



# Set-valued risk measures as backward stochastic difference inclusions and equations

Çağın Ararat<sup>1</sup> · Zachary Feinstein<sup>2</sup>

Received: 14 December 2019 / Accepted: 23 October 2020

© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2020

**Abstract** Scalar dynamic risk measures for univariate positions in continuous time are commonly represented via backward stochastic differential equations. In the multivariate setting, dynamic risk measures have been defined and studied as families of set-valued functionals in the recent literature. There are two possible extensions of scalar backward stochastic differential equations for the set-valued framework: (1) backward stochastic differential inclusions, which evaluate the risk dynamics on the selectors of acceptable capital allocations; or (2) set-valued backward stochastic differential equations, which evaluate the risk dynamics on the full set of acceptable capital allocations as a singular object. In this work, the discrete-time setting is investigated with difference inclusions and difference equations in order to provide insights for such differential representations for set-valued dynamic risk measures in continuous time.

**Keywords** Set-valued risk measure · Dynamic risk measure · Difference inclusion · Set-valued difference equation · Time-consistency

**Mathematics Subject Classification (2010)** 26E25 · 28B20 · 34A60 · 39A50 · 46A20 · 91B30

**JEL Classification** D81 · G32

---

✉ Ç. Ararat  
[cararat@bilkent.edu.tr](mailto:cararat@bilkent.edu.tr)

Z. Feinstein  
[zfeinste@stevens.edu](mailto:zfeinste@stevens.edu)

<sup>1</sup> Department of Industrial Engineering, Bilkent University, Ankara, Turkey

<sup>2</sup> School of Business, Stevens Institute of Technology, Hoboken, NJ, USA

# 1 Introduction

## 1.1 Literature review

The seminal work of Artzner et al. [3] introduced coherent risk measures in an axiomatic framework to provide capital requirements for financial portfolios. The coherence axiom was relaxed to convexity in Föllmer and Schied [24], [25, Chap. 4]. Each of these works considers risk measured at a single time point for frictionless claims. These convex and coherent risk measures were placed in a dynamic system in which the capital requirements for portfolios and contingent claims are updated in time as new information becomes available. The relation of these risks over time is a key property of study; this so-called time-consistency property has been studied in e.g. Bion-Nadal [8], Cheridito et al. [11], Detlefsen and Scandolo [14], Riedel [46], Ruszczyński and Shapiro [48].

For the purposes of this work, the relevant literature is more specialised. It has been noted for almost as long as dynamic risk measures have been studied that certain nonlinear ( $g$ -)expectations have all the properties of a time-consistent convex risk measure, and vice versa. In this way, convex risk measures can be represented as solutions to backward stochastic differential equations (see e.g. Barrieu and El Karoui [4, 5], Jiang [32], Peng [44], Rosazza Gianin [47]). This representation allows the detailed study of dynamic risk measures in continuous time as well as their efficient computation. We take our motivation for this work from Stadje [49] who considers the problem of representing a *discrete*-time dynamic risk measure as the solution of a backward stochastic difference equation. That work further provides the  $L^2$ -convergence of the drivers and solutions of backward stochastic difference equations to the corresponding backward stochastic differential equation with desired driver, as the number of time steps (on a finite time horizon) grows to infinity.

Scalar risk measures have been extended to consider multivariate portfolios as well. Often, and with particular relevance to this work, they are studied as set-valued risk measures. Such set-valued risk measures were first introduced in a coherent setting in Jouini et al. [33]. This was extended to consider convex risk measures in Hamel et al. [27]. Such risk measures have been applied, primarily, in two settings:

- *Portfolios with frictions*: Due to transaction costs, valuing a portfolio by a numéraire leads to a non-unique value (e.g. mark-to-market or liquidation value). Each choice of valuation by a numéraire results in an intrinsic misspecification of risk. For example, mark-to-market valuation provides an overestimate of the true value of the portfolio and thus an underestimate of the risk; liquidation valuation provides an underestimate of the true value of the portfolio and thus an overestimate of the risk. A market with frictions is the prototypical case for set-valued risk measures as studied in Hamel et al. [27]. Specifically, due to transaction and market impact costs, there is a misspecification of value in using any numéraire to hedge a risky portfolio. In order to compensate for these frictions, capital requirements in that specific numéraire must be larger than the current value of the hedging portfolio. Set-valued risk measures approach this problem by considering the set of all risk-compensating portfolios, for which there is unlikely to be a uniquely minimal choice, so as to more accurately capture the riskiness of the original portfolio. In order to provide a numéraire-free

representation, a portfolio is represented by a random vector in physical units, i.e., the  $i$ th component of the random vector denotes the number of asset  $i$  held in the portfolio.

– *Systemic risk*: Due to the interaction between banks, the capital requirements for any institution will ultimately depend on the health of all other institutions. This is presented in a set-valued setting in Biagini et al. [7], Feinstein et al. [23]. Specifically, capital *cannot* be freely transferred between banks, but due to e.g. default contagion (see Eisenberg and Noe [15]), the capital requirements for one institution intrinsically depend on the health of all other banks. Due to this coupling of the health of all banks in the system, capital requirements for the different institutions cannot be determined independently. In particular, this generates an efficient frontier of capital requirements for all banks in the financial system, and as such, such objects are naturally described by the set of all acceptable (system-wide) capital requirements. The collection of wealths for the financial system is presented as a random vector, i.e., the  $i$ th component of the random vector denotes the wealth of bank  $i$  in the system.

There are analogues of many scalar risk measures in the set-valued framework. For instance, superhedging was considered in Hamel et al. [27], Löhne and Rudloff [41], and shortfall and divergence risk measures were considered in Ararat et al. [2].

The typical analogue of time-consistency in the set-valued setting is “multiportfolio time-consistency” as first defined in Feinstein and Rudloff [16]. This was shown to be equivalent to a variation of the time-consistency property in Ben Tahar and Lépinette [6] for the random set approach in Feinstein and Rudloff [18]. Additional properties for multiportfolio time-consistency were studied in e.g. Feinstein and Rudloff [17, 20–22] and have been utilised directly for computing such risk measures in discrete time in Feinstein and Rudloff [19]. Though not the focus of this work, multiportfolio time-consistency has been extended to risk measures for processes in Chen and Hu [10].

Time-consistency for risk measures more generally relates to the dynamic programming principle from optimisation. Optimal control problems are typically formulated using backward stochastic differential equations in continuous time. In finance, the relation between time-consistency, the dynamic programming principle and backward stochastic differential equations was highlighted in e.g. Karnam et al. [34], Lépinette and Molchanov [40]. By considering multiportfolio time-consistency and the relation to the dynamic programming principle as espoused in Feinstein and Rudloff [19], we wish to further investigate this relationship for set-valued processes.

## 1.2 Motivation

The backward stochastic differential equation (BSDE) representation of scalar risk measures in continuous time is fundamental to our understanding of how risks propagate over time. We briefly recall this representation. Under the augmented filtration of a standard  $m$ -dimensional Brownian motion  $W = (W_t)_{t \in [0, T]}$ , let us consider the solution  $(Y^X, Z^X)$  of the BSDE

$$Y_t^X = -X + \int_t^T g(s, Z_s^X) ds - \int_t^T Z_s^X dW_s, \quad t \in [0, T], \quad (1.1)$$

for each terminal value  $X$  in the  $L^2$ -space of  $\mathcal{F}_T$ -measurable (univariate) random variables. As shown in Rosazza Gianin [47], the family of these solutions indexed by the terminal value gives rise to a time-consistent dynamic risk measure via the mapping  $X \mapsto \rho_t(X) := Y_t^X$ , provided that the driver function  $g: \Omega \times [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  is convex in the last variable and satisfies further regularity properties (e.g. Lipschitz in the last variable) to ensure the existence of a unique solution for the BSDE. Conversely, a given dynamic time-consistent convex risk measure  $\rho = (\rho_t)_{t \in [0, T]}$  for which  $\rho_0$  satisfies two additional conditions (i.e., strict monotonicity and dominatedness, see Rosazza Gianin [47, Proposition 20] for details) can be represented as  $\rho_t(X) = Y_t^X$  through a family of BSDEs with a common driver  $g$  as described above.

In contrast, no such representation is known for set-valued dynamic risk measures for random vectors. As a starting point, one can restrict attention to set-valued risk measures that can be written as the sum of a vector-valued function and the cone of positive random vectors, which is essentially equivalent to defining risk measures as vector-valued functions. In this case, as studied in Xu [50], a natural candidate that extends (1.1) is a multidimensional BSDE which has the same form as (1.1), but with a vector-valued driver function  $g$  and a vector-valued terminal value  $X$ . However, the only vector-valued functions that can be seen as risk measures (i.e., monotone, translatable, convex) are the trivial ones whose coordinates are risk measures on individual coordinate spaces; this was observed in Xu [50] in the dynamic setting via BSDE arguments and a more direct proof in the static setting was given recently in Molchanov and Mühlemann [43]. Hence, multidimensional BSDEs have very limited use for the representation of dynamic risk measures in the multivariate framework. This motivates us to consider more general structures that we describe next.

In set-valued stochastic analysis, stochastic differential inclusions and set-valued stochastic differential equations appear as natural ways of extending stochastic differential equations to the set-valued framework. In the forward case, with the developments in set-valued stochastic integration, a theory of such inclusions and set-valued equations was developed in the last two decades; the reader is referred to Kisielewicz [37, Chap. 4] and the references therein for a detailed treatment. For set-valued risk measures, we consider the versions of these two structures that move backward in time:

- *Backward stochastic differential inclusions*: The vector-valued selectors of a set-valued process propagate backward in time with the path of the specific solution being of fundamental importance; this is a suitable choice if one is interested in characterising a path moving inside the dynamic set-valued risk measure.

- *Set-valued backward stochastic differential equations*: The entire set-valued process propagates backward in time with no specific selector itself deemed to be important; this is a more sophisticated structure as it characterises the entire set-valued process that corresponds to the dynamic set-valued risk measure. It is the understanding of these representations that motivates us in our current work.

Despite the considerable literature on forward inclusions and set-valued equations, at least in the case of driver functions with compact values, their backward counterparts are quite new in the literature. The few works dealing with backward stochastic differential inclusions are due to Kisielewicz [35, 36], who also assumed compact

values for the driver. To the best of our knowledge, set-valued backward stochastic differential equations have not been studied in the literature yet. Hence, though we are motivated by the continuous-time representations of risk measures described above, such differential inclusions and set-valued differential equations fall beyond the scope of this work.

In order to begin understanding the continuous-time setting, we investigate the discrete-time approximation. As such, in this work, we consider backward stochastic difference inclusions (Sect. 4) and set-valued backward stochastic difference equations (Sect. 5) in order to understand fundamental questions about the representation of dynamic risk measures via a differential system, e.g. what mathematical form such a representation may take and which approach (inclusion or set-valued equation) is more appropriate. This permits us to both begin the study of continuous-time multivariate risk measures and at the same time direct the theoretical research towards usable applications.

It should be noted that the fundamental BSDE representation in (1.1), which works under the filtration of a multidimensional Brownian motion, already covers a rich class of applications; see for instance Rosazza Gianin [47, Sect. 3]. In our discrete-time framework, we consider a multidimensional Bernoulli random walk as a standard discretisation of Brownian motion. This is in line with the approach in Stadje [49] for univariate risk measures. The multidimensional random walk models the underlying source of randomness and admits a predictable representation property in a generalised sense. This property is the key to having the representations via difference inclusions in Sect. 4 and via set-valued difference equations in Sect. 5. In Remark 3.3, we outline a more general setting where the Bernoulli random walk is replaced with a multidimensional square-integrable martingale in discrete time. In this case, a more abstract version of the predictable representation property is available and can be used in order to extend the results of this paper.

Beyond the immediate applications for risk measures, BSDEs are strongly related with the dynamic programming principle. The mean–variance and mean–risk problems are well known to generally be time-inconsistent and do *not* follow the dynamic programming principle, as noted in Karnam et al. [34]. However, by considering a multiobjective formulation, Kováčová and Rudloff [38] have found that the dynamic programming principle in the mean–risk problem is satisfied in discrete time. Therefore, the insights we gain on expanding the BSDE representation of the dynamic programming principle to multivariate problems through either backward stochastic differential inclusions or set-valued backward stochastic differential equations is of wider interest and importance.

This work has two primary contributions we wish to highlight. First, we present backward stochastic difference inclusion and set-valued backward stochastic difference equation representations for set-valued dynamic risk measures. These are found in Sects. 4 and 5, respectively. Second, because neither inclusion nor set-valued equation representations have been studied in applications, we use set-valued risk measures in order to determine which generalisation of BSDEs is a more promising avenue for considering the dynamic programming principle. Specifically, as highlighted in this work, backward stochastic differential inclusions appear to be the proper methodology for future studies of the dynamic programming principle.

## 2 Background

Fix  $T > 0$  and let  $\mathbb{T}$  be a finite subset of  $[0, T]$  with  $T \in \mathbb{T}$ . Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{T}}, \mathbb{P})$  be a filtered probability space and  $d \geq 1$  an integer. Let  $\mathcal{G}$  be a sub- $\sigma$ -algebra of  $\mathcal{F}$ . We denote by  $L_d^0(\mathcal{G})$  the space of all  $\mathcal{G}$ -measurable random variables  $X: \Omega \rightarrow \mathbb{R}^d$ , where two elements are distinguished only up to  $\mathbb{P}$ -almost sure equality. For  $p \in [1, +\infty)$ , we denote by  $L_d^p(\mathcal{G})$  the set of all  $X \in L_d^0(\mathcal{G})$  for which  $\mathbb{E}[|X|^p] < +\infty$ , where  $|\cdot|$  is a fixed norm on  $\mathbb{R}^d$ . We denote by  $L_d^\infty(\mathcal{G})$  the set of all  $X \in L_d^0(\mathcal{G})$  for which  $|X|$  is  $\mathbb{P}$ -essentially bounded. For ease of notation, **throughout the rest of this paper**, we define  $L_{d,+}^p(\mathcal{G}) := \{X \in L_d^p(\mathcal{G}): \mathbb{P}[X \in \mathbb{R}_+^d] = 1\}$ .

For  $C, D \subseteq L_d^0(\mathcal{F}_T)$ , the set  $C + D := \{X + Y: X \in C, Y \in D\}$  is the Minkowski sum of  $C, D$ . If  $Z \in L_1^0(\mathcal{F}_T)$ , then  $ZC := CZ := \{ZX: X \in C\}$  is the set of pointwise products.

Let  $p \in [1, +\infty]$ . As defined in Feinstein and Rudloff [16, 17], a *conditional risk measure at time  $t \in \mathbb{T}$*  is a mapping

$$R_t: L_d^p(\mathcal{F}_T) \rightarrow \mathcal{P}_+(L_d^p(\mathcal{F}_t)) := \{C \subseteq L_d^p(\mathcal{F}_t): C = C + L_{d,+}^p(\mathcal{F}_t)\}$$

of random vectors into upper sets. As described above, we consider risk measures as functions of random vectors; the interpretation of these vectors depends on the context of the problem, i.e., markets with frictions, systemic risk, etc. If the input can be aggregated into a single random variable ( $d = 1$ ) without loss of information, then univariate risk measures (Riedel [46]) can be applied instead. Throughout this work, we focus solely on *normalised convex risk measures*, i.e., those that are

1) *normalised*:  $R_t(0) \neq \emptyset$ ,  $\mathbb{P}[\tilde{R}_t(0) = \mathbb{R}^d] = 0$  (where  $\tilde{R}_t(0)$  is an  $\mathcal{F}_t$ -measurable random set such that  $R_t(0) = \{u \in L_d^p(\mathcal{F}_t): \mathbb{P}[u \in \tilde{R}_t(0)] = 1\}$  and which exists due to the properties below and Molchanov [42, Theorem 2.1.10]), and for any  $X \in L_d^p(\mathcal{F}_T)$ ,  $R_t(X) = R_t(0) + R_t(X)$ ;

2) *translative*:  $R_t(X + u) = R_t(X) - u$  for any  $X \in L_d^p(\mathcal{F}_T)$  and  $u \in L_d^p(\mathcal{F}_t)$ ;

3) *monotone*:  $R_t(X) \subseteq R_t(Y)$  for any  $X, Y \in L_d^p(\mathcal{F}_T)$  with  $Y - X \in L_{d,+}^p(\mathcal{F}_T)$ ;

4) *conditionally convex*:

$$R_t(\lambda X + (1 - \lambda)Y) \supseteq \lambda R_t(X) + (1 - \lambda)R_t(Y)$$

for any  $X, Y \in L_d^p(\mathcal{F}_T)$  and  $\lambda \in L_1^p(\mathcal{F}_t)$  with  $\mathbb{P}[0 \leq \lambda \leq 1] = 1$ ;

5) *closed*: the set

$$\text{graph } R_t := \{(X, u) \in L_d^p(\mathcal{F}_T) \times L_d^p(\mathcal{F}_t): u \in R_t(X)\}$$

is closed in the product topology, and as a consequence,  $R_t(X)$  is closed for every  $X \in L_d^p(\mathcal{F}_T)$ . A *dynamic risk measure*  $R = (R_t)_{t \in \mathbb{T}}$  is a sequence of conditional risk measures over time.

Let  $R = (R_t)_{t \in \mathbb{T}}$  be a dynamic risk measure. By construction (see e.g. Feinstein and Rudloff [16, Proposition 2.8] and Feinstein and Rudloff [18, Lemma 3.6]),  $R$  is *decomposable*, i.e.,

$$R_t(\mathbb{1}_A X + \mathbb{1}_{A^c} Y) = \mathbb{1}_A R_t(X) + \mathbb{1}_{A^c} R_t(Y)$$

for every  $X, Y \in L_d^p(\mathcal{F}_T)$ ,  $A \in \mathcal{F}_t$ ,  $t \in \mathbb{T}$ , where  $\mathbb{1}_A$  denotes the probabilistic indicator function of the event  $A$ .

*Remark 2.1* Although we present all results in this paper for set-valued risk measures, these results apply equally for utility functionals  $U_t : L_d^p(\mathcal{F}_T) \ni X \mapsto -R_t(X)$ . In Hamel and Wang [29], a (nontranslative) set-valued utility functional is defined based on expected utility and the associated utility maximisation problem is studied. An axiomatic treatment of (translative) set-valued utility functionals defined on a set of closed convex random sets is given in Molchanov and Mühlemann [43] under the name *superlinear set-valued expectations*; coherent set-valued risk measures for random vectors can be embedded into that framework. As our constructions rely on some earlier results on set-valued risk measures (e.g. Feinstein and Rudloff [16, 18]), we consider these functionals in the current work for consistency with the literature most relevant for us.

With the introduction of dynamic risk measures, the manner in which risks change over time is fundamentally important. As studied in Ben Tahar and Lépinette [6], Chen and Hu [10], Feinstein and Rudloff [16, 19], multiportfolio time-consistency provides the appropriate definition from a mathematical and computational perspective. A dynamic risk measure  $R = (R_t)_{t \in \mathbb{T}}$  is said to be *multiportfolio time-consistent* if for any  $X \in L_d^p(\mathcal{F}_T)$  and  $\mathbb{Y} \subseteq L_d^p(\mathcal{F}_T)$ ,

$$R_s(X) \subseteq R_s[\mathbb{Y}] := \bigcup_{Y \in \mathbb{Y}} R_s(Y) \quad \implies \quad R_t(X) \subseteq R_t[\mathbb{Y}]$$

for all times  $t, s \in \mathbb{T}$  with  $t < s$ . To clarify the notation, the functions  $R_t(\cdot)$  are applied to random vectors, whereas the  $R_t[\cdot]$  are applied to sets of random vectors as the union over each element as defined above. Multiportfolio time-consistency states that if every risk compensating portfolio for  $X$  at time  $s$  also compensates for the risk of some portfolio  $Y \in \mathbb{Y}$ , then that same relation holds for any time  $t < s$ . That is, the risk of *baskets* of portfolios is consistent over time. This version of time-consistency is related to the dynamic programming principle through the recursive relation

$$R_t(X) = R_{t,s}[-R_s(X)] \quad (2.1)$$

for every  $X \in L_d^p(\mathcal{F}_T)$ ,  $t, s \in \mathbb{T}$  with  $t < s$ . This recursive formulation is constructed via the stepped risk measure  $R_{t,s} : L_d^p(\mathcal{F}_s) \rightarrow \mathcal{P}_+(L_d^p(\mathcal{F}_t))$  which is defined as the restriction of the conditional risk measure  $R_t$  to the domain  $L_d^p(\mathcal{F}_s) \subseteq L_d^p(\mathcal{F}_T)$ , i.e.,

$$R_{t,s}(u) = R_t(u)$$

for every  $u \in L_d^p(\mathcal{F}_s)$ .

To consider local versions of the difference equations and inclusions, it is preferable to consider the random set approach to conditional risk measures. By the equivalence between closed decomposable sets of random vectors and closed random sets (see e.g. Molchanov [42, Theorem 2.1.10]) as applied to set-valued risk measures in Feinstein and Rudloff [18], there exists a random set  $\tilde{R}_t(X)$  (i.e.,  $\tilde{R}_t(X)(\omega) \subseteq \mathbb{R}^d$  for every  $\omega \in \Omega$ ) such that

$$R_t(X) = \{u \in L_d^p(\mathcal{F}_t) : \mathbb{P}[u \in \tilde{R}_t(X)] = 1\}$$

for every  $X \in L_d^p(\mathcal{F}_T)$ ,  $t \in \mathbb{T}$ . As the rest of the work is concerned solely with a finitely generated filtration, we abuse notation to define

$$R_t(X)(\omega) := \tilde{R}_t(X)(\omega)$$

for every  $X \in L_d^p(\mathcal{F}_T)$ ,  $t \in \mathbb{T}$ ,  $\omega \in \Omega$ . These definitions are given likewise for the stepped risk measures  $R_{t,s}$  with  $t, s \in \mathbb{T}$  and  $t < s$ .

### 3 Discrete-time setting

In this section, we present the mathematical notation and setting for the discrete-time stochastic processes of interest. In particular, we introduce an  $m$ -dimensional Bernoulli random walk and consider the predictable representation property with respect to this process. We present these results for completeness and to assist in notation later in this paper; more details can be found in Protter [45, Sect. IV.3]. We wish to highlight the need for the orthogonal martingale term when considering an  $m$ -dimensional Bernoulli random walk with  $m > 1$ , which complicates the setting from e.g. Stadje [49]. Such considerations are similarly required if a different random walk is chosen as the driving process; see Remark 3.3 for more details. We focus on the Bernoulli random walk both because it presents a simple mathematical setting that allows an explicit orthogonal process and because of its ubiquity as a discrete-time representation whose limit is Brownian motion.

Let  $T > 0$  be a fixed time. Let  $m, K \geq 1$  be integers and consider the discrete time set  $\mathbb{T} = \{t_0, \dots, t_K\}$ , where  $0 = t_0 < t_1 < \dots < t_{K-1} < t_K = T$ . Let us write  $\Delta t_k := t_k - t_{k-1}$  for  $k \in \{1, \dots, K\}$ . Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space on which there exist independent  $m$ -dimensional random vectors  $B_1, \dots, B_K$  with values in  $\{-1, +1\}^m$ , where the components of  $B_k = (B_k^1, \dots, B_k^m)^\top$  are possibly correlated symmetric Rademacher random variables for  $k \in \{1, \dots, K\}$ . We define an  $m$ -dimensional random walk  $M = (M_t^1, \dots, M_t^m)_{t \in \mathbb{T}}$  by  $M_{t_0}^i := 0$  and

$$M_{t_k}^i := M_{t_{k-1}}^i + \sqrt{\Delta t_k} B_k^i \quad (3.1)$$

for  $k \in \{1, \dots, K\}$  and  $i \in \{1, \dots, m\}$ . For notational simplicity, set

$$\Delta M_{t_k} := M_{t_k} - M_{t_{k-1}} = \sqrt{\Delta t_k} B_k.$$

Let  $(\mathcal{F}_t^M)_{t \in \mathbb{T}}$  be the natural filtration of  $M$ , that is,  $\mathcal{F}_{t_0}^M := \{\emptyset, \Omega\}$  and

$$\mathcal{F}_{t_k}^M := \sigma(M_{t_1}, \dots, M_{t_k}) = \sigma(B_1, \dots, B_k)$$

for  $k \in \{1, \dots, K\}$ .

Let  $k \in \{1, \dots, K\}$ . Since  $\mathcal{F}_{t_k}^M$  is generated by finitely many events, every real-valued random variable that is measurable with respect to  $\mathcal{F}_{t_k}^M$  is bounded. Hence,  $L_d^0(\mathcal{F}_{t_k}^M) = L_d^p(\mathcal{F}_{t_k}^M)$  for  $p \in [1, +\infty]$ . To simplify notation for the remainder of this work, we define

$$\mathcal{L}_{t_k}^d := L_d^0(\mathcal{F}_{t_k}^M).$$



When  $m = 1$ , every  $X \in \mathcal{L}_{t_k}^d$  can be written as

$$X = \xi + \psi \Delta M_{t_k} \quad (3.2)$$

for some  $\xi, \psi \in \mathcal{L}_{t_{k-1}}^d$  (see e.g. Föllmer and Schied [25, Theorem 5.38]). This is called the *predictable representation property* of the one-dimensional random walk  $M$ . For  $m \geq 2$ , the analogous representation

$$\xi + \sum_{i=1}^m \psi^i \Delta M_{t_k}^i \quad (3.3)$$

with  $\xi, \psi^1, \dots, \psi^m \in \mathcal{L}_{t_{k-1}}^d$  fails to equal  $X \in \mathcal{L}_{t_k}^d$  in general.

Instead of (3.3), a proper generalisation of (3.2) should take into account the “cross-terms” created by the  $m$  components of the vector-valued random walk  $M$  (i.e.,  $M^1, \dots, M^m$ ). To that end, let us denote by  $\mathcal{I}$  the set of all nonempty subsets of  $\{1, \dots, m\}$ . For each  $I \in \mathcal{I}$ , we define a process  $M^I = (M_t^I)_{t \in \mathbb{T}}$  by  $M_{t_0}^I := 0$  and

$$\Delta M_{t_k}^I := M_{t_k}^I - M_{t_{k-1}}^I = \sqrt{\Delta t_k} B_k^I, \quad \text{where } B_k^I := \prod_{i \in I} B_k^i$$

for  $k \in \{1, \dots, K\}$ . In addition to the earlier processes  $M^{\{i\}} = M^i$ ,  $i \in \{1, \dots, m\}$ , this definition creates  $2^m - m - 1$  new processes. Moreover, when the components of  $B_k$  are independent for each  $k \in \{1, \dots, K\}$ , all the  $2^m - 1$  processes are martingales that are orthogonal. The new form of predictable representation is stated in the next lemma.

**Lemma 3.1** *Let  $k \in \{1, \dots, K\}$ . Every  $Y \in \mathcal{L}_{t_k}^d$  can be written as*

$$Y = \xi + \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I$$

for some  $\xi, \psi^I \in \mathcal{L}_{t_{k-1}}^d$ ,  $I \in \mathcal{I}$ .

**Remark 3.2** For  $d = 1$ , the result appears as Lemma 6.1 of a preprint version of Cheridito et al. [12]; its proof there is based on a spanning argument for the finite-dimensional vector space  $\mathcal{L}_{t_k}^1$  assuming  $k = 1$ . We provide a more elementary and complete proof below with an explicit derivation of the predictable representation. Unlike the proof in the preprint version of Cheridito et al. [12], our proof is also valid without assuming that the components of the random walk are independent.

**Proof of Lemma 3.1** Assume that  $d = 1$ . As a first step, let  $Y \in L_1^0(\sigma(B_k))$ . Note that  $\sigma(B_k)$  is generated by  $2^m$  events of the form  $\{B_k = b\}$  with vectors  $b$  of the form  $b = (b^1, \dots, b^m) \in \{-1, +1\}^m$  that partition  $\Omega$ . Hence we may write

$$Y = \sum_{b \in \{-1, +1\}^m} c_b \mathbb{1}_{\{B_k = b\}} = \sum_{b \in \{-1, +1\}^m} c_b \prod_{i=1}^m \mathbb{1}_{\{B_k^i = b^i\}}$$

for some  $c_b \in \mathbb{R}$ , with  $b \in \{-1, +1\}^m$ . Note that  $\mathbb{1}_{\{B_k^i = b^i\}} = (1 + b^i B_k^i)/2$  for  $i \in \{1, \dots, m\}$  and  $b \in \{-1, +1\}^m$ . Hence,

$$Y = \sum_{b \in \{-1, +1\}^m} c_b \prod_{i=1}^m \frac{1 + b^i B_k^i}{2}.$$

One can rewrite the above sum in the form of a polynomial of  $B_k^1, \dots, B_k^m$ . The constant term of this polynomial is

$$\xi := \frac{1}{2^m} \sum_{b \in \{-1, +1\}^m} c_b.$$

On the other hand, each functional term of the polynomial is of the form  $\tilde{\psi}^I B_k^I$ , where  $I \in \mathcal{I}$  and

$$\tilde{\psi}^I := \frac{1}{2^m} \sum_{b \in \{-1, +1\}^m} c_b \prod_{i \in I} b^i.$$

Taking  $\psi^I := \tilde{\psi}^I / \sqrt{\Delta t_k}$  for  $I \in \mathcal{I}$ , we obtain

$$Y = \xi + \sum_{I \in \mathcal{I}} \tilde{\psi}^I B_k^I = \xi + \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I.$$

If  $k = 1$ , the conclusion of the lemma follows by the first step since  $\mathcal{F}_{t_1}^M = \sigma(B_1)$ . Next, suppose that  $k \in \{2, \dots, K\}$  and let  $Y \in \mathcal{L}_{t_k}^1$ . Hence, we may write

$$Y = f \circ (B_1, \dots, B_k)$$

for some (Borel-measurable) function  $f: \{-1, +1\}^{km} \rightarrow \mathbb{R}$ . In particular,

$$Y = \sum_{b_1 \in \{-1, +1\}^m} \dots \sum_{b_{k-1} \in \{-1, +1\}^m} f \circ (b_1, \dots, b_{k-1}, B_k) \mathbb{1}_{\{B_1 = b_1, \dots, B_{k-1} = b_{k-1}\}}. \quad (3.4)$$

Let  $b_1, \dots, b_{k-1} \in \{-1, +1\}^m$ . Since  $f \circ (b_1, \dots, b_{k-1}, B_k) \in L_1^0(\sigma(B_k))$ , by the first step, there exist  $g(b_1, \dots, b_{k-1}) \in \mathbb{R}$ ,  $h^I(b_1, \dots, b_{k-1}) \in \mathbb{R}$  for each  $I \in \mathcal{I}$  such that

$$f \circ (b_1, \dots, b_{k-1}, B_k) = g(b_1, \dots, b_{k-1}) + \sum_{I \in \mathcal{I}} h^I(b_1, \dots, b_{k-1}) \Delta M_{t_k}^I. \quad (3.5)$$

Then (3.4) and (3.5) imply that

$$Y = g \circ (B_1, \dots, B_{k-1}) + \sum_{I \in \mathcal{I}} h^I \circ (B_1, \dots, B_{k-1}) \Delta M_{t_k}^I.$$

The functions  $g, h^I: \{-1, +1\}^{(k-1)m} \rightarrow \mathbb{R}$ ,  $I \in \mathcal{I}$ , are Borel-measurable since they are defined on finite sets. Hence by taking

$$\xi := g \circ (B_1, \dots, B_{k-1}) \in \mathcal{L}_{t_{k-1}}^1, \quad \psi^I := h^I \circ (B_1, \dots, B_{k-1}) \in \mathcal{L}_{t_{k-1}}^1, \quad I \in \mathcal{I},$$

the conclusion of the lemma follows for  $d = 1$ .

For arbitrary  $d \geq 1$ , applying the result for  $d = 1$  to each component of  $Y \in \mathcal{L}_{t_k}^d$  gives the claimed representation.  $\square$

**Remark 3.3** One can consider a more general setting in which  $M$  is a square-integrable martingale such as a symmetric random walk with an arbitrary distribution for its increments; e.g.  $M$  can be a symmetric Gaussian random walk. In this case, the simple predictable representation in (3.3) fails to hold in general even for  $m = 1$ . Nevertheless, when the  $m$  components of  $M$  are independent, by the results in Protter [45, Sect. IV.3], one has the more general form

$$X = \xi + \sum_{i=1}^m \psi^i \Delta M_{t_k}^i + \Delta N_{t_k}, \quad (3.6)$$

where  $\xi, \psi^1, \dots, \psi^m \in \mathcal{L}_{t_{k-1}}^d$ ,  $\Delta N_{t_k} = N_{t_k} - N_{t_{k-1}}$  and  $N = (N_t^1, \dots, N_t^d)_{t \in \mathbb{T}}$  is a square-integrable martingale orthogonal to  $M$ , i.e.,  $\mathbb{E}[\Delta M_{t_k}^i \Delta N_{t_k}^j | \mathcal{F}_{t_{k-1}}] = 0$  for every combination of indices  $i \in \{1, \dots, m\}$ ,  $j \in \{1, \dots, d\}$ ,  $k \in \{1, \dots, K\}$ . The representation in Lemma 3.1 above can be seen as a special case of (3.6) for the Bernoulli random walk, but it also provides the explicit structure of the orthogonal term as

$$\Delta N_{t_k} = \sum_{I \in \mathcal{I} \setminus \{\{1\}, \dots, \{m\}\}} \psi^I \Delta M_{t_k}^I. \quad (3.7)$$

When  $M$  is a general square-integrable martingale, Jacod et al. [31, Proposition 2.1] has an explicit formula for the integrands  $\psi^1, \dots, \psi^m$ ; however, the structure in (3.7) for the orthogonal term is very specific to the Bernoulli case. Similarly, for a Gaussian random walk, the orthogonal term in (3.6) is referred to as the *martingale representation error* in Akahori et al. [1] and is made explicit via a discrete version of the Clark–Ocone formula [1, Theorem 2.1]. The Bernoulli random walk makes it possible to express the sum of the stochastic integral (with respect to  $M$ ) and the orthogonal terms in (3.6) compactly as  $\sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I$ , which simplifies notation in the backward stochastic difference inclusions and set-valued backward stochastic difference equations to follow. On the other hand, when passing to continuous time, we may use a sequence of Bernoulli random walks in order to approximate Brownian motion, which is the underlying process of the fundamental BSDE representations in continuous time (see (1.1)). For these reasons, we prefer to work within the Bernoulli framework and use the predictable representation in Lemma 3.1.

**Remark 3.4** Within this paper, we consider a discrete-time model driven by a Bernoulli random walk. This lends itself to financial settings due to its convergence to Brownian motion as  $\Delta t$  tends to 0. We wish to note that for such a convergence in the multidimensional case, other processes may perform better. We refer to e.g. the He-type multinomial approximation; see He [30] and Grépat and Kabanov [26].

We conclude this section with an illustrative example to be developed further in the later sections.

**Example 3.5** We consider a simple market model with proportional transaction costs. For notational simplicity, we set  $t_k = k$  so that  $\Delta t_k = 1$  for  $k \in \{1, \dots, K\}$ . The market is driven by  $m = d - 1$  symmetric and correlated random walks, and it has  $d$  assets. Asset 0 is a riskless bond with no transaction costs; its (deterministic) price process

$(A_t)_{t \in \mathbb{T}}$  is given by  $A_t = A_0(1+r)^t$ ,  $t \in \{1, \dots, K\}$ , where  $A_0 > 0$  is the initial bond price and  $r > -1$  is the periodic interest rate. To simplify the presentation, let us assume that  $A_0 = 1$  and  $r = 0$  so that  $A_t = 1$  for  $t \in \{1, \dots, K\}$ . In addition, for  $i \in \{1, \dots, d-1\}$ , asset  $i$  is a risky stock with proportional transaction costs; its mid-market price process  $(S_t^i)_{t \in \mathbb{T}}$ , denoted in terms of the numéraire of the bond, follows a Cox–Ross–Rubinstein model given by

$$S_t^i = S_0^i e^{\sigma^i M_t^i}, \quad t \in \{1, \dots, K\},$$

where  $S_0^i > 0$  is the initial mid-market price and  $\sigma^i > 0$  is the volatility coefficient. Furthermore, there is a fixed transaction cost rate  $\lambda^i \geq 0$  for the stock, which is the same for all periods for simplicity. Hence, the bid and ask prices at time  $t \in \mathbb{T}$  are given by

$$S_t^{i,b} = S_t^i(1 - \lambda^i), \quad S_t^{i,a} = S_t^i(1 + \lambda^i),$$

respectively. Let us define the price (row) vectors

$$S_t^b := (S_t^{1,b}, \dots, S_t^{d-1,b}), \quad S_t^a := (S_t^{1,a}, \dots, S_t^{d-1,a}).$$

For simplicity, we assume that all transactions are performed via the bond, that is, a direct conversion between any two risky assets is not allowed. For  $t \in \mathbb{T}$  and  $\omega \in \Omega$ , we denote by  $K_t(\omega)$  the solvency cone of all  $d$ -dimensional portfolio vectors that can be exchanged into nonnegative portfolio vectors. In our case (see Feinstein and Rudloff [19, Sect. 7] and Löhne and Rudloff [41, Sect. 5]),  $K_t(\omega)$  is the convex cone generated by the columns of the matrix

$$\begin{pmatrix} -S_t^b(\omega) & S_t^a(\omega) \\ I_{d-1} & -I_{d-1} \end{pmatrix}, \quad (3.8)$$

where  $I_{d-1}$  is the  $(d-1) \times (d-1)$  identity matrix. From the structure of the Cox–Ross–Rubinstein model, it is clear that  $K_t(\omega)$  depends on  $\omega$  only through  $M_t(\omega)$ , the realisation of the underlying  $m$ -dimensional random walk. Hence, by a slight abuse of notation, we write

$$K_t(\omega) = K_t(c) \quad (3.9)$$

whenever  $M_t(\omega) = c$ .

When  $d = 2$ , we write  $\sigma = \sigma^1$ ,  $S_t^b = S_t^{1,b}$ ,  $S_t^a = S_t^{1,a}$  and from (3.8), we have

$$K_t(\omega) = \text{co} \left\{ \begin{pmatrix} -S_t^b(\omega) \\ 1 \end{pmatrix}, \begin{pmatrix} S_t^a(\omega) \\ -1 \end{pmatrix} \right\},$$

where  $\text{co}$  denotes the convex conic hull operator; when  $d = 3$ , we have

$$K_t(\omega) = \text{co} \left\{ \begin{pmatrix} -S_t^{1,b}(\omega) \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -S_t^{2,b}(\omega) \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} S_t^{1,a}(\omega) \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} S_t^{2,a}(\omega) \\ 0 \\ -1 \end{pmatrix} \right\}.$$

For  $t \in \mathbb{T}$ , let us define

$$L_d^0(\mathcal{F}_t^M, K_t) := \{Y \in \mathcal{L}_t^d : \mathbb{P}[\{\omega \in \Omega : Y(\omega) \in K_t(\omega)\}] = 1\},$$

the set of all measurable selectors of the random set  $K_t$ . Let

$$C_{t,T} := - \sum_{s=t}^T L_d^0(\mathcal{F}_s^M, K_s),$$

which is the set of all self-financing portfolios in this market. Then for a multidimensional claim  $X \in \mathcal{L}_T^d$ , the set

$$\text{SHP}_t(X) := \{Y \in \mathcal{L}_t^d : X \in Y + C_{t,T}\}$$

is the set of all superhedging portfolios of  $X$  at time  $t \in \mathbb{T}$ . Letting

$$R_t^{\text{SHP}}(X) := \text{SHP}_t(-X),$$

the family  $R^{\text{SHP}} = (R_t^{\text{SHP}})_{t \in \mathbb{T}}$  is a coherent multiportfolio time-consistent dynamic risk measure as studied in Feinstein and Rudloff [17, Example 5.4].

## 4 Backward stochastic difference inclusion

In this section, we show that a given dynamic set-valued risk measure in discrete time gives rise to a backward stochastic difference inclusion (BSΔI). Since we do not consider the continuous-time limits of these risk measures, we derive a BSΔI without scaling and tilting the original risk measure as was done in Stadje [49]. BSΔIs present a recursive formulation for the computation of the set-valued risk measure as a collection of selectors. As such, the proposed BSΔI provides a methodology for studying singular capital allocation strategies over time to manage risk. Because a risk manager would only ever implement a single strategy, such a construction provides exactly the dynamics an investor would need to consider. Therefore, in this section, we are interested in the backward stochastic difference inclusion that encodes the dynamic programming principle of multiportfolio time-consistency.

As in Sect. 3, we work with the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t^M)_{t \in \mathbb{T}}, \mathbb{P})$ , where  $M$  is the  $m$ -dimensional random walk defined in (3.1). Let us consider a multiportfolio time-consistent dynamic set-valued convex risk measure  $R = (R_t)_{t \in \mathbb{T}}$  with one-step conditional risk measures  $R_{t_{k-1}, t_k} : \mathcal{L}_{t_k}^d \rightarrow \mathcal{P}_+(\mathcal{L}_{t_{k-1}}^d)$ ,  $k \in \{1, \dots, K\}$ . For the terminal risk measure  $R_{t_K} : \mathcal{L}_{t_K}^d \rightarrow \mathcal{P}_+(\mathcal{L}_{t_K}^d)$ , we have  $R_{t_K}(X) = -X + R_{t_K}(0)$  for  $X \in \mathcal{L}_{t_K}^d$ .

We first relate  $R$  to a BSΔI with a *nonlocal* driver. To that end, let us introduce the domain

$$\mathbb{D} := \{(t, \psi) : t \in \{t_0, \dots, t_{K-1}\}, \psi = (\psi^I)_{I \in \mathcal{I}} \in (\mathcal{L}_t^d)^{\mathcal{I}}\} \quad (4.1)$$

and define the set-valued driver  $G : \mathbb{D} \rightarrow 2^{\mathcal{L}^d}$  by

$$G(t_{k-1}, \psi) := \frac{1}{\Delta t_k} R_{t_{k-1}, t_k} \left( - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I \right) \quad (4.2)$$

for each  $(t_{k-1}, \psi) \in \mathbb{D}$  with  $k \in \{1, \dots, K\}$ . Note that  $G$  is adapted in the sense that  $G(t, \psi) \in \mathcal{P}_+(\mathcal{L}_t^d)$  whenever  $(t, \psi) \in \mathbb{D}$ .

**Proposition 4.1** Let  $k \in \{1, \dots, K\}$  and  $Y_{t_k} \in \mathcal{L}_{t_k}^d$ . Consider the one-step BSΔI

$$Y_{t_{k-1}} \in Y_{t_k} + G(t_{k-1}, \psi_{t_{k-1}}) \Delta t_k - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \quad (4.3)$$

for some  $\psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$ . Then the set  $R_{t_{k-1}, t_k}(-Y_{t_k})$  coincides with the reachable set of (4.3), that is,

$$R_{t_{k-1}, t_k}(-Y_{t_k}) = \{Y_{t_{k-1}} \in \mathcal{L}_{t_{k-1}}^d : (4.3) \text{ holds for some } \psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}\}.$$

*Proof* Let  $Y_{t_{k-1}} \in R_{t_{k-1}, t_k}(-Y_{t_k})$ . By Lemma 3.1,  $Y_{t_k}$  has the predictable representation

$$Y_{t_k} = \xi_{t_{k-1}} + \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I$$

for some  $\xi_{t_{k-1}}, \psi_{t_{k-1}}^I \in \mathcal{L}_{t_{k-1}}^d, I \in \mathcal{I}$ . By (4.2), we have

$$\begin{aligned} Y_{t_{k-1}} - Y_{t_k} &\in R_{t_{k-1}, t_k}(-Y_{t_k}) - Y_{t_k} \\ &= R_{t_{k-1}, t_k} \left( -\xi_{t_{k-1}} - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \right) - \xi_{t_{k-1}} - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \\ &= R_{t_{k-1}, t_k} \left( -\sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \right) - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \\ &= G(t_{k-1}, \psi_{t_{k-1}}) \Delta t_k - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I. \end{aligned}$$

Hence (4.3) holds.

Conversely, let  $Y_{t_{k-1}} \in \mathcal{L}_{t_{k-1}}^d$  satisfy the BSΔI (4.3) for some  $\psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$ . By the BSΔI, there exists  $V_{t_{k-1}} \in G(t_{k-1}, \psi_{t_{k-1}})$  such that

$$Y_{t_{k-1}} = Y_{t_k} + V_{t_{k-1}} \Delta t_k - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I,$$

that is,  $Y_{t_k}$  has the predictable representation

$$Y_{t_k} = \xi_{t_{k-1}} + \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I,$$

where  $\xi_{t_{k-1}} := Y_{t_{k-1}} - V_{t_{k-1}} \Delta t_k \in \mathcal{L}_{t_{k-1}}^d$ . Recalling the definition of the driver in (4.2), the BSΔI yields

$$\begin{aligned} Y_{t_{k-1}} &\in Y_{t_k} + R_{t_{k-1}, t_k} \left( -\sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \right) - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \\ &= Y_{t_k} + R_{t_{k-1}, t_k}(-Y_{t_k} + \xi_{t_{k-1}}) - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \\ &= Y_{t_k} + R_{t_{k-1}, t_k}(-Y_{t_k}) - \xi_{t_{k-1}} - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \\ &= Y_{t_k} + R_{t_{k-1}, t_k}(-Y_{t_k}) - Y_{t_k} = R_{t_{k-1}, t_k}(-Y_{t_k}), \end{aligned}$$

which completes the proof.  $\square$

**Corollary 4.2** Let  $X \in \mathcal{L}_{t_K}^d$ . If  $(Y_t)_{t \in \mathbb{T}}$  is a process such that

$$Y_{t_K} \in R_{t_K}(X) \quad \text{and} \quad Y_{t_{k-1}} \in R_{t_{k-1}, t_k}(-Y_{t_k}), \quad k \in \{1, \dots, K\}, \quad (4.4)$$

then there exists  $\psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$  for  $k \in \{1, \dots, K\}$  such that the BSΔI

$$Y_{t_{k-1}} \in Y_{t_k} + G(t_{k-1}, \psi_{t_{k-1}}) \Delta t_k - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I, \quad k \in \{1, \dots, K\},$$

$$Y_{t_K} \in -X + R_{t_K}(0)$$

holds. Conversely, if there exist an adapted process  $(Y_t)_{t \in \mathbb{T}}$  and  $\psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$  for  $k \in \{1, \dots, K\}$  such that the above BSΔI holds, then the recursive relation (4.4) holds. In each case, the multi-step version of the BSΔI

$$Y_{t_k} \in Y_{t_K} + \sum_{\ell=k+1}^K G(t_{\ell-1}, \psi_{t_{\ell-1}}) \Delta t_{\ell} - \sum_{\ell=k+1}^K \sum_{I \in \mathcal{I}} \psi_{t_{\ell-1}}^I \Delta M_{t_{\ell}}^I, \quad k \in \{0, \dots, K-1\},$$

$$Y_{t_K} \in -X + R_{t_K}(0)$$

holds as well.

*Proof* By translativity,  $R_{t_K}(X) = -X + R_{t_K}(0)$ . From this and Proposition 4.1, the two claims about the one-step BSΔI follow immediately. The claim about the multi-step BSΔI follows by iterating the one-step version. Indeed, for  $k \in \{0, \dots, K-1\}$ , we have

$$\begin{aligned} Y_{t_k} &= Y_{t_K} + \sum_{\ell=k+1}^K (Y_{t_{\ell-1}} - Y_{t_{\ell}}) \\ &\in Y_{t_K} + \sum_{\ell=k+1}^K G(t_{\ell-1}, (\psi_{t_{\ell-1}}^I)_{I \in \mathcal{I}}) \Delta t_{\ell} - \sum_{\ell=k+1}^K \sum_{I \in \mathcal{I}} \psi_{t_{\ell-1}}^I \Delta M_{t_{\ell}}^I, \end{aligned}$$

as desired.  $\square$

Note that Corollary 4.2 shows that each process  $(Y_t)_{t \in \mathbb{T}}$  that is a solution of the one-step BSΔI is a “path” in the dynamic risk measure evaluated at  $X$ , and vice versa. Moreover, such processes are also solutions of the multi-step BSΔI. The next corollary provides a partial converse to the latter statement.

**Corollary 4.3** Let  $X \in \mathcal{L}_{t_K}^d$  and  $k \in \{0, \dots, K\}$ . Consider the multi-step BSΔI

$$Y_{t_k} \in Y_{t_K} + \sum_{\ell=k+1}^K G(t_{\ell-1}, \psi_{t_{\ell-1}}) \Delta t_{\ell} - \sum_{\ell=k+1}^K \sum_{I \in \mathcal{I}} \psi_{t_{\ell-1}}^I \Delta M_{t_{\ell}}^I. \quad (4.5)$$

Then the set  $R_{t_k}(X)$  coincides with the reachable set of (4.5), that is,

$$R_{t_k}(X) = \left\{ Y_{t_k} \in \mathcal{L}_{t_k}^d : \begin{array}{l} (4.5) \text{ holds for some } Y_{t_K} \in -X + R_{t_K}(0), \\ \psi_{t_k} \in (\mathcal{L}_{t_k}^d)^{\mathcal{I}}, \dots, \psi_{t_{K-1}} \in (\mathcal{L}_{t_{K-1}}^d)^{\mathcal{I}} \end{array} \right\}. \quad (4.6)$$

*Proof* For  $k \in \{0, \dots, K-1\}$ , let us denote by  $\bar{R}_{t_k}(X)$  the reachable set on the right-hand of (4.6). We show that  $\bar{R}_{t_k}(X) = R_{t_k}(X)$  for  $k \in \{0, \dots, K-1\}$  by backward induction on  $k$ . For the base case, we have  $\bar{R}_{t_{K-1}}(X) = R_{t_{K-1}}(X)$  directly from Corollary 4.2. For the inductive case, let  $k \in \{1, \dots, K-1\}$  and assume that  $\bar{R}_{t_k}(X) = R_{t_k}(X)$ . Consider the inclusion (4.5) at time  $t_{k-1}$ , i.e., the inclusion

$$Y_{t_{k-1}} \in Y_{t_K} + \sum_{\ell=k}^K G(t_{\ell-1}, \psi_{t_{\ell-1}}) \Delta t_{\ell} - \sum_{\ell=k}^K \sum_{I \in \mathcal{I}} \psi_{t_{\ell-1}}^I \Delta M_{t_{\ell}}^I. \quad (4.7)$$

Then

$$\begin{aligned} & \bar{R}_{t_{k-1}}(X) \\ &= \left\{ Y_{t_{k-1}} \in \mathcal{L}_{t_{k-1}}^d : \begin{array}{l} (4.7) \text{ holds for some } Y_{t_K} \in -X + R_{t_K}(0), \\ \psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}, \psi_{t_k} \in (\mathcal{L}_{t_k}^d)^{\mathcal{I}}, \dots, \psi_{t_{K-1}} \in (\mathcal{L}_{t_{K-1}}^d)^{\mathcal{I}} \end{array} \right\} \\ &= \left\{ Y_{t_{k-1}} \in \mathcal{L}_{t_{k-1}}^d : \begin{array}{l} Y_{t_{k-1}} \in Y_{t_k} + G(t_{k-1}, \psi_{t_{k-1}}) \Delta t_k - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \\ \text{holds for some } Y_{t_k} \in R_{t_k}(X), \psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}} \end{array} \right\} \\ &= \{Y_{t_{k-1}} \in R_{t_{k-1}, t_k}(-Y_{t_k}) : Y_{t_k} \in R_{t_k}(X)\} \\ &= R_{t_{k-1}, t_k}[-R_{t_k}(X)] = R_{t_{k-1}}(X). \end{aligned}$$

In this calculation, the second equality follows from applying the induction hypothesis, the third from Proposition 4.1, and the last from multiportfolio time-consistency.  $\square$

The driver function  $G$  defined in (4.2) can be considered *nonlocal* for the following two reasons: its second argument is a  $(d \times |\mathcal{I}|)$ -dimensional random vector (rather than a deterministic vector), and its output is a set of random vectors (rather than a deterministic set). We first aim to rewrite the BSΔI in Corollary 4.2 using a *semi-local* driver  $g$  whose second argument is a deterministic vector, but whose output is still a set of random vectors. Let us define  $g: \{t_0, \dots, t_{K-1}\} \times (\mathbb{R}^d)^{\mathcal{I}} \rightarrow 2^{\mathcal{L}_{t_k}^d}$  by

$$g(t_{k-1}, z) = \frac{1}{\Delta t_k} R_{t_{k-1}, t_k} \left( - \sum_{I \in \mathcal{I}} z^I \Delta M_{t_k}^I \right) \quad (4.8)$$

for  $k \in \{0, \dots, K-1\}$ ,  $z = (z^I)_{I \in \mathcal{I}} \in (\mathbb{R}^d)^{\mathcal{I}}$ .

To connect the two drivers  $G, g$ , we also define a special type of composition of  $g$ , through its second argument, with a random vector. Let  $\ell \in \{0, \dots, K\}$  and denote by  $\mathcal{A}_{t_{\ell}} \subseteq 2^{\Omega}$  the partition of  $\Omega$  that generates  $\mathcal{F}_{t_{\ell}}^M$ , which is of size  $2^{\ell m}$ . Let  $\psi = (\psi^I)_{I \in \mathcal{I}} \in (\mathcal{L}_{t_{\ell}}^d)^{\mathcal{I}}$ . Given  $I \in \mathcal{I}$  and  $A \in \mathcal{A}_{t_{\ell}}$ , note that  $\psi^I$  is constant on  $A$  so that by a slight abuse of notation, we can define

$$\psi^I(A) := \psi^I(\omega)$$

for some  $\omega \in A$  (chosen arbitrarily in  $A$ ) and write  $\psi(A) = (\psi^I(A))_{I \in \mathcal{I}} \in (\mathbb{R}^d)^{\mathcal{I}}$ .



For  $k \in \{1, \dots, K\}$  and  $\psi \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$ , we define

$$g \circ (t_{k-1}, \psi) := \sum_{A \in \mathcal{A}_{t_{k-1}}} g(t_{k-1}, \psi(A)) \mathbb{1}_A.$$

The next lemma states that this composition coincides with the nonlocal driver  $G$  in (4.2).

**Lemma 4.4** For  $k \in \{1, \dots, K\}$  and  $\psi = (\psi^I)_{I \in \mathcal{I}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$ , we have

$$G(t_{k-1}, \psi) = g \circ (t_{k-1}, \psi).$$

*Proof* By the decomposability of  $R$ , we have

$$\begin{aligned} G(t_{k-1}, \psi) &= \frac{1}{\Delta t_k} R_{t_{k-1}, t_k} \left( - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I \right) \\ &= \frac{1}{\Delta t_k} R_{t_{k-1}, t_k} \left( \sum_{A \in \mathcal{A}_{t_{k-1}}} \left( - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I \right) \mathbb{1}_A \right) \\ &= \frac{1}{\Delta t_k} R_{t_{k-1}, t_k} \left( \sum_{A \in \mathcal{A}_{t_{k-1}}} \left( - \sum_{I \in \mathcal{I}} \psi^I(A) \Delta M_{t_k}^I \right) \mathbb{1}_A \right) \\ &= \frac{1}{\Delta t_k} \sum_{A \in \mathcal{A}_{t_{k-1}}} R_{t_{k-1}, t_k} \left( - \sum_{I \in \mathcal{I}} \psi^I(A) \Delta M_{t_k}^I \right) \mathbb{1}_A \\ &= \sum_{A \in \mathcal{A}_{t_{k-1}}} g(t_{k-1}, \psi(A)) \mathbb{1}_A = g \circ (t_{k-1}, \psi), \end{aligned}$$

from which the result follows.  $\square$

Thanks to Lemma 4.4, the one-step BSΔI in Corollary 4.2 can be rewritten as

$$Y_{t_{k-1}} \in Y_{t_k} + g \circ (t_{k-1}, \psi_{t_{k-1}}) \Delta t_k - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I, \quad k \in \{1, \dots, K\},$$

$$Y_{t_K} \in -X + R_{t_K}(0).$$

Note that this BSΔI is a *functional* inclusion where one random vector is included in a set of random vectors. In Proposition 4.6 below, we present an alternative, defining BSΔI that is a *random* inclusion. This alternative BSΔI has a completely local driver  $\hat{g} : \Omega \times \{t_0, \dots, t_{K-1}\} \times (\mathbb{R}^d)^{\mathcal{I}} \rightarrow \mathcal{P}_+(\mathbb{R}^d)$  defined by

$$\hat{g}(\omega, t_{k-1}, z) := g(t_{k-1}, z)(\omega) = \frac{1}{\Delta t_k} R_{t_{k-1}, t_k} \left( - \sum_{I \in \mathcal{I}} z^I \Delta M_{t_k}^I \right)(\omega)$$

for  $\omega \in \Omega, k \in \{1, \dots, K\}, z = (z^I)_{I \in \mathcal{I}} \in (\mathbb{R}^d)^{\mathcal{I}}$ . To begin with, we formulate the connection between the drivers  $G$  and  $\hat{g}$  in the next lemma.

**Lemma 4.5** Let  $k \in \{1, \dots, K\}$  and  $\psi = (\psi^I)_{I \in \mathcal{I}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$ . For each  $\omega \in \Omega$ , we have

$$G(t_{k-1}, \psi)(\omega) = \hat{g}(\omega, t_{k-1}, \psi(\omega)).$$

*Proof* Let  $\omega \in \Omega$ . We claim that

$$G(t_{k-1}, \psi)(\omega) = \left( \sum_{A \in \mathcal{A}_{t_{k-1}}} g(t_{k-1}, \psi(A)) \mathbb{1}_A \right)(\omega) = \hat{g}(\omega, t_{k-1}, \psi(\omega)).$$

The first equality is from the calculation in the proof of Lemma 4.4. To prove the second, let  $\bar{A}$  be the unique set in  $\mathcal{A}_{t_{k-1}}$  for which  $\omega \in \bar{A}$ . Let

$$u \in \left( \sum_{A \in \mathcal{A}_{t_{k-1}}} g(t_{k-1}, \psi(A)) \mathbb{1}_A \right)(\omega). \quad (4.9)$$

Hence

$$\begin{aligned} u \mathbb{1}_{\{\omega\}} &\in \sum_{A \in \mathcal{A}_{t_{k-1}}} g(t_{k-1}, \psi(A)) \mathbb{1}_A \mathbb{1}_{\{\omega\}} = g(t_{k-1}, \psi(\bar{A})) \mathbb{1}_{\bar{A}} \mathbb{1}_{\{\omega\}} \\ &= g(t_{k-1}, \psi(\bar{A})) \mathbb{1}_{\{\omega\}}. \end{aligned}$$

This shows that  $u \in g(t_{k-1}, \psi(\bar{A}))(\omega) = g(t_{k-1}, \psi(\omega))(\omega) = \hat{g}(\omega, t_{k-1}, \psi(\omega))$ . Conversely, let  $u \in \hat{g}(\omega, t_{k-1}, \psi(\omega)) = g(t_{k-1}, \psi(\bar{A}))(\omega)$ . So

$$u \mathbb{1}_{\{\omega\}} \in g(t_{k-1}, \psi(\bar{A})) \mathbb{1}_{\{\omega\}} = \left( \sum_{A \in \mathcal{A}_{t_{k-1}}} g(t_{k-1}, \psi(A)) \mathbb{1}_A \right) \mathbb{1}_{\{\omega\}},$$

which shows that (4.9) holds for  $u$ .  $\square$

With Lemma 4.5, we are ready to formulate the random BSΔI.

**Proposition 4.6** Let  $X \in \mathcal{L}_{t_K}^d$ . Consider a process  $(Y_t)_{t \in \mathbb{T}}$  satisfying the relation (4.4), i.e.,  $Y_{t_K} \in R_{t_K}(X)$  and  $Y_{t_{k-1}} \in R_{t_{k-1}, t_k}(-Y_{t_k})$  for  $k \in \{1, \dots, K\}$ . Then there exist  $\psi_{t_k} = (\psi_{t_k}^I)_{I \in \mathcal{I}} \in (\mathcal{L}_{t_k}^d)^{\mathcal{I}}$  for  $k \in \{1, \dots, K\}$  such that the random BSΔI

$$Y_{t_{k-1}}(\omega) \in Y_{t_k}(\omega) + \hat{g}(\omega, t_{k-1}, \psi_{t_{k-1}}(\omega)) \Delta t_k - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I(\omega) \Delta M_{t_k}^I(\omega),$$

$$Y_{t_K}(\omega) \in -X(\omega) + R_{t_K}(0)(\omega)$$

holds for every  $\omega \in \Omega$ ,  $k \in \{1, \dots, K\}$ . Conversely, if there exist an adapted process  $(Y_t)_{t \in \mathbb{T}}$  and  $\psi_{t_{k-1}} \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$  for  $k \in \{1, \dots, K\}$  such that the above random BSΔI holds for every  $\omega \in \Omega$ , then  $Y_{t_K} \in R_{t_K}(X)$  and  $Y_{t_{k-1}} \in R_{t_{k-1}, t_k}(-Y_{t_k})$  for  $k \in \{1, \dots, K\}$ . In each case, the multi-step version of the random BSΔI

$$Y_{t_k}(\omega) \in Y_{t_K}(\omega) + \sum_{\ell=k+1}^K \hat{g}(\omega, t_{\ell-1}, \psi_{t_{\ell-1}}(\omega)) \Delta t_{\ell} - \sum_{\ell=k+1}^K \sum_{I \in \mathcal{I}} \psi_{t_{\ell-1}}^I(\omega) \Delta M_{t_{\ell}}^I(\omega)$$

holds for every  $k \in \{0, \dots, K-1\}$  and  $\omega \in \Omega$ .

*Proof* Let  $k \in \{0, \dots, K-1\}$ . We rewrite the BSΔI in Corollary 4.2 as

$$Y_{t_{k-1}} - Y_{t_k} + \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I \Delta M_{t_k}^I \in G(t_{k-1}, \psi_{t_{k-1}}) \Delta t_k.$$

By the decomposability of  $R$ , it follows that the set on the right is decomposable. Hence the above inclusion is equivalent to

$$Y_{t_{k-1}}(\omega) - Y_{t_k}(\omega) + \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I(\omega) \Delta M_{t_k}^I(\omega) \in G(t_{k-1}, \psi_{t_{k-1}})(\omega) \Delta t_k,$$

which is equivalent to the random one-step inclusion in the statement of the proposition, by the definition of  $\hat{g}$ . Similarly, it can be checked that the multi-step inclusion in Corollary 4.2 is equivalent to the random multi-step inclusion in the statement of the proposition. All the claims follow immediately from these equivalences and Corollary 4.2.  $\square$

*Remark 4.7* To consider the local analogue of Corollary 4.3, we utilise a notation from Feinstein and Rudloff [19], Löhne and Rudloff [41] on the *finitely generated* filtration considered herein. As this filtration is generated by the Bernoulli random walk, we can view our processes as taking values on a tree. With that in mind, we present this local analogue of Corollary 4.3 using the concept and notation of a tree. Recall that  $\mathcal{A}_{t_\ell} \subseteq 2^\Omega$  denotes the set of all atoms of  $\mathcal{F}_{t_\ell}^M$  for  $\ell \in \{0, \dots, K\}$ , and for a given  $Z \in \mathcal{L}_{t_\ell}^d$ , we denote by  $Z(A)$  the constant value of  $Z$  on  $A \in \mathcal{A}_{t_\ell}$ . Additionally, when  $\ell \leq K-1$ , we denote the “successor” nodes at time  $t_{\ell+1}$  of  $A_\ell \in \mathcal{A}_{t_\ell}$  by

$$\text{succ}(A_\ell) = \{A_{\ell+1} \in \mathcal{A}_{t_{\ell+1}} : A_{\ell+1} \subseteq A_\ell\}.$$

Fix  $X \in \mathcal{L}_{t_K}^d$ . In this atomised notation, the one-step random BSΔI from Proposition 4.6 for  $k \in \{1, \dots, K\}$  can be rewritten as

$$\begin{aligned} Y_{t_{k-1}}(A_{k-1}) \in & \bigcap_{A_k \in \text{succ}(A_{k-1})} \left( Y_{t_k}(A_k) + \hat{g}(A_{k-1}, t_{k-1}, \psi_{t_{k-1}}(A_{k-1})) \Delta t_{k-1} \right. \\ & \left. - \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I(A_{k-1}) \Delta M_{t_k}^I(A_k) \right), \\ Y_{t_K}(A_K) \in & -X(A_K) + R_{t_K}(0)(\omega_K) \end{aligned}$$

for every  $A_{k-1} \in \mathcal{A}_{t_{k-1}}$ . In this way, we can consider the multi-step inclusion at time  $t_k$  with  $k \in \{0, \dots, K-1\}$  and  $A_k \in \mathcal{A}_{t_k}$  as

$$\begin{aligned} y \in & \bigcap_{\substack{A_{k+1} \in \text{succ}(A_k), \dots, \\ A_K \in \text{succ}(A_{K-1})}} \left( Y_{t_K}(A_K) + \sum_{\ell=k+1}^K \hat{g}(A_{\ell-1}, t_{\ell-1}, \psi_{t_{\ell-1}}(A_{\ell-1})) \Delta t_\ell \right. \\ & \left. - \sum_{\ell=k+1}^K \sum_{I \in \mathcal{I}} \psi_{t_{\ell-1}}^I(A_{\ell-1}) \Delta M_{t_\ell}^I(A_\ell) \right) \end{aligned} \quad (4.10)$$

for some  $\psi_{t_{\ell-1}} \in (\mathcal{L}_{t_{\ell-1}}^d)^{\mathcal{I}}$ ,  $\ell \in \{k+1, \dots, K\}$ . (This representation is an example of the conditional core introduced in Lépinette and Molchanov [39]; we conjecture that the conditional core may be useful for extending (4.10) to e.g. the continuous-time limit.) Similarly to Corollary 4.3, given  $A_k \in \mathcal{A}_{t_k}$ , it can be checked that the set  $R_{t_k}^{A_k}(X) := R_{t_k}(X)(\omega)$ ,  $\omega \in A_k$ , coincides with the reachable set of (4.10), that is,

$$R_{t_k}^{A_k}(X) = \left\{ y \in \mathbb{R}^d : \begin{array}{l} \text{(4.10) holds for some } Y_{t_K}(A_K) \in -X(A_K) + R_{t_K}^{A_K}(0), \\ \psi_{t_k}(A_k), \dots, \psi_{t_{K-1}}(A_{K-1}) \in (\mathbb{R}^d)^{\mathcal{I}}, \\ \text{for all } A_{k+1} \in \text{succ}(A_k), \dots, A_K \in \text{succ}(A_{K-1}) \end{array} \right\}.$$

We conclude this section with two examples on superhedging as a follow-up to the model introduced in Example 3.5.

*Example 4.8* In the setting of Example 3.5, let us calculate the semi-local driver  $g^{\text{SHP}}$  of the superhedging risk measure  $R^{\text{SHP}}$  when  $d = 2$  and  $m = 1$ . Following (4.8) and the definitions in Example 3.5, for  $t \in \{0, \dots, T-1\}$  and  $z \in \mathbb{R}^2$ , we have

$$\begin{aligned} g^{\text{SHP}}(t, z) &= R_{t,t+1}^{\text{SHP}}(-z \Delta M_{t+1}) = \text{SHP}_t(z B_{t+1}) \\ &= \left\{ Y \in \mathcal{L}_t^2 : Y \in z B_{t+1} + \sum_{s=t}^T L_d^0(\mathcal{F}_s^M, K_s) \right\} \\ &= L_d^0(\mathcal{F}_t^M, K_t) + \left\{ Y \in \mathcal{L}_t^2 : Y \in z B_{t+1} + \sum_{s=t+1}^T L_d^0(\mathcal{F}_s^M, K_s) \right\} \\ &= L_d^0(\mathcal{F}_t^M, K_t) + \left\{ Y \in \mathcal{L}_t^2 : Y = z B_{t+1} + \sum_{s=t+1}^T \left( b_s \begin{pmatrix} -S_s^b \\ 1 \end{pmatrix} + a_s \begin{pmatrix} S_s^a \\ -1 \end{pmatrix} \right), \right. \\ &\quad \left. a_s, b_s \in \mathcal{L}_{s,+}^2, s \in \{t+1, \dots, T\} \right\}. \end{aligned}$$

Note that for  $s \in \{t+1, \dots, T\}$ , we have

$$S_s^b = S_t^b e^{\sigma B_{t+1}} e^{\sigma(M_s - M_{t+1})}, \quad S_s^a = S_t^a e^{\sigma B_{t+1}} e^{\sigma(M_s - M_{t+1})}.$$

Let us denote by  $\hat{g}^{\text{SHP}}$  the local driver. By the structure of the solvency cones, it follows that  $\hat{g}^{\text{SHP}}(\omega, t, z)$  depends on  $\omega$  only through  $M_t(\omega)$ . Hence similarly to (3.9), we write

$$\hat{g}^{\text{SHP}}(\omega, t, z) = \hat{g}^{\text{SHP}}(c, t, z)$$

whenever  $M_t(\omega) = c$ . The above calculations for the semi-local driver  $g^{\text{SHP}}$  imply that

$$\begin{aligned} \hat{g}^{\text{SHP}}(c, t, z) &= K_t(c) + \left( \left( z + \sum_{s=t+1}^T \bigcap_{b_{t+2}, \dots, b_s \in \{-1, +1\}} K_s(c + 1 + b_{t+2} + \dots + b_s) \right) \right. \\ &\quad \left. \cap \left( -z + \sum_{s=t+1}^T \bigcap_{b_{t+2}, \dots, b_s \in \{-1, +1\}} K_s(c - 1 + b_{t+2} + \dots + b_s) \right) \right). \end{aligned}$$

Note that for  $s \in \{t+1, \dots, T\}$ , the intersection over  $b_{t+2}, \dots, b_s \in \{-1, +1\}$  has  $s - t$  different cones due to the dependence on the sum  $b_{t+2} + \dots + b_s$ .

As a special case, assume that  $T = 2$ . Then the general expression for  $\hat{g}^{\text{SHP}}$  yields

$$\begin{aligned} \hat{g}^{\text{SHP}}(0, 0, z) &= K_0 + \left( \left( z + K_1(1) + (K_2(2) \cap K_2(0)) \right) \right. \\ &\quad \left. \cap \left( -z + K_1(-1) + (K_2(0) \cap K_2(-2)) \right) \right), \\ \hat{g}^{\text{SHP}}(1, 1, z) &= K_1(1) + \left( (z + K_2(2)) \cap (-z + K_2(0)) \right), \\ \hat{g}^{\text{SHP}}(-1, 1, z) &= K_1(-1) + \left( (z + K_2(0)) \cap (-z + K_2(-2)) \right). \end{aligned}$$

In this example, the driver is the intersection of two “affine” set-valued functions of  $z$ . In the univariate setting  $d = m = 1$  with no transaction costs, it is worth noting that the hedging example considered in Rosazza Gianin [47, Sect. 3] has a driver that is linear in  $z$ .

**Example 4.9** The calculations in Example 4.8 can be extended to the case with  $d = 3$  and  $m = 2$ . In addition to the principal random walks  $M^1, M^2$ , we also need the process  $M^{\{1,2\}}$  whose increments are of the form  $\Delta M_t^{\{1,2\}} = B_t^{\{1,2\}} = B_t^1 B_t^2$ . Let  $t \in \{0, 1, \dots, T-1\}$  and  $z = (z^1, z^2, z^{12}) \in (\mathbb{R}^3)^T$ . For the semi-local driver, similar calculations as in Example 4.8 yield

$$g^{\text{SHP}}(t, z) = \left\{ Y \in \mathcal{L}_t^3 : Y \in z^1 B_{t+1}^1 + z^2 B_{t+1}^2 + z^{12} B_{t+1}^{\{1,2\}} + \sum_{s=t}^T L_d^0(\mathcal{F}_s^M, K_s) \right\}.$$

For  $u \in \mathbb{N}$ , let us define the set

$$\mathbb{S}_u := \left\{ \sum_{s=1}^u b_s : b_s \in \{-1, +1\}, s \in \{1, \dots, u\} \right\}.$$

For instance,  $\mathbb{S}_1 = \{-1, +1\}$ ,  $\mathbb{S}_2 = \{-2, 0, +2\}$ ,  $\mathbb{S}_3 = \{-3, -1, +1, +3\}$ . For convenience, we set  $\mathbb{S}_0 = \{0\}$ . Let  $\omega \in \Omega$ . To express the local driver  $\hat{g}^{\text{SHP}}$ , we write

$$K_t(\omega) = K_t(c^1, c^2), \quad \hat{g}^{\text{SHP}}(\omega, t, z) = \hat{g}^{\text{SHP}}(c^1, c^2, t, z)$$

whenever  $M_t^1(\omega) = c^1$  and  $M_t^2(\omega) = c^2$ . Then  $\hat{g}^{\text{SHP}}$  is given by

$$\begin{aligned} & \hat{g}^{\text{SHP}}(c^1, c^2, t, z) \\ &= K_t(c^1, c^2) \\ &+ \left( \left( z^1 + z^2 + z^{12} + \sum_{s=t+1}^T \bigcap_{b^1, b^2 \in \mathbb{S}_{s-t-1}} K_s(c^1 + 1 + b^1, c^2 + 1 + b^2) \right) \right. \\ &\quad \cap \left( z^1 - z^2 - z^{12} + \sum_{s=t+1}^T \bigcap_{b^1, b^2 \in \mathbb{S}_{s-t-1}} K_s(c^1 + 1 + b^1, c^2 - 1 + b^2) \right) \\ &\quad \cap \left( -z^1 + z^2 - z^{12} + \sum_{s=t+1}^T \bigcap_{b^1, b^2 \in \mathbb{S}_{s-t-1}} K_s(c^1 - 1 + b^1, c^2 + 1 + b^2) \right) \\ &\quad \left. \cap \left( -z^1 - z^2 + z^{12} + \sum_{s=t+1}^T \bigcap_{b^1, b^2 \in \mathbb{S}_{s-t-1}} K_s(c^1 - 1 + b^1, c^2 - 1 + b^2) \right) \right). \end{aligned}$$

As a special case, for  $T = 2$ , we obtain

$$\begin{aligned} & \hat{g}^{\text{SHP}}(0, 0, 0, z) \\ &= K_0 + \left( \left( z^1 + z^2 + z^{12} + K_1(1, 1) \right. \right. \\ &\quad \left. \left. + (K_2(2, 2) \cap K_2(2, 0) \cap K_2(0, 2) \cap K_2(0, 0)) \right) \right. \\ &\quad \cap \left( z^1 - z^2 - z^{12} + K_1(1, -1) \right. \\ &\quad \left. + (K_2(2, 0) \cap K_2(2, -2) \cap K_2(0, 0) \cap K_2(0, -2)) \right) \\ &\quad \cap \left( -z^1 + z^2 - z^{12} + K_1(-1, 1) \right. \\ &\quad \left. + (K_2(0, 2) \cap K_2(-2, 2) \cap K_2(0, 0) \cap K_2(-2, 0)) \right) \\ &\quad \left. \cap \left( -z^1 - z^2 + z^{12} + K_1(-1, -1) \right. \right. \\ &\quad \left. \left. + (K_2(0, 0) \cap K_2(0, -2) \cap K_2(-2, 0) \cap K_2(-2, -2)) \right) \right) \end{aligned}$$

for  $t = 0$  and

$$\begin{aligned} & \hat{g}^{\text{SHP}}(1, 1, 1, z) \\ &= K_1(1, 1) + \left( (z^1 + z^2 + z^{12} + K_2(2, 2)) \cap (z^1 - z^2 - z^{12} + K_2(2, 0)) \right. \\ &\quad \left. \cap (-z^1 + z^2 - z^{12} + K_2(0, 2)) \cap (-z^1 - z^2 + z^{12} + K_2(0, 0)) \right), \end{aligned}$$

$$\begin{aligned}
& \hat{g}^{\text{SHP}}(1, -1, 1, z) \\
&= K_1(1, -1) + \left( (z^1 + z^2 + z^{12} + K_2(2, 0)) \cap (z^1 - z^2 - z^{12} + K_2(2, -2)) \right. \\
&\quad \cap (-z^1 + z^2 - z^{12} + K_2(0, 0)) \\
&\quad \left. \cap (-z^1 - z^2 + z^{12} + K_2(0, -2)) \right), \\
& \hat{g}^{\text{SHP}}(-1, 1, 1, z) \\
&= K_1(-1, 1) + \left( (z^1 + z^2 + z^{12} + K_2(0, 2)) \cap (z^1 - z^2 - z^{12} + K_2(0, 0)) \right. \\
&\quad \cap (-z^1 + z^2 - z^{12} + K_2(-2, 2)) \\
&\quad \left. \cap (-z^1 - z^2 + z^{12} + K_2(-2, 0)) \right), \\
& \hat{g}^{\text{SHP}}(-1, -1, 1, z) \\
&= K_1(-1, -1) + \left( (z^1 + z^2 + z^{12} + K_2(0, 0)) \cap (z^1 - z^2 - z^{12} + K_2(0, -2)) \right. \\
&\quad \cap (-z^1 + z^2 - z^{12} + K_2(-2, 0)) \\
&\quad \left. \cap (-z^1 - z^2 + z^{12} + K_2(-2, -2)) \right)
\end{aligned}$$

for  $t = 1$ .

## 5 Set-valued backward stochastic difference equation

In this section, we show that a given dynamic set-valued risk measure in discrete time gives rise to a set-valued backward stochastic difference equation (SV-BSΔE). In contrast to the BSΔI presented in the previous section, which is evaluated for the selectors of the solution set, an SV-BSΔE provides the dynamics for the entire set as a “singular” object. As shown in (2.1), multiportfolio time-consistency provides a recursive relation for the *set* of acceptable capital allocations. Therefore, since the dynamic programming principle for set-valued risk measures is defined for the full set rather than selectors, we wish to consider the SV-BSΔE which encodes the defining relation of multiportfolio time-consistency for dynamic risk measures.

We wish to highlight that the general theory of set-valued difference and differential equations is typically restricted to the space of compact and convex sets. However, as previously discussed, risk measures are naturally upper sets and thus require further consideration. As far as the authors are aware, the method for constructing a set-valued difference (or differential in the limit) equation via an intersection of half-spaces is novel to this work. We wish to note that set-valued (Dini) directional derivatives have been studied with respect to halfspaces in e.g. Crespi et al. [13], Hamel and Schrage [28].

As in Sect. 3, we work with the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t^M)_{t \in \mathbb{T}}, \mathbb{P})$ , where  $M$  is the  $m$ -dimensional Bernoulli random walk defined in (3.1). Con-

sider a multiportfolio time-consistent dynamic set-valued convex risk measure  $R = (R_t)_{t \in \mathbb{T}}$  with one-step conditional risk measures  $R_{t_{k-1}, t_k} : \mathcal{L}_{t_k}^d \rightarrow \mathcal{P}_+(\mathcal{L}_{t_{k-1}}^d)$ ,  $k \in \{1, \dots, K\}$ . For the terminal risk measure  $R_{t_K} : \mathcal{L}_{t_K}^d \rightarrow \mathcal{P}_+(\mathcal{L}_{t_K}^d)$ , we have  $R_{t_K}(X) = -X + R_{t_K}(0)$  for  $X \in \mathcal{L}_{t_K}^d$ .

As with our discussion of the BSΔI, we relate  $R$  to an SV-BSΔE with a *nonlocal* driver. To that end, let us introduce the domain

$$\mathbb{D}_E := \{(t_k, \psi, w) : k \in \{0, \dots, K-1\}, \psi = (\psi^I)_{I \in \mathcal{I}} \in (\mathcal{L}_{t_k}^d)^{\mathcal{I}}, w \in \mathcal{L}_{t_{k+1}, +}^d\}.$$

Note that this domain has an additional variable compared to the domain for the BSΔI introduced by (4.1) in the previous section. With this, we define the set-valued driver  $G_E : \mathbb{D}_E \rightarrow 2^{\mathcal{L}^d}$  by

$$G_E(t_{k-1}, \psi, w) := \frac{1}{\Delta t_k} R_{t_{k-1}, t_k} \left[ - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I - \Gamma_{t_k}(w) \right] \quad (5.1)$$

for each  $(t_{k-1}, \psi, w) \in \mathbb{D}_E$  with  $k \in \{1, \dots, K\}$  and where

$$\Gamma_{t_k}(w) := \{u \in \mathcal{L}_{t_k}^d : w^\top u \geq 0\}. \quad (5.2)$$

Note that  $G_E$  is adapted so that  $G_E(t, \psi, w) \in \mathcal{P}_+(\mathcal{L}_t^d)$  whenever  $(t, \psi, w) \in \mathbb{D}_E$ .

The following lemma is required for the results about SV-BSΔEs. We separate it from the main results for readability.

**Lemma 5.1** *Let  $X \in \mathcal{L}_{t_K}^d$  and define*

$$H_{t_k}(X, w_{t_k}) := \text{cl} \left( R_{t_k}(X) + \Gamma_{t_k}(w_{t_k}) \right)$$

*for  $w_{t_k} \in \mathcal{L}_{t_k, +}^d$ , where the conditional halfspace  $\Gamma_{t_k}(w_{t_k})$  is defined by (5.2). Then for any  $k \in \{1, \dots, K\}$ , we have the relation*

$$R_{t_{k-1}}(X) = \bigcap_{w_{t_k} \in \mathcal{L}_{t_k, +}^d} R_{t_{k-1}, t_k}[-H_{t_k}(X, w_{t_k})]. \quad (5.3)$$

*Proof* First notice that  $H_{t_k}(X, w_{t_k}) = \{u \in \mathcal{L}_{t_k}^d : w_{t_k}^\top u \geq \rho_{t_k}(X, w_{t_k})\}$  for the scalarisation

$$\rho_{t_k}(X, w_{t_k}) := \text{ess inf}_{u \in R_{t_k}(X)} w_{t_k}^\top u.$$

Therefore, by Feinstein and Rudloff [18, Lemma 3.18],

$$R_{t_k}(X) = \bigcap_{w_{t_k} \in \mathcal{L}_{t_k, +}^d} H_{t_k}(X, w_{t_k}).$$



Next, we prove the “ $\subseteq$ ” part of (5.3). By multiportfolio time-consistency, monotonicity and the above notes,

$$\begin{aligned} R_{t_{k-1}}(X) &= R_{t_{k-1}, t_k}[-R_{t_k}(X)] = R_{t_{k-1}, t_k} \left[ \bigcap_{w_{t_k} \in \mathcal{L}_{t_k, +}^d} -H_{t_k}(X, w_{t_k}) \right] \\ &\subseteq \bigcap_{w_{t_k} \in \mathcal{L}_{t_k, +}^d} R_{t_{k-1}, t_k}[-H_{t_k}(X, w_{t_k})]. \end{aligned}$$

Finally, we prove the “ $\supseteq$ ” part of (5.3). To get a contradiction, suppose that there exists  $u \notin R_{t_{k-1}}(X)$  with

$$u \in \bigcap_{w_{t_k} \in \mathcal{L}_{t_k, +}^d} R_{t_{k-1}, t_k}[-H_{t_k}(X, w_{t_k})].$$

In particular, for every  $w_{t_k} \in \mathcal{L}_{t_k, +}^d$ , there exists  $Z(w_{t_k}) \in H_{t_k}(X, w_{t_k})$  such that

$$u \in R_{t_{k-1}, t_k}(-Z(w_{t_k})). \quad (5.4)$$

Before continuing, we first need to introduce some notation. Let  $\mathcal{M}$  be the space of all probability measures that are absolutely continuous with respect to  $\mathbb{P}$ . For any  $\mathbb{Q} \in \mathcal{M}^d$ ,  $v \in \mathcal{L}_{t_{k-1}}^d$  and  $\ell \in \{k, \dots, K\}$ , define  $w_{t_{k-1}}^{t_\ell}(\mathbb{Q}, v) \in \mathcal{L}_{t_\ell}^d$  by

$$w_{t_{k-1}}^{t_\ell}(\mathbb{Q}, v)_i = \begin{cases} v_i \mathbb{E}[\frac{d\mathbb{Q}_i}{d\mathbb{P}} \mid \mathcal{F}_{t_\ell}^M] & \text{on } \{\mathbb{E}[\frac{d\mathbb{Q}_i}{d\mathbb{P}} \mid \mathcal{F}_{t_{k-1}}^M] > 0\}, \\ v_i & \text{on } \{\mathbb{E}[\frac{d\mathbb{Q}_i}{d\mathbb{P}} \mid \mathcal{F}_{t_{k-1}}^M] = 0\}. \end{cases}$$

Define the space of set-valued dual variables

$$\mathcal{W}_{t_{k-1}} := \{(\mathbb{Q}, v) \in \mathcal{M}^d \times \mathcal{L}_{t_{k-1}, +}^d : w_{t_{k-1}}^{t_K}(\mathbb{Q}, v) \in \mathcal{L}_{t_K, +}^d\}$$

and the projection

$$\mathcal{W}_{t_{k-1}}(v) := \{\mathbb{Q} \in \mathcal{M}^d : (\mathbb{Q}, v) \in \mathcal{W}_{t_{k-1}}\}.$$

Note that for any choice of dual variables  $(\mathbb{Q}, v) \in \mathcal{W}_{t_{k-1}}$ , we have  $w_{t_{k-1}}^{t_k}(\mathbb{Q}, v) \in \mathcal{L}_{t_k, +}^d$  so that  $Z(w_{t_{k-1}}^{t_k}(\mathbb{Q}, v)) \in H_{t_k}(X, w_{t_{k-1}}^{t_k}(\mathbb{Q}, v))$ , that is,

$$w_{t_{k-1}}^{t_k}(\mathbb{Q}, v)^\top Z(w_{t_{k-1}}^{t_k}(\mathbb{Q}, v)) \geq \rho_{t_k}(X, w_{t_{k-1}}^{t_k}(\mathbb{Q}, v)) \quad (5.5)$$

for every  $(\mathbb{Q}, v) \in \mathcal{W}_{t_{k-1}}$ . From Feinstein and Rudloff [18, Lemma 3.18] and [22, Theorem 5.3],

$$u \notin R_{t_{k-1}}(X) = \bigcap_{v \in \mathcal{L}_{t_{k-1}, +}^d} H_{t_{k-1}}(X, v)$$

$$\Rightarrow \exists v \in \mathcal{L}_{t_{k-1}, +}^d : \mathbb{P}[v^\top u < \rho_{t_{k-1}}(X, v)] > 0$$

$$\Rightarrow \exists v \in \mathcal{L}_{t_{k-1}, +}^d :$$

$$\mathbb{P} \left[ v^\top u < \operatorname{ess\,sup}_{\mathbb{Q} \in \mathcal{W}_{t_{k-1}}(v)} \left( -\alpha_{t_{k-1}, t_k}(\mathbb{Q}, v) + \mathbb{E}[\rho_{t_k}(X, w_{t_{k-1}}^{t_k}(\mathbb{Q}, v)) \mid \mathcal{F}_{t_{k-1}}^M] \right) \right] > 0.$$

This implies that there exists some  $(Q^*, v) \in \mathcal{W}_{t_{k-1}}$  such that

$$\mathbb{P}\left[v^\top u < -\alpha_{t_{k-1}, t_k}(Q^*, v) + \mathbb{E}[\rho_{t_k}(X, w_{t_{k-1}}^{t_k}(Q^*, v)) \mid \mathcal{F}_{t_{k-1}}^M]\right] > 0;$$

fix this choice of dual variables. By (5.5), it then follows that

$$\begin{aligned} & \mathbb{P}\left[v^\top u < -\alpha_{t_{k-1}, t_k}(Q^*, v) + \mathbb{E}[\rho_{t_k}(X, w_{t_{k-1}}^{t_k}(Q^*, v)) \mid \mathcal{F}_{t_{k-1}}^M]\right] > 0 \\ \Rightarrow & \mathbb{P}\left[v^\top u < -\alpha_{t_{k-1}, t_k}(Q^*, v) + \mathbb{E}[w_{t_{k-1}}^{t_k}(Q^*, v)^\top Z(w_{t_{k-1}}^{t_k}(Q^*, v)) \mid \mathcal{F}_{t_{k-1}}^M]\right] > 0 \\ \Rightarrow & \mathbb{P}\left[v^\top u < -\alpha_{t_{k-1}, t_k}(Q^*, v) + v^\top \mathbb{E}^{Q^*}[Z(w_{t_{k-1}}^{t_k}(Q^*, v)) \mid \mathcal{F}_{t_{k-1}}^M]\right] > 0 \\ \Rightarrow & \mathbb{P}\left[v^\top u < \operatorname{ess\,sup}_{Q \in \mathcal{W}_{t_{k-1}}(v)} \left(-\alpha_{t_{k-1}, t_k}(Q, v) + v^\top \mathbb{E}^Q[Z(w_{t_{k-1}}^{t_k}(Q^*, v)) \mid \mathcal{F}_{t_{k-1}}^M]\right)\right] > 0 \\ \Rightarrow & \mathbb{P}\left[v^\top u < \rho_{t_{k-1}}^v\left(-Z(w_{t_{k-1}}^{t_k}(Q^*, v))\right)\right] > 0 \\ \Rightarrow & u \notin R_{t_{k-1}}\left(-Z(w_{t_{k-1}}^{t_k}(Q^*, v))\right), \end{aligned}$$

which is a contradiction to (5.4).  $\square$

The next proposition provides an SV-BS $\Delta E$  associated to  $R$  which is analogous to the BS $\Delta I$  in Proposition 4.6. The “difference” in the SV-BS $\Delta E$  is formulated by the notion of geometric difference between sets: for  $C, D \subseteq \mathcal{L}^d$ , we define their *geometric difference* by

$$C - \bullet D := \{u \in \mathcal{L}^d : u + D \subseteq C\}.$$

Fix  $k \in \{1, \dots, K\}$  and suppose that  $C, D \subseteq \mathcal{L}_{t_k}^d$ . We often want to consider the difference modified by the halfspace

$$\left(\operatorname{cl}(C + \Gamma_{t_k}(w)) - \bullet D\right) \cap \mathcal{L}_{t_{k-1}}^d$$

for some  $w \in \mathcal{L}_{t_k, +}^d$ . It can be checked that

$$\begin{aligned} \left(\operatorname{cl}(C + \Gamma_{t_k}(w)) - \bullet D\right) \cap \mathcal{L}_{t_{k-1}}^d &= \left(\operatorname{cl}(C + \Gamma_{t_k}(w)) - \bullet \operatorname{cl}(D + \mathcal{L}_{t_k, +}^d)\right) \cap \mathcal{L}_{t_{k-1}}^d \\ &= \left(\operatorname{cl}(C + \Gamma_{t_k}(w)) - \bullet \operatorname{cl}(D + \Gamma_{t_k}(w))\right) \cap \mathcal{L}_{t_{k-1}}^d. \end{aligned}$$

**Proposition 5.2** *Let  $X \in \mathcal{L}_{t_K}^d$ . If  $(\mathcal{Y}_t)_{t \in \mathbb{T}}$  is a process such that  $\mathcal{Y}_{t_k} = R_{t_k}(X)$  for  $k \in \{0, \dots, K\}$ , then the SV-BS $\Delta E$*

$$\begin{aligned} \mathcal{Y}_{t_{k-1}} &= \bigcap_{\substack{w_{t_k} \in \mathcal{L}_{t_k, +}^d, \\ \psi \in \Psi_{t_{k-1}}(w_{t_k})}} \left( G_E(t_{k-1}, \psi, w_{t_k}) \Delta t_k \right. \\ &\quad \left. + (\operatorname{cl}[\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k})] - \bullet [\sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \mathcal{L}_{t_k, +}^d]) \cap \mathcal{L}_{t_{k-1}}^d \right), \\ \mathcal{Y}_{t_K} &= -X + R_{t_K}(0) \end{aligned}$$

holds, where

$$\begin{aligned} & \Psi_{t_{k-1}}(w_{t_k}) \\ &:= \bigcup_{\xi \in \mathcal{L}_{t_{k-1}}^d} \left\{ \psi \in (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}} : \xi + \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \Gamma_{t_k}(w_{t_k}) = \text{cl}(\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k})) \right\}. \end{aligned} \quad (5.6)$$

In the above SV-BSΔE, we make the convention that the intersection over  $\Psi_{t_{k-1}}(w_{t_k})$  gives  $\mathcal{L}_{t_{k-1}}^d$  in case  $\Psi_{t_{k-1}}(w_{t_k}) = \emptyset$ .

Conversely, if there exist a set-valued process  $(\mathcal{Y}_t)_{t \in \mathbb{T}}$  and  $\Psi_{t_{k-1}}(w_{t_k}) \subseteq (\mathcal{L}_{t_{k-1}}^d)^{\mathcal{I}}$  for each  $k \in \{1, \dots, K\}$  and  $w_{t_k} \in \mathcal{L}_{t_k, +}^d$  such that the above SV-BSΔE holds, then  $\mathcal{Y}_{t_k} = R_{t_k}(X)$  for  $k \in \{0, \dots, K\}$ .

*Proof* Let  $(\mathcal{Y}_t)_{t \in \mathbb{T}}$  be a process such that  $\mathcal{Y}_{t_k} = R_{t_k}(X)$  for  $k \in \{0, \dots, K\}$ . By construction of the risk measure, the terminal condition of the SV-BSΔE trivially holds. Consider now  $k \in \{1, \dots, K\}$ . If  $\mathcal{Y}_{t_k} = \emptyset$ , then by multiportfolio time-consistency,  $\mathcal{Y}_{t_{k-1}} = \emptyset$  as well. Thus the SV-BSΔE is satisfied trivially prior to  $t_k$ ; for the remainder of this proof, we therefore assume that  $\mathcal{Y}_{t_k} \neq \emptyset$ .

Utilising Lemma 5.1, where  $H_{t_k}(X, w_{t_k}) = \text{cl}(\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k}))$ , we get

$$\mathcal{Y}_{t_{k-1}} = \bigcap_{w_{t_k} \in \mathcal{L}_{t_k, +}^d} R_{t_{k-1}, t_k}[-H_{t_k}(X, w_{t_k})].$$

Let

$$W_{t_k}(X) := \{w \in \mathcal{L}_{t_k, +}^d : H_{t_k}(X, w_{t_k})(\omega) \neq \mathbb{R}^d, \forall \omega \in \Omega\}. \quad (5.7)$$

We claim that

$$\mathcal{Y}_{t_{k-1}} = \bigcap_{w_{t_k} \in W_{t_k}(X)} R_{t_{k-1}, t_k}[-H_{t_k}(X, w_{t_k})]. \quad (5.8)$$

The “ $\subseteq$ ” part of (5.8) is clear. To prove the “ $\supseteq$ ” part, let  $w_{t_k} \in \mathcal{L}_{t_k, +}^d \setminus W_{t_k}(X)$ . Recall from the previous section that  $\mathcal{A}_{t_k}$  denotes the partition of  $\Omega$  that generates  $\mathcal{F}_{t_k}^M$ . Given  $A \in \mathcal{A}_{t_k}$ , note that  $\mathcal{Y}_{t_k}(\omega)$  is the same nonempty subset of  $\mathbb{R}^d$  for all  $\omega \in A$ ; let us denote this set by  $\mathcal{Y}_{t_k}(A)$ . First, we assume that  $\mathcal{Y}_{t_k}(A) \neq \mathbb{R}^d$ . Since  $\mathcal{Y}_{t_k}(A) \notin \{\emptyset, \mathbb{R}^d\}$ , there exists  $w^A \in \mathbb{R}_+^d \setminus \{0\}$  such that

$$\inf_{u \in \mathcal{Y}_{t_k}(A)} ((w^A)^\top u) \in \mathbb{R}.$$

Similarly, let us denote  $H_{t_k}(X, w_{t_k})(A) = H_{t_k}(X, w_{t_k})(\omega)$  for all  $\omega \in A$ ,  $A \in \mathcal{A}_{t_k}$ . Since  $w_{t_k} \notin W_{t_k}(X)$ , the set  $\mathcal{B} = \{A \in \mathcal{A}_{t_k} : H_{t_k}(X, w_{t_k})(A) = \mathbb{R}^d\}$  is nonempty. Let

$$\tilde{w}_{t_k} := \sum_{A \in \mathcal{B}} w^A \mathbb{1}_A + \sum_{A \in \mathcal{A}_{t_k} \setminus \mathcal{B}} w_{t_k} \mathbb{1}_A \in W_{t_k}(X).$$

Then it can be checked that  $H_{t_k}(X, w_{t_k}) \supseteq H_{t_k}(X, \tilde{w}_{t_k})$ . By monotonicity, it follows that

$$R_{t_{k-1}, t_k}[-H_{t_k}(X, w_{t_k})] \supseteq R_{t_{k-1}, t_k}[-H_{t_k}(X, \tilde{w}_{t_k})].$$

Hence the “ $\supseteq$ ” part of (5.8) follows as long as  $\mathcal{Y}_{t_k}(A) \neq \mathbb{R}^d$  for every  $A \in \mathcal{A}_{t_k}$ . This must be true due to the construction of  $\mathcal{Y}_{t_k}$ . In particular,

$$\mathcal{Y}_{t_k} = R_{t_k}(X) \subseteq R_{t_k}(\|X\|_\infty) = R_{t_k}(0) - \|X\|_\infty$$

due to monotonicity and translativity of the risk measure with

$$\|X\|_\infty := \left( \max_{\omega \in \Omega} |X_1(\omega)|, \dots, \max_{\omega \in \Omega} |X_d(\omega)| \right).$$

Therefore, for any  $A \in \mathcal{A}_{t_k}$  and  $\omega \in A$ ,  $\mathcal{Y}_{t_k}(A) = \mathbb{R}^d$  implies that  $R_{t_k}(0)(\omega) = \mathbb{R}^d$ . However, this violates that the risk measure is normalised, and so  $\mathcal{Y}_{t_k}(A) \neq \mathbb{R}^d$  for every  $A \in \mathcal{A}_{t_k}$ .

Let  $w_{t_k} \in W_{t_k}(X)$ . By the construction of  $H_{t_k}(X, w_{t_k})$  and Lemma 3.1,  $H_{t_k}(X, w_{t_k})$  has the predictable representation

$$H_{t_k}(X, w_{t_k}) = \left( \xi_{t_{k-1}}(w_{t_k}) + \sum_{I \in \mathcal{I}} \psi_{t_{k-1}}^I(w_{t_k}) \Delta M^I(t_k) \right) + \Gamma_{t_k}(w_{t_k})$$

for some  $\xi_{t_{k-1}}(w_{t_k}), \psi_{t_{k-1}}^I(w_{t_k}) \in \mathcal{L}_{t_{k-1}}^d, I \in \mathcal{I}$ . Consider the set  $\Psi_{t_{k-1}}(w_{t_k})$  of all constructors of  $H_{t_k}(X, w_{t_k})$  defined by (5.6). By construction, for  $\psi \in \Psi_{t_{k-1}}(w_{t_k})$ , there exists some  $\xi_{t_{k-1}}(w_{t_k}, \psi) \in \mathcal{L}_{t_{k-1}}^d$  such that

$$H_{t_k}(X, w_{t_k}) = \xi_{t_{k-1}}(w_{t_k}, \psi) + \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \Gamma_{t_k}(w_{t_k}).$$

Therefore,

$$\begin{aligned} & R_{t_{k-1}, t_k}[-H_{t_k}(X, w_{t_k})] \\ &= \bigcap_{\psi \in \Psi_{t_{k-1}}(w_{t_k})} R_{t_{k-1}, t_k} \left[ -\xi_{t_{k-1}}(w_{t_k}, \psi) - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I - \Gamma_{t_k}(w_{t_k}) \right] \\ &= \bigcap_{\psi \in \Psi_{t_{k-1}}(w_{t_k})} R_{t_{k-1}, t_k} \left[ -\xi_{t_{k-1}}(w_{t_k}, \psi) - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I - \Gamma_{t_k}(w_{t_k}) - \Gamma_{t_{k-1}}(w_{t_k}) \right] \\ &= \bigcap_{\psi \in \Psi_{t_{k-1}}(w_{t_k})} \left( R_{t_{k-1}, t_k} \left[ -\sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I - \Gamma_{t_k}(w_{t_k}) \right] \right. \\ &\quad \left. + \xi_{t_{k-1}}(w_{t_k}, \psi) + \Gamma_{t_{k-1}}(w_{t_k}) \right). \end{aligned}$$

Let  $\psi \in \Psi_{t_{k-1}}(w_{t_k})$ . By the construction of  $\Gamma_{t_{k-1}}(w_{t_k}) = \Gamma_{t_k}(w_{t_k}) \cap \mathcal{L}_{t_{k-1}}^d$  and the set-valued subtraction, we have

$$\Gamma_{t_{k-1}}(w_{t_k}) = (\Gamma_{t_k}(w_{t_k}) - \bullet \Gamma_{t_k}(w_{t_k})) \cap \mathcal{L}_{t_{k-1}}^d$$

since  $\Gamma_{t_k}(w_{t_k})$  is a convex cone. This allows us to recover the representation

$$\begin{aligned} & R_{t_{k-1}, t_k} \left[ - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I - \Gamma_{t_k}(w_{t_k}) \right] + \xi_{t_{k-1}}(w_{t_k}, \psi) + \Gamma_{t_{k-1}}(w_{t_k}) \\ &= G_E(t_{k-1}, \psi, w_{t_k}) \Delta t_k + \xi_{t_{k-1}}(w_{t_k}, \psi) \\ &\quad + \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \Gamma_{t_{k-1}}(w_{t_k}) - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I \\ &= G_E(t_{k-1}, \psi, w_{t_k}) \Delta t_k + H_{t_k}(X, w_{t_k}) - \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I \\ &= G_E(t_{k-1}, \psi, w_{t_k}) \Delta t_k + \left( H_{t_k}(X, w_{t_k}) - \bullet \left( \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \Gamma_{t_k}(w_{t_k}) \right) \right) \cap \mathcal{L}_{t_{k-1}}^d \\ &= G_E(t_{k-1}, \psi, w_{t_k}) \Delta t_k \\ &\quad + \left( \text{cl}(\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k})) - \bullet \left( \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \Gamma_{t_k}(w_{t_k}) \right) \right) \cap \mathcal{L}_{t_{k-1}}^d \\ &= G_E(t_{k-1}, \psi, w_{t_k}) \Delta t_k \\ &\quad + \left( \text{cl}(\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k})) - \bullet \left( \sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \mathcal{L}_{t_k, +}^d \right) \right) \cap \mathcal{L}_{t_{k-1}}^d. \end{aligned}$$

Combining the above calculations yields

$$\begin{aligned} & \mathcal{Y}_{t_{k-1}} \\ &= \bigcap_{\substack{w_{t_k} \in W_{t_k}(X), \\ \psi \in \Psi_{t_{k-1}}(w_{t_k})}} \left( G_E(t_{k-1}, \psi, w_{t_k}) \Delta t_k \right. \\ &\quad \left. + (\text{cl}(\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k})) - \bullet (\sum_{I \in \mathcal{I}} \psi^I \Delta M_{t_k}^I + \mathcal{L}_{t_k, +}^d)) \cap \mathcal{L}_{t_{k-1}}^d \right). \end{aligned}$$

Note that for  $w_{t_k} \in \mathcal{L}_{t_k, +}^d \setminus W_{t_k}(X)$ , the definition in (5.6) gives  $\Psi_{t_{k-1}}(w_{t_k}) = \emptyset$  as desired, since no constructor of  $H_{t_k}(X, w_{t_k})$  exists in this case. Hence with our convention that the intersection in the SV-BSΔE over an empty index set gives  $\mathcal{L}_{t_k}^d$ , it follows that the SV-BSΔE is satisfied.

Finally, the converse follows by the same logic as above since the construction of  $\Psi_{t_{k-1}}(w_{t_k})$  guarantees that it is part of a predictable representation of  $H_{t_k}(X, w_{t_k})$ .  $\square$

**Corollary 5.3** Fix some positive vector  $\mathbf{r} \in \mathbb{R}_{++}^d$ . Let  $X \in \mathcal{L}_{t_K}^d$ . Let  $(\mathcal{Y}_t)_{t \in \mathbb{T}}$  be a process such that  $\mathcal{Y}_{t_k} = R_{t_k}(X)$  for  $k \in \{0, \dots, K\}$ . Then for each  $k \in \{1, \dots, K\}$ ,

there exists a nonempty set  $W_{t_k} \subseteq \mathcal{L}_{t_k}^d$  such that for each  $w_{t_k} \in W_{t_k}$ , there exists a unique pair  $(\hat{\xi}_{t_{k-1}}(w_{t_k}), \hat{\psi}_{t_{k-1}}(w_{t_k})) \in \mathcal{L}_{t_{k-1}} \times (\mathcal{L}_{t_{k-1}})^{\mathcal{I}}$  satisfying

$$\left( \hat{\xi}_{t_{k-1}}(w_{t_k}) + \sum_{I \in \mathcal{I}} \hat{\psi}_{t_{k-1}}^I(w_{t_k}) \Delta M_{t_k}^I \right) \mathbf{r} + \Gamma_{t_k}(w_{t_k}) = \text{cl}(\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k})), \quad (5.9)$$

and there exists no such pair for  $w_{t_k} \in \mathcal{L}_{t_k,+}^d \setminus W_{t_k}$ . Moreover, the reformulated SV-BSΔE

$$\mathcal{Y}_{t_{k-1}} = \bigcap_{w_{t_k} \in W_{t_k}} \left( G_E(t_{k-1}, \hat{\psi}_{t_{k-1}}(w_{t_k}) \mathbf{r}, w_{t_k}) \Delta t_k + (\text{cl}(\mathcal{Y}_{t_k} + \Gamma_{t_k}(w_{t_k})) - \bullet (\sum_{I \in \mathcal{I}} \hat{\psi}_{t_{k-1}}^I(w_{t_k}) \mathbf{r} \Delta M_{t_k}^I + \mathcal{L}_{t_k,+}^d)) \cap \mathcal{L}_{t_{k-1}}^d \right),$$

$$\mathcal{Y}_{t_K} = -X + R_{t_K}(0)$$

holds.

*Proof* Let us take  $W_{t_k} = W_{t_k}(X)$ , where  $W_{t_k}(X)$  is defined by (5.7). Fixing the direction  $\mathbf{r} \in \mathbb{R}_{++}^d$  guarantees the uniqueness of the predictable representation of  $H_{t_k}(X, w_{t_k})$  for  $w_{t_k} \in W_{t_k}$  in the form (5.9). As noted in the proof of Proposition 5.2, no predictable representation exists for  $H_{t_k}(X, w_{t_k})$  when  $w_{t_k} \in \mathcal{L}_{t_k,+}^d \setminus W_{t_k}$ . It follows that the reformulated SV-BSΔE is equivalent to that given in Proposition 5.2.  $\square$

*Remark 5.4* The SV-BSΔEs given in Proposition 5.2 and all subsequent results of this section can utilise the dual variables  $(\mathbb{Q}, w_{t_{k-1}}) \in \mathcal{W}_{t_{k-1}}$  in place of  $w_{t_k} \in \mathcal{L}_{t_k,+}^d$  by considering  $w_{t_k} := w_{t_{k-1}}^{t_k}(\mathbb{Q}, w_{t_{k-1}})$ . This follows from an application of Feinstein and Rudloff [17, Lemma A.1].

*Remark 5.5* While in the BSΔI framework, we introduce semi-local and local versions, this does not appear to be possible for the SV-BSΔE setup. Specifically, the Minkowski difference restricted to an earlier time point as taken in the SV-BSΔE can not readily be defined  $\omega$ -wise. Conceptually, this can be viewed as akin to the work of Ben Tahar and Lépinette [6] insofar as the recursive formulation of multiportfolio time-consistency is defined with respect to the selectors rather than the random sets.

We conclude this section with a continuation of Examples 3.5 and 4.8 on the superhedging risk measure.

*Example 5.6* In the setting of Example 3.5, let us calculate the set-valued driver  $G_E^{\text{SHP}}$  of the superhedging risk measure  $R^{\text{SHP}}$  when  $d = 2$  and  $m = 1$ . Recall that as provided in Example 3.5 and presented in Example 4.8, for  $s \in \{t + 1, \dots, T\}$ , we have

$$S_s^b = S_t^b e^{\sigma B_{t+1}} e^{\sigma(M_s - M_{t+1})}, \quad S_s^a = S_t^a e^{\sigma B_{t+1}} e^{\sigma(M_s - M_{t+1})}.$$

As with Examples 4.8 and 4.9, we can generalise to larger number of assets and random walks, but focus on this simple case for illustrative purposes. Following (5.1)

and the definitions in Example 3.5, for  $t \in \{0, \dots, T-1\}$ ,  $\psi \in \mathcal{L}_t^2$  and  $w \in \mathcal{L}_{t+1,+}^2$ , we have

$$\begin{aligned}
 & G_E^{\text{SHP}}(t, \psi, w) \\
 &= R_{t,t+1}^{\text{SHP}}[-\psi \Delta M_{t+1} - \Gamma_{t+1}(w)] = \text{SHP}_t[\psi B_{t+1} + \Gamma_{t+1}(w)] \\
 &= \left\{ Y \in \mathcal{L}_t^2 : Y \in \psi B_{t+1} + u + \sum_{s=t}^T L_d^0(\mathcal{F}_s^M, K_s), u \in \mathcal{L}_{t+1}^2, w^\top u \geq 0 \right\} \\
 &= L_d^0(\mathcal{F}_t^M, K_t) + \left\{ Y \in \mathcal{L}_t^2 : Y \in \psi B_{t+1} + u + \sum_{s=t+1}^T L_d^0(\mathcal{F}_s^M, K_s), \right. \\
 &\quad \left. u \in \mathcal{L}_{t+1}^2, w^\top u \geq 0 \right\} \\
 &= L_d^0(\mathcal{F}_t^M, K_t) + \left\{ Y \in \mathcal{L}_t^2 : Y = \psi B_{t+1} + u + \sum_{s=t+1}^T \left( b_s \begin{pmatrix} -S_s^b \\ 1 \end{pmatrix} + a_s \begin{pmatrix} S_s^a \\ -1 \end{pmatrix} \right), \right. \\
 &\quad \left. a_s, b_s \in \mathcal{L}_{s,+}^2, s \in \{t+1, \dots, T\}, \right. \\
 &\quad \left. u \in \mathcal{L}_{t+1}^2, w^\top u \geq 0 \right\}.
 \end{aligned}$$

As stated previously, unlike in the BSΔI setup of Sect. 4, no *local* driver can be given for the SV-BSΔE formulation of the superhedging risk measure.

## 6 Discussion

In this work, we have introduced two backward representations for multiportfolio time-consistent dynamic set-valued risk measures: a BSΔI and an SV-BSΔE. Though both of these representations provide an equivalent dynamic risk measure, these formulations provide important insights for considering dynamic risk measures in continuous time where  $\mathbb{T} = [0, T]$ . Specifically, in continuous time, either a backward stochastic differential inclusion (BSDI) or a set-valued backward stochastic differential equation (SV-BSDE) could potentially be used to characterise a dynamic risk measure. This has yet to be examined in the literature. Our work on studying difference inclusions and equations provides the initial insights for when these concepts are appropriate to be applied. In this way, we gain knowledge of the (likely) best approach for studying risk measures in continuous time. In fact, based on the preceding analysis, it is the opinion of the authors that BSDIs are the proper methodology to consider. While presented only for risk measures, our results provide insights for the dynamic programming principle in multivariate problems more generally.

By studying both the BSΔI and SV-BSΔE, we begin to understand when these approaches are appropriate. Namely, the inclusion appears to be the appropriate method if we care specifically about singular paths. For instance, risk measures are used to

compute capital requirements. Ultimately, the singular, implemented capital investment over time is the important result for a practitioner rather than the entire set of acceptable requirements. In contrast, a set-valued equation appears to be the appropriate method if we care about the “mass” of the set itself over time, rather than any specific value in that set. Such concepts are important beyond the immediate study of risk measures, for instance for the mean–risk problem in Kováčová and Rudloff [38], where the dynamic programming principle holds for the multiobjective version but *not* the traditional scalar approach. These results indicate that the BSDI is likely to be the appropriate approach for that problem; this is left for future study.

Following from the interpretation of risk measures, we talk about the set of acceptable capital requirements. Thus the individual requirements are the important notion themselves in our setting. This leads to the notion that BSDIs appear to be the appropriate methodology for us to consider. Though our recursive formulation of multiperiod time-consistency and the dynamic programming principle are defined with respect to the risk measure of sets (i.e.,  $R_t(X) = R_{t,s}[-R_s(X)]$ ), this is ultimately defined element-wise. We conjecture that SV-BSDEs would be the only available method if set-valued portfolios (e.g. Cascos and Molchanov [9]) or, generally, functions of sets were considered themselves.

While we are able to construct an SV-BSΔE for this setting, this does not permit an immediate integral (summation in the discrete setting) representation when incrementing from time  $t$  to  $T$  directly. To pass to the continuous-time BSDI limit of the BSΔI, we propose following the approach of Stadje [49] in which the risk measures are “scaled and tilted” first. We conjecture that convergence of the BSΔI should be sought for the space of reachable sets as constructed in Corollary 4.3; this is in contrast to studying the individual paths. However, we leave that consideration for future works.

**Acknowledgements** The authors would like to thank two anonymous referees and the Associate Editor for their detailed feedback which was helpful in improving the manuscript. This work was initiated and an early version of it was presented at the two meetings on Dynamic Multivariate Programming (2018, 2019) hosted and funded by the Vienna University of Economics and Business. The authors would like to thank Birgit Rudloff for organising these meetings as well as all the participants for fruitful discussions.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

1. Akahori, J., Amaba, T., Okuma, K.: A discrete-time Clark–Ocone formula and its application to an error analysis. *J. Theor. Probab.* **30**, 932–960 (2017)
2. Ararat, Ç., Hamel, A.H., Rudloff, B.: Set-valued shortfall and divergence risk measures. *Int. J. Theor. Appl. Finance* **20**, 1750026 (2017)
3. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. *Math. Finance* **9**, 203–228 (1999)
4. Barrieu, P., El Karoui, N.: Optimal derivative design under dynamic risk measures. In: Ying, G., Zhang, Q. (eds.) *Mathematics of Finance*, pp. 13–26. Am. Math. Soc., Providence (2004)
5. Barrieu, P., El Karoui, N.: Pricing, hedging and optimally designing derivatives via minimization of risk measures. In: Carmona, R. (ed.) *Indifference Pricing: Theory and Applications*, pp. 77–146. Princeton University Press, Princeton (2009)



6. Ben Tahar, I., Lépinette, E.: Vector-valued coherent risk measure processes. *Int. J. Theor. Appl. Finance* **17**, 1450011–1–28 (2014)
7. Biagini, F., Fouque, J.P., Frittelli, M., Meyer-Brandis, T.: A unified approach to systemic risk measures via acceptance sets. *Math. Finance* **9**, 329–367 (2019)
8. Bion-Nadal, J.: Conditional risk measures and robust representation of convex risk measures. Ecole Polytechnique CMAP preprint no. 557 (2004). Available online at <http://www.cmap.polytechnique.fr/preprint/repository/557.pdf>
9. Cascos, I., Molchanov, I.: Multivariate risk measures: a constructive approach based on selections. *Math. Finance* **26**, 867–900 (2016)
10. Chen, Y., Hu, Y.: Time consistency for set-valued dynamic risk measures for bounded discrete-time processes. *Math. Financ. Econ.* **12**, 305–333 (2018)
11. Cheridito, P., Delbaen, F., Kupper, M.: Dynamic monetary risk measures for bounded discrete-time processes. *Electron. J. Probab.* **11**, 57–106 (2006)
12. Cheridito, P., Horst, U., Kupper, M., Pirvu, T.A.: Equilibrium pricing in incomplete markets under translation invariant preferences. *Math. Oper. Res.* **41**, 174–195 (2016)
13. Crespi, G.P., Hamel, A.H., Schrage, C.: A Minty variational principle for set optimization. *J. Math. Anal. Appl.* **423**, 770–796 (2015)
14. Detlefsen, K., Scandolo, G.: Conditional and dynamic convex risk measures. *Finance Stoch.* **9**, 539–561 (2005)
15. Eisenberg, L., Noe, T.H.: Systemic risk in financial systems. *Manag. Sci.* **47**, 236–249 (2001)
16. Feinstein, Z., Rudloff, B.: Time consistency of dynamic risk measures in markets with transaction costs. *Quant. Finance* **13**, 1473–1489 (2013)
17. Feinstein, Z., Rudloff, B.: Multi-portfolio time consistency for set-valued convex and coherent risk measures. *Finance Stoch.* **19**, 67–107 (2015)
18. Feinstein, Z., Rudloff, B.: A comparison of techniques for dynamic multivariate risk measures. In: Hamel, A.H., et al. (eds.) *Set Optimization and Applications – the State of the Art. From Set Relations to Set-valued Risk Measures*, pp. 3–41. Springer, Berlin (2015)
19. Feinstein, Z., Rudloff, B.: A recursive algorithm for multivariate risk measures and a set-valued Bellman’s principle. *J. Glob. Optim.* **68**, 47–69 (2017)
20. Feinstein, Z., Rudloff, B.: A supermartingale relation for multivariate risk measures. *Quant. Finance* **18**, 1971–1990 (2018)
21. Feinstein, Z., Rudloff, B.: Scalar multivariate risk measures with a single eligible asset. Working paper (2019). Available online at <http://arxiv.org/abs/1807.10694>
22. Feinstein, Z., Rudloff, B.: Time consistency for scalar multivariate risk measures. Working paper (2020). Available online at <http://arxiv.org/abs/1810.04978>
23. Feinstein, Z., Rudloff, B., Weber, S.: Measures of systemic risk. *SIAM J. Financ. Math.* **8**, 672–708 (2017)
24. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. *Finance Stoch.* **6**, 429–447 (2002)
25. Föllmer, H., Schied, A.: *Stochastic Finance: An Introduction in Discrete Time*, 3rd edn. de Gruyter, Berlin (2011)
26. Grépat, J., Kabanov, Yu.: On a multi-asset version of the Kusuoka limit theorem of option superreplication under transaction costs. *Finance Stoch.* **25**, 167–187 (2021)
27. Hamel, A.H., Heyde, F., Rudloff, B.: Set-valued risk measures for conical market models. *Math. Financ. Econ.* **5**, 1–28 (2011)
28. Hamel, A.H., Schrage, C.: Directional derivatives and subdifferential of set-valued convex functions. *Pac. J. Optim.* **10**, 667–689 (2014)
29. Hamel, A.H., Wang, S.Q.: A set optimization approach to utility maximization under transaction costs. *Decis. Econ. Finance* **40**, 257–275 (2017)
30. He, H.: Convergence from discrete to continuous time contingent claim prices. *Rev. Financ. Stud.* **3**, 523–546 (1990)
31. Jacod, J., Méléard, S., Protter, P.: Explicit form and robustness of martingale representations. *Ann. Probab.* **28**, 1747–1780 (2000)
32. Jiang, L.: Convexity, translation invariance and subadditivity for  $g$ -expectations and related risk measures. *Ann. Appl. Probab.* **18**, 245–258 (2008)
33. Jouini, E., Meddeb, M., Touzi, N.: Vector-valued coherent risk measures. *Finance Stoch.* **8**, 531–552 (2004)

34. Karnam, C., Ma, J., Zhang, J.: Dynamic approaches for some time-inconsistent optimization problems. *Ann. Appl. Probab.* **27**, 3435–3477 (2017)
35. Kisielewicz, M.: Backward stochastic differential inclusions. *Dyn. Syst. Appl.* **16**, 121–140 (2007)
36. Kisielewicz, M.: Weak compactness of weak solutions to backward stochastic differential inclusions. *Dyn. Syst. Appl.* **17**, 351–370 (2008)
37. Kisielewicz, M.: *Stochastic Differential Inclusions and Applications*. Springer, New York (2013)
38. Kováčová, G., Rudloff, B.: Time consistency of the mean-risk problem. *Oper. Res.*, forthcoming (2020). Available online at <https://arxiv.org/abs/1806.10981>
39. Lépinette, E., Molchanov, I.: Conditional cores and conditional convex hulls of random sets. *J. Math. Anal. Appl.* **478**, 368–392 (2019)
40. Lépinette, E., Molchanov, I.: Risk arbitrage and hedging to acceptability under transaction costs. *Finance Stoch.* **25**, 101–132 (2021)
41. Löhne, A., Rudloff, B.: An algorithm for calculating the set of superhedging portfolios in markets with transaction costs. *Int. J. Theor. Appl. Finance* **17**, 1450012 (2014)
42. Molchanov, I.: *Theory of Random Sets*, 2nd edn. Springer, London (2017)
43. Molchanov, I., Mühlemann, A.: Nonlinear expectations of random sets. *Finance Stoch.* **25**, 5–41 (2021)
44. Peng, S.: Dynamical evaluations. *C. R. Acad. Sci. Paris, Ser. I* **339**, 585–589 (2004)
45. Protter, P.: *Stochastic Integration and Differential Equations*, 2nd edn. Springer, Berlin (2005), version 2.1
46. Riedel, F.: Dynamic coherent risk measures. *Stoch. Process. Appl.* **112**, 185–200 (2004)
47. Rosazza Gianin, E.: Risk measures via  $g$ -expectations. *Insur. Math. Econ.* **39**, 19–34 (2006)
48. Ruszczyński, A., Shapiro, A.: Conditional risk mappings. *Math. Oper. Res.* **31**, 544–561 (2006)
49. Stadje, M.: Extending dynamic convex risk measures from discrete time to continuous time: a convergence approach. *Insur. Math. Econ.* **47**, 391–404 (2010)
50. Xu, Y.: Multidimensional dynamic risk measure via conditional  $g$ -expectation. *Math. Finance* **26**, 638–673 (2016)