Filomat 37:11 (2023), 3493–3512
https://doi.org/10.2298/FIL2311493B

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

A multidimensional stable limit theorem

Mátyás Barczy[a], Gyula Pap[b]

[a]*ELKH-SZTE Analysis and Applications Research Group, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary*
[b]*Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary*

Abstract. We establish multidimensional analogues of one-dimensional stable limit theorems due to Häusler and Luschgy (2015) for so called explosive processes. As special cases we present multidimensional stable limit theorems involving multidimensional normal-, Cauchy- and stable distributions as well.

1. Introduction and main results

Stable convergence and mixing convergence have been frequently used in limit theorems in probability theory and statistics. Historically the notion of mixing convergence was introduced first, and it can be traced back at least to Rényi [16], see also Rényi [17] and [19]. The more general concept of stable convergence is also due to Rényi [18]. Stable convergence should not be mistaken for weak convergence to a stable distribution. Recently, Häusler and Luschgy [11] have given an up to date and rigorous exposition of the mathematical theory of stable convergence, and they provided many applications in different areas to demonstrate the usefulness of this mode of convergence as well. In many classical limit theorems, such as in the classical central limit theorem, not only convergence in distribution, but stable convergence holds as well, see, e.g., Examples 3.13 and 3.14 in Häusler and Luschgy [11]. Stable convergence comes into play in the description of limit points of random sequences, in limit theorems with random indices, there is a version of the classical Δ-method with stable convergence as well, see, e.g., Chapter 4 in Häusler and Luschgy [11]. Stable convergence has a central role in limit theorems for martingale difference arrays, and one can find its nice applications in describing the asymptotic behaviour of some estimators (such as conditional least squares estimator) of some parameters of autoregressive and moving average processes and supercritical Galton–Watson processes (for a detailed description, see Chapters 9 and 10 in Häusler and Luschgy [11]). For a short survey on the role of stable convergence in limit theorems for semimartingales, see Podolskij and Vetter [15]. In numerical probability, especially, in studying discretized processes, in approximation of stochastic integrals and stochastic differential equations, and in high frequency statistics, stable convergence also plays an essential role, see the recent books Aït-Sahalia and Jacod [1] and Jacod

2020 *Mathematics Subject Classification*. Primary 60F05; Secondary 60B10.
Keywords. Stable convergence; Mixing convergence; Stable limit theorem; Multidimensional normal distribution; Multidimensional Cauchy distribution; Multidimensional stable distribution.
Received: 02 June 2022; Accepted: 22 December 2022
Communicated by Dora Seleši
Mátyás Barczy was supported by the project TKP2021-NVA-09. Project no. TKP2021-NVA-09 has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.
Email address: barczy@math.u-szeged.hu (Mátyás Barczy)

and Protter [13]. Very recently, Basse-O'Connor et al. [4, part (i) of Theorem 2.1 and part (i) of Proposition 2.3] have proved new limit theorems with stable convergence for some variational functionals of stationary increments Lévy driven moving averages in the high frequency setting.

Recently, Crimaldi et al. [9, Definition 3] have extended the notion of stable convergence: they have introduced the notion of stable convergence of random variables with respect to a so-called conditioning system towards a kernel, by replacing the single sub-σ-field appearing in the definition of (the original) stable convergence with a family of sub-σ-fields (called a conditioning system). Then, as a generalization of the previously mentioned concept, Crimaldi et al. [9, Definition 4] have introduced the notion of stable convergence of random variables in the strong sense with respect to a conditioning system, where not only the single sub-σ-field appearing in the definition of (the original) stable convergence is replaced by a conditioning system, but also the type of convergence for the conditional expectations with respect to the members of the conditional system in question is strengthened to convergence in L_1. Moreover, as a further generalization, Crimaldi [8, Definition 2.1] have defined the notion of almost sure conditional convergence of random variables with respect to a conditional system towards a kernel. If such a convergence holds, then the conditional expectations with respect to the members of the conditional system in question converge almost surely to a random variable.

Let $\mathbb{Z}_+$, $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the set of non-negative integers, positive integers, real numbers, non-negative real numbers and positive real numbers, respectively. The imaginary unit is denoted by i. The Borel σ-algebra on $\mathbb{R}^d$ is denoted by $\mathcal{B}(\mathbb{R}^d)$, where $d \in \mathbb{N}$. Further, let $\log^+(x) := \log(x)\mathbb{1}_{\{x\geqslant 1\}} + 0 \cdot \mathbb{1}_{\{0\leqslant x<1\}}$ for $x \in \mathbb{R}_+$. Convergence in a probability, in L_1, in L_2 and in distribution under a probability measure $\mathbb{P}$ will be denoted by $\xrightarrow{\mathbb{P}}$, $\xrightarrow{L_1(\mathbb{P})}$, $\xrightarrow{L_2(\mathbb{P})}$ and $\xrightarrow{\mathcal{D}(\mathbb{P})}$, respectively. For an event A with $\mathbb{P}(A) > 0$, let $\mathbb{P}_A(\cdot) := \mathbb{P}(\cdot \mid A) = \mathbb{P}(\cdot \cap A)/\mathbb{P}(A)$ denote the conditional probability measure given A. Let $\mathbb{E}_{\mathbb{P}}$ denote the expectation under a probability measure $\mathbb{P}$. Almost sure equality under a probability measure $\mathbb{P}$ and equality in distribution will be denoted by $\overset{\mathbb{P}\text{-a.s.}}{=}$ and $\overset{\mathcal{D}}{=}$, respectively. Every random variable will be defined on a (suitable) probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For a random variable $\xi : \Omega \to \mathbb{R}^d$, the distribution of ξ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is denoted by $\mathbb{P}^\xi$. The notions of stable and mixing convergence and some of their important properties used in the present paper are recalled in Appendix A.

First, we will recall a one-dimensional stable limit theorem due to Häusler and Luschgy [11, Theorem 8.2] for so called explosive processes. The increments of these processes are in general not asymptotically negligible and do not satisfy the conditional Lindeberg condition, so they are not in the scope of stable martingale central limit theorems. For such explosive processes, Häusler and Luschgy [11] developed the following limit theorem (Theorem 1.1) which states stable (mixing) convergence of the appropriately scaled explosive process in question, and they successfully applied it for proving stable (mixing) convergence of conditional least squares estimator of the autoregressive parameter of supercritical autoregressive processes of order 1 (see Häusler and Luschgy [11, Example 8.10 and Theorem 9.2]) and that of Lotka-Nagaev estimator, conditional least squares estimator and Harris estimator of the offspring mean of supercritical Galton-Watson branching processes conditionally on non-extinction (see Häusler and Luschgy [11, Corollaries 10.2, 10.4 and 10.6]).

Theorem 1.1 (Häusler and Luschgy [11, Theorem 8.2]) *Let $(X_n)_{n\in\mathbb{Z}_+}$ and $(A_n)_{n\in\mathbb{Z}_+}$ be real-valued stochastic processes defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and adapted to a filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}_+}$. Suppose that $A_n \geqslant 0$, $n \in \mathbb{N}$, and that there exists $n_0 \in \mathbb{N}$ such that $A_n > 0$ for each $n \geqslant n_0$. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $(0, \infty)$ such that $a_n \to \infty$ as $n \to \infty$, and let $G \in \mathcal{F}_\infty := \sigma(\bigcup_{n\in\mathbb{Z}_+} \mathcal{F}_n)$ such that $\mathbb{P}(G) > 0$. Assume that the following conditions are satisfied:*

(HLi) *there exists a non-negative, $\mathcal{F}_\infty$-measurable random variable $\eta : \Omega \to \mathbb{R}$ such that $\mathbb{P}(G \cap \{\eta^2 > 0\}) > 0$ and*

$$\frac{A_n}{a_n^2} \xrightarrow{\mathbb{P}_G} \eta^2 \qquad \text{as } n \to \infty,$$

(HLii) $(\frac{X_n}{a_n})_{n\in\mathbb{N}}$ *is stochastically bounded in* $\mathbb{P}_{G\cap\{\eta^2>0\}}$*-probability, i.e.,*

$$\lim_{K\to\infty}\sup_{n\in\mathbb{N}}\mathbb{P}_{G\cap\{\eta^2>0\}}\left(\frac{|X_n|}{a_n}>K\right)=0,$$

(HLiii) *there exists* $p\in(1,\infty)$ *such that*

$$\lim_{n\to\infty}\frac{a_{n-r}^2}{a_n^2}=\frac{1}{p^r}\qquad\text{for each } r\in\mathbb{N},$$

(HLiv) *there exists a probability measure* μ *on* $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ *with* $\int_{\mathbb{R}}\log^+(|x|)\,\mu(\mathrm{d}x)<\infty$ *such that*

$$\mathbb{E}_{\mathbb{P}}\left(\exp\left\{\mathrm{i}t\frac{\Delta X_n}{A_n^{1/2}}\right\}\,\Big|\,\mathcal{F}_{n-1}\right)\xrightarrow{\mathbb{P}_{G\cap\{\eta^2>0\}}}\int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}\mu(x)\qquad\text{as } n\to\infty$$

for all $t\in\mathbb{R}$, *where* $\Delta X_n:=X_n-X_{n-1}$, $n\in\mathbb{N}$, *and* $\Delta X_0:=0$.

Then

$$\frac{X_n}{A_n^{1/2}}\to\sum_{j=0}^{\infty}p^{-j/2}Z_j\qquad\mathcal{F}_\infty\text{-mixing under }\mathbb{P}_{G\cap\{\eta^2>0\}}\text{ as } n\to\infty,\tag{1}$$

and

$$\frac{X_n}{a_n}\to\eta\sum_{j=0}^{\infty}p^{-j/2}Z_j\qquad\mathcal{F}_\infty\text{-stably under }\mathbb{P}_{G\cap\{\eta^2>0\}}\text{ as } n\to\infty,\tag{2}$$

where $(Z_j)_{j\in\mathbb{Z}_+}$ *denotes a* $\mathbb{P}$*-independent and identically distributed sequence of real-valued random variables being* $\mathbb{P}$*-independent of* $\mathcal{F}_\infty$ *with* $\mathbb{P}(Z_0\in B)=\mu(B)$ *for all* $B\in\mathcal{B}(\mathbb{R})$.

Remark 1.2 *(i) The series* $\sum_{j=0}^{\infty}p^{-j/2}Z_j=\sum_{j=0}^{\infty}(p^{1/2})^{-j}Z_j$ *in* (1) *and* (2) *is absolutely convergent* $\mathbb{P}$*-almost surely, since* $p^{1/2}>1$, $\mathbb{E}_{\mathbb{P}}(\log^+(|Z_0|))<\infty$ *(by condition (HLiv) of Theorem 1.1), and one can apply Lemma 8.1 in Häusler and Luschgy [11].*

(ii) We note that in condition (HLi) of Theorem 1.1 the $\mathcal{F}_\infty$*-measurability of* η *is supposed, but in condition (i) of Theorem 8.2 in Häusler and Luschgy [11] it is not supposed. However, in the proof of Theorem 8.2 in Häusler and Luschgy [11, page 148] it is written that the* $\mathcal{F}_\infty$*-measurability of* η *can be assumed without loss of generality. Note also that if the probability space* $(\Omega,\mathcal{F}_\infty,\mathbb{P}_G)$ *is complete, then the* $\mathcal{F}_\infty$*-measurability of* η *follows itself from the convergence* $\frac{A_n}{a_n^2}\xrightarrow{\mathbb{P}_G}\eta^2$ *as* $n\to\infty$ *involved in condition (HLi) of Theorem 1.1. Indeed, then there exists a subsequence* $(n_k)_{k\in\mathbb{N}}$ *such that* $A_{n_k}/a_{n_k}^2\xrightarrow{\mathbb{P}_G\text{-a.s.}}\eta^2$ *as* $k\to\infty$. *Since* $A_{n_k}/a_{n_k}^2$ *is* $\mathcal{F}_\infty$*-measurable for each* $k\in\mathbb{N}$ *and* $(\Omega,\mathcal{F}_\infty,\mathbb{P}_G)$ *is complete, by a standard measure theoretical argument, we have* η^2 *is* $\mathcal{F}_\infty$*-measurable. The continuity of the square-root function together with* $\eta\geqslant 0$ *yield the* $\mathcal{F}_\infty$*-measurability of* η, *as desired.*

(iii) The $\mathcal{F}_\infty$*-measurability of* η *yields that* η *and* Z_j, $j\in\mathbb{N}$, *are* $\mathbb{P}$*-independent in Theorem 1.1. Further, we have* $\mathbb{P}_G(\eta>0)=\mathbb{P}_G(\eta^2>0)>0$ *and* $\mathbb{P}_{G\cap\{\eta^2>0\}}(\eta>0)=1$, *where we used that* η *is non-negative.* □

By $\|\boldsymbol{x}\|$ and $\|\boldsymbol{A}\|$, we denote the Euclidean norm of a vector $\boldsymbol{x}\in\mathbb{R}^d$ and the induced matrix norm of a matrix $\boldsymbol{A}\in\mathbb{R}^{d\times d}$, respectively. By $\langle\boldsymbol{x},\boldsymbol{y}\rangle$, we denote the Euclidean inner product of vectors $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$. The null vector and the null matrix will be denoted by $\boldsymbol{0}$. By $\varrho(\boldsymbol{A})$, we denote the spectral radius of $\boldsymbol{A}\in\mathbb{R}^{d\times d}$. Moreover, $\boldsymbol{I}_d\in\mathbb{R}^{d\times d}$ denotes the $d\times d$ identity matrix, and if $\boldsymbol{A}\in\mathbb{R}^{d\times d}$ is symmetric and positive semidefinite, then $\boldsymbol{A}^{1/2}$ denotes the unique symmetric, positive semidefinite square root of $\boldsymbol{A}$. If $\boldsymbol{V}\in\mathbb{R}^{d\times d}$ is symmetric and positive semidefinite, then $\mathcal{N}_d(\boldsymbol{0},\boldsymbol{V})$ denotes the d-dimensional normal distribution with mean vector $\boldsymbol{0}\in\mathbb{R}^d$ and covariance matrix $\boldsymbol{V}$.

In order to formulate our multidimensional stable limit theorems, we need the following result, which is a multidimensional generalization of Lemma 8.1 in Häusler and Luschgy [11], and it is interesting on its own right.

Lemma 1.3 *Let $(\boldsymbol{Z}_j)_{j\in\mathbb{Z}_+}$ be a $\mathbb{P}$-independent and identically distributed sequence of $\mathbb{R}^d$-valued random vectors. Let $\boldsymbol{P} \in \mathbb{R}^{d\times d}$ be an invertible matrix with $\varrho(\boldsymbol{P}) < 1$. Then the following assertions are equivalent:*

(i) $\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{Z}_0\|)) < \infty$.

(ii) $\sum_{j=0}^{\infty} \|\boldsymbol{P}^j \boldsymbol{Z}_j\| < \infty$ *$\mathbb{P}$-almost surely.*

(iii) $\sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j$ *converges $\mathbb{P}$-almost surely in $\mathbb{R}^d$.*

(iv) $\boldsymbol{P}^j \boldsymbol{Z}_j \to \boldsymbol{0}$ *as $j \to \infty$ $\mathbb{P}$-almost surely.*

The proof of Lemma 1.3 and the proofs of all the forthcoming results can be found in Section 2. We note that from the proof of Lemma 1.3 it turns out that for the implications (i) ⇒ (ii) ⇒ (iii) ⇒ (iv), we do not need the invertibility of $\boldsymbol{P}$, we only need it for (iv) ⇒ (i).

For an $\mathbb{R}^d$-valued stochastic process $(\boldsymbol{U}_n)_{n\in\mathbb{Z}_+}$, the increments $\Delta\boldsymbol{U}_n$, $n \in \mathbb{Z}_+$, are defined by $\Delta\boldsymbol{U}_0 := \boldsymbol{0}$ and $\Delta\boldsymbol{U}_n := \boldsymbol{U}_n - \boldsymbol{U}_{n-1}$ for $n \in \mathbb{N}$.

Our main result is the following multidimensional analogue of Theorem 8.2 in Häusler and Luschgy [11] (see also Theorem 1.1).

Theorem 1.4 *Let $(\boldsymbol{U}_n)_{n\in\mathbb{Z}_+}$ and $(\boldsymbol{B}_n)_{n\in\mathbb{Z}_+}$ be $\mathbb{R}^d$-valued and $\mathbb{R}^{d\times d}$-valued stochastic processes, respectively, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and adapted to a filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}_+}$. Suppose that $\boldsymbol{B}_n$ is invertible for sufficiently large $n \in \mathbb{N}$. Let $(\boldsymbol{Q}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}^{d\times d}$ such that $\boldsymbol{Q}_n \to \boldsymbol{0}$ as $n \to \infty$ and $\boldsymbol{Q}_n$ is invertible for sufficiently large $n \in \mathbb{N}$. Let $G \in \mathcal{F}_\infty := \sigma(\bigcup_{k=0}^{\infty} \mathcal{F}_k)$ with $\mathbb{P}(G) > 0$. Assume that the following conditions are satisfied:*

(i) *there exists an $\mathbb{R}^{d\times d}$-valued, $\mathcal{F}_\infty$-measurable random matrix $\boldsymbol{\eta} : \Omega \to \mathbb{R}^{d\times d}$ such that $\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\}) > 0$ and*

$$\boldsymbol{Q}_n \boldsymbol{B}_n^{-1} \xrightarrow{\mathbb{P}_G} \boldsymbol{\eta} \qquad \text{as } n \to \infty,$$

(ii) *$(\boldsymbol{Q}_n \boldsymbol{U}_n)_{n\in\mathbb{N}}$ is stochastically bounded in $\mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}}$-probability, i.e.,*

$$\lim_{K\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}}(\|\boldsymbol{Q}_n \boldsymbol{U}_n\| > K) = 0,$$

(iii) *there exists an invertible matrix $\boldsymbol{P} \in \mathbb{R}^{d\times d}$ with $\varrho(\boldsymbol{P}) < 1$ such that*

$$\boldsymbol{B}_n \boldsymbol{B}_{n-r}^{-1} \xrightarrow{\mathbb{P}_G} \boldsymbol{P}^r \qquad \text{as } n \to \infty \text{ for each } r \in \mathbb{N},$$

(iv) *there exists a probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $\int_{\mathbb{R}^d} \log^+(\|\boldsymbol{x}\|)\, \mu(\mathrm{d}\boldsymbol{x}) < \infty$ such that*

$$\mathbb{E}_{\mathbb{P}}\Big(\mathrm{e}^{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{B}_n \Delta\boldsymbol{U}_n\rangle} \,|\, \mathcal{F}_{n-1}\Big) \xrightarrow{\mathbb{P}_{G\cap\{\exists\boldsymbol{\eta}^{-1}\}}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{x}\rangle}\, \mu(\mathrm{d}\boldsymbol{x}) \qquad \text{as } n \to \infty$$

for all $\boldsymbol{\theta} \in \mathbb{R}^d$.

Then

$$\boldsymbol{B}_n \boldsymbol{U}_n \to \sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-mixing under } \mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{3}$$

and

$$\boldsymbol{Q}_n \boldsymbol{U}_n \to \boldsymbol{\eta} \sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{4}$$

where $(\boldsymbol{Z}_j)_{j\in\mathbb{Z}_+}$ denotes a $\mathbb{P}$-independent and identically distributed sequence of $\mathbb{R}^d$-valued random vectors being $\mathbb{P}$-independent of $\mathcal{F}_\infty$ with $\mathbb{P}(\boldsymbol{Z}_0 \in B) = \mu(B)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$.

Remark 1.5 *(i) The series* $\sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j$ *in* (3) *and in* (4) *is absolutely convergent* $\mathbb{P}$*-almost surely, since* $\boldsymbol{P}$ *is invertible,* $\varrho(\boldsymbol{P}) < 1$, $\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{Z}_0\|)) < \infty$ *(by condition (iv) of Theorem 1.4), and one can apply Lemma 1.3.*

(ii) The random variable $\boldsymbol{\eta}$ *and the sequence* $(\boldsymbol{Z}_j)_{j\in\mathbb{Z}_+}$ *are* $\mathbb{P}$*-independent in Theorem 1.4, since* $\boldsymbol{\eta}$ *is* $\mathcal{F}_\infty$*-measurable and the sequence* $(\boldsymbol{Z}_j)_{j\in\mathbb{Z}_+}$ *is* $\mathbb{P}$*-independent of* $\mathcal{F}_\infty$*. Further, we have* $\mathbb{P}_G(\exists\,\boldsymbol{\eta}^{-1}) > 0$ *and* $\mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}}(\exists\,\boldsymbol{\eta}^{-1}) = 1$.

(iii) The proof of Theorem 1.4 (which can be found in Section 2) follows the method of that of Theorem 8.2 in Häusler and Luschgy [11]. However, a natural question also occurs, namely, would it be possible to prove Theorem 1.4 using the Cramér-Wold theorem for stable convergence (see, e.g., Häusler and Luschgy [11, Corollary 3.19])? We do not know the answer to this question. The Cramér-Wold theorem for stable convergence states that, given $\mathbb{R}^d$*-valued random variables* $\boldsymbol{X}_n$, $n \in \mathbb{N}$, *and* $\boldsymbol{X}$, $\boldsymbol{X}_n$ *converges* $\mathcal{G}$*-stably to* $\boldsymbol{X}$ *as* $n \to \infty$ *if and only if for all* $\boldsymbol{u} \in \mathbb{R}^d$, *the real-valued random variables* $\langle \boldsymbol{u}, \boldsymbol{X}_n\rangle$ *converges* $\mathcal{G}$*-stably to the real-valued random variable* $\langle \boldsymbol{u}, \boldsymbol{X}\rangle$ *as* $n \to \infty$ *(where we used the setup given in Definition A.1). Here we only note that even in the proofs of multivariate central limit theorems with scaling matrices not converging to a fixed positive definite matrix, not only the Cramér-Wold theorem (for convergence in distribution) comes into play, but a key lemma originated to Bolthausan [7] and its generalization due to Biscio et al. [6, Lemma 3.2], for more details see Biscio et al. [6].* □

In the next remark we reformulate condition (iii) of Theorem 1.4 in the one-dimensional case.

Remark 1.6 *In case of* $d = 1$ *(so not using boldface style in this case), if condition (i) of Theorem 1.4 and* $\mathbb{P}(\exists\,\eta^{-1}) = \mathbb{P}(\eta \neq 0) = 1$ *hold, then condition (iii) of Theorem 1.4 is equivalent to the following condition:*

$$\text{there exists } P \in (-1,1)\setminus\{0\} \text{ such that } Q_n Q_{n-r}^{-1} \to P^r \text{ as } n\to\infty \text{ for each } r\in\mathbb{N}. \tag{5}$$

Indeed, if conditions (i) and (iii) of Theorem 1.4 with $d = 1$ *and* $\mathbb{P}(\exists\,\eta^{-1}) = 1$ *hold, then there exists* $P \in (-1,1)\setminus\{0\}$ *such that for each* $r \in \mathbb{N}$, *we have*

$$Q_n Q_{n-r}^{-1} = Q_n B_n^{-1} B_n B_{n-r}^{-1} B_{n-r} Q_{n-r}^{-1} \xrightarrow{\mathbb{P}_G} \eta P^r \eta^{-1} = P^r \qquad \text{as } n \to \infty.$$

Since $Q_n Q_{n-r}^{-1}$ *is non-random, we have* (5)*. Conversely, if condition (i) of Theorem 1.4 with* $d = 1$, $\mathbb{P}(\exists\,\eta^{-1}) = \mathbb{P}(\eta \neq 0) = 1$, *and* (5) *hold, then there exists* $P \in (-1,1)\setminus\{0\}$ *such that for each* $r \in \mathbb{N}$, *we have*

$$B_n B_{n-r}^{-1} = B_n Q_n^{-1} Q_n Q_{n-r}^{-1} Q_{n-r} B_{n-r}^{-1} \xrightarrow{\mathbb{P}_G} \eta^{-1} P^r \eta = P^r \qquad \text{as } n \to \infty,$$

i.e., condition (iii) of Theorem 1.4 with $d = 1$ *holds. Finally, note that, with the notation* $a_n := Q_n^{-1}$, *condition* (5) *implies that for each* $r \in \mathbb{N}$ *we have*

$$\frac{a_{n-r}^2}{a_n^2} = Q_n^2 Q_{n-r}^{-2} \to P^{2r} = \frac{1}{((P^2)^{-1})^r} \qquad \text{as } n \to \infty,$$

which is nothing else but condition (iii) of Theorem 8.2 in Häusler and Luschgy [11] (see also condition (HLiii) of Theorem 1.1) with $p := (P^2)^{-1} \in (1,\infty)$*. In Remark 1.7, we give a more detailed comparison of Theorem 8.2 in Häusler and Luschgy [11] (see also Theorem 1.1) and Theorem 1.4.* □

In the next remark we investigate the connection between Theorem 8.2 in Häusler and Luschgy [11] (see also Theorem 1.1) and Theorem 1.4.

Remark 1.7 *Theorem 1.4 gives back Theorem 8.2 in Häusler and Luschgy [11] (see also Theorem 1.1) provided that* $\mathbb{P}(\eta > 0) = 1$ *in condition (i) of Theorem 8.2 in Häusler and Luschgy [11]. Indeed, let* $(X_n)_{n\in\mathbb{Z}_+}$ *and* $(A_n)_{n\in\mathbb{Z}_+}$ *be real-valued stochastic processes defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and adapted to a filtration* $(\mathcal{F}_n)_{n\in\mathbb{Z}_+}$*. Suppose that* $A_n \geqslant 0$, $n \in \mathbb{N}$, *and that there exists* $n_0 \in \mathbb{N}$ *such that* $A_n > 0$ *for each* $n \geqslant n_0$*. Let* $(a_n)_{n\in\mathbb{N}}$ *be a sequence in* $(0,\infty)$ *such that* $a_n \to \infty$ *as* $n \to \infty$, *and let* $G \in \mathcal{F}_\infty$ *with* $\mathbb{P}(G) > 0$ *such that the conditions (HLi) together with* $\mathbb{P}(\eta > 0) = 1$, *(HLii), (HLiii) and (HLiv) of Theorem 1.1 hold. Note that in this case* $\mathbb{P}_{G\cap\{\eta^2>0\}} = \mathbb{P}_G$, *since* $\mathbb{P}(\eta > 0) = 1$ *implies that* $\mathbb{P}(\eta^2 > 0) = 1$*. In Theorem 1.4, let us make the following choices* $\boldsymbol{U}_n := X_n$,

$n \in \mathbb{Z}_+$, $\boldsymbol{B}_n := A_n^{-1/2}$, $n \geqslant n_0$, $\boldsymbol{Q}_n := a_n^{-1}$, $n \in \mathbb{N}$, and $\boldsymbol{P} := p^{-1/2}$, where $p \in (1,\infty)$ is given in (HLiii) of Theorem 1.1. Then (HLi) of Theorem 1.1, the non-negativity of η and the continuity of the square-root function yield that $\boldsymbol{Q}_n\boldsymbol{B}_n^{-1} = \frac{A_n^{1/2}}{a_n} \xrightarrow{\mathbb{P}_G} \eta$ as $n \to \infty$, i.e., condition (i) of Theorem 1.4 is satisfied. Further, (HLi) of Theorem 1.1 together with $\mathbb{P}(\eta > 0) = 1$, (HLiii) of Theorem 1.1 and the continuity of the square-root function imply that for each $r \in \mathbb{N}$, we have

$$\boldsymbol{B}_n\boldsymbol{B}_{n-r}^{-1} = \frac{A_{n-r}^{1/2}}{A_n^{1/2}} = \frac{A_{n-r}^{1/2}}{a_{n-r}}\frac{a_n}{A_n^{1/2}}\frac{a_{n-r}}{a_n} \xrightarrow{\mathbb{P}_G} \eta \cdot \frac{1}{\eta} \cdot \frac{1}{p^{r/2}} = \boldsymbol{P}^r \qquad \text{as } n \to \infty,$$

i.e., condition (iii) of Theorem 1.4 holds. Conditions (HLii) and (HLiv) of Theorem 1.1 readily yield conditions (ii) and (iv) of Theorem 1.4, respectively. So we can apply Theorem 1.4 and we have (1) *and* (2)*, as desired.* □

Next, we present a multidimensional stable central limit theorem, which is a multidimensional counterpart of Corollary 8.5 in Häusler and Luschgy [11].

Corollary 1.8 *Let us assume that the conditions of Theorem 1.4 hold with $\mu := \mathbb{P}^{\mathcal{N}_d(\boldsymbol{0},\boldsymbol{D})}$, where $\mathbb{P}^{\mathcal{N}_d(\boldsymbol{0},\boldsymbol{D})}$ denotes the distribution of a d-dimensional normally distributed random variable with mean vector $\boldsymbol{0} \in \mathbb{R}^d$ and covariance matrix $\boldsymbol{D} \in \mathbb{R}^{d\times d}$. Then*

$$\boldsymbol{B}_n\boldsymbol{U}_n \to \boldsymbol{Z} \qquad \mathcal{F}_\infty\text{-mixing under } \mathbb{P}_{G\cap\{\exists\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{6}$$

and

$$\boldsymbol{Q}_n\boldsymbol{U}_n \to \boldsymbol{\eta}\boldsymbol{Z} \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{P}_{G\cap\{\exists\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{7}$$

where $\boldsymbol{Z}$ denotes a d-dimensional normally distributed random vector with mean vector $\boldsymbol{0} \in \mathbb{R}^d$ and covariance matrix $\sum_{j=0}^\infty \boldsymbol{P}^j\boldsymbol{D}(\boldsymbol{P}^j)^\top$, and $\boldsymbol{Z}$ is $\mathbb{P}$-independent of $\mathcal{F}_\infty$.

In Corollary 1.8, $\boldsymbol{\eta}$ and $\boldsymbol{Z}$ are $\mathbb{P}$-independent, since $\boldsymbol{\eta}$ is $\mathcal{F}_\infty$-measurable (supposed in condition (i) of Theorem 1.4).

Next, we will formulate a corollary of Theorem 1.4 involving multidimensional stable distributions, in particular, a multidimensional Cauchy distribution. For this, first we recall the notion of a multidimensional stable distribution. A d-dimensional random variable $\boldsymbol{\zeta} := (\zeta_1,\ldots,\zeta_d)$ is said to be stable if for any $a_1, a_2 \in \mathbb{R}_{++}$ there exist $b \in \mathbb{R}_{++}$ and $\boldsymbol{c} \in \mathbb{R}^d$ such that

$$a_1\boldsymbol{\zeta}^{(1)} + a_2\boldsymbol{\zeta}^{(2)} \stackrel{\mathcal{D}}{=} b\boldsymbol{\zeta} + \boldsymbol{c}, \tag{8}$$

where $\boldsymbol{\zeta}^{(1)}$ and $\boldsymbol{\zeta}^{(2)}$ are independent copies of $\boldsymbol{\zeta}$. It is known that $\boldsymbol{\zeta}$ is stable if and only if there exists $\alpha \in (0,2]$ such that for each $n \geqslant 2$, $n \in \mathbb{N}$ there exists $\boldsymbol{c}_n \in \mathbb{R}^d$ satisfying $\boldsymbol{\zeta}^{(1)} + \cdots + \boldsymbol{\zeta}^{(n)} \stackrel{\mathcal{D}}{=} n^{\frac{1}{\alpha}}\boldsymbol{\zeta} + \boldsymbol{c}_n$, where $\boldsymbol{\zeta}^{(1)}, \boldsymbol{\zeta}^{(2)}, \ldots, \boldsymbol{\zeta}^{(n)}$ are independent copies of $\boldsymbol{\zeta}$. The index α is called the index of stability or the characteristic exponent of $\boldsymbol{\zeta}$. In what follows, let $S_{d-1} := \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\| = 1\}$ be the unit surface in $\mathbb{R}^d$. We say that $\boldsymbol{\zeta}$ is symmetric stable if it is stable and $\mathbb{P}(\boldsymbol{\zeta} \in A) = \mathbb{P}(-\boldsymbol{\zeta} \in A)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$. It known that a d-dimensional random variable $\boldsymbol{\zeta}$ is symmetric α-stable with index $\alpha \in (0,2)$ if and only if there exists a unique symmetric finite measure Π on $(S_{d-1}, \mathcal{B}(S_{d-1}))$ (where the property symmetric means that $\Pi(A) = \Pi(-A)$ for any $A \in \mathcal{B}(S_{d-1})$) such that

$$\mathbb{E}_{\mathbb{P}}\Big(\exp(\mathrm{i}\langle\boldsymbol{\theta},\boldsymbol{\zeta}\rangle)\Big) = \exp\left\{-\int_{S_{d-1}} \big|\langle\boldsymbol{\theta},\boldsymbol{x}\rangle\big|^\alpha\,\Pi(\mathrm{d}\boldsymbol{x})\right\}, \qquad \boldsymbol{\theta} \in \mathbb{R}^d,$$

see, e.g., Sato [20, Theorem 14.13]. The measure Π is called the spectral measure of $\boldsymbol{\zeta}$. We say that a d-dimensional random variable $\boldsymbol{\zeta}$ has a d-dimensional Cauchy distribution with parameter $(\boldsymbol{0}, \boldsymbol{I}_d)$, if its density function takes the form

$$f_{\boldsymbol{\zeta}}(\boldsymbol{x}) = \frac{\Gamma\left(\frac{1+d}{2}\right)}{\pi^{\frac{1+d}{2}}}\Big(1 + \|\boldsymbol{x}\|^2\Big)^{-\frac{1+d}{2}}, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

see, e.g., Kotz and Nadarajah [14, Section 2.2, page 41] or Sato [20, Example 2.12]. It is known that if $\boldsymbol{\zeta}$ has a d-dimensional Cauchy distribution with parameter $(\mathbf{0}, \boldsymbol{I}_d)$, then the characteristic function of $\boldsymbol{\zeta}$ takes the form $\mathbb{E}_{\mathbb{P}}(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta},\boldsymbol{\zeta}\rangle}) = \mathrm{e}^{-\|\boldsymbol{\theta}\|}$, $\boldsymbol{\theta} \in \mathbb{R}^d$, and $\boldsymbol{\zeta}$ is symmetric 1-stable, see, e.g., Sato [20, Theorem 14.14].

Corollary 1.9 *Let us assume that the conditions of Theorem 1.4 hold with $\mu := \mathbb{P}^{\boldsymbol{\zeta}}$, where $\boldsymbol{\zeta}$ is a d-dimensional symmetric α-stable random variable with characteristic exponent $\alpha \in (0,2)$ and spectral measure Π. Then*

$$\boldsymbol{B}_n \boldsymbol{U}_n \to \boldsymbol{Z} \qquad \mathcal{F}_\infty\text{-mixing under } \mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{9}$$

and

$$\boldsymbol{Q}_n \boldsymbol{U}_n \to \boldsymbol{\eta}\boldsymbol{Z} \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{10}$$

where $\boldsymbol{Z}$ denotes a d-dimensional random vector $\mathbb{P}$-independent of $\mathcal{F}_\infty$ with a characteristic function

$$\mathbb{E}_{\mathbb{P}}(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta},\boldsymbol{Z}\rangle}) = \exp\left\{-\int_{S_{d-1}} \sum_{j=0}^{\infty} \left|\langle(\boldsymbol{P}^j)^\top\boldsymbol{\theta}, \boldsymbol{x}\rangle\right|^\alpha \Pi(\mathrm{d}\boldsymbol{x})\right\}, \qquad \boldsymbol{\theta} \in \mathbb{R}^d. \tag{11}$$

In particular, if $\boldsymbol{\zeta}$ has a d-dimensional Cauchy distribution with parameter $(\mathbf{0}, \boldsymbol{I}_d)$, then $\boldsymbol{Z}$ has a characteristic function

$$\mathbb{E}_{\mathbb{P}}(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta},\boldsymbol{Z}\rangle}) = \exp\left\{-\sum_{j=0}^{\infty} \|(\boldsymbol{P}^j)^\top\boldsymbol{\theta}\|\right\}, \qquad \boldsymbol{\theta} \in \mathbb{R}^d. \tag{12}$$

In Corollary 1.9, $\boldsymbol{\eta}$ and $\boldsymbol{Z}$ are $\mathbb{P}$-independent, since $\boldsymbol{\eta}$ is $\mathcal{F}_\infty$-measurable. Corollary 1.9 in the special case when $\boldsymbol{\zeta}$ has a d-dimensional Cauchy distribution with parameter $(\mathbf{0}, \boldsymbol{I}_d)$ can be considered as a multidimensional counterpart of Exercise 8.1 in Häusler and Luschgy [11].

Finally, we formulate a slight generalization of Theorem 1.4 in case of $G = \Omega$, by weakening its condition (iv) a little bit. This generalization can be considered as a multidimensional analogue of Corollary 8.8 in Häusler and Luschgy [11].

Corollary 1.10 *Let us suppose that the conditions of Theorem 1.4 are satisfied with $G := \Omega$ except its condition (iv) which is replaced by*

(iv′) there exists a probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $\int_{\mathbb{R}^d} \log^+(\|\boldsymbol{x}\|)\,\mu(\mathrm{d}\boldsymbol{x}) < \infty$, and an $\mathcal{F}_\infty$-measurable, $\mathbb{R}^{d\times d}$-valued discrete random variable $\boldsymbol{S}$ such that

$$\mathbb{E}_{\mathbb{P}}\Big(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta},\boldsymbol{B}_n\Delta\boldsymbol{U}_n\rangle} \,|\, \mathcal{F}_{n-1}\Big) \xrightarrow{\mathbb{P}_{\{\exists\eta^{-1}\}}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta},\boldsymbol{S}\boldsymbol{x}\rangle}\,\mu(\mathrm{d}\boldsymbol{x}) \text{ as } n \to \infty \text{ for all } \boldsymbol{\theta} \in \mathbb{R}^d.$$

Then

$$\boldsymbol{B}_n \boldsymbol{U}_n \to \sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{S} \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{P}_{\{\exists\,\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{13}$$

and

$$\boldsymbol{Q}_n \boldsymbol{U}_n \to \boldsymbol{\eta} \sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{S} \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{P}_{\{\exists\,\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty, \tag{14}$$

where $(\boldsymbol{Z}_j)_{j\in\mathbb{Z}_+}$ denotes a $\mathbb{P}$-independent and identically distributed sequence of $\mathbb{R}^d$-valued random vectors $\mathbb{P}$-independent of $\mathcal{F}_\infty$ with $\mathbb{P}(\boldsymbol{Z}_0 \in B) = \mu(B)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$.

In Corollary 1.10, $\boldsymbol{\eta}$ and $(\boldsymbol{Z}_j)_{j\in\mathbb{Z}_+}$ are $\mathbb{P}$-independent (see part (ii) of Remark 1.5). For an application of Corollary 1.10 with $d = 1$, see the proof of Theorem 9.1 in Häusler and Luschgy [11], where the authors prove stable convergence of conditional least squares estimator of the autoregressive parameter of supercritical autoregressive processes of order 1.

Finally, we note that in a companion paper Barczy and Pap [3], we use our main result Theorem 1.4 for studying the asymptotic behaviour of least squares estimator of the autoregressive parameters of some supercritical Gaussian autoregressive processes of order 2 using random scaling. In another companion paper Barczy [2], we also use Theorem 1.4 for proving stable convergence of conditional least squares estimators of drift parameters for supercritical continuous state and continuous time branching processes with immigration based on discrete time observations.

2. Proofs

Proof of Lemma 1.3. (i) $\Rightarrow$ (ii). We have $\varrho(\boldsymbol{P}) = \lim_{k\to\infty} \|\boldsymbol{P}^k\|^{1/k}$ by the Gelfand formula, see, e.g., Horn and Johnson [12, Corollary 5.6.14]. Hence there exists $k_0 \in \mathbb{N}$ such that

$$\|\boldsymbol{P}^k\|^{1/k} \leqslant \varrho(\boldsymbol{P}) + \frac{1-\varrho(\boldsymbol{P})}{2} = \frac{1+\varrho(\boldsymbol{P})}{2} < 1 \qquad \text{for each } k \geqslant k_0, \tag{15}$$

since $\varrho(\boldsymbol{P}) < 1$. Choose $c \in \left(1, \frac{2}{1+\varrho(\boldsymbol{P})}\right)$. Then (i) implies

$$\begin{aligned}\sum_{j=k_0}^{\infty} \mathbb{P}(\|\boldsymbol{Z}_j\| > c^j) &= \sum_{j=k_0}^{\infty} \mathbb{P}(\|\boldsymbol{Z}_0\| > c^j) = \sum_{j=k_0}^{\infty} \mathbb{P}(\log^+(\|\boldsymbol{Z}_0\|) > j\log^+(c)) \\ &= \sum_{j=k_0}^{\infty} \mathbb{P}\left(\frac{\log^+(\|\boldsymbol{Z}_0\|)}{\log(c)} > j\right) < \infty,\end{aligned}$$

where we used that $\log^+(c) = \log(c) > 0$ and $\sum_{n=1}^{\infty} \mathbb{P}(\xi \geqslant n) \leqslant \mathbb{E}_{\mathbb{P}}(\xi)$ for any non-negative random variable ξ. By the Borel–Cantelli lemma,

$$\mathbb{P}\Big(\limsup_{j\to\infty}\{\|\boldsymbol{Z}_j\| > c^j\}\Big) = 0, \qquad \text{and hence} \qquad \mathbb{P}\Big(\liminf_{j\to\infty}\{\|\boldsymbol{Z}_j\| \leqslant c^j\}\Big) = 1,$$

i.e., for $\mathbb{P}$-a.a. $\omega \in \Omega$, there exists $j_0(\omega) \in \mathbb{N}$ such that $\|\boldsymbol{Z}_j(\omega)\| \leqslant c^j$ for each $j \geqslant j_0(\omega)$. Consequently, for $\mathbb{P}$-a.a. $\omega \in \Omega$, we have

$$\sum_{j=k_0\vee j_0(\omega)}^{\infty} \|\boldsymbol{P}^j \boldsymbol{Z}_j(\omega)\| \leqslant \sum_{j=k_0\vee j_0(\omega)}^{\infty} \|\boldsymbol{P}^j\| \cdot \|\boldsymbol{Z}_j(\omega)\| \leqslant \sum_{j=k_0}^{\infty} \left(\frac{1+\varrho(\boldsymbol{P})}{2}\right)^j c^j < \infty,$$

since $\frac{1+\varrho(\boldsymbol{P})}{2} c \in (0,1)$. It yields (ii).

The implications (ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (iv) are obvious.

(iv) $\Rightarrow$ (i). We have $\mathbb{P}(\limsup_{j\to\infty}\{\|\boldsymbol{P}^j\boldsymbol{Z}_j\| > 1\}) = 0$, and hence, by the Borel–Cantelli lemma and the independence of $(\boldsymbol{Z}_j)_{j\in\mathbb{Z}_+}$, we get

$$\sum_{j=0}^{\infty} \mathbb{P}(\|\boldsymbol{P}^j \boldsymbol{Z}_j\| > 1) < \infty.$$

Using that the determinant of $\boldsymbol{P}$ coincides with the product of its eigenvalues, the invertibility of $\boldsymbol{P}$ implies that $\boldsymbol{P}$ does not have an eigenvalue 0, and, in particular, we get $\varrho(\boldsymbol{P}) > 0$. The eigenvalues of $\boldsymbol{P}^{-1}$ are the reciprocals of the eigenvalues of $\boldsymbol{P}$, hence $\varrho(\boldsymbol{P}^{-1}) \geqslant \frac{1}{\varrho(\boldsymbol{P})}$, implying $\|\boldsymbol{P}^{-1}\| \geqslant \varrho(\boldsymbol{P}^{-1}) \geqslant \frac{1}{\varrho(\boldsymbol{P})} > 1$. Thus for each $j \in \mathbb{Z}_+$, we have

$$\begin{aligned}\mathbb{P}(\|\boldsymbol{P}^j \boldsymbol{Z}_j\| > 1) &= \mathbb{P}(\|\boldsymbol{P}^{-1}\|^j \|\boldsymbol{P}^j \boldsymbol{Z}_0\| > \|\boldsymbol{P}^{-1}\|^j) \geqslant \mathbb{P}(\|(\boldsymbol{P}^{-1})^j \boldsymbol{P}^j \boldsymbol{Z}_0\| > \|\boldsymbol{P}^{-1}\|^j) \\ &= \mathbb{P}(\|\boldsymbol{Z}_0\| > \|\boldsymbol{P}^{-1}\|^j) = \mathbb{P}(\log^+(\|\boldsymbol{Z}_0\|) > j \log^+(\|\boldsymbol{P}^{-1}\|)).\end{aligned}$$

Consequently, $\sum_{j=0}^{\infty} \mathbb{P}(\log^+(\|\boldsymbol{Z}_0\|) > j \log(\|\boldsymbol{P}^{-1}\|)) < \infty$, yielding

$$\mathbb{E}_{\mathbb{P}}\left(\frac{\log^+(\|\boldsymbol{Z}_0\|)}{\log(\|\boldsymbol{P}^{-1}\|)}\right) < \infty$$

and hence (i), where we used that $\log^+(\|\boldsymbol{P}^{-1}\|) = \log(\|\boldsymbol{P}^{-1}\|) > 0$ and $\mathbb{E}_{\mathbb{P}}(\xi) \leqslant 1 + \sum_{n=1}^{\infty} \mathbb{P}(\xi > n)$ for any non-negative random variable ξ. □

Proof of Theorem 1.4.
Step 1: Let $\mathbb{Q} := \mathbb{P}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}}$, and for each $n \in \mathbb{Z}_+$, put

$$L_n := \frac{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\} \mid \mathcal{F}_n)}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})}.$$

Then $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ and $\mathbb{P}_G$ as well, and, for each $n \in \mathbb{N}$, L_n is a well-defined and $\mathcal{F}_n$-measurable random variable, since $\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\}) > 0$. Note that $(L_n)_{n \in \mathbb{Z}_+}$ is the density process of $\mathbb{Q}$ with respect to $\mathbb{P}$, that is, $L_n = \frac{\mathrm{d}\,\mathbb{Q}|_{\mathcal{F}_n}}{\mathrm{d}\,\mathbb{P}|_{\mathcal{F}_n}}$ for every $n \in \mathbb{Z}_+$, where $\mathbb{Q}|_{\mathcal{F}_n}$ and $\mathbb{P}|_{\mathcal{F}_n}$ denote the restriction of $\mathbb{Q}$ and $\mathbb{P}$ onto $(\Omega, \mathcal{F}_n)$, respectively. Indeed, for all $A \in \mathcal{F}_n$, we have

$$\mathbb{Q}|_{\mathcal{F}_n}(A) = \mathbb{Q}(A) = \frac{\mathbb{P}(A \cap G \cap \{\exists \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})},$$

and, by the definition of conditional expectation with respect to the σ-algebra $\mathcal{F}_n$,

$$\begin{aligned}\int_A L_n(\omega)\, \mathbb{P}|_{\mathcal{F}_n}(\mathrm{d}\omega) &= \int_A \frac{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\} \mid \mathcal{F}_n)}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})}(\omega)\, \mathbb{P}|_{\mathcal{F}_n}(\mathrm{d}\omega) \\ &= \frac{1}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})} \int_A (\mathbb{E}_{\mathbb{P}}(\mathbb{1}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}} \mid \mathcal{F}_n))(\omega)\, \mathbb{P}(\mathrm{d}\omega) \\ &= \frac{1}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})} \int_A \mathbb{1}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}}(\omega)\, \mathbb{P}(\mathrm{d}\omega) = \frac{\mathbb{P}(A \cap G \cap \{\exists \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})},\end{aligned}$$

yielding that $\mathbb{Q}|_{\mathcal{F}_n}(A) = \int_A L_n(\omega)\, \mathbb{P}|_{\mathcal{F}_n}(\mathrm{d}\omega)$, $A \in \mathcal{F}_n$, as desired. Then, by Lévy's upwards theorem (see, e.g., Theorem A.6), we get

$$L_n \xrightarrow{L_1(\mathbb{P})} \frac{\mathbb{E}_{\mathbb{P}}(\mathbb{1}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}} \mid \mathcal{F}_\infty)}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})} = \frac{\mathbb{1}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}}}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})} = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \qquad \text{as } n \to \infty, \tag{16}$$

$$L_n \xrightarrow{\mathbb{P}\text{-a.s.}} \frac{\mathbb{E}_{\mathbb{P}}(\mathbb{1}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}} \mid \mathcal{F}_\infty)}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})} = \frac{\mathbb{1}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}}}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})} = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \qquad \text{as } n \to \infty, \tag{17}$$

where the second equality in (16) (and in (17)) holds, since for all $A \in \mathcal{F}$,

$$\mathbb{Q}(A) = \mathbb{P}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}}(A) = \frac{\mathbb{P}(A \cap G \cap \{\exists \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})},$$

and

$$\int_A \frac{\mathbb{1}_{G \cap \{\exists \boldsymbol{\eta}^{-1}\}}(\omega)}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})}\, \mathbb{P}(\mathrm{d}\omega) = \frac{\mathbb{P}(A \cap G \cap \{\exists \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists \boldsymbol{\eta}^{-1}\})}.$$

Next, we check that $\boldsymbol{Z}_j$, $j \in \mathbb{Z}_+$, and $\mathcal{F}_\infty$ are independent under $\mathbb{Q}$ as well. Indeed, since $\boldsymbol{Z}_j$, $j \in \mathbb{Z}_+$, and $\mathcal{F}_\infty$ are independent under $\mathbb{P}$ (by assumption) and $G \cap \{\exists\, \boldsymbol{\eta}^{-1}\} \in \mathcal{F}_\infty$ (since $G \in \mathcal{F}_\infty$ and $\boldsymbol{\eta}$ is $\mathcal{F}_\infty$-measurable), we have for each $m \in \mathbb{N}$, $B_0, B_1, \ldots, B_m \in \mathcal{B}(\mathbb{R}^d)$ and $A \in \mathcal{F}_\infty$,

$$\begin{aligned}
&\mathbb{Q}(\{\boldsymbol{Z}_0 \in B_0\} \cap \{\boldsymbol{Z}_1 \in B_1\} \cap \cdots \cap \{\boldsymbol{Z}_m \in B_m\} \cap A) \\
&\quad = \frac{\mathbb{P}(\{\boldsymbol{Z}_0 \in B_0\} \cap \{\boldsymbol{Z}_1 \in B_1\} \cap \cdots \cap \{\boldsymbol{Z}_m \in B_m\} \cap A \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \\
&\quad = \frac{\mathbb{P}(\{\boldsymbol{Z}_0 \in B_0\} \cap \{\boldsymbol{Z}_1 \in B_1\} \cap \cdots \cap \{\boldsymbol{Z}_m \in B_m\})\mathbb{P}(A \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \\
&\quad = \frac{\mathbb{P}(\boldsymbol{Z}_0 \in B_0)\mathbb{P}(\boldsymbol{Z}_1 \in B_1) \cdots \mathbb{P}(\boldsymbol{Z}_m \in B_m)\mathbb{P}(A \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})},
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{Q}(\boldsymbol{Z}_0 \in B_0)\mathbb{Q}(\boldsymbol{Z}_1 \in B_1) \cdots \mathbb{Q}(\boldsymbol{Z}_m \in B_m)\mathbb{Q}(A) \\
&\quad = \frac{\mathbb{P}(\{\boldsymbol{Z}_0 \in B_0\} \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \cdots \frac{\mathbb{P}(\{\boldsymbol{Z}_m \in B_m\} \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \cdot \frac{\mathbb{P}(A \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \\
&\quad = \mathbb{P}(\{\boldsymbol{Z}_0 \in B_0\}) \cdots \mathbb{P}(\{\boldsymbol{Z}_m \in B_m\}) \frac{\mathbb{P}(A \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})},
\end{aligned}$$

where we used that

$$\mathbb{P}(\{\boldsymbol{Z}_j \in B_j\} \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\}) = \mathbb{P}(\boldsymbol{Z}_j \in B_j)\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\}), \qquad j \in \{0, 1, \ldots, m\}.$$

It yields that

$$\begin{aligned}
&\mathbb{Q}(\{\boldsymbol{Z}_0 \in B_0\} \cap \{\boldsymbol{Z}_1 \in B_1\} \cap \cdots \cap \{\boldsymbol{Z}_m \in B_m\} \cap A) \\
&\qquad = \mathbb{Q}(\boldsymbol{Z}_0 \in B_0)\mathbb{Q}(\boldsymbol{Z}_1 \in B_1) \cdots \mathbb{Q}(\boldsymbol{Z}_m \in B_m)\mathbb{Q}(A),
\end{aligned}$$

as desired.

For each $\boldsymbol{\theta} \in \mathbb{R}^d$, let us introduce the notation

$$\varphi_\mu(\boldsymbol{\theta}) := \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{x}\rangle} \,\mu(\mathrm{d}\boldsymbol{x}) = \mathbb{E}_{\mathbb{P}}(\mathrm{e}^{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{Z}_0\rangle}) = \mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{Z}_0\rangle}), \tag{18}$$

since the distributions of $\boldsymbol{Z}_0$ under $\mathbb{P}$ and $\mathbb{Q}$ coincide. Indeed, by the independence of $\mathcal{F}_\infty$ and $\boldsymbol{Z}_0$ under $\mathbb{P}$, for all $B \in \mathcal{B}(\mathbb{R}^d)$, we have

$$\mathbb{Q}(\boldsymbol{Z}_0 \in B) = \frac{\mathbb{P}(\{\boldsymbol{Z}_0 \in B\} \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} = \frac{\mathbb{P}(\boldsymbol{Z}_0 \in B)\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} = \mathbb{P}(\boldsymbol{Z}_0 \in B),$$

as desired. Note that the function $\varphi_\mu : \mathbb{R}^d \to \mathbb{C}$ defined in (18) is nothing else but the characteristic function of $\boldsymbol{Z}_0$ under $\mathbb{P}$ (or $\mathbb{Q}$).

Step 2: Next, we show that for each $r \in \mathbb{Z}_+$, we have

$$\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j} \to \sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-mixing under } \mathbb{Q} \text{ as } n \to \infty. \tag{19}$$

Let $r \in \mathbb{Z}_+$ be fixed in this step. Since $\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j$ and $\mathcal{F}_\infty$ are independent under $\mathbb{Q}$, we need to check that

$$\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j} \to \sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{Q} \text{ as } n \to \infty, \tag{20}$$

see the discussion after Definition 3.15 in Häusler and Luschgy [11] (or Definition A.1). For this, by Corollary 3.19 in Häusler and Luschgy [11] (see, also Theorem A.3) with $\mathcal{G} := \mathcal{F}_\infty$ and $\mathcal{E} := \bigcup_{n\in\mathbb{Z}_+} \mathcal{F}_n$, it is enough to show that

$$\int_\Omega \mathbb{1}_F \exp\left\{\mathrm{i}\left\langle \boldsymbol{\theta}, \sum_{j=0}^r \boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j}\right\rangle\right\} \mathrm{d}\mathbb{Q} \to \int_\Omega \mathbb{1}_F \exp\left\{\mathrm{i}\left\langle \boldsymbol{\theta}, \sum_{j=0}^r \boldsymbol{P}^j \boldsymbol{Z}_j\right\rangle\right\} \mathrm{d}\mathbb{Q} \tag{21}$$

as $n \to \infty$ for all $\boldsymbol{\theta} \in \mathbb{R}^d$ and $F \in \mathcal{E}$. Indeed, $\mathcal{E} \subset \mathcal{F}_\infty$, $\mathcal{E}$ is closed under finite intersections, $\Omega \in \mathcal{E}$ and $\sigma(\mathcal{E}) = \mathcal{F}_\infty$. Now we turn to prove (21). For all $\boldsymbol{\theta} \in \mathbb{R}^d$ and $F \in \mathcal{E}$, we have

$$\exp\left\{\mathrm{i}\left\langle \boldsymbol{\theta}, \sum_{j=0}^r \boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j}\right\rangle\right\} = \prod_{j=0}^r \mathrm{e}^{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j}\rangle}$$

and

$$\begin{aligned}\int_\Omega \mathbb{1}_F \exp\left\{\mathrm{i}\left\langle \boldsymbol{\theta}, \sum_{j=0}^r \boldsymbol{P}^j \boldsymbol{Z}_j\right\rangle\right\} \mathrm{d}\mathbb{Q} &= \mathbb{Q}(F) \prod_{j=0}^r \mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{P}^j \boldsymbol{Z}_j\rangle}) = \mathbb{Q}(F) \prod_{j=0}^r \varphi_\mu((\boldsymbol{P}^\top)^j \boldsymbol{\theta}) \\ &= \int_F \prod_{j=0}^r \varphi_\mu((\boldsymbol{P}^\top)^j \boldsymbol{\theta})\, \mathrm{d}\mathbb{Q},\end{aligned}$$

where we used that $\boldsymbol{Z}_j$, $j \in \mathbb{Z}_+$, and $\mathcal{F}_\infty$ are independent under $\mathbb{Q}$, $\boldsymbol{Z}_j$, $j \in \mathbb{Z}_+$, are identically distributed under $\mathbb{Q}$, and the notation (18). Hence, fixing $\boldsymbol{\theta} \in \mathbb{R}^d$ arbitrarily, and using the notation $A_{n,j} := \exp\{\mathrm{i}\langle \boldsymbol{\theta}, \boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j}\rangle\}$, $C_j := \varphi_\mu((\boldsymbol{P}^\top)^j \boldsymbol{\theta})$ and $g_{n,r} := \prod_{j=0}^r C_j - \prod_{j=0}^r A_{n,j}$ for $n \in \mathbb{N}$ and $j \in \{0, \ldots, r\}$, convergence (21) means that $\int_F g_{n,r}\, \mathrm{d}\mathbb{Q} \to 0$ as $n \to \infty$ for all $F \in \mathcal{E}$. By $|g_{n,r}| \leqslant 2$ and (16), we get

$$\left| \int_F g_{n,r}\, \mathrm{d}\mathbb{Q} - \int_F L_{n-r-1} g_{n,r}\, \mathrm{d}\mathbb{P} \right| \leqslant 2 \int_F \left| \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} - L_{n-r-1} \right| \mathrm{d}\mathbb{P} \to 0$$

as $n \to \infty$. Consequently, in order to show (21), it is enough to verify that $\lim_{n\to\infty} \int_F L_{n-r-1} g_{n,r}\, \mathrm{d}\mathbb{P} = 0$. The condition $F \in \mathcal{E}$ yields the existence of $n_0 \in \mathbb{Z}_+$ such that $F \in \mathcal{F}_{n_0}$, and consequently $F \in \mathcal{F}_n$ for $n \geqslant n_0$. For each $n \in \mathbb{N}$ and $j \in \{0, \ldots, r\}$, put

$$D_{n,j} := \begin{cases} \prod_{k=1}^r A_{n,k} & \text{if } j = 0, \\ \left(\prod_{k=0}^{j-1} C_k\right)\left(\prod_{k=j+1}^r A_{n,k}\right) & \text{if } 1 \leqslant j \leqslant r-1, \\ \prod_{k=0}^{r-1} C_k & \text{if } j = r. \end{cases}$$

Then for each $n \in \mathbb{N}$, we have

$$\begin{aligned} g_{n,r} &= \prod_{k=0}^r C_k - \prod_{k=0}^r A_{n,k} \\ &= \prod_{k=0}^r C_k - \left(\prod_{k=0}^{r-1} C_k\right) A_{n,r} + \sum_{j=1}^{r-1} \left[\left(\prod_{k=0}^{j} C_k\right)\left(\prod_{k=j+1}^{r} A_{n,k}\right) - \left(\prod_{k=0}^{j-1} C_k\right)\left(\prod_{k=j}^{r} A_{n,k}\right) \right] \\ &\quad + C_0 \left(\prod_{k=1}^r A_{n,k}\right) - \prod_{k=0}^r A_{n,k} \\ &= \sum_{j=0}^r D_{n,j}(C_j - A_{n,j}), \end{aligned}$$

see also Lemma 8.4 in Häusler and Luschgy [11]. Moreover, for each $n \in \mathbb{N}$ and $j \in \{0, \ldots, r\}$, we have $|D_{n,j}| \leqslant 1$, and the $\mathcal{F}_{n-j}$-measurability of $A_{n,j}$ yields that $D_{n,j}$ is $\mathcal{F}_{n-j-1}$-measurable. Further, for each $n \geqslant n_0 + r + 1$, the random variable $\mathbb{1}_F L_{n-r-1}$ is $\mathcal{F}_{n-r-1}$-measurable, and hence $\mathcal{F}_{n-j-1}$-measurable for each $j \in \{0, \ldots, r\}$. Indeed, since $n - r - 1 \geqslant n_0$ and $F \in \mathcal{F}_{n_0}$, we have $F \in \mathcal{F}_{n-r-1}$, i.e., $\mathbb{1}_F$ is $\mathcal{F}_{n-r-1}$-measurable, so the $\mathcal{F}_{n-r-1}$-measurability of L_{n-r-1} yields that $\mathbb{1}_F L_{n-r-1}$ is $\mathcal{F}_{n-r-1}$-measurable. By the definition of conditional expectation, for each $n \geqslant n_0 + r + 1$, we obtain

$$\begin{aligned}
\left|\int_F L_{n-r-1} g_{n,r} \, \mathrm{d}\mathbb{P}\right| &= \left|\sum_{j=0}^r \int_F L_{n-r-1} D_{n,j}(C_j - A_{n,j}) \, \mathrm{d}\mathbb{P}\right| = \left|\sum_{j=0}^r \mathbb{E}_{\mathbb{P}}\left(\mathbb{1}_F L_{n-r-1} D_{n,j}(C_j - A_{n,j})\right)\right| \\
&= \left|\sum_{j=0}^r \mathbb{E}_{\mathbb{P}}\left(\mathbb{E}_{\mathbb{P}}(\mathbb{1}_F L_{n-r-1} D_{n,j}(C_j - A_{n,j}) \,|\, \mathcal{F}_{n-j-1})\right)\right| \\
&= \left|\sum_{j=0}^r \mathbb{E}_{\mathbb{P}}\left(\mathbb{1}_F L_{n-r-1} D_{n,j}(C_j - \mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1}))\right)\right| \\
&= \left|\sum_{j=0}^r \int_F L_{n-r-1} D_{n,j}(C_j - \mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1})) \, \mathrm{d}\mathbb{P}\right|.
\end{aligned}$$

Since $L_n \leqslant 1/\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})$, $|C_j| \leqslant 1$, $|A_{n,j}| \leqslant 1$, and $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \mathbb{1}_{G \cap \{\exists\, \boldsymbol{\eta}^{-1}\}} / \mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})$ (see the second equality in (16)) for each $n \geqslant n_0 + r + 1$, we have

$$\begin{aligned}
\left|\int_F L_{n-r-1} g_{n,r} \, \mathrm{d}\mathbb{P}\right| &\leqslant \sum_{j=0}^r \int_\Omega L_{n-r-1} |C_j - \mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1})| \, \mathrm{d}\mathbb{P} \\
&\leqslant \sum_{j=0}^r \int_{G \cap \{\exists\, \boldsymbol{\eta}^{-1}\}} \frac{1}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} |C_j - \mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1})| \, \mathrm{d}\mathbb{P} \\
&\quad + \sum_{j=0}^r \int_{\Omega \setminus (G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} L_{n-r-1}\Big(|C_j| + \mathbb{E}_{\mathbb{P}}(|A_{n,j}| \,|\, \mathcal{F}_{n-j-1})\Big) \, \mathrm{d}\mathbb{P} \\
&\leqslant \sum_{j=0}^r \int_\Omega |C_j - \mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1})| \, \mathrm{d}\mathbb{Q} + 2 \sum_{j=0}^r \int_{\Omega \setminus (G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} L_{n-r-1} \, \mathrm{d}\mathbb{P}.
\end{aligned}$$

For each $j \in \{0, \ldots, r\}$, condition (iv) yields

$$\int_\Omega |C_j - \mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1})| \, \mathrm{d}\mathbb{Q} \to 0 \qquad \text{as } n \to \infty. \tag{22}$$

Indeed, since $|\,\mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1})| \leqslant 1$, the family $\{\mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1}) : n \in \mathbb{N}\}$ is uniformly integrable under $\mathbb{Q}$ for each $j \in \{0, \ldots, r\}$, and, by (iv), $\mathbb{E}_{\mathbb{P}}(A_{n,j} \,|\, \mathcal{F}_{n-j-1}) \xrightarrow{\mathbb{Q}} C_j$ as $n \to \infty$ for each $j \in \{0, \ldots, r\}$, so the momentum convergence theorem yields (22). Further, using (17) and that $0 \leqslant L_{n-r-1} \leqslant 1/\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})$, the dominated convergence theorem yields that

$$\int_{\Omega \setminus (G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} L_{n-r-1} \, \mathrm{d}\mathbb{P} \to \int_{\Omega \setminus (G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \frac{\mathbb{1}_{G \cap \{\exists\, \boldsymbol{\eta}^{-1}\}}}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \, \mathrm{d}\mathbb{P} = \mathbb{Q}(\Omega \setminus (G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})) = 0$$

as $n \to \infty$. Consequently, we conclude $\lim_{n\to\infty} \int_F L_{n-r-1} g_{n,r} \, \mathrm{d}\mathbb{P} = 0$ for all $F \in \mathcal{E}$, and hence (21), which, as it was explained, implies (19).

Step 3: Next, we check that for each $r \in \mathbb{Z}_+$,

$$\boldsymbol{B}_n(\boldsymbol{U}_n - \boldsymbol{U}_{n-r-1}) \to \sum_{j=0}^r \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-mixing under } \mathbb{Q} \text{ as } n \to \infty. \tag{23}$$

For each $r \in \mathbb{Z}_+$ and $j \in \{0, \ldots, r\}$, we have

$$\boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j} - \boldsymbol{B}_n \Delta \boldsymbol{U}_{n-j} = (\boldsymbol{P}^j - \boldsymbol{B}_n \boldsymbol{B}_{n-j}^{-1}) \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j} \xrightarrow{\mathbb{Q}} \boldsymbol{0} \qquad \text{as } n \to \infty. \tag{24}$$

Indeed, $\boldsymbol{B}_n$ is invertible for sufficiently large $n \in \mathbb{N}$, and $\boldsymbol{P}^j - \boldsymbol{B}_n \boldsymbol{B}_{n-j}^{-1} \xrightarrow{\mathbb{Q}} \boldsymbol{0}$ as $n \to \infty$, since for all $\varepsilon > 0$, by condition (iii),

$$\begin{aligned}
\mathbb{Q}(\|\boldsymbol{P}^j - \boldsymbol{B}_n \boldsymbol{B}_{n-j}^{-1}\| > \varepsilon) &= \frac{\mathbb{P}(\{\|\boldsymbol{P}^j - \boldsymbol{B}_n \boldsymbol{B}_{n-j}^{-1}\| > \varepsilon\} \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \\
&\leqslant \frac{\mathbb{P}(\{\|\boldsymbol{P}^j - \boldsymbol{B}_n \boldsymbol{B}_{n-j}^{-1}\| > \varepsilon\} \cap G)}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \\
&= \mathbb{P}_G(\|\boldsymbol{P}^j - \boldsymbol{B}_n \boldsymbol{B}_{n-j}^{-1}\| > \varepsilon) \frac{\mathbb{P}(G)}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \to 0 \qquad \text{as } n \to \infty.
\end{aligned}$$

Further, by (19) with $r = 0$ and using the fact that $\mathcal{F}_\infty$-mixing convergence under $\mathbb{Q}$ yields convergence in distribution under $\mathbb{Q}$, we have $\boldsymbol{B}_n \Delta \boldsymbol{U}_n \xrightarrow{\mathcal{D}(\mathbb{Q})} \boldsymbol{Z}_0$ as $n \to \infty$, and especially, for each $j \in \{0, \ldots, r\}$, $\boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j} \xrightarrow{\mathcal{D}(\mathbb{Q})} \boldsymbol{Z}_0$ as $n \to \infty$. By Slutsky's lemma, we have (24). Hence for each $r \in \mathbb{Z}_+$, we have

$$\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{B}_{n-j} \Delta \boldsymbol{U}_{n-j} - \sum_{j=0}^{r} \boldsymbol{B}_n \Delta \boldsymbol{U}_{n-j} \xrightarrow{\mathbb{Q}} \boldsymbol{0} \qquad \text{as } n \to \infty.$$

Consequently, since $\sum_{j=0}^{r} \boldsymbol{B}_n \Delta \boldsymbol{U}_{n-j} = \boldsymbol{B}_n (\boldsymbol{U}_n - \boldsymbol{U}_{n-r-1})$, $n \in \mathbb{N}$, by (20) and part (a) of Theorem 3.18 in Häusler and Luschgy [11] (see also Theorem A.2), for each $r \in \mathbb{Z}_+$, we have

$$\boldsymbol{B}_n (\boldsymbol{U}_n - \boldsymbol{U}_{n-r-1}) \to \sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{Q} \text{ as } n \to \infty.$$

Since $\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j$ and $\mathcal{F}_\infty$ are independent under $\mathbb{Q}$ (following from the $\mathbb{Q}$-independence of $\boldsymbol{Z}_j$, $j \in \mathbb{Z}_+$, and $\mathcal{F}_\infty$, which was proved in Step 1), by the discussion after Definition 3.15 in Häusler and Luschgy [11] (see also Definition A.1), we have (23).

Step 4: Now we turn to prove (3). Lemma 1.3, the invertibility of $\boldsymbol{P}$, $\varrho(\boldsymbol{P}) < 1$, the condition $\int_{\mathbb{R}^d} \log^+(\|\boldsymbol{x}\|) \, \mu(\mathrm{d}\boldsymbol{x}) < \infty$ and the fact that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ (see Step 1) yield the $\mathbb{P}$-almost sure and the $\mathbb{Q}$-almost sure absolute convergence of the series $\sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j$. Especially,

$$\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j \to \sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \text{as } r \to \infty \ \mathbb{Q}\text{-almost surely,}$$

and hence

$$\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j \xrightarrow{\mathcal{D}(\mathbb{Q})} \sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \text{as } r \to \infty.$$

Consequently, using that $\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j$ and $\mathcal{F}_\infty$ are independent under $\mathbb{Q}$ for each $r \in \mathbb{Z}_+$, by Exercise 3.4 in Häusler and Luschgy [11], we have

$$\sum_{j=0}^{r} \boldsymbol{P}^j \boldsymbol{Z}_j \to \sum_{j=0}^{\infty} \boldsymbol{P}^j \boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-mixing under } \mathbb{Q} \text{ as } r \to \infty. \tag{25}$$

Since $\boldsymbol{B}_n\boldsymbol{U}_n - \boldsymbol{B}_n(\boldsymbol{U}_n - \boldsymbol{U}_{n-r-1}) = \boldsymbol{B}_n\boldsymbol{U}_{n-r-1}$, and $\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{Z}_j$ and $\mathcal{F}_\infty$ are independent under $\mathbb{Q}$ (following from the fact that $\boldsymbol{Z}_j$, $j \in \mathbb{Z}_+$, and $\mathcal{F}_\infty$ are independent under $\mathbb{Q}$, which we checked in Step 1), by (23), (25) and Theorem 3.21 in Häusler and Luschgy [11] (see, also Theorem A.4), we obtain (3) if we can check

$$\lim_{r\to\infty}\limsup_{n\to\infty}\mathbb{Q}(\|\boldsymbol{B}_n\boldsymbol{U}_{n-r-1}\| > \varepsilon) = 0 \tag{26}$$

for all $\varepsilon \in (0,\infty)$. Since $\boldsymbol{B}_n$ and $\boldsymbol{Q}_n$ are invertible for sufficiently large $n \in \mathbb{N}$, and $\boldsymbol{P}$ is invertible, for each $r \in \mathbb{Z}_+$ and for sufficiently large $n \in \mathbb{N}$, we have

$$\|\boldsymbol{B}_n\boldsymbol{U}_{n-r-1}\| \leqslant \|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\|.$$

Since for each $r \in \mathbb{Z}_+$, $\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} \xrightarrow{\mathbb{Q}} \boldsymbol{P}^{r+1}$ as $n \to \infty$ (see Step 3), and

$$\|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} - \boldsymbol{I}_d\| \leqslant \|\boldsymbol{P}^{-r-1}\|\|\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} - \boldsymbol{P}^{r+1}\|,$$

we have $\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} \xrightarrow{\mathbb{Q}} \boldsymbol{I}_d$ as $n \to \infty$ for each $r \in \mathbb{Z}_+$. Hence for all $\widetilde{\varepsilon} > 0$, $\kappa > 0$ and $r \in \mathbb{Z}_+$, we have

$$\mathbb{Q}(\|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} - \boldsymbol{I}_d\| \geqslant \widetilde{\varepsilon}) < \kappa \qquad \text{for sufficiently large } n \in \mathbb{N}. \tag{27}$$

Consequently, with the notation $G_{n,r,\widetilde{\varepsilon}} := \{\|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} - \boldsymbol{I}_d\| < \widetilde{\varepsilon}\}$, for all $\varepsilon, \widetilde{\varepsilon}, \delta, \kappa \in (0,\infty)$, $r \in \mathbb{Z}_+$, and for sufficiently large $n \in \mathbb{N}$, we have

$$\begin{aligned}
&\mathbb{Q}(\|\boldsymbol{B}_n\boldsymbol{U}_{n-r-1}\| > \varepsilon)\\
&\leqslant \mathbb{Q}\Big(\|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \varepsilon\Big)\\
&= \mathbb{Q}\Big(\Big\{\|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \varepsilon\Big\} \cap G_{n,r,\widetilde{\varepsilon}}\Big)\\
&\quad + \mathbb{Q}\Big(\Big\{\|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \varepsilon\Big\} \cap G_{n,r,\widetilde{\varepsilon}}^c\Big)\\
&\leqslant \mathbb{Q}\Big(\Big\{\|\boldsymbol{P}^{r+1}\|\cdot\|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} - \boldsymbol{I}_d\|\cdot\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\|\cdot\|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2}\Big\} \cap G_{n,r,\widetilde{\varepsilon}}\Big)\\
&\quad + \mathbb{Q}\Big(\Big\{\|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2}\Big\} \cap G_{n,r,\widetilde{\varepsilon}}\Big)\\
&\quad + \mathbb{Q}\Big(\Big\{\|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \varepsilon\Big\} \cap G_{n,r,\widetilde{\varepsilon}}^c\Big)\\
&\leqslant \mathbb{Q}\Big(\|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2\widetilde{\varepsilon}}\Big)\\
&\quad + \mathbb{Q}\Big(\|\boldsymbol{P}^{r+1}\| \cdot \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2}\Big)\\
&\quad + \mathbb{Q}\Big(\|\boldsymbol{P}^{-r-1}\boldsymbol{B}_n\boldsymbol{B}_{n-r-1}^{-1} - \boldsymbol{I}_d\| \geqslant \widetilde{\varepsilon}\Big),
\end{aligned}$$

where $G_{n,r,\widetilde{\varepsilon}}^c$ denotes the complement of $G_{n,r,\widetilde{\varepsilon}}$. Since, by (15), $\|\boldsymbol{P}^{r+1}\| \leqslant \left(\frac{1+\varrho(\boldsymbol{P})}{2}\right)^{r+1}$ for sufficiently large $r \in \mathbb{N}$, using also (27), for all $\varepsilon, \delta, \kappa \in (0,\infty)$, $\widetilde{\varepsilon} \in (0,1)$, and for sufficiently large $r \in \mathbb{N}$, there exists a sufficiently large $n(r) \in \mathbb{N}$ (here $n(r)$ may depend on $\widetilde{\varepsilon}$ and κ as well, but we do not denote this dependence) such that for $n \geqslant n(r)$, we have

$$
\begin{aligned}
&\mathbb{Q}(\|\boldsymbol{B}_n \boldsymbol{U}_{n-r-1}\| > \varepsilon) \\
&\leqslant \mathbb{Q}\Big(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2\widetilde{\varepsilon}}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}\Big) \\
&\quad + \mathbb{Q}\Big(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}\Big) + \kappa \\
&\leqslant 2\mathbb{Q}\Big(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}\Big) + \kappa \\
&= 2\mathbb{Q}\Big(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}, \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \leqslant \delta\Big) \\
&\quad + 2\mathbb{Q}\Big(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| \cdot \|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}, \|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| > \delta\Big) + \kappa \\
&\leqslant 2\mathbb{Q}\Big(\|\boldsymbol{Q}_{n-r-1}\boldsymbol{U}_{n-r-1}\| > \frac{\varepsilon}{2\delta}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}\Big) + 2\mathbb{Q}(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| > \delta) + \kappa.
\end{aligned}
$$

So for all $\varepsilon, \delta, \kappa \in (0, \infty)$ and for sufficiently large $r \in \mathbb{N}$, there exists a sufficiently large $n(r) \in \mathbb{N}$ such that for $n \geqslant n(r)$, we have

$$
\begin{aligned}
&\mathbb{Q}(\|\boldsymbol{B}_n \boldsymbol{U}_{n-r-1}\| > \varepsilon) \\
&\leqslant 2\sup_{\ell\in\mathbb{N}} \mathbb{Q}\Big(\|\boldsymbol{Q}_\ell \boldsymbol{U}_\ell\| > \frac{\varepsilon}{2\delta}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}\Big) + 2\mathbb{Q}(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| > \delta, \|\boldsymbol{\eta}^{-1}\| \leqslant \delta/2) \\
&\quad + 2\mathbb{Q}(\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| > \delta, \|\boldsymbol{\eta}^{-1}\| > \delta/2) + \kappa \\
&\leqslant 2\sup_{\ell\in\mathbb{N}} \mathbb{Q}\Big(\|\boldsymbol{Q}_\ell \boldsymbol{U}_\ell\| > \frac{\varepsilon}{2\delta}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}\Big) + 2\mathbb{Q}\Big(\big|\|\boldsymbol{B}_{n-r-1}\boldsymbol{Q}_{n-r-1}^{-1}\| - \|\boldsymbol{\eta}^{-1}\|\big| > \delta/2\Big) \\
&\quad + 2\mathbb{Q}(\|\boldsymbol{\eta}^{-1}\| > \delta/2) + \kappa,
\end{aligned}
$$

where we used that $\mathbb{Q}(\exists\, \boldsymbol{\eta}^{-1}) = 1$. Similarly as we have seen in Step 3, condition (i) implies $\boldsymbol{Q}_n \boldsymbol{B}_n^{-1} \xrightarrow{\mathbb{Q}} \boldsymbol{\eta}$ as $n \to \infty$. Indeed, since $\mathbb{P}(G) > 0$, for all $\gamma > 0$, we have

$$
\begin{aligned}
\mathbb{Q}(\|\boldsymbol{Q}_n \boldsymbol{B}_n^{-1} - \boldsymbol{\eta}\| > \gamma) &= \frac{\mathbb{P}(\{\|\boldsymbol{Q}_n \boldsymbol{B}_n^{-1} - \boldsymbol{\eta}\| > \gamma\} \cap G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \\
&\leqslant \frac{\mathbb{P}(\{\|\boldsymbol{Q}_n \boldsymbol{B}_n^{-1} - \boldsymbol{\eta}\| > \gamma\} \cap G)}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \\
&= \mathbb{P}_G(\|\boldsymbol{Q}_n \boldsymbol{B}_n^{-1} - \boldsymbol{\eta}\| > \gamma)\frac{\mathbb{P}(G)}{\mathbb{P}(G \cap \{\exists\, \boldsymbol{\eta}^{-1}\})} \to 0 \qquad \text{as } n \to \infty.
\end{aligned}
$$

Since $\boldsymbol{Q}_n$ is invertible for sufficiently large $n \in \mathbb{N}$, $\mathbb{Q}(\exists\, \boldsymbol{\eta}^{-1}) = 1$ and the norm function is continuous, we get $\|\boldsymbol{B}_n \boldsymbol{Q}_n^{-1}\| \xrightarrow{\mathbb{Q}} \|\boldsymbol{\eta}^{-1}\|$ as $n \to \infty$. Thus, for all $\varepsilon, \delta, \kappa \in (0, \infty)$ and for sufficiently large $r \in \mathbb{N}$, we obtain

$$
\limsup_{n\to\infty} \mathbb{Q}(\|\boldsymbol{B}_n \boldsymbol{U}_{n-r-1}\| > \varepsilon) \leqslant 2\sup_{\ell\in\mathbb{N}} \mathbb{Q}\Big(\|\boldsymbol{Q}_\ell \boldsymbol{U}_\ell\| > \frac{\varepsilon}{2\delta}\Big(\frac{2}{1+\varrho(\boldsymbol{P})}\Big)^{r+1}\Big) + 2\mathbb{Q}(\|\boldsymbol{\eta}^{-1}\| > \delta/2) + \kappa.
$$

Using condition (ii) and that $\frac{2}{1+\varrho(\boldsymbol{P})} > 1$, for all $\varepsilon, \delta, \kappa \in (0, \infty)$, we get

$$
\limsup_{r\to\infty} \limsup_{n\to\infty} \mathbb{Q}(\|\boldsymbol{B}_n \boldsymbol{U}_{n-r-1}\| > \varepsilon) \leqslant 2\mathbb{Q}(\|\boldsymbol{\eta}^{-1}\| > \delta/2) + \kappa.
$$

We have $\mathbb{Q}(\|\boldsymbol{\eta}^{-1}\| > \delta/2) \to 0$ as $\delta \to \infty$, hence, taking $\limsup_{\delta\to\infty}$ and $\limsup_{\kappa\downarrow 0}$, we obtain (26) for all $\varepsilon \in (0, \infty)$, and then we conclude (3).

Step 5: Now we turn to prove (4). As we have seen in Step 4, condition (i) implies $\boldsymbol{Q}_n\boldsymbol{B}_n^{-1} \xrightarrow{\mathbb{Q}} \boldsymbol{\eta}$ as $n \to \infty$. Hence, since $\boldsymbol{\eta}$ is $\mathcal{F}_\infty$-measurable, by (3) (which was proved in Step 4) and parts (b) and (c) of Theorem 3.18 in Häusler and Luschgy [11] (see, also Theorem A.2), we have

$$\boldsymbol{Q}_n\boldsymbol{U}_n = (\boldsymbol{Q}_n\boldsymbol{B}_n^{-1})(\boldsymbol{B}_n\boldsymbol{U}_n) \to \boldsymbol{\eta}\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{Z}_j \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{Q} = \mathbb{P}_{G\cap\{\exists\boldsymbol{\eta}^{-1}\}} \text{ as } n \to \infty,$$

yielding (4). □

Proof of Corollary 1.8. First, note that $\log^+(\|\boldsymbol{x}\|) \leqslant \|\boldsymbol{x}\|$, $\boldsymbol{x} \in \mathbb{R}^d$, so

$$\int_{\mathbb{R}^d}\log^+(\|\boldsymbol{x}\|)\,\mu(\mathrm{d}\boldsymbol{x}) \leqslant \int_{\mathbb{R}^d}\|\boldsymbol{x}\|\,\mu(\mathrm{d}\boldsymbol{x}) < \infty,$$

and then we can indeed apply Theorem 1.4 and $\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{Z}_0\|)) < \infty$. It remains to check that $\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{Z}_j$ is a d-dimensional normally distributed random variable with mean vector $\boldsymbol{0} \in \mathbb{R}^d$ and covariance matrix $\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{D}(\boldsymbol{P}^j)^\top$. Since $\boldsymbol{P}$ is invertible, $\varrho(\boldsymbol{P}) < 1$ and $\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{Z}_0\|)) < \infty$, by Lemma 1.3, we have that the series $\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{Z}_j$ is absolutely convergent $\mathbb{P}$-a.s., and hence, by the continuity theorem, we get

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta},\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{Z}_j\rangle}\right) &= \lim_{r\to\infty}\mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta},\sum_{j=0}^{r}\boldsymbol{P}^j\boldsymbol{Z}_j\rangle}\right) = \lim_{r\to\infty}\prod_{j=0}^{r}\mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{\mathrm{i}\langle(\boldsymbol{P}^j)^\top\boldsymbol{\theta},\boldsymbol{Z}_j\rangle}\right)\\ &= \lim_{r\to\infty}\prod_{j=0}^{r}\mathrm{e}^{-\frac{1}{2}\langle\boldsymbol{D}(\boldsymbol{P}^j)^\top\boldsymbol{\theta},(\boldsymbol{P}^j)^\top\boldsymbol{\theta}\rangle} = \mathrm{e}^{-\frac{1}{2}\left\langle\left(\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{D}(\boldsymbol{P}^j)^\top\right)\boldsymbol{\theta},\boldsymbol{\theta}\right\rangle}, \qquad \boldsymbol{\theta} \in \mathbb{R}^d,\end{aligned}$$

where the series $\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{D}(\boldsymbol{P}^j)^\top$ is absolutely convergent, since, by (15),

$$\sum_{j=0}^{\infty}\|\boldsymbol{P}^j\boldsymbol{D}(\boldsymbol{P}^j)^\top\| \leqslant \sum_{j=0}^{\infty}\|\boldsymbol{P}^j\|\|\boldsymbol{D}\|\|(\boldsymbol{P}^j)^\top\| \leqslant \|\boldsymbol{D}\|\sum_{j=0}^{k_0-1}\|\boldsymbol{P}^j\|^2 + \|\boldsymbol{D}\|\sum_{j=k_0}^{\infty}\left(\frac{1+\varrho(\boldsymbol{P})}{2}\right)^{2j} < \infty,$$

where k_0 is appearing in (15). So $\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{Z}_j$ is a d-dimensional normally distributed random variable with mean vector $\boldsymbol{0} \in \mathbb{R}^d$ and covariance matrix $\sum_{j=0}^{\infty}\boldsymbol{P}^j\boldsymbol{D}(\boldsymbol{P}^j)^\top$, as desired. □

Proof of Corollary 1.9. First, note that the integral appearing in (11) is convergent, since, by Cauchy-Schwarz's inequality and (15), for all $\boldsymbol{\theta} \in \mathbb{R}^d$,

$$\begin{aligned}&\int_{S_{d-1}}\sum_{j=0}^{\infty}\left|\langle(\boldsymbol{P}^j)^\top\boldsymbol{\theta},\boldsymbol{x}\rangle\right|^\alpha\,\Pi(\mathrm{d}\boldsymbol{x}) \leqslant \int_{S_{d-1}}\sum_{j=0}^{\infty}\left\|(\boldsymbol{P}^j)^\top\boldsymbol{\theta}\right\|^\alpha\|\boldsymbol{x}\|^\alpha\,\Pi(\mathrm{d}\boldsymbol{x})\\ &\leqslant \left\|\boldsymbol{\theta}\right\|^\alpha\int_{S_{d-1}}\sum_{j=0}^{\infty}\|\boldsymbol{P}^j\|^\alpha\|\boldsymbol{x}\|^\alpha\,\Pi(\mathrm{d}\boldsymbol{x})\\ &\leqslant \left\|\boldsymbol{\theta}\right\|^\alpha\int_{S_{d-1}}\sum_{j=0}^{k_0-1}\|\boldsymbol{P}^j\|^\alpha\|\boldsymbol{x}\|^\alpha\,\Pi(\mathrm{d}\boldsymbol{x}) + \left\|\boldsymbol{\theta}\right\|^\alpha\int_{S_{d-1}}\sum_{j=k_0}^{\infty}\left(\frac{1+\varrho(\boldsymbol{P})}{2}\right)^{\alpha j}\|\boldsymbol{x}\|^\alpha\,\Pi(\mathrm{d}\boldsymbol{x})\\ &= \left\|\boldsymbol{\theta}\right\|^\alpha\sum_{j=0}^{k_0-1}\|\boldsymbol{P}^j\|^\alpha\,\Pi(S_{d-1}) + \left\|\boldsymbol{\theta}\right\|^\alpha\sum_{j=k_0}^{\infty}\left(\frac{1+\varrho(\boldsymbol{P})}{2}\right)^{\alpha j}\Pi(S_{d-1}) < \infty,\end{aligned}$$

where k_0 is appearing in (15) and we also used that $\frac{1+\varrho(\boldsymbol{P})}{2} \in (0,1)$ and $\Pi(S_{d-1}) < \infty$.

Next, we check that $\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{\zeta}\|)) < \infty$. We have

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{\zeta}\|)) &= \mathbb{E}_{\mathbb{P}}(\log(\|\boldsymbol{\zeta}\|)\mathbb{1}_{\{\|\boldsymbol{\zeta}\|\geqslant 1\}}) = \int_0^\infty \mathbb{P}(\log(\|\boldsymbol{\zeta}\|)\mathbb{1}_{\{\|\boldsymbol{\zeta}\|\geqslant 1\}} \geqslant y)\,\mathrm{d}y \\
&= \int_0^\infty \mathbb{P}(\log(\|\boldsymbol{\zeta}\|) \geqslant y, \|\boldsymbol{\zeta}\| \geqslant 1)\,\mathrm{d}y = \int_0^\infty \mathbb{P}(\|\boldsymbol{\zeta}\| \geqslant \mathrm{e}^y, \|\boldsymbol{\zeta}\| \geqslant 1)\,\mathrm{d}y \\
&= \int_0^1 \mathbb{P}(\|\boldsymbol{\zeta}\| \geqslant \mathrm{e}^y)\,\mathrm{d}y + \int_1^\infty \mathbb{P}(\|\boldsymbol{\zeta}\| \geqslant \mathrm{e}^y)\,\mathrm{d}y \\
&\leqslant 1 + \int_{\mathrm{e}}^\infty \mathbb{P}(\|\boldsymbol{\zeta}\| \geqslant z)\frac{1}{z}\,\mathrm{d}z.
\end{aligned}
$$

Since $\boldsymbol{\zeta}$ has a d-dimensional stable distribution, it belongs to its own domain of attraction, and then it is known that the function $\mathbb{R}_{++} \ni z \mapsto \mathbb{P}(\|\boldsymbol{\zeta}\| \geqslant z)$ is regularly varying with tail index α. As a consequence, the function $\mathbb{R}_{++} \ni z \mapsto z^\alpha \mathbb{P}(\|\boldsymbol{\zeta}\| \geqslant z) =: L(z)$ is slowly varying. Hence there exists $z_0 \in (\mathrm{e}, \infty)$ such that $z^{-\frac{\alpha}{2}} L(z) \leqslant 1$ for all $z \in [z_0, \infty)$, see, e.g., Bingham et al. [5, Proposition 1.3.6.(v)]. Consequently, we have

$$
\begin{aligned}
\int_{\mathrm{e}}^\infty \mathbb{P}(\|\boldsymbol{\zeta}\| \geqslant z)\frac{1}{z}\,\mathrm{d}z &= \int_{\mathrm{e}}^{z_0} z^{-\alpha}L(z)\frac{1}{z}\,\mathrm{d}z + \int_{z_0}^\infty z^{-\alpha}L(z)\frac{1}{z}\,\mathrm{d}z \\
&\leqslant \int_{\mathrm{e}}^{z_0} z^{-\alpha}L(z)\frac{1}{z}\,\mathrm{d}z + \int_{z_0}^\infty z^{-\frac{\alpha}{2}}\frac{1}{z}\,\mathrm{d}z \leqslant \int_{\mathrm{e}}^{z_0} \frac{1}{z}\,\mathrm{d}z + \int_{z_0}^\infty z^{-1-\frac{\alpha}{2}}\,\mathrm{d}z < \infty,
\end{aligned}
$$

since $z^{-\alpha}L(z) = \mathbb{P}(\|\boldsymbol{\zeta}\| > z) \leqslant 1$, $z \in \mathbb{R}_{++}$.

Hence one can indeed apply Theorem 1.4 and $\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{Z}_0\|)) < \infty$. It remains to check that the characteristic function of $\sum_{j=0}^\infty \boldsymbol{P}^j \boldsymbol{Z}_j$ is given by (11). Since $\boldsymbol{P}$ is invertible, $\varrho(\boldsymbol{P}) < 1$, and $\mathbb{E}_{\mathbb{P}}(\log^+(\|\boldsymbol{Z}_0\|)) < \infty$, by Lemma 1.3, we have that $\sum_{j=0}^\infty \boldsymbol{P}^j \boldsymbol{Z}_j$ is absolutely convergent $\mathbb{P}$-a.s., and hence, by the continuity theorem, we get

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta}, \sum_{j=0}^\infty \boldsymbol{P}^j \boldsymbol{Z}_j\rangle}\right) &= \lim_{r\to\infty} \prod_{j=0}^r \mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{\mathrm{i}\langle(\boldsymbol{P}^j)^\top\boldsymbol{\theta}, \boldsymbol{Z}_j\rangle}\right) = \lim_{r\to\infty} \exp\left\{-\sum_{j=0}^r \int_{S_{d-1}} |\langle(\boldsymbol{P}^j)^\top\boldsymbol{\theta}, \boldsymbol{x}\rangle|^\alpha \,\Pi(\mathrm{d}\boldsymbol{x})\right\} \\
&= \exp\left\{-\sum_{j=0}^\infty \int_{S_{d-1}} |\langle(\boldsymbol{P}^j)^\top\boldsymbol{\theta}, \boldsymbol{x}\rangle|^\alpha \,\Pi(\mathrm{d}\boldsymbol{x})\right\}, \qquad \boldsymbol{\theta} \in \mathbb{R}^d,
\end{aligned}
$$

yielding (11).

In the special case when $\boldsymbol{\zeta}$ has a d-dimensional Cauchy distribution with parameter $(\boldsymbol{0}, \boldsymbol{I}_d)$, we have $\int_{S_{d-1}} |\langle(\boldsymbol{P}^j)^\top\boldsymbol{\theta}, \boldsymbol{x}\rangle|^\alpha \,\Pi(\mathrm{d}\boldsymbol{x}) = \|(\boldsymbol{P}^j)^\top\boldsymbol{\theta}\|$, $\boldsymbol{\theta} \in \mathbb{R}^d$, $j \in \mathbb{Z}_+$, yielding (12). □

Proof of Corollary 1.10. Let $\{\boldsymbol{s}_k : k \in \mathbb{N}\}$ be the range of $\boldsymbol{S}$, let $G_k := \{\boldsymbol{S} = \boldsymbol{s}_k\}$, $k \in \mathbb{N}$, and $I := \{k \in \mathbb{N} : \mathbb{P}(G_k \cap \{\exists\, \boldsymbol{\eta}^{-1}\}) > 0\}$. Since $\mathbb{P}(\exists\, \boldsymbol{\eta}^{-1}) > 0$ (due to $G = \Omega$), we have that I is not the empty set. Further, since $\mathbb{P}_{G_k\cap\{\exists\,\boldsymbol{\eta}^{-1}\}}$ is absolutely continuous with respect to $\mathbb{P}_{\{\exists\,\boldsymbol{\eta}^{-1}\}}$, by (iv′), and using that convergence in $\mathbb{P}_{\{\exists\,\boldsymbol{\eta}^{-1}\}}$-probability yields convergence in $\mathbb{P}_{G_k\cap\{\exists\,\boldsymbol{\eta}^{-1}\}}$-probability (which can be checked similarly as in case of $\mathbb{P}_G$ and $\mathbb{P}_{G\cap\{\exists\,\boldsymbol{\eta}^{-1}\}}$ as we have seen in the proof of Step 3 of Theorem 1.4), we have for each $k \in I$ and $\boldsymbol{\theta} \in \mathbb{R}^d$,

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}}\Big(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta}, \boldsymbol{B}_n\Delta\boldsymbol{U}_n\rangle} \,|\, \mathcal{F}_{n-1}\Big) \xrightarrow{\mathbb{P}_{G_k\cap\{\exists\boldsymbol{\eta}^{-1}\}}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta}, \boldsymbol{S}\boldsymbol{x}\rangle}\,\mu(\mathrm{d}\boldsymbol{x}) &= \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta}, \boldsymbol{s}_k\boldsymbol{x}\rangle}\,\mu(\mathrm{d}\boldsymbol{x}) = \mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{s}_k^\top\boldsymbol{\theta}, \boldsymbol{Z}_0\rangle}\right) \\
&= \mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{\mathrm{i}\langle\boldsymbol{\theta}, \boldsymbol{s}_k\boldsymbol{Z}_0\rangle}\right)
\end{aligned}
$$

as $n \to \infty$. Moreover, $\mathbb{E}_{\mathbb{P}}(\log^+(\|s_k Z_0\|)) < \infty$, since

$$\begin{aligned}\log^+(\|s_k Z_0\|) &= \log(\|s_k Z_0\|)\mathbb{1}_{\{\|s_k Z_0\| \geqslant 1\}} \leqslant \log(\|s_k\|\|Z_0\|)\mathbb{1}_{\{\|s_k\|\|Z_0\| \geqslant 1\}} \\ &\leqslant \log(\|s_k\|)\mathbb{1}_{\{s_k \neq \mathbf{0}\}}\mathbb{1}_{\{\|s_k\|\|Z_0\| \geqslant 1\}} + \log(\|Z_0\|)\mathbb{1}_{\{\|Z_0\| \geqslant 1\}} + \log(\|Z_0\|)\mathbb{1}_{\{\frac{1}{\|s_k\|} \leqslant \|Z_0\| < 1\}}\mathbb{1}_{\{s_k \neq \mathbf{0}\}},\end{aligned}$$

which yields that

$$\mathbb{E}_{\mathbb{P}}(\log^+(\|s_k Z_0\|)) \leqslant \log(\|s_k\|)\mathbb{1}_{\{s_k \neq \mathbf{0}\}} + \mathbb{E}_{\mathbb{P}}(\log^+(\|Z_0\|)) < \infty.$$

Hence, by Theorem 1.4, for each $k \in I$, we have

$$B_n U_n \to \sum_{j=0}^{\infty} P^j s_k Z_j = \sum_{j=0}^{\infty} P^j S Z_j \qquad \mathcal{F}_\infty\text{-mixing under } \mathbb{P}_{G_k \cap \{\exists\, \eta^{-1}\}} \text{ as } n \to \infty,$$

and

$$Q_n U_n \to \eta \sum_{j=0}^{\infty} P^j s_k Z_j = \eta \sum_{j=0}^{\infty} P^j S Z_j \qquad \mathcal{F}_\infty\text{-stably under } \mathbb{P}_{G_k \cap \{\exists\, \eta^{-1}\}} \text{ as } n \to \infty.$$

Note that, since $G = \Omega$, we have $\mathbb{P}(\exists\, \eta^{-1}) > 0$ and for all $A \in \mathcal{F}$,

$$\begin{aligned}\mathbb{P}_{\{\exists\, \eta^{-1}\}}(A) &= \sum_{k=1}^{\infty} \mathbb{P}_{\{\exists\, \eta^{-1}\}}(A \cap G_k) = \sum_{k=1}^{\infty} \frac{\mathbb{P}(A \cap G_k \cap \{\exists\, \eta^{-1}\})}{\mathbb{P}(\exists\, \eta^{-1})} \\ &= \sum_{k \in I} \frac{\mathbb{P}(A \cap G_k \cap \{\exists\, \eta^{-1}\})}{\mathbb{P}(G_k \cap \{\exists\, \eta^{-1}\})} \frac{\mathbb{P}(G_k \cap \{\exists\, \eta^{-1}\})}{\mathbb{P}(\exists\, \eta^{-1})} = \sum_{k \in I} \mathbb{P}_{G_k \cap \{\exists\, \eta^{-1}\}}(A)\mathbb{P}_{\{\exists\, \eta^{-1}\}}(G_k),\end{aligned}$$

so we have

$$\mathbb{P}_{\{\exists\, \eta^{-1}\}} = \sum_{k \in I} \mathbb{P}_{\{\exists\, \eta^{-1}\}}(G_k)\mathbb{P}_{G_k \cap \{\exists\, \eta^{-1}\}},$$

where $\sum_{k \in I} \mathbb{P}_{\{\exists\, \eta^{-1}\}}(G_k) = 1$. Finally, Proposition 3.24 in Häusler and Luschgy [11] (see also Theorem A.5) yields the statement. $\square$

Appendix

Appendix A. Stable convergence and Lévy's upwards theorem

First, we recall the notions of stable and mixing convergence.

Definition A.1 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-field. Let $(X_n)_{n \in \mathbb{N}}$ and X be $\mathbb{R}^d$-valued random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$, where $d \in \mathbb{N}$.*

(i) *We say that X_n converges $\mathcal{G}$-stably to X as $n \to \infty$, if the conditional distribution $\mathbb{P}^{X_n | \mathcal{G}}$ of X_n given $\mathcal{G}$ converges weakly to the conditional distribution $\mathbb{P}^{X | \mathcal{G}}$ of X given $\mathcal{G}$ as $n \to \infty$ in the sense of weak convergence of Markov kernels. It equivalently means that*

$$\lim_{n \to \infty} \mathbb{E}_{\mathbb{P}}(\xi\, \mathbb{E}_{\mathbb{P}}(h(X_n) \,|\, \mathcal{G})) = \mathbb{E}_{\mathbb{P}}(\xi\, \mathbb{E}_{\mathbb{P}}(h(X) \,|\, \mathcal{G}))$$

for all random variables $\xi : \Omega \to \mathbb{R}$ with $\mathbb{E}_{\mathbb{P}}(|\xi|) < \infty$ and for all bounded and continuous functions $h : \mathbb{R}^d \to \mathbb{R}$.

(ii) *We say that X_n converges $\mathcal{G}$-mixing to X as $n \to \infty$, if X_n converges $\mathcal{G}$-stably to X as $n \to \infty$, and $\mathbb{P}^{X | \mathcal{G}} = \mathbb{P}^X$ $\mathbb{P}$-almost surely, where $\mathbb{P}^X$ denotes the distribution of X on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ under $\mathbb{P}$. Equivalently,* *we can say that* X_n *converges* $\mathcal{G}$*-mixing to* X *as* $n \to \infty$*, if* X_n *converges* $\mathcal{G}$*-stably to* X *as* $n \to \infty$*, and* $\sigma(X)$ *and* $\mathcal{G}$ *are independent, which equivalently means that*

$$\lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}(\xi\, \mathbb{E}_{\mathbb{P}}(h(X_n)\,|\,\mathcal{G})) = \mathbb{E}_{\mathbb{P}}(\xi)\, \mathbb{E}_{\mathbb{P}}(h(X))$$

for all random variables $\xi : \Omega \to \mathbb{R}$ *with* $\mathbb{E}_{\mathbb{P}}(|\xi|) < \infty$ *and for all bounded and continuous functions* $h : \mathbb{R}^d \to \mathbb{R}$*.*

In Definition A.1, $\mathbb{P}^{X_n\,|\,\mathcal{G}}$, $n \in \mathbb{N}$, and $\mathbb{P}^{X\,|\,\mathcal{G}}$ are the $\mathbb{P}$-almost surely unique $\mathcal{G}$-measurable Markov kernels from $(\Omega, \mathcal{F})$ to $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that for each $n \in \mathbb{N}$,

$$\int_G \mathbb{P}^{X_n\,|\,\mathcal{G}}(\omega, B)\, \mathbb{P}(\mathrm{d}\omega) = \mathbb{P}(X_n^{-1}(B) \cap G) \qquad \text{for all } G \in \mathcal{G},\ B \in \mathcal{B}(\mathbb{R}^d).$$

and

$$\int_G \mathbb{P}^{X\,|\,\mathcal{G}}(\omega, B)\, \mathbb{P}(\mathrm{d}\omega) = \mathbb{P}(X^{-1}(B) \cap G) \qquad \text{for all } G \in \mathcal{G},\ B \in \mathcal{B}(\mathbb{R}^d),$$

respectively. For the notion of weak convergence of Markov kernels towards a Markov kernel, see Häusler and Luschgy [11, Definition 2.2]. For more details on stable convergence, see Häusler and Luschgy [11, Chapter 3 and Appendix A]. In particular, it turns out that X_n converges $\mathcal{G}$-stably to X as $n \to \infty$ if and only if $\lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}(\xi h(X_n)) = \mathbb{E}_{\mathbb{P}}(\xi h(X))$ for all $\mathcal{G}$-measurable random variables $\xi : \Omega \to \mathbb{R}$ with $\mathbb{E}_{\mathbb{P}}(|\xi|) < \infty$ and for all bounded and continuous functions $h : \mathbb{R}^d \to \mathbb{R}$ (following from Theorem 3.17 in Häusler and Luschgy [11]). Furthermore, X_n converges $\mathcal{G}$-mixing to X as $n \to \infty$ if and only if $\lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}(\xi h(X_n)) = \mathbb{E}_{\mathbb{P}}(\xi)\, \mathbb{E}_{\mathbb{P}}(h(X))$ for all $\mathcal{G}$-measurable random variables $\xi : \Omega \to \mathbb{R}$ with $\mathbb{E}_{\mathbb{P}}(|\xi|) < \infty$ and for all bounded and continuous functions $h : \mathbb{R}^d \to \mathbb{R}$ (following from Corollary 3.3 in Häusler and Luschgy [11]).

Next, we recall four results about stable convergence of random variables, which play important roles in the proofs of Theorem 1.4 and Corollary 1.10.

Theorem A.1 [Häusler and Luschgy [11, Theorem 3.18]] *Let* X_n*,* $n \in \mathbb{N}$*,* X*,* Y_n*,* $n \in \mathbb{N}$*, and* Y *be* $\mathbb{R}^d$*-valued random variables on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and* $\mathcal{G} \subset \mathcal{F}$ *be a sub-*σ*-field. Assume that* $X_n \to X$ $\mathcal{G}$*-stably as* $n \to \infty$*.*

(a) *If* $\|X_n - Y_n\| \xrightarrow{\mathbb{P}} 0$ *as* $n \to \infty$*, then* $Y_n \to X$ $\mathcal{G}$*-stably as* $n \to \infty$*.*

(b) *If* $Y_n \xrightarrow{\mathbb{P}} Y$ *as* $n \to \infty$*, and* Y *is* $\mathcal{G}$*-measurable, then* $(X_n, Y_n) \to (X, Y)$ $\mathcal{G}$*-stably as* $n \to \infty$*.*

(c) *If* $g : \mathbb{R}^d \to \mathbb{R}^d$ *is a Borel-measurable function such that* $\mathbb{P}^X(\{x \in \mathbb{R}^d : g \text{ is not continuous at } x\}) = 0$*, then* $g(X_n) \to g(X)$ $\mathcal{G}$*-stably as* $n \to \infty$*. Here recall that* $\mathbb{P}^X$ *denotes the distribution of* X *on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *under* $\mathbb{P}$*.*

Theorem A.2 [Häusler and Luschgy [11, Corollary 3.19]] *Let* X_n*,* $n \in \mathbb{N}$*, and* X *be* $\mathbb{R}^d$*-valued random variables on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and* $\mathcal{G} \subset \mathcal{F}$ *be a sub-*σ*-field. Let* $\mathcal{E} \subset \mathcal{G}$ *be closed under finite intersections such that* $\Omega \in \mathcal{E}$ *and* $\sigma(\mathcal{E}) = \mathcal{G}$*, where* $\sigma(\mathcal{E})$ *denotes the* σ*-algebra generated by* $\mathcal{E}$*. Then the following statements are equivalent:*

(i) $X_n \to X$ $\mathcal{G}$*-stably as* $n \to \infty$*,*
(ii) $\lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}(\mathbb{1}_F \mathrm{e}^{\mathrm{i}\langle u, X_n\rangle}) = \mathbb{E}_{\mathbb{P}}(\mathbb{1}_F \mathrm{e}^{\mathrm{i}\langle u, X\rangle})$ *for all* $F \in \mathcal{E}$ *and* $u \in \mathbb{R}^d$*,*
(iii) $\langle u, X_n\rangle \to \langle u, X\rangle$ $\mathcal{G}$*-stably for all* $u \in \mathbb{R}^d$*.*

Theorem A.3 [Häusler and Luschgy [11, Theorem 3.21]] *Let* $X_{n,r}$*,* X_r*,* $n, r \in \mathbb{N}$*,* X*, and* Y_n*,* $n \in \mathbb{N}$*, be* $\mathbb{R}^d$*-valued random variables on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and* $\mathcal{G} \subset \mathcal{F}$ *be a sub-*σ*-field. Assume that*

(i) *for each* $r \in \mathbb{N}$*, we have* $X_{n,r} \to X_r$ $\mathcal{G}$*-stably as* $n \to \infty$*,*
(ii) $X_r \to X$ $\mathcal{G}$*-stably as* $r \to \infty$*,*

(iii) $\lim_{r\to\infty} \limsup_{n\to\infty} \mathbb{P}(\|\boldsymbol{X}_{n,r} - \boldsymbol{Y}_n\| > \varepsilon) = 0$ *for all* $\varepsilon > 0$.

Then $\boldsymbol{Y}_n \to \boldsymbol{X}$ $\mathcal{G}$*-stably as* $n \to \infty$.

Theorem A.4 [Häusler and Luschgy [11, Proposition 3.24]] *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. Suppose that* $\mathbb{P} = \sum_{i=1}^{\infty} p_i \mathbb{Q}_i$, *where* $\mathbb{Q}_i$, $i \in \mathbb{N}$, *is a probability measure on* $(\Omega, \mathcal{F})$ *and* $p_i \in [0, 1]$, $i \in \mathbb{N}$, *satisfying* $\sum_{i=1}^{\infty} p_i = 1$. *Let* $\boldsymbol{X}_n$, $n \in \mathbb{N}$, *and* $\boldsymbol{X}$ *be* $\mathbb{R}^d$*-valued random variables on* $(\Omega, \mathcal{F}, \mathbb{P})$. *If* $\boldsymbol{X}_n$ *converges* $\mathcal{G}$*-stably to* $\boldsymbol{X}$ *under* $\mathbb{Q}_i$ *as* $n \to \infty$ *for each* $i \in \mathbb{N}$ *satisfying* $p_i > 0$, *then* $\boldsymbol{X}_n$ *converges* $\mathcal{G}$*-stably to* $\boldsymbol{X}$ *under* $\mathbb{P}$ *as* $n \to \infty$.

Finally, we recall Lévy's upwards theorem used in the proof of Theorem 1.4.

Theorem A.5 [Lévy's upwards theorem] *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, and let* ξ *be a real-valued random variable such that* $\mathbb{E}_{\mathbb{P}}(|\xi|) < \infty$ *and* $(\mathcal{F}_n)_{n\in\mathbb{Z}_+}$ *be a filtration with* $\mathcal{F}_\infty := \sigma\Big(\bigcup_{n\in\mathbb{Z}_+} \mathcal{F}_n\Big)$. *Then*

$$\mathbb{E}_{\mathbb{P}}(\xi \,|\, \mathcal{F}_n) \xrightarrow{\mathbb{P}\text{-a.s.}} \mathbb{E}_{\mathbb{P}}(\xi \,|\, \mathcal{F}_\infty) \quad \text{as } n \to \infty, \qquad \text{and} \qquad \mathbb{E}_{\mathbb{P}}(\xi \,|\, \mathcal{F}_n) \xrightarrow{L_1(\mathbb{P})} \mathbb{E}_{\mathbb{P}}(\xi \,|\, \mathcal{F}_\infty) \quad \text{as } n \to \infty.$$

We note that Theorem A.5 sometimes is called Lévy's zero-one law as well, since if $\xi = \mathbb{1}_A$, where $A \in \mathcal{F}_\infty$, then it yields that $\mathbb{P}(A \,|\, \mathcal{F}_n) \xrightarrow{\mathbb{P}\text{-a.s.}} \mathbb{1}_A$ as $n \to \infty$, where the limit can be zero or one.

Acknowledgements

We would like to thank the referee for the comments that helped us improve the paper.

References

[1] Aït-Sahalia, Y. and Jacod, J. (2014). High-Frequency Financial Econometrics. Princeton University Press, Princeton.
[2] Barczy, M. (2022). Stable convergence of conditional least squares estimators for supercritical continuous state and continuous time branching processes with immigration. ArXiv 2207.14056
[3] Barczy, M. and Pap, G. (2021). Mixing convergence of LSE for supercritical Gaussian AR(2) processes using random scaling. ArXiv 2101.01590
[4] Basse-O'Connor, A., Heinrich, C. and Podolskij, M. (2019). On limit theory for functionals of stationary increments Lévy driven moving averages. Electronic Journal of Probability **24(79)** 1–42.
[5] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). Regular variation. Cambridge University Press, Cambridge.
[6] Biscio, C. A. N., Poinas, A. and Waagepetersen, R. (2018). A note on gaps in proofs of central limit theorems. Statistics and Probability Letters **135** 7–10.
[7] Bolthausen, E. (1982). On the central limit theorem for stationary mixing random fields. The Annals of Probability **10(4)** 1047–1050.
[8] Crimaldi, I. (2009). An almost sure conditional convergence result and an application to a generalized Pólya urn. International Mathematical Forum **4(23)** 1139–1156.
[9] Crimaldi, I., Letta, G. and Pratelli, L. (2007). A strong form of stable convergence. In Séminaire de Probabilités XL, 203–225, Lecture Notes in Mathematics, 1899, Springer, Berlin.
[10] Gradshteyn, I. S. and Ryzhik, I. M. (2015). Table of Integrals, Series, and Products, 7th edition. Academic Press, San Diego.
[11] Häusler, E. and Luschgy, H. (2015). Stable Convergence and Stable Limit Theorems. Springer, Cham.
[12] Horn, R. A. and Johnson, Ch. R. (1985). Matrix Analysis. Cambridge University Press, Cambridge.
[13] Jacod, J. and Protter, Ph. E. (2012). Discretization of Processes. Springer, Heidelberg.
[14] Kotz, S. and Nadarajah, S. (2004). Multivariate *t*-Distributions and Their Applications. Cambridge University Press, Cambridge.
[15] Podolskij, M. and Vetter, M. (2010). Understanding limit theorems for semimartingales: a short survey. Statistica Neerlandica **64(3)** 329–351.
[16] Rényi, A. (1950). Contributions to the theory of independent random variables. Acta Mathematica Academiae Scientiarum Hungaricae **1** 99–108.
[17] Rényi, A. (1958). On mixing sequences of sets. Acta Mathematica Academiae Scientiarum Hungaricae **9(1–2)** 215–228.
[18] Rényi, A. (1963). On stable sequences of events. Sankhyā. Series A **25** 293–302.
[19] Rényi, A. and Révész, P. (1958). On mixing sequences of random variables. Acta Mathematica Academiae Scientiarum Hungaricae **9(3–4)** 389–393.
[20] Sato, K.-i.(1999). Lévy processes and infinitely divisible distributions. Cambridge University Press, Cambridge, 1999.