Modern Stochastics: Theory and Applications 9 (3) (2022) 265–277
https://doi.org/10.15559/22-VMSTA204

On the denseness of the subset of discrete distributions in a certain set of two-dimensional distributions

Dmitriy Borzykh[a,*], **Alexander Gushchin**[a,b]

[a]*National Research University Higher School of Economics, Pokrovsky Boulevard 11, 109028 Moscow, Russia*
[b]*Steklov Mathematical Institute, Gubkina 8, 119991 Moscow, Russia*

borzykh.dmitriy@gmail.com (D. Borzykh), gushchin@mi-ras.ru (A. Gushchin)

Received: 1 December 2021, Accepted: 16 March 2022, Published online: 25 March 2022

Abstract In the article [Theory of Probability & Its Applications 62(2) (2018), 216–235], a class $\mathbb{W}$ of terminal joint distributions of integrable increasing processes and their compensators was introduced. In this paper, it is shown that the discrete distributions lying in $\mathbb{W}$ form a dense subset in the set $\mathbb{W}$ for ψ-weak topology with a gauge function ψ of linear growth.

Keywords Increasing process, compensator, terminal joint distribution, Doob–Meyer decomposition, dense set of distributions
2010 MSC 60G44, 60G07

1 Introduction

A class $\mathbb{W}$ of probability measures defined on $(\mathbb{R}^2_+, \mathcal{B}(\mathbb{R}^2_+))$ is introduced in [4]. It consists of all probability measures satisfying

1) $\int_{\mathbb{R}^2_+} (x + y)\, \mu(dx, dy) < \infty$,

2) $\int_{\mathbb{R}^2_+} x\, \mu(dx, dy) = \int_{\mathbb{R}^2_+} y\, \mu(dx, dy)$,

3) $\forall c \geq 0\ \int_{\{y \leq c\}} x\, \mu(dx, dy) \leq \int_{\mathbb{R}^2_+} (y \wedge c)\, \mu(dx, dy)$.

*Corresponding author.

It is also proved in [4] that a measure μ belongs to $\mathbb{W}$ if and only if there is an integrable increasing process $(X_t)_{t\in\mathbb{R}_+}$ with a compensator $(A_t)_{t\in\mathbb{R}_+}$ such that $\operatorname{Law}(X_\infty, A_\infty) = \mu$. The main idea behind the existence of such process X is as follows, see [4, Proposition 3.6, Remark 3.4, and Proposition 3.4]. Let a measure $\mu \in \mathbb{W}$ be given. Then one can find nonnegative random variables V, W, Z on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\operatorname{Law}(V+Z,\ W+Z) = \mu,$$

and $\mathbb{E}[V\mathbb{1}\{W \le t\}] = \mathbb{E}[W \wedge t]$ for all $t \in \mathbb{R}_+$. This means that $\operatorname{Law}(V, W) \in \mathbb{W}$ with equality in 3). Now it is easy to introduce a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ (one can take the single jump filtration introduced in [5]) such that $W \wedge t$, $t \in \mathbb{R}_+$, is the compensator of $V\mathbb{1}\{W \le t\}$. This is a special case of Theorem 1 below though the direct proof is much easier [4, Proposition 3.4]. To complete the construction one needs to incorporate Z into these processes, e.g.,

$$X_t = \begin{cases} V\mathbb{1}\{W \le t/(1-t)\}, & \text{if } t < 1,\\ V + (t\wedge 2 - 1)Z, & \text{if } t \geqslant 1,\end{cases}$$

$$A_t = \begin{cases} \frac{t}{1-t} \wedge W, & \text{if } t < 1,\\ W + (t\wedge 2 - 1)Z, & \text{if } t \geqslant 1.\end{cases}$$

The procedure of constructing a triple V, W and Z for a measure $\mu \in \mathbb{W}$ is available from the proof of Proposition 3.6 [4]. For measures $\mu \in \mathbb{W}$ with finite support, calculations are explicit. For example, let $\mu = \frac{1}{2}\delta_{(1,\,3/2)} + \frac{1}{2}\delta_{(3/2,\,1)}$. Then $\mu \in \mathbb{W}$. Put $\Omega := \{0, 1\} \times [0; 1]$, $\mathcal{F} := \mathcal{B}(\Omega)$,

$$\mathbb{P}(\{0\} \times B) = \frac{1}{2}\Lambda(B), \quad \mathbb{P}(\{1\} \times B) = \frac{1}{2}\Lambda(B),$$

where $B \in \mathcal{B}([0; 1])$ and Λ is the Lebesgue measure on $[0; 1]$. Then one can take

$$\begin{aligned} V(0,u) &:= \frac{1}{2-u}, & V(1,u) &:= 0,\\ W(0,u) &:= \frac{u}{4-2u}, & W(1,u) &:= 1/2,\\ Z(0,u) &:= 1 - \frac{u}{4-2u}, & Z(1,u) &:= 1.\end{aligned}$$

Our aim here is to show that measures with finite support and lying in $\mathbb{W}$ are dense in $\mathbb{W}$ in some natural topology of the space of Borel measures $\mu(dx, dy)$ on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ with integrable marginal distributions.

We will make a few remarks that are intended to demonstrate that our task is nontrivial. If $\mathbb{E}[V\mathbb{1}\{W \le t\}] = \mathbb{E}[W \wedge t]$ for all $t \ge 0$, then it is easy to see that the distribution of W cannot be discrete. Note that in the example above W has a continuous component, though the distribution μ has a finite support. This means that the subset of $\mathbb{W}$ consisting of measures with equality in 3) does not contain measures with finite support. In particular, if $V = 1$, then W has the exponential distribution with mean 1. Example 2.1 in [4] says that if we consider a two-point mass ν that is obtained from the exponential distribution by averaging over the sets $[0;\, a]$ and $(a;\, +\infty)$, where $a > 0$, then the measure $\delta_{\{1\}} \otimes \nu$ does not belong to $\mathbb{W}$.

The problem considered here is an important technical tool in studying some stochastic orders on the plane. In particular, let $\mu_0(dx, dy) \preccurlyeq \mu_1(dx, dy)$ if there are stochastic processes $X = (X_t)_{t\in[0;1]}$ and $Y = (Y_t)_{t\in[0;1]}$ such that X is adapted and Y is predictable on some stochastic basis, X_t and Y_t are integrable for all t, $X_t - X_0$ is an increasing process with the compensator $Y_t - Y_0$, and $\mathrm{Law}(X_0, Y_0) = \mu_0$ and $\mathrm{Law}(X_1, Y_1) = \mu_1$. In these terms $\mu \in \mathbb{W}$ is equivalent to $\delta_{(0,0)} \preccurlyeq \mu$. Our main result allows us to find necessary and sufficient conditions for $\mu_0 \preccurlyeq \mu_1$ expressed in appropriate terms [1].

Let us also mention a connection between our problem and the Skorokhod embedding problem. Namely, let B be a standard Brownian motion and T a finite stopping time on some stochastic basis. Put $\overline{B}_T := \sup_{t\le T} B_t$. Then, if $\mathbb{E}|B_T| < \infty$ and $\mathbb{E}[\overline{B}_T] < \infty$, $\mathrm{Law}(\overline{B}_T - B_T, \overline{B}_T) \in \mathbb{W}$. However, not every distribution μ from $\mathbb{W}$ can be represented in this way. The necessary and sufficient condition on μ is that $\mathbb{R}_+ \ni t \mapsto \int_{\mathbb{R}^2_+} [(y \wedge t) - x\mathbb{1}_{\{y\le t\}}]\, \mu(dx,\, dy)$ be an increasing function and $\mu((0;\infty) \times \{0\}) = 0$, see [8] and [6]. A discretisation of the Skorokhod embedding problem is considered, in particular, in [2].

To state the problem we need a proper topological setting. Consider a *gauge function* $\psi(x, y) := 1 + |x| + |y|$, $x,\ y \in \mathbb{R}$, and define the class $C_\psi(\mathbb{R}^2)$ of continuous test functions $f : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\forall f \in C_\psi(\mathbb{R}^2)\ \ \exists c \in \mathbb{R}\ \ \forall x, y \in \mathbb{R} \quad |f(x, y)| \le c \cdot \psi(x, y).$$

Denote also by $\mathcal{M}_1^\psi(\mathbb{R}^2)$ the set of all Borel measures on $\mathbb{R}^2$ such that

$$\int_{\mathbb{R}^2} \psi(x, y)\, \mu(dx,\, dy) < \infty.$$

The coarsest topology on $\mathcal{M}_1^\psi(\mathbb{R}^2)$ for which all mappings

$$\mathcal{M}_1^\psi(\mathbb{R}^2) \ni \mu \mapsto \int_{\mathbb{R}^2} f(x, y)\, \mu(dx,\, dy), \quad f \in C_\psi(\mathbb{R}^2),$$

are continuous, is called the *ψ-weak topology* on $\mathcal{M}_1^\psi(\mathbb{R}^2)$.

We refer to [3, §A.6] for the definition and further properties of ψ-weak topologies with arbitrary gauge functions ψ. Here we only mention that in our case the ψ-weak topology is metrizable and separable. Moreover, $\mathcal{M}_1^\psi(\mathbb{R}^2)$ is a Polish space with this topology. A sequence $(\mu_n)_{n=1}^\infty \subseteq \mathcal{M}_1^\psi(\mathbb{R}^2)$ converges to a measure $\mu \in \mathcal{M}_1^\psi(\mathbb{R}^2)$ in the ψ-weak topology if and only if $(\mu_n)_{n=1}^\infty$ converges to μ in the usual weak topology and the uniform integrability condition

$$\sup_{n\in\mathbb{N}} \int_{\{|x|+|y|>c\}} (|x| + |y|)\, \mu_n(dx,\, dy) \to 0 \quad \text{as} \ \ c \to \infty$$

is satisfied.

The choice of the ψ-weak topology with the prescribed function ψ is natural in this problem. On the one hand, the space $\mathcal{M}_1^{\psi}(\mathbb{R}^2)$ contains the set $\mathbb{W}$. On the other hand, it is this topology that insures the closedness of the set $\mathbb{W}$ in $\mathcal{M}_1^{\psi}(\mathbb{R}^2)$, while the usual weak topology does not provide the closedness.

The proof of our main result is based on Theorem 5 in [5]. For the reader's convenience, we present this theorem here with some refinement in the statement. We do not explain here how a single jump filtration is defined because the only fact concerning these filtrations that we use is that such a filtration exists.

Theorem 1. *Let V and γ be random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}$ and $[0; +\infty]$, respectively, such that $\mathbb{E}(|V|\mathbb{1}_{\{\gamma \le t\}}) < \infty$ for any $t \in \mathcal{T} := \big\{t \in \mathbb{R}_+ : \mathbb{P}\{\gamma \ge t\} > 0\big\}$. Define, for $t \in \mathcal{T}$,*

$$K(t) := \mathbb{E}[V|\gamma = t], \quad \text{and} \quad F(t) := \int_{(0;t]} \overline{G}(s-)^{-1} K(s)\, dG(s),$$

where $G(s) := \mathbb{P}\{\gamma \le s\}$ and $\overline{G}(s) := \mathbb{P}\{\gamma > s\}$. Then the process $F(t \wedge \gamma)$, $t \in \mathbb{R}_+$, is the compensator of the process $V\mathbb{1}_{\{\gamma \le t\}}$ with respect to the single jump filtration generated by γ and $\mathcal{F}$.

The only difference between this theorem and Theorem 5 in [5] refers to the case where γ is bounded by a constant $t_G < \infty$ with probability one and $\mathbb{P}\{\gamma = t_G\} > 0$. In this case, in Theorem 5 in [5] F is defined by left-continuity at t_G, so the difference between the definitions is

$$\int_{\{t_G\}} \overline{G}(s-)^{-1} K(s)\, dG(s) = K(t_G).$$

Now it is easy to see that the formulae for the compensators from both theorems coincide. Thus, our definition of F at t_G allows us to combine two terms in the formula for the compensator in Theorem 5 in [5] and to obtain a unified expression.

2 Main result

Let us consider two subsets of the set $\mathbb{W}$, namely, the subset of *simple measures* $\mathbb{W}_{\text{simp}}$ and the subset of *discrete measures* $\mathbb{W}_{\text{disc}}$. We say that $\mu \in \mathbb{W}_{\text{simp}}$ (correspondingly, $\mu \in \mathbb{W}_{\text{disc}}$) if $\mu \in \mathbb{W}$ and

$$\mu(dx, dy) = \sum_{j \in J} p_j \cdot \delta_{[x_j a_j]}(dx, dy),$$

where J is a finite set (correspondingly, J is at most countable), $p_j \ge 0$, $\sum_{j\in J} p_j = 1$, and $\delta_{[x_j a_j]}(dx, dy)$ is the Dirac measure at point $\left[\begin{smallmatrix} x_j \\ a_j \end{smallmatrix}\right] \in \mathbb{R}^2$. It is clear that $\mathbb{W}_{\text{simp}} \subseteq \mathbb{W}_{\text{disc}} \subseteq \mathbb{W}$.

The following theorem is the main result of this paper.

Theorem 2. *The set $\mathbb{W}_{\text{simp}}$ is dense in the set $\mathbb{W}$ in the ψ-weak topology. In other words, for any probability measure $\mu \in \mathbb{W}$, there exists a sequence of simple measures $(\nu_n)_{n=1}^{\infty} \subseteq \mathbb{W}_{\text{simp}}$ converging to μ in the ψ-weak topology, i.e., for any test* *function* $f \in C_\psi(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} f d\nu_n \to \int_{\mathbb{R}^2} f d\mu \quad as\ n \to \infty.$$

The proof will be performed in two steps:

Step 1: Given $\mu \in \mathbb{W}$, construct a sequence $(\mu_n)_{n=1}^\infty \subseteq \mathbb{W}_{\text{disc}}$ that converges to μ in the ψ-weak topology.

Step 2: Using the sequence $(\mu_n)_{n=1}^\infty$ from Step 1, construct a sequence $(\nu_n)_{n=1}^\infty \subseteq \mathbb{W}_{\text{simp}}$ converging to μ in the same topology. Notice that most of the arguments of the proof are the same as at Step 1.

Proof. Fix a measure $\mu \in \mathbb{W}$. Due to [4, Proposition 3.6, Remark 3.4, and Proposition 3.4], one can find a stochastic basis $\mathbb{B} := (\Omega,\ \mathcal{F},\ \mathbb{P},\ (\mathcal{G}_t)_{t\in\mathbb{R}_+})$ and nonnegative integrable random variables V, W, Z such that

$$\text{Law}(V+Z,\ W+Z) = \mu, \tag{1}$$

and the process $V\mathbb{1}\{W \le t\}$, $t \in \mathbb{R}_+$, is an integrable increasing process with the compensator $W \wedge t$, $t \in \mathbb{R}_+$. Therefore, $\mathbb{E}[V\mathbb{1}\{W \le t\}] = \mathbb{E}[W \wedge t]$ for all $t \in \mathbb{R}_+$, i.e., condition 3) from the definition of $\mathbb{W}$ holds as equality. In particular, $\text{Law}(V,\ W) \in \mathbb{W}$.

Step 1. It is enough to construct three sequences of nonnegative random variables $(\hat{V}^{(n)})_{n=1}^\infty$, $(\hat{W}^{(n)})_{n=1}^\infty$, and $(\hat{Z}^{(n)})_{n=1}^\infty$ on the probability space $(\Omega,\ \mathcal{F},\ \mathbb{P})$ which satisfy the following conditions:

(i) $\text{Law}(\hat{V}^{(n)},\ \hat{W}^{(n)}) \in \mathbb{W}_{\text{disc}}$ for all $n \in \mathbb{N}$;

(ii) $\text{Law}(\hat{V}^{(n)} + \hat{Z}^{(n)},\ \hat{W}^{(n)} + \hat{Z}^{(n)}) \in \mathbb{W}_{\text{disc}}$ for all $n \in \mathbb{N}$;

(iii) $\hat{V}^{(n)} \to V$ and $\hat{Z}^{(n)} \to Z$ as $n \to \infty$ $\mathbb{P}$-a. s.;

(iv) $(\hat{V}^{(n)})_{n=1}^\infty$ and $(\hat{Z}^{(n)})_{n=1}^\infty$ are uniformly integrable;

(v) $\hat{W}^{(n)} \to W$ as $n \to \infty$ $\mathbb{P}$-a. s.;

(vi) $(\hat{W}^{(n)})_{n=1}^\infty$ is uniformly integrable.

Indeed, let $f \in C_\psi(\mathbb{R}^2)$, then $|f(x,y)| \le c\psi(x,y)$ for some $c \ge 0$ and

$$\left|f(\hat{V}^{(n)} + \hat{Z}^{(n)},\ \hat{W}^{(n)} + \hat{Z}^{(n)})\right| \le c\Big(1 + \left|\hat{V}^{(n)} + \hat{Z}^{(n)}\right| + \left|\hat{W}^{(n)} + \hat{Z}^{(n)}\right|\Big). \tag{2}$$

Now, the expression in the LHS of (2) under the sign of modulus is uniformly integrable due to (iv) and (vi) and converges $\mathbb{P}$-a. s. to $f(V+Z,\ W+Z)$ due to (iii), (v) and continuity of f. Taking expectations, we arrive at the claim of Step 1 in view of (ii) and (1).

In what follows we will use a result from [5]. Let us introduce some notation from that paper. Denote

- $G(t) := \mathbb{P}\{W \le t\}$, $t \in \mathbb{R}$;

- $\overline{G}(t) := \mathbb{P}\{W > t\} = 1 - G(t)$;
- $t_G := \sup\{t \in \mathbb{R}_+ \colon G(t) < 1\}$;
- $\mathcal{T} := \big\{t \in \mathbb{R}_+ \colon \mathbb{P}\{W \geq t\} > 0\big\}$.

Introduce the following cases:

- **Case A1:** $t_G < \infty$ and $\mathbb{P}\{W = t_G\} = 0$;
- **Case A2:** $t_G = \infty$;
- **Case B:** $\mathbb{P}\{W = t_G < \infty\} > 0$.

Case **A1** + **A2** and Case **B** correspond to Case **A** and Case **B** respectively in [5].

It is easy to see that $\mathbb{P}\{W \notin \mathcal{T}\} = 0$ and $\mathcal{T} = [0;\ t_G)$ in Cases **A1** and **A2**, and $\mathcal{T} = [0;\ t_G]$ in Case **B**.

Let us define auxiliary random variables

$$V^{(n)} := \sum_{i=1}^{\infty} \tfrac{i\Delta}{2^n} \mathbb{1}\Big\{\tfrac{(i-1)\Delta}{2^n} < V \leq \tfrac{i\Delta}{2^n}\Big\},$$
$$W^{(n)} := \sum_{j=1}^{\infty} \tfrac{j\Delta}{2^n} \mathbb{1}\Big\{\tfrac{(j-1)\Delta}{2^n} < W \leq \tfrac{j\Delta}{2^n}\Big\},$$
$$Z^{(n)} := \sum_{k=1}^{\infty} \tfrac{k\Delta}{2^n} \mathbb{1}\Big\{\tfrac{(k-1)\Delta}{2^n} < Z \leq \tfrac{k\Delta}{2^n}\Big\},$$

where $\Delta := 1$ if $t_G = \infty$, and $\Delta := t_G$ if $t_G < \infty$.

Similarly to how we defined $G(t)$, $\overline{G}(t)$, t_G and $\mathcal{T}$ for W, let us define $G^{(n)}(t)$, $\overline{G}^{(n)}(t)$, $t_{G^{(n)}}$, and $\mathcal{T}^{(n)}$ for $W^{(n)}$:

- $G^{(n)}(t) := \mathbb{P}\{W^{(n)} \leq t\}$, $t \in \mathbb{R}$;
- $\overline{G}^{(n)}(t) := \mathbb{P}\{W^{(n)} > t\} = 1 - G^{(n)}(t)$;
- $t_{G^{(n)}} := \sup\{t \in \mathbb{R}_+ \colon G^{(n)}(t) < 1\}$;
- $\mathcal{T}^{(n)} := \big\{t \in \mathbb{R}_+ \colon \mathbb{P}\{W^{(n)} \geq t\} > 0\big\}$.

It is easy to see that, for any $n \in \mathbb{N}$, $t_{G^{(n)}} = t_G$ and $\mathbb{P}\{W^{(n)} \notin \mathcal{T}^{(n)}\} = 0$; $\mathcal{T}^{(n)} = [0;\ t_G]$ in Cases **A1** and **B**; and $\mathcal{T}^{(n)} = [0;\ t_G)$ in Case **A2**.

Finally, let us introduce

- $K^{(n)}(t) := \mathbb{E}\big[V^{(n)} |\ W^{(n)} = t\big]$, where $t \in \mathcal{T}^{(n)}$;
- $F^{(n)}(t) := \int_{(0;\ t]} \frac{K^{(n)}(s)}{\overline{G}^{(n)}(s-)}\, dG^{(n)}(s)$, where $t \in \mathcal{T}^{(n)}$.

In fact, since $W^{(n)}$ is discrete and takes values $\frac{j\Delta}{2^n} \in \mathcal{T}^{(n)}$, only the values of $K^{(n)}(t)$ for $t = \frac{j\Delta}{2^n} \in \mathcal{T}^{(n)}$ are essential.

Fix $n \in \mathbb{N}$. Introduce the single jump filtration generated by $W^{(n)}$ and $\mathcal{F}$, see [5]. In accordance with Theorem 1 the process $V^{(n)}\mathbb{1}\{W^{(n)} \le t\}$, $t \in \mathbb{R}_+$, has the compensator $F^{(n)}(W^{(n)} \wedge t)$, $t \in \mathbb{R}_+$, with respect to this filtration.

Denote by $\hat{V}^{(n)}$ and $\hat{W}^{(n)}$ the values of this increasing process and its compensator at time ∞:

$$\hat{V}^{(n)} := \lim_{t\to\infty} V^{(n)}\mathbb{1}\{W^{(n)} \le t\} = V^{(n)},$$
$$\hat{W}^{(n)} := \lim_{t\to\infty} F^{(n)}(W^{(n)} \wedge t) = F^{(n)}(W^{(n)}).$$

Put also $\hat{Z}^{(n)} := Z^{(n)}$. Obviously, $\hat{V}^{(n)}$, $\hat{W}^{(n)}$, and $\hat{Z}^{(n)}$ are discrete random variables. By the construction, $\mathrm{Law}(\hat{V}^{(n)}, \hat{W}^{(n)}) \in \mathbb{W}$. Thus, taking into account Lemma 1 in Section 3, $\mathrm{Law}(\hat{V}^{(n)} + \hat{Z}^{(n)}, \hat{W}^{(n)} + \hat{Z}^{(n)}) \in \mathbb{W}$. It follows that our construction satisfies (i) and (ii). Conditions (iii) and (iv) are clear from the definitions.

The proof of condition (v) is the main point of Step 1. We defer the proof to the next section, see Lemma 4.

Assuming (v), we complete the proof of Step 1. We have

$$\mathbb{E}[\hat{W}^{(n)}] = \mathbb{E}[\hat{V}^{(n)}] \to \mathbb{E}[V] = \mathbb{E}[W] \quad \text{as} \quad n \to \infty,$$

and (vi) follows because $\hat{W}^{(n)}$ are nonnegative and integrable.

Step 2. In contrast to Step 1 we define auxiliary random variables $V^{[n]}$, $W^{[n]}$ and $Z^{[n]}$ taking finite number of values instead of discrete random variables $V^{(n)}$, $W^{(n)}$ and $Z^{(n)}$:

$$V^{[n]} := \sum_{i=1}^{r(n)} \tfrac{i\Delta}{2^n}\mathbb{1}\Big\{\tfrac{(i-1)\Delta}{2^n} < V \le \tfrac{i\Delta}{2^n}\Big\},$$
$$W^{[n]} := \sum_{j=1}^{r(n)} \tfrac{j\Delta}{2^n}\mathbb{1}\Big\{\tfrac{(j-1)\Delta}{2^n} < W \le \tfrac{j\Delta}{2^n}\Big\} + \tfrac{(r(n)+1)\Delta}{2^n}\mathbb{1}\Big\{W > \tfrac{r(n)\Delta}{2^n}\Big\},$$
$$Z^{[n]} := \sum_{k=1}^{r(n)} \tfrac{k\Delta}{2^n}\mathbb{1}\Big\{\tfrac{(k-1)\Delta}{2^n} < Z \le \tfrac{k\Delta}{2^n}\Big\},$$

where $\Delta := 1$ if $t_G = \infty$, and $\Delta := t_G$ if $t_G < \infty$, and $r \colon \mathbb{N} \to \mathbb{N}$ is a sequence such that

$$\forall n \in \mathbb{N} \;\; r(n) \ge 2^n \quad \text{and} \quad \tfrac{r(n)}{2^n} \uparrow \infty \quad \text{as} \quad n \to \infty. \tag{3}$$

The choice of $r(n)$ will be specified in Lemma 5.

Similarly to Step 1, we define

- $G^{[n]}(t) := \mathbb{P}\{W^{[n]} \le t\}$, $t \in \mathbb{R}$;
- $\overline{G}^{[n]}(t) := \mathbb{P}\{W^{[n]} > t\} = 1 - G^{[n]}(t)$;
- $t_{G^{[n]}} := \sup\{t \in \mathbb{R}_+ \colon G^{[n]}(t) < 1\}$;

- $\mathcal{T}^{[n]} := \big\{t \in \mathbb{R}_+ \colon \mathbb{P}\{W^{[n]} \geq t\} > 0\big\}$;
- $K^{[n]}(t) := \mathbb{E}\big[V^{[n]} | \, W^{[n]} = t\big]$, where $t \in \mathcal{T}^{[n]}$;
- $F^{[n]}(t) := \int_{(0;\, t]} \frac{K^{[n]}(s)}{\overline{G}^{[n]}(s-)} \, dG^{[n]}(s)$, where $t \in \mathcal{T}^{[n]}$.

Finally, we put $\hat{V}^{[n]} := V^{[n]}$, $\hat{Z}^{[n]} := Z^{[n]}$, $\hat{W}^{[n]} := F^{[n]}(W^{[n]})$. As in Step 1, $\hat{W}^{[n]}$ is a terminal value of the compensator of the process $V^{[n]}\mathbb{1}\{W^{[n]} \leq t\}$, $t \in \mathbb{R}_+$.

Our aim is to prove the statements (i)–(vi) from Step 1 for $\hat{V}^{[n]}$, $\hat{W}^{[n]}$ and $\hat{Z}^{[n]}$ instead of $\hat{V}^{(n)}$, $\hat{W}^{(n)}$ and $\hat{Z}^{(n)}$. The proof of (i)–(iv), (vi) is similar to Step 1. We only mention that $\hat{V}^{[n]}$, $\hat{W}^{[n]}$ and $\hat{Z}^{[n]}$ take values in finite sets and our assumptions on $r(n)$ imply (iii). The only significant difference between Steps 1 and 2 is in the proof of (v). In the course of the proof we will replace Lemmas 3 and 4 by Lemmas 5 and 6 respectively. □

3 Auxiliary results

Lemma 1. *If* $\mathrm{Law}(V,\ W) \in \mathbb{W}$*, then* $\mathrm{Law}(V + Z,\ W + Z) \in \mathbb{W}$ *for any nonnegative integrable random variable* Z*.*

Proof. The first two conditions in the definition of $\mathbb{W}$ are evidently satisfied for $V + Z$ and $W + Z$. Rewrite the third condition for V and W in the form

$$\mathbb{E}\big[(V - W + c)\mathbb{1}\{W \leq c\}\big] \leq c, \quad c \geq 0. \tag{4}$$

The same condition for $V + Z$ and $W + Z$ is written in this form as

$$\mathbb{E}\big[(V - W + c)\mathbb{1}\{W + Z \leq c\}\big] \leq c, \quad c \geq 0. \tag{5}$$

It remains to note that $\{W + Z \leq c\} \subseteq \{W \leq c\}$ and $V - W + c$ is nonnegative on the set $\{W \leq c\}$. Hence,

$$(V - W + c)\mathbb{1}\{W + Z \leq c\} \leq (V - W + c)\mathbb{1}\{W \leq c\},$$

and (5) follows from (4). □

Lemma 2. *Let* $\mathrm{Law}(V, W) \in \mathbb{W}$ *and condition 3) in the definition of* $\mathbb{W}$ *hold as equality for all* $c \geq 0$*. Then, for all* $0 \leq a < b < \infty$*,*

$$\mathbb{E}\big[V\mathbb{1}\{a < W \leq b\}\big] = \int_{(a;\, b]} \overline{G}(s-)\, ds.$$

Proof. By Fubini's theorem,

$$\int_{[0;\, t]} \overline{G}(s-)\, ds = \int_{[0;\, t]} \mathbb{P}\{W \geq s\}\, ds = \int_{[0;\, t]} \mathbb{E}\big[\mathbb{1}\{W \geq s\}\big]\, ds =$$
$$= \mathbb{E}\left[\int_{[0;\, t]} \mathbb{1}\{W \geq s\}\, ds\right] = \mathbb{E}[W \wedge t] = \mathbb{E}\big[V\mathbb{1}\{W \leq t\}\big],$$

where the last equality holds by assumption of the lemma. The claim follows. □

Lemma 3. *For all $t \in \mathcal{T}$, $F^{(n)}(t) \to t$ as $n \to \infty$.*

Proof. Let us introduce the set $\mathfrak{T}$ as follows:

- $\mathfrak{T} := \left\{ \frac{j\Delta}{2^n} : j \in \{0, \dots, 2^n - 1\}, \ n \in \mathbb{N} \right\}$ in Case **A1**,
- $\mathfrak{T} := \left\{ \frac{j\Delta}{2^n} : j \in \{0, 1, 2 \dots\}, \ n \in \mathbb{N} \right\}$ in Case **A2**,
- $\mathfrak{T} := \left\{ \frac{j\Delta}{2^n} : j \in \{0, \dots, 2^n\}, \ n \in \mathbb{N} \right\}$ in Case **B**.

Fix a point $t = \frac{l_0\Delta}{2^{n_0}} \in \mathfrak{T}$. Let $n \geq n_0$ and $l := \frac{t}{\Delta}2^n$. Then $\frac{l\Delta}{2^n} = t$. Consider an arbitrary point $s = \frac{j\Delta}{2^n} \in \mathfrak{T}$, where $j \in \{1, \dots, l\}$. Since $0 \leq V^{(n)} - V \leq \frac{\Delta}{2^n}$, we have

$$0 \leq \mathbb{E}\big[V^{(n)} - V \,|\, W^{(n)} = s\big] \leq \frac{\Delta}{2^n}.$$

Therefore,

$$\mathbb{E}\big[V \,|\, W^{(n)} = s\big] \leq K^{(n)}(s) = \mathbb{E}\big[V^{(n)} \,|\, W^{(n)} = s\big] \leq \frac{\Delta}{2^n} + \mathbb{E}\big[V \,|\, W^{(n)} = s\big]. \tag{6}$$

By Lemma 2,

$$\begin{aligned} \mathbb{E}\big[V \,|\, W^{(n)} = s\big] &= \frac{\mathbb{E}\big[V \mathbb{1}\{W^{(n)} = s\}\big]}{\mathbb{P}\{W^{(n)} = s\}} = \\ &= \frac{\mathbb{E}\big[V \mathbb{1}\{\frac{(j-1)\Delta}{2^n} < W \leq \frac{j\Delta}{2^n}\}\big]}{\mathbb{P}\{\frac{(j-1)\Delta}{2^n} < W \leq \frac{j\Delta}{2^n}\}} = \frac{\int_{(\frac{(j-1)\Delta}{2^n};\, \frac{j\Delta}{2^n}]} \overline{G}(s-)\, ds}{G\big(\frac{j\Delta}{2^n}\big) - G\big(\frac{(j-1)\Delta}{2^n}\big)}. \end{aligned} \tag{7}$$

It follows from (6) and (7) that

$$\frac{\int_{(\frac{(j-1)\Delta}{2^n};\, \frac{j\Delta}{2^n}]} \overline{G}(s-)\, ds}{G\big(\frac{j\Delta}{2^n}\big) - G\big(\frac{(j-1)\Delta}{2^n}\big)} \leq K^{(n)}(s) \leq \frac{\Delta}{2^n} + \frac{\int_{(\frac{(j-1)\Delta}{2^n};\, \frac{j\Delta}{2^n}]} \overline{G}(s-)\, ds}{G\big(\frac{j\Delta}{2^n}\big) - G\big(\frac{(j-1)\Delta}{2^n}\big)}. \tag{8}$$

Taking into account that $G^{(n)}\big(\frac{j\Delta}{2^n}\big) = G\big(\frac{j\Delta}{2^n}\big)$, $G^{(n)}\big(\frac{(j-1)\Delta}{2^n}\big) = G\big(\frac{(j-1)\Delta}{2^n}\big)$ and $\overline{G}^{(n)}\big(\frac{j\Delta}{2^n}-\big) = \overline{G}\big(\frac{(j-1)\Delta}{2^n}\big)$, we have

$$\begin{aligned} F^{(n)}(t) &= \int_{(0;\, t]} \frac{K^{(n)}(s)}{\overline{G}^{(n)}(s-)}\, dG^{(n)}(s) = \\ &= \sum_{j=1}^{l} \frac{K^{(n)}\big(\frac{j\Delta}{2^n}\big)}{\overline{G}^{(n)}\big(\frac{j\Delta}{2^n}-\big)} \Big(G^{(n)}\big(\frac{j\Delta}{2^n}\big) - G^{(n)}\big(\frac{(j-1)\Delta}{2^n}\big)\Big) = \\ &= \sum_{j=1}^{l} K^{(n)}\big(\frac{j\Delta}{2^n}\big) \frac{G\big(\frac{j\Delta}{2^n}\big) - G\big(\frac{(j-1)\Delta}{2^n}\big)}{\overline{G}\big(\frac{(j-1)\Delta}{2^n}\big)}. \end{aligned} \tag{9}$$

Denote

$$S_1^{(n)}(t) := \frac{\Delta}{2^n} \sum_{j=1}^{l} \frac{G\big(\frac{j\Delta}{2^n}\big) - G\big(\frac{(j-1)\Delta}{2^n}\big)}{\overline{G}\big(\frac{(j-1)\Delta}{2^n}\big)},$$

$$S_2^{(n)}(t) := \sum_{j=1}^{l} \frac{\int_{\left(\frac{(j-1)\Delta}{2^n};\, \frac{j\Delta}{2^n}\right]} \overline{G}(s-)\, ds}{\overline{G}\left(\frac{(j-1)\Delta}{2^n}\right)}.$$

Then it follows from (8) and (9) that

$$S_2^{(n)}(t) \le F^{(n)}(t) \le S_1^{(n)}(t) + S_2^{(n)}(t).$$

We want to prove that $S_1^{(n)}(t) \to 0$ and $S_2^{(n)}(t) \to t$ as $n \to \infty$.

- In Case **A1** put $t' := \Delta \frac{2l_0+1}{2^{n_0+1}}$. It is easy to see that point t' satisfies $t < t' < t_G = \Delta$. Also note that $\mathfrak{g} := \overline{G}(t') > 0$, because $t' < t_G$.
- In Case **A2** put $t'' := \Delta \frac{l_0+1}{2^{n_0}}$. It is easy to see that point t'' satisfies $t < t'' < t_G = \infty$. Also note that $\mathfrak{g} := \overline{G}(t'') > 0$, because $t'' < t_G$.
- In Case **B** it is easy to verify that $\mathfrak{g} := \overline{G}(t_G-) > 0$.

Let us prove that $S_1^{(n)}(t) \to 0$ as $n \to \infty$. Since the function $\overline{G}(s)$ is decreasing we have $\overline{G}\left(\frac{(j-1)\Delta}{2^n}\right) \ge \mathfrak{g} > 0$ for $j = 1, \dots, l$, and

$$0 \le S_1^{(n)}(t) = \frac{\Delta}{2^n} \sum_{j=1}^{l} \frac{G\left(\frac{j\Delta}{2^n}\right) - G\left(\frac{(j-1)\Delta}{2^n}\right)}{\overline{G}\left(\frac{(j-1)\Delta}{2^n}\right)} \le \frac{\Delta}{\mathfrak{g}2^n} \to 0, \quad n \to \infty.$$

In order to prove that $S_2^{(n)}(t) \to t$ as $n \to \infty$ let us rewrite $S_2^{(n)}(t)$ in the form

$$S_2^{(n)}(t) = \int_{[0;\, t)} \sum_{j=1}^{l} \frac{\overline{G}(s)}{\overline{G}\left(\frac{(j-1)\Delta}{2^n}\right)} \mathbb{1}_{\left[\frac{(j-1)\Delta}{2^n};\, \frac{j\Delta}{2^n}\right)}(s)\, ds. \tag{10}$$

Note that the integrand in (10) is bounded and converges to 1 a.e. with respect to the Lebesgue measure as $n \to \infty$. Thus, by the dominated convergence theorem, $S_2^{(n)}(t) \to t$ as $n \to \infty$.

We have proved that, for any $t \in \mathfrak{T}$,

$$F^{(n)}(t) \to t \quad \text{as} \quad n \to \infty. \tag{11}$$

A simple argument shows that this convergence holds for any $t \in \mathcal{T}$. Let $t \in \mathcal{T} \setminus \mathfrak{T}$. Given $\varepsilon > 0$, there are points $t_*,\ t^* \in \mathfrak{T}$ such that

$$t_* < t < t^* \quad \text{and} \quad t^* - t_* < \varepsilon/2. \tag{12}$$

Then $F^{(n)}(t_*) \le F^{(n)}(t) \le F^{(n)}(t^*)$. Hence,

$$t_* \le \liminf_{n\to\infty} F^{(n)}(t) \le \limsup_{n\to\infty} F^{(n)}(t) \le t^*,$$

which implies (11). $\square$

Lemma 4. $\hat{W}^{(n)} = F^{(n)}(W^{(n)}) \to W$ *as* $n \to \infty$ $\mathbb{P}$*-a.s.*

Proof. According to Lemma 3, $F^{(n)}(t) \to t$ as $n \to \infty$ for all $t \in \mathcal{T}$. We will use a well-known statement that if a sequence of increasing functions on an interval of the real line converges pointwise to a continuous function, then the convergence is uniform on compact subintervals (see, e.g., [7, Proposition 2.1]). Thus, $F^{(n)}(t)$ converges uniformly to t on $\mathcal{T}$ in Case **B**, and on any interval $[0; b] \subseteq \mathcal{T}$, where $b < t_G$, in Cases **A1** and **A2**.

Recall that $W^{(n)}$ converges to W for all $\omega \in \Omega$ by the construction. In Case **B**, for all $n \in \mathbb{N}$, $\mathbb{P}(W \in [0; t_G]) = \mathbb{P}(W^{(n)} \in [0; t_G]) = 1$. Thus, for all $n \in \mathbb{N}$, we have $\mathbb{P}$-a.s.

$$\begin{aligned}\left|F^{(n)}(W^{(n)}) - W\right| &\le \left|F^{(n)}(W^{(n)}) - W^{(n)}\right| + |W^{(n)} - W| \le \\ &\le \sup_{t\in[0;\, t_G]} |F^{(n)}(t) - t| + |W^{(n)} - W|,\end{aligned}$$

which implies $\hat{W}^{(n)} = F^{(n)}(W^{(n)}) \to W$ as $n \to \infty$ $\mathbb{P}$-a.s.

In Cases **A1** and **A2** fix ω such that $W(\omega) \in \mathcal{T} = [0; t_G)$. Take $b = b(\omega)$ such that $W(\omega) < b < t_G$. Then $W^{(n)}(\omega) < b$ for n large enough, and we have $F^{(n)}(W^{(n)}(\omega)) \to W(\omega)$ as $n \to \infty$ in view of the above inequality with t_G replaced by b. Since $\mathbb{P}(W \in \mathcal{T}) = 1$, the requested convergence holds $\mathbb{P}$-a.s. $\square$

Lemma 5. *There exists an increasing sequence* $r(n) \in \mathbb{N}$*,* $n \in \mathbb{N}$*, satisfying* (3) *and such that* $F^{[n]}(t) \to t$ *as* $n \to \infty$ *for all* $t \in \mathcal{T}$*.*

Proof. Let $\mathfrak{T}$ be defined as in the proof of Lemma 3. Fix a point $t = \frac{l_0\Delta}{2^{n_0}} \in \mathfrak{T}$. Note that $n_0 = n_0(t)$. Let $n \ge n_0(t)$ and $l := \frac{t}{\Delta}2^n$. Then $\frac{l\Delta}{2^n} = t$. Let $s = \frac{j\Delta}{2^n} \in \mathfrak{T}$, where $j \in \{1, \dots, l\}$. It follows from (9) and from

$$K^{(n)}(s) = \sum_{i=1}^{\infty} \frac{i\Delta}{2^n} \frac{\mathbb{P}\Big(\big\{\frac{(i-1)\Delta}{2^n} < V \le \frac{i\Delta}{2^n}\big\} \bigcap \big\{\frac{(j-1)\Delta}{2^n} < W \le \frac{j\Delta}{2^n}\big\}\Big)}{\mathbb{P}\big\{\frac{(j-1)\Delta}{2^n} < W \le \frac{j\Delta}{2^n}\big\}}$$

that

$$F^{(n)}(t) = \sum_{j=1}^{l}\sum_{i=1}^{\infty} \mathfrak{a}(i, j, n), \tag{13}$$

where $F^{(n)}(t)$ is the function that was defined in Step 1 and

$$\mathfrak{a}(i, j, n) := \frac{i\Delta}{2^n} \frac{\mathbb{P}\Big(\big\{\frac{(i-1)\Delta}{2^n} < V \le \frac{i\Delta}{2^n}\big\} \bigcap \big\{\frac{(j-1)\Delta}{2^n} < W \le \frac{j\Delta}{2^n}\big\}\Big)}{\mathbb{P}\big\{W > \frac{(j-1)\Delta}{2^n}\big\}}.$$

Since, for a fixed $n \ge n_0(t)$, the series (13) with nonnegative terms converges, for any $n \ge n_0(t)$, one can find $r(t, n)$ such that

$$\sum_{j=1}^{l}\sum_{i=1}^{\infty} \mathfrak{a}(i, j, n) - \sum_{j=1}^{l}\sum_{i=1}^{r(t,n)} \mathfrak{a}(i, j, n) < 1/n. \tag{14}$$

For $n < n_0(t)$ put $r(t,n) := 2^n$. Without loss of generality, we may assume that a sequence $r(t,n)$, $n \in \mathbb{N}$, satisfies conditions (3), i.e., $r(t,1) \le \cdots \le r(t,n) \le \cdots$ for all $n \in \mathbb{N}$, $2^n \le r(t,n)$ and $\frac{r(t,n)}{2^n} \uparrow \infty$ as $n \to \infty$.

In order to construct a sequence $r(n)$, $n \in \mathbb{N}$, which guarantees that the statement of this lemma is true, we shall use a version of Cantor's diagonalization procedure. Enumerate somehow elements of the countable set $\mathfrak{T}$, so that $\mathfrak{T} = \{t_k\}_{k=1}^{\infty}$. Put $r(n) := \max\{r(t_1,n),\dots,r(t_n,n)\}$, $n \ge 1$. Then the sequence $r(n)$ has the property

$$\forall k \in \mathbb{N} \ \ \forall n \ge k \quad r(t_k,n) \le r(n). \tag{15}$$

Fix $k \in \mathbb{N}$ (thus we fix $t_k \in \mathfrak{T}$). Then by (14) and (15), for all $n \ge \max(k,\ n_0(t_k))$, we have

$$\begin{aligned} &\sum_{j=1}^{l}\sum_{i=1}^{\infty}\mathfrak{a}(i,j,n) - \sum_{j=1}^{l}\sum_{i=1}^{r(n)}\mathfrak{a}(i,j,n) \overset{(15)}{\le} \\ \overset{(15)}{\le} &\sum_{j=1}^{l}\sum_{i=1}^{\infty}\mathfrak{a}(i,j,n) - \sum_{j=1}^{l}\sum_{i=1}^{r(t_k,n)}\mathfrak{a}(i,j,n) \overset{(14)}{\le} \frac{1}{n}. \end{aligned} \tag{16}$$

Now let us check that the sequence $r(n)$ constructed above meets the requirements of the lemma. Since $\frac{r(n)}{2^n} \uparrow \infty$ as $n \to \infty$, for any $t \in \mathfrak{T}$, one can find $N_0(t) \ge n_0(t)$ such that $t \le \frac{r(n)\Delta}{2^n}$ for all $n \ge N_0(t)$. Let $n \ge N_0(t)$. The function $F^{[n]}(t)$ has the representation

$$F^{[n]}(t) = \sum_{j=1}^{l}\sum_{i=1}^{r(n)}\mathfrak{a}(i,j,n). \tag{17}$$

Again, let us fix $k \in \mathbb{N}$. It follows from (13), (17) and (16) that if $n \ge \max(k, N_0(t_k))$ then

$$F^{(n)}(t_k) - F^{[n]}(t_k) \overset{(13),(17)}{=} \sum_{j=1}^{l}\sum_{i=1}^{\infty}\mathfrak{a}(i,j,n) - \sum_{j=1}^{l}\sum_{i=1}^{r(n)}\mathfrak{a}(i,j,n) \overset{(16)}{\le} \frac{1}{n}.$$

Due to Lemma 4, it follows that $F^{[n]}(t_k) \to t_k$ as $n \to \infty$. Using the same arguments as at the end of the proof of Lemma 3, we obtain that the convergence takes place not only on $\mathfrak{T}$ but on $\mathcal{T}$ as well. □

Lemma 6. *$\hat{W}^{[n]} = F^{[n]}(W^{[n]}) \to W$ as $n \to \infty$ $\mathbb{P}$-a.s.*

The proof of the lemma is based on Lemma 5 and is similar to the proof of Lemma 4.

Acknowledgments

We thank two anonymous referees for constructive comments and suggestions for improving the presentation.

Funding

The article was prepared within the framework of the Basic Research Program at HSE University.

References

[1] Borzykh, D.A.: Joint distributions of generalized integrable increasing processes and their generalized compensators. Theory Probab. Appl. (2022).

[2] Cox, A.M.G., Kinsley, S.M.: Discretisation and duality of optimal Skorokhod embedding problems. Stoch. Process. Appl. **129**(7), 2376–2405 (2019). MR3958436. https://doi.org/10.1016/j.spa.2018.07.008

[3] Föllmer, H., Schied, A.: Stochastic Finance. An introduction in discrete time. De Gruyter Graduate. Fourth revised and extended edition of [MR1925197]. De Gruyter, Berlin (2016). 596 pp. MR3859905. https://doi.org/10.1515/9783110463453

[4] Gushchin, A.A.: The joint law of terminal values of a nonnegative submartingale and its compensator. Theory Probab. Appl. **62**(2), 216–235 (2018). MR3649035. https://doi.org/10.1137/S0040585X97T988575

[5] Gushchin, A.A.: Single jump filtrations and local martingales. Mod. Stoch. Theory Appl. **7**(2), 135–156 (2020). MR4120612

[6] Gushchin, A.A.: The joint law of a max-continuous local submartingale and its maximum. Theory Probab. Appl. **65**(4), 545–557 (2021). MR4167880. https://doi.org/10.4213/tvp5339

[7] Resnick, S.I.: Heavy-tail Phenomena. Probabilistic and statistical modeling. Springer Series in Operations Research and Financial Engineering. Springer (2007), 404 pp. MR2271424

[8] Rogers, L.C.G.: The joint law of the maximum and terminal value of a martingale. Probab. Theory Relat. Fields **95**(4), 451–466 (1993). MR1217446. https://doi.org/10.1007/BF01196729