



Mathematical Statistics  
Stockholm University

**Approximations of multi-period liability values by  
simple formulas**

Nils Engler  
Filip Lindskog

**Research Report 2023:1**

ISSN 1650-0377

**Postal address:**

Mathematical Statistics  
Dept. of Mathematics  
Stockholm University  
SE-106 91 Stockholm  
Sweden

**Internet:**

<http://www.math.su.se>



# Approximations of multi-period liability values by simple formulas

Nils Engler and Filip Lindskog

January 2023

## Abstract

This paper is motivated by computational challenges arising in multi-period valuation in insurance. Aggregate insurance liability cashflows typically correspond to stochastic payments several years into the future. However, insurance regulation requires that capital requirements are computed for a one-year horizon, by considering cashflows during the year and end-of-year liability values. This implies that liability values must be computed recursively, backwards in time, starting from the year of the most distant liability payments. Solving such backward recursions with paper and pen is rarely possible, and numerical solutions give rise to major computational challenges.

The aim of this paper is to provide explicit and easily computable expressions for multi-period valuations that appear as limit objects for a sequence of multi-period models that converge in terms of conditional weak convergence. Such convergence appears naturally if we consider large insurance portfolios such that the liability cashflows, appropriately centered and scaled, converge weakly as the size of the portfolio tends to infinity.

# 1 Introduction and motivation

This paper is motivated by problems arising in multi-period valuation in insurance, but the applicability of the results extends beyond the insurance setting. Aggregate insurance liability cashflows typically correspond to payments several years into the future. However, insurance regulation requires that capital requirements are calculated for a one-year horizon, considering cashflows during the year and end-of-year liability values. This implies that liability values must be computed recursively, backwards in time, starting from the year of the most distant liability payments. Solving such backward recursions leads to major computational challenges. Consequently, simple and explicit (standard) formulas with questionable conceptual soundness have become the preferred alternative. With this paper we say that there are alternatives to commonly encountered formulas that give rise to explicit formulas that are easy to use in practice and that retain both economic interpretability and the conceptual soundness of the original principles for insurance valuation. Mathematically, the explicit formulas follow from combining the original principles for insurance valuation with widely applicable large portfolio asymptotics.

We begin by presenting the motivating insurance problem. Let  $(C^{n_0}, \mathbb{F}^{n_0})$  be a stochastic processes in discrete time with  $(C_t^{n_0})_{t=1}^T$  denoting the discounted payments at times  $t$  due to insurance claims, and  $\mathbb{F}^{n_0} = (\mathcal{F}_t^{n_0})_{t=0}^T$ , with  $\mathcal{F}_0^{n_0} = \{\Omega, \emptyset\}$ , denoting a filtration to which  $C^{n_0}$  is adapted.  $C^{n_0}$  describes discounted claim payments over time in a so-called runoff situation, when no new contracts are written. The number  $n_0$  is intended as a measure of volume or exposure, for instance the number of contracts that may generate future claims payments. We want to assign a value to this liability cashflow. In order to do this we consider a sequence  $(C^n, \mathbb{F}^n)_{n \geq 1}$  of such stochastic processes, with  $\|C^n\| \rightarrow \infty$  almost surely as  $n \rightarrow \infty$ , where  $\|\cdot\|$  denotes the Euclidean norm on  $\mathbb{R}^T$ . An example could be  $C^n$  given by

$$C_t^n = \sum_{k=1}^{M_n} 1_{\{D_k=t\}} Z_k, \quad t = 1, \dots, T,$$

where  $M_n$  denotes the total number of claims payments and  $M_n \rightarrow \infty$  almost surely as  $n \rightarrow \infty$ ,  $D_k$  denotes the time and  $Z_k$  denotes the size of the  $k$ th claims payment. For this example, the filtration could be the one generated by the discounted cashflow  $C^n$  or a larger filtration also including information about the number of payments  $\sum_{k=1}^{M_n} 1_{\{D_k=t\}}$  at each point in time, or more.

We assume nonrandom sequences with terms  $a_n \in (0, \infty)$  and  $b_n \in \mathbb{R}^T$  such that there is convergence in distribution

$$X^n = a_n^{-1}(C^n - b_n) \xrightarrow{d} X \quad \text{as } n \rightarrow \infty.$$

The value  $V_0^n(C^n)$  of the insurance liabilities in a multi-period model is the result of applying a suitable functional to the pair  $(C^n, \mathbb{F}^n)$ . Natural valuation functionals satisfy the property  $V_0^n(C^n) = a_n V_0^n(X^n) + \sum_{s=1}^T b_{n,s}$ . Convergence  $X^n \xrightarrow{d} X$  suggests  $V_0^n(X^n) \rightarrow V_0(X)$  and therefore the approximation

$$V_0^n(C^n) \approx a_n V_0(X) + \sum_{s=1}^T b_{n,s}. \quad (1)$$

However, the flow of information over time is an essential ingredient in valuation and it is not true that convergence  $X^n \xrightarrow{d} X$  implies the convergence  $V_0^n(X^n) \rightarrow V_0(X)$  (even if  $\mathbb{F}^n$

is taken to be the filtration generated by  $C^n$ ). Consequently, one of the main objectives of the present paper is to determine sharp sufficient conditions for the convergence of  $V_0^n(X^n)$  to  $V_0(X)$  in terms of an appropriate mode of convergence of  $(X^n, \mathbb{F}^n)$  to  $(X, \mathbb{F})$ , similar to so-called extended weak convergence introduced by Aldous in [2] and conditional weak convergence studied by Sweeting in [13]. The importance of the approximation (1) is because  $V_0^n(C^n)$  is typically very difficult to compute numerically whereas  $V_0(X)$  is easier to compute numerically and, more importantly, in the case of a Gaussian limit model, is given by an explicit expression in terms of conditional variances of components of  $X$  (note that Gaussian vectors have the rare feature that conditional variances of one component given a subset of components are nonrandom).

The paper is organized as follows. Section 2 introduces notation and basic properties of conditional distributions and risk measures. Section 3 presents the main contents of the paper and, following a general presentation of the mathematical setup, gives economic motivation of key quantities in Section 3.1 and presents the main results in Section 3.2. All proofs together with auxiliary results are found in Section 4.

## 2 Preliminaries

$\mathbb{N} = \{1, 2, \dots\}$ ,  $\mathbb{R}$  denotes the real numbers and  $\mathbb{R}_+ = [0, \infty)$ . Whenever relevant, for  $T \in \mathbb{N}$ , a vector  $x \in \mathbb{R}^T$  is assumed to be column vector and its transpose  $x^T$  a row vector. For  $x = (x_1, \dots, x_T)^T \in \mathbb{R}^T$ ,  $\|x\|^2 = x^T x$  and we let  $x_{\leq t} = (x_1, \dots, x_t)^T$  and  $x_{> t} = (x_{t+1}, \dots, x_T)^T$ . For  $d \in \mathbb{N}$  and a Borel set  $A \in \mathcal{B}(\mathbb{R}^d)$ ,  $\mathcal{P}(A)$  denotes the set of probability measures on  $A$ . For  $\mu \in \mathcal{P}(A)$ ,  $\text{supp}(\mu)$  denotes its support. We consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . For a  $\sigma$ -algebra  $\mathcal{F}_t \subset \mathcal{F}$ ,  $L^0(\mathcal{F}_t, \mathbb{P})$  denotes the vector space of all real-valued  $\mathcal{F}_t$ -measurable random variables, and for  $p \in (0, \infty)$ ,  $L^p(\mathcal{F}_t, \mathbb{P})$  denotes the subset  $\{Y \in L^0(\mathcal{F}_t, \mathbb{P}) : \mathbb{E}[|Y|^p] < \infty\}$ .

For an  $\mathbb{R}^d$ -valued random variable  $Y$  we let  $\mathcal{L}(Y)$  denote its distribution, i.e. the induced probability measure  $\mathbb{P}(Y \in \cdot)$  on  $\mathbb{R}^d$ . Given  $\sigma$ -algebras  $\mathcal{F}_t \subset \mathcal{F}$  and a random variable  $Y$ , a regular conditional distribution  $Q_{\mathcal{F}_t, Y}$  is a version of  $\mathbb{P}(Y \in \cdot \mid \mathcal{F}_t)$  which forms a probability kernel from  $(\Omega, \mathcal{F}_t)$  to  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . If  $\mathcal{F}_t = \sigma(Z)$  for an  $\mathbb{R}^d$ -valued random variable  $Z$ , then  $Q_{\mathcal{F}_t, Y}$  is an  $\mathcal{F}_t$ -measurable random measure  $\omega \mapsto \kappa(Z(\omega), \cdot)$  on  $\mathbb{R}^d$ , where  $\kappa$  is a probability kernel from  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  to  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ , see Theorem 6.3 in Kallenberg [8]. The notation  $\mathcal{L}(Y \mid Z = z)$  means  $\kappa(z, \cdot)$ . In particular,  $z \mapsto \mathcal{L}(Y \mid Z = z)$  is well defined on  $\text{supp}(\mathcal{L}(Z))$ .

We define the conditional  $p$ -quantile of  $Y \in L^0(\mathcal{F}, \mathbb{P})$  given  $\mathcal{F}_t$  as the random variable  $F_{Y|\mathcal{F}_t}^{-1}(p)$  given by

$$\omega \mapsto \min\{m \in \mathbb{R} : Q_{\mathcal{F}_t, Y}(\omega, (-\infty, m]) \geq p\}.$$

If  $Q_{\mathcal{F}_t, Y}(\omega, \cdot) = \kappa(Z(\omega), \cdot)$ , then  $F_{Y|\mathcal{F}_t}^{-1}(p)$  means  $\min\{m \in \mathbb{R} : \kappa(z, (-\infty, m]) \geq p\}$  and is well defined on  $\text{supp}(\mathcal{L}(Z))$ . Similar to ordinary quantiles, the conditional quantiles have the property  $F_{aY+b|\mathcal{F}_t}^{-1}(p) = aF_{Y|\mathcal{F}_t}^{-1}(p) + b$  if  $a \in \mathbb{R}_+$  and  $b \in L^0(\mathcal{F}_t, \mathbb{P})$  (in particular if  $b \in \mathbb{R}$ ). Monetary conditional risk measures appear naturally for multi-period valuations, see e.g. Chapter 11 in Föllmer and Schied [6]. The risk measures Value-at-Risk and Average Value-at-Risk, conditional on  $\mathcal{F}_t$  and for  $u \in (0, 1)$ , are given by

$$\begin{aligned} \text{V@R}_u(Y \mid \mathcal{F}_t) &= F_{-Y|\mathcal{F}_t}^{-1}(1-u), \\ \text{AV@R}_u(Y \mid \mathcal{F}_t) &= \frac{1}{u} \int_0^u \text{V@R}_v(Y \mid \mathcal{F}_t) dv. \end{aligned}$$

In [6],  $V@R_u(Y | \mathcal{F}_t)$  is defined as  $\text{essinf}\{m_t \in L^0(\mathcal{F}_t, \mathbb{P}) : \mathbb{P}(Y + m_t < 0 | \mathcal{F}_t) \leq u\}$  but the two definitions of  $V@R_u(Y | \mathcal{F}_t)$  are equal almost surely. For  $p \geq 1$ ,  $V@R_u(Y | \mathcal{F}_t) \in L^p(\mathcal{F}_t, \mathbb{P})$  if  $Y \in L^p(\mathcal{F}, \mathbb{P})$ , and similarly for  $AV@R_u(Y | \mathcal{F}_t)$ . Let  $\mathcal{P}([0, 1])'$  denote the subset of  $\mathcal{P}([0, 1])$  consisting of measures  $\mu \in \mathcal{P}([0, 1])$  that either have a bounded density with respect to Lebesgue measure or satisfy  $\text{supp}(\mu) \subset [a, b]$  for some  $0 < a < b < 1$ . Both  $V@R_u(Y | \mathcal{F}_t)$  and  $AV@R_u(Y | \mathcal{F}_t)$  can be expressed as

$$\rho(Y | \mathcal{F}_t) = \int_0^1 F_{-Y|\mathcal{F}_t}^{-1}(p) \mu(dp), \quad \mu \in \mathcal{P}([0, 1])'. \quad (2)$$

$V@R_u(Y | \mathcal{F}_t)$  corresponds to  $\mu(dp) = \delta_{1-u}(dp)$  (a unit point mass at  $1 - u$ ) and  $AV@R_u(Y | \mathcal{F}_t)$  corresponds to  $\mu(dp) = u^{-1}I\{p \in [1 - u, 1]\}(dp)$  (a bounded density). Conditional risk measures of the form (2) satisfy  $\rho(aY + b | \mathcal{F}_t) = a\rho(Y | \mathcal{F}_t) - b$  if  $a \in \mathbb{R}_+$  and  $b$  is  $\mathcal{F}_t$ -measurable, called positive homogeneity and conditional cash additivity. Moreover, they are monotone:  $\tilde{Y} \geq Y$  implies  $\rho(\tilde{Y} | \mathcal{F}_t) \leq \rho(Y | \mathcal{F}_t)$ . Throughout the paper (in)equalities between random variables should be interpreted in the almost sure sense.

### 3 Convergence of multi-period valuations

Fix  $T \in \mathbb{N}$  and let  $X, X^1, X^2, \dots$  be random vectors in  $\mathbb{R}^T$  and let  $Y, Y^1, Y^2, \dots$  be random vectors in  $(\mathbb{R}^d)^T$ ,  $d \in \mathbb{N}$ . Suppose that  $\mathcal{L}(X^n, Y^n) \xrightarrow{w} \mathcal{L}(X, Y)$  as  $n \rightarrow \infty$ , where  $\mathcal{L}(X, Y)$  is Gaussian. Set  $\mathcal{F}_t = \sigma((X, Y)_{\leq t})$ ,  $\mathcal{F}_t^n = \sigma((X^n, Y^n)_{\leq t})$ .  $(X, Y)$  and  $(X^n, Y^n)$  are adapted discrete-time stochastic processes with respect to the filtrations  $(\mathcal{F}_t)_{t=0}^T$  and  $(\mathcal{F}_t^n)_{t=0}^T$ , respectively, where  $\mathcal{F}_0 = \mathcal{F}_0^n = \{\Omega, \emptyset\}$ . For each  $n$ ,  $X^n = (X_t^n)_{t=1}^T$  corresponds to a discounted incremental cashflow in a multi-period model with  $T$  periods and time points  $0, 1, \dots, T$ . For each  $n$ ,  $Y^n = (Y_t^n)_{t=1}^T$  corresponds to a stochastic process that provides additional information, additional to the information provided by the process  $X^n$ . Taking  $Y^n$  to be nonrandom means that no such additional information is considered and that  $(\mathcal{F}_t^n)_{t=0}^T$  is the natural filtration generated by  $X^n$ . The discounted value at time  $t$  of the cashflow occurring after time  $t$  is denoted by  $V_t^n(X^n)$ . Since no cashflows occur after time  $T$ ,  $V_T^n(X^n) = 0$ . The process of values of the cashflow  $X^n$  is  $(V_t^n(X^n))_{t=0}^T$ . Both  $X_t^n$  and  $V_t^n(X^n)$  are  $\mathcal{F}_t^n$ -measurable. The processes  $(V_t(X))_{t=0}^T$  and  $(V_t^n(X^n))_{t=0}^T$  are defined backward recursively by

$$V_T(X) = 0, \quad V_t(X) = \varphi_t(X_{t+1} + V_{t+1}(X)), \quad t < T, \quad (3)$$

$$V_T^n(X^n) = 0, \quad V_t^n(X^n) = \varphi_t^n(X_{t+1}^n + V_{t+1}^n(X^n)), \quad t < T, \quad (4)$$

where  $\varphi_t : L^1(\mathcal{F}_T) \rightarrow L^1(\mathcal{F}_t)$  and  $\varphi_t^n : L^1(\mathcal{F}_T^n) \rightarrow L^1(\mathcal{F}_t^n)$  are mappings that are made precise below. In particular, the mappings  $\varphi_t$  and  $\varphi_t^n$  will satisfy

$$\varphi_t(aY + \tilde{Y}) = a\varphi_t(Y) + \tilde{Y}, \quad \varphi_t^n(aY + \tilde{Y}) = a\varphi_t^n(Y) + \tilde{Y} \quad \text{for } a \in \mathbb{R}_+, \tilde{Y} \in L^1(\mathcal{F}_t). \quad (5)$$

These properties are called positive homogeneity (the effect of multiplication by a non-negative scalar) and conditional cash additivity. As a consequence, for  $a \in \mathbb{R}_+$ ,  $b \in \mathbb{R}^T$ ,

$$V_t(aX + b) = aV_t(X) + \sum_{s=t+1}^T b_s, \quad V_t^n(aX^n + b) = aV_t^n(X^n) + \sum_{s=t+1}^T b_s. \quad (6)$$

Theorems 1 and 2 presented below essentially say that if  $\mathcal{L}(X^n, Y^n) \xrightarrow{w} \mathcal{L}(X, Y)$  as  $n \rightarrow \infty$  and also in terms of conditional distributions (see e.g. (15)) together with a uniform

integrability assumption (see e.g. (16)), then  $\lim_{n \rightarrow \infty} V_0^n(X^n) = V_0(X)$ , where the values  $V_0^n(X^n)$  and  $V_0(X)$  are defined with respect to the filtrations generated by  $(X^n, Y^n)$  and  $(X, Y)$ , respectively. The essential point is that  $V_0^n(X^n)$  is typically impossible to compute analytically and only with significant difficulties numerically, whereas  $V_0(X)$  has an explicit expression that is straightforward to compute in terms of the mean vector and covariance matrix of the Gaussian weak limit (see e.g. (20)). Therefore, the theorems we present enable the use of conceptually sound valuation techniques without the significant efforts otherwise needed to obtain numerical solutions to backward recursions.

### 3.1 Economic motivation

Current regulatory frameworks for the insurance industry prescribe so-called cost-of-capital valuation of liability cashflows. Such approaches to valuation consider capital requirements and the costs stemming from the financing of buffer capital. In the multi-period setting, with capital requirements determined one period at the time, the randomness of future capital requirements and associated costs drive the liability valuation.

Multiperiod cost-of-capital valuation is studied in Salzmann and Wüthrich [12], Möhr [9], Pelsser and Salahnejhad Ghalehjooghi [10] and Engsner et al. [5]. A common theme is that multiperiod valuations is constructed through backward induction of one-period valuations. Pelsser and Salahnejhad Ghalehjooghi [10] study continuous-time limits of multiperiod valuations defined in terms of one-step valuations, similar to those considered below and including cost-of-capital valuation. In [10] the convergence takes place as the length of the time periods tends to zero and the number of time periods tends to infinity. Clearly, this is a different kind of convergence than that studied here. Multiperiod valuations are so-called time-consistent by their construction through the backward induction of one-period valuations. For mathematical properties of general multiperiod valuations, we refer to Cheridito et al [4], Artzner et al. [3] and Jobert and Rogers [7].

We now explain the basic ingredients of multi-period cost-of-capital valuation. Let  $V_0$  be the value at time 0 of the liability cashflow  $X$ . This amount should be interpreted as the capital that needs to be transferred along with the liability to another external agent in order for the external agent (or capital provider) to accept managing the liability runoff and the associated capital costs. Let  $V_t$  denote the value at time  $t$  of the liability cashflows  $X_{>t}$ . The capital  $V_t$  is reserved at time  $t$  for managing the liability. However, regulation requires the capital  $R_t = \rho(-X_{t+1} - V_{t+1} | \mathcal{F}_t) > V_t$  to be set aside at time  $t$ . The difference  $R_t - V_t$  is made available by a capital provider requiring an excess expected return  $1 + \eta_t$  on the provided capital between time  $t$  and  $t + 1$ . The acceptability criterion under which the capital provider accepts to provide capital gives the equation

$$\mathbb{E}[R_t - X_{t+1} - V_{t+1} | \mathcal{F}_t] = (1 + \eta_t)(R_t - V_t). \quad (7)$$

Solving for  $V_t$  yields  $V_t = \varphi_t(X_{t+1} + V_{t+1})$ , where

$$\varphi_t(Y) = \frac{1}{1 + \eta_t} \mathbb{E}[Y | \mathcal{F}_t] + \frac{\eta_t}{1 + \eta_t} \rho(-Y | \mathcal{F}_t). \quad (8)$$

A commonly used conditional risk measure for variables with finite variance is  $\rho^{\text{SD}}(Y | \mathcal{F}_t) = \mathbb{E}[-Y | \mathcal{F}_t] + c^{\text{SD}} \text{var}(Y | \mathcal{F}_t)^{1/2}$ . Note that it can be argued that  $\rho^{\text{SD}}(Y | \mathcal{F}_t)$  is an inappropriate choice because it violates the monotonicity property:  $\tilde{Y} \geq Y$  does not imply  $\rho^{\text{SD}}(Y | \mathcal{F}_t) \geq \rho^{\text{SD}}(\tilde{Y} | \mathcal{F}_t)$ . For this choice of conditional risk measure, (8) takes

the form

$$\varphi_t(Y) = \mathbb{E}[Y \mid \mathcal{F}_t] + c_t \text{var}(Y \mid \mathcal{F}_t)^{1/2}, \quad (9)$$

where  $c_t = c^{\text{SD}} \eta_t / (1 - \eta_t)$ , and  $\varphi_t(Y) \in L^2(\mathcal{F}_t)$  if  $Y \in L^2(\mathcal{F}_T)$ .

It can be argued that the capital provider, seen as the share holder of the company, has limited liability and is not required to continue injecting capital if the the value  $R_t$  of available asset turns out insufficient to match the value  $X_{t+1} + V_{t+1}$  of the liability towards the policy holders. For discussions on limited liability in the context of cost-of-capital valuation we refer to Albrecher et al. [1] and Möhr [9]. In the setting with limited liability, (7) is replaced by

$$\mathbb{E}[(R_t - X_{t+1} - V_{t+1})^+ \mid \mathcal{F}_t] = (1 + \eta_t)(R_t - V_t) \quad (10)$$

and (8) is replaced by

$$\varphi_t(Y) = \rho(-Y \mid \mathcal{F}_t) - \frac{1}{1 + \eta_t} \mathbb{E}[(\rho(-Y \mid \mathcal{F}_t) - Y)^+ \mid \mathcal{F}_t]. \quad (11)$$

Regardless of whether we consider cost-of-capital valuation with or without limited liability, it follows that  $\varphi_t(aY + \tilde{Y}) = a\varphi_t(Y) + \tilde{Y}$  for  $a \in \mathbb{R}_+$  and  $\tilde{Y} \in L^1(\mathcal{F}_t)$  if the conditional risk measure satisfies  $\rho(aY + \tilde{Y} \mid \mathcal{F}_t) = a\rho(Y \mid \mathcal{F}_t) - \tilde{Y}$ . Moreover,  $\varphi_t$  inherits monotonicity from the conditional risk measure: if  $\tilde{Y} \geq Y$  implies  $\rho(Y \mid \mathcal{F}_t) \geq \rho(\tilde{Y} \mid \mathcal{F}_t)$ , then  $\varphi_t(\tilde{Y}) \geq \varphi_t(Y)$ .

An alternative to an acceptability criterion based on expected excess return and a cost-of-capital rate is an acceptability criterion saying that a risk averse capital provider provides capital if the payoff resulting from providing capital is preferred, in terms of expected utility, to simply rolling this capital forward by investing it in a riskless bond (or more generally, investing it in the numeraire asset). In this setting, the acceptability criterion (10) is replaced by

$$\mathbb{E}[u_t((R_t - X_{t+1} - V_{t+1})^+) \mid \mathcal{F}_t] = u_t(R_t - V_t), \quad (12)$$

where  $u_t$  is an increasing and concave (utility) function. Consequently, (11) is replaced by

$$\varphi_t(Y) = \rho(-Y \mid \mathcal{F}_t) - u_t^{-1} \left( \mathbb{E} \left[ u_t((\rho(-Y \mid \mathcal{F}_t) - Y)^+) \mid \mathcal{F}_t \right] \right). \quad (13)$$

Also in this case  $\varphi_t$  inherits monotonicity and the property  $\varphi_t(Y + \tilde{Y}) = \varphi_t(Y) + \tilde{Y}$ ,  $\tilde{Y} \in L^1(\mathcal{F}_t)$ , from a monotone and conditionally cash additive conditional risk measure. However, the property  $\varphi_t(aY + \tilde{Y}) = a\varphi_t(Y) + \tilde{Y}$ ,  $a \in \mathbb{R}_+$ , requires that  $u_t$  is chosen a power utility function  $u_t(x) = \alpha_t x^{\beta_t}$ , where  $\alpha_t > 0$  and  $\beta_t \in (0, 1]$ . In this case,  $\varphi_t$  in (13) takes the form

$$\varphi_t(Y) = \rho(-Y \mid \mathcal{F}_t) - \mathbb{E} \left[ ((\rho(-Y \mid \mathcal{F}_t) - Y)^+)^{\beta_t} \mid \mathcal{F}_t \right]^{1/\beta_t}. \quad (14)$$

The mappings  $\varphi_t : L^1(\mathcal{F}_T) \rightarrow L^1(\mathcal{F}_t)$  given by (8), (11) or (14) all satisfy  $\varphi_t(Y + \tilde{Y}) = \varphi_t(Y) + \tilde{Y}$  whenever  $\tilde{Y} \in L^1(\mathcal{F}_t)$ . Therefore, (3) can be expressed as, with  $\circ$  denoting composition,

$$V_T(X) = 0, \quad V_t(X) = \varphi_t \circ \dots \circ \varphi_{T-1} \left( \sum_{s=t+1}^T X_s \right), \quad t < T.$$

These properties also hold for the mappings  $\varphi_t$  in (9) with  $L^1$  replaced by  $L^2$ .



### 3.2 Main results

The main results consist of Theorems 1, 2 and 3. Theorems 1 and 2 presents conditions under which we have convergence for multi-period values (that are typically not computable) to a computable explicit limit. Theorem 3 presents a monotonicity result for the limit expressions in terms of a partial order between filtrations.

**Theorem 1.** *Let  $X, X^1, X^2, \dots$  be random vectors in  $\mathbb{R}^T$  and let  $Y, Y^1, Y^2, \dots$  be random vectors in  $(\mathbb{R}^d)^T$ ,  $d \in \mathbb{N}$ . Suppose that  $\mathcal{L}(X^n, Y^n) \xrightarrow{w} \mathcal{L}(X, Y)$  as  $n \rightarrow \infty$ , where  $\mathcal{L}(X, Y)$  is Gaussian, and that  $(\mathcal{L}(\|X^n\|))_{n \in \mathbb{N}}$  is uniformly integrable. Suppose also that, for each  $t$  and each convergent sequence  $(x^n, y^n)_{\leq t} \rightarrow (x, y)_{\leq t}$  with  $(x^n, y^n)_{\leq t} \in \text{supp}(\mathcal{L}((X^n, Y^n)_{\leq t}))$ ,*

$$\mathcal{L}\left((X^n, Y^n) \mid (X^n, Y^n)_{\leq t} = (x^n, y^n)_{\leq t}\right) \xrightarrow{w} \mathcal{L}\left((X, Y) \mid (X, Y)_{\leq t} = (x, y)_{\leq t}\right) \quad \text{as } n \rightarrow \infty, \quad (15)$$

$$\left(\mathcal{L}(\|X^n\| \mid (X^n, Y^n)_{\leq t} = (x^n, y^n)_{\leq t})\right)_{n \in \mathbb{N}} \quad \text{is uniformly integrable.} \quad (16)$$

For all  $t$ , let  $V_t(X)$  and  $V_t^n(X^n)$  be given by (3) and (4) with  $\varphi_t$  and  $\varphi_t^n$  given by either (8) and

$$\varphi_t^n(Y) = \frac{1}{1 + \eta_t} \mathbb{E}[Y \mid \mathcal{F}_t^n] + \frac{\eta_t}{1 + \eta_t} \rho(-Y \mid \mathcal{F}_t^n), \quad \eta_t \in \mathbb{R}_+, \quad (17)$$

or by (11) and

$$\varphi_t^n(Y) = \rho(-Y \mid \mathcal{F}_t^n) - \frac{1}{1 + \eta_t} \mathbb{E}\left[(\rho(-Y \mid \mathcal{F}_t^n) - Y)^+ \mid \mathcal{F}_t^n\right], \quad \eta_t \in \mathbb{R}_+, \quad (18)$$

or by (14) and

$$\varphi_t^n(Y) = \rho(-Y \mid \mathcal{F}_t^n) - \mathbb{E}\left[\left((\rho(-Y \mid \mathcal{F}_t^n) - Y)^+\right)^{\beta_t} \mid \mathcal{F}_t^n\right]^{1/\beta_t}, \quad \beta_t \in (0, 1], \quad (19)$$

where  $\mathcal{F}_t = \sigma((X, Y)_{\leq t})$  and  $\mathcal{F}_t^n = \sigma((X^n, Y^n)_{\leq t})$ , and where  $\rho(-Y \mid \mathcal{F}_t)$  and  $\rho(-Y \mid \mathcal{F}_t^n)$  are of the form (2). Then  $\lim_{n \rightarrow \infty} V_0^n(X^n) = V_0(X)$ , where

$$V_0(X) = \mathbb{E}\left[\sum_{t=1}^T X_t\right] + \sum_{t=1}^T \varphi_{t-1}(\varepsilon_t) \left( \text{Var}\left(\sum_{u=t}^T X_u \mid (X, Y)_{\leq t-1}\right) - \text{Var}\left(\sum_{u=t}^T X_u \mid (X, Y)_{\leq t}\right) \right)^{1/2}, \quad (20)$$

where  $\varepsilon_t$  is standard normally distributed and independent of  $\mathcal{F}_{t-1}$ .

**Remark 1.** *Notice that the conditional variances in (20) are nonrandom since the (joint) distribution of  $(X, Y)$  is Gaussian. Notice also  $\mathcal{F}_t = \sigma((X, Y)_{\leq t})$  generates the natural filtration of  $X$  if  $Y$  is chosen as nonrandom (a degenerate Gaussian process). We emphasize that if the  $\eta_t$  in (8) and (11) do not depend on  $t$ , and similarly if the  $\beta_t$  in (19) do not depend on  $t$ , then  $\varphi_{t-1}(\varepsilon_t) = \varphi_0(\varepsilon_1)$  in (20) does not depend on  $t$ .*

**Remark 2.** *Joint weak convergence  $\mathcal{L}(X^n, Y^n) \xrightarrow{w} \mathcal{L}(X, Y)$  as  $n \rightarrow \infty$  does not imply conditional weak convergence such as the convergence in (15). A counterexample as well as sufficient conditions for conditional weak convergence are presented in [13].*

**Remark 3.** For  $\varepsilon$  is standard normally distributed and independent of  $\mathcal{F}_{t-1}$ ,

$$\mathbb{V}\textcircled{\text{R}}_u(\varepsilon \mid \mathcal{F}_{t-1}) = \Phi^{-1}(1-u), \quad \text{AV}\textcircled{\text{R}}_u(\varepsilon \mid \mathcal{F}_{t-1}) = \frac{1}{u} \varphi \left( \Phi^{-1}(1-u) \right),$$

where  $\varphi$  and  $\Phi$  here denote the standard normal density and distribution function, respectively. Hence, with  $\rho_0(\varepsilon)$  denoting either  $\mathbb{V}\textcircled{\text{R}}_u(\varepsilon \mid \mathcal{F}_{t-1})$  or  $\text{AV}\textcircled{\text{R}}_u(\varepsilon \mid \mathcal{F}_{t-1})$  with values above that do not depend on  $t$ , we get

$$\varphi_{t-1}(\varepsilon) = \frac{\eta_{t-1}}{1 + \eta_{t-1}} \rho_0(\varepsilon)$$

in the case of (8), and

$$\varphi_{t-1}(\varepsilon) = \rho_0(\varepsilon) - \frac{1}{1 + \eta_{t-1}} \left( \rho_0(\varepsilon) \Phi(\rho_0(\varepsilon)) + \varphi(\rho_0(\varepsilon)) \right) \leq \frac{\eta_{t-1}}{1 + \eta_{t-1}} \rho_0(\varepsilon)$$

in the case of (11). In case of (14), the expectation  $\mathbb{E}[(\rho_0(\varepsilon) - \varepsilon)^+]^{\beta_{t-1}}$  has to be computed numerically in order to compute  $\varphi_{t-1}(\varepsilon)$ .

**Remark 4.** As stated in Theorem 1, the result holds when  $\varphi_t$  and  $\varphi_t^n$  are given by (8) and (17). However, the theorem is actually proven for mappings of a more general kind, namely, for  $\lambda_t \in [0, 1]$  and  $\mu_t^1, \mu_t^2 \in \mathcal{P}([0, 1])'$ , the mappings  $\varphi_t : L^1(\mathcal{F}_T) \rightarrow L^1(\mathcal{F}_t)$ ,  $\varphi_t^n : L^1(\mathcal{F}_T^n) \rightarrow L^1(\mathcal{F}_t^n)$  given by (21) and (22) below:

$$\varphi_t(Y) = \lambda_t \int F_{Y|\mathcal{F}_t}^{-1}(p) \mu_t^1(dp) + (1 - \lambda_t) \int F_{-Y|\mathcal{F}_t}^{-1}(p) \mu_t^2(dp), \quad (21)$$

$$\varphi_t^n(Y) = \lambda_t \int F_{Y|\mathcal{F}_t^n}^{-1}(p) \mu_t^1(dp) + (1 - \lambda_t) \int F_{-Y|\mathcal{F}_t^n}^{-1}(p) \mu_t^2(dp). \quad (22)$$

Note that (8) and (17) correspond to  $\mu_t^1(dp) = dp$  and  $\lambda_t = (1 + \eta_t)^{-1}$ .

**Remark 5.** Both (11), (18) and (14), (19) are special cases of

$$\varphi_t(Y) = \rho(-Y \mid \mathcal{F}_t) - \gamma_t \mathbb{E} \left[ ((\rho(-Y \mid \mathcal{F}_t) - Y)^+)^{\beta_t} \mid \mathcal{F}_t \right]^{1/\beta_t}, \quad (23)$$

$$\varphi_t^n(Y) = \rho(-Y \mid \mathcal{F}_t^n) - \gamma_t \mathbb{E} \left[ ((\rho(-Y \mid \mathcal{F}_t^n) - Y)^+)^{\beta_t} \mid \mathcal{F}_t^n \right]^{1/\beta_t}, \quad (24)$$

as seen by choosing  $(\gamma_t, \beta_t) = ((1 + \eta_t)^{-1}, 1)$  in case of (11), (18), and  $\gamma_t = 1$  in case of (14), (19).

Our next result shows convergence of multi-period values in the case of where the one period valuation mapping is defined as a sum of a conditional expectation and a constant times a conditional standard deviation. Let  $c_t \in \mathbb{R}_+$  and  $\varphi_t : L^2(\mathcal{F}_T) \rightarrow L^2(\mathcal{F}_t)$ ,  $\varphi_t^n : L^2(\mathcal{F}_T^n) \rightarrow L^2(\mathcal{F}_t^n)$  be given by

$$\varphi_t(Y) = \mathbb{E}[Y \mid \mathcal{F}_t] + c_t \text{var}[Y \mid \mathcal{F}_t]^{1/2}, \quad (25)$$

$$\varphi_t^n(Y) = \mathbb{E}[Y \mid \mathcal{F}_t^n] + c_t \text{var}[Y \mid \mathcal{F}_t^n]^{1/2}. \quad (26)$$

$(V_t(X))_{t=0}^T$  and  $(V_t^n(X^n))_{t=0}^T$  are defined backward recursively by (3) and (4), and (5) and (6) hold.

**Theorem 2.** Let  $X, X^1, X^2, \dots$  be random vectors in  $\mathbb{R}^T$  and let  $Y, Y^1, Y^2, \dots$  be random vectors in  $(\mathbb{R}^d)^T$ ,  $d \in \mathbb{N}$ . Suppose that  $\mathcal{L}(X^n, Y^n) \xrightarrow{w} \mathcal{L}(X, Y)$  as  $n \rightarrow \infty$ , where  $\mathcal{L}(X, Y)$  is Gaussian, and such that  $(\mathcal{L}(\|X^n\|^2))_{n \in \mathbb{N}}$  is uniformly integrable. Suppose also that, for each  $t$  and each convergent sequence  $(x^n, y^n)_{\leq t} \rightarrow (x, y)_{\leq t}$  with  $(x^n, y^n)_{\leq t} \in \text{supp}(\mathcal{L}((X^n, Y^n)_{\leq t}))$ ,

$$\mathcal{L}\left((X^n, Y^n) \mid (X^n, Y^n)_{\leq t} = (x^n, y^n)_{\leq t}\right) \xrightarrow{w} \mathcal{L}\left((X, Y) \mid (X, Y)_{\leq t} = (x, y)_{\leq t}\right) \quad \text{as } n \rightarrow \infty, \quad (27)$$

$$\left(\mathcal{L}(\|X^n\|^2 \mid (X^n, Y^n)_{\leq t} = (x^n, y^n)_{\leq t})\right)_{n \in \mathbb{N}} \quad \text{is uniformly integrable.} \quad (28)$$

Let  $V_t(X)$  and  $V_t^n(X^n)$  be given by (3) and (4) with  $\varphi_t$  and  $\varphi_t^n$  given by (25) and (26) with  $\mathcal{F}_t = \sigma((X, Y)_{\leq t})$  and  $\mathcal{F}_t^n = \sigma((X^n, Y^n)_{\leq t})$ . Then  $\lim_{n \rightarrow \infty} V_0^n(X^n) = V_0(X)$ , where

$$V_0(X) = \mathbb{E} \left[ \sum_{t=1}^T X_t \right] + \sum_{t=1}^T \varphi_{t-1}(\varepsilon_t) \left( \text{Var} \left( \sum_{u=t}^T X_u \mid (X, Y)_{\leq t-1} \right) - \text{Var} \left( \sum_{u=t}^T X_u \mid (X, Y)_{\leq t} \right) \right)^{1/2},$$

where  $\varepsilon_t$  is standard normally distributed and independent of  $\mathcal{F}_{t-1}$ .

The final theorem considers the Gaussian limit models appearing in Theorems 1 and 2 and establishes a partial ordering of values  $V_0(X)$  assigned to a given cashflow  $X$  depending on a partial order between filtrations to which  $X$  is adapted. One may guess that a larger filtration allows for more accurate predictions over time of future cashflows and thereby reduces the value  $V_0(X)$ . We show that this holds if the parameters of the one-step valuation mappings  $\varphi_t$  do not depend on  $t$  and if the conditional variance of the aggregate cashflow  $\sum_{s=1}^T X_s$  decays in a convex manner over time as more information becomes available.

We consider filtrations  $\mathbb{F} = (\mathcal{F}_t)_{t=0}^T$  with  $\mathcal{F}_0 = \{\Omega, \emptyset\}$  and  $\mathcal{F}_t = \sigma((X, Y)_{\leq t})$ , for  $t \geq 1$ , where  $(X, Y)$  is jointly Gaussian. We assume that the one-step valuation mappings  $\varphi_t$  do not depend on  $t$  such that  $\varphi_{t-1}(\varepsilon_t) = \varphi_0(\varepsilon_1)$  in (20). We write

$$V_0(X, \mathbb{F}) = \mathbb{E} \left[ \sum_{t=1}^T X_t \right] + \varphi_0(\varepsilon_1) \sum_{t=1}^T \left( \text{Var} \left( \sum_{u=t}^T X_u \mid \mathcal{F}_{t-1} \right) - \text{Var} \left( \sum_{u=t}^T X_u \mid \mathcal{F}_t \right) \right)^{1/2}.$$

The following theorem says that for filtrations  $\mathbb{F}$  and  $\mathbb{G}$  of the above kind (generated by a process that is jointly Gaussian together with the cashflow process  $X$ ), if  $t \mapsto \text{Var}(\sum_{s=1}^T X_s \mid \mathcal{F}_t)$  is convex, then  $V_0(X, \mathbb{F}) \geq V_0(X, \mathbb{G})$  whenever  $\mathcal{F}_t \subseteq \mathcal{G}_t$  for every  $t$ . The convexity assumption is necessary.

**Theorem 3.** Let  $(X, Y)_{t=1}^T$  and  $(X, Z)_{t=1}^T$  be Gaussian processes. Let  $\mathbb{F} = (\mathcal{F}_t)_{t=0}^T$  and  $\mathbb{G} = (\mathcal{G}_t)_{t=0}^T$  be filtrations with  $\mathcal{F}_0 = \mathcal{G}_0 = \{\Omega, \emptyset\}$  and  $\mathcal{F}_t = \sigma((X, Y)_{\leq t})$  and  $\mathcal{G}_t = \sigma((X, Z)_{\leq t})$ , for  $t \geq 1$ . If  $\mathcal{F}_t \subseteq \mathcal{G}_t$  for every  $t$  and if  $t \mapsto \text{Var}(\sum_{s=1}^T X_s \mid \mathcal{F}_t)$  is convex, then  $V_0(X, \mathbb{F}) \geq V_0(X, \mathbb{G})$ .

**Remark 6.** We emphasize that the Gaussian assumption in Theorem 3 means that the terms

$$c_t = \text{Var} \left( \sum_{s=1}^T X_s \mid \mathcal{F}_{t-1} \right) - \text{Var} \left( \sum_{s=1}^T X_s \mid \mathcal{F}_t \right)$$

form a nonrandom sequence  $(c_t)_{t=1}^T$  and that the convexity property is equivalent to  $c_1 \geq c_2 \geq \dots \geq c_T$ .

## 4 Proofs and auxiliary results

We use the notation  $Z_{\leq t}^n = (X^n, Y^n)_{\leq t}$ ,  $Z_{\leq t} = (X, Y)_{\leq t}$  and  $z_{\leq t}^n = (x^n, y^n)_{\leq t}$ ,  $z_{\leq t} = (x, y)_{\leq t}$ .

*Proof of Theorem 1 assuming (21) and (22).* As explained in Remark 4, (8) and (17) are special cases of (21) and (22).

For  $t = 1, \dots, T$ , consider mappings  $\psi_t^n : \text{supp}(\mathcal{L}(Z_{\leq t}^n)) \rightarrow \mathbb{R}$  given by  $\psi_T^n(z_{\leq T}^n) = \sum_{s=1}^T x_s^n$  and

$$\psi_t^n(z_{\leq t}^n) = \lambda_t \int F_{\psi_{t+1}^n(Z_{\leq t+1}^n) | Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \mu_t^1(dp) + (1 - \lambda_t) \int F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) | Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \mu_t^2(dp), \quad (29)$$

and let

$$\psi_0^n = \lambda_0 \int F_{\psi_1^n(Z_1^n)}^{-1}(p) \mu_0^1(dp) + (1 - \lambda_0) \int F_{-\psi_1^n(Z_1^n)}^{-1}(p) \mu_0^2(dp), \quad (30)$$

where  $\mu_t^1, \mu_t^2 \in \mathcal{P}([0, 1])'$ . We define  $\psi_t : \mathbb{R}^t \rightarrow \mathbb{R}$  and  $\psi_0$  analogously without the superscript  $n$ . Note that

$$\begin{aligned} \psi_t^n(Z_{\leq t}^n) &= \sum_{s=1}^t X_s^n + V_t^n(X^n) = \varphi_t^n(\psi_{t+1}^n(Z_{\leq t+1}^n)), \quad t \geq 1, \quad \psi_0^n = V_0^n(X) = \varphi_0^n(\psi_1^n(Z_{\leq 1}^n)), \\ \psi_t(Z_{\leq t}) &= \sum_{s=1}^t X_s + V_t(X) = \varphi_t(\psi_{t+1}(Z_{\leq t+1})), \quad t \geq 1, \quad \psi_0 = V_0(X) = \varphi_0(\psi_1(Z_{\leq 1})). \end{aligned}$$

Therefore, the proof is complete once we show the convergence  $\lim_{n \rightarrow \infty} \psi_0^n = \psi_0$ .

The argument of the proof is backwards induction. We will show that for  $t = T, \dots, 1$ ,

**UI(t):**  $(\mathcal{L}(\psi_t^n(Z_{\leq t}^n) | Z_{\leq s}^n = z_{\leq s}^n))_n$  is uniformly integrable for each  $s = 1, \dots, t-1$  and  $z_{\leq s}^n \rightarrow z_{\leq s}$ ,

**CC(t):**  $(\psi_t^n)_n$  is continuously convergent, i.e.  $\psi_t^n(z_{\leq t}^n) \rightarrow \psi_t(z_{\leq t})$ , whenever  $z_{\leq t}^n \rightarrow z_{\leq t}$ .

**Induction base:** **UI(T)** and **CC(T)** hold.

For  $s \in \{1, \dots, T-1\}$  and  $z_{\leq s}^n \rightarrow z_{\leq s}$ , **UI(T)** follows from (16) since norms on Euclidean spaces are equivalent and  $|\psi_T^n(z_{\leq T}^n)| = |\sum_{s=1}^T x_s^n| \leq \sum_{s=1}^T |x_s^n|$ . Moreover,

$$\psi_T^n(z_{\leq T}^n) = \sum_{s=1}^T x_s^n \rightarrow \sum_{s=1}^T x_s = \psi_T(z_{\leq T})$$

whenever  $z_{\leq T}^n \rightarrow z_{\leq T}$ , yielding **CC(T)**.

**Induction step:** **UI(t+1)**, **CC(t+1)** together imply **UI(t)**, **CC(t)**.

Fix  $s \in \{1, \dots, t-1\}$  and sequence  $(z_{\leq s}^n)_n$  with  $z_{\leq s}^n \rightarrow z_{\leq s}$ . Let  $\kappa_{s+1, t+1}^n(z_{\leq s}^n, \cdot)$  and  $\kappa_{s+1, t+1}(z_{\leq s}, \cdot)$  be regular versions of the conditional distributions

$$P((Z_{s+1}^n, \dots, Z_{t+1}^n)^\top \in \cdot | Z_{\leq s}^n = z_{\leq s}^n) \quad \text{and} \quad P((Z_{s+1}, \dots, Z_{t+1})^\top \in \cdot | Z_{\leq s} = z_{\leq s}).$$

By assumption (15),  $\kappa_{s+1, t+1}^n(z_{\leq s}^n, \cdot) \xrightarrow{w} \kappa_{s+1, t+1}(z_{\leq s}, \cdot)$ . Since  $(\psi_{t+1}^n(z_{\leq s}^n, \cdot))$  is continuously convergent by the induction assumption **CC(t+1)**, by the generalized continuous mapping theorem ([8, Theorem 4.27, Exercise 27]),

$$\kappa_{s+1, t+1}^n(z_{\leq s}^n, \cdot) \circ (\psi_{t+1}^n(z_{\leq s}^n, \cdot))^{-1} \xrightarrow{w} \kappa_{s+1, t+1}(z_{\leq s}, \cdot) \circ (\psi_{t+1}(z_{\leq s}, \cdot))^{-1},$$

i.e.

$$\mathcal{L}(\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq s}^n = z_{\leq s}^n) \xrightarrow{w} \mathcal{L}(\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq s} = z_{\leq s}). \quad (31)$$

Lemma 3 shows that the limiting distribution in (31) is Gaussian. By Lemma 2, there exists a constant  $c > 0$ , such that

$$|\psi_t^n(z)| \leq c \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z].$$

Hence, it is sufficient to show uniform integrability of

$$\left( \mathcal{L}(\mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n] \mid Z_{\leq s}^n = z_{\leq s}^n) \right)_n. \quad (32)$$

Setting  $\widetilde{W}^n = \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n]$  and  $W^n = |\psi_{t+1}^n(Z_{\leq t+1}^n)|$ , this follows from Lemma 4, using **UI**( $t+1$ ) and **CC**( $t+1$ ). Hence, **UI**( $t$ ) holds.

We proceed by showing **CC**( $t$ ), i.e. the convergence of

$$\psi_t^n(z_{\leq t}^n) = \lambda_t \int F_{\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \mu_t^1(dp) \quad (33)$$

$$+ (1 - \lambda_t) \int F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \mu_t^2(dp). \quad (34)$$

We will prove the convergence assuming that  $\mu_t^1$  has a bounded density and that  $\text{supp}(\mu_t^2) \subset [a, b] \subset (0, 1)$ . Other possibilities for  $\mu_t^1, \mu_t^2 \in \mathcal{P}([0, 1])'$  are handled by the same arguments as those shown below.

We start by the integral in (33) and will show convergence by an application of Pratt's Lemma [11]. Let  $w : (0, 1) \rightarrow \mathbb{R}$ , bounded by some  $c' > 0$ , be the density of  $\mu_t^1$ , i.e.  $\mu_t^1(dp) = w(p)dp$ . Let

$$f_n(p) = F_{\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) w(p).$$

From (31) follows convergence

$$f_n(p) \rightarrow F_{\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(p) w(p)$$

for almost every  $p \in (0, 1)$ . By combining (31) and the continuous mapping theorem applied to the absolute value function follows convergence of upper and lower bounds  $l_n(p) \leq f_n(p) \leq u_n(p)$  for almost every  $p \in (0, 1)$ :

$$\begin{aligned} l_n(p) &= c' F_{-|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \rightarrow c' F_{-|\psi_{t+1}(Z_{\leq t+1})| \mid Z_{\leq t} = z_{\leq t}}^{-1}(p) = l(p) \\ u_n(p) &= c' F_{|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \rightarrow c' F_{|\psi_{t+1}(Z_{\leq t+1})| \mid Z_{\leq t} = z_{\leq t}}^{-1}(p) = u(p). \end{aligned}$$

The induction assumption **UI**( $t+1$ ) together with

$$\mathcal{L}(|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z_{\leq t}^n) \xrightarrow{w} \mathcal{L}(|\psi_{t+1}(Z_{\leq t+1})| \mid Z_{\leq t} = z_{\leq t}),$$

allow us to conclude from [8, Theorem 4.11] that

$$\begin{aligned} \int l_n(p) dp &= -c' \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z_{\leq t}^n] \\ &\rightarrow -c' \mathbb{E}[|\psi_{t+1}(Z_{\leq t+1})| \mid Z_{\leq t} = z_{\leq t}] = \int l(p) dp \end{aligned}$$

and

$$\begin{aligned}\int u_n(p)dp &= c' \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z_{\leq t}^n] \\ &\rightarrow c' \mathbb{E}[|\psi_{t+1}(Z_{\leq t+1})| \mid Z_{\leq t} = z_{\leq t}] = \int u(p)dp.\end{aligned}$$

Hence, by Pratt's Lemma [11, Theorem 1] we have convergence of the integral in (33), i.e.

$$\begin{aligned}\int F_{\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p)\mu_1(dp) &= \int f_n(p)dp \rightarrow \int f(p)dp \\ &= \int F_{\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(p)\mu_1(dp).\end{aligned}$$

We now consider the integral in (34). By Lemma 3, the mapping  $p \mapsto F_{-\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(p)$  is continuous for all  $p \in (0, 1)$ . Therefore, from (31) follows pointwise convergence

$$F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \rightarrow F_{-\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(p)$$

for all  $p \in (0, 1)$ . In particular, the convergence holds at the points  $a, b \in (0, 1)$ . Therefore, using the fact that  $p \mapsto F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p)$  and  $p \mapsto F_{-\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(p)$  are increasing, we can find  $d > 0$ , such that

$$c'' = \max \left( \left| F_{-\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(a) \right|, \left| F_{-\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(b) \right| \right) + d$$

uniformly bounds the integrand, i.e.

$$\left| F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \right| \leq c''$$

for all  $p \in [a, b]$  and for all  $n \in \mathbb{N}$ . Hence, the bounded convergence theorem yields

$$\int F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p)\mu_t^2(dp) \rightarrow \int F_{-\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}}^{-1}(p)\mu_t^2(dp).$$

Hence, we have shown  $\mathbf{CC}(t)$  and the proof of the induction step is complete.

It remains to show convergence of  $\psi_0^n$ . By the continuous convergence of  $(\psi_1^n)$  and the generalized continuous mapping theorem ([8, Theorem 4.27]), it follows that  $\mathcal{L}(\psi_1^n(Z_1^n)) \xrightarrow{w} \mathcal{L}(\psi_1(Z_1))$ , where the limit is Gaussian (cf. Lemma 3). Therefore the convergence of

$$\psi_0^n = \lambda_0 \int F_{\psi_1^n(Z_1^n)}^{-1}(p)\mu_0^1(dp) + (1 - \lambda_0) \int F_{-\psi_1^n(Z_1^n)}^{-1}(p)\mu_0^2(dp)$$

follows from arguments completely analogous to those verifying the induction step. Lemma 3 together with the variance decomposition

$$\begin{aligned}\text{Var} \left( \mathbb{E} \left[ \sum_{v=u}^T X_v \mid Z_{\leq u} \right] \mid Z_{\leq u-1} \right) &= \text{Var} \left( \sum_{v=u}^T X_v \mid Z_{\leq u-1} \right) \\ &\quad - \mathbb{E} \left[ \text{Var} \left( \sum_{v=u}^T X_v \mid Z_{\leq u} \right) \mid Z_{\leq u-1} \right] \\ &= \text{Var} \left( \sum_{v=u}^T X_v \mid Z_{\leq u-1} \right) - \text{Var} \left( \sum_{v=u}^T X_v \mid Z_{\leq u} \right)\end{aligned}$$

completes the proof.  $\square$

*Proof of Theorem 1 assuming (23) and (24).* The proofs is similar to the above proof of Theorem 1 assuming (21) and (22).

Fix  $\mu_0, \dots, \mu_{T-1} \in \mathcal{P}([0, 1])'$ . Let  $\nu_T^n(z_{\leq T}^n) = \sum_{s=1}^T x_s^n$  and  $\psi_T^n(z_{\leq T}^n) = \sum_{s=1}^T x_s^n$ . For  $t = 1, \dots, T-1$ , define mappings  $\nu_t^n, \psi_t^n : \text{supp}(\mathcal{L}(Z_{\leq t}^n)) \rightarrow \mathbb{R}$  by

$$\nu_t^n(z_{\leq t}^n) = \int F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) | Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \mu_t(dp), \quad (35)$$

$$\psi_t^n(z_{\leq t}^n) = \nu_t^n(z_{\leq t}^n) - \gamma_t \mathbb{E} \left[ \left( (\nu_t^n(z_{\leq t}^n) - \psi_{t+1}^n(Z_{\leq t+1}^n))^+ \right)^{\beta_t} \mid Z_{\leq t}^n = z_{\leq t}^n \right]^{1/\beta_t}. \quad (36)$$

Let

$$\nu_0^n = \int F_{-\psi_1^n(Z_1^n)}^{-1}(p) \mu_0(dp), \quad \psi_0^n = \nu_0^n - \gamma_0 \mathbb{E} \left[ \left( (\nu_0^n - \psi_1^n(Z_1^n))^+ \right)^{\beta_0} \right]^{1/\beta_0}. \quad (37)$$

Define  $\nu_t, \psi_t : \mathbb{R}^t \rightarrow \mathbb{R}$  and  $\nu_0, \psi_0$  analogously without the superscript  $n$ . The proof is complete once we show the convergence  $\lim_{n \rightarrow \infty} \psi_0^n = \psi_0$ .

The argument of the proof is backwards induction. We will show that for  $t = T, \dots, 1$ ,

**UI(t):**  $(\mathcal{L}(\psi_t^n(Z_{\leq t}^n) \mid Z_{\leq s}^n = z_{\leq s}^n))_n$  is uniformly integrable for each  $s = 1, \dots, t-1$  and  $z_{\leq s}^n \rightarrow z_{\leq s}$ ,

**CC(t):**  $(\psi_t^n)_n$  is continuously convergent, i.e.  $\psi_t^n(z_{\leq t}^n) \rightarrow \psi_t(z_{\leq t})$ , whenever  $z_{\leq t}^n \rightarrow z_{\leq t}$ .

**Induction base:** **UI(T)** and **CC(T)** hold. The argument verifying the induction base is identical to that in the proofs of Theorem 1 assuming (21) and (22).

**Induction step:** **UI(t+1)**, **CC(t+1)** together imply **UI(t)**, **CC(t)**. The argument verifying that

$$\mathcal{L}(\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq s}^n = z_{\leq s}^n) \xrightarrow{w} \mathcal{L}(\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq s} = z_{\leq s}). \quad (38)$$

holds is identical to that in the proofs of Theorem 1 assuming (21) and (22). By Lemma 5, there exists a constant  $c > 0$ , such that

$$|\psi_t^n(z)| \leq c \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z].$$

Hence, to verify **UI(t)** it is sufficient to show uniform integrability of

$$\left( \mathcal{L} \left( \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n] \mid Z_{\leq s}^n = z_{\leq s}^n \right) \right)_{n \in \mathbb{N}}$$

and after setting  $W^n = |\psi_{t+1}^n(Z_{\leq t+1}^n)|$ ,  $\widetilde{W}^n = \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n]$ , **UI(t)** follows from **UI(t+1)** and (38) by applying Lemma 4.

We proceed by showing **CC(t)**, i.e. the convergence of (36). We first consider (35). If  $\mu_t$  admits a bounded density, convergence  $\rho_t^n(z_{\leq t}^n) \rightarrow \rho_t(z_{\leq t})$  can be shown by Pratt's Lemma, together with (38) and **UI(t+1)** (cf. Proof of Theorem 1 assuming (21) and (22)). If  $\mu_t$  satisfies  $\text{supp}(\mu_t) \subset [a, b]$  with  $0 < a < b < 1$ , we first note that we have pointwise convergence  $F_{-\psi_{t+1}^n(Z_{\leq t}^n) | Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p) \rightarrow F_{-\psi_{t+1}(Z_{\leq t}) | Z_{\leq t} = z_{\leq t}}^{-1}(p)$  for all  $p \in (0, 1)$  since the limiting distribution in (38) is Gaussian (cf. Lemma 3). There exists  $c > 0$  such that

$$\sup_n \sup_{p \in [a, b]} |F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) | Z_{\leq t}^n = z_{\leq t}^n}^{-1}(p)| \leq \sup_n c \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z_{\leq t}^n]$$

and the right-hand side is finite. The inequality is due to properties of quantile functions (cf. Proof of Lemma 2). The finiteness is because of the convergence of the conditional expectation which is due to (38) and  $\mathbf{UI}(t+1)$ . Hence, dominated convergence yields  $\nu_t^n(z_{\leq t}^n) \rightarrow \nu_t(z_{\leq t})$ .

We now show convergence of the conditional expectation in (36) which can be written in terms of

$$\mathbb{E} \left[ h_n(\psi_{t+1}^n(Z_{\leq t+1}^n)) \mid Z_{\leq t}^n = z_{\leq t}^n \right],$$

where  $h_n(w) = ((\nu_t^n(z_{\leq t}^n) - w)^+)^{\beta_t}$  converges continuously towards  $h(w) = ((\nu_t(z_{\leq t}) - w)^+)^{\beta_t}$ . Moreover, since

$$|h_n(w)| \leq \left( \sup_n (\nu_t^n(z_{\leq t}^n)) + |w| \right)^{\beta_t}, \quad \sup_n (\nu_t^n(z_{\leq t}^n)) < \infty,$$

and  $\beta_t \in (0, 1]$ , uniform integrability of  $\mathcal{L}(h_n(\psi_{t+1}^n(Z_{\leq t+1}^n)) \mid Z_{\leq t}^n = z_{\leq t}^n)$  follows from  $\mathbf{UI}(t+1)$ . This means convergence

$$\mathbb{E} \left[ h_n(\psi_{t+1}^n(Z_{\leq t+1}^n)) \mid Z_{\leq t}^n = z_{\leq t}^n \right] \rightarrow \mathbb{E} \left[ h(\psi_{t+1}(Z_{\leq t+1})) \mid Z_{\leq t} = z_{\leq t} \right]$$

which implies  $\mathbf{CC}(t)$  and the proof of the induction step is complete.

Finally, the proof of the convergence  $\lim_{n \rightarrow \infty} \psi_0^n = \psi_0$  follows by arguments analogous to those in the proof of Theorem 1 assuming (21) and (22).  $\square$

**Lemma 1.** *Let  $W$  be a random variable with quantile function  $F^{-1}(p) = \min\{x \in \mathbb{R} : F_W(x) \geq p\}$ . If  $\mu \in \mathcal{P}([0, 1])'$ , then there exists  $c \in (0, \infty)$  such that*

$$\left| \int F_W^{-1}(p) \mu(dp) \right| \leq c \mathbb{E}[|W|].$$

*Proof.* We have the general bounds  $l, u$  given by

$$l = \int F_{-|W|}^{-1}(p) \mu(dp) \leq \int F_W^{-1}(p) \mu(dp) \leq \int F_{|W|}^{-1}(p) \mu(dp) = u.$$

Moreover,

$$F_{-|W|}^{-1}(p) = -F_{|W|}^{-1}((1-p)^+) = - \lim_{v \downarrow 1-p} F_{|W|}^{-1}(v).$$

For any  $p \in (0, 1)$ ,

$$\begin{aligned} p F_{|W|}^{-1}(1-p) &\leq \int_{1-p}^1 F_{|W|}^{-1}(v) dv \leq \mathbb{E}[|W|], \\ p F_{-|W|}^{-1}(p) &\geq \int_0^p F_{-|W|}^{-1}(v) dv \geq -\mathbb{E}[|W|]. \end{aligned}$$

Consequently, if  $\text{supp}(\mu) \subset [a, b] \subset (0, 1)$ , then

$$l \geq -\frac{1}{a} \mathbb{E}[|W|], \quad u \leq \frac{1}{1-b} \mathbb{E}[|W|], \quad \left| \int F_W^{-1}(p) \mu(dp) \right| \leq \max\left(\frac{1}{a}, \frac{1}{1-b}\right) \mathbb{E}[|W|].$$

If  $\mu$  has a density bounded by  $c' \in (0, \infty)$ , then

$$l \geq -c' \mathbb{E}[|W|], \quad u \leq c' \mathbb{E}[|W|], \quad \left| \int F_W^{-1}(p) \mu(dp) \right| \leq c' \mathbb{E}[|W|].$$

$\square$



**Lemma 2.** *If  $\psi_t^n$  is given by (29) and (30), then exists  $c \in (0, \infty)$  such that*

$$|\psi_0^n| \leq c \mathbb{E}[|\psi_1^n(Z_1^n)|] \quad \text{and} \quad |\psi_t^n(z)| \leq c \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z]$$

for any  $z \in \text{supp}(\mathcal{L}(Z_{\leq t}^n))$ .

*Proof.* We proof the inequality for  $t \geq 1$ . The argument for  $t = 0$  is identical. Since

$$|\psi_t^n(z)| \leq \left| \int F_{\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z}^{-1}(p) \mu_t^1(dp) \right| + \left| \int F_{-\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z}^{-1}(p) \mu_t^2(dp) \right|,$$

we can apply Lemma 1, bounding the right-hand side by  $c \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t}^n)| \mid Z_{\leq t}^n = z]$  for some  $c$ .  $\square$

The following result is similar to Proposition 6 in [5].

**Lemma 3.** *Let  $(Z_t)_{t=1}^T = (X_t, Y_t)_{t=1}^T$  be a Gaussian process, let  $\mathcal{F}_0 = \{\Omega, \emptyset\}$ , and let  $\mathcal{F}_t = \sigma(Z_{\leq t})$  for  $t \geq 1$ . Let  $V_T = 0$  and, for  $t \in \{0, 1, \dots, T-1\}$ , let  $V_t = \varphi_t(X_{t+1} + V_{t+1})$ , where  $\varphi_t$  is positive homogeneous and conditionally cash additive. Suppose that for each  $t$ ,*

$$\text{if } \mathbb{P}(W \in \cdot \mid \mathcal{F}_t) = \mathbb{P}(\widetilde{W} \in \cdot \mid \mathcal{F}_t), \quad \text{then } \varphi_t(W) = \varphi_t(\widetilde{W}),$$

and if  $W$  is independent of  $\mathcal{F}_t$ , then  $\varphi_t(W)$  is nonrandom. Then

$$V_t = \mathbb{E} \left[ \sum_{u=t+1}^T X_u \mid \mathcal{F}_t \right] + \sum_{u=t+1}^T \varphi_{u-1}(\varepsilon_u) \text{Var} \left( \mathbb{E} \left[ \sum_{v=u}^T X_v \mid \mathcal{F}_u \right] \mid \mathcal{F}_{u-1} \right)^{1/2},$$

where  $\varepsilon_u$  is standard normal and independent of  $\mathcal{F}_{u-1}$ .

*Proof.* The lemma is proved by backward induction. We prove the induction step. Suppose that  $X_{t+1} + V_{t+1}$  and  $Z_{\leq t}$  are jointly Gaussian. Then there exists a standard normal  $\varepsilon_{t+1}$  independent of  $\mathcal{F}_t$  such that

$$\mathbb{P}(X_{t+1} + V_{t+1} \in \cdot \mid \mathcal{F}_t) = \mathbb{P} \left( \mathbb{E}[X_{t+1} + V_{t+1} \mid \mathcal{F}_t] + \varepsilon_{t+1} \text{Var}(X_{t+1} + V_{t+1} \mid \mathcal{F}_t)^{1/2} \in \cdot \mid \mathcal{F}_t \right)$$

and it follows from properties of the multivariate normal distribution that the conditional variance  $\text{Var}(X_{t+1} + V_{t+1} \mid \mathcal{F}_t)$  is nonrandom. Hence, by positive homogeneity and conditional cash additivity,

$$\varphi_t(X_{t+1} + V_{t+1}) = \mathbb{E}[X_{t+1} + V_{t+1} \mid \mathcal{F}_t] + \varphi_t(\varepsilon_{t+1}) \text{Var}(X_{t+1} + V_{t+1} \mid \mathcal{F}_t)^{1/2}.$$

By the induction assumption (in the nontrivial case  $t+1 < T$ ),

$$V_{t+1} = \mathbb{E} \left[ \sum_{u=t+2}^T X_u \mid \mathcal{F}_{t+1} \right] + \sum_{u=t+2}^T \varphi_{u-1}(\varepsilon_u) \text{Var} \left( \mathbb{E} \left[ \sum_{v=u}^T X_v \mid \mathcal{F}_u \right] \mid \mathcal{F}_{u-1} \right)^{1/2}.$$

Summing up the terms and using the tower property of conditional expectations complete the proof.  $\square$

**Lemma 4.** *Let  $s \in \{1, \dots, t-1\}$  and let  $(W^n)_{n \in \mathbb{N}}$  be a sequence of non-negative random variables such that*

$$\mathcal{L}(W^n \mid Z_{\leq s}^n = z_{\leq s}^n) \xrightarrow{w} \mathcal{L}(W \mid Z_{\leq s} = z_{\leq s}) \quad \text{as } n \rightarrow \infty \quad (39)$$

and

$$\left( \mathcal{L}(W^n \mid Z_{\leq s}^n = z_{\leq s}^n) \right)_{n \in \mathbb{N}} \quad \text{is uniformly integrable.} \quad (40)$$

If  $\widetilde{W}^n = \mathbb{E}[W^n \mid Z_{\leq t}^n]$ , then  $(\mathcal{L}(\widetilde{W}^n \mid Z_{\leq s}^n = z_{\leq s}^n))_{n \in \mathbb{N}}$  is uniformly integrable.

*Proof.* First note that

$$\begin{aligned} \mathbb{E}[\widetilde{W}^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n] &= \mathbb{E}[\mathbb{E}[W^n \mid Z_{\leq t}^n] I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n] \\ &= \mathbb{E}[\mathbb{E}[W^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq t}^n] \mid Z_{\leq s}^n] \\ &= \mathbb{E}[W^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n], \end{aligned}$$

i.e.  $\mathbb{E}[\widetilde{W}^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n = z_{\leq s}^n] = \mathbb{E}[W^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n = z_{\leq s}^n]$ . For each  $\gamma_r > 0$ ,

$$\begin{aligned} &\mathbb{E}[W^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n = z_{\leq s}^n] \\ &= \mathbb{E}[W^n I\{\widetilde{W}^n \geq r\} I\{W^n \leq \gamma_r\} \mid Z_{\leq s}^n = z_{\leq s}^n] \\ &\quad + \mathbb{E}[W^n I\{\widetilde{W}^n \geq r\} I\{W^n > \gamma_r\} \mid Z_{\leq s}^n = z_{\leq s}^n] \\ &\leq \gamma_r \mathbb{P}(\widetilde{W}^n \geq r \mid Z_{\leq s}^n = z_{\leq s}^n) + \mathbb{E}[W^n I\{W^n > \gamma_r\} \mid Z_{\leq s}^n = z_{\leq s}^n]. \end{aligned} \quad (41)$$

By (39) and (40),

$$\mathbb{E}[\widetilde{W}^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n = z_{\leq s}^n] < \infty \quad \text{for all } n \in \mathbb{N},$$

and similarly with  $W^n$  instead of  $\widetilde{W}^n$ . Therefore, uniform integrability of  $(\mathcal{L}(\widetilde{W}^n \mid Z_{\leq s}^n = z_{\leq s}^n))_{n \in \mathbb{N}}$  is equivalent to (cf. [8] p. 67)

$$\lim_{r \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{E}[\widetilde{W}^n I\{\widetilde{W}^n \geq r\} \mid Z_{\leq s}^n = z_{\leq s}^n] = 0$$

which follows if we show that there exists  $\gamma_r$  such that  $\lim_{r \rightarrow \infty} \limsup_{n \rightarrow \infty} [(41)] = 0$ . By the uniform integrability assumption (40) and the finiteness of all the terms,

$$\lim_{r \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{E}[W^n I\{W^n > \gamma_r\} \mid Z_{\leq s}^n = z_{\leq s}^n] = 0.$$

Therefore, it only remains to show that

$$\lim_{r \rightarrow \infty} \limsup_{n \rightarrow \infty} \gamma_r \mathbb{P}(\widetilde{W}^n \geq r \mid Z_{\leq s}^n = z_{\leq s}^n) = 0$$

for some sequence  $\gamma_r \rightarrow \infty$ . By Markov's inequality,

$$\begin{aligned} \mathbb{P}(\widetilde{W}^n \geq r \mid Z_{\leq s}^n = z_{\leq s}^n) &\leq \frac{\mathbb{E}[\widetilde{W}^n \mid Z_{\leq s}^n = z_{\leq s}^n]}{r} \\ &= \frac{\mathbb{E}[W^n \mid Z_{\leq s}^n = z_{\leq s}^n]}{r} \\ &\rightarrow \frac{\mathbb{E}[W \mid Z_{\leq s} = z_{\leq s}]}{r}, \end{aligned}$$

where the convergence in the last step follows from Lemma 4.11 in [8], using (39) and (40). Letting e.g.  $\gamma_r = \sqrt{r}$ , the first term in (41) can therefore be made arbitrarily small by choosing  $r$  sufficiently large.  $\square$

**Lemma 5.** *If  $\nu_t^n, \psi_t^n$  are given by (35), (36) and (37), then exists  $c \in (0, \infty)$  such that*

$$|\psi_0^n| \leq c \mathbb{E}[|\psi_1^n(Z_1^n)|] \quad \text{and} \quad |\psi_t^n(z)| \leq c \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z]$$

for any  $z \in \text{supp}(\mathcal{L}(Z_{\leq t}^n))$ .

*Proof of Lemma 5.* Consider  $t \geq 1$ . Note that, by Lemma 1, there exists  $c' > 0$ , such that

$$|\nu_t^n(z)| \leq c' \mathbf{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z].$$

Hence,

$$\begin{aligned} |\psi_t^n(z)| &\leq c' \mathbf{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z] \\ &\quad + \gamma_t \mathbf{E} \left[ \left( (c' \mathbf{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z] - \psi_{t+1}^n(Z_{\leq t+1}^n))^+ \right)^{\beta_t} \mid Z_{\leq t}^n = z \right]^{1/\beta_t}. \end{aligned}$$

By Jensen's inequality, the second summand above is bounded by

$$\begin{aligned} &\gamma_t \mathbf{E} \left[ (c' \mathbf{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z] - \psi_{t+1}^n(Z_{\leq t+1}^n))^+ \mid Z_{\leq t}^n = z \right] \\ &\leq \gamma_t \mathbf{E} \left[ c' \mathbf{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z] + |\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z \right] \\ &= (c' + 1) \gamma_t \mathbf{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z]. \end{aligned}$$

Hence, the statement holds with  $c = c' + \gamma_t(c' + 1)$ . The inequality  $|\psi_0^n| \leq c \mathbf{E}[|\psi_1^n(Z_1^n)|]$  is shown by the same arguments.  $\square$

*Proof of Theorem 2.* The proof proceeds in a similar manner as the proof of Theorem 1. For  $t = 1, \dots, T$ , consider mappings  $\psi_t^n : \text{supp}(\mathcal{L}(Z_{\leq t}^n)) \rightarrow \mathbb{R}$  given by

$$\psi_t^n(z_{\leq t}^n) = \begin{cases} \sum_{s=1}^T x_s^n, & t = T, \\ \mathbf{E}[\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n] + c_t \text{Var}(\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n)^{1/2}, & t < T, \end{cases}$$

and

$$\psi_0^n = \mathbf{E}[\psi_1^n(Z_1^n)] + c_0 \text{Var}(\psi_1^n(Z_1^n))^{1/2}.$$

We define  $\psi_t : \mathbb{R}^t \rightarrow \mathbb{R}$  and  $\psi_0$  analogously without the superscript  $n$ . The proof is complete once we show the convergence  $\lim_{n \rightarrow \infty} \psi_0^n = \psi_0$ .

The argument of the proof is backwards induction. We will show that for  $t = T, \dots, 1$ ,

**UI(t):**  $(\mathcal{L}(\psi_t^n(Z_{\leq t}^n)^2 \mid Z_{\leq s}^n = z_{\leq s}^n))_n$  is uniformly integrable for each  $s = 1, \dots, t-1$  and  $z_{\leq s}^n \rightarrow z_{\leq s}$ ,

**CC(t):**  $(\psi_t^n)_n$  is continuously convergent, i.e.  $\psi_t^n(z_{\leq t}^n) \rightarrow \psi_t(z_{\leq t})$ , whenever  $z_{\leq t}^n \rightarrow z_{\leq t}$ .

**Induction base:** **UI(T)** and **CC(T)** hold.

For  $s \in \{1, \dots, T-1\}$  and  $z_{\leq s}^n \rightarrow z_{\leq s}$ , **UI(T)** follows immediately from (28) since  $\psi_T^n(z_{\leq T}^n) = \sum_{s=1}^T x_s^n$ . Moreover,

$$\psi_T^n(z_{\leq T}^n) = \sum_{s=1}^T x_s^n \rightarrow \sum_{s=1}^T x_s = \psi_T(z_{\leq T})$$

whenever  $z_{\leq T}^n \rightarrow z_{\leq T}$ , yielding **CC(T)**.

**Induction step:** **UI(t+1)**, **CC(t+1)** together imply **UI(t)**, **CC(t)**. Fix  $s \in \{1, \dots, t-1\}$  and a sequence  $(z_{\leq s}^n)_n$  with  $z_{\leq s}^n \rightarrow z_{\leq s}$ . By the same argument as in the proof of Theorem 1,

$$\mathcal{L}(\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq s}^n = z_{\leq s}^n) \xrightarrow{w} \mathcal{L}(\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq s} = z_{\leq s}) \quad (42)$$

which implies

$$\mathcal{L}((\psi_{t+1}^n(Z_{\leq t+1}^n))^2 \mid Z_{\leq s}^n = z_{\leq s}^n) \xrightarrow{w} \mathcal{L}((\psi_{t+1}(Z_{\leq t+1}))^2 \mid Z_{\leq s} = z_{\leq s}). \quad (43)$$

Lemma 3 shows that the limiting distribution in (42) is Gaussian. We can bound

$$\begin{aligned} \psi_t^n(z)^2 &\leq 2 \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)| \mid Z_{\leq t}^n = z]^2 + 2c_t^2 \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)|^2 \mid Z_{\leq t}^n = z] \\ &\leq (2 + 2c_t^2) \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)|^2 \mid Z_{\leq t}^n = z]. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbb{E}[|\psi_t^n(Z_{\leq t}^n)|^2 \mid Z_{\leq s}^n = z_{\leq s}^n] &\leq (2 + 2c_t^2) \mathbb{E}[\mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)|^2 \mid Z_{\leq t}^n] \mid Z_{\leq s}^n = z_{\leq s}^n] \\ &= (2 + 2c_t^2) \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)|^2 \mid Z_{\leq s}^n = z_{\leq s}^n]. \end{aligned}$$

Hence, it is sufficient to show uniform integrability of

$$\left( \mathcal{L}(\mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)|^2 \mid Z_{\leq t}^n] \mid Z_{\leq s}^n = z_{\leq s}^n) \right)_n. \quad (44)$$

Setting  $\widetilde{W}^n = \mathbb{E}[|\psi_{t+1}^n(Z_{\leq t+1}^n)|^2 \mid Z_{\leq t}^n]$  and  $W^n = |\psi_{t+1}^n(Z_{\leq t+1}^n)|^2$ , this follows from Lemma 4 together with (43) and **UI**( $t+1$ ). Hence, **UI**( $t$ ) holds.

We proceed by showing **CC**( $t$ ). Let  $z_{\leq t}^n \rightarrow z_{\leq t}$ . We need to show convergence of

$$\begin{aligned} \psi_t^n(z_{\leq t}^n) &= \mathbb{E}[\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n] \\ &\quad + c_t \left( \mathbb{E}[(\psi_{t+1}^n(Z_{\leq t+1}^n))^2 \mid Z_{\leq t}^n = z_{\leq t}^n] - \mathbb{E}[\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n]^2 \right)^{1/2}. \end{aligned}$$

Indeed, by (42) and (43),

$$\begin{aligned} \mathcal{L}(\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n) &\xrightarrow{w} \mathcal{L}(\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}), \\ \mathcal{L}((\psi_{t+1}^n(Z_{\leq t+1}^n))^2 \mid Z_{\leq t}^n = z_{\leq t}^n) &\xrightarrow{w} \mathcal{L}((\psi_{t+1}(Z_{\leq t+1}))^2 \mid Z_{\leq t} = z_{\leq t}), \end{aligned}$$

and by [8, Lemma 4.11] applied to **UI**( $t+1$ ), it follows that

$$\mathbb{E}[\psi_{t+1}^n(Z_{\leq t+1}^n) \mid Z_{\leq t}^n = z_{\leq t}^n] \rightarrow \mathbb{E}[\psi_{t+1}(Z_{\leq t+1}) \mid Z_{\leq t} = z_{\leq t}]$$

and

$$\mathbb{E}[(\psi_{t+1}^n(Z_{\leq t+1}^n))^2 \mid Z_{\leq t}^n = z_{\leq t}^n] \rightarrow \mathbb{E}[(\psi_{t+1}(Z_{\leq t+1}))^2 \mid Z_{\leq t} = z_{\leq t}],$$

yielding **CC**( $t$ ) which completes the proof of the induction step.

As a result of the induction,

$$\mathcal{L}(\psi_1^n(Z_1^n)) \xrightarrow{w} \mathcal{L}(\psi_1(Z_1))$$

and  $(\psi_1^n(Z_1^n))_n, ((\psi_1^n(Z_1^n))^2)_n$  are uniformly integrable. Therefore,

$$\begin{aligned} \psi_0^n &= \mathbb{E}[\psi_1^n(Z_1^n)] + c_0 \left( \mathbb{E}[(\psi_1^n(Z_1^n))^2] - \mathbb{E}[\psi_1^n(Z_1^n)]^2 \right)^{1/2} \\ &\rightarrow \mathbb{E}[\psi_1(Z_1)] + c_0 \left( \mathbb{E}[(\psi_1(Z_1))^2] - \mathbb{E}[\psi_1(Z_1)]^2 \right)^{1/2} = \psi_0. \end{aligned}$$

□

**Lemma 6.** Let  $c \in \mathbb{R}_+^T$  satisfy  $\sum_{t=1}^T c_t = 1$  and  $c_{t+1} \leq c_t$  for every  $t$ . Then  $c$  solves

$$\max_{d \in \mathcal{D}_{c,T}} \sum_{t=1}^T d_t^{1/2}, \quad \mathcal{D}_{c,T} = \left\{ d \in \mathbb{R}_+^T : \sum_{t=1}^T d_t = 1, \sum_{s=1}^t d_s \geq \sum_{s=1}^t c_s \text{ for } t = 1, \dots, T \right\}. \quad (45)$$

*Proof.* Since  $\mathcal{D}_{c,T}$  is a convex set and the objective function is concave, we have a convex optimization problem. Consequently, it is sufficient to verify that  $c \in \mathcal{D}_{c,T}$  is a local optimum. By assumption either  $c_1 = 1$  and  $c_t = 0$  for  $t \geq 2$  or there exists  $t_0 \in \{2, \dots, T\}$  such that  $c_1 \geq \dots \geq c_{t_0} > 0$  and  $c_{t_0+1} = \dots = c_T = 0$ . The first case has the trivial maximizer  $d_1 = 1$  and  $d_t = 0$  for  $t \geq 2$ . Hence, it is sufficient to consider only the second case for a fixed  $c$  and  $t_0 \in \{2, \dots, T\}$ . In this case, (45) is equivalent to, with  $c = (c_1, \dots, c_{t_0})$ ,

$$\max_{d \in \mathcal{D}_c} \sum_{t=1}^{t_0} d_t^{1/2}, \quad \mathcal{D}_c = \left\{ d \in \mathbb{R}_+^{t_0} : \sum_{t=1}^{t_0} d_t = 1, \sum_{s=1}^t d_s \geq \sum_{s=1}^t c_s \text{ for } t = 1, \dots, t_0 \right\}. \quad (46)$$

Since  $\mathcal{D}_c$  is a convex set and the objective function is concave, we have a convex optimization problem. Consequently, it is sufficient to verify that  $c \in \mathcal{D}_c$  is a local optimum. Let

$$\mathcal{X} = \left\{ x \in \mathbb{R}^{t_0} : \sum_{t=1}^{t_0} x_t = 0, \sum_{s=1}^t x_s \geq 0 \text{ for } t = 1, \dots, t_0 \right\}$$

and note that  $\mathcal{D}_c \subset c + \mathcal{X}$ . Consider the set of vectors  $\{b_1, \dots, b_{t_0-1}\} \subset \mathbb{R}^{t_0}$  given by

$$b_{i,i} = 1, b_{i,i+1} = -1, b_{i,j} = 0 \text{ for } j \notin \{i, i+1\}.$$

We claim that  $\mathcal{X} = \text{span}_+\{b_1, \dots, b_{t_0-1}\}$ , where

$$\text{span}_+\{b_1, \dots, b_{t_0-1}\} = \left\{ \sum_{k=1}^{t_0-1} \lambda_k b_k : \lambda_1, \dots, \lambda_{t_0-1} \geq 0 \right\}.$$

To show that  $\text{span}_+\{b_1, \dots, b_{t_0-1}\} \subseteq \mathcal{X}$  it is sufficient to note that

$$\begin{aligned} \sum_{s=1}^{t_0} \sum_{k=1}^{t_0-1} \lambda_k b_{k,s} &= \sum_{k=1}^{t_0-1} \lambda_k \sum_{s=1}^{t_0} b_{k,s} = 0, \\ \sum_{s=1}^t \sum_{k=1}^{t_0-1} \lambda_k b_{k,s} &= \sum_{k=1}^{t_0-1} \lambda_k \sum_{s=1}^t b_{k,s} = \lambda_t \geq 0. \end{aligned}$$

We now show that  $\mathcal{X} \subseteq \text{span}_+\{b_1, \dots, b_{t_0-1}\}$ . Take  $x \in \mathcal{X}$  and set  $\lambda_1 = x_1$  (noting that  $x_1 \geq 0$ ) and  $\lambda_k = \lambda_{k-1} + x_k$  for  $k \geq 2$ . Hence,  $\lambda_k \geq 0$  for every  $k$  and  $\sum_{k=1}^{t_0-1} \lambda_k b_k = x$ .

Let  $g(d) = \sum_{t=1}^{t_0} d_t^{1/2}$  and note that  $g$  is well defined and concave on  $\mathcal{D}_c$ , and has a well defined gradient at  $c$ :

$$\nabla g(c)^\top = \frac{1}{2}(c_1^{-1/2}, \dots, c_{t_0}^{-1/2}).$$

Take  $d \in \mathcal{D}_c$  and notice that  $d = c + x$  for  $x \in \mathcal{X}$ . Since  $g$  is concave,

$$g(d) - g(c) \leq \nabla g(c)^\top x = \sum_{k=1}^{t_0-1} \lambda_k \nabla g(c)^\top b_k = \frac{1}{2} \sum_{k=1}^{t_0-1} \lambda_k (c_k^{-1/2} - c_{k+1}^{-1/2}) \leq 0,$$

where the last inequality holds since, for every  $k$ ,  $\lambda_k \geq 0$  and  $c_k^{-1/2} - c_{k+1}^{-1/2} \leq 0$ . Hence, we have shown that  $g(d) \leq g(c)$  and the proof is complete.  $\square$

*Proof of Theorem 3.* For  $t = 1, \dots, T$ , set

$$c_t = \text{Var} \left( \sum_{s=1}^T X_s \right)^{-1} \left( \text{Var} \left( \sum_{s=1}^T X_s \mid \mathcal{F}_{t-1} \right) - \text{Var} \left( \sum_{s=1}^T X_s \mid \mathcal{F}_t \right) \right),$$

$$d_t = \text{Var} \left( \sum_{s=1}^T X_s \right)^{-1} \left( \text{Var} \left( \sum_{s=1}^T X_s \mid \mathcal{G}_{t-1} \right) - \text{Var} \left( \sum_{s=1}^T X_s \mid \mathcal{G}_t \right) \right).$$

By construction,  $\sum_{t=1}^T c_t = \sum_{t=1}^T d_t = 1$ . Since  $\mathcal{F}_t \subseteq \mathcal{G}_t$  for every  $t$ ,  $\sum_{s=t+1}^T c_s \geq \sum_{s=t+1}^T d_s$  which is equivalent to  $\sum_{s=1}^t d_s \geq \sum_{s=1}^t c_s$  for every  $t$  since  $\sum_{t=1}^T c_t = \sum_{t=1}^T d_t$ . Applying Lemma 6 completes the proof.  $\square$

## Acknowledgements

F. Lindskog would like to acknowledge financial support from the Swedish Research Council, Project 2020-05065, and from Länsförsäkringars Forskningsfond, Project P9.20.

## References

- [1] Albrecher, H.-J., Eisele, K.-T., Steffensen, M., Wüthrich, M. V. (2022). On the cost-of-capital rate under incomplete market valuation, *Journal of Risk and Insurance* 89, 1139–1158.
- [2] Aldous, D.J. (1981). Weak convergence and general theory of processes, Unpublished draft of monograph; Department of Statistics, University of California, Berkeley, CA 94720.
- [3] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., Ku, H. (2007). Coherent multiperiod risk adjusted values and Bellman’s principle, *Annals of Operations Research* 152, 5–22.
- [4] Cheridito, P., Delbaen, F. and Kupper, M. (2006). Dynamic monetary risk measures for bounded discrete-time processes, *Electronic Journal of Probability* 11, 57–106.
- [5] Engsner, H., Lindholm, M., Lindskog, F. (2017). Insurance valuation: a computable multi-period cost-of-capital approach, *Insurance: Mathematics and Economics* 72, 250–264.
- [6] Föllmer, H., Schied, A. *Stochastic Finance: An Introduction in Discrete Time*. De Gruyter, Berlin, Boston, 2016.
- [7] Jobert, A. and Rogers, L.C.G. (2008). Valuations and dynamic convex risk measures, *Mathematical Finance* 18(1), 1–22.
- [8] Kallenberg, O. *Foundations of Modern Probability*, 2nd edn. Springer, 2002.
- [9] Möhr, C. (2011). Market-consistent valuation of insurance liabilities by cost of capital, *ASTIN Bulletin* 41(2), 315–341.
- [10] Pelsser, A. and Salahnejhad Ghalehjooghi, A. (2016). Time-consistent actuarial valuations, *Insurance: Mathematics and Economics* 66, 97–112.

- [11] Pratt, J.W. (1960). On interchanging limits and integrals, *Annals of Mathematical Statistics* 31(1), 74–77.
- [12] Salzmann, R. and Wüthrich, M.V. (2010). Cost-of-Capital margin for general insurance liability runoff, *ASTIN Bulletin* 40(2), 415–451.
- [13] Sweeting, T.J. (1989). On conditional weak convergence, *Journal of Theoretical Probability* 2, 461–474.