THEORY PROBAB. APPL.
Vol. 65, No. 4, pp. 656–664

CONVERGENCE OF CERTAIN CLASSES OF RANDOM FLIGHTS IN THE KANTOROVICH METRIC*

V. D. KONAKOV† AND A. R. FALALEEV‡

(*Translated by A. R. Alimov*)

Abstract. A random walk of a particle in $\mathbf{R}^d$ is considered. The weak convergence of various transformations of trajectories of random flights with Poisson switching times was studied by Davydov and Konakov in [*Random walks in nonhomogeneous Poisson environment*, in Modern Problems of Stochastic Analysis and Statistics, Springer, 2017, pp. 3–24], who also built a diffusion approximation of the process of random flights. The goal of the present paper is to prove a stronger convergence with respect to the Kantorovich distance. Three types of transformations are considered. The cases of exponential and superexponential growth of the switching time transformation function are quite simple—in these cases the required result follows from the fact that the limit processes lie within the unit ball. In the case of a power-like growth of the transformation function, the convergence follows from combinatorial arguments and properties of the Kantorovich metric.

Key words. Kantorovich metric, random walk of a particle, convergence of transformations of paths of random flights, Doob's maximal inequality

DOI. 10.1137/S0040585X97T990204

1. Introduction. Random flights have many physical applications. As an example, we mention a Lorentz process, which is a stochastic process defined by a particle moving according to Newton's law of motion through static scatterers distributed in the space in accordance with some probability measure. Consider the Boltzmann-Grad limit: the density of the scatterers increases to infinity as the diameter of the scatterers decreases to zero so that the average free path of the particle remains constant. A Lorentz process is known to converge to a stochastic process in the weak* topology of regular Borel measures on the space of trajectories. The limit process is a Markov process if and only if the scaled density of the scatterers converges in probability to its average value. In this case, the limit process is a (spatially inhomogeneous) random flight process.

Let $(\mathcal{X}, \mathbf{d})$ be a Polish space and let $p \in [1, \infty)$. The Kantorovich space of order p is defined as

$$P_p(\mathcal{X}) := \left\{ \mu \in P(\mathcal{X});\ \int_{\mathcal{X}} d(x_0, x)^p \, \mu(dx) < +\infty \right\}$$

for some (and hence for any) $x_0 \in \mathcal{X}$, where $P(\mathcal{X})$ is the set of all probability measures on $\mathcal{X}$.

For any two probability measures μ, ν on $\mathcal{X}$, the Kantorovich distance of order p between μ and ν is defined by

$$\begin{aligned} W_p(\mu, \nu) &= \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}} d(x, y)^p \, d\pi(x, y) \right)^{1/p} \\ &= \inf\bigl\{ [\mathbf{E} \mathbf{d}(X, Y)^p]^{1/p},\ \mathrm{law}(X) = \mu,\ \mathrm{law}(Y) = \nu \bigr\}. \end{aligned}$$

It is well known that the Kantorovich distance W_p metrizes the weak convergence in $P_p(\mathcal{X})$ (see [1, Theorem 6.9]). However, it is worth pointing out that the "weak convergence" in the sense of [1] is stronger than the classical weak convergence [4] (see [1, Definition 6.8]). These two types of convergence are equivalent if the metric d is bounded;

*Received by the editors October 14, 2019; revised December 25, 2019. This work was supported by the Russian Science Foundation (grant 20-11-20119). Originally published in the Russian journal *Teoriya Veroyatnostei i ee Primeneniya*, 65 (2020), pp. 829–840.
https://doi.org/10.1137/S0040585X97T990204

†National Research University Higher School of Economics, Moscow, Russia (VKonakov@hse.ru).
‡Lomonosov Moscow State University, Moscow, Russia (mrfal9@gmail.com).

however, in general, they are different. In [3], the weak convergence result was established in the classical sense. This is why the convergence in the Kantorovich metric should be justified, because it does not automatically follow from the weak convergence established in [3].

Consider a random walk of a particle in $\mathbf{R}^d$ defined by two independent sequences of random variables (r.v.'s) T_k and ε_k. The sequence ε_k, which consists of independent r.v.'s distributed on the unit sphere S^{d-1}, controls the direction of motion of the particle. The sequence T_k, which is such that $T_k \geqslant 0$, $T_k \leqslant T_{k+1}$ for any k, can be interpreted as a sequence of times when the flight direction of the particle changes. The particle starts from the origin and moves in the direction ε_1 up to time T_1 when it changes the direction to ε_2 and moves in this direction for a period of duration $T_2 - T_1$, and so on. The velocity is constant on all intervals. The position of the particle at time t is denoted by $X(t)$. The paper [3] gives conditions for the process $\{Y_T, T > 0\}$,

$$Y_T(t) = \frac{1}{B(T)} X(tT), \qquad t \in [0,1],$$

to converge weakly in $\mathbf{C}[0,1]$: $Y_T \Rightarrow Y$, $T \to \infty$, and $B(T) \to \infty$.

The switching times are assumed to form a Poisson process $\mathbb{T} = (T_k)$ in $\mathbf{R}_+$. In the homogeneous case, the process $X(t)$ is a random walk, because the intervals $T_{k+1} - T_k$ are independent and Y is a Wiener process. However, the situation is more challenging in the case of an inhomogeneous Poisson process, because the increments $T_{k+1} - T_k$ cease to be independent.

Nevertheless, the form of the limiting process was found and some Poisson switching time transformation functions were shown to be weakly convergent. Let $T_k = f(\Gamma_k)$, where (Γ_k) is a standard homogeneous Poisson process in $\mathbf{R}_+$ of intensity 1. In this case,

$$(\Gamma_k) = (\gamma_1 + \gamma_2 + \cdots + \gamma_k),$$

where (γ_k) are standard independent identically distributed exponential r.v.'s and $f(x)$ is a regular function of polynomial (exponential, superexponential) growth. We also assume that $\mathbf{E}\varepsilon_1 = 0$.

Consider the process

$$Z_n(t) = Y_{T_n}(t).$$

For $T = T_n$, the trajectories $\{Z_n(t),\, t \in [0,1]\}$ are continuous broken lines with vertices at the points $\{(t_{n,k}, S_k/B_n),\, k = 0, 1, \ldots, n\}$, where $t_{n,k} = T_k/T_n$, $T_0 = 0$, $B_n = B(T_n)$, $S_k = \sum_{i=1}^{k} \varepsilon_i (T_i - T_{i-1})$.

We formulate the main result of the first part of [3] as follows.

THEOREM 1. *The following results hold under the above assumptions.*

(1) *If the function f grows polynomially* $(f(t) = t^\alpha$, $\alpha > 1/2)$, *we set* $B(T) = T^{(2\alpha-1)/(2\alpha)}$. *Then the process Z_n converges weakly to Y, where Y is a Gaussian process*

$$Y(t) = \sqrt{2\alpha} \int_0^t s^{(\alpha-1)/(2\alpha)}\, dw(s),$$

where w is a process of Brownian motion such that the covariance matrix $w(1)$ *coincides with the covariance matrix* ε_1.

(2) *If the function f grows exponentially* $(f(t) = e^{t\beta}$, $\beta > 0)$, *we set* $B(T) = T$. *Then the process Z_n converges weakly to Y, where Y is a continuous piecewise-linear process with vertices at the points* $(t_k, Y(t_k))$,

$$t_k = e^{-\beta \Gamma_{k-1}}, \quad \Gamma_0 = 0,$$
$$Y(t_k) = \sum_{i=k}^{\infty} \varepsilon_i (e^{-\beta \Gamma_{i-1}} - e^{-\beta \Gamma_i}), \quad Y(0) = 0.$$

(3) *In the case of a superexponential growth of the function f, we assume that f is an increasing continuous function such that*

$$\lim_{t\to\infty} \frac{f'(t)}{f(t)} = +\infty.$$

Let $B(T) = T$. Then $T_n/T_{n+1} \to 0$ in probability and $Z_n \Rightarrow Y$, where the limit process degenerates as

$$Y(t) = \varepsilon_1 t, \qquad t \in [0,1].$$

Recall that the goal of the present paper is to prove a stronger convergence, namely, the convergence with respect to the Kantorovich distance. Throughout the proofs given below, constants may assume different values in different relations.

2. The main result. In what follows, we put $\mathcal{X} = \mathbf{C}[0,1]$ and $\mathbf{d}(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|$. For a continuous random process $X(t)$, $t \in [0,1]$, we denote by μ_X the measure in $\mathbf{C}[0,1]$ corresponding to this process.

THEOREM 2. *Let $(\mathcal{X}, \mathbf{d})$ be a Polish space and let $p \in [1,\infty)$. Then*

$$W_p(\mu_{X_n}, \mu_Y) \to 0,$$

where the process X_n in cases (1)–(3) *is a broken line with vertices at the points $(t_{n,k}, X_n(t_{n,k}))$.*

In case (1),

$$t_{n,k} = \left(\frac{\Gamma_k}{\Gamma_n}\right)^{\alpha}, \quad X_n(t_{n,k}) = n^{1/2-\alpha} \sum_{i=1}^{k} \varepsilon_i (\Gamma_i^{\alpha} - \Gamma_{i-1}^{\alpha}),$$
$$\Gamma_0^{\alpha} = 0, \qquad k = 1, \dots, n.$$

The limit process $Y(t)$ is a Gaussian process with integral representation

$$Y(t) = \sqrt{2\alpha} \int_0^t s^{(\alpha-1)/(2\alpha)}\, dw(s),$$

where $w(s)$ is a Brownian motion with the covariance matrix $w(1)$, which is equal to the covariance matrix ε_1.

In case (2),

$$t_{n,k} = e^{-\beta(\Gamma_n - \Gamma_k)}, \quad X_n(t_{n,k}) = e^{-\beta\Gamma_n} \sum_{i=1}^{k} \varepsilon_i (e^{\beta\Gamma_i} - e^{\beta\Gamma_{i-1}}),$$
$$\Gamma_0 = 0, \qquad k = 1, \dots, n.$$

The limit process $Y(t)$ is a continuous piecewise-linear process with a countable number of vertices $(t_k, Y(t_k))$, $k = 1, 2, \dots$, $t_k = e^{-\beta\Gamma_{k-1}}$, $\Gamma_0 = 0$,

$$Y(t_k) = \sum_{i=k}^{\infty} \varepsilon_i (e^{-\beta\Gamma_{i-1}} - e^{-\beta\Gamma_i}).$$

In case (3),

$$t_{n,k} = \frac{f(T_k)}{f(T_n)}, \quad X_n(t_{n,k}) = \frac{1}{f(T_n)} \sum_{i=1}^{k} \varepsilon_i (f(\Gamma_i) - f(\Gamma_{i-1})),$$
$$\Gamma_0 = 0, \qquad k = 1, \dots, n.$$

The limit process $Y(t)$ degenerates as

$$Y(t) = \varepsilon_1 t, \qquad t \in [0,1].$$

As for the intuitive explanation of limit processes, in the polynomial case such an explanation was actually given in [3]. The original process was approximated by a process in the form of a broken line constructed from a sequence of partial sums of independent (non-identically distributed) r.v.'s. In this case, Prokhorov's theorem applies. The exponential case can be verified by a direct calculation, and the superexponential case appears as the degeneration of the exponential case, when the entire broken line up to its next-to-last link is concentrated in an arbitrarily small neighborhood of the origin, and the last link degenerates to a ray emanating from the origin.

3. Auxiliary definitions and results.

DEFINITION 1 (the weak convergence in P_p). *Let* $(\mathcal{X}, \mathbf{d})$ *be a Polish space and let* $p \in [1, \infty)$. *Next, let* (μ_k), $k \in \mathbf{N}$, *be a sequence of probability measures in* $P_p(X)$ *and let* $\mu \in P_p(\mathcal{X})$. *We say that "*μ_k *converges weakly in* $P_p(X)$*" if any of the following equivalent conditions holds for some* (*and hence for any*) $x_0 \in X$:

(i) $\mu_k \Rightarrow \mu$, $k \to \infty$, *and*

$$\int d(x_0, x)^p \, d\mu_k(x) \to \int d(x_0, x)^p d\mu(x);$$

(ii) $\mu_k \Rightarrow \mu$, $k \to \infty$, *and*

$$\limsup_{k\to\infty} \int d(x_0, x)^p \, d\mu_k(x) \leqslant \int d(x_0, x)^p \, d\mu(x);$$

(iii) $\mu_k \Rightarrow \mu$, $k \to \infty$, *and*

$$\lim_{R\to\infty} \limsup_{k\to\infty} \int_{d(x_0,x)\geqslant R} d(x_0, x)^p \, d\mu_k(x) = 0;$$

(iv) *for any continuous function* φ *such that* $|\varphi(x)| \leqslant C(1 + d(x_0, x)^p)$, $C \in \mathbf{R}_+$, *we have*

$$\int \varphi(x) \, d\mu_k(x) \to \int \varphi(x) \, d\mu(x).$$

THEOREM 3 (the distance W_p metrizes $P_p(\mathcal{X})$, [1, Theorem 6.9]). *Let* $(\mathcal{X}, \mathbf{d})$ *be a Polish space and let* $p \in [1, \infty)$. *Then the Kantorovich distance* W_p *metrizes the "weak convergence in* $P_p(\mathcal{X})$*." In other words, if* $(\mu_k)_{k\in\mathbf{N}}$ *is a sequence of probability measures in* $P_p(\mathcal{X})$ *and* μ *is a measure in* $P_p(\mathcal{X})$, *then the conditions*

$$\mu_k \text{ “converges weakly in } P_p(\mathcal{X})\text{” to } \mu$$

and

$$W_p(\mu_k, \mu) \to 0$$

are equivalent.

For a proof we require additional estimates.

THEOREM 4 (Doob's maximal inequality [2]). *If* X_k *is a martingale or a positive submartingale indexed by a finite set* $k \in (0, 1, \ldots, N)$, *then, for any* $p \geqslant 1$ *and* $\lambda > 0$,

$$\lambda^p \mathbf{P}\Big[\sup_{0\leqslant k\leqslant N} |X_k| \geqslant \lambda\Big] \leqslant \mathbf{E}[|X_N|^p].$$

The following estimates can be found in [3].

LEMMA 1. *Let* $\alpha > 0$ *and* $m \geqslant 1$. *Then, for any* $x > 0$, $h > 0$,

$$(x + h)^\alpha - x^\alpha = \sum_{k=1}^{m} a_k h^k x^{\alpha-k} + R(x, h),$$

where

$$a_k = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!},$$
$$|R(x,h)| \leqslant |a_{m+1}|h^{m+1}\max\{x^{\alpha-(m+1)},(x+h)^{\alpha-(m+1)}\}.$$

LEMMA 2. *Let $\alpha \geqslant 0$. Then, as $k\to\infty$,*

$$\left(1+\frac{\alpha}{k}\right)^k = e^{\alpha} + O\left(\frac{1}{k}\right).$$

LEMMA 3. *Let Γ be the Euler gamma function. Then, as $k\to\infty$,*

$$\frac{\Gamma(k+\alpha)}{\Gamma(k)} = k^{\alpha} + O(k^{\alpha-1}).$$

LEMMA 4. *For any real β and $k\to\infty$,*

$$\mathbf{E}\Gamma_k^{\beta} = k^{\beta} + O(k^{\beta-1}).$$

LEMMA 5. *Let $\alpha \geqslant 0$. If $k\to\infty$, then the following formulas hold*:

$$\Gamma_{k+1}^{\alpha} - \Gamma_k^{\alpha} = \alpha\gamma_{k+1}\Gamma_k^{\alpha-1} + \rho_k,$$

where $|\rho_k| = O(k^{\alpha-2})$ in probability, and

$$\mathbf{E}|\Gamma_{k+1}^{\alpha} - \Gamma_k^{\alpha}|^2 = 2\alpha^2 k^{2\alpha-2} + O(k^{2\alpha-3}).$$

The following result is a consequence of Lemma 5.

COROLLARY 1. *The following equality holds:*

$$\sum_{1}^{n-1}\mathbf{E}|\Gamma_{k+1}^{\alpha} - \Gamma_k^{\alpha}|^2 = \frac{2\alpha^2}{2\alpha-1}n^{2\alpha-1} + O(n^{2\alpha-2}).$$

4. Proof of Theorem 2. There are three cases to consider.

The case of exponential growth. The switching time transformation function reads as $f(t) = e^{t\beta}$, $\beta > 0$, $B(T) = T$, and the process Z_n converges weakly to Y, where Y is a continuous piecewise-linear process with vertices at the points $(t_k, Y(t_k))$,

$$t_k = e^{-\beta\Gamma_{k-1}}, \quad \Gamma_0 = 0,$$
$$Y(t_k) = \sum_{i=k}^{\infty}\varepsilon_k(e^{-\beta\Gamma_{i-1}} - e^{-\beta\Gamma_i}), \quad Y(0) = 0.$$

For $T = T_n$, the trajectories $\{Z_n(t),\, t\in[0,1]\}$ are continuous broken lines with vertices at the points $\{(t_{n,k}, S_k/B_n),\, k = 0,1,\dots,n\}$, where $t_{n,k} = T_k/T_n$, $T_0 = 0$, $B_n = B(T_n)$, $S_k = \sum_{i=1}^{k}\varepsilon_i(T_i - T_{i-1})$.

So, the trajectories of the process are broken lines with vertices at the points $(t_{n,k}, X_n(t_{n,k}))$,

$$X_n(t_{n,k}) = \frac{1}{e^{\beta\Gamma_n}}\sum_{i=1}^{k}\varepsilon_i(e^{\beta\Gamma_i} - e^{\beta\Gamma_{i-1}}).$$

We have $X_n(\,\cdot\,) \stackrel{\mathcal{L}}{=} Y_n(\,\cdot\,)$ (see [3, p. 11]), where $Y_n(\,\cdot\,)$ is the broken line with vertices $(\tau_{n,k}, Y_n(\tau_{n,k}))$, $(\tau_{n,k})\!\downarrow$, $\tau_{n,1} = 1$, $\tau_{n,k} = e^{-\beta(\gamma_1+\cdots+\gamma_{k-1})}$, $k = 2,\dots,n$,

$$Y_n(\tau_{n,k}) = \sum_{i=k}^{n-1}\varepsilon_i(e^{-\beta\Gamma_{i-1}} - e^{-\beta\Gamma_i}) + \varepsilon_n e^{-\beta\Gamma_{n-1}},$$

$Y_n(0) = 0$, and $\Gamma_0 = 0$.

Since $Y_n(\tau_{n,k})$ is a sum of nonnegative terms multiplied by the random vector ε_i, $|\varepsilon_i| = 1$, we have

$$\max_{k=1,\dots,n} |Y_n(\tau_{n,k})| \leqslant \sum_{i=1}^{n-1} (e^{-\beta\Gamma_{i-1}} - e^{-\beta\Gamma_i}) + e^{-\beta\Gamma_{n-1}} = 1.$$

Therefore, for $R > 1$,

$$\mu_n(\mathbf{d}(\mathbf{0}, x) \geqslant R) = \mathbf{P}\Big(\max_{0\leqslant t\leqslant 1} |Y_n(t)| \geqslant R\Big) = 0.$$

Hence

$$\lim_{R\to\infty} \overline{\lim_{n\to\infty}} \int_{\mathbf{d}(\mathbf{0},x)\geqslant R} \mathbf{d}^p(\mathbf{0}, x)\, d\mu_n(x) = 0.$$

This proves the convergence $W_p(\mu_n, \mu) \to 0$ for any $p > 1$.

The case of superexponential growth. In this case, putting $B_n = B(T_n) = T_n$, we have

$$\max_{k=1,\dots,n} |X_n(t_{n,k})| \leqslant \sum_{k=1}^{n} \frac{T_k - T_{k-1}}{T_n} = \frac{T_n}{T_n} = 1.$$

Therefore, for $R > 1$

$$\mu_n(\mathbf{d}(\mathbf{0}, x) \geqslant R) = \mathbf{P}\Big(\max_{0\leqslant t\leqslant 1} |X_n(t)| \geqslant R\Big) = 0.$$

Then

$$\lim_{R\to\infty} \overline{\lim_{n\to\infty}} \int_{\mathbf{d}(\mathbf{0},x)\geqslant R} \mathbf{d}^p(0, x)\, d\mu_n(x) = 0.$$

This proves the convergence $W_p(\mu_n, \mu) \to 0$ for any $p > 1$.

The case of polynomial growth. Note that if ε_j is a uniformly distributed r.v. on the unit ball in $\mathbf{R}^d$, then $\langle \varepsilon_i, e_j \rangle$ is a one-dimensional r.v. distributed symmetrically with respect to zero. Since the odd moments of such an r.v. are 0, we have

$$T_k = \Gamma_k^\alpha, \quad \alpha > \frac{1}{2}, \quad t_{n,k} = \frac{T_k}{T_n} = \left(\frac{\Gamma_k}{\Gamma_n}\right)^\alpha, \quad B_n = n^{\alpha-1/2},$$

$$\Gamma_0^\alpha = 0, \quad X_n(t_{n,k}) = \frac{1}{B_n} \sum_{i=1}^{k} \varepsilon_i (\Gamma_i^\alpha - \Gamma_{i-1}^\alpha).$$

This gives us the upper estimate

$$\begin{aligned}
&\mathbf{P}\left(\max_{k=1,\dots,n} \left| \frac{1}{B_n} \sum_{i=1}^{k} \langle \varepsilon_i, e_j \rangle (\Gamma_i^\alpha - \Gamma_{i-1}^\alpha) \right| \geqslant 3R\right) \\
&\quad\leqslant \mathbf{P}\left(\max_{k=1,\dots,n} \left| \frac{1}{B_n} \sum_{i=1}^{k} \langle \varepsilon_i, e_j \rangle (\Gamma_i^\alpha - \Gamma_{i-1}^\alpha) - \frac{\alpha}{B_n} \sum_{i=1}^{k} \langle \varepsilon_i, e_j \rangle \gamma_i \Gamma_{i-1}^{\alpha-1} \right| > R\right) \\
&\qquad + \mathbf{P}\left(\max_{k=1,\dots,n} \left| \frac{\alpha}{B_n} \sum_{i=1}^{k} \langle \varepsilon_i, e_j \rangle \gamma_i \Gamma_{i-1}^{\alpha-1} - \frac{\alpha}{B_n} \sum_{i=1}^{k} \langle \varepsilon_i, e_j \rangle \gamma_i (i-1)^{\alpha-1} \right| > R\right) \\
&\qquad + \mathbf{P}\left(\max_{k=1,\dots,n} \left| \frac{\alpha}{B_n} \sum_{i=1}^{k} \langle \varepsilon_i, e_j \rangle \gamma_i (i-1)^{\alpha-1} \right| > R\right) = \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}$$

Estimate for I. We use Doob's maximal inequality with $\lambda = B_n R$, $p = 2N$. Let $\mathfrak{M}_n = \sigma(\gamma_1, \dots, \gamma_n)$ be the filtration generated by $(\gamma_1, \dots, \gamma_n)$. Then the process

$$A_k^\alpha = \sum_{i=1}^{k} \langle \varepsilon_i, e_j \rangle (\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha \gamma_i \Gamma_{i-1}^{\alpha-1})$$

becomes a conditional martingale. By Doob's maximal inequality

$$I = \mathbf{E}\Big(\mathbf{P}\Big(\max_{k=1,\dots,n} |A_k^\alpha| > B_n R\Big) \Bigm| \mathfrak{M}_n\Big) \leqslant \frac{1}{R^{2N} n^{2N\alpha - N}} \mathbf{E}(A_k^\alpha)^{2N}$$
$$= \frac{n^{N-2N\alpha}}{R^{2N}} \sum_{k_1+\cdots+k_n=2N} \frac{(2N)!}{k_1!\cdots k_n!} \prod_{i=1}^{n} \mathbf{E}(\langle \varepsilon_i, e_j\rangle)^{k_i} \mathbf{E}(\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha\gamma_i \Gamma_{i-1}^{\alpha-1})^{k_i}.$$

Note that if there is at least one odd number among k_i, then the corresponding term in the sum is zero due to the symmetry of the distribution ε_i with respect to 0.

Using (19) in [3], we have, for $1/2 < \alpha < 2$,

$$\big|\mathbf{E}(\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha\gamma_i\Gamma_{i-1}^{\alpha-1})^{k_i}\big| \leqslant \mathbf{E}|\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha\gamma_i\Gamma_{i-1}^{\alpha-1}|^{k_i}$$
$$\leqslant C(\alpha, N)\mathbf{E}(\gamma_i)^{2k_i}\mathbf{E}\Gamma_{i-1}^{k_i(\alpha-2)} \leqslant C(\alpha, N)(2k_i)!\, i^{(\alpha-2)k_i}.$$

Therefore, since $\alpha < 2$, we have

$$\left|\prod_{i=1}^{n} \mathbf{E}(\langle\varepsilon_i, e_j\rangle)^{k_i}\mathbf{E}(\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha\gamma_i\Gamma_{i-1}^{\alpha-1})^{k_i}\right| \leqslant C(\alpha, N)\prod_{i=1}^{n} i^{k_i(\alpha-2)} \leqslant C(\alpha, N).$$

For $\alpha \geqslant 2$, an appeal to the Cauchy–Schwarz inequality shows that

$$\mathbf{E}\big(\gamma_i^{2k_i}\Gamma_i^{(\alpha-2)k_i}\big) \leqslant \sqrt{\mathbf{E}(\gamma_i)^{4k_i}}\sqrt{\mathbf{E}\Gamma_i^{(2\alpha-4)k_i}} \leqslant C(\alpha, N) i^{(\alpha-2)k_i}$$

and

$$\big|\mathbf{E}(\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha\gamma_i\Gamma_{i-1}^{\alpha-1})^{k_i}\big| \leqslant \mathbf{E}|\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha\gamma_i\Gamma_{i-1}^{\alpha-1}|^{k_i} \leqslant C(\alpha, N) i^{(\alpha-2)k_i}.$$

Therefore, by Lemma 5

$$\left|\frac{(2N)!}{k_1!\cdots k_n!}\prod_{i=1}^{n}\mathbf{E}(\langle\varepsilon_i, e_j\rangle)^{k_i}\mathbf{E}(\Gamma_i^\alpha - \Gamma_{i-1}^\alpha - \alpha\gamma_i\Gamma_{i-1}^{\alpha-1})^{k_i}\right|$$
$$\leqslant C(\alpha, N)\prod_{i=1}^{n} i^{(\alpha-2)k_i} \leqslant C(\alpha, N)\prod_{i=1}^{n} n^{(\alpha-2)k_i} = C(\alpha, N) n^{2N\alpha - 4N}.$$

Let us estimate the number of nonzero terms in the sum $\sum_{k_1+\cdots+k_n=2N}$. The constraint $C(N)n^N$ on the number of terms is a consequence of simple combinatorial arguments. We have

$$\mathrm{I} \leqslant \frac{C(\alpha, N)}{R^{2N} n^{2N\alpha-N}} n^{2N\alpha - 4N} n^N < \frac{C(\alpha, N)}{R^{2N}}. \tag{4.1}$$

This estimates is sufficient for verifying assertion (iii) in Definition 1.

Estimate for II. Arguing as before and taking into account the independence γ_i and Γ_{i-1}^α, we get

$$\mathrm{II} \leqslant \frac{\alpha^{2N}}{R^{2N} n^{2N\alpha-N}} \sum_{k_1+\cdots+k_n=2N} D_{\overline{k}},$$

where

$$D_{\overline{k}} = \frac{(2N)!}{k_1!\cdots k_n!}\prod_{i=2}^{n}\mathbf{E}(\langle\varepsilon_i, e_j\rangle)^{k_i}\mathbf{E}(\gamma_i)^{k_i}\mathbf{E}\big(\Gamma_{i-1}^{\alpha-1} - (i-1)^{\alpha-1}\big)^{k_i},$$

$\overline{k} = (k_1, \dots, k_n)$. Let us estimate the expectation. We have

$$\mathbf{E}\big(\Gamma_{i-1}^{\alpha-1} - (i-1)^{\alpha-1}\big)^{k_i} = \sum_{m=0}^{k_i} C_{k_i}^m (-1)^{k_i-m}(i-1)^{(k_i-m)(\alpha-1)}\mathbf{E}\Gamma_{i-1}^{(\alpha-1)m}$$
$$= \sum_{m=0}^{k_i}(-1)^{k_i-m}(i-1)^{(k_i-m)(\alpha-1)}C_{k_i}^m\big[(i-1)^{(\alpha-1)m} + O_m\big((i-1)^{(\alpha-1)m-1}\big)\big]$$
$$= (i-1)^{k_i(\alpha-1)}\left[\sum_{m=0}^{k_i} C_{k_i}^m(-1)^{k_i-m} + \sum_{m=0}^{k_i} C_{k_i}^m(-1)^{k_i-m}O_m\big((i-1)^{-1}\big)\right],$$

and further,

$$\left|\prod_{i=2}^{n}\mathbf{E}(\langle\varepsilon_i,e_j\rangle)^{k_i}\mathbf{E}(\gamma_i)^{k_i}\mathbf{E}\big(\Gamma_{i-1}^{\alpha-1}-(i-1)^{\alpha-1}\big)^{k_i}\right|\leqslant C(N,\alpha)\prod_{i=1}^{n}i^{k_i(\alpha-1)}$$
$$\leqslant C(N,\alpha)\prod_{i=1}^{n}n^{k_i(\alpha-1)}=C(N,\alpha)n^{2N(\alpha-1)};$$
$$\mathrm{II}\leqslant C(N,\alpha)\frac{\alpha^{2N}}{R^{2N}n^{2N\alpha-N}}n^N n^{2N(\alpha-1)-1}\leqslant C(N,\alpha)\frac{\alpha^{2N}}{R^{2N}}. \tag{4.2}$$

Next,

$$\mathrm{III}\leqslant\frac{\alpha^{2N}}{R^{2N}n^{2N\alpha-N}}\sum_{k_1+\cdots+k_n=2N}\frac{(2N)!}{k_1!\cdots k_n!}\prod_{i=2}^{n}\mathbf{E}(\langle\varepsilon_i,e_j\rangle)^{k_i}\mathbf{E}(\gamma_i)^{k_i}(i-1)^{k_i(\alpha-1)},$$

where

$$\left|\prod_{i=2}^{n}\mathbf{E}(\langle\varepsilon_i,e_j\rangle)^{k_i}\mathbf{E}(\gamma_i)^{k_i}(i-1)^{k_i(\alpha-1)}\right|\leqslant C(N,\alpha)n^{2N(\alpha-1)}.$$

We finally get

$$\mathrm{III}\leqslant C(N,\alpha)\frac{1}{R^{2N}n^{2N\alpha-N}}n^N n^{2N(\alpha-1)}=\frac{C(N,\alpha)}{R^{2N}}. \tag{4.3}$$

The multidimensional cases can be reduced to the one-dimensional case as follows:

$$\mathbf{P}\left(\max_{k=1,2,\dots,n}\left|\sum_{i=1}^{k}\varepsilon_i(\Gamma_i^\alpha-\Gamma_{i-1}^\alpha)\right|\geqslant B_nR\right)$$
$$=\mathbf{P}\left(\max_{k=1,2,\dots,n}\left|\sum_{j=1}^{d}\sum_{i=1}^{k}\langle\varepsilon_i,e_j\rangle e_j(\Gamma_i^\alpha-\Gamma_{i-1}^\alpha)\right|\geqslant B_nR\right)$$
$$\leqslant\mathbf{P}\left(\exists\, j^*,\,1\leqslant j^*\leqslant d,\ \max_{k=1,2,\dots,n}\left|\sum_{i=1}^{k}\langle\varepsilon_i,e_{j^*}\rangle e_{j^*}(\Gamma_i^\alpha-\Gamma_{i-1}^\alpha)\right|\geqslant\frac{B_nR}{d}\right)$$
$$\leqslant\sum_{j=1}^{d}\mathbf{P}\left(\max_{k=1,2,\dots,n}\sum_{i=1}^{k}|\langle\varepsilon_i,e_j\rangle|(\Gamma_i^\alpha-\Gamma_{i-1}^\alpha)\geqslant\frac{B_nR}{d}\right). \tag{4.4}$$

Let us verify the hypotheses of Theorem 6.9 in [1]. For this purpose, using the above estimates (4.1)–(4.4), we have

$$\int_{d(0,x)>R}\mathbf{d}^p(\mathbf{0},x)\,d\mu_n(x)=\sum_{i=0}^{\infty}\int_{(i+1)R\leqslant\mathbf{d}(\mathbf{0},x)<(i+2)R}\mathbf{d}^p(\mathbf{0},x)\,d\mu_n(x)$$
$$\leqslant R^p\sum_{i=0}^{\infty}(i+2)^p\cdot\mu_n\big(\mathbf{d}(\mathbf{0},x)\geqslant(i+1)R\big)$$
$$=R^p\sum_{i=0}^{\infty}(i+2)^p\cdot\mathbf{P}\Big(\max_{0\leqslant t\leqslant 1}|X_n(t)|\geqslant(i+1)R\Big)$$
$$=R^p\sum_{i=0}^{\infty}(i+2)^p\cdot\mathbf{P}\left(\max_{k=1,\dots,n}\left|\frac{1}{B_n}\sum_{i=1}^{k}\varepsilon_i(\Gamma_i^\alpha-\Gamma_{i-1}^\alpha)\right|\geqslant(i+1)R\right)$$
$$\leqslant R^p\sum_{i=0}^{\infty}(i+2)^p\cdot\sum_{j=1}^{d}\mathbf{P}\left(\max_{k=1,2,\dots,n}\sum_{i=1}^{k}|\langle\varepsilon_i,e_j\rangle|(\Gamma_i^\alpha-\Gamma_{i-1}^\alpha)\geqslant\frac{B_n(i+1)R}{d}\right)$$
$$\leqslant\frac{C(N,\alpha,d)}{R^{2N-p}}.$$

Thus, condition (iii) of Definition 1 holds for $\alpha > 1/2$,

$$\lim_{R\to\infty} \overline{\lim_{n\to\infty}} \int_{\mathbf{d}(\mathbf{0},x)\geqslant R} \mathbf{d}^p(\mathbf{0}, x)\, d\mu_n(x) = 0.$$

This completes the proof of the convergence in the Kantorovich metric in the cases $\alpha > 1/2$ and $p \in [1, \infty)$.

Acknowledgment. The authors are grateful to Y. Davydov for useful remarks and discussions that were helpful in the preparation of this paper.

REFERENCES

[1] C. Villani, *Optimal Transport, Old and New*, Grundlehren Math. Wiss. 338, Springer-Verlag, Berlin, 2009, https://doi.org/10.1007/978-3-540-71050-9.
[2] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.
[3] Y. Davydov and V. Konakov, *Random walks in nonhomogeneous Poisson environment*, in Modern Problems of Stochastic Analysis and Statistics, Selected Contributions in Honor of Valentin Konakov, Springer Proc. Math. Stat. 208, Springer, Cham, 2017, pp. 3–24, https://doi.org/10.1007/978-3-319-65313-6_1.
[4] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Grundlehren Math. Wiss. 293, Springer-Verlag, Berlin, 1999, https://doi.org/10.1007/978-3-662-06400-9.