \right) \Pi(dx), \end{aligned} \quad (4.2)$$

whenever the integral on the right-hand side is finite.  $\square$

*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 a significant effort reads

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

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

$$dW_t^{\mathbb{Q}} = dW_t + \lambda_* \sigma dt,$$

and a new compensated Poisson jump measure

$$\widehat{N}^{\mathbb{Q}}(dt, dx) = \widehat{N}(dt, dx) + (1 - e^{-\lambda_*(e^x - 1)}) \Pi(dx) dt,$$

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

$$\begin{aligned} \mathcal{L}(e^{vX}) &= \int_0^\cdot \left( \alpha v + \frac{\sigma^2}{2} (v^2 - 2\lambda_*v) \right) ds \\ &\quad + \int_0^\cdot \sigma v dW_s^{\mathbb{Q}} + \int_0^\cdot \int_{\mathbb{R}} (e^{vx} - 1) \widehat{N}^{\mathbb{Q}}(dx, ds) \\ &\quad + \int_0^\cdot \int_{|x| \leq 1} \left( e^{-\lambda_*(e^x - 1)} (e^{vx} - 1) - vx \right) \Pi(dx) ds \\ &\quad + \int_0^\cdot \int_{|x| > 1} e^{-\lambda_*(e^x - 1)} (e^{vx} - 1) \Pi(dx) ds, \end{aligned} \quad (4.3)$$

provided the  $\mathbb{Q}$ -drift of  $\mathcal{L}(e^{vX})$  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  $h(x) = x\mathbf{1}_{|x|\leq 1}$  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).  $\square$

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

$$S^{(1)} = S_0^{(1)} \mathcal{E}(X^{(1)}) = S_0^{(1)} + \int_0^\cdot S_{t-}^{(1)} dX_t^{(1)}; \quad S^{(2)} = S_0^{(2)} \mathcal{E}(X^{(2)}) = S_0^{(2)} + \int_0^\cdot S_{t-}^{(2)} dX_t^{(2)}.$$

In financial economics one interprets  $\mathcal{E}(X)$  as the value of a closed fund with initial investment of \$1 following a trading strategy whose cumulative rate of return equals  $X$ . For the sake of simplicity, we assume the existence of some risk-free asset that pays zero interest rate. We furthermore assume that the Lévy measure  $\Pi$  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

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

see Margrabe (1978). Here the expectation is taken with respect to the valuation measure with the risk-free asset as numéraire, which explains why  $X$  is a martingale.

Let  $\mathbf{Q}_k$  be the valuation measure with  $S^{(k)}$  as a numéraire, that is,  $d\mathbf{Q}_k/d\mathbf{P} = S_T^{(k)}/S_0^{(k)}$  for each  $k \in \{1, 2\}$ . Then one obtains an alternative expression for the price of the Margrabe option, namely<sup>8</sup>

$$p = S_0^{(1)} \mathbf{E}^{\mathbf{Q}_1} \left[ \left( 1 - \frac{S_T^{(2)}}{S_T^{(1)}} \right)^+ \right]. \quad (4.4)$$

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

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

for certain values  $v \in \mathbb{C}$ . Let us fix such  $v$ . In the absence of default (of either asset), the computation of (4.5) is achieved by evaluating the expected rate of change of  $V = \mathbf{1}_{\{S_T^{(2)} > 0\}} (S^{(2)}/S^{(1)})^v$  under the measure  $\mathbf{Q}_1$ , in analogy to Example 4.1. This is an easy exercise in the simplified stochastic calculus: for a semimartingale  $Y$  with  $Y > 0$  and  $Y_- > 0$  one obtains

$$\frac{dY_t^v}{Y_{t-}^v} = \frac{(Y_{t-} + dY_t)^v - Y_{t-}^v}{Y_{t-}^v} = \left( 1 + \frac{dY_t}{Y_{t-}} \right)^v - 1,$$

<sup>8</sup>Note  $\mathbf{Q}^{(k)}$  is not necessarily equivalent but only absolutely continuous with respect to  $\mathbf{P}$  since the event  $\{S_T^{(k)} = 0\}$  is allowed to have positive probability under  $\mathbf{P}$ . Nonetheless, the formula (4.4) remains valid because the option payoff is zero whenever  $S_T^{(1)}$  is zero. For other applications of defaultable numéraires see, for example, Fisher et al. (2019) and references therein.

which yields

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

Composition with  $\mathcal{L}(S^{(2)}/S^{(1)}) = ((1 + \text{id}_2)/(1 + \text{id}_1) - 1) \circ (\mathcal{L}(S^{(1)}), \mathcal{L}(S^{(2)}))$  then gives

$$\mathcal{L}(V) = \left( \left( \frac{1 + \text{id}_2}{1 + \text{id}_1} \right)^v - 1 \right) \circ X. \quad (4.6)$$

Representation (4.6) together with  $\mathcal{L}(S^{(1)}) = X^{(1)} = \text{id}_1 \circ X$  yields

$$\mathcal{L}(V) + [\mathcal{L}(V), \mathcal{L}(S^{(1)})] = (1 + \text{id}_1) \left( \left( \frac{1 + \text{id}_2}{1 + \text{id}_1} \right)^v - 1 \right) \circ X. \quad (4.7)$$

Let us denote the function appearing on the right-hand side of (4.7) by  $\psi$ . In conclusion, without default (neither  $S^{(1)}$  nor  $S^{(2)}$  is allowed to hit zero) one obtains from Corollary C.3(ii) that

$$\mathbb{E}^{\mathbb{Q}_1} \left[ \mathbf{1}_{\{S_T^{(2)} > 0\}} \left( \frac{S_T^{(2)}}{S_T^{(1)}} \right)^v \right] = \left( \frac{S_0^{(2)}}{S_0^{(1)}} \right)^v \exp(b^{\psi \circ X} T). \quad (4.8)$$

In the presence of default (i.e., if either  $S^{(1)}$  or  $S^{(2)}$  may hit zero),  $V$  is no longer a  $\mathbb{P}$ -semimartingale. However,

$$S_{\uparrow}^{(1)} = S_0^{(1)} \mathcal{E}(\mathbf{1}_{\text{id}_1 \neq -1} \text{id}_1 \circ X)$$

is  $\mathbb{Q}_1$ -indistinguishable from  $S^{(1)}$  with  $S_{\uparrow}^{(1)} > 0$  and  $S_{\uparrow-}^{(1)} > 0$ ,  $\mathbb{P}$ -a.s. Therefore, the process

$$\begin{aligned} V_{\uparrow} = \mathbf{1}_{\{S^{(2)} > 0\}} \left( \frac{S^{(2)}}{S^{(1)}} \right)^v &= \left( \frac{S_0^{(2)}}{S_0^{(1)}} \right)^v \mathcal{E}(-\mathbf{1}_{\text{id}_2 = -1} \circ X) \left( \frac{\mathcal{E}(\text{id}_2 \mathbf{1}_{\text{id}_2 \neq -1} \circ X)}{\mathcal{E}(\text{id}_1 \mathbf{1}_{\text{id}_1 \neq -1} \circ X)} \right)^v \\ &= \left( \frac{S_0^{(2)}}{S_0^{(1)}} \right)^v \mathcal{E} \left( \left( \left( \frac{1 + \text{id}_2 \mathbf{1}_{\text{id}_2 \neq -1}}{1 + \text{id}_1 \mathbf{1}_{\text{id}_1 \neq -1}} \right)^v \mathbf{1}_{\text{id}_2 \neq -1} - 1 \right) \circ X \right) \end{aligned}$$

is a  $\mathbb{P}$ -semimartingale  $\mathbb{Q}_1$ -indistinguishable from  $V$ . Corollary C.3(ii) shows that (4.8) goes through with a modified jump transformation function

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

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

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

Next, formula (2.14) with  $h(x) = x$  for all  $x \in \mathbb{R}$  yields

$$\begin{aligned} b^{\psi \circ X} &= \frac{1}{2} (\sigma_1^2 - 2\sigma_{12} + \sigma_2^2) v(v-1) \\ &\quad + \int_{\mathbb{R}^2} \left( (1 + x_1) \left( \left( \frac{1 + x_2}{1 + x_1} \right)^v \mathbf{1}_{x_2 \neq -1} - 1 \right) + vx_1 - vx_2 \right) \Pi(dx_1, dx_2) \\ &= \frac{1}{2} (\sigma_1^2 - 2\sigma_{12} + \sigma_2^2) v(v-1) - \lambda_2^{\mathbb{Q}_1} + v (\lambda_2^{\mathbb{Q}_1} - \lambda_1^{\mathbb{Q}_2}) \\ &\quad + \int_{(-1, \infty)^2} \left( (1 + x_1) \left( \left( \frac{1 + x_2}{1 + x_1} \right)^v - 1 \right) + vx_1 - vx_2 \right) \Pi(dx_1, dx_2), \end{aligned}$$

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

$$\lambda_2^{\mathbb{Q}_1} = \int_{\mathbb{R}^2} (1 + x_1) \mathbf{1}_{x_2 = -1} \Pi(dx_1, dx_2)$$

signifies the arrival intensity of default of asset 2 under the probability measure  $\mathbf{Q}_1$  and  $\lambda_1^{\mathbf{Q}_2}$  has the converse meaning.<sup>9</sup> Observe that without default  $\kappa(v) = b^{\psi \circ X}$  can be interpreted as the cumulant function of  $\ln s_1^{(2)}/s_1^{(1)} - \ln s_0^{(2)}/s_0^{(1)}$  under  $\mathbf{Q}_1$ .

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  $\Pi$  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  $(e^{\text{id}_1} - 1, e^{\text{id}_2} - 1)$ . Once the integrals have been evaluated one obtains

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

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

$$(1-x)^+ = \mathbf{1}_{x=0} + \mathbf{1}_{x>0} \int_{\beta+i\mathbb{R}} g(v)x^v dv, \quad x \geq 0,$$

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

$$\begin{aligned} \frac{p}{S_0^{(1)}} &= \mathbf{Q}_1 [S_T^{(2)} = 0] + \int_{\beta+i\mathbb{R}} g(v) \mathbf{E}^{\mathbf{Q}_1} \left[ \mathbf{1}_{\{S_T^{(2)} > 0\}} \left( \frac{S_T^{(2)}}{S_T^{(1)}} \right)^v \right] dv \\ &= e^{\kappa(0)T} + \int_{\beta+i\mathbb{R}} \left( \frac{S_0^{(2)}}{S_0^{(1)}} \right)^v \frac{1}{2\pi i} \frac{1}{v(v-1)} e^{\kappa(v)T} dv, \end{aligned}$$

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

## 5. JUMPS AT PREDICTABLE TIMES

This section illustrates the simplified stochastic calculus in a discrete-time model. 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. For simplicity we assume these times can be enumerated in an increasing sequence without an accumulation point. In general, the jumps could occur at all rational times, and if we dropped the independent increments assumption, also at random predictable times. These advanced features are handled in full generality in the companion papers [Černý and Ruf \(2020a,b\)](#).

**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

<sup>9</sup>The coefficient  $\lambda_2^{\mathbf{Q}_1}$  is the drift rate of the process  $\mathbf{1}_{\text{id}_2 = -1} \circ X$  under  $\mathbf{Q}_1$ ; see [Theorem C.2\(ii\)](#).

of  $\Delta X_k$  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  $R = \mathcal{L}(e^X)$ .

To evaluate the expected utility

$$\mathbb{E} \left[ e^{-\lambda R_t} \right], \quad t \geq 0,$$

we recall from Example 3.4 the representation of the cumulative percentage change in  $e^{-\lambda R}$ . Specifically, from (3.10) one obtains  $\mathcal{L}(e^{-\lambda R}) = \eta \circ X$  with

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

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

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

for all  $t \geq 0$ . □

**Example 5.2** (Minimal entropy martingale measure). We now compute the minimal entropy martingale measure in the setting of the previous example for some given time horizon  $T > 0$ . Optimizing (5.1) over  $\lambda$ , 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  $e^{-\lambda_* R_T} / \mathbb{E}[e^{-\lambda_* R_T}]$  gives the density  $d\mathbb{Q}/d\mathbb{P}$  of the minimal entropy martingale measure  $\mathbb{Q}$ . As in Example 4.1, we seek the expected value of  $e^{vX_t}$  under  $\mathbb{Q}$ , for fixed  $t \in [0, T]$  and  $v \in \mathbb{C}$ , provided the expectation is finite. Because  $\mathcal{L}(e^{vX}) = \xi \circ X$  with

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

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

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

Here  $\mathbb{E}^{\mathbb{Q}}[e^{vX_t}]$  considered as a function of  $v \in \mathbb{C}$  gives the moment generating function of  $X_t = \ln S_t$  under  $\mathbb{Q}$  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  $\mathbb{P}$ -characteristics are given as the primitive input to the problem. The process  $X$ , which is trivially representable, is transformed by several applications of the composite rules (2.1)–(2.2) to another process  $Y$  which is also  $X$ -representable. In many situations the required end product is

the  $\mathbb{P}$ -drift of  $Y$ . These examples include i) the construction of partial integro-differential equations from martingale criteria (e.g., [Večer 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  $dX_t/X_{t-}$  to the logarithmic return  $d \ln X_t$  always takes the form  $d \ln X_t = \ln(1 + dX_t/X_{t-})$ . 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  $\mathbb{Q}$  absolutely continuous with respect to  $\mathbb{P}$ . The need to switch measures comes particularly from mathematical finance as illustrated in [Examples 4.1 and 4.3](#), 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  $d\mathbb{Q}/d\mathbb{P}$  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  $\mathbb{Q}$ -drift of  $V$  equals the  $\mathbb{P}$ -drift of  $V + [V, \mathcal{L}(M)]$ . Since it is 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 stochastic calculus less intellectually taxing than classical approaches firmly rooted in Brownian motion.



## REFERENCES

- Aït-Sahalia, Y. and F. Matthys (2019). Robust consumption and portfolio policies when asset prices can jump. *Journal of Economic Theory* 179, 1–56.
- 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 & Stochastics* 12(3), 381–410.
- Biagini, S. and A. Černý (2011). Admissible strategies in semimartingale portfolio selection. *SIAM Journal on Control and Optimization* 49(1), 42–72.
- Björk, T. (2009). *Arbitrage Theory in Continuous Time* (3rd ed.). Oxford University Press, Oxford.
- Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 163–175.
- Cai, N. and S. Kou (2012). Pricing Asian options under a hyper-exponential jump diffusion model. *Operations Research* 60(1), 64–77.
- Černý, A. and J. Ruf (2020a). Simplified calculus for semimartingales: Multiplicative compensators and changes of measure. Available from [ssrn.com/abstract=3633622](https://ssrn.com/abstract=3633622).
- Černý, A. and J. Ruf (2020b). Simplified stochastic calculus via semimartingale representations. Available from [ssrn.com/abstract=3633638](https://ssrn.com/abstract=3633638).
- Černý, A. and J. Ruf (2020c). Supplement to: Simplified stochastic calculus with applications in Economics and Finance. Available from [ssrn.com/abstract\\_id=3752072](https://ssrn.com/abstract_id=3752072).
- Duffie, D. (2001). *Dynamic Asset Pricing Theory* (3rd ed.). Princeton: Princeton University Press.
- Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine processes and applications in finance. *The Annals of Applied Probability* 13(3), 984–1053.
- Émery, M. (1978). Stabilité des solutions des équations différentielles stochastiques application aux intégrales multiplicatives stochastiques. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 41(3), 241–262.
- Feng, L. and V. Linetsky (2008). Pricing options in jump-diffusion models: An extrapolation approach. *Operations Research* 56(2), 304–325.
- Fisher, T., S. Pulido, and J. Ruf (2019). Financial models with defaultable numéraires. *Mathematical Finance* 29(1), 117–136.
- Fujiwara, T. and Y. Miyahara (2003). The minimal entropy martingale measures for geometric Lévy processes. *Finance and Stochastics* 7(4), 509–531.
- Grigoriu, M. (2002). *Stochastic Calculus: Applications in Science and Engineering*. Springer, New York.
- Hong, Y. and X. Jin (2018). Semi-analytical solutions for dynamic portfolio choice in jump-diffusion models and the optimal bond-stock mix. *European Journal of Operational Research* 265(1), 389 – 398.
- Hubalek, F., J. Kallsen, and L. Krawczyk (2006). Variance-optimal hedging for processes with stationary independent increments. *The Annals of Applied Probability* 16(2), 853–885.
- Itô, K. (1951). On a formula concerning stochastic differentials. *Nagoya Mathematics Journal* 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. *The Annals of Applied Probability* 17(5-6), 1615–1638.
- Kallsen, J. (2000). Optimal portfolios for exponential Lévy processes. *Mathematical Methods of Operations Research* 51, 357–374.
- Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus* (2nd ed.). Springer, New York.
- Larsson, M. and J. Ruf (2020). Convergence of local supermartingales. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*. to appear.
- 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 Mathematics*, 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. *Industrial Management Review* 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 Analysis* 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. *Quantitative 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  $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbf{P})$ . 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

$$\int_0^\cdot \zeta_t dX_t = \int_0^\cdot \zeta'_t dX'_t - \int_0^\cdot \zeta''_t dX''_t + i \left( \int_0^\cdot \zeta''_t dX'_t + \int_0^\cdot \zeta'_t dX''_t \right).$$

We write  $\overline{\mathbb{C}}^n = \mathbb{C}^n \cup \{\text{NaN}\}$  for some ‘non-number’  $\text{NaN} \notin \bigcup_{n \in \mathbb{N}} \mathbb{C}^n$  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  $\xi$  we shall always assume that  $\xi(\text{NaN}) = \text{NaN}$ . If  $\psi : \overline{\Omega}_{\mathbb{C}}^n \rightarrow \overline{\mathbb{C}}^m$ , 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  $\mathbb{C}$  replaced by  $\mathbb{R}$ .

Provided they exist, we write  $D\xi$  and  $D^2\xi$  for the complex derivatives of  $\xi : \overline{\Omega}_{\mathbb{C}}^d \rightarrow \overline{\mathbb{C}}^n$ , resp., the real derivatives of  $\xi : \overline{\Omega}_{\mathbb{R}}^d \rightarrow \overline{\mathbb{R}}^n$ . 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  $\mathfrak{J}_{0\mathbb{C}}^{d,n}$  denote the set of all predictable functions  $\xi : \overline{\Omega}_{\mathbb{C}}^d \rightarrow \overline{\mathbb{C}}^n$  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)  $\sup_{|x| \leq R} |D^2\xi(x)|$  is locally bounded.
- (3)  $D\xi(0)$  is locally bounded.

We write  $\mathfrak{J}_{0\mathbb{C}} = \bigcup_{k,r \in \mathbb{N}} \mathfrak{J}_{0\mathbb{C}}^{k,r}$ . The subclass  $\mathfrak{J}_{0\mathbb{R}}$  of predictable functions  $\xi : \overline{\Omega}_{\mathbb{R}}^d \rightarrow \overline{\mathbb{R}}^n$  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  $(\omega, t) \in \Omega \times [0, \infty)$ .  $\square$

Let us provide some context to the previous definition. Most of the time, we are interested in ‘real-valued’ transformations  $\xi$ , which map an  $\mathbb{R}^d$ -valued semimartingale to an  $\mathbb{R}^n$ -valued semimartingale. The core class  $\mathfrak{J}_{0\mathbb{R}}$  is perfectly suited for this purpose. The Émery formula, as stated in (1.18), works also for complex-valued  $\xi$  if we interpret  $D\xi(0)$  and  $D^2\xi(0)$  as complex derivatives. Such generalization of the theory from  $\mathbb{R}$  to  $\mathbb{C}$ , albeit limited by forcing  $\xi$  to be analytic at 0, is helpful when computing characteristic functions, for example. This leads to the definition of  $\mathfrak{J}_{0\mathbb{C}}$ , which is now a proper subclass within a larger core class  $\mathfrak{J}_0$  of complex-valued universal representing functions nesting also  $\mathfrak{J}_{0\mathbb{R}}$ .

We do not define  $\mathfrak{J}_0$  itself in this paper but it can be shown that  $\mathfrak{J}_0^{d,n}$  has a one-to-one correspondence with  $\mathfrak{J}_{0\mathbb{R}}^{2d,2n}$ . A more general Émery formula is available for  $\mathfrak{J}_0$  but will not be needed in this paper. We hence refer the interested reader to Černý and Ruf (2020b) for more details. All computations in this paper can be performed either within  $\mathfrak{J}_{0\mathbb{R}}$  or within  $\mathfrak{J}_{0\mathbb{C}}$  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 (1.18) are well-defined.

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

$$\int_0^\cdot \sum_{i,j=1}^d D_{ij}^2\xi_t(0) d[X^{(i)}, X^{(j)}]_t^c$$

*are well-defined. If, additionally,  $\xi$  is compatible with  $X$ , then (1.20) 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,

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

as the last sum has only finitely many summands.  $\square$

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

**Proposition A.3** (Representation of stochastic integrals). *Let  $\zeta$  be a locally bounded predictable  $\mathbb{R}^{n \times d}$ -valued (resp.,  $\mathbb{C}^{n \times d}$ -valued) process. Then the predictable function  $\xi = \zeta \text{id}$  belongs to  $\mathfrak{J}_{0\mathbb{R}}^{d,n}$  (resp.,  $\mathfrak{J}_{0\mathbb{C}}^{d,n}$ ) and for any  $\mathbb{R}^d$ -valued (resp.,  $\mathbb{C}^d$ -valued) semimartingale  $X$  one has*

$$\int_0^\cdot \zeta_t dX_t = (\zeta \text{id}) \circ X. \quad (\text{A.1})$$

*Proof.* We start with the complex-valued case. As  $D\xi = \zeta$  and  $D^2\xi = 0$ , we have that  $\xi$  belongs to  $\mathfrak{J}_{0\mathbb{C}}$  and is compatible with any  $\mathbb{C}^d$ -valued semimartingale  $X$ . The Émery formula (1.18) now yields (A.1). The real-valued proof proceeds analogously.  $\square$

**Proposition A.4** (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  $f : \mathcal{U} \rightarrow \mathbb{R}^n$  be a twice continuously differentiable function (resp., let  $f : \mathcal{U} \rightarrow \mathbb{C}^n$  be an analytic function), and let*

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

*Then  $\xi^{f,X} \in \mathfrak{J}_{0\mathbb{R}}^{d,n}$  (resp.,  $\xi^{f,X} \in \mathfrak{J}_{0\mathbb{C}}^{d,n}$ ) is compatible with  $X$  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  $D^2\xi^{f,X}(0) = D^2f(X_-)$ . As both  $Df(X_-)$  and  $D^2f(X_-)$  are finite-valued predictable processes, they are locally bounded by Larsson and Ruf (2020, 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

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

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

For  $\xi \in \mathfrak{J}_{0\mathbb{R}}$ , Lemma A.2 and the Émery formula (1.18) 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  $\xi \in \mathfrak{J}_{0\mathbb{C}}$  the result follows by identifying  $\mathbb{C}^d$  with  $\mathbb{R}^{2d}$  and using the real-valued statement we have just proved.  $\square$

**Proposition A.5** (Composition of universal representing functions). *The space  $\mathfrak{J}_{0\mathbb{R}}$  is closed under dimensionally correct composition, i.e., if  $\xi \in \mathfrak{J}_{0\mathbb{R}}^{d,n}$  and  $\psi \in \mathfrak{J}_{0\mathbb{R}}^{n,m}$  then  $\psi \circ \xi \in \mathfrak{J}_{0\mathbb{R}}^{d,m}$ . An analogous statement holds for  $\mathfrak{J}_{0\mathbb{C}}$ .*

*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, for all  $|x| \leq \delta_{\psi \circ \xi}$  we now have

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

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

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

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

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  $\sum_{n \in \mathbb{N}} \delta_{\psi \circ \xi}^{(n)} \mathbf{1}_{[[\tau_{n-1}, \tau_n]]}$ .  $\square$

## 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  $h$  is a truncation function for a semimartingale  $X$  if  $h$  is time-constant and deterministic, compatible with  $X$ ,  $\sum_{0 < t \leq \cdot} |\Delta X_t - h(\Delta X_t)| < \infty$ , and if

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

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.  $\square$

**Proposition B.2** (Universal truncation functions). *If a bounded time-constant deterministic function  $h$  equals identity on an open neighbourhood of 0 then it is a truncation function for any compatible semimartingale  $X$ . Furthermore, one then has  $h \in \mathfrak{I}_{0\mathbb{R}} \cup \mathfrak{I}_{0\mathbb{C}}$  and*

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

*Proof.* Clearly  $h \in \mathfrak{I}_{0\mathbb{R}} \cup \mathfrak{I}_{0\mathbb{C}}$ . Lemma A.2 with  $\xi = h$  now yields  $\sum_{0 < t \leq \cdot} |h(\Delta X_t) - \Delta X_t| < \infty$ . Next, observe that  $\Delta X[h]_t = h(\Delta X_t)$  is bounded, therefore  $X[h]$  is special by Jacod and Shiryaev (2003, I.4.24). Finally, the Émery formula (1.18) yields

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

which completes the proof.  $\square$

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

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

and

$$\begin{aligned} \xi \circ X &= \int_0^\cdot D\xi_t(0)dX[h]_t + \frac{1}{2} \int_0^\cdot \sum_{i,j=1}^d D_{ij}^2 \xi_t(0) d[X^{(i)}, X^{(j)}]_t^c \\ &\quad + \sum_{0 < t \leq \cdot} (\xi_t(\Delta X_t) - D\xi_t(0)h(\Delta X_t)). \end{aligned} \tag{B.1}$$

*Proof.* First, the triangle inequality gives

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

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

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

and the Émery formula (1.18) now yield the second part of the claim.  $\square$

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  $\mathbb{P}$ -martingale. If  $\mathbb{Q}$  is another probability measure absolutely continuous with respect to  $\mathbb{P}$  and  $X$  is  $\mathbb{Q}$ -special, we denote the corresponding  $\mathbb{Q}$ -compensator by  $B_{\mathbb{Q}}^X$ . Finally, we denote by  $\nu^X$  the predictable  $\mathbb{P}$ -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 semimartingale is PII if it has independent increments. The following result for PII semimartingales relates drifts to expected values, and hence shall be very useful. It is proved in Černý and Ruf (2020a, Proposition 2.14 and Theorem 4.1). At this point, we remind the reader that the stochastic exponential  $\mathcal{E}(X)$  of a one-dimensional semimartingale  $X$  is given as the (unique) solution of the stochastic differential equation

$$\mathcal{E}(X) = 1 + \int_0^\cdot \mathcal{E}(X)_{t-} dX_t. \tag{B.2}$$

**Theorem B.4.** *Assume  $\xi \in \mathfrak{I}_{0\mathbb{R}} \cup \mathfrak{I}_{0\mathbb{C}}$  is compatible with  $X$ . If  $\xi$  is deterministic and  $X$  is PII, then  $\xi \circ X$ , too, is PII. Furthermore, if  $\xi \circ X$  is special one has*

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

$$\mathbb{E}[\mathcal{E}(\xi \circ X)_t] = \mathcal{E}(B^{\xi \circ X})_t, \quad t \geq 0. \tag{B.4}$$

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  $B^{X[h]} = \int_0^\cdot b^{X[h]} dt$ ,  $[X, X]^c = \int_0^\cdot c^X dt$ , and  $\nu$  can be written in disintegrated form as  $\nu = \int_0^\cdot \int F^X(dx) dt$ .  $\square$

**Theorem B.6.** *Let  $X$  be a semimartingale and let  $h$  be a truncation function for  $X$ . Assume that  $\xi \in \mathfrak{I}_{0\mathbb{R}} \cup \mathfrak{I}_{0\mathbb{C}}$  is compatible with  $X$  and that  $\xi \circ X$  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*

$$\begin{aligned} B_t^{\xi \circ X} &= \sum_{k=1}^{\lfloor t \rfloor} \mathbf{E}_{k-}[\xi_k(\Delta X_k)], & t \geq 0; \\ \mathcal{E}(B^{\xi \circ X})_t &= \prod_{k=1}^{\lfloor t \rfloor} \mathbf{E}_{k-}[1 + \xi_k(\Delta X_k)], & t \geq 0. \end{aligned} \quad (\text{B.5})$$

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

$$\begin{aligned} b^{\xi \circ X} &= D\xi(0)b^{X[h]} + \frac{1}{2} \sum_{i,j=1}^d D_{ij}^2 \xi(0) c_{ij}^X + \int_{\mathbb{R}^d} (\xi(x) - D\xi(0)h(x)) F^X(dx); \\ B^{\xi \circ X} &= \int_0^\cdot b_t^{\xi \circ X} dt; & \mathcal{E}(B^{\xi \circ X}) &= \exp\left(\int_0^\cdot b_t^{\xi \circ X} dt\right). \end{aligned}$$

*Proof.* By (B.1), we have

$$\begin{aligned} B^{\xi \circ X} &= \int_0^\cdot D\xi_t(0) dB_t^{X[h]} + \frac{1}{2} \int_0^\cdot \sum_{i,j=1}^d D_{ij}^2 \xi_t(0) d[X^{(i)}, X^{(j)}]_t^c \\ &\quad + \int_0^\cdot \int_{\mathbb{R}^d} (\xi_t(x) - D\xi_t(0)h(x)) \nu^X(dt, dx), \end{aligned}$$

yielding the statement.  $\square$

## 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  $\mathbb{P}$ -semimartingale such that*

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

*is a uniformly integrable  $\mathbb{P}$ -martingale with  $M \geq 0$ . Define the probability measure  $\mathbb{Q}$  by*

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

*For a  $\mathbb{Q}$ -semimartingale  $V$  and a  $\mathbb{P}$ -semimartingale  $V_\uparrow$ ,  $\mathbb{Q}$ -indistinguishable from  $V$ , the following are equivalent.*

- (1)  $V$  is  $\mathbb{Q}$ -special.
- (2)  $V_\uparrow + [V_\uparrow, N]$  is  $\mathbb{P}$ -special.

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

$$B_{\mathbb{Q}}^V = B^{V_\uparrow + [V_\uparrow, N]}, \quad \mathbb{Q}\text{-almost surely.}$$

The proof is quite classical and we do not reproduce it here. For details see Černý and Ruf (2020a, Proposition 5.2).

**Theorem C.2** (Girsanov's theorem – representations). *Assume  $\eta, \xi \in \mathfrak{I}_{0\mathbb{R}}$  (resp.,  $\mathfrak{I}_{0\mathbb{C}}$ ) with  $\eta \geq -1$  are compatible with a semimartingale  $X$  such that  $\eta \circ X$  is special and  $\Delta B^{\eta \circ X} > -1$ .*

Assume further that  $M = \mathcal{E}(\eta \circ X) / \mathcal{E}(B^{\eta \circ X})$  is a real-valued uniformly integrable  $\mathbb{P}$ -martingale and define the probability measure  $\mathbb{Q}$  by

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

Assume also that  $\xi \circ X$  is  $\mathbb{Q}$ -special. Then the following statements hold.

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

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

- (ii) If  $X$  is an Itô  $\mathbb{P}$ -semimartingale then  $\xi \circ X$  is an Itô  $\mathbb{Q}$ -semimartingale and

$$\begin{aligned} b_{\mathbb{Q}}^{\xi \circ X} &= b^{(1+\eta)\xi \circ X} = D\xi(0)b^{X[h]} + \frac{1}{2} \sum_{i,j=1}^d (D_{ij}^2 \xi(0) + 2D_i \xi(0)D_j \eta(0)) c_{ij}^X \\ &\quad + \int_{\mathbb{R}^d} (\xi(x)(1 + \eta(x)) - D\xi(0)h(x)) F^X(dx). \end{aligned}$$

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

$$N = \left( \frac{1 + \eta(\text{id}_1)}{1 + \text{id}_2} - 1 \right) \circ (X, B^{\eta \circ X}).$$

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

$$V + [V, N] = \xi(\text{id}_1) \frac{1 + \eta(\text{id}_1)}{1 + \text{id}_2} \circ (X, B^{\eta \circ X}) = \xi \frac{1 + \eta}{1 + \Delta B^{\eta \circ X}} \circ X.$$

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

**Corollary C.3.** *With the notation and assumptions as in Theorem C.2 above, if  $X$  is PII under  $\mathbb{P}$  stopped at a finite time and  $\eta$  is deterministic then  $M = \mathcal{E}(\eta \circ X) / \mathcal{E}(B^{\eta \circ X})$  is a uniformly integrable martingale. Furthermore, if  $\xi$ , too, is deterministic then  $\xi \circ X$  is PII under  $\mathbb{Q}$  and the following statements hold for all  $t \geq 0$ .*

- (i) If  $X$  is a discrete-time process under  $\mathbb{P}$  then

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

- (ii) If  $X$  is an Itô  $\mathbb{P}$ -semimartingale then

$$\mathbb{E}^{\mathbb{Q}}[\mathcal{E}(\xi \circ X)_t] = \exp \left( \int_0^t b_u^{(1+\eta)\xi \circ X} du \right).$$

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