Risk-Averse Optimal Control

Alois Pichler* Ruben Schlotter†

July 3, 2020

Abstract

This paper extends dynamic control problems from a risk-neutral to a risk-averse setting. By consistently nesting coherent risk measures it is possible to establish the limit of risk-averse multistage optimal control problems in continuous time. For the limiting case we elaborate a new dynamic programming principle and give risk-averse Hamilton–Jacobi–Bellman equations by generalizing the infinitesimal generator. In doing so we provide a constructive explanation of the driver "g" in g-expectation, a dynamic risk measure based on backwards stochastic differential equations.

Moreover we demonstrate that the Entropic Value-at-Risk is the natural and universal candidate for a coherent risk measure in the context of optimal control.

Keywords: Risk measures, Optimal control, Stochastic processes

Classification: 90C15, 60B05, 62P05

1 Introduction

This paper discusses dynamic control problems in a risk-averse environment. For this purpose it is *crucial* to compose (or nest) risk measures, as exactly nested risk measures allow quantifying risk via dynamic programming equations which appear eminently in optimal control problems.

More specifically, this paper formulates risk-averse optimal control problems by extending the discrete time setting to a continuous time framework. Here, a decision maker incurs an uncertain stream of costs $c(\cdot)$ over time and his goal is to manage and minimize the accumulated costs incurred. The risk-averse decision maker intends to guard against undesired scenarios in particular. The nested risk measure $\rho_{0:T}$ governs the risk over the finite time horizon $[0,T]$ and the optimization problem in consideration thus is

$$\inf_{u(\cdot)\in\mathcal{U}} \rho_{0:T}\left(\int_0^T c\big(t,u(t)\big)\,\mathrm{d}t\right),$$

where $\mathcal{U}$ collects feasible control policies. We will see that the nested risk measure $\rho_{0:T}$ allows to guard against risk in every instant of time, which is the typical objective of a risk manager permanently hedging against risk. Non-nested risk measures, in contrast, only assess the accumulated position at terminal time T, giving no chance to the risk manager to intervene.

Nested risk measures were introduced in Ruszczyński and Shapiro (2006). Pichler and Shapiro elaborate that exactly these risk measures allow expanding the associated control problem to a risk-averse framework. New dynamic programming equations, derived below, then reflect the risk-averse character. For a discussion on risk measures and dynamic optimization we refer to De Lara and Leclère (2016). Applications can be found in Philpott and de Matos (2012), Maggioni et al. (2012) or Guigues and Römisch (2012), e.g., where stochastic dual dynamic programming methods are addressed.

*Funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 416228727 – SFB 1410.

†Both authors: Technische Universität Chemnitz, 09126 Chemnitz, Germany. Contact: ruben.schlotter@math.tu-chemnitz.de

Çavuş and Ruszczyński (2014) and Fan and Ruszczyński (2014) study *risk-averse* control problems in discrete time. Pflug and Pichler (2016) introduce extended conditional risk measures, which generate dynamically consistent risk measures in discrete time. Extending these ideas, Dentcheva and Ruszczyński (2018) analyze the value functions in a continuous time Markov chain model and derive dynamic programming equations. For dynamic equations in the context of Markov decision processes see Haskell and Jain (2015) and Haskell et al. (2016). Additionally, Ruszczyński and Yao (2015) discuss a risk-averse optimal control problem in continuous time using g-expectations and derive an Hamilton–Jacobi–Bellman equation.

g-expectations are dynamic risk measures based on backward stochastic differential equations introduced in Pardoux and Peng (1990). The relationship between risk measures and g-expectations was extensively studied in Coquet et al. (2002) and Peng (2004) among many others (see also the reference therein or Rosazza Gianin (2006)).

This paper, in contrast, derives the driver $g(\cdot)$ of g-expectations explicitly by nesting classical risk measures. Our approach thus provides a novel and elementary understanding of the risk-averse evolution equations and g-expectations based on nested risk measures. Moreover, our approach elaborates a clear relation between static risk measures, dynamic risk measures in discrete time and the limiting risk measure in continuous time.

Within this study, a special focus is given to nested risk measures based on the Entropic Value-at-Risk defined by

$$\mathrm{EV@R}_\beta(Y) := \sup\left\{\mathbb{E}\, YZ\colon Z \geq 0,\ \mathbb{E}\, Z = 1 \text{ and } \mathbb{E}\, Z\log Z \leq \beta\right\}, \tag{1}$$

where $\beta \geq 0$ is the coefficient of risk aversion (or risk level) and $\mathbb{E}\, Z\log Z$ is the Shannon entropy (see also Breuer and Csiszár (2013) and Pichler and Schlotter (2019) for a discussion of entropy and risk measures). We give explicit formulas for the nested Entropic Value-at-Risk of the Wiener process and investigate more general diffusion processes. Other risk measures can be nested as well, but their composite counterpart is often degenerate. It is of independent interest that the Average Value-at-Risk (also known as Conditional Value-at-Risk), the most prominent risk measure, degenerates when nested naively.

The infinitesimal generator is the essential tool for classical stochastic optimal control in continuous time, see Fleming and Soner (2006), e.g. In this classical, risk-neutral setting, the infinitesimal generator of an Itô process is a linear second order differential operator. We introduce the risk-averse analogue, called risk generator and derive explicit expressions for Itô processes. Here, the risk generator is a second order operator, which is *nonlinear* in the first derivative. While g-expectations are nonlinear expectations (cf. Peng (2004)) sui generis, we *derive* risk aversion from the nonlinear generator. We further derive risk-averse Hamilton–Jacobi–Bellman equations, which extend the classical, risk-neutral dynamic programming equations to the risk-averse setting.

Based on the derivation of the risk generator we conclude that the Entropic Value-at-Risk is the natural choice for a nested risk measure in continuous time by adapting the risk levels in a natural way. Furthermore, our approach provides an interpretation of the Hamilton–Jacobi–Bellman equation given in Ruszczyński and Yao (2015).

Outline of the paper. Section 2 introduces conditional and nested risk measures and provides the general mathematical setup in a discrete time setting. Section 3 deals with the extension to continuous time. Section 4 provides risk evaluations for Itô processes and introduces the risk generator of a stochastic process. We close with a discussion on other risk measures besides the Entropic Value-at-Risk. Section 5 introduces the risk-averse stochastic control problem and derives the corresponding risk-averse Hamilton–Jacobi–Bellman partial differential equation.

2 Notation and preliminaries

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{T}}, P)$ and associate $t \in \mathcal{T}$ with *stage* or *time*. The entropic space is

$$E := \left\{ Y : \Omega \to \mathbb{R} \,\middle|\, \mathbb{E}\, e^{\ell |Y|} < \infty \text{ for all } \ell \in \mathbb{R} \right\}$$

and we shall assume throughout the paper that the process $X = (X_t)_{t\in\mathcal{T}}$ is adapted to the filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$ with marginals $X_t \in E$ for all $t \in \mathcal{T}$. We shall write $X_t \lhd \mathcal{F}_t$ to indicate that X_t is measurable with respect to $\mathcal{F}_t$.

We further assume that our filtered probability space is Polish. As this paper develops a theory for risk aversion of Itô processes based on Brownian motion we may work on the classical Wiener space without loss of generality.

2.1 Conditional risk measures and EV@R

We recall the definition of *law invariant, coherent risk measures* $\rho: L \to \mathbb{R}$ defined on some vector space L of $\mathbb{R}$-valued random variables first. They satisfy the following axioms introduced by Artzner et al. (1999):

A1. Monotonicity: $\rho(Y) \le \rho(Y')$, provided that $Y \le Y'$ almost surely;
A2. Translation equivariance: $\rho(Y + c) = \rho(Y) + c$ for $c \in \mathbb{R}$;
A3. Convexity: $\rho\big((1-\lambda)\, Y + \lambda\, Y'\big) \le (1-\lambda)\, \rho(Y) + \lambda\, \rho(Y')$ for $\lambda \in [0,1]$;
A4. Positive homogeneity: $\rho(\lambda\, Y) = \lambda\, \rho(Y)$ for $\lambda \ge 0$;

A5. Law invariance: $\rho(Y) = \rho(Y')$, whenever Y and Y' have the same law, i.e., $P(Y \le y) = P(Y' \le y)$ for all $y \in \mathbb{R}$.

Remark 1. Any functional $\rho: L \to \mathbb{R}$ satisfying the Axioms A1–A4 can be represented by

$$\rho(Y) = \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q\, Y$$

for a convex set of probability measures $\mathcal{Q}$ absolutely continuous with respect to P (cf. Delbaen (2002)). We consider the conditional risk measures ρ^t with respect to the sigma algebra $\mathcal{F}_t$ defined by

$$\rho^t\, (Y\,|\mathcal{F}_t\,) := \operatorname*{ess\,sup}_{Q\in\mathcal{Q}} \mathbb{E}_Q\, [Y\,|\,\mathcal{F}_t\,]\,. \tag{2}$$

Note that ρ^t satisfies conditional versions of the Axioms A1–A5. For further details, we refer the interested reader to Ruszczyński and Shapiro (2006) and Riedel (2004). For the essential supremum of a set of random variables as in (2) we refer to Karatzas and Shreve (1998, Appendix A). We remark that $\rho^t\, (Y\,|\mathcal{F}_t\,) \lhd \mathcal{F}_t$ and the axioms A2 and A4 extend to $\mathcal{F}_t$-measurable random variables.

Definition 2 (Entropic Value-at-Risk)**.** The *Entropic Value-at-Risk* of a random variable $Y \in E$ at risk level $\beta \ge 0$ is (cf. (1))

$$\text{EV@R}_\beta(Y) = \sup\left\{\mathbb{E}\, YZ : Z \ge 0,\ \mathbb{E}\, Z = 1,\ \mathbb{E}\, Z \log Z \le \beta\right\}. \tag{3}$$

Similarly, for the risk level $0 \le \beta \lhd \mathcal{F}_s$, we define the *conditional Entropic Value-at-Risk* $\text{EV@R}_\beta(\cdot \mid \mathcal{F}_s)$ as

$$\text{EV@R}_\beta\, (Y\,|\,\mathcal{F}_s\,) := \operatorname{ess\,sup}\left\{\mathbb{E}\,[YZ\,|\,\mathcal{F}_s\,] : 0 \le Z,\ \mathbb{E}\,[Z\,|\,\mathcal{F}_s\,] = 1,\ \mathbb{E}\,[Z \log Z\,|\,\mathcal{F}_s\,] \le \beta\right\}. \tag{4}$$

Remark 3. For a random variable $Y \lhd \mathcal{F}_t$, the density Z in the defining equation (4) may be chosen to satisfy $Z \lhd \mathcal{F}_t$. Indeed, the density $Z_t := \mathbb{E}\left[Z \mid \mathcal{F}_t\right]$ satisfies all constraints as well, as can be seen by applying the conditional Jensen inequality

$$\mathbb{E}\left[Z_t \log Z_t \mid \mathcal{F}_s\right] \leq \mathbb{E}\left[\mathbb{E}\left[Z \log Z \mid \mathcal{F}_t\right] \mid \mathcal{F}_s\right] = \mathbb{E}\left[Z \log Z \mid \mathcal{F}_s\right].$$

The tower property of the expectation finally insures the assertion.

For future reference we provide a closed form for the Entropic Value-at-Risk for Gaussian random variables.

Proposition 4 (EV@R of Gaussians)**.** *For a normally distributed random variable $Y \sim \mathcal{N}(\mu, \sigma^2)$ with $\sigma \geq 0$ and $a, b \in \mathbb{R}$ it holds that*

$$\mathrm{EV@R}_\beta(a + b\,Y) = a + b\,\mu + \sigma\ |b|\ \sqrt{2\beta}.$$

Proof. Let $Y \sim \mathcal{N}(\mu, \sigma^2)$, $\beta \in [0, \infty)$ and consider the alternative representation of the Entropic Value-at-Risk (cf. Ahmadi-Javid (2012))

$$\mathrm{EV@R}_\beta(Y) = \inf_{\ell > 0} \frac{1}{\ell}\left(\beta + \log \mathbb{E}\, e^{\ell Y}\right). \tag{5}$$

It holds that $\mathbb{E}\, e^{\ell Y} = \exp\left(\mu\ell + \frac{1}{2}\ell^2\sigma^2\right)$ and thus $\frac{1}{\ell}\beta + \frac{1}{\ell}\log\left(e^{\mu\ell + \frac{1}{2}\ell^2\sigma^2}\right) = \frac{1}{\ell}\beta + \mu + \frac{1}{2}\ell\sigma^2$,which attains its infimum at $\ell^* = \frac{1}{\sigma}\sqrt{2\beta}$. The Entropic Value-at-Risk thus is

$$\mathrm{EV@R}_\beta(Y) = \frac{1}{\ell^*}\beta + \mu + \frac{1}{2}\ell^*\sigma^2 = \mu + \sigma\sqrt{2\beta}.$$

Finally notice that $a + b\,Y \sim \mathcal{N}(a + b\,\mu, b^2\,\sigma^2)$ and hence

$$\mathrm{EV@R}_\beta(a + b\,Y) = a + b\,\mu + |b|\ \sigma\ \sqrt{2\beta},$$

the assertion. □

2.2 Nested risk measures

Nested risk measures are compositions of conditional risk measures. This section elaborates general properties of nested risk measures. The results are then discussed in more detail for the Entropic Value-at-Risk. Throughout, we will always consider risk on the time interval $[0, T]$, $T > 0$, and denote by $\mathcal{P} := \{0 = t_0 < t_1 < \cdots < t_n = T\}$ a finite partition of the interval $[0,\, T]$ including its endpoints. With $\Delta t_i := t_{i+1} - t_i$ we denote the time step and $\|\mathcal{P}\| := \max_{0 \leq i \leq n-1} \Delta t_i$ is the mesh size of the partition $\mathcal{P}$.

Definition 5 (Nested risk measures)**.** Let $\mathcal{P}$ be a partition of the interval $[0, T]$ and let $Y \lhd \mathcal{F}_T$. For a collection of conditional risk measures $(\rho^t)_{t \in \mathcal{P}}$ and $i < n$, the nested risk measure is

$$\rho^{t_i:t_n}\left(Y \middle| \mathcal{F}_{t_i}\right) := \rho^{t_i}\left(\rho^{t_{i+1}} \ldots \left(\rho^{t_{n-1}}\left(Y \middle| \mathcal{F}_{t_{n-1}}\right) \ldots \middle| \mathcal{F}_{t_{i+1}}\right) \middle| \mathcal{F}_{t_i}\right). \tag{6}$$

Remark 6 (Risk martingales). Nested risk measures naturally follow a martingale like pattern. Indeed, the stochastic process $Y_t := \rho^{t:t_n}(Y \mid \mathcal{F}_t)$ satisfies

$$\begin{aligned}\rho^{t_i:t_n}\left(Y_{t_{i+1}} \middle| \mathcal{F}_{t_i}\right) &= \rho^{t_i}\left(\rho^{t_{i+1}:t_n}\left(Y \middle| \mathcal{F}_{t_{i+1}}\right) \middle| \mathcal{F}_{t_i}\right)\\ &= \rho^{t_i}\left(\rho^{t_{i+1}} \cdots \left(\rho^{t_{n-1}}\left(Y \middle| \mathcal{F}_{t_{n-1}}\right) \cdots \middle| \mathcal{F}_{t_{i+1}}\right) \middle| \mathcal{F}_{t_i}\right)\\ &= Y_{t_i}.\end{aligned}$$

We call this process $(Y_t)_{t \in \mathcal{P}}$ a *risk martingale* with respect to the family of risk measures $(\rho^t)_{t \in \mathcal{P}}$.

Often, the risk evaluation $\rho^{0:T}(X_T)$ of the terminal value X_T of some stochastic process X is of interest. The terminal value X_T can then be represented as the sum of its increments $\Delta X_{t_j} := X_{t_{j+1}} - X_{t_j}$ as

$$X_T = X_{t_i} + \sum_{j=i}^{n-1} \Delta X_{t_j}.$$

As a consequence of translation equivariance we have the following useful proposition.

Proposition 7. *Suppose that $(X_t)_{t\in\mathcal{P}}$ is a discrete time stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\in\mathcal{P}}$. The nested risk measure (6) is*

$$\rho^{t_i:T}\left(X_T \,\middle|\, \mathcal{F}_{t_i}\right) = X_{t_i} + \rho^{t_i}\left(\Delta X_{t_i} + \rho^{t_{i+1}}\left(\ldots \Delta X_{t_{n-2}} + \rho^{t_{n-1}}\left(\Delta X_{t_{n-1}} \,\middle|\, \mathcal{F}_{t_{n-1}}\right) \ldots \,\middle|\, \mathcal{F}_{t_{i+1}}\right) \,\middle|\, \mathcal{F}_{t_i}\right). \tag{7}$$

It is therefore sufficient to study conditional risk evaluations of increments. The next section exploits this observation by giving explicit formulas in important cases.

2.3 The nested Entropic Value-at-Risk

In what follows we consider the nested Entropic Value-at-Risk, but we adjust the risk level to span the respective time interval. For convenience of the reader and future reference we reemphasize and state these details in the following definition.

Definition 8 (Nested Entropic Value-at-Risk)**.** Let $\mathcal{P}$ be a partition of $[0,T]$ and $Y \lhd \mathcal{F}_T$. For a vector of risk levels $\beta := \left(\beta_{t_i}\,\Delta t_i, \ldots, \beta_{t_{n-1}}\,\Delta t_{n-1}\right)$, the nested Entropic Value-at-Risk is

$$\mathrm{nEV@R}_\beta^{t_i:t_n}\left(Y \,\middle|\, \mathcal{F}_{t_i}\right) := \mathrm{EV@R}_{\beta_{t_i}\cdot\Delta t_i}\left(\ldots \mathrm{EV@R}_{\beta_{t_{n-1}}\cdot\Delta t_{n-1}}\left(Y \,\middle|\, \mathcal{F}_{t_{n-1}}\right) \ldots \,\middle|\, \mathcal{F}_{t_i}\right). \tag{8}$$

To emphasize the dependence on the partition we will also write $\mathrm{nEV@R}_\beta^{\mathcal{P}}(Y \mid \mathcal{F}_{t_i})$ for (8). Furthermore, for the trivial sigma algebra $\mathcal{F}_0 = \{\emptyset, \Omega\}$, we simply write $\mathrm{nEV@R}_\beta^{0:T}(Y)$.

As a corollary to Proposition 4 we provide an explicit formula for the nEV@R for a Wiener process evaluated at discrete time points.

Proposition 9 (Nested EV@R for the Gaussian random walk)**.** *Let $W = (W_t)_{t\in\mathcal{P}}$ be a Wiener process evaluated on the partition $\mathcal{P}$. Furthermore, let $\beta := \left(\beta_{t_0}\,\Delta t_0, \ldots, \beta_{t_{n-1}}\,\Delta t_{n-1}\right)$ be a vector of risk levels. Then the nested Entropic Value-at-Risk is*

$$\mathrm{nEV@R}_\beta^{\mathcal{P}}(W_T) = \sum_{i=0}^{n-1} \Delta t_i \sqrt{2\beta_{t_i}}. \tag{9}$$

Proof. Note that the increments of W satisfy $W_{t_{i+1}} - W_{t_i} \sim \mathcal{N}(0, t_{i+1} - t_i)$ and, by Proposition 4, the conditional Entropic Value-at-Risk is

$$\mathrm{EV@R}_{\beta_{t_i}\,\Delta t_i}\left(W_{t_{i+1}} \,\middle|\, W_{t_i}\right) = W_{t_i} + \sqrt{\Delta t_i}\,\sqrt{2\beta_{t_i}\Delta t_i}.$$

Iterating as in Equation (7) shows

$$\mathrm{nEV@R}_\beta^{\mathcal{P}}(W_T) = \sum_{i=0}^{n-1} \sqrt{\Delta t_i}\,\sqrt{2\beta_{t_i}\Delta t_i} = \sum_{i=0}^{n-1} \Delta t_i \sqrt{2\beta_{t_i}},$$

the assertion. □

Remark 10 (Parametrization). Comparing the explicit formula (9) with Proposition 4 we observe the surprising consistency property

$$\mathrm{nEV@R}_{\beta}^{\mathcal{P}}(W_T) = \mathrm{EV@R}_{\beta_0\cdot T}(W_T)$$

for the risk levels $\beta = (\beta_0\,\Delta t_0, \dots, \beta_0\,\Delta t_{n-1})$ with constant risk level β_0. This is a consequence of the parametrization chosen in (3), the definition of the Entropic Value-at-Risk.

3 Quantification of risk in continuous time

The previous section considers nested risk measures in discrete time on partitions $\mathcal{P} = \{0 = t_0 < \dots < t_n = T\}$ of the interval $\mathcal{T} \coloneqq [0, T]$. In what follows, we refine the partitions to obtain a limit of the nested risk measures in continuous time.

To this end, we first extend the vector of risk levels to continuous time by involving a piecewise constant function β called the *risk rate* and extend the definition of the nested Entropic Value-at-Risk to risk rates $\beta(\cdot)$.

Definition 11. Let $\beta\colon \mathcal{T} \to [0, \infty)$ be the simple function

$$\beta(t) = \sum_{i=0}^{n-1} \beta_{t_i}\, \mathbb{1}_{[t_i, t_{i+1})}(t).$$

Then the *nested Entropic Value-at-Risk* is defined as

$$\mathrm{nEV@R}_{\beta(\cdot)}(Y) \coloneqq \mathrm{nEV@R}_{\widehat{\beta}}^{\mathcal{P}}(Y),$$

where the vector of risk levels is $\widehat{\beta} \coloneqq (\beta(t_0)\Delta t_0, \dots, \beta(t_{n-1})\Delta t_{n-1})$.

Given a risk rate $\beta(\cdot)$ we now investigate the relationship of nEV@R for different simple functions. We assume throughout that the risk rate $\beta(\cdot)$ is Riemann integrable.

Theorem 12 (Closed under pairwise minimization)**.** *Let $Y \lhd \mathcal{F}_T$ be a random variable and let $\beta,\ \beta'\colon \mathcal{T} \to [0, \infty)$ be piecewise constant risk rates. Then the nested entropic Value-at-Risk is closed under pairwise minimization, i.e.,*

$$\mathrm{nEV@R}_{\min\left(\beta(\cdot), \beta'(\cdot)\right)}(Y) \le \min\left\{\mathrm{nEV@R}_{\beta(\cdot)}(Y),\ \mathrm{nEV@R}_{\beta'(\cdot)}(Y)\right\}$$

holds true.

Proof. It is enough to consider a constant risk rate $\beta'(\cdot) = \beta_1$ and a piecewise constant risk rate

$$\beta(t) = \beta_0\, \mathbb{1}_{[t_0, t_1)}(t) + \beta_1\, \mathbb{1}_{[t_1, t_2)}(t)$$

with $\beta_0 \le \beta_1$ and hence $\min\{\beta(\cdot), \beta'(\cdot)\} = \beta(\cdot)$. Consider the infimum representation as in Proposition 4,

$$\mathrm{EV@R}_{\beta_1\,\Delta t_1}\left(Y \,\middle|\, \mathcal{F}_{t_1}\right) = \inf_{\ell} \frac{1}{\ell}\left(\beta_1 \Delta t_1 + \log \mathbb{E}\left[e^{\ell Y} \,\middle|\, \mathcal{F}_{t_1}\right]\right),$$

of the conditional Entropic Value-at-Risk. By nesting we obtain

$$\mathrm{nEV@R}_{\beta(\cdot)}(Y) = \inf_{x} \frac{1}{x}\left(\beta_0 \Delta t_0 + \log \mathbb{E}\left[\exp\left(x\left(\inf_{\ell} \frac{1}{\ell}\left(\beta_1 \Delta t_1 + \log \mathbb{E}\left[e^{\ell Y} \,\middle|\, \mathcal{F}_{t_1}\right]\right)\right)\right) \,\middle|\, \mathcal{F}_{t_0}\right]\right).$$

Choosing $\ell = x$ gives the upper bound

$$\text{nEV@R}_{\beta(\cdot)}(Y) \le \inf_x \frac{1}{x}\left(\beta_0\Delta t_0 + \beta_1\Delta t_1 + \log\mathbb{E}\left(\mathbb{E}\left[e^{xY}\,\middle|\,\mathcal{F}_{t_1}\right]\,\middle|\,\mathcal{F}_{t_0}\right)\right) = \text{EV@R}_{\beta_0 T}(Y).$$

Because $\beta'(\cdot) \ge \beta(\cdot)$ is constant, it follows that

$$\text{nEV@R}_{\beta(\cdot)}(Y) \le \text{EV@R}_{\beta_0 T}(Y) = \text{nEV@R}_{\beta'(\cdot)}(Y).$$

The general case follows by induction using the monotonicity property A1 of conditional risk measures. □

We are now ready to extend the nested Entropic Value-at-Risk to continuous time and demonstrate that the extension is well-defined.

Definition 13 (Nested Entropic Value-at-Risk in continuous time)**.** Let $T > 0$, $t \in [0, T)$ and $Y \in E$. The *nested Entropic Value-at-Risk in continuous time* for the Riemann integrable risk rate $\beta\colon [t, T] \to [0, \infty)$ is the upper envelope

$$\text{nEV@R}^{t:T}_{\beta(\cdot)}(Y \mid \mathcal{F}_t) \coloneqq \operatorname*{ess\,inf}_{\widetilde{\beta}(\cdot)\ge\beta(\cdot)} \text{nEV@R}_{\widetilde{\beta}(\cdot)}(Y \mid \mathcal{F}_t), \tag{10}$$

where the infimum is among simple functions $\widetilde{\beta}(\cdot) \ge \beta(\cdot)$.

Remark 14. We have the lower bound

$$\text{nEV@R}^{t:T}_{\beta(\cdot)}(Y \mid \mathcal{F}_t) \ge \mathbb{E}[Y \mid \mathcal{F}_t]$$

and hence the essential infimum (10) is well defined and an element of E. Moreover, simple functions are closed under pairwise minimization. Therefore, by Theorem A.3 in Karatzas and Shreve (1998, Appendix A), there exists a sequence of decreasing simple functions $\big(\beta_n(\cdot)\big)_n$ such that

$$\text{nEV@R}^{t:T}_{\beta(\cdot)}(Y \mid \mathcal{F}_t) = \lim_{n\to\infty} \text{nEV@R}_{\beta_n(\cdot)}(Y \mid \mathcal{F}_t) \qquad \text{almost surely.}$$

We may assume, without loss of generality, that $\|\mathcal{P}_n\| = \max_{0\le i\le n-1}\Delta t_i \to 0$ as $n \to \infty$, where $\mathcal{P}_n = \{t_0 < t_1 < \cdots < t_k = T\}$ is the mesh corresponding to the simple function $\beta_n(\cdot) = \sum_{i=1}^k \beta_i \mathbb{1}_{[t_i, t_{i+1})}(\cdot)$.

4 Itô processes and nested risk measures

The preceding sections introduce nested risk measures in discrete time and subsequently extend the nested Entropic Value-at-Risk to continuous time using the monotonicity property of Theorem 12. Continuing the ideas of Proposition 7 we now consider nested risk measures on increments of a stochastic process in continuous time. A large class of such processes is given by Itô processes driven by Brownian motion. We now focus our attention on this important class.

As a motivating example, consider the Brownian motion, where Proposition 9 makes an explicit formula available.

Proposition 15. *The nested Entropic Value-at-Risk of the Wiener process W on $\mathcal{T} = [0, T]$ for a risk rate $\beta : \mathcal{T} \to [0, \infty)$ is*

$$\text{nEV@R}^{0:T}_{\beta(\cdot)}(W_T) = \int_0^T \sqrt{2\beta(t)}\,\mathrm{d}t.$$

Proof. Let $(\beta_n)_n$ be the sequence of risk rates of Remark 14. It follows from (9) that

$$\mathrm{nEV@R}_{\beta_n(\cdot)}(W_T) = \sum_{i=0}^{n-1} \sqrt{\Delta t_i}\,\sqrt{2\beta_n(t_i)\,\Delta t_i} = \sum_{i=0}^{n-1} \Delta t_i\,\sqrt{2\beta_n(t_i)}.$$

Taking the limit demonstrates

$$\mathrm{nEV@R}_{\beta(\cdot)}^{0:T}(W_T) = \int_0^T \sqrt{2\beta(t)}\,\mathrm{d}t,$$

the assertion. □

In what follows we demonstrate that the nested Entropic Value-at-Risk is well defined and finite in general cases of Itô processes. For important stochastic processes following a linear stochastic differential equation we can give explicit formulas. To this end consider the general stochastic differential equation

$$\begin{aligned} \mathrm{d}X_t &= b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t, \quad 0 < t \le T, \\ X_0 &= x_0, \end{aligned} \tag{11}$$

where $b,\ \sigma\colon \mathcal{T} \times \mathbb{R} \to \mathbb{R}$ are measurable functions. The next lemma recalls conditions for the solution of the stochastic differential equation (11) to exist.

Lemma 16 (cf. Øksendal (2003, Theorem 5.2.1))**.** *Let $t \in [0, T]$ and $x, y \in \mathbb{R}$ and suppose that*

$$|b(t, x)| + |\sigma(t, x)| \le C(1 + |x|)$$

and

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \le C\,|x - y|\,.$$

Then the stochastic differential equation (11) has a unique solution for all initial values $x_0 \in \mathbb{R}$.

Without loss of generality we assume throughout that the solution X of (11) has continuous paths. In this case we may choose the canonical representation $\Omega = C(\mathcal{T})$ with $X(t, \omega) = \omega(t)$. The next theorem extends the explicit formula obtained in Proposition 15 to a large class of linear stochastic differential equations.

Theorem 17. *For $\mathcal{T} = [0, T]$ let $X = (X_t)_{t \in \mathcal{T}}$ be a linear diffusion process driven by the stochastic differential equation*

$$\mathrm{d}X_t = \left(A(t)X_t + a(t)\right)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W_t, \quad X_0 = x_0, \tag{12}$$

where the functions A, a, σ and $\beta\colon \mathcal{T} \to [0, \infty)$ are bounded. Then the nested Entropic Value-at-Risk is given explicitly by

$$\mathrm{nEV@R}_{\beta(\cdot)}^{0:T}(X_T) = e^{\int_0^T A(s)\,\mathrm{d}s} X_0 + \int_0^T a(u) e^{\int_u^T A(s)\,\mathrm{d}s}\,\mathrm{d}u + \int_0^T e^{\int_u^T A(s)\,\mathrm{d}s}\,|\sigma(u)|\,\sqrt{2\beta(u)}\,\mathrm{d}u.$$

Proof. The solution of the linear stochastic differential equation (12) is given by (see, e.g., Karatzas and Shreve (1991, Section 5.6))

$$X_t = e^{\int_r^t A(s)\,\mathrm{d}s} \left\{ X_r + \int_r^t a(u) e^{-\int_r^u A(s)\,\mathrm{d}s}\,\mathrm{d}u + \int_r^t \sigma(u) e^{-\int_r^u A(s)\,\mathrm{d}s}\,\mathrm{d}W_u \right\}.$$

Set $\Phi(r,t) := \exp\left\{\int_r^t A(s)\,\mathrm{d}s\right\}$ and consider the sequence of risk rates $(\beta_n(\cdot))_n$ of Remark 14. By translation equivariance it follows that

$$\begin{aligned}&\mathrm{EV@R}_{\beta_n(t_i)\,\Delta t_i}\left(X_{t_{i+1}} \,\middle|\, \mathcal{F}_{t_i}\right) = \\ &\quad \Phi(t_i, t_{i+1})\, X_{t_i} + \int_{t_i}^{t_{i+1}} a(u)\,\Phi(u, t_{i+1})\,\mathrm{d}u + \mathrm{EV@R}_{\beta_n(t_i)\,\Delta t_i}\left(\int_{t_i}^{t_{i+1}} \Phi(u,t_{i+1})\sigma(u)\,\mathrm{d}W_u \,\middle|\, \mathcal{F}_{t_i}\right).\end{aligned}$$

The random variable

$$\int_{t_i}^{t_{i+1}} \Phi(u,t_{i+1})\sigma(u)\,\mathrm{d}W_u \tag{13}$$

is Gaussian with mean zero and variance $\int_{t_i}^{t_{i+1}} \Phi^2(u,t_{i+1})\sigma^2(u)\,\mathrm{d}u$. From Proposition 4 we conclude that

$$\begin{aligned}&\mathrm{EV@R}_{\beta_n(t_i)\,\Delta t_i}\left(X_{t_{i+1}} \,\middle|\, \mathcal{F}_{t_i}\right) \\ &\quad = \Phi(t_i,t_{i+1})X_{t_i} + \int_{t_i}^{t_{i+1}} a(u)\Phi(u,t_{i+1})\,\mathrm{d}u + \left(\int_{t_i}^{t_{i+1}} \Phi^2(u,t_{i+1})\sigma^2(u)\,\mathrm{d}u\cdot 2\beta_n(t_i)\,\Delta t_i\right)^{\frac{1}{2}}.\end{aligned}$$

Repeating the same steps gives

$$\begin{aligned}&\mathrm{EV@R}_{\beta_n(t_{i-1})\,\Delta t_{i-1}}\left(\mathrm{EV@R}_{\beta_n(t_i)\,\Delta t_i}\left(X_{t_{i+1}} \,\middle|\, \mathcal{F}_{t_i}\right) \,\middle|\, \mathcal{F}_{t_{i-1}}\right) \\ &\quad = \Phi(t_{i-1},t_{i+1})X_{t_{i-1}} + \int_{t_{i-1}}^{t_{i+1}} a(u)\Phi(u,t_{i+1})\,\mathrm{d}u + \left(\int_{t_i}^{t_{i+1}} \Phi^2(u,t_{i+1})\sigma^2(u)\,\mathrm{d}u\cdot 2\beta_n(t_i)\,\Delta t_i\right)^{\frac{1}{2}} \\ &\qquad + \Phi(t_i,t_{i+1})\cdot\left(\int_{t_{i-1}}^{t_i} \Phi^2(u,t_i)\sigma^2(u)\,\mathrm{d}u\cdot 2\beta_n(t_{i-1})\,\Delta t_{i-1}\right)^{\frac{1}{2}}.\end{aligned}$$

Iterating this argument and nesting with respect to n stages we obtain the explicit formula

$$\mathrm{nEV@R}_{\beta_n(\cdot)}(X_T) = \Phi(0,T)X_0 + \int_0^T a(u)\Phi(u,T)\,\mathrm{d}u + \sum_{i=0}^{n-1}\left(\int_{t_i}^{t_{i+1}} e^{2\int_u^{t_n} A(s)\,\mathrm{d}s}\sigma^2(u)\,\mathrm{d}u\cdot 2\beta_n(t_i)\,\Delta t_i\right)^{\frac{1}{2}}$$

for the nested Entropic Value-at-Risk of a linear diffusion. Using the linear approximation

$$\int_{t_i}^{t_{i+1}} e^{2\int_u^{t_n} A(s)\,\mathrm{d}s}\sigma^2(u)\,\mathrm{d}u = e^{2\int_{t_i}^{t_n} A(s)\,\mathrm{d}s}\sigma^2(t_i)\,\Delta_{t_i} + o(\Delta_{t_i})$$

we first have

$$\lim_{n\to\infty}\sum_{i=0}^{n-1}\left(\int_{t_i}^{t_{i+1}} e^{2\int_u^{t_n} A(s)\,\mathrm{d}s}\sigma^2(u)\,\mathrm{d}u\cdot 2\beta(t_i)\,\Delta t_i\right)^{\frac{1}{2}} = \lim_{n\to\infty}\sum_{i=0}^{n-1} e^{\int_{t_i}^{t_n} A(s)\,\mathrm{d}s}\,|\sigma(t_i)|\,\Delta t_i\cdot\sqrt{2\beta_n(t_i)}.$$

In the limit we obtain

$$\mathrm{nEV@R}_\beta^{0:T}(X_T) = e^{\int_0^T A(s)\,\mathrm{d}s}X_0 + \int_0^T a(u)e^{\int_u^T A(s)\,\mathrm{d}s}\mathrm{d}u + \int_0^T e^{\int_u^T A(s)\,\mathrm{d}s}\,|\sigma(u)|\,\sqrt{2\beta(u)}\,\mathrm{d}u$$

and thus the assertion. □

The Ornstein–Uhlenbeck process is a well-known example of a process satisfying the assumptions of Theorem 17 above. For this process the nested Entropic Value-at-Risk simplifies further.

Example 18 (Ornstein–Uhlenbeck)**.** Consider the Ornstein–Uhlenbeck process X_t following the stochastic differential equation with constant coefficients

$$\mathrm{d}X_t = \theta(\mu - X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad X_0 = x_0. \tag{14}$$

The closed form solution of (14) is

$$X_t = e^{-t\theta}x_0 + \mu(1 - e^{-\theta t}) + \int_0^t \sigma e^{-\theta(t-s)}\,\mathrm{d}W_s.$$

From Theorem 17 we obtain the explicit formula by setting $A(t) = -\theta$, $a(t) = \theta\mu$ and $\sigma(t) = \sigma$ in (12) and thus

$$\mathrm{nEV@R}_\beta^{0:T}(X_T) = e^{-T\theta}x_0 + \mu(1 - e^{-\theta T}) + \int_0^T e^{-\theta(T-t)}\sigma\sqrt{2\beta(t)}\,\mathrm{d}t.$$

Remark 19. It is essential for the proof of Theorem 17 that the diffusion coefficient $\sigma(\cdot)$ is independent of the state variable x. Otherwise, the stochastic integral in (13) is *not* Gaussian. However, the nested Entropic Value-at-Risk is well-defined for Itô processes where $\sigma(\cdot)$ depends on the state. Due to the strong integrability conditions for random variables in E further assumptions besides those in Lemma 16 have to be imposed as the next example illustrates.

Example 20. Let $(X_t)_t$ be the Wald martingale

$$X_t = \exp\left\{-\frac{\sigma^2}{2}t + \sigma W_t\right\},$$

then X_t is log normally distributed for which the moment generating function

$$m_Y(\ell) := \mathbb{E}\,e^{\ell X_t}$$

is not defined for $\ell > 0$ and $\mathrm{EV@R}_\beta(X_t) = \infty$.

We conclude that for a general Itô process $(X_t)_t$ the diffusion coefficient σ must not be unbounded, in general.

4.1 The risk generator

A fundamental tool in the classical theory of risk-neutral stochastic optimal control is the infinitesimal generator, a differential operator describing the evolution of the system. This subsection introduces the generator in the presence of risk, which extends the notion of the infinitesimal generator of Markov processes by replacing the expectation by a risk measure. This enables us to formulate and solve risk-averse control problems.

Throughout this section we consider the interval $\mathcal{T} = [t, T]$ and the Itô process $(X_s)_{s\in\mathcal{T}}$ given by the stochastic differential equation

$$\begin{aligned} \mathrm{d}X_s &= b(s, X_s)\,\mathrm{d}s + \sigma(s, X_s)\,\mathrm{d}W_s, \qquad s \in \mathcal{T}, \\ X_t &= x. \end{aligned} \tag{15}$$

For such processes we can define the risk generator.

Definition 21 (Risk generator)**.** Let $(X_t)_{t\in\mathcal{T}}$ be the solution of (15) on $\mathcal{T}$ with initial condition $X_t = x$. For the risk rate $\beta\colon \mathcal{T} \to \mathbb{R}$, the *risk generator based on the Entropic Value-at-Risk* is

$$\mathcal{R}_\beta\Phi(t,x) \coloneqq \lim_{h\downarrow 0}\frac{1}{h}\left(\mathrm{EV@R}_{\beta(t)\cdot h}\left(\Phi(t+h,X_{t+h})\mid X_t = x\right) - \Phi(t,x)\right),$$

for those functions $\Phi\colon \mathcal{T}\times\mathbb{R}\to\mathbb{R}$ for which the limit exists.

For normally distributed random variables Proposition 4 gives explicit representations for EV@R and we may calculate the risk generator in the case where X_t is given by (15). In light of Example 20 we impose the following condition on the diffusion coefficient σ.

Assumption 22 (Hölder continuity)**.** *There exists a* $\widetilde{C} > 0$ *and an* $\alpha > 0$ *such that*

$$|\sigma(u,X_u) - \sigma(s,X_s)| \le \widetilde{C}\,|u-s|^\alpha, \qquad s,u\in\mathcal{T}$$

and furthermore $\sigma\colon \mathcal{T}\times\mathbb{R}\to[0,\infty)$ *is bounded uniformly.*

Proposition 23 (Risk generator)**.** *Let* $(X_s)_{s\in\mathcal{T}}$ *be the solution of* (15) *on* $\mathcal{T}$ *with initial condition* $X_t = x$. *Let* $\Phi(\cdot,\cdot)$ *be continuously differentiable in the first variable and twice continuously differentiable in the second, i.e.,* $\Phi\in C^{1,2}(\mathcal{T}\times\mathbb{R})$. *If* $\frac{\partial\Phi}{\partial x}$ *is bounded, the risk generator based on* EV@R *satisfies*

$$\begin{aligned}\mathcal{R}_\beta\Phi(t,x) = {}& \frac{\partial\Phi}{\partial t}(t,x) + b(t,x)\frac{\partial\Phi}{\partial x}(t,x) + \frac{\sigma^2(t,x)}{2}\frac{\partial^2\Phi}{\partial x^2}(t,x) \\ &+ \sqrt{2\beta(t)}\cdot\left|\sigma(t,x)\frac{\partial\Phi}{\partial x}(t,x)\right|.\end{aligned} \tag{16}$$

Remark 24. The risk generator $\mathcal{R}_\beta$ can be decomposed as the sum of the classical generator *plus* the nonlinear term $\sqrt{2\beta}\cdot\left|\sigma\frac{\partial\Phi}{\partial x}\right|$. The additional risk term is a directed drift term, where the uncertain drift $\frac{\partial\Phi}{\partial x}(t,X_t)$ is scaled with volatility σ and the coefficient $\sqrt{2\beta(\cdot)}$, which expresses risk aversion. For absent risk, $\beta = 0$, we obtain the classical risk-neutral infinitesimal generator. Furthermore, if $\sigma = 0$, i.e., no randomness occurs in the model, the generator reduces to a first order differential operator describing the dynamics of a deterministic system, where risk does not apply.

Proof of Proposition 23. By assumption $\Phi\in C^{1,2}(\mathcal{T}\times\mathbb{R})$ and hence we may apply Itô's formula

$$\Phi(t+h,X_{t+h}) - \Phi(t,X_t) = \int_t^{t+h}\left(\frac{\partial\Phi}{\partial t} + b\frac{\partial\Phi}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2\Phi}{\partial x^2}\right)(s,X_s)\mathrm{d}s + \int_t^{t+h}\left(\sigma\frac{\partial\Phi}{\partial x}\right)(s,X_s)\mathrm{d}W_s.$$

For convenience we set $f_1(t,x) \coloneqq \left(\frac{\partial\Phi}{\partial t} + b\frac{\partial\Phi}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2\Phi}{\partial x^2}\right)(t,x)$ and $f_2(t,x) \coloneqq \left(\sigma\frac{\partial\Phi}{\partial x}\right)(t,x)$. In this setting $\mathcal{R}_\beta$ rewrites as

$$\mathcal{R}_\beta\Phi(t,x) = \lim_{h\downarrow 0}\frac{1}{h}\,\mathrm{EV@R}_{\beta(t)\cdot h}\left[\int_t^{t+h} f_1(s,X_s)\,\mathrm{d}s + \int_t^{t+h} f_2(s,X_s)\,\mathrm{d}W_s\,\middle|\, X_t = x\right].$$

We need to show (16) for each fixed (t,x), i.e., the inequality

$$\left|\mathcal{R}_\beta\Phi(t,x) - f_1(t,x) - \sqrt{2\beta(t)}\,|f_2(t,x)|\right| \le 0. \tag{17}$$

Note that $\mathrm{EV@R}_{\beta(t)\cdot h}\left[\frac{1}{h}\int_t^{t+h} f_2(t,x)\mathrm{d}W_s\,\middle|\, X_t = x\right] = \sqrt{2\beta(t)}\,|f_2(t,x)|$ because $f_2(t,x)$ is deterministic and hence (17) is equivalent to

$$\lim_{h\downarrow 0}\frac{1}{h}\left|\mathrm{EV@R}_{\beta(t)\cdot h}\left[\int_t^{t+h} f_1(s,X_s) - f_1(t,x)\mathrm{d}s + \int_t^{t+h} f_2(s,X_s) - f_2(t,x)\,\mathrm{d}W_s\,\middle|\, X_t = x\right]\right| = 0.$$

Using convexity of EV@R and the triangle inequality we have

$$0 \leq \lim_{h\downarrow 0} \left| \mathrm{EV@R}_{\beta(t)\cdot h} \left[\frac{1}{h} \int_t^{t+h} f_1(s, X_s) \mathrm{d}s - f_1(t,x) \middle| \, X_t = x \right] \right| + \tag{18}$$

$$+ \lim_{h\downarrow 0} \left| \mathrm{EV@R}_{\beta(t)\cdot h} \left[\frac{1}{h} \int_t^{t+h} f_2(s, X_s) - f_2(t,x)\, \mathrm{d}W_s \middle| \, X_t = x \right] \right|. \tag{19}$$

We continue by looking at each term separately. Note that $s \mapsto f_1(s, X_s) - f_1(t, x)$ is continuous almost surely and hence

$$\frac{1}{h} \int_t^{t+h} f_1(s, X_s)\, \mathrm{d}s - f_1(t,x) \tag{20}$$

converges almost surely to zero. Furthermore, for $h \leq 1$,

$$\mathrm{EV@R}_{\beta\cdot h} \left(\frac{1}{h} \int_t^{t+h} |f_1(s, X_s)|\, \mathrm{d}s \right) \leq \mathrm{EV@R}_{\beta} \left(\frac{1}{h} \int_t^{t+h} 1 + |X_s|\, \mathrm{d}s \right) \leq 1 + \mathrm{EV@R}_{\beta} \left(\int_t^{t+1} |X_s|\, \mathrm{d}s \right)$$

and hence (20) is uniformly bounded in the EV@R norm $\|\cdot\| = \mathrm{EV@R}(|\cdot|)$. Using dominated convergence, this implies convergence of (18) to zero.

Furthermore, the stochastic integral $M_h := \int_t^{t+h} f_2(s, X_s) - f_2(t,x) \mathrm{d}W_s$ in (19) is a continuous martingale with quadratic variation

$$\langle M \rangle_h = \int_t^{t+h} \left(f_2(s, X_s) - f_2(t,x) \right)^2 \mathrm{d}s.$$

Using Assumption 22, we bound the quadratic variation by

$$\langle M \rangle_h \leq \frac{\widetilde{C} \cdot h^{1+\alpha}}{1+\alpha}, \tag{21}$$

where $\widetilde{C}$ is deterministic. Recall the infimum representation of the Entropic Value-at-Risk in (5),

$$\frac{1}{h} \mathrm{EV@R}_{\beta(t)\cdot h} \left(M_h \,|\, X_t = x \right) = \inf_{\ell > 0} \frac{1}{h\ell} \left(\beta(t) \cdot h + \log \left[\mathbb{E} \exp\left(\ell M_h \right) |\, X_t = x \right] \right),$$

but M_h satisfies Novikov's condition and thus $1 = \mathbb{E} \exp\left(\ell M_h - \ell^2 \frac{\langle M \rangle_h}{2} \right)$ holds. Together with (21) we obtain

$$\mathbb{E} \left[e^{\ell M_h} \middle| \, X_t = x \right] \leq \exp\left(\frac{\ell^2}{2} \cdot \frac{\widetilde{C} \cdot h^{1+\alpha}}{1+\alpha} \right).$$

It follows similarly to Proposition 4 that

$$\frac{1}{h} \mathrm{EV@R}_{\beta(t)\cdot h} \left(M_h \,|\, X_t = x \right) \leq \inf_{\ell > 0} \frac{1}{h\ell} \left(\beta(t) \cdot h + \frac{\ell^2}{2} \cdot \frac{\widetilde{C} \cdot h^{1+\alpha}}{1+\alpha} \right) = \sqrt{2\beta(t)} \cdot \sqrt{\frac{\widetilde{C} h^{\alpha}}{1+\alpha}}, \tag{22}$$

where the infimum is attained at $\ell^* = \left(\frac{2\beta(t)(1+\alpha)}{\widetilde{C} h^{\alpha}} \right)^{\frac{1}{2}}$. We conclude that

$$\lim_{h\downarrow 0} \frac{1}{h} \left| \mathrm{EV@R}_{\beta(t)\cdot h} \left[\int_t^{t+h} f_1(s, X_s) - f_1(t,x) \mathrm{d}s + \int_t^{t+h} f_2(s, X_s) - f_2(t,x)\, \mathrm{d}W_s \middle| \, X_t = x \right] \right| = 0,$$

which shows the assertion. □

The following result is an analogue to the fundamental theorem of calculus and it also generalizes the classical Dynkin formula (see Fleming and Soner (2006)) to the risk-averse setting.

Lemma 25 (Dynkin's formula)**.** *Let $(X_s)_{s\in\mathcal{T}}$ be the solution of (15) on $\mathcal{T}$ with initial condition $X_t = x$. For $\Phi \in C^{1,2}(\mathcal{T}\times\mathbb{R})$ such that $\frac{\partial\Phi}{\partial x}$ is bounded, the risk-averse Dynkin formula*

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\Phi(r,X_r)\right) = \Phi(t,x) + \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r \mathcal{R}_\beta\Phi(s,X_s)\,\mathrm{d}s\right) \tag{23}$$

holds for any $t \le r \le T$.

Proof. The rather technical proof builds on the arguments in the proof of Proposition 23 as well as Theorem 12. We have moved the detailed proof to the appendix. □

4.2 The relation of risk measures to g-expectations

The previous sections focus on developing a dynamic nested version of Entropic Value-at-Risk and a risk-averse analogue of the generator is introduced. We now relate the dynamic risk measures introduced above to solutions of certain backwards stochastic differential equations, called g-expectation (cf. Rosazza Gianin (2006)) . Let $X = (X_t)_t$ be the forward process given by (15) and let (Y, Z) be the solution of the following backwards stochastic differential equation

$$Y_t = \int_0^T c(s,X_s)\,\mathrm{d}s - \int_t^T g(s,Z_s)\,\mathrm{d}s + \int_t^T Z_s\,\mathrm{d}W_s, \tag{24}$$

where the cost c is accumulated over the entire time horizon. Rosazza Gianin (2006, Proposition 19) shows that if g is convex and positively homegeneous in the second component and satisfies some regularity assumptions, then

$$\rho^{t:T}\left(\int_0^T c(s,X_s)\,\mathrm{d}s\right) = Y_t$$

describes a dynamic risk measure, the so-called *(conditional) g-expectation*.

Moreover, there is an intimate relationship between solutions of backwards stochastic differential equations and partial differential equations (see Zhang (2017) for details). In fact, $V(t,x) := Y_t^{X_t=x}$ solves the partial differential equation

$$0 = \frac{\partial V}{\partial t}(t,x) + c(t,x) + b(t,x)\cdot\frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2(t,x)\cdot\frac{\partial^2 V}{\partial x^2} + g\left(t,\,\sigma(t,x)\cdot\frac{\partial V}{\partial x}\right). \tag{25}$$

Comparing (25) with the risk generator $\mathcal{R}_\beta$ it follows that g-expectation defines nested risk measures in continuous time and for the Entropic Value-at-Risk we have

$$g(t,z) := \sqrt{2\beta(t)}\cdot|z|\,.$$

We now show that all coherent risk measures, for which the risk generator exists, lead to a nonlinearity of the form

$$\mu(t)\cdot\left|\sigma(t,x)\cdot\frac{\partial V}{\partial x}\right|, \qquad \mu \ge 0.$$

By Kusuokas representation theorem it is sufficient to consider the Average Value-at-Risk, the most prominent coherent risk measure. We demonstrate that the nested Average Value-at-Risk degenerates either to the expectation or the essential supremum provided that the limit over all partitions in (6) is taken without properly rescaling the risk levels. Furthermore, we provide the proper rescaling such that the nested Average Value-at-Risk does not degenerate in the continuous time setting.

The Average Value-at-Risk at risk level α is given by (cf. Ogryczak and Ruszczyński (2002))

$$\mathrm{AV@R}_\alpha(Y) := \min_{x\in\mathbb{R}} x + \frac{1}{1-\alpha}\,\mathbb{E}\,(Y-x)_+\,.$$

For normally distributed $Y \sim \mathcal{N}(\mu,\sigma^2)$ we have the explicit formula $\mathrm{AV@R}_\alpha(Y) = \mu + \sigma\frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}$, where φ and Φ are the density and cumulative distribution function of the standard normal distribution, respectively. The conditional Average Value-at-Risk is given by

$$\mathrm{AV@R}_\alpha\,(Y\,|\,\mathcal{F}_s) := \min_{x\in\mathbb{R}}\, x + \frac{1}{1-\alpha}\,\mathbb{E}\left[(Y-x)_+ \mid \mathcal{F}_s\right].$$

Similarly to the Entropic Value-at-Risk, we now consider a risk rate $\alpha\colon \mathcal{T} \to [0,1]$ and introduce the nested Average Value-at-Risk for a vector of risk levels $(\alpha(t_0),\dots,\alpha(t_{n-1}))$ on a partition $\mathcal{P}$ by

$$\mathrm{nAV@R}_\alpha^{\mathcal{P}}\left(Y\,\middle|\,\mathcal{F}_{t_i}\right) := \mathrm{AV@R}_{\alpha(t_i)}\left(\dots\,\mathrm{AV@R}_{\alpha(t_{n-1})}\left(Y\,\middle|\,\mathcal{F}_{t_{n-1}}\right)\dots\middle|\,\mathcal{F}_{t_i}\right).$$

For a Brownian motion evaluated on $\mathcal{P}$ (cf. Corollary 9) the nested Average Value-at-Risk evaluates to

$$\mathrm{nAV@R}_\alpha^{\mathcal{P}}(W_T) = \sum_{i=0}^{n-1}\sqrt{\Delta t_i}\,\frac{\varphi\left(\Phi^{-1}\left(\alpha(t_i)\right)\right)}{1-\alpha(t_i)}. \tag{26}$$

The following example show that natural choices for risk rates $\alpha(\cdot)$ lead to a degenerate risk evaluation.

Example 26. We consider constant risk levels $\alpha(t_i) = \alpha_0 \in (0,1)$. Then the limit of the right side of (26) above is

$$\lim_{n\to\infty}\sum_{i=0}^{n-1}\sqrt{\Delta t_i}\,\frac{\varphi\left(\Phi^{-1}(\alpha_0)\right)}{1-\alpha_0} = \infty.$$

On the other hand the analysis in Xin and Shapiro (2011) suggests to choose the risk rate $\alpha(t_i) := \Delta t_i \in (0,1)$. Then

$$\sum_{i=0}^{n-1}\sqrt{\Delta t_i}\,\frac{\varphi\left(\Phi^{-1}(\alpha(t_i))\right)}{1-\alpha(t_i)} = \sum_{i=0}^{n-1}\sqrt{\Delta t_i}\frac{1}{(1-\Delta t_i)}\cdot\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\Phi^{-1}(\Delta t_i)\right)^2\right\},$$

but for p close to zero the asymptotic relation $\Phi^{-1}(p) \sim -\sqrt{-2\log p}$ holds (see the stable reference http://dlmf.nist.gov/7.17.iii) and thus

$$\sum_{i=0}^{n-1}\sqrt{\Delta t_i}\frac{1}{(1-\Delta t_i)}\cdot\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\Phi^{-1}(\Delta t_i)\right)^2\right\} \sim \sum_{i=0}^{n-1}\frac{\sqrt{\Delta t_i}\;\Delta t_i}{(1-\Delta t_i)\sqrt{2\pi}},$$

which tends to zero as $\Delta t_i \to 0$.

We now derive the correct asymptotic behavior of α, such that the nested Average-Value-at-Risk does not degenerate. To this end, we construct a *modified risk rate* $\mathsf{A}^\alpha(t,h)$ depending on time t as well as the step size h.

Theorem 27. *Consider a Brownian motion* $W = (W_t)_{t\in[0,T]}$ *on the interval* $[0,T]$*. For a risk rate* $\alpha\colon [0,T] \to [0,\infty)$ *we set*[1]

$$\mathsf{A}^\alpha(t,h) \coloneqq \Phi\left(-\sqrt{-\log(2\pi h\cdot\alpha(t))}\right).$$

Then the nested Average Value-at-Risk for the Brownian motion is given by

$$\mathrm{nAV@R}^{0:T}_{\mathsf{A}^\alpha(\cdot)}\left(\int_0^T \mathrm{d}W_s\right) = \int_0^T \sqrt{\alpha(s)}\,\mathrm{d}s.$$

Proof. Let φ be the density of the standard normal distribution, and Φ be its cumulative distribution function. Denote by $\mathcal{P} = (t_0,\dots,t_n)$ a partition of the interval $[0,T]$ and set $\Delta t_i = t_{i+1} - t_i$. For $\mathsf{A}^\alpha(t,h)$ as in the theorem we obtain

$$\varphi\left(\Phi^{-1}(\mathsf{A}^\alpha(t_i,\Delta t_i))\right) = \sqrt{\Delta t_i\,\alpha(t_i)}$$

and hence (26) shows that

$$\mathrm{nAV@R}^{\mathcal{P}}_\alpha(W_T) = \sum_{i=0}^{n-1}\Delta t_i\,\frac{\alpha(t_i)}{1-\alpha(t_i)}.$$

Proceeding exactly as in Definition 13 and Remark 14 shows that

$$\mathrm{nAV@R}^{0:T}_{\mathsf{A}^\alpha(\cdot)}\left(\int_0^T \mathrm{d}W_s\right) = \int_0^T \sqrt{\alpha(s)}\,\mathrm{d}s$$

and thus the assertion. □

The following proposition shows that the risk generator with respect to the modified Average Value-at-Risk with risk rate $\mathsf{A}^\alpha(\cdot)$ coincides with the risk generator based on the Entropic Value-at-Risk up to a scaling factor. Its proof is similar to the proof of Proposition 23.

Proposition 28 (Risk generator for AV@R)**.** *Let* $(X_s)_{s\in\mathcal{T}}$ *be the solution of* (15) *on* $\mathcal{T} = [t,T]$ *with initial condition* $X_t = x$ *satisfying Assumption 22. For* $\Phi \in C^{1,2}(\mathcal{T}\times\mathbb{R})$ *such that* $\frac{\partial\Phi}{\partial x}$ *is bounded, the risk generator based on* AV@R *satisfies*

$$\mathcal{R}\Phi(t,x) = \frac{\partial\Phi}{\partial t}(t,x) + b(t,x)\cdot\frac{\partial\Phi}{\partial x}(t,x) + \frac{1}{2}\sigma^2(t,x)\cdot\frac{\partial^2\Phi}{\partial x^2}(t,x) + \sqrt{\alpha(t)}\left|\sigma(t,x)\frac{\partial\Phi}{\partial x}(t,x)\right|.$$

The preceding proposition implies that all nested coherent risk measures lead to the same g-expectation up to a scaling factor. Rosazza Gianin (2006, Section 3) discusses the particular choice and interpretation of the driver g from an economic perspective whereas our results connect the appropriate choice of g to static risk measures by modifying the risk levels in a consistent manner.

Moreover, our results demonstrate that from the perspective of dynamic programming equations it is not important which coherent risk measure is considered as the risk generators are essentially the same. However, the adapted risk levels for the Entropic Value-at-Risk can be interpreted intuitively as a fixed risk level related to the time horizon considered, whereas the modified risk levels A^α for the Average Value-at-Risk do not allow for such an immediate interpretation. For this reason, we consider the nested Entropic Value-at-Risk as natural choice for a coherent risk measure in continuous time.

[1] The asymptotic expansion

$$\mathsf{A}^\alpha(t,h) \sim \sqrt{h}\left(\frac{1}{\sqrt{-\log(2\pi h\,\alpha(t))}} - \frac{1}{\sqrt{-\log(2\pi h\,\alpha(t))}^3} + \frac{3}{\sqrt{-\log(2\pi h\,\alpha(t))}^5}\cdots\right)$$

holds for $h \to 0$

5 The risk-averse control problem

The preceding section develops a risk-averse extension of the infinitesimal generator. Moreover, a risk-averse Dynkin formula is shown. Using these results, we now formulate a risk-averse optimal control problem and derive associated Hamilton–Jacobi–Bellman equations.

Consider the set of controls

$$\mathcal{U}[0,T] \coloneqq \{u\colon \mathcal{T}\times\Omega \to U \mid u \text{ is adapted } \}$$

and $U\subset\mathbb{R}$. For any initial condition $(t,x)\in[0,T)\times\mathbb{R}$ and control $u\in\mathcal{U}[t,T]$ we consider the controlled stochastic process $(X_s^{t,x,u})_s$ given by (compare to (15))

$$\begin{aligned} \mathrm{d}X_s^{t,x,u} &= b\big(s,X_s^{t,x,u},u(s)\big)\,\mathrm{d}s+\sigma\big(s,X_s^{t,x,u},u(s)\big)\,\mathrm{d}W_s, \qquad s\in[t,T], \\ X_t^{t,x,u} &= x. \end{aligned} \tag{27}$$

The aim is to evaluate the risk of the accumulated cost over time, therefore we consider a *cost rate* $c\colon [0,T]\times\mathbb{R}\times U\to\mathbb{R}$ and a *terminal cost* $\Psi\colon\mathbb{R}\to\mathbb{R}$ so that the total cost accumulated over $[t,T]$ is

$$\int_t^T c\big(s,X_s^{t,x,u},u(s)\big)\,\mathrm{d}s+\Psi\big(X_T^{t,x,u}\big).$$

For $u\in\mathcal{U}[t,T]$ and adapted $\beta\colon [0,T]\to[0,\infty)$ we define the *controlled value function* V^u by

$$V^u(t,x)\coloneqq \mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(\int_t^T c\big(s,X_s^{t,x,u},u(s)\big)\,\mathrm{d}s+\Psi\big(X_T^{t,x,u}\big)\,\middle|\,X_t=x\right).$$

For arbitrary $u\in\mathcal{U}[0,T]$ the controlled value function V^u may not exist. We follow Fleming and Soner (2006, p. 141) and introduce the set of *admissible controls*.

Definition 29 (Admissible control)**.** $\mathcal{U}[t,T]$ is called an *admissible control system* if it satisfies the following conditions.

(i) For $u\in\mathcal{U}[t,T]$, the function $u\colon [t,T]\times\Omega\to U$ is an adapted process with respect to the Brownian filtration.

(ii) For any initial value $x\in\mathbb{R}$ and $u\in\mathcal{U}[t,T]$ the stochastic differential equation (27) admits a unique solution and $V^u(t,x)$ is well defined.

From now on $\mathcal{U}[t,T]$ always denotes an admissible control system and we define the *optimal value function* $V\colon [0,\,T]\times\mathbb{R}\to\mathbb{R}$ as

$$\begin{aligned} V(t,x) &\coloneqq \inf_{u\in\mathcal{U}[t,T]} V^u(t,x) \\ &= \inf_{u\in\mathcal{U}[t,T]} \mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(\int_t^T c(s,X_s^{t,x,u},u(s))\,\mathrm{d}s+\Psi(X_T^{t,x,u})\,\middle|\,X_t=x\right). \end{aligned} \tag{28}$$

The risk-averse control problem can now be formulated as:

given $(t,x)\in[0,T)\times\mathbb{R}$, find an admissible control $u^*\in\mathcal{U}[t,T]$ such that

$$V^{u^*}(t,x)=\inf_{u\in\mathcal{U}[t,T]} V^u(t,x). \tag{29}$$

The following proposition guarantees that condition (ii) in Definition 29 is satisfied. It is an extension of Lemma 16 as well as Lemma 23.

Proposition 30. *Let $s \in [t,T]$ and $x_1, x_2 \in \mathbb{R}$ and $u_1, u_2 \in U$. Suppose there exists a constant $C > 0$ such that*

$$|b(s,x_1,u_1)| + |\sigma(s,x_1,u_1)| + |c(s,x_1,u_1)| + |\Psi(x_1)| \le C(1 + |x_1| + |u_1|)$$

and

$$\begin{aligned}|b(s,x_1,u_1) - b(s,x_2,u_2)| + |\sigma(s,x_1,u_1) - \sigma(s,x_2,u_2)| + |c(s,x_1,u_1) - c(s,x_2,u_2)|\\ \le C\left(|x_1 - x_2| + |u_1 - u_2|\right)\end{aligned}$$

hold. Then the stochastic differential equation (27) has a unique solution. Moreover, if for $u \in U$, the diffusion coefficient $\sigma(\cdot, u(\cdot))$ satisfies Assumption 22, the controlled value function $V^u(t,x)$ is well defined and deterministic.

5.1 Principle of dynamic programming

We show that the risk-averse optimal value function $V(\cdot,\cdot)$ defined in (28) satisfies an analogue of the dynamic programming principle. Furthermore, we introduce the risk-averse Hamilton–Jacobi–Bellman equations and show that the optimal value function solves these equations in the sense of viscosity solutions. Additionally, we provide a verification theorem, showing that a classical solution to the risk-averse Hamilton–Jacobi–Bellman equation is the optimal value function (28).

Lemma 31 (Dynamic programming principle)**.** *Let $(t,x) \in [0,T) \times \mathbb{R}$ and $r \in (t,T]$ and suppose that $\mathcal{U}[t,T]$ is an admissible control system, then it holds that*

$$V(t,x) = \inf_{u \in \mathcal{U}(t,r)} \mathrm{nEV@R}_{\beta(\cdot)}^{t:r}\left(\int_t^r c(s, X_s^{t,x,u}, u_s)\,\mathrm{d}s + V(r, X_r^{t,x,u})\middle| X_t = x\right). \tag{30}$$

Proof. For every $\varepsilon > 0$ there exists a $\widetilde{u}(\cdot) \in \mathcal{U}[t,T]$ such that $V(t,x) + \varepsilon \ge V^{\widetilde{u}}(t,x)$. Using the recursive property of the nested risk measures we obtain

$$\begin{aligned}V(t,x) + \varepsilon &\ge V^{\widetilde{u}}(t,x)\\ &= \mathrm{nEV@R}_{\beta(\cdot)}^{t:r}\left(\mathrm{nEV@R}_{\beta(\cdot)}^{r:T}\left(\int_t^T c(s, X_s^{t,x,\widetilde{u}}, \widetilde{u}_s)\,\mathrm{d}s + \Psi(X_T^{t,x,\widetilde{u}})\middle| \mathcal{F}_r\right)\middle| X_t = x\right).\end{aligned}$$

For each $r \in (t,T]$ the inequality

$$\mathrm{nEV@R}_{\beta(\cdot)}^{r:T}\left(\int_r^T c(s, X_s^{r,x(r),\widetilde{u}}, \widetilde{u}_s)\,\mathrm{d}s + \Psi(X_T^{r,x(r),\widetilde{u}})\middle| \mathcal{F}_r\right) \ge V(r, X_r^{t,x,\widetilde{u}})$$

holds almost surely and thus

$$V(t,x) + \varepsilon \ge \inf_{u \in \mathcal{U}(t,r)} \mathrm{nEV@R}_{\beta(\cdot)}^{t:r}\left(\int_t^r c(s, X_s^{t,x,u}, u_s)\,\mathrm{d}s + V(r, X_r^{t,x,u})\middle| X_t = x\right).$$

As $\varepsilon > 0$ can be chosen arbitrarily we have shown the inequality $\ge$ in (30).

To see the converse inequality consider a fixed $\varepsilon > 0$ and let $\bar{u} \in \mathcal{U}[t,r]$ be an ε-optimal solution to (30), that is

$$\begin{aligned}&\inf_{u \in \mathcal{U}(t,r)} \mathrm{nEV@R}_{\beta(\cdot)}^{t:r}\left(\int_t^r c(s, X_s^{t,x,u}, u_s)\,\mathrm{d}s + V(r, X_r^{t,x,u})\middle| X_t = x_t\right) + \varepsilon\\ &\qquad \ge \mathrm{nEV@R}_{\beta(\cdot)}^{t:r}\left(\int_t^r c(s, X_s^{t,x,\bar{u}}, \bar{u}_s)\,\mathrm{d}s + V(r, X_r^{t,x,\bar{u}})\middle| X_t = x\right).\end{aligned}$$

For every $y \in \mathbb{R}$, let $\widetilde{u}(y) \in \mathcal{U}[r,T]$ be such that $V(r,y)+\varepsilon \geq V^{\widetilde{u}(y)}(r,y)$. We may assume that the mapping $y \mapsto \widetilde{u}(y)$ is measurable (measurable selection theorem) and construct the control function

$$u_s^0 = \begin{cases} \bar{u}_s & s \in [t,r) \\ \widetilde{u}_s(X_r^{t,x,\bar{u}}) & s \in [r,T] \end{cases}.$$

Using monotonicity and the recursive property of the nested risk measure we get

$$\begin{aligned} &\mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r c(s, X_s^{t,x,\bar{u}}, \bar{u}_s)\,\mathrm{d}s + V(r, X_r^{t,x,\bar{u}}) \middle| X_t = x\right) \\ &\quad\geq \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r c(s, X_s^{t,x,\bar{u}}, \bar{u}_s)\,\mathrm{d}s + V^{\widetilde{u}_s(X_r^{t,x,\bar{u}})}(r, X_r^{t,x,\bar{u}}) \middle| X_t = x\right) - \varepsilon \\ &\quad= \mathrm{nEV@R}^{t:T}_{\beta(\cdot)}\left(\int_t^T c(s, X_s^{t,x,u^0}, u_s^0)\,\mathrm{d}s + \Psi(X_T^{t,x,u^0}) \middle| X_t = x\right) - \varepsilon \\ &\quad= V^{u_0}(t,x) - \varepsilon. \end{aligned}$$

Combining the last inequalities we get

$$\inf_{u \in \mathcal{U}(t,r)} \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r c(s, X_s^{t,x,u}, u_s)\,\mathrm{d}s + V(r, X_r^{t,x,u}) \middle| X_t = x\right) + \varepsilon \geq V(t,x) - \varepsilon$$

and as $\varepsilon > 0$ was arbitrary. The assertion follows. □

5.2 Hamilton–Jacobi–Bellman equations

The dynamic programming principle (30) suggests to consider quantities of the form

$$\mathrm{nEV@R}^{t:T}_{\beta(\cdot)}\left(\int_t^T c(s, X_s)\mathrm{d}s + V(T, X_T) - V(t,x) \middle| X_t = x\right),$$

where $c(\cdot,\cdot)$ is a cost functional and $V(\cdot,\cdot)$ a terminal cost functional. The next theorem extends Proposition 23 to this case.

Theorem 32. *Let $(X_s)_{s\in\mathcal{T}}$ be the solution of (15) on $\mathcal{T}$ with initial condition $X_t = x$. Let $c(\cdot,\cdot)$, $V(\cdot,\cdot) \in C^{1,2}(\mathcal{T} \times \mathbb{R})$ and let V satisfy the conditions of Proposition 23, then it holds that*

$$\lim_{h \downarrow 0} \frac{1}{h} \mathrm{nEV@R}^{t:t+h}_{\beta(\cdot)}\left(\int_t^{t+h} c(s, X_s)\mathrm{d}s + V(t+h, X_{t+h}) - V(t,x) \middle| X_t = x\right) = c(t,x) + \mathcal{R}_\beta V(t,x). \tag{31}$$

Proof. From convexity of coherent risk measures it follows that the left side of (31) can be bounded from above by

$$\lim_{h\downarrow 0} \mathrm{nEV@R}^{t:t+h}_{\beta(\cdot)}\left(\frac{1}{h}\int_t^{t+h} c(s,X_s)\mathrm{d}s \middle| X_t = x\right) + \lim_{h\downarrow 0} \mathrm{nEV@R}^{t:t+h}_{\beta(\cdot)}\left(\frac{V(t+h, X_{t+h}) - V(t,x)}{h} \middle| X_t = x\right),$$

where the first part converges to $c(t,x)$ following the arguments in the proof of Proposition 23.

The second term can be rewritten as a limit over a sequence of risk rates $(\beta_n(\cdot))_n \colon [t, t+h] \to [0,\infty)$ as in Remark 14, i.e.,

$$\lim_{\beta_n} \mathrm{EV@R}_{\beta_n(t_0)\,\Delta n}\left(\dots \mathrm{EV@R}_{\beta_n(t_{n-1})\,\Delta n}\left(\frac{V(t+h, X_{t+h}) - V(t,x)}{h} \middle| \mathcal{F}_{t_{n-1}}\right) \dots \middle| \mathcal{F}_{t_0}\right). \tag{32}$$

We now show that the iterated limit (32) converges uniformly to $\mathcal{R}_\beta V(t,x)$. To this end, employing Theorem 12 gives an upper bound for

$$\mathrm{nEV@R}_{\beta_n(\cdot)}\left(\frac{V(t+h,X_{t+h})-V(t,x)}{h}\,\middle|\, X_t=x\right)-\mathcal{R}_\beta V(t,x) \tag{33}$$

of the form

$$\left|\mathrm{EV@R}_{\beta(t_0)\cdot C\cdot h}\left(\frac{V(t+h,X_{t+h})-V(t,x)}{h}\right)-\mathcal{R}_\beta V(t,x)\right|.$$

Here C is chosen such that $C\cdot\beta(t_0)\geq\beta(\cdot)$ on $[t,t+h]$. The analysis in the proof of Proposition 23 implies (compare with Equation (22)) the existence an ε depending on C, h and the inital value t_0 such that

$$\left|\mathrm{EV@R}_{C\cdot\beta(t_0)\,h}\left(\frac{V(t+h,X_{t+h})-V(t,x)}{h}\right)-\mathcal{R}_\beta V(t,x)\right|\leq\varepsilon(C\cdot h,t_0).$$

Exchanging the order of the arguments in (33) we get

$$\mathcal{R}_\beta V(t,x)-\mathrm{nEV@R}_{\beta_n(\cdot)}\left(\frac{V(t+h,X_{t+h})-V(t,x)}{h}\,\middle|\, X_t=x\right)\geq-\varepsilon(C\cdot h,t_0).$$

We may conclude that for fixed $\varepsilon'>0$ there exists an $h'>0$ such that for every smaller $h>0$

$$\left|\mathrm{nEV@R}_{\beta_n(\cdot)}\left(\frac{V(t+h,X_{t+h})-V(t,x)}{h}\,\middle|\, X_t=x\right)-\mathcal{R}_\beta V(t,x)\right|\leq\varepsilon'$$

holds independently of the risk rate $\beta_n(\cdot)$. Therefore, the limits can be interchanged. By Definition 21 we then obtain

$$\begin{aligned}\lim_{h\downarrow 0}\mathrm{nEV@R}^{t:t+h}_{\beta(\cdot)}\left(\frac{V(t+h,X_{t+h})-V(t,x)}{h}\,\middle|\,\mathcal{F}_{t_0}\right)&=\lim_{h\downarrow 0}\mathrm{EV@R}_{\beta(t_0)\,h}\left(\frac{V(t+h,X_{t+h})-V(t,x)}{h}\right)\\&=\mathcal{R}_\beta V(t,x).\end{aligned}$$

Now, applying the triangle inequality to

$$\lim_{h\downarrow 0}\left|\mathrm{nEV@R}^{t:t+h}_{\beta(\cdot)}\left(\frac{1}{h}\left(\int_t^{t+h}c(s,X_s)\mathrm{d}s+V(t+h,X_{t+h})-V(t,x)\right)\middle|\,\mathcal{F}_t\right)-c(t,x)-\mathcal{R}_\beta V(t,x)\right|,$$

the assertion follows immediately. □

Formally taking the limit $r\longrightarrow t$ in the dynamic programming principle

$$0=\inf_{u\in\mathcal{U}(t,r)}\frac{1}{r-t}\,\mathrm{nEV@R}^{t:r}_{\beta}\left(\int_t^r c(s,X_s^{t,x,u},u(s))\,\mathrm{d}s+V(r,X_r^{t,x,u})-V(t,x)\middle|\,X_t=x\right)$$

together with Theorem 32 shows that

$$0=\inf_{u\in U}c(t,x,u)+\mathcal{R}_\beta V(t,x).$$

This suggests to consider the partial differential equation on the space $C_b^{1,2}\left([t,T]\times\mathbb{R}\right)$ for $(t,x)\in[0,T]\times\mathbb{R}$

$$\frac{\partial V}{\partial t}(t,x)=\mathcal{H}\left(t,x,\frac{\partial V}{\partial x},\frac{\partial^2 V}{\partial x^2}\right), \tag{34}$$

with terminal condition $v(T,x) = \Psi(x)$ and Hamiltonian $\mathcal{H}\colon [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ given by

$$\mathcal{H}(t,x,g,H) \coloneqq \sup_{u\in U}\left\{-c(t,x,u) - g\cdot b(t,x,u) - H\cdot\frac{1}{2}\sigma^2(t,x,u) - |g|\cdot\sqrt{2\beta(t)}\sigma(t,x,u)\right\}.$$

The formal derivation suggests that the value function

$$V(t,x) = \inf_{u\in\mathcal{U}[t,T]} \mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(\int_t^T c(s,X_s^{t,x,u},u(s))\,\mathrm{d}s + \Psi(X_T^{t,x,u})\middle| X_t = x\right) \tag{35}$$

solves the nonlinear partial differential equation (34).

Remark 33. Using backwards stochastic differential equations (BSDEs) an equation similar to (34) was derived in Ruszczyński and Yao (2015). However, our analysis in Section 3 and Section 4 provides a more elementary approach towards risk averse dynamic control. Moreover, we clarify the relationship between static risk measures and their dynamic extensions by linking the risk level to the corresponding time period.

The value function (35) may not be regular enough and a more general concept of solutions for (34) is needed. Therefore, the concept of viscosity solution was introduced by Crandall and Lions (1983). We recall the definition in the next subsection and elaborate that the classical theory of viscosity solutions developed for the risk-neutral setting is sufficient for the risk-averse case as well.

5.3 Viscosity solutions

We show that the optimal value function $V(\cdot,\cdot)$ defined in (28) solves equation (34) and vice versa a solution of (34) is the optimal value function of problem (28). In order to discuss solutions of the partial differential equation (34) we recall the concept of viscosity solutions.

Definition 34 (Viscosity solution)**.** A function $v\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ satisfying $v(T,x)=\Psi(x)$ for all $x\in\mathbb{R}$ is called a *viscosity solution* of (34) if the following two conditions are met:

- v is a viscosity *sub*solution, i.e., for every $w\in C_b^{1,2}([0,T]\times\mathbb{R})$ such that $w\geq v$ on $[0,T]\times\mathbb{R}$ and $\min_{(t,x)}\{w(t,x)-v(t,x)\}=0$, the inequality

$$0\geq -\frac{\partial w}{\partial t}(\bar t,\bar x) + \mathcal{H}(\bar t,\bar x,\frac{\partial w}{\partial x},\frac{\partial^2 w}{\partial x^2})$$

  holds for every $(\bar t,\bar x)\in[0,T]\times\mathbb{R}$ such that $w(\bar t,\bar x)=v(\bar t,\bar x)$.

- v is a viscosity *super*solution, i.e., for every $w\in C_b^{1,2}([0,T]\times\mathbb{R})$ such that $w\leq v$ on $[0,T]\times\mathbb{R}$ and $\min_{(t,x)}\{v(t,x)-w(t,x)\}=0$, the inequality

$$0\leq -\frac{\partial w}{\partial t}(\bar t,\bar x) + \mathcal{H}(\bar t,\bar x,\frac{\partial w}{\partial x},\frac{\partial^2 w}{\partial x^2})$$

  holds for every $(\bar t,\bar x)\in[0,T]\times\mathbb{R}$ such that $w(\bar t,\bar x)=v(\bar t,\bar x)$.

The following theorems highlights the relation between the optimal value function $V(\cdot,\cdot)$ defined in (28) and the partial differential equation (34). We first show that the optimal value function $V(\cdot,\cdot)$ solves the Hamilton–Jacobi–Bellman partial differential equation (34) in the sense of viscosity solutions.

Theorem 35. *Suppose the assumptions of Proposition 30 as well as Assumption 22 are satisfied and suppose the control set U is compact. Then the optimal value function $V(\cdot,\cdot)$ is a viscosity solution of the equations* (34).

Proof. The proof follows familiar arguments and thus can be found in the appendix. □

We finally demonstrate that a classical solution of (34) is the optimal value function of the optimal control problem (29). This provides a converse statement to Theorem 35.

Theorem 36 (Verification theorem)**.** *Suppose the assumptions of Proposition 30 as well as Assumption 22 are satisfied. Let* $L \in C_b^{1,2}\big([0,T] \times \mathbb{R}\big)$ *be bounded and satisfy the partial differential equation* (34)*, then* $L(t,x) \le V^u(t,x)$ *for all* $u \in \mathcal{U}[t,T]$ *and all* $(t,x) \in [0,T] \times \mathbb{R}$*. Moreover, if a control* $u^* \in \mathcal{U}[0,T]$ *exists such that for almost all* $(s,\omega) \in [0,T] \times \Omega$ *the relation*

$$u_s^* \in \operatorname*{arg\,min}_{v \in U} \Big\{ c(s, X_s^{t,x,u^*}, v) + b(s, X_s^{t,x,u^*}, v)\,\frac{\partial L}{\partial x}(t, X_s^{t,x,u^*}, v) + \frac{1}{2}\sigma^2(t, X_s^{t,x,u^*}, v)\,\frac{\partial^2 L}{\partial x^2}(t, X_s^{t,x,u^*}, v) + \sqrt{2\beta(t)}\,\sigma(t, X_s^{t,x,u^*}, v)\left|\frac{\partial L}{\partial x}(t, X_s^{t,x,u^*}, v)\right| \Big\} \tag{36}$$

holds, then $L(t,x) = V(t,x) = V^{u^*}(t,x)$ *for all* $(t,x) \in [0,T] \times \mathbb{R}$*.*

Proof. Let $(t,x) \in [0,T] \times \mathbb{R}$ and consider a control $\widetilde{u} \in \mathcal{U}[t,T]$. It holds by assumption that

$$0 = -\frac{\partial L}{\partial t}(t,x) + \mathcal{H}\Big(t, x, \frac{\partial L}{\partial x}, \frac{\partial^2 L}{\partial x^2}\Big). \tag{37}$$

Equation (37) shows that for the fixed control $\widetilde{u} \in \mathcal{U}[t,T]$ and all $s \in [t,T]$

$$0 \le c(s, X_s^{t,x,\widetilde{u}}, \widetilde{u}_s) + \mathcal{R}_\beta L(s, X_s^{t,x,\widetilde{u}}). \tag{38}$$

It now follows that

$$L(t,x) \le \mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(L(t,x) + \int_t^T \mathcal{R}_\beta L(s, X_s^{t,x,\widetilde{u}})\,\mathrm{d}s + \int_t^T c(s, X_s^{t,x,\widetilde{u}}, \widetilde{u}_s)\mathrm{d}s \right). \tag{39}$$

In the spirit of Lemma 25 we now show that the right side of (39) is equal to

$$\mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(\int_t^T c(s, X_s^{t,x,\widetilde{u}}, \widetilde{u}_s)\,\mathrm{d}s + L(T, X_T^{t,x,\widetilde{u}}) \right). \tag{40}$$

In fact, taking the difference of the right side of (39) and (40) we obtain an upper bound

$$\mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(L(t,x) + \int_t^T \mathcal{R}_\beta L(s, X_s^{t,x,\widetilde{u}})\,\mathrm{d}s - L(T, X_T^{t,x,\widetilde{u}}) \right). \tag{41}$$

Using Itô's Lemma on $L(T, X_T^{t,x,\widetilde{u}})$ we get

$$(41) = \mathrm{nEV@R}_{\beta(\cdot)}^{t:r}\left(\int_t^r \left(\sigma \frac{\partial L}{\partial x}\right)(s, X_s)\,\mathrm{d}W_s - \int_t^r \sqrt{2\beta}\left|\sigma \frac{\partial L}{\partial x}\right|(s, X_s)\,\mathrm{d}s \right).$$

This is equation (43) in the proof of Lemma 25. Following exactly the same steps we show assertion (40), i.e.,

$$L(t,x) \le \mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(\int_t^T c(s, X_s^{t,x,\widetilde{u}}, \widetilde{u}_s)\,\mathrm{d}s + L(T, X_T^{t,x,\widetilde{u}}) \right) = V^{\widetilde{u}}(t,x),$$

which concludes the first part of the assertion. Now suppose a control $u^* \in \mathcal{U}[0,T]$ exists such that (36) is satisfied, then the inequality (38) becomes an equality, where again the above steps show that

$$L(t,x) = \mathrm{nEV@R}_{\beta(\cdot)}^{t:T}\left(\int_t^T c(s, X_s^{t,x,u^*}, u_s^*)\,\mathrm{d}s + \Psi(X_T^{t,x,u^*}) \right) = V^{u^*}(t,x),$$

concluding the proof. □

6 Summary

This paper introduces nested risk measures in continuous time, explains the “g” in g-expectation and derives risk-averse Hamilton–Jacobi–Bellman equations.

Nested risk measures in continuous time are constructed as suitable limits from discrete time risk measures. We demonstrate that the natural building block for nesting is the Entropic Value-at-Risk. The risk levels have to be adjusted to the time period – otherwise, the nested risk measures degenerate.

Conditional risk measures are associated with a risk-generator, a nonlinear generalization of the infinitesimal generator. We relate nested risk measures with dynamic risk measures based on backwards stochastic differential equations, called g-expectation. Our constructive approach explains the driver “g” of g-expectations and provides a novel and elementary understanding of the risk-averse evolution equations and g-expectations based on nested risk measures. The new Hamilton–Jacobi–Bellman equations involve an new, additional drift term accounting for risk-aversion.

The approach presented elaborates a clear relation between static risk measures, dynamic risk measures in discrete time and the limiting risk measure in continuous time. In this way, the construction implies a consistent numerical procedure for solving risk-averse Hamilton–Jacobi–Bellman equations.

References

A. Ahmadi-Javid. Entropic Value-at-Risk: A New Coherent Risk Measure. *Journal of Optimization Theory and Applications*, 155(3):1105–1123, 2012. doi:10.1007/s10957-011-9968-2.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent Measures of Risk. *Mathematical Finance*, 9: 203–228, 1999. doi:10.1111/1467-9965.00068.

T. Breuer and I. Csiszár. Measuring Distribution Model Risk. *Mathematical Finance*, 2013. doi:10.1111/mafi.12050.

Ö. Çavuş and A. Ruszczyński. Risk-averse control of undiscounted transient markov models. *SIAM Journal on Control and Optimization*, 52(2):3935–3966, 2014. doi:10.1137/13093902X.

F. Coquet, Y. Hu, J. Mémin, and S. Peng. Filtration-consistent nonlinear expectations and related g-expectations. *Probability Theory and Related Fields*, 123:1–27, 2002. doi:10.1007/s004400100172.

M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277(1):1–1, jan 1983. doi:10.1090/s0002-9947-1983-0690039-8.

M. De Lara and V. Leclère. Building up time-consistency for risk measures and dynamic optimization. *European Journal of Operational Research*, 249:177–187, 2016. doi:10.1016/j.ejor.2015.03.046.

F. Delbaen. Coherent risk measures on general probability spaces. In *Essays in Honour of Dieter Sondermann*, pages 1–37. Springer-Verlag, Berlin, 2002.

D. Dentcheva and A. Ruszczyński. Time-coherent risk measures for continuous-time Markov chains. *SIAM Journal on Financial Mathematics*, 9(2):690–715, 2018. doi:10.1137/16m1063794.

J. Fan and A. Ruszczyński. Process-based risk measures for observable and partially observable discrete-time controlled systems. 2014. URL `http://arxiv.org/abs/1411.2675v1`.

W. H. Fleming and H. Soner. *Controlled Markov Processes and Viscosity Solutions*. 2006. ISBN 0-387-26045-5.

V. Guigues and W. Römisch. Sampling-based decomposition methods for multistage stochastic programs based on extended polyhedral risk measures. *SIAM Journal on Optimization*, 22(2):286–312, 2012. doi:10.1137/100811696.

W. B. Haskell and R. Jain. A convex analytic approach to risk-aware markov decision processes. *SIAM Journal on Control and Optimization*, 53(3):1569–1598, 2015. doi:10.1137/140969221.

W. B. Haskell, R. Jain, and D. Kalathil. Empirical dynamic programming. *Mathematics of Operations Research*, 41(2):402–429, 2016. doi:10.1287/moor.2015.0733.

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer, 1991. URL `http://books.google.com/books?id=ATNy_Zg3PSsC`.

I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Springer New York, 1998. doi:10.1007/978-1-4939-6845-9.

F. Maggioni, E. Allevi, and M. Bertocchi. Bounds in multistage linear stochastic programming. *Journal of Optimization Theory and Applications*, 163(1):200–229, 2012. doi:10.1007/s10957-013-0450-1.

W. Ogryczak and A. Ruszczyński. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13(1):60–78, 2002. doi:10.1137/S1052623400375075.

B. Øksendal. *Stochastic Differential Equations*. Springer Berlin Heidelberg, 2003. doi:10.1007/978-3-642-14394-6.

E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990. doi:10.1016/0167-6911(90)90082-6.

S. Peng. *Nonlinear Expectations, Nonlinear Evaluations and Risk Measures*. Springer, Berlin, Heidelberg, 2004. doi:10.1007/978-3-540-44644-6_4.

G. Ch. Pflug and A. Pichler. Time-consistent decisions and temporal decomposition of coherent risk functionals. *Mathematics of Operations Research*, 41(2):682–699, 2016. doi:10.1287/moor.2015.0747.

A. B. Philpott and V. L. de Matos. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218(2):470–483, 2012. doi:10.1016/j.ejor.2011.10.056.

A. Pichler and R. Schlotter. Entropy based risk measures. *European Journal of Operational Research*, jan 2019. doi:10.1016/j.ejor.2019.01.016.

A. Pichler and A. Shapiro. Risk averse stochastic programming: time consistency and optimal stopping.

F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185–200, 2004. doi:10.1016/j.spa.2004.03.004.

E. Rosazza Gianin. Risk measures via g-expectation. *Insurance Mathematics and Economics*, 39:19–34, 2006. doi:10.1016/j.insmatheco.2006.01.002.

A. Ruszczyński and A. Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31(3):544–561, 2006. doi:10.1287/moor.1060.0204.

A. Ruszczyński and J. Yao. A risk-averse analogue of the Hamilton–Jacobi–Bellman equation. In *Proceedings of the Conference on Control and its Applications*, chapter 62, pages 462–468. Society for Industrial and Applied Mathematics (SIAM), 2015. doi:10.1137/1.9781611974072.63.

L. Xin and A. Shapiro. Bounds for nested law invariant coherent risk measures. *Operation Research Letters*, 2011. doi:10.1016/j.orl.2012.09.002.

J. Zhang. *Backward Stochastic Differential Equations*. Springer New York, 2017. doi:10.1007/978-1-4939-7256-2.

A Appendix

Proof of a risk-averse Dynkin Formula

For convenience of the reader we restate the Lemma 25.

Lemma 37 (Dynkin's formula)**.** *Let $(X_s)_{s\in\mathcal{T}}$ be the solution of (15) on $\mathcal{T}$ with initial condition $X_t = x$. For $\Phi \in C^{1,2}(\mathcal{T}\times\mathbb{R})$ such that $\frac{\partial\Phi}{\partial x}$ is bounded, the risk-averse Dynkin formula*

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\Phi(r,X_r)\right) = \Phi(t,x) + \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r \mathcal{R}_\beta\Phi(s,X_s)\,\mathrm{d}s\right) \tag{42}$$

holds for any $t \le r \le T$.

Proof. The left side of (42) rewrites as

$$\Phi(t,x) + \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r \left(\frac{\partial\Phi}{\partial t} + b\frac{\partial\Phi}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2\Phi}{\partial x^2}\right)(s,X_s)\,\mathrm{d}s + \int_t^r \left(\sigma\frac{\partial\Phi}{\partial x}\right)(s,X_s)\,\mathrm{d}W_s\right)$$

using Itô's formula. Using the representation of the risk generator $\mathcal{R}_\beta$ in Equation (16) we can write the assertion as

$$\begin{aligned}&\mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r \left(\frac{\partial\Phi}{\partial t} + b\frac{\partial\Phi}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2\Phi}{\partial x^2}\right)(s,X_s)\,\mathrm{d}s + \int_t^r \left(\sigma\frac{\partial\Phi}{\partial x}\right)(s,X_s)\,\mathrm{d}W_s\right)\\ &= \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r \left(\frac{\partial\Phi}{\partial t} + b\frac{\partial\Phi}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2\Phi}{\partial x^2}\right)(s,X_s)\,\mathrm{d}s + \int_t^r \sqrt{2\beta(s)}\cdot\left|\left(\sigma\frac{\partial\Phi}{\partial x}\right)(s,X_s)\right|\,\mathrm{d}s\right).\end{aligned}$$

the representation of the risk generator in Proposition 23 shows that

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\Phi(r,X_r)\right) - \Phi(t,x) - \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r \mathcal{R}_\beta\Phi(s,X_s)\,\mathrm{d}s\right)$$

can be bounded by

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_t^r \sigma(s,X_s)\frac{\partial\Phi}{\partial x}(s,X_s)\,\mathrm{d}W_s - \int_t^r \sqrt{2\beta}\left|\sigma(s,X_s)\frac{\partial\Phi}{\partial x}(s,X_s)\right|\,\mathrm{d}s\right). \tag{43}$$

We will show that (43) is less than zero. For ease of notation, we now omit the arguments whenever there is no ambiguity. Let $n \in \mathbb{N}$. It follows from convexity that

$$(43) \le \sum_{i=0}^{n-1} \mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\int_{t_i}^{t_{i+1}} \sigma\frac{\partial\Phi}{\partial x}\,\mathrm{d}W_s - \int_{t_i}^{t_{i+1}} \sqrt{2\beta}\left|\sigma\frac{\partial\Phi}{\partial x}\right|\,\mathrm{d}s\right). \tag{44}$$

Moreover, monotonicity of nEV@R as well as Theorem 12 show that the summands of (44) are bounded by

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)}\left(\mathrm{EV@R}_{\beta_{t_i}\Delta t_i}\left[\int_{t_i}^{t_{i+1}} \sigma\frac{\partial\Phi}{\partial x}\,\mathrm{d}W_s - \Delta t_i\sqrt{2\beta_{t_i}}\left|\sigma\frac{\partial\Phi}{\partial x}(t_i,X_{t_i})\right| + o(\Delta t_i)\,\middle|\,\mathcal{F}_{t_i}\right]\right), \tag{45}$$

where $\beta_{t_i} \geq \beta(s)$ for all $s \in [t_i, t_{i+1}]$. We demonstrate that the inner conditional risk measure converges to zero fast enough. To this end, we argue similarly to the proof of Proposition 23 and split the stochastic integral in two parts as

$$\int_{t_i}^{t_{i+1}} \sigma \frac{\partial \Phi}{\partial x}(s, X_s)\, \mathrm{d}W_s = \int_{t_i}^{t_{i+1}} \sigma \frac{\partial \Phi}{\partial x}(t_i, X_{t_i})\, \mathrm{d}W_s + \int_{t_i}^{t_{i+1}} \sigma \frac{\partial \Phi}{\partial x}(s, X_s) - \sigma \frac{\partial \Phi}{\partial x}(t_i, X_{t_i})\, \mathrm{d}W_s.$$

Thus (45) can be bounded by

$$\begin{aligned} &\mathrm{nEV@R}^{t:r}_{\beta(\cdot)} \left(\mathrm{EV@R}_{\beta_{t_i} \Delta t_i} \left[\int_{t_i}^{t_{i+1}} \sigma \frac{\partial \Phi}{\partial x}(t_i, X_{t_i})\, \mathrm{d}W_s - \Delta t_i \sqrt{2\beta_{t_i}} \left| \sigma \frac{\partial \Phi}{\partial x}(t_i, X_{t_i}) \right| \,\middle|\, \mathcal{F}_{t_i} \right] \right) \\ &\quad + \mathrm{nEV@R}^{t:r}_{\beta(\cdot)} \left(\mathrm{EV@R}_{\beta_{t_i} \Delta t_i} \left[\int_{t_i}^{t_{i+1}} \sigma \frac{\partial \Phi}{\partial x}(s, X_s) - \sigma \frac{\partial \Phi}{\partial x}(t_i, X_{t_i})\, \mathrm{d}W_s + o(\Delta t_i) \,\middle|\, \mathcal{F}_{t_i} \right] \right). \end{aligned}$$

The first part is equal to zero and the argument in the proof of Proposition 23 shows that the second part tends to zero faster than linearly. In conclusion we obtain

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)} \left(\int_{t_i}^{t_{i+1}} \sigma \frac{\partial \Phi}{\partial x}\, \mathrm{d}W_s - \int_{t_i}^{t_{i+1}} \sqrt{2\beta} \left| \sigma \frac{\partial \Phi}{\partial x} \right| \mathrm{d}s \,\middle|\, \mathcal{F}_t \right) \leq \mathrm{nEV@R}^{t:r}_{\beta(\cdot)} \left(0 + o(\Delta t_i) \mid \mathcal{F}_t \right)$$

and hence taking the limit shows that

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)} \left(\int_t^r \sigma(s, X_s) \frac{\partial \Phi}{\partial x}(s, X_s)\, \mathrm{d}W_s - \int_t^r \sqrt{2\beta} \left| \sigma(s, X_s) \frac{\partial \Phi}{\partial x}(s, X_s) \right| \mathrm{d}s \right) \leq 0. \tag{46}$$

To accept that the left side of (46) is equal to zero, reverse the order of the arguments and repeat the above steps for

$$\mathrm{nEV@R}^{t:r}_{\beta(\cdot)} \left(\int_t^r \sqrt{2\beta} \left| \sigma(s, X_s) \frac{\partial \Phi}{\partial x}(s, X_s) \right| \mathrm{d}s - \int_t^r \sigma(s, X_s) \frac{\partial \Phi}{\partial x}(s, X_s)\, \mathrm{d}W_s \right) \leq 0,$$

which concludes the proof. □

Proof of Theorem 35

Again we restate the Theorem 35 for the convenience of the reader.

Theorem 38. *Suppose the assumptions of Proposition 30 as well as Assumption 22 are satisfied and suppose the control set U is compact. Then the optimal value function $V(\cdot,\cdot)$ is a viscosity solution of the equations (34).*

Proof. Let $w \in C_b^{1,2}([0,T] \times \mathbb{R})$ be such that $w \geq V$ on $[0,T] \times \mathbb{R}$ and

$$\min_{(t,x)} \{ w(t,x) - V(t,x) \} = 0.$$

Consider a point (t', x') such that $w(t', x') = V(t', x')$, let $h > 0$ and consider a constant control $u_s = v$ on $[t', t'+h]$. From Lemma 31 it follows that

$$\begin{aligned} V(t', x') &\leq \mathrm{nEV@R}^{t':t'+h}_{\beta(\cdot)} \left(\int_{t'}^{t'+h} c(s, X_s^{t',x',v}, v)\, \mathrm{d}s + V(t'+h, X_{t'+h}^{t',x',v}) \,\middle|\, X_{t'} = x' \right) \\ &\leq \mathrm{nEV@R}^{t':t'+h}_{\beta(\cdot)} \left(\int_{t'}^{t'+h} c(s, X_s^{t',x',v}, v)\, \mathrm{d}s + w(t'+h, X_{t'+h}^{t',x',v}) \,\middle|\, X_{t'} = x' \right). \end{aligned}$$

It follows from translation equivariance that

$$0 \leq \mathrm{nEV@R}_{\beta}^{t':t'+h}\left(\int_{t'}^{t'+h} c(s, X_s^{t',x',v}, v)\,\mathrm{d}s + w(t'+h, X_{t'+h}^{t',x',v}) - w(t',x') \middle| X_{t'} = x'\right). \tag{47}$$

By assumption $w \in C_b^{1,2}$ and hence Itô's formula holds. Furthermore Theorem 32 can be applied and hence

$$\frac{1}{h}\,\mathrm{nEV@R}_{\beta(\cdot)}^{t':t'+h}\left(\int_{t'}^{t'+h} c(s, X_s^{t',x',v}, v)\,\mathrm{d}s + w(t'+h, X_{t'+h}^{t',x',v}) - w(t',x') \middle| X_{t'} = x'\right)$$

converges to

$$c(t,x,v) + \frac{\partial w}{\partial t}(t,x) + b(t,x,v)\frac{\partial w}{\partial x}(t,x,v) + \frac{1}{2}\sigma^2(t,x,v)\cdot\frac{\partial^2 w}{\partial x^2}(t,x,v) + \sqrt{2\beta(t)}\sigma(t,x,v)\left|\frac{\partial w}{\partial x}(t,x,v)\right|.$$

Since the constant control v was arbitrary it follows that

$$0 \leq \frac{\partial w}{\partial t}(t,x) - \mathcal{H}\left(t, x, \frac{\partial w}{\partial x}, \frac{\partial^2 w}{\partial x^2}\right)$$

and hence V is a viscosity subsolution.

Now, let $w \in C_b^{1,2}([t, T \times \mathbb{R})$ be such that $w \leq V$ on $[0,T]\times\mathbb{R}$, and $\min_{(t,x)}[V(t,x) - w(t,x)] = 0$. Consider a point $(t',x') \in [0,T)\times\mathbb{R}$ such that $w(t',x') = V(t',x')$. Let $u(\cdot) \in U[t', t'+h]$ be an εh-optimal control in (34) at $(t,x) = (t',x')$. Proceeding exactly as in the derivation of (47), we obtain the inequality:

$$\mathrm{nEV@R}_{\beta(\cdot)}^{t':t'+h}\left(\int_{t'}^{t'+h} c(s, X_s^{t',x',v}, u)\,\mathrm{d}s + w(t'+h, X_{t'+h}^{t',x',u}) - w(t',x') \middle| X_{t'} = x'\right) \leq \varepsilon h.$$

Therefore we also have

$$\min_{u \in U[t',t'+h]} \frac{1}{h}\,\mathrm{nEV@R}_{\beta(\cdot)}^{t':t'+h}\left(\int_{t'}^{t'+h} c(s, X_s^{t',x',u}, u)\,\mathrm{d}s + w(t'+h, X_{t'+h}^{t',x',u}) - w(t',x') \middle| X_{t'} = x'\right) \leq \varepsilon$$

and letting h tend to zero, we see with Theorem 32 that

$$\inf_{u \in U}\left\{c(t,x,u) + \frac{\partial V}{\partial t}(t,x) + b(t,x,u)\frac{\partial V}{\partial x}(t,x) + \frac{\sigma^2(t,x,u)}{2}\frac{\partial^2 V}{\partial x^2}(t,x) + \sqrt{2\beta(t)}\left|\sigma(t,x,u)\frac{\partial V}{\partial x}(t,x)\right|\right\} \leq \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we conclude that V is a viscosity supersolution, which completes the proof. □