A central limit theorem for stochastic recursive sequences of topical operators
Glenn Merlet

To cite this version:

HAL Id: hal-00082113
https://hal.archives-ouvertes.fr/hal-00082113v3
Submitted on 30 Oct 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A CENTRAL LIMIT THEOREM FOR STOCHASTIC RECURSIVE SEQUENCES OF TOPICAL OPERATORS

By Glenn Merlet

CEREMADE, Université Paris–Dauphine

Let \((A_n)_{n\in\mathbb{N}}\) be a stationary sequence of topical (i.e., isotone and additively homogeneous) operators. Let \(x(n, x_0)\) be defined by 
\[x(0, x_0) = x_0\text{ and } x(n+1, x_0) = A_n x(n, x_0).\]
It can model a wide range of systems including train or queuing networks, job-shop, timed digital circuits or parallel processing systems.

When \((A_n)_{n\in\mathbb{N}}\) has the memory loss property, \((x(n, x_0))_{n\in\mathbb{N}}\) satisfies a strong law of large numbers. We show that it also satisfies the CLT if \((A_n)_{n\in\mathbb{N}}\) fulfills the same mixing and integrability assumptions that ensure the CLT for a sum of real variables in the results by P. Billingsley and I. Ibragimov.

1. Model. An operator \(A: \mathbb{R}^d \rightarrow \mathbb{R}^d\) is called additively homogeneous if it satisfies \(A(x + a1) = Ax + a1\) for all \(x \in \mathbb{R}^d\) and \(a \in \mathbb{R}\), where \(1\) is the vector \((1, \ldots, 1)'\) in \(\mathbb{R}^d\). It is called isotone if \(x \leq y\) implies \(Ax \leq Ay\), in which the order is the product order on \(\mathbb{R}^d\). It is called topical if it is isotone and homogeneous. The set of topical operators on \(\mathbb{R}^d\) will be denoted by \(\text{Top}_d\).

We recall that the action of matrices with entries in the semiring \(\mathbb{R}_{\max} = (\mathbb{R} \cup \{-\infty\}, \max, +)\) on \(\mathbb{R}^d_{\max}\) is defined by \((Ax)_i = \max_j (A_{ij} + x_j)\). When matrix \(A\) has no \(-\infty\)-row, this formula defines a topical operator, also denoted by \(A\). Such operators are called max-plus operators and operators composition corresponds to the product of matrices in the max-plus semiring.

Let \((A_n)_{n\in\mathbb{N}}\) be a sequence of random topical operators on \(\mathbb{R}^d\). A stochastic recursive sequence (SRS) driven by stochastic recursive sequence is a sequence \((X_n)_{n\in\mathbb{N}}\) satisfying equation \(X_{n+1} = A_n X_n\). To study such sequences,
we define \( (x(n,x_0))_{n \in \mathbb{N}} \) by
\[
\begin{align*}
  x(0,x_0) &= x_0, \\
  x(n+1,x_0) &= A_n x(n,x_0).
\end{align*}
\]

(1)

This class of system can model a wide range of situations. A review of applications can be found in the last section of [5]. When the \( x(n, \cdot) \)'s are daters, the isotonicity assumption expresses the causality principle, whereas the additive homogeneity expresses the possibility to change the origin of time. (See Gunawardena and Keane [18], where topical functions were introduced.) The max-plus case has, for instance, been applied to model queueing networks (Mairesse [26], Heidergott [20]), train networks (Heidergott and De Vries [21] and Braker [9]) or job-shop (Cohen, Dubois, Quadrat and Viot [11]). It also computes the daters of some task resources models (Gaubert and Mairesse [16]) and timed Petri nets including events graphs (Baccelli [1]) and 1-bounded Petri nets (Gaubert and Mairesse [15]). The role of the max operation is to synchronize different events. For developments on the max-plus modeling power, see Baccelli, Cohen, Olser and Quadrat [2] or Heidergott, Olser and van der Woude [22].

To clarify things, let us introduce a simple example.

**Example 1.1.** Our process assembles two parts. The \( n \)th time it is done, it takes time \( a_3(n) \). The parts are prepared separately, which respectively takes times \( a_1(n) \) and \( a_2(n) \). Then, they are sent from the preparation places to the assemblage place, which takes times \( t_1(n) \) and \( t_2(n) \) respectively. Once the assemblage place has finished an operation, it asks for new parts. At that time, if a preparation place has a ready part, it sends it and starts preparing another one. Otherwise, it finishes the one it is processing, sends it, and immediately starts preparing another one. This is summed up in Figure 1.

We denote by \((X_n)_1\) and \((X_n)_2\) the starting date of preparation of the \( n \)th part of each type and by \((X_n)_3\) the starting date of the \( (n-1) \)th assembly.

Sequence \((X_n)_{n \in \mathbb{N}}\) is ruled by equations
\[
\begin{align*}
  (X_{n+1})_1 &= \max((X_n)_1 + a_1(n),(X_n)_3), \\
  (X_n)_2 &= ((X_n)_2 + a_1(n),(X_n)_3)
\end{align*}
\]

and
\[
\begin{align*}
  (X_{n+1})_3 &= a_3(n) + \max((X_n)_1 + t_1(n),(X_n)_2 + t_2(n)) \\
  &= \max((X_n)_1 + t_1(n) + a_3(n-1) + a_1(n), \\
  &\quad (X_n)_2 + t_2(n) + a_3(n-1) + a_2(n), \\
  &\quad (X_n)_3 + t_1(n) \lor t_2(n) + a_3(n-1)).
\end{align*}
\]
in which we recognize equation (1), with $A_n$ defined by the action in the max-plus algebra of

$$
A(n) = \begin{pmatrix}
  a_1(n) & -\infty & a_3(n) \\
  -\infty & a_2(n) & a_3(n) \\
  a_1(n) + t_1(n) & a_2(n) + t_2(n) & t_1(n) \lor t_2(n) + a_3(n-1)
\end{pmatrix}.
$$

We assume that the sequence $(A(n))_{n \in \mathbb{N}}$ is stationary and ergodic.

We will focus on the asymptotic behavior of $(x(n, \cdot))_{n \in \mathbb{N}}$. It follows from Theorem 2.1, due to Vincent, that $(\frac{1}{n} \bigvee_i x_i(n, X^0))_{n \in \mathbb{N}}$ converges to a limit $\gamma$.

In many cases, if the modeled system is closed, then every sequence of coordinate $(x_i(n, X^0))_{n \in \mathbb{N}}$ also tends to $\gamma$, by Theorem 2.2. The so-called cycle time $\gamma$ is the inverse of the network’s throughput or the inverse of the production system’s output, as in Example 1.1. Therefore, there have been many attempts to estimate it (Cohen [12], Gaujal and Jean-Marie [17], Resig et al. [30]). Even when the $A_n$’s are i.i.d. and take only finitely many values, approximating $\gamma$ is NP-hard (Blondel, Gaubert and Tsitsiklis [7]). Hong and his coauthors have obtained [3, 4, 14] analyticity of $\gamma$ as a function of the law of $A_1$. They did so under the so-called memory loss property (MLP) introduced by Mairesse to ensure some stability of $(x(n, \cdot))_{n \in \mathbb{N}}$ (see [26]).

We prove another type of stability under the same assumptions. If $(A_n)_{n \in \mathbb{N}}$ has the MLP, then $(x(n, \cdot))_{n \in \mathbb{N}}$ actually satisfies a central limit theorem (CLT) under the same mixing and integrability hypotheses as the real variables in the CLT for sum of stationary variables by Billingsley [6] and Ibragimov [24].

As far as we know, two CLTs have already been proved for this type of sequences: one in [30] and the other in [28]. The most obvious improvement is that both assumed that the $A_n$’s are i.i.d. Moreover, in [30], the hypothesis is difficult to check and the $A_n$’s are max-plus operators defined by almost surely bounded matrices. In [28] the main hypothesis is also the
MLP, but integrability hypotheses were stronger, except for a subclass of topical operators.

The remainder of this article is divided into two sections. In Section 2 we define the memory loss property, present some law of large number type results and state our central limit theorems. In Section 3 we prove the theorems. First, we state the CLT for subadditive processes by Ishitani [25], then we check that \((\forall i, x_i(n, 0))_{n \in \mathbb{N}}\) satisfies each of its hypotheses. To this aim, we use Mairesse’s construction of the stationary version of the SRS, as well as different results from ergodic theory, depending on the hypothesis. We eventually deduce the results on \((x(n, \cdot))_{n \in \mathbb{N}}\) from those on \((\bigvee_i x_i(n, 0))_{n \in \mathbb{N}}\).

2. Presentation.

2.1. Memory loss property. Dealing with homogeneous operators, it is natural to introduce the quotient space of \(\mathbb{R}^d\) by the equivalence relation \(\sim\) defined by \(x \sim y\) if \(x - y\) is proportional to \(1 = (1, \ldots, 1)^t\). This space will be called projective space and denoted by \(\mathbb{P}\mathbb{R}^d_{\max}\). Moreover, \(x\) will be the equivalence class of \(x\).

The function \(x \mapsto (x_i - x_j)_{i < j}\) embeds \(\mathbb{P}\mathbb{R}^d_{\max}\) onto a subspace of \(\mathbb{R}^{(d(d-1))/2}\) with dimension \(d-1\). The infinity norm of \(\mathbb{R}^{(d(d-1))/2}\) therefore induces a distance on \(\mathbb{P}\mathbb{R}^d_{\max}\), which will be denoted by \(\delta\). A direct computation shows that \(\delta(\tau, \gamma) = \bigvee_i (x_i - y_i) + \bigvee_i (y_i - x_i)\). By a slight abuse, we will also write \(\delta(x, y)\) for \(\delta(\tau, \gamma)\). The projective norm of \(x\) will be \(|x|_p = \delta(x, 0) = \bigvee_i x_i - \bigwedge_i x_i\).

Let us recall two well-known facts about topical operators. First, a topical operator is nonexpanding with respect to the infinity norm (Crandall and Tartar [13]). Second, the operator it defines from \(\mathbb{P}\mathbb{R}^d_{\max}\) to itself is nonexpanding for \(\delta\) (Mairesse [26]).

The key property for our proofs is below:

**Definition 2.1 (MLP).**

1. A topical operator \(A\) is said to have rank 1 if it defines a constant operator on \(\mathbb{P}\mathbb{R}^d_{\max} : Ax\) does not depend on \(x \in \mathbb{R}^d\).
2. A sequence \((A_n)_{n \in \mathbb{N}}\) of Top\(_d\)-valued random variables is said to have the memory loss property (MLP) if there exists an \(N\) such that \(A_N \cdots A_1\) has rank 1 with positive probability.

This notion has been introduced by Mairesse [26], with the \(A_n\)’s as max-plus operators. In this case, the denomination rank 1 is natural.

We have proved in [27] that this property is generic for i.i.d. max-plus operators: it is fulfilled when the support of the law of \(A_1\) is not included in the union of finitely many affine hyperplanes, and in the discrete case the atoms of the probability measure are linearly related.
This result applied to Example 1.1 states that the sequence \((A_n)_{n \in \mathbb{N}}\) has the MLP provided that the support of \((a_1(2), a_2(2), t_1(2), t_2(2), a_3(1))\) is not included in a union of finitely many affine hyperplanes of \(\mathbb{R}^5\). This is not completely straightforward because the matrix \(A(1)\) is defined by only 5 variables, but the detailed result (see Remark 5.1 in [27]) shows that the linear forms on \(\mathbb{R}^{3 \times 3}\) that define the hyperplanes are not canceled by \(A(1)\), because of the \(-\infty\) entries.

In [28], we have proved that if the \(A_n\)'s are i.i.d. and the sequence has the MLP, then \((x(n, X^0))_{n \in \mathbb{N}}\) satisfies the same limit theorem as a sum of i.i.d. real variables. Here we prove that it still satisfies the CLT if the \(A_n\)'s are mixing quick enough. Quick enough means that the \(A_n\)'s satisfy the same integrability and mixing hypothesis as the real variables in the CLT for the sum of stationary variables by Billingsley [6] and Ibragimov [24]. Moreover, this proves the CLT under weaker integrability condition than in [28].

2.2. Law of large numbers. There have been many papers about the law of large numbers for products of random max-plus matrices since its introduction by Cohen [12]. We can, for instance, cite Baccelli [1], the most recent paper by Bousch and Mairesse [8] and Merlet [29] (in French). The latter article gives results for a larger class of topical operators, called uniformly topical.

Vincent [31] proved a law of large number for topical operators that will do in our case. He noticed that \((\bigvee_i x_i(n, 0))_{n \in \mathbb{N}}\) [resp. \((\bigwedge_i x_i(n, 0))_{n \in \mathbb{N}}\)] is subadditive (resp. superadditive), which leads to the following:

**Theorem 2.1** (Vincent [31]). Let \((A_n)_{n \in \mathbb{N}}\) be a stationary ergodic sequence of topical operators and \(X^0\) an \(\mathbb{R}^d\)-valued random variable. If \(A_{1,0}\) and \(X^0\) are integrable, then there exist \(\gamma\) and \(\overline{\gamma}\) in \(\mathbb{R}\) such that
\[
\lim_{n} \frac{\bigvee_i x_i(n, X^0)}{n} = \gamma \quad \text{a.s. and in } \mathbb{L}^1,
\]
\[
\lim_{n} \frac{\bigwedge_i x_i(n, X^0)}{n} = \overline{\gamma} \quad \text{a.s. and in } \mathbb{L}^1.
\]

Baccelli and Mairesse give a condition to ensure \(\overline{\gamma} = \gamma\), hence, the convergence of \((x(n, X^0)/n)_{n \in \mathbb{N}}\):

**Theorem 2.2** (Baccelli and Mairesse [5]). Let \((A_n)_{n \in \mathbb{N}}\) be a stationary ergodic sequence of topical operators and \(X^0\) an \(\mathbb{R}^d\)-valued random variable such that \(A_{1,0}\) and \(X^0\) are integrable. If there exists an \(N\), such that \(A_N \cdots A_1\) has a bounded projective image with positive probability, then there exists \(\gamma\) in \(\mathbb{R}\) such that
\[
\lim_{n} \frac{x(n, X^0)}{n} = \gamma 1 \quad \text{a.s. and in } \mathbb{L}^1.
\]
That being the case, $\gamma$ is called the Lyapunov exponent of the sequence. Since matrices with rank 1 have a bounded projective image, any ergodic sequence $(A_n)_{n\in\mathbb{N}}$ with the MLP fulfills the hypotheses of Theorem 2.2.

2.3. Statements of the results. Let us state the definitions of mixing to be used in the sequel.

**Definition 2.2 (Mixing).** We denote by $F_n$ the $\sigma$-algebra generated by the $A_k$’s for $k \leq n$ and by $F^*_n$ the one generated by the $A_k$’s for $k \geq n$. We define $\alpha_n$ and $\phi_n$ by the following:

1. $\phi(F,G) = \sup \{ |P(A \cap B) - P(A)P(B)| : A \in F, B \in G \}$ and $\phi_n = \sup_k \phi(F_k, F^{k+n})$.
2. $\alpha(F,G) = \sup \{ |P(A \cap B) - P(A)P(B)| : A \in F, B \in G \}$ and $\alpha_n = \sup_k \alpha(F_k, F^{k+n})$.

**Theorem 2.3.** If $(A_n)_{n\in\mathbb{N}}$ has the MLP and satisfies one of the following hypotheses:

A. $A_10 \in L^2$ and $\sum_{n=1}^{\infty} \sqrt{\phi_n} < +\infty$,
B. $A_10 \in L^{2+\delta}$ and $\sum_{n=1}^{\infty} \alpha_n^{\delta/(2+\delta)} < +\infty$ for some $\delta > 0$,
C. $A_10 \in L^\infty$ and $\sum_{n=1}^{\infty} \alpha_n < +\infty$,

then

$$\frac{1}{\sqrt{n}}(x(n, X^0) - n\gamma 1) \xrightarrow{\mathcal{L}} \mathcal{N}1,$$

where $\mathcal{N}$ is a random variable with zero-mean Gaussian law (or Dirac measure in $0$) whose variance does not depend on $X^0$, and $\xrightarrow{\mathcal{L}}$ denotes the convergence in law.

Moreover, if $X^0$ is integrable, then the variance $\sigma$ of $\mathcal{N}$ is given by

$$\lim_{n \to +\infty} \frac{1}{\sqrt{n}} \mathbb{E} \left| \sqrt{i} x_i(n, X^0) - n\gamma \right| = \left( \frac{2\sigma^2}{\pi} \right)^{1/2},$$

and $\sigma = 0$ if and only if the sequence $(x(n, X^0) - n\gamma 1)_{n\in\mathbb{N}}$ is tight.

**Remark 2.1 (I.i.d. case).** When the $A_n$ are i.i.d., I gave more precise results about $\sigma$ in [28]. In that case, if $\psi$ is a topical function from $\mathbb{R}^d$ to $\mathbb{R}$, such that $\sup_x |\psi(A_1 x) - \psi(x)|$ has a second moment or if $A_1 0$ has a $(6 + \varepsilon)$th moment and $X^0$ has a $(3 + \varepsilon)$th moment, then:

- $\sigma^2 = \lim_{n} \frac{1}{n} \mathbb{E}(\psi(x(n, X^0)) - n\gamma)^2$,
- $\sigma = 0$ iff there is a $\theta \in \text{Top}_d$ with rank 1 such that, for any $A$ in the support $S_A$ of $A_1$ and any $\theta'$ with rank 1 in the semi-group $T_A$ generated by $S_A$, we have

$$\theta A \theta' = \theta \theta' + \gamma 1.$$
I also proved that if there is such a $\theta$, then every $\theta \in T_A$ with rank 1 has this property.

Moreover, when the $A_n$ are defined by matrices in the max-plus algebra, $\sigma$ is positive provided that the support of $A_1$ is not included in a union of finitely many hyperplanes of $\mathbb{R}^{d \times d}$.

In this paper's framework it is not possible to express $\sigma^2$ as a limit like in the i.i.d. case, because the stationary random variables in Ishitani’s proof of Theorem 3.1 are not necessarily $L^2$ (see Section 3.8).

3. Proofs.

3.1. Results of Ishitani. We use the results of Ishitani [25] for mixing subadditive processes, which we state now:

Let $(\Omega, \mathcal{F}, T, \mathbb{P})$ be an ergodic measurable dynamical system, and $(\mathcal{F}_b^a)_{a,b \in \mathbb{N}}$ a family of sub $\sigma$-algebras of $\mathcal{F}$, such that $\mathcal{F}_b^{a+1} = T^{-1} \mathcal{F}_a^b$, and for any $a \leq c \leq d \leq b$, $\mathcal{F}_c^d \subset \mathcal{F}_a^b$. The family $(x_{st})_{s<t}$ of random variables is adapted if, for any $s,t$, $x_{st}$ is $\mathcal{F}_t^s$-measurable. It is subadditive (resp. submultiplicative) for any $s<t<u$, $x_{su} \leq x_{st} + x_{tu}$ (resp. $x_{su} \leq x_{st} \cdot x_{tu}$).

Theorem 3.1 (Ishitani [25] and Hall and Heyde [19]). Assume $(x_{st})_{s<t}$ is adapted and subadditive. We set $\mathcal{F}_n = \mathcal{F}_0^n$ and $\mathcal{F}_n^\infty = \mathcal{F}_n^\infty$, and define $\alpha_n$ and $\phi_n$ like in Definition 2.2. We set $(p, \theta)$ as follows:

(a) $(p, \theta) = (2, 2)$ if $\sum_{n=1}^{\infty} \sqrt{\phi_n} < +\infty$.
(b) $(p, \theta) = (2 + \delta, \delta)$ if $\sum_{n=1}^{\infty} \alpha_n^{\delta/(2+\delta)} < +\infty$ for some $\delta > 0$.
(c) $(p, \theta) = (+\infty, 1)$ if $\sum_{n=1}^{\infty} \alpha_n < +\infty$.

If the following hypotheses are satisfied:

1. $\lim_{t \to \infty} \frac{\mathbb{E}(x_{0t}) - \gamma}{\sqrt{t}} = 0$, where $\gamma = \inf_1 \mathbb{E}(x_{0t})$,
2. $\forall t \in \mathbb{N}, |x_{0t} - x_{1t}| \leq \Psi$, where $\Psi \in \mathbb{L}^p$,
3. $\sum_n \sup_t \|x_{0t} - x_{1t} - \mathbb{E}(x_{0t} - x_{1t} | \mathcal{F}_0^t)\|_\theta < \infty$,

then

$$\frac{1}{\sqrt{n}} (x_{0n} - n\gamma) \overset{L}{\to} \mathcal{N},$$

where $\mathcal{N}$ is a zero-mean Gaussian law (or a Dirac measure in 0).

Moreover, the variance $\sigma$ of $\mathcal{N}$ is given by

$$\lim_{n \to +\infty} \frac{1}{\sqrt{n}} \mathbb{E}|x_{0n} - n\gamma| = \left(\frac{2\sigma^2}{\pi}\right)^{1/2}.$$
In the sequel we take \( \Omega = \text{Top}_d^Z \), \( T \) the shift and \( \mathbb{P} \) such that the law of \((A_n)_{n \in \mathbb{N}} \) is the image of \( \mathbb{P} \) by the projection on the positive coordinates. From now on, \( A_n \) is the projection on the \( n \)-th coordinate, and we denote \( A_0 \) by \( A \), so that \( A_n = A \circ T^n \).

For any \( s < t \), we set \( x_{st} = V_i(A_{t-1} \cdots A_s 0)_i \), and \( F^t_s = \sigma(A_s, \ldots, A_{t-1}) \), so that \((x_{st})_{s,t \in \mathbb{N}}\) is adapted to \((F^t_s)_{a,b \in \mathbb{N}}\). Vincent has noticed in [31] that \((x_{st})_{s,t \in \mathbb{N}}\) is subadditive. From now on we check that it satisfies hypotheses 1–3 with \((p, \theta) = (2, 2)\) under hypothesis A, \((p, \theta) = (2 + \delta, \frac{\delta}{2 + \delta})\) under hypothesis B and \((p, \theta) = (\infty, 1)\) under hypothesis C.

Since \( x \mapsto V_i(A_{t-1} \cdots A_1 x)_i \) is topical, \( \bigwedge_i (A_0)_i \leq A0 \leq \bigvee_i (A_0)_i \) implies

\[
\bigwedge_i (A_0)_i \leq A_0 - \bigvee_i (A_0)_i.
\]

Therefore, we can take \( \Psi = |A0|_\infty \) and hypothesis 2 of Theorem 3.1 is checked. In the sequel we check the other two hypotheses.

3.2. Bound on \( \mathbb{E}(x_{0t}) - t\gamma \). It is well known and easy to check that, for any \( A \in \text{Top}_d \) and \( x \in \mathbb{R}^d \), the quantity \( \bigvee_i (Ax)_i - \bigvee_i x_i \) only depends on \( A \) and \( \varpi \). We denote it by \( \xi(A, \varpi) \). With this notation, we have

\[
\bigvee_i x_i(n, X^0)_i - \bigvee_i (X^0)_i = \sum_{k=0}^{n-1} \xi(A_k, \varpi(k, X^0))_i.
\]

It follows from the main theorem of [26]—which can be extended without difficulty from max-plus to topical operators—that there is a choice \( Y \) of \( X^0 \), such that \( \varpi(n, Y) = \varpi \circ T^n \). In this case, we see that \( \xi(A_k, \varpi(k, Y)) = \xi(A, Y) \circ T^k \), therefore, \( \bigvee_i x_i(n, Y)_i - \bigvee_i Y_i \) is the partial sum of the stationary sequence \( \xi(A, Y) \circ T^k \) \( k \in \mathbb{N} \).

Let us assume for a while that \( Y \) is integrable. Then, so is \( \xi(A, Y) \), because \( A0 + \bigwedge_i Y_i \leq AY \leq A0 + \bigvee_i Y_i \) implies

\[
|\xi(A, Y)| \leq \left| \bigvee_i (A0)_i \right| + |Y|_\varpi.
\]

Therefore, it follows from equation (3) with \( X^0 = Y \) that

\[
\mathbb{E} \left( \bigvee_i x_i(n, Y) \right) - \mathbb{E} \left( \bigvee_i Y_i \right) = n\mathbb{E}(\xi(A, Y)).
\]

Since topical functions are nonexpanding, we have \( |\bigvee_i x_i(n, Y) - x_{0n}| \leq \|Y\|_\infty \), therefore, \( |\mathbb{E}(x_{0n}) - n\mathbb{E}(\xi(A, Y))| \leq 2\mathbb{E}(\|Y\|_\infty) \). In that case \( \gamma = \mathbb{E}(\xi(A, Y)) \) and hypothesis 1 of Theorem 3.1 follows from the integrability of \( Y \).

The end of the subsection is devoted to the proof of the bounds that will give this integrability.
First, we recall from Mairesse’s proof that there is almost surely an \( n \in \mathbb{N} \) such that \( rk(A_{-1} \cdots A_{-n}) = 1 \) and that, for such an \( n \), \( Y = A_{-1} \cdots A_{-n} 0 \). In the sequel we denote by \( N \) the smallest such \( n \).

Since \( (V_i x_i(n, 0))_{n \in \mathbb{N}} \) [resp. \( (A_i x_i(n, 0))_{n \in \mathbb{N}} \)] is subadditive (resp. superadditive), we have, for any \( n \in \mathbb{N} \) and \( i \in [1, d] \),

\[
\sum_{k=1}^{n} \bigwedge_i (A_{-k} 0)_i \leq (A_{-1} \cdots A_{-n} 0)_i \leq \sum_{k=1}^{n} \bigvee_i (A_{-k} 0)_i,
\]

therefore, \( |A_{-1} \cdots A_{-n} 0|_p \leq \sum_{k=1}^{n} |A_{-k} 0|_p \) and

\[
|Y|_p \leq \sum_{n \in \mathbb{N}} 1_{\{N=n\}} \sum_{k=1}^{n} |A_{-k} 0|_p
= \sum_{k \in \mathbb{N}^*} 1_{\{N \geq k\}} |A_{-k} 0|_p
= \sum_{k \in \mathbb{N}^*} 1_{\{rk(A_{-1} \cdots A_{-k+1}) \neq 1\}} |A_{-k} 0|_p.
\]

Finally, we get

\[
(5) \quad \|Y\|_1 \leq \|Y\|_\theta \leq \sum_k \|1_{\{rk(A_{-1} \cdots A_{-k}) \neq 1\}} |A_0|_p \|_\theta.
\]

The finiteness of the right part of this inequality with 1 instead of \( \theta \) would be enough to check hypothesis 1, but the finiteness of this quantity also ensures hypothesis 3, as will be shown in the next section. Finally, Sections 3.4 to 3.6 will be devoted to the proof of the finiteness under each hypothesis of Theorem 2.3.

3.3. Bound on \( \|x_{0t} - x_{1t} - \mathbb{E}(x_{0t} - x_{1t} | F_0^n)\|_\theta \). We denote by \( \Delta_{t_n}^t \) the quantity \( |x_{0t} - x_{1t} - \mathbb{E}(x_{0t} - x_{1t} | A_0, \ldots, A_n)| \). If \( t \leq n \), then \( \Delta_{n}^t = 0 \). From now on, we assume \( t \geq n \).

First, it follows from equation (2) and the \( F_0^n \)-measurability of \( A_0 \) that

\[
(6) \quad \bigwedge_i (A_0)_i \leq \mathbb{E}(x_{0t} - x_{1t} | F_0^n) \leq \bigvee_i (A_0)_i
\]

and \( \Delta_{n}^t \leq |(A_0)|_p \).

Second, if \( rk(A_{-n+1} \cdots A_1) = 1 \), then

\[
x_{0t} - x_{1t} - (x_{0n} - x_{1n}) = \xi(A_{-1} \cdots A_n, A_{-n+1} \cdots A_0)
= -\xi(A_{-1} \cdots A_n, A_{-n+1} \cdots A_1) = 0,
\]

where \( \xi \) is the same function as in equation (3). Therefore, we have

\[
(8) \quad 1_{\{rk(A_{-n+1} \cdots A_1) = 1\}} (x_{0t} - x_{1t}) = 1_{\{rk(A_{-n+1} \cdots A_1) = 1\}} (x_{0n} - x_{1n})
\]
and
\[
\mathbb{1}_{\{rk(A_{n-1} \cdots A_1) = 1\}} \mathbb{E}(x_{0t} - x_{1t} | \mathcal{F}_0^n) = \mathbb{E}(\mathbb{1}_{\{rk(A_{n-1} \cdots A_1) = 1\}} (x_{0t} - x_{1t}) | \mathcal{F}_0^n)
\]
\[
= \mathbb{E}(\mathbb{1}_{\{rk(A_{n-1} \cdots A_1) = 1\}} (x_{0n} - x_{1n}) | \mathcal{F}_0^n)
\]
\[
= \mathbb{1}_{\{rk(A_{n-1} \cdots A_1) = 1\}} (x_{0n} - x_{1n}).
\]  
(9)

Equations (8) and (9) together imply that \( \mathbb{1}_{\{rk(A_{n-1} \cdots A_1) = 1\}} \Delta_n^t = 0 \), and finally, we have

\[
\Delta_n^t = 1_{\{rk(A_{n-1} \cdots A_1) \neq 1\}} \Delta_n^t \leq 1_{\{rk(A_{n-1} \cdots A_1) \neq 1\}} |(A0)|_{P}.
\]  
(10)

It follows from equations (5) and (10) that \((x_{st})\) satisfies hypotheses 1 and 3 of Theorem 3.1, provided that

\[
\sum_{n=1}^{\infty} \|1_{\{rk(A_{n-1} \cdots A_1) \neq 1\}} |(A0)|_{P} \|_{\theta} < \infty.
\]  
(11)

The next three subsections will prove that relation (11) is satisfied, under each of the hypotheses of Theorem 2.3.

3.4. Finiteness under hypothesis A. From the definition of \(\phi\), we see that, for \(X \in L^1(\mathcal{F})\) and \(Y \in L^\infty(\mathcal{G})\), \(|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq \phi(\mathcal{F}, \mathcal{G})\|X\|_1\|Y\|_\infty\).

We apply this inequality with \(X = |A0|_{P}^2\) and \(Y = 1_{\{rk(A_{n-1} \cdots A_{n/2+1}) \neq 1\}}\), where \(n/2\) is the integer part of the half of \(n\), and we take the square root. We get

\[
\|1_{\{rk(A_{n-1} \cdots A_1) \neq 1\}} |(A0)|_{P}\|_2 \\
\leq \|1_{\{rk(A_{n-1} \cdots A_{n/2+1}) \neq 1\}} |(A0)|_{P}\|_2 \\
\leq \sqrt{\mathbb{P}(rk(A_{n/2} \cdots A_1) \neq 1)} \||(A0)|_{P}\|_2 + \sqrt{\phi_{n/2}} \||(A0)|_{P}\|_2.
\]

The \(\sqrt{\phi_{n/2}}\)'s are summable by hypothesis A. Let us see that the \(\sqrt{\mathbb{P}(rk(A_{n/2} \cdots A_1) \neq 1)}\)'s are too. For any integers \(n\) and \(n_0\), we have the following inequality:

\[
\mathbb{P}(rk(A_{n+2n_0} \cdots A_1) \neq 1) \\
\leq \mathbb{E}(1_{\{rk(A_{n} \cdots A_1) \neq 1\}} 1_{\{rk(A_{n+2n_0} \cdots A_{n+n_0+1}) \neq 1\}}) \\
\leq (\phi_{n_0} + \mathbb{E}(1_{\{rk(A_{n+2n_0} \cdots A_{n+n_0+1}) \neq 1\}})) \mathbb{E}(1_{\{rk(A_{n} \cdots A_1) \neq 1\}}) \\
\leq (\phi_{n_0} + \mathbb{P}(rk(A_{n_0} \cdots A_1) \neq 1)) \mathbb{P}(rk(A_{n} \cdots A_1) \neq 1).
\]

Taking \(n_0\) big enough, we have \((\phi_{n_0} + \mathbb{P}(rk(A_{n_0} \cdots A_1) \neq 1)) < 1\), hence \(\mathbb{P}(rk(A_{n} \cdots A_1) \neq 1)\) decreases exponentially fast and \(\sum_{n \in \mathbb{N}} \sqrt{\mathbb{P}(rk(A_{n/2} \cdots A_1) \neq 1)} < \infty\). This concludes the proof of hypotheses 1 and 3 under hypothesis A.
3.5. Finiteness under hypothesis B. Let us take $X = |A_0|^{(2+\delta)/(1+\delta)}$, $Y = 1_{\{rk(A_{n-1}\cdots A_1)\neq 1\}}$ and $\varrho = \frac{1}{2+\delta}$ in the mixing inequality (see, e.g., [19]) which states for any $X \in L^1(\mathcal{F})$ and $Y \in L^\infty(\mathcal{G})$

$$|E(XY) - E(X)E(Y)| \leq 6\varrho^{-1/q}(\mathcal{F}, \mathcal{G})\|X\|_q\|Y\|_\infty$$

and let us elevate it to the power $\frac{1}{2+\delta}$. We get

$$\|1_{\{rk(A_{n-1}\cdots A_1)\neq 1\}}|A_0|_p\|^{(2+\delta)/(1+\delta)} \leq \|1_{\{rk(A_{n-1}\cdots A_{n/2+1})\neq 1\}}|A_0|_p\|^{(2+\delta)/(1+\delta)}$$

$$\leq \mathbb{P}(rk(A_{n/2}\cdots A_1) \neq 1)(1+\delta)/(2+\delta) ||A_0|_p\|^{(2+\delta)/(1+\delta)}$$

$$+ (6\varrho_{n/2})^{\delta/(2+\delta)} ||A_0|_p\|_{2+\delta}.$$  

The $(6\varrho_{n/2})^{\delta/(2+\delta)}$’s are summable by hypothesis B. To see that the

$$\mathbb{P}(rk(A_{n/2}\cdots A_1) \neq 1)(1+\delta)/(2+\delta)$$ too, we apply the following lemma from [23] with $\lambda = \frac{2+\delta}{\delta}$ and $M_{st} = 1_{\{rk(A_{n-1}\cdots A_1)\neq 1\}}$.

**Lemma 3.1 (Hennion [23])**. Let $(M_{st})_{s<t}$ be submultiplicative and adapted with values in $[0, 1]$ such that $\lim_{n} \mathbb{E}(M_{0n}) = 0$. If $\sum_{n} \alpha_n^{1/\lambda} < \infty$, then there exists $c \in \mathbb{R}$, such that

$$\mathbb{E}(M_{0n}) \leq c \left(\frac{\ln^2 n}{n}\right)^{\lambda}.$$  

This concludes the proof of hypotheses 1 and 3 under hypothesis B.

3.6. Finiteness under hypothesis C. We notice that

$$\sum_{k} \|1_{\{rk(A_{k-1}\cdots A_1)\neq 1\}}|A_0|_p\|_{1} \leq \sum_{k} \|1_{\{rk(A_{k-1}\cdots A_1)\neq 1\}}\|_{1} ||A_0|_p\|_{\infty}$$

$$= ||A_0|_p\|_{\infty} \sum_{k} \mathbb{P}(R \geq k)$$

$$= ||A_0|_p\|_{\infty} \mathbb{E}(R),$$

where $R = \min\{n|rk(A_{n-1}\cdots A_1) = 1\}$. Moreover, if $\mathbb{P}(rk(A_{n_0}\cdots A_1) \neq 1) < 1$, then $R - n_0$ is bounded from above by the hitting time of $\{rk(A_{n_0}\cdots A_1) = 1\}$. The integrability of $R$ will follow from the next theorem due to Chazottes.

**Theorem 3.2 (Chazottes [10])**. Let $(\Omega, \mathcal{F}, \mathbb{P}, T)$ be a measurable dynamical system, $B \in \mathcal{F}$ a set with positive probability, and $1_B$ its indicator function. If the mixing coefficients $\alpha_n$ of the sequence $(1_B \circ T^n)_{n \in \mathbb{N}}$ satisfy $\sum_{n} \alpha_n < \infty$, then the hitting time of $B$ is integrable.
To apply the theorem, we notice that, when \( B = \{ \text{rk}(A_{n_0} \cdots A_1) = 1 \} \), every \( \alpha_n \) defined by \( (1_B \circ T^n)_{n \in \mathbb{N}} \) is less than the \( \alpha_{n-n_0} \) defined by \( (A_n)_{n \in \mathbb{N}} \). This ensures the hypothesis of Theorem 3.2 and concludes the proof of hypotheses 1 and 3 under hypothesis C.

3.7. Conclusion of the proof. In the last six subsections we have proved that, under hypothesis A, B or C of Theorem 2.3, \( (\bigvee_i x_i(n,0))_{n \in \mathbb{N}} \) satisfies the hypotheses of Theorem 3.1. Therefore, we have

\[
\frac{1}{\sqrt{n}} \left( \bigvee_i x_i(n,0) - n\gamma \right) \xrightarrow{L} \mathcal{N}.
\]

Since topical functions are nonexpanding

\[
\frac{1}{\sqrt{n}} \left| \bigvee_i x_i(n,0) - \frac{1}{\sqrt{n}} \bigvee_i x_i(n,0) \right| \leq \frac{1}{\sqrt{n}} ||X^0||_{\infty} \xrightarrow{P} 0,
\]

therefore, \( \frac{1}{\sqrt{n}} (\bigvee_i x_i(n,0) - n\gamma) \xrightarrow{L} \mathcal{N} \).

\( \delta_n = (x(n,X^0) - \bigvee_i x_i(n,X^0))_1 \) is a function of \( x(n,X^0) \), which is converging in law (and even in total variation) by the main theorem of [26], therefore, \( \frac{1}{\sqrt{n}} \delta_n \xrightarrow{P} 0 \) and \( \frac{1}{\sqrt{n}} (x(n,X^0) - n\gamma 1) \xrightarrow{L} \mathcal{N}1 \), which concludes the proof of the convergence in law.

Inequality (12) also implies that

\[
\frac{1}{\sqrt{n}} E \left| \bigvee_i x_i(n,X^0) - \bigvee_i x_i(n,0) \right| \leq \frac{1}{\sqrt{n}} E ||X^0||_{\infty} \xrightarrow{P} 0,
\]

so that

\[
\lim_{n \to +\infty} \frac{1}{\sqrt{n}} E \left| \bigvee_i x_i(n,X^0) - n\gamma \right| = \left( \frac{2\sigma^2}{\pi} \right)^{1/2}
\]

follows from \( \lim_{n \to +\infty} \frac{1}{\sqrt{n}} E |x_{0n} - n\gamma| = \left( \frac{2\sigma^2}{\pi} \right)^{1/2}. \)

3.8. Tightness. Without loss of generality, we assume \( \gamma = 0 \). (Otherwise, just replace \( A_n \) by \( A_n - \gamma \).) One part of the equivalence is obvious. To prove the other part, we have to go into the proof of Theorem 3.1. Ishitani constructs a random variable \( Z \) (named \( y_{01} \) in [25]) and approximates \( x_{0n} \) by the Birkhoff sum \( S_n = \sum_{k=0}^{n-1} Z \circ T^k (y_{0n} \text{ in [25]}) \). Then he shows that \( (S_n)_{n \in \mathbb{N}} \) fulfills the hypotheses of Billingsley–Ibragimov’s CLT.

In Billingsley–Ibragimov’s CLT, the asymptotic variance is zero if and only if \( Z \) is a coboundary, that is, if there is a random variable \( f \) such that \( Z = f \circ T - f \). (See, e.g., [19].)
Let us assume we are in this situation and identify $Z$. It is built as a kind of Cesaro type limit of the sequence $(x_{0n} - x_{1n})_{n \in \mathbb{N}}$. But in our situation equation (7) says that this sequence is ultimately constant and that $X_n$ is equal to the limit as soon as $rk(A_n \cdots A_1) = 1$.

Let us denote by $R$ the smallest such $n$ and by $\psi$ the topical function that maps $x \in \mathbb{R}^d$ to $\bigvee_i x_i$. The random variable $R$ is almost surely finite because of ergodicity and MLP. With notation, we have

$$Z = x_{0R} - x_{1R} = \psi(A_R \cdots A_00) - \psi(A_R \cdots A_10) \quad \text{a.s.}$$

and, for any integer $n$ such that $rk(A_n \cdots A_1) = 1$,

$$f \circ T - f = \psi(A_n \cdots A_00) - \psi(A_n \cdots A_10).$$

In the sequel we deduce the tightness from equation (13). As a first and main step, let us show that $(\psi(A_R \cdots A_n0))_{n \in \mathbb{N}}$ is tight. For any $k \in \mathbb{N}$, since $rk(A_R \cdots A_1) = 1$, $rk(A_R \cdots A_{-k}) = 1$ and equation (13) holds for $n = R \circ T^k + k$. Compounded by $T^{-k}$, it becomes

$$f \circ T^{-k+1} - f \circ T^{-k} = \psi(A_R \cdots A_{-k}0) - \psi(A_R \cdots A_{-k+1}0).$$

Summing over $k$, we get

$$f \circ T - f \circ T^{-n} = \psi(A_R \cdots A_{-n}0) - \psi(A_R \cdots A_00),$$

from which the tightness of $(\psi(A_R \cdots A_n0))_{n \in \mathbb{N}}$ is obvious.

The tightness of $(A_{-1} \cdots A_{-n}0)_{n \in \mathbb{N}}$ is obvious too, because the sequence converges in law.

From those two tightnesses, we successively deduce the tightness of the following sequences:

- $(\psi(A_{-1} \cdots A_{-n}0))_{n \in \mathbb{N}}$, because equation (4) implies that
  $$|\psi(A_R \cdots A_{-n}0) - \psi(A_{-1} \cdots A_{-n}0)| = |\xi(A_R \cdots A_{0}, \overline{A_{-1} \cdots A_{-n}0})| \leq |\psi(A_R \cdots A_00)| + |\overline{A_{-1} \cdots A_{-n}0}|.$$  

- $((\psi(A_{-1} \cdots A_{-n}0, A_{-1} \cdots A_{-n}0)))_{n \in \mathbb{N}}$, again because $(A_{-1} \cdots A_{-n}0)_{n \in \mathbb{N}}$ is tight.

- $(A_{-1} \cdots A_{-n}0)_{n \in \mathbb{N}}$, because $x \mapsto (\psi(x), \overline{F})$ is a bi-Lipschitz homeomorphism from $\mathbb{R}^d$ to $\mathbb{R} \times \mathbb{R}_{\max}$ (see [28]).

- $(x(n,0))_{n \in \mathbb{N}}$, because, for any $n \in \mathbb{N}$, the random variables $(A_{-1} \cdots A_{-n}0)$ and $x(n,0)$ have the same law.

- Eventually $(x(n, X^0))_{n \in \mathbb{N}}$, because the $A_n$ are nonexpanding and, therefore, we have $\|x(n, X^0) - x(n,0)\|_\infty \leq \|X^0\|_\infty$.

**Acknowledgments.** This work was done while I was a JSPS post-doctoral fellow at Keio university. I gratefully thanks H. Ishitani for introducing me to the article [25] and for interesting discussions, and M. Keane for his useful rereading.
REFERENCES


CEREMADE
Université Paris-Dauphine
Place du Maréchal de Lattre de Tassigny
75775 Paris cedex 16
France
E-mail: glenn.merlet@gmail.com