Vector-valued Almost Sure Invariance Principles
For (Non)stationary And Random Dynamical Systems

Yaofeng Su*
July 6, 2021

Abstract
We study the limit behavior of (non)stationary and random chaotic dynamical systems. Several (vector-valued) almost sure invariance principles for (non)stationary dynamical systems and quenched (vector-valued) almost sure invariance principles for random dynamical systems are proved. We also apply our results to stationary chaotic dynamical systems, which admit Young towers, and to (non)uniformly expanding non-stationary and random dynamical systems with intermittencies or uniform spectral gaps. It implies that the systems under study tend to a Brownian motion under various scalings.

Contents
1 Introduction 2
2 Definitions, Notations, Main Theorems 3
3 Applications 9
4 A Purely Probabilistic Theorem 13
  4.1 Several Inequalities 15
  4.2 Proof of Theorem 4.1 25
5 Proof of Theorem 2.6 33

*Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA.
yfsu@math.uh.edu; yaoefeng.su@math.gatech.edu
1 Introduction

The paper deals with strong statistical properties of (non)stationary and random dynamical systems. Such problems naturally arise e.g. in a framework of the non-equilibrium statistical physics. A non-stationary dynamical system is generated by consecutive applications of maps $T_k \circ T_{k-1} \circ \cdots \circ T_1$ acting on a phase space $X$, where the maps $T_i : X \to X$ are allowed to vary with $i$. If $T_k = T_1$ for all $k \geq 1$, then we get a stationary dynamical system. In difference, in a random dynamical system the maps $T_\omega$ are picked from a probability space $(\Omega, \mathbb{P})$ in accordance with the probability distribution $\mathbb{P}$. In this case, trajectories in $X$ are formed by a sequence of concatenation of maps $T_{\sigma^n \omega} \circ \cdots \circ T_\omega$ where $\sigma$ is a shift of $\Omega$. As time evolves, $\sigma$ updates the current configuration and the dynamics $T_\omega$ on $X$. Various statistical properties of non-stationary and random dynamical systems were studied in [BB84, OSY09, HNTV17, CR07, NTV18, DFGTV18a]. In this paper we consider a more refined property, called a vector-valued almost sure invariance principle (VASIP), for a variety of (non)stationary and random dynamical systems.

Suppose $(X_k)_{k \geq 1}$ is a sequence of zero-mean random vectors. We say $(X_k)_{k \geq 1}$ satisfies the VASIP if there exists a sequence of independent zero-mean Gaussian random vectors $(G_k)_{k \geq 1}$ (perhaps on an extended probability space) such that the difference between $\sum_{k \leq n} X_k$ and $\sum_{k \leq n} G_k$ is negligible in comparison to $\left[ \text{Var}(\sum_{k \leq n} X_k) \right]^{\frac{1}{2}}$. Here $\lim_{n \to \infty} \text{Var}(\sum_{k \leq n} X_k) = \infty$.

One of our main results is a purely probabilistic Theorem 4.1 which proves the VASIP for $(X_k)_{k \geq 1}$ adapted to a decreasing filtration, in case when the growth rate of $\text{Var}(\sum_{k \leq n} X_k)$ could be sufficiently fast and polynomial.

We apply our Theorem 4.1 to prove the VASIP for non-stationary dynamical systems (Theorem 2.6) and random dynamical systems (Theorem 2.9).

As applications (section 3), we apply our Theorems 2.6 and 2.9

1. For a large class of non-stationary dynamical systems which were studied in [CR07, HNTV17, NTV18, HS18, HL18, NPT19, Su19a].

2. And for random dynamical systems considered in [DFGTV18a, DFGTV18b, NTV18, NPT19].

3. And for stationary Young towers in [MN05, MN09, You99].
Also our results on the VASIP improve the ones obtained in [Su19a, HNTV17, DFGTV18a, MN05, MN09, Gou10]. Namely,

1. The technique used in [Su19a, HNTV17, DFGTV18a] only works for random variables $X_k$ while ours works for random vectors $X_k$.

2. The papers [Gou10, HNTV17, DFGTV18a] require the system to have a very strong hyperbolicity. They are just corollaries of our results.

3. The papers [MN05, MN09, Gou10] dealt with different types of stationary systems (strong or weak hyperbolicity) using different techniques. They are just corollaries of our results.

4. The systems in [Su19a, HNTV17, DFGTV18a] have to assume a fast decay rate of $\lim_{n \to \infty} \sup_k \text{Cov}(X_k, X_{k+n}) = 0$ (also called a decay of correlation), while we require a slower rate only. In particular, [Su19a] considered the composition of a small set of Pomeau-Manneville type maps, obtained by perturbing the slope at the indifferent fixed point 0, while the results of the present paper hold for a larger (and actually optimal) family of such maps.

Some weaker than VASIP results for the same class as the one studied in [Su19a] were obtained in [NTV18, HS18, HL18, NPT19].

5. In the random setting, [NTV18, HS18, HL18, NPT19] considered the random composition of the same family of Pomeau-Manneville type maps and required the shift $\sigma$ on $\Omega$ to be Bernoulli, rather than our assumption that $\sigma$ is just ergodic.

The structure of the paper is the following one. In the next section 2 we introduce the necessary notations and formulate the main Theorems 2.6 and 2.9. The proofs of these theorems can be found in sections 5 and 6. In section 3 some corollaries and applications of our Theorems 2.6 and 2.9 are considered. Section 4 deals with a proof of a (purely probabilistic) Theorem 4.1. In section 5 and 6 this Theorem 4.1 is applied to prove the main Theorems 2.6 and 2.9. Section 7 contains the proofs of corollaries formulated in the section 3. Computations of the parameters considered in Theorems 2.6 and 4.1 are given in Appendix (section 8).

2 Definitions, Notations, Main Theorems

Consider a probability space $(X, \mathcal{B}, \mu)$ with $\mu$ as a reference probability and $\mathcal{B}$ as a $\sigma$-algebra of $X$, a map (also called a dynamics) $T : X \to X$ is called non-singular if

$$\mu(A) = 0 \iff \mu(T^{-1}A) = 0 \text{ for all } A \in \mathcal{B}. \,$$
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be another probability space with a probability \(\mathbb{P}\) and a \(\sigma\)-algebra \(\mathcal{F}\). In this paper, we consider the following (non)stationary and random dynamical systems only.

**Definition 2.1 (Non-stationary, stationary and random dynamical systems).**

\((X, \mathcal{B}, (T_k)_{k \geq 1}, \mu)\) is called a non-stationary dynamical system if \((T_k)_{k \geq 1}\) are non-singular maps on \(X\).

In contrast, a stationary dynamical system means that \(T_k = T_1\) for all \(k \geq 1\) and \((T_1)_\mu := \mu \circ T_1^{-1} = \mu\).

\((X, \mathcal{B}, (T_\omega)_{\omega \in \Omega}, \mu, \sigma, (\mu_\omega)_{\omega \in \Omega})\) is called a random dynamical system if

1. \(\sigma : \Omega \to \Omega\) is an invertible, ergodic probability-preserving map on \((\Omega, \mathbb{P})\).
2. The probability \(\mu_\omega\) is absolutely continuous w.r.t. \(\mu\).
3. \((T_\omega)_{\omega \in \Omega}\) are non-singular maps on \(X\) with respect to (w.r.t.) \(\mu\).
4. \((T_\omega)_\mu = \mu_\sigma\) a.e. \(\omega \in \Omega\).

\((\mu_\omega)_{\omega \in \Omega}\) and \((h_\omega)_{\omega \in \Omega}\) are called quasi-invariant probabilities and quasi-invariant densities respectively.

For any \(n, m, k \in \mathbb{N}\), we denote:

\[
T_{m+n} := T_n \circ T_{n-1} \circ \cdots \circ T_m, \quad T_{m-1} := \text{Id},
\]

\[
T_n := T_1^n = T_n \circ T_1 \circ \cdots \circ T_1, \quad T_0 := \text{Id},
\]

\[
T_{\omega}^k := T_{\sigma^{-k}\omega} \circ T_{\sigma^{-k-2}\omega} \circ \cdots \circ T_\omega, \quad T_{\omega}^0 := \text{Id}.
\]

The transfer operators (Perron-Frobenius operators) \(P_k\) (resp. \(P_\omega\)) associated to \(T_k\) (resp. \(T_\omega\)) are defined by the duality relations:

\[
\int g \cdot P_k(f) d\mu = \int g \circ T_k \cdot f d\mu \quad \text{for all } f \in L^1(\mu), g \in L^\infty(\mu),
\]

\[
\int g \cdot P_\omega(f) d\mu = \int g \circ T_\omega \cdot f d\mu \quad \text{for all } f \in L^1(\mu), g \in L^\infty(\mu).
\]  \(2.1\)

Observe that (2.1) implies

\[P_\omega h_\omega = h_{\sigma\omega} \text{ in } L^1(\mu) \text{ for a.e. } \omega \in \Omega.\]

Similar to \(T_{m+n}\), \(T^n\) and \(T^k\), we denote:

\[
P_{m+n} := P_n \circ P_{n-1} \circ \cdots \circ P_m, \quad P_{m-1} := \text{Id},
\]

\[
P_n := P_1^n = P_n \circ P_{n-1} \circ \cdots \circ P_1, \quad P_0 := \text{Id},
\]

\[
P_{\omega}^k := P_{\sigma^{-k}\omega} \circ P_{\sigma^{-k-2}\omega} \circ \cdots \circ P_\omega, \quad P_{\omega}^0 := \text{Id}.
\]
Notation 2.2.

1. $C_a$ denotes a constant that depends only on $a$.

2. $a_n \approx_w b_n$ (resp. “$a_n \lesssim_w b_n$”) means that there is a constant $C_w \geq 1$ such that $C_w^{-1} \cdot b_n \leq a_n \leq C_w \cdot b_n$ for all $n \in \mathbb{N}$ (resp. $a_n \leq C_w \cdot b_n$ for all $n \in \mathbb{N}$).

3. 1 denotes the constant function 1 on $X$.

4. For any $m \in \mathbb{N}$, scalar function $f$ and $L^1$-matrix $[f_{ij}]$ (i.e. $f_{ij} \in L^1(X)$ for all $i, j \geq 1$), define:

\[
    f \cdot P_m ([f_{ij}]) = P_m ([f_{ij}]) \cdot f := [f \cdot P_m (f_{ij})],
\]
\[
    f \cdot P_w ([f_{ij}]) = P_w ([f_{ij}]) \cdot f := [f \cdot P_w (f_{ij})].
\]

Next, we define the VASIP and quenched VASIP for the (non)stationary and random dynamical systems defined in the Definition 2.1.

Definition 2.3 (VASIP and Quenched VASIP).

For a non-stationary dynamical system $(X, \mathcal{B}, (T_k)_{k \geq 1}, \mu)$, consider a sequence of observables $\{\phi_k \in L^\infty(X, \mu; \mathbb{R}^d) : k \in \mathbb{N}\}$ satisfying

\[
    \sup_k ||\phi_k||_\infty < \infty, \int \phi_k \circ T^k d\mu = 0 \text{ for all } k \geq 1.
\]

We denote a $d \times d$ variance matrix by

\[
\sigma_n^2 := \int (\sum_{k=1}^n \phi_k \circ T^k) \cdot (\sum_{k=1}^n \phi_k \circ T^k)^T d\mu.
\]

and the least eigenvalue of $\sigma_n^2$ by

\[
\lambda(\sigma_n^2) := \inf_{||u||=1, u \in \mathbb{R}^d} \int (u^T \cdot \sum_{k=1}^n \phi_k \circ T^k)^2 d\mu.
\]

We say $(\phi_k \circ T^k)_{k \geq 1}$ satisfies a VASIP if there exists a constant $\epsilon \in (0, 1)$ and zero-mean $d$-dimensional independent Gaussian random vectors $(G_k)_{k \geq 1}$ defined on some extended probability space of $(X, \mathcal{B}, \mu)$ such that:

\[
\sum_{k=1}^n \phi_k \circ T^k - \sum_{k=1}^n G_k = o(\lambda(\sigma_n^2)^{1-\epsilon}) \text{ almost surely (a.s.)}, \quad (2.2)
\]
\[
\sigma_n^2 = \sum_{k=1}^n \tilde{E}(G_k \cdot G_k^T) + o(\lambda(\sigma_n^2)^{1-\epsilon}), \quad (2.3)
\]
\[ \lambda(\sigma_n^2) \to \infty, \]  
\hspace{1cm} (2.4)

where \( \tilde{E} \) in (2.3) is the expectation w.r.t. the probability \( \tilde{P} \) of the extended probability space of \((X, \mathcal{B}, \mu)\).

For a random dynamical system \((X, \mathcal{B}, (\tau_\omega)_{\omega \in \Omega}, \mu, \Omega, \mathcal{F}, \tilde{P}, \sigma, (\mu_\omega)_{\omega \in \Omega})\), consider a sequence of observables \(\{\phi_\omega \in L^\infty(X, \mu; \mathbb{R}^d) : \omega \in \Omega\}\) satisfying

\[
\sup_{\omega \in \Omega} ||\phi_\omega||_{\infty} < \infty, \int \phi_\omega d\mu_\omega = 0 \text{ for any } \omega \in \Omega,
\]

we denote a \(d \times d\) variance matrix by

\[
\sigma_n^2(\omega) := \int \left( \sum_{k=1}^n \phi_{\sigma^k \omega} \circ T_k^\omega \right) \cdot \left( \sum_{k=1}^n \phi_{\sigma^k \omega} \circ T_k^\omega \right)^T d\mu_\omega.
\]

We say \((\phi_{\sigma^k \omega} \circ T_k^\omega)_{k \geq 1, \omega \in \Omega} \) satisfies a quenched VASIP if, for a.e. \(\omega \in \Omega\), there exists a constant \(\epsilon \in (0, 1)\) and zero-mean \(d\)-dimensional independent Gaussian random vectors \((G_k^\omega)_{k \geq 1}\) defined on some extended probability space of \((X, \mathcal{B}, \mu_\omega)\) such that:

\[
\sum_{k=1}^n \phi_{\sigma^k \omega} \circ T_k^\omega - \sum_{k=1}^n G_k^\omega = o(n^{\frac{1-\epsilon}{2}}) \text{ a.s.,} \quad (2.5)
\]

\[
\sigma_n^2(\omega) = \sum_{k=1}^n \tilde{E}_\omega [G_k^\omega \cdot (G_k^\omega)^T] + o(n^{1-\epsilon}), \quad (2.6)
\]

\[
\sigma_n^2(\omega) \approx w n \cdot I_d \times d, \quad (2.7)
\]

where \(\tilde{E}_\omega\) in (2.6) is the expectation w.r.t. the probability \(\tilde{P}_\omega\) of the extended probability space of \((X, \mathcal{B}, \mu_\omega)\).

**Remark 2.4.** If \(\sigma_n^2 = n \cdot \Sigma^2 + o(n^{1-\epsilon})\) holds for a constant \(\epsilon \in (0, 1)\) and a positive definite \(d \times d\) matrix \(\Sigma^2\), then (2.2) becomes

\[
\sum_{k=1}^n \phi_k \circ T^k - \Sigma \cdot B_n = o(n^{\frac{1-\epsilon}{2}}) \text{ a.s.,}
\]

where \(B\) is a standard \(d\)-dimensional Brownian motion (see Lemma 8.4). Then our VASIP coincides with the standard VASIP for stationary dynamical systems proved in [MN09].

**Remark 2.5.** If \(d = 1\), then \(G_k\) can be embedded into a one-dimensional standard Brownian motion \((B_t)_{t \geq 0}\). Using (2.3), (2.2) becomes

\[
\sum_{k=1}^n \phi_k \circ T^k - B_{\sigma_n^2} = o(\sigma_n^{1-\epsilon}) \text{ a.s.}
\]
This implies statistical limit theorems such as the self-norming Central Limit Theorem (CLT) and the self-norming Law of the Iterated Logarithm (LIL):

\[
\lim_{n \to \infty} \int e^{-t \cdot \sum_{k=1}^{n} \phi_k \circ T_k^k} \sigma_n d\mu = e^{-t^2/2} \text{ for all } t \in \mathbb{R},
\]

\[
\limsup_{n \to \infty} \sum_{k=1}^{n} \phi_{k} \circ T_k \sqrt{\frac{\sigma_n}{n \log \log n}} = 1, \mu\text{-a.s.}
\]

Similarly, we have the quenched CLT and the quenched LIL, that is, there is a constant \( \Sigma^2 > 0 \) such that for a.e. \( \omega \in \Omega \),

\[
\lim_{n \to \infty} \int e^{-t \cdot \sum_{k=1}^{n} \phi_{k} \circ T_k^k \sigma_k^2} \sigma_n d\mu_{\omega} = e^{-t^2/2} \text{ for all } t \in \mathbb{R},
\]

\[
\limsup_{n \to \infty} \sum_{k=1}^{n} \phi_{k} \circ T_k^k \sqrt{\frac{\sigma_n}{n \log \log n}} = \Sigma, \mu_{\omega}\text{-a.s.}
\]

Now we can present our main results.

\textbf{Theorem 2.6 (Non-stationary Dynamical Systems).}

Assume a non-stationary dynamical system \((X, \mathcal{B}, (T_k)_{k \geq 1}, \mu)\) and a sequence of observables \((\phi_k)_{k \in \mathbb{N}}\) from the Definition 2.3, they satisfy the conditions \((A1)-(A3)\) below:

\[
\sup_{i \geq 1} \left\| P_{i+1}(\phi_i \cdot P^i 1) \right\| \leq n^{1-1/\alpha},
\]

\[
\sup_{i \geq 1} \int |P_{i+1}(\phi_i \cdot P^i 1) - \int \phi_i T_i \circ \phi_{i+1}^T T_i^i \circ T^i d\mu| \cdot P^i 1 |d\mu \leq n^{1-1/\alpha},
\]

\[
\sup_{i \geq 1, j \geq 0} \int |P_{i+j+1}(\phi_i \cdot P^i 1) \cdot \phi_{i+j}^T - P_{i+j}^i 1 \cdot \phi_{i+j}^T \circ P^i_1 \phi_{i+j} d\mu| \leq n^{1-1/\alpha},
\]

where \( \alpha \in (0, 1/2) \), \( \cdot \) is the Euclidean norm for vectors or matrices.

Then there is \( \gamma \in (0, 1) \) which depends on \( d, \alpha \) only (will be given in Appendix, Lemma 8.1), such that if \( \lambda(\sigma_n^2) \gtrsim n^\gamma \), then \((\phi_k \circ T_k)_{k \geq 1}\) satisfies the VASIP.

\textbf{Remark 2.7.} We assume that \( \alpha < 1/2 \) throughout this paper.

\textbf{Remark 2.8.} For a stationary dynamical system, that is, \( \phi_k := \phi, T_k = T \) for all \( k \geq 1 \), \((T)_*\mu = \mu \) and \( \int \phi d\mu = 0 \). In this case, \( \sigma_n^2 \approx n \cdot 1_{d \times d} \). We denote the transfer operator of \( T \) by \( P \) (see (2.1)). Then \( P^i 1 = 1 \) almost surely for any \( i \geq 1 \) and the assumptions \((A1)-(A3)\) become:

\[
\int |P^n(\phi)| d\mu \lesssim n^{1-1/\alpha},
\]
Theorem 2.9 (Random Dynamical Systems).

Assume a random dynamical system $\{X, B, (T_\omega)_{\omega \in \Omega}, \mu, \Omega, F, \mathbb{P}, \sigma, (\mu_\omega)_{\omega \in \Omega}\}$ and a sequence of observables $\{\phi_\omega \in L^\infty(X, \mu; \mathbb{R}^d) : \omega \in \Omega\}$ from the Definition 2.3, they satisfy the conditions (A1')-(A3') below:

\[
\int |P^n_{\sigma^i\omega}(\phi_{\sigma^i\omega} \cdot h_{\sigma^i\omega})|d\mu \leq Cn^{-1/\alpha}, \quad \text{(A1')}
\]

\[
\int |P^n_{\sigma^i\omega}([\phi_{\sigma^i\omega} \cdot \phi_{\sigma^i\omega}^T - \int \phi_{\sigma^i\omega} \cdot \phi_{\sigma^i\omega}^T d\mu_{\sigma^i\omega}] \cdot h_{\sigma^i\omega})|d\mu \leq Cn^{-1/\alpha}, \quad \text{(A2')}
\]

\[
\int |P^n_{\sigma^i+j\omega}\{[P^j_{\sigma^i\omega}(\phi_{\sigma^i\omega}h_{\sigma^i\omega})\phi_{\sigma^i+j\omega}^T - h_{\sigma^i+j\omega} \int P^j_{\sigma^i\omega}(\phi_{\sigma^i\omega}h_{\sigma^i\omega})\phi_{\sigma^i+j\omega}^T d\mu]\}|d\mu \leq Cn^{-1/\alpha}, \quad \text{(A3')}
\]

where $\alpha < 1/2$, $| \cdot |$ is the Euclidean norm for vectors or matrices, $C > 0$ is a constant which does not depend on $i, j, n, \omega$.

Then there are two linear subspaces (do not depend on $\omega$): $W_1, W_2 \subset \mathbb{R}^d$, $\mathbb{R}^d = W_1 \oplus W_2$ with projections $\pi_1 : W_1 \oplus W_2 \rightarrow W_1, \pi_2 : W_1 \oplus W_2 \rightarrow W_2$ such that

\[
\int |P^n(\phi \cdot \phi^T - \int \phi \cdot \phi^T d\mu)|d\mu \lesssim n^{1-1/\alpha},
\]

\[
\sup_{j \geq 0} \int |P^n[P_j(\phi) \cdot \phi^T - \int P_j(\phi) \cdot \phi^T d\mu]|d\mu \lesssim n^{1-1/\alpha}.
\]
1. \((\pi_1 \circ \sigma_k \circ T^k)_{k \geq 1, \omega \in \Omega}\) satisfies the quenched VASIP.

2. \((\pi_2 \circ \sigma_k \circ T^k)_{k \geq 1, \omega \in \Omega}\) is a coboundary, that is, there is \(\psi \in L^1(\Omega \times X, d\mu d\mathbb{P})\) such that:

\[
\pi_2 \circ \sigma_\omega (T_\omega x) = \psi(\sigma(\omega), T_\omega(x)) - \psi(\omega, x) \text{ a.e. } (\omega, x),
\]

where \(\int 1_A d\mu_d\mathbb{P} := \int \mu(A_\omega) d\mathbb{P}\) and \(A_\omega := \{x \in X : (\omega, x) \in A\}\) for any measurable set \(A \subseteq \Omega \times X\).

**Remark 2.10.**

1. Conditions \(A3\), \(A3'\), \(A6\) can be verified by the methods of invariant cones and tower extensions, which are shown in our Corollaries 3.1, 3.8 and 3.10.

2. The quasi-invariant density \(h_\omega\) is not required to be bounded away from zero.

3. \(\sigma: \Omega \to \Omega\) is ergodic only.

4. Note that \(\preceq\) in conditions \(A1'\)-\(A3'\) does not depend on all \(\omega \in \Omega\). A weaker case is studied in another paper \([Su19b]\).

## 3 Applications

The paper \([LSV99]\) considered a Pomeau-Manneville type map: for \(\beta > 0\),

\[
T_\beta(x) = \begin{cases} 
  x + 2\beta x^{1+\beta}, & 0 \leq x \leq 1/2 \\
  2x - 1, & 1/2 < x \leq 1
\end{cases}
\]

Equation (3.1)

It is proved that, for each \(\beta \in (0, 1)\), \(T_\beta\) preserves an absolutely continuous invariant probability. \([Gou04]\) proved that the CLT (2.8) holds for \(T_\beta\) only when \(\beta \in (0, 1/2)\). Now we consider the non-stationary case:

**Corollary 3.1** (Polynomially mixing non-stationary systems).

Consider a non-stationary dynamical system \(([0, 1], B, (T_k)_{k \geq 1}, m)\) in \([NTV18]\), where \(m\) is the Lebesgue measure, \(T_k := T_\beta_k\) (\(0 < \beta_k < \alpha < 1/2\)) for all \(k \geq 1\).

Assume that observables \((\phi_k)_{k \in \mathbb{N}} \subseteq \text{Lip}([0, 1]; \mathbb{R}^d)\) satisfy \(\sup_k \|\phi_k\|_{\text{Lip}} < \infty\) and \(\int \phi_k \circ T_k d\mu = 0\). Then there is \(\gamma \in (0, 1)\) (the same as in the Theorem 2.6), such that if \(\lambda(\sigma_n^\alpha) \succsim n^{\gamma}\), then \((\phi_i \circ T_i)_{i \geq 1}\) satisfies the VASIP.

If \(d = 1\), then there is \(\gamma_1 \in (0, 1)\) which depends on \(\alpha\) only (\(\gamma_1\) will be given in Appendix, Lemma 8.2), such that if \(\lambda(\sigma_n^\alpha) \succsim n^{\gamma_1}\), then \((\phi_i \circ T_i)_{i \geq 1}\) satisfies the self-norming CLT (2.8).
Remark 3.2. [NTV18] proved that, if $\alpha < 1/8$ and $\sigma_n^2$ grows with a sufficiently fast polynomial rate, then the self-norming CLT (2.8) holds for the observables $\phi_k \in C^1[0,1]$. [NPT19] extended it to $\alpha \in (0, 1/2)$ using Stein’s methods.

Corollary 3.3 (Exponentially mixing non-stationary systems).

Consider a non-stationary dynamical system $(X, \mathcal{B}, (T_k)_{k \geq 1}, \mu)$, assume that $(\mathcal{V}, || \cdot ||_\mathcal{V})$ is a $(P_k)_{k \geq 1}$-invariant Banach algebra contained in $(L^1, || \cdot ||_{L^1})$ and satisfies the following assumptions: there are constants $A > 0$, $\rho \in (0, 1)$ such that

1. $1 \in \mathcal{V}$.
2. $|| \cdot ||_\infty \leq A \cdot || \cdot ||_\mathcal{V}$.
3. For any $n, m \in \mathbb{N}$ and any $v \in \mathcal{V}$,
   $$||P_{m+1}v||_\mathcal{V} \leq A \cdot ||v||_\mathcal{V}.$$
4. For any $n, m \in \mathbb{N}$ and any $v \in \mathcal{V}_0 := \{v \in \mathcal{V} : \int v d\mu = 0\}$, we have
   $$||P_{m+1}v||_\mathcal{V} \leq A \cdot \rho^n \cdot ||v||_\mathcal{V}.$$

Assume that observables $(\phi_k)_{k \in \mathbb{N}} \subset \mathcal{V}$ satisfy $\int \phi_k \circ T^k d\mu = 0$ and $\sup_k ||\phi_k||_{\mathcal{V}} < \infty$. Then there is $\gamma \in (0, 1)$ (the same as in the Theorem 2.6), such that if $\lambda(\sigma_n^2) \gtrsim n^\gamma$, then $(\phi_k \circ T^k)_{k \geq 1}$ satisfies the VASIP.

Remark 3.4. We apply now this result for some dynamical systems considered in [CR07, HNTV17]:

1. Non-stationary observations on Axiom A dynamical systems in Corollary 6.2 of [HNTV17]: let $\mathcal{V} = C^{0,\beta}$ (a $\beta$-Hölder space), $T_k := T$ with $T_* \mu = \mu$, $\phi_k : X \to \mathbb{R}^d$ with $\sup_k ||\phi_k||_{C^{0,\beta}} < \infty$ and $\lambda(\sigma_n^2) \gtrsim n^{\max\{\gamma, \sqrt{7}/4\}}$ where $\gamma$ is in our Theorem 2.6. Here $\alpha$ in Theorem 2.6 is chosen to be an arbitrarily small positive number.

2. The systems in section 7 of [HNTV17] are essentially the same, so we just consider the “perturbed expanding maps $(T_k := T_{\epsilon_k})_{k \geq 1}$ of a fixed expanding map $T$ on the circle” in Theorem 7.4 to present our VASIP result: $\mathcal{V} := BV$, $d\mu := hdm$ is the SRB measure for $T$, $\phi_k := \phi - \int \phi \circ T^k d\mu : S^1 \to \mathbb{R}^d$ where $\phi \in \mathcal{V}$ is not a coboundary for $T$ and $\int \phi d\mu = 0$. By our Corollary 3.3 and Lemma 7.1 of [HNTV17], $(\phi_k \circ T^k := \phi \circ T^k - \int \phi \circ T^k d\mu)_{k \geq 1}$ has the VASIP. Moreover,

$$\sum_{k=1}^n \int \phi \circ T^k d\mu = \sum_{k=1}^n \int \phi \cdot P^k dm = \sum_{k=1}^n \int \phi \cdot [P^k - P^0] dm.$$
\[ \sum_{k=1}^{n} ||\phi||_{\mathcal{V}} \cdot ||P_k^1(h) - P_k^k(h)||_{L^1} \]

where \( P_k^k \) and \( P \) are the transfer operators of \( T_k^k \) and \( T \) respectively. Then by Lemma 2.13 in [CR07], we have

\[ \sup_n \left| \sum_{k=1}^{n} \int \phi \circ T_k^1 d\mu \right| = O(1). \]

So we have the same statement of the VASIP for \((\phi \circ T_k^k)_{k \geq 1}\) as in the Theorem 7.4 of [HNTV17].

Remark 3.5. Observe that conditions (Min) in [CR07] and (LB) in [HNTV17] are not required here. This observation applies to the stationary dynamical systems in [Gou10] which obtained a similar result: the VASIP holds without assuming conditions (Min) and (LB) (e.g., Rychlik maps [Ryc83] are this type of systems). Then [LM13] obtained the VASIP for interval maps with singularities, by proving the VASIP for induced Rychlik maps and employing tower techniques of [MT04] (see the Propositions 3.1 and 4.1 in [LM13]).

Corollary 3.6 (Exponentially mixing random systems).

Consider a random dynamical system \((X, \mathcal{B}, (T_\omega)_{\omega \in \Omega}, \mu, \Omega, \mathcal{F}, \mathbb{P}, \sigma)\) defined in the Definition 2.1, assume that \((B, || \cdot ||_B)\) is a \((P_\omega)_{\omega \in \Omega}\)-invariant Banach algebra contained in \((L^1, || \cdot ||_{L^1})\) and satisfies the following assumptions: there are constants \( A > 0, \rho \in (0, 1) \) such that

1. \( 1 \in B \).

2. The map \((\omega, x) \rightarrow (\omega, H(\omega, \cdot))(x)\) is \( \mathbb{P} \otimes \mu\)-measurable for every \( \mathbb{P} \otimes \mu\)-measurable function \( H \) such that \( H(\omega, \cdot) \in L^1(X, \mu) \) for a.e. \( \omega \in \Omega \).

3. \( || \cdot ||_{\infty} \leq A \cdot || \cdot ||_B \).

4. For any \( n \in \mathbb{N}, \omega \in \Omega \) and \( v \in B \),

\[ ||P_\omega^n v||_B \leq A \cdot ||v||_B. \]

5. For any \( n \in \mathbb{N}, \omega \in \Omega \) and any \( v \in B_0 := \{ v \in B : \int v d\mu = 0 \} \), we have

\[ ||P_\omega^n v||_B \leq A \cdot \rho^n \cdot ||v||_B. \]

Then there are functions \( h_\omega \in L^1 \) and quasi-invariant probabilities \( d\mu_\omega := h_\omega d\mu \) such that for a.e. \( \omega \in \Omega \), \( P_\omega h_\omega = h_\sigma \omega \), \( \sup_\omega ||h_\omega||_B < \infty \). Moreover, assume that observables \((\phi_\omega)_{\omega \in \Omega} \subset B\) satisfy \( \int \phi_\omega d\mu_\omega = 0 \) and \( \sup_\omega ||\phi_\omega||_B < \infty \). Then there are two linear subspaces (do not depend on \( \omega \)): \( W_1, W_2 \subset \mathbb{R}^d \), \( \mathbb{R}^d = W_1 \oplus W_2 \) with projections \( \pi_1 : W_1 \oplus W_2 \to W_1, \pi_2 : W_1 \oplus W_2 \to W_2 \) such that
1. \((\pi_1 \circ \phi_{\sigma^k, \omega} \circ T^k_{\omega})_{k \geq 1, \omega \in \Omega}\) satisfies the quenched VASIP.

2. \((\pi_2 \circ \phi_{\sigma^k, \omega} \circ T^k_{\omega})_{k \geq 1, \omega \in \Omega}\) is a coboundary: there is \(\psi \in L^2(\Omega \times X, d\mu, d\mathbb{P})\) such that:

\[
\pi_2 \circ \phi_{\sigma^k}(T^k_\omega x) = \psi(\sigma(\omega), T^k_\omega(x)) - \psi(\omega, x) \text{ a.e. } (\omega, x).
\]

**Remark 3.7.** [DFGTV18a] and [DFGTV18b] consider the same random dynamical systems, which satisfy the conditions of our Corollary 3.6. (H3), (H4) in [DFGTV18a] and (C4) in [DFGTV18b] are not required here. Our Corollary 3.6 works for the random dynamical systems in [DFGTV18a, DFGTV18b] including Random piecewise expanding maps in higher dimensions and Random Lasota-Yorke maps.

**Corollary 3.8** (Polynomially mixing random systems).

Consider a random system \(([0,1], \mathcal{B}, (T_\omega)_{\omega \in \Omega}, m, \Omega, \mathcal{F}, \mathbb{P}, \sigma)\) where \(m\) is the Lebesgue measure on \([0,1]\), \(\Omega := [0, 1/2]^\mathbb{Z}\), \(T_\omega := T_{\omega_0}\) are the Pomeau-Manneville type maps (3.1) which are picked from \(\{T_\beta : \beta \in [0,1/2]\}\) and \(\sigma : \Omega \to \Omega\) is an invertible ergodic left shift preserving a probability \(\mathbb{P}\) on \(\Omega\).

Then there are functions \(h_\omega \in L^1(m)\) and quasi-invariant probabilities \(d\mu_\omega := h_\omega dm\) such that \(P_\omega h_\omega = h_\sigma\) for a.e. \(\omega \in \Omega\). Moreover, Assume that observables \((\phi_\omega)_{\omega \in \Omega} \subset \text{Lip}([0,1]; \mathbb{R}^d)\) satisfy \(\sup_{\omega} ||\phi_\omega||_{\text{Lip}} < \infty\) and \(\int \phi_\omega d\mu_\omega = 0\). Then there are two linear subspaces (do not depend on \(\omega\)): \(W_1, W_2 \subset \mathbb{R}^d\), \(\mathbb{R}^d = W_1 \oplus W_2\) with projections \(\pi_1 : W_1 \oplus W_2 \to W_1, \pi_2 : W_1 \oplus W_2 \to W_2\) such that

1. \((\pi_1 \circ \phi_{\sigma^k, \omega} \circ T^k_{\omega})_{k \geq 1, \omega \in \Omega}\) satisfies the quenched VASIP.

2. \((\pi_2 \circ \phi_{\sigma^k, \omega} \circ T^k_{\omega})_{k \geq 1, \omega \in \Omega}\) is a coboundary: there is \(\psi \in L^1([0,1], d\mu, d\mathbb{P})\) such that:

\[
\pi_2 \circ \phi_{\sigma^k}(T^k_\omega x) = \psi(\sigma(\omega), T^k_\omega(x)) - \psi(\omega, x) \text{ a.e. } (\omega, x).
\]

**Remark 3.9.** \(\sigma : \Omega \to \Omega\) is ergodic only, which is weaker than the results in [NTV18, HS18, HL18, NPT19].

**Corollary 3.10** (Stationary dynamical systems).

Assume that a stationary dynamical system \((X, \mathcal{B}, T, \mu)\) (that is, \(T_* \mu = \mu\)) and a zero-mean observable \(\phi : X \to \mathbb{R}^d\) satisfy \((A_4)\), then there is a \(d \times d\) positive semi-definite matrix \(\Sigma^2\) and \(\epsilon \in (0,1)\), such that \(o_n^2 = n \cdot \Sigma^2 + o(n^{1-\epsilon})\). If the conditions \((A_4)-(A_6)\) are all satisfied, then there are two linear subspaces: \(W_1, W_2 \subset \mathbb{R}^d\) such that \(\mathbb{R}^d = W_1 \oplus W_2\) with projections \(\pi_1 : W_1 \oplus W_2 \to W_1, \pi_2 : W_1 \oplus W_2 \to W_2\), such that

1. \((\pi_1 \circ \phi \circ T^k)_{k \geq 1}\) satisfies the VASIP.
2. \((\pi_2 \circ \phi \circ T^k)_{k \geq 1}\) is a coboundary, that is, there is \(\psi \in L^1(X, d\mu)\) such that:
\[
\pi_2 \circ \phi(Tx) = \psi(Tx) - \psi(x) \text{ a.e.}
\]

In particular, if a dynamical system can be described by a Young tower \([You99]\), that is, \((\Delta, \mathcal{B}, F, v)\) with \(v \circ F^{-1} = v\), \(dv = \frac{dm}{dm} dm\) is exact, \(m\) is a reference measure on \(\Delta\), a return map \(R\) is defined on the base of the tower: \(\Delta_0 = \bigcup_{i \geq 1} \Delta_{0,i}\) such that \(R|_{\Delta_{0,i}} \equiv R_i \in \mathbb{N}, \gcd\{R_i\} = 1\), \(\int_{\Delta_0} Rdm < \infty\) and \(\Delta = \{(x, n) \in \Delta_0 \times \mathbb{N}_0 : n < R(x)\}\). \(F^R : \Delta_0 \to \Delta_0\) is a Gibbs-Markov map, satisfying
\[
\frac{|JF^R(x)|}{JF^R(y)} - 1 \lesssim \beta^{s(F^R(x), F^R(y))} \tag{3.2}
\]
where \(J\) is the Jacobian w.r.t. \(m\), \(\beta \in (0, 1)\), \(s(x, y)\) is the separation time defined on \(\Delta_0 \times \Delta_0\):
\[
s(x, y) := \min\{n \geq 0 : (F^R)^n(x), (F^R)^n(y) \text{ lie in distinct } \Delta_{0,i}\}.
\]
Meanwhile, we endow a metric \(d\) on \(\Delta\): for any \(z_1 = (x_1, n_1) \in \Delta, z_2 = (x_2, n_2) \in \Delta,
\[
d(z_1, z_2) := \begin{cases} \beta^{s(x_1, x_2)}, & n_1 = n_2 \\ 1, & n_1 \neq n_2 \end{cases}. \tag{3.3}
\]

Then for the stationary Young tower \((\Delta, \mathcal{B}, F, v)\) and any zero-mean observable \(\phi \in \text{Lip}(\Delta)\), all conditions \((A_4)-(A_6)\) are all satisfied. On the other hand, the stationary dynamical systems such as Pomeau-Manneville maps, Viana maps considered in \([MN05]\) and \([MN09]\) can be described by Young towers. Therefore we recover the VASIP for those systems.

Remark 3.11. Unlike our direct verification of conditions \((A_4)-(A_6)\), Melbourne and Nicol \([MN05, MN09]\) generalized the ideas from \([BP79, KP80]\) and used the Markov partitions of Young towers to prove the VASIP (e.g., see Theorem 2.8 in \([MN05]\)).

4 A Purely Probabilistic Theorem

Before giving the proofs of Theorems 2.6 and 2.9, we start with a purely probabilistic theorem.

Theorem 4.1. Let \((X_k)_{k \geq 1}\) be a sequence of random vectors in \(\mathbb{R}^d\) on a probability space \((X, \mathcal{B}, \mu)\) and \((\mathcal{E}_k)_{k \geq 1}\) be a decreasing filtration (i.e. \(\mathcal{E}_{k+1} \subseteq \mathcal{E}_k\) for all \(k \geq 1\)) such that \(X_k\) is \(\mathcal{E}_k\)-measurable.
We denote the conditional expectation w.r.t. \( \mathcal{E}_n \) and \( \mu \) by:
\[
\mathbb{E}_n(\cdot) := \mathbb{E}(\cdot | \mathcal{E}_n).
\]

In particular, the expectation (that is, the conditional expectation w.r.t. \( \{\emptyset, X\} \) and \( \mu \)) is denoted by:
\[
\mathbb{E}(\cdot) := \int (\cdot) d\mu.
\]

If \((X_k)_{k \geq 1}\) satisfy the conditions \((A0”)-(A3”)\) below:
\[
\sup_k ||X_k||_\infty < \infty \text{ and } \mathbb{E}X_k = 0, \quad \text{(A0“)}
\]
\[
\sup_{i \geq 1} \mathbb{E}||\mathbb{E}_{n+i}X_i|| \lesssim n^{1-1/\alpha}, \quad \text{(A1“)}
\]
\[
\sup_{i \geq 1} \mathbb{E}||\mathbb{E}_{n+i}[X_i \cdot X_i^T - \mathbb{E}(X_i \cdot X_i^T)]|| \lesssim n^{1-1/\alpha}, \quad \text{(A2“)}
\]
\[
\sup_{i \geq 1, j \geq 0} \mathbb{E}||\mathbb{E}_{n+i+j}[X_i \cdot X_{i+j}^T - \mathbb{E}(X_i \cdot X_{i+j}^T)]|| \lesssim n^{1-1/\alpha}, \quad \text{(A3“)}
\]
where \(\alpha \in (0, 1/2)\), \(\cdot\) is the Euclidean norm for vectors or matrices.

Define the \(d \times d\)-matrix
\[
\sigma_n^2 := \mathbb{E}\left(\sum_{i=1}^n X_i \cdot \left(\sum_{i=1}^n X_i\right)^T\right),
\]
and the least eigenvalue of \(\sigma_n^2\) by
\[
\lambda(\sigma_n^2) := \inf_{|u|=1, u \in \mathbb{R}^d} \int (u^T \cdot \sum_{k=1}^n X_k)^2 d\mu.
\]

Then there is \(\gamma \in (0, 1)\) which depends on \(d, \alpha\) only (will be given in Appendix, Lemma 8.1), such that if \(\lambda(\sigma_n^2) \gtrsim n^\gamma\), then there exists a constant \(\epsilon’ \in (0, 1)\) and zero-mean \(d\)-dimensional independent Gaussian random vectors \((G_k)_{k \geq 1}\) defined on some extended probability space of \((X, \mathcal{B}, \mu)\), they satisfy
\[
\sum_{k=1}^n X_k - \sum_{k=1}^n G_k = o((\lambda(\sigma_n^2))^{1-\epsilon’}) \text{ a.s.,} \quad (4.1)
\]
\[
\sigma_n^2 = \sum_{k=1}^n \bar{\mathbb{E}} G_k \cdot G_k^T + o((\lambda(\sigma_n^2))^{1-\epsilon’}), \quad (4.2)
\]
where \(\bar{\mathbb{E}}\) is the expectation w.r.t. the probability \(\bar{P}\) of the extended probability space of \((X, \mathcal{B}, \mu)\).

In the following subsections, we focus on the proof of Theorem 4.1.
4.1 Several Inequalities

In this subsection, we will obtain several inequalities derived from the conditions (A0”)-(A3”), which are needed to prove the Theorem 4.1.

Lemma 4.2. If the conditions (A0”) and (A1”) are satisfied, then there is a constant $C > 0$ such that for all $n, m \in \mathbb{N}$

$$| \int (\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k^T) d\mu | \leq Cn.$$

Proof.

$$\int (\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k^T) d\mu = \int \sum_{k=m}^{n+m-1} X_k \cdot X_k^T + \sum_{m \leq i < j \leq n+m-1} X_i \cdot X_j^T$$

$$+ (\sum_{m \leq i < j \leq n+m-1} X_i \cdot X_j^T)^T d\mu.$$

By (A0”), the equality above becomes

$$= O(n) + \sum_{m \leq i < j \leq n+m-1} \int (\mathbb{E}X_i) \cdot X_j^T + (\mathbb{E}X_i \cdot X_j^T)^T d\mu$$

$$\preceq n + \sum_{m \leq i < j \leq n+m-1} \int |(\mathbb{E}X_i)| d\mu.$$

By (A1”) and $\alpha \in (0, \frac{1}{2})$, the inequality above becomes

$$\preceq n + \sum_{m \leq i < j \leq n+m} (j - i)^{1-1/\alpha} = n + \sum_{1 \leq i < j \leq n} (j - i)^{1-1/\alpha}$$

$$\preceq n + \sum_{j=2}^{n} \sum_{1 \leq i < j} (j - i)^{1-1/\alpha} = O(n).$$

All constants in $\preceq, O(\cdot)$ do not depend on $m, n$. \hfill \Box

Lemma 4.3. If (A1”) is satisfied, then the following holds:

$$\sup_{n,m \geq 1} \mathbb{E} | \sum_{k=m}^{n+m-1} X_k | = O(1).$$

Proof. By (A1”),

$$\mathbb{E} \sum_{k=m}^{n+m-1} X_k \leq \sum_{k=m}^{n+m-1} \mathbb{E} | X_k | \preceq \sum_{k=m}^{n+m-1} (m + n - k)^{1-1/\alpha} = O(1).$$

All constants in $\preceq, O(\cdot)$ do not depend on all $m, n$. The last equality holds because of $1/\alpha - 1 > 1$. \hfill \Box
Lemma 4.4. If the conditions (A0’)-(A3”) are all satisfied, then there is a constant $C > 0$ such that for any $n, m \geq 1$

$$\mathbb{E}|\mathbb{E}_{n+m}[(\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k)^T] - \mathbb{E}[(\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k)^T]| \leq Cn^{\frac{\alpha}{2}}.$$ 

Proof.

\[
\begin{align*}
&\mathbb{E}|\mathbb{E}_{n+m}[(\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k)^T] - \mathbb{E}[(\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k)^T]| \\
&\leq \mathbb{E}|\mathbb{E}_{n+m}(\sum_{k=m}^{n+m-1} X_k \cdot X_k^T) - \mathbb{E}(\sum_{k=m}^{n+m-1} X_k \cdot X_k^T)| \\
&\quad + \mathbb{E}|\mathbb{E}_{n+m}(\sum_{m \leq i < j \leq n+m-1} X_i \cdot X_j^T) - \mathbb{E}(\sum_{m \leq i < j \leq n+m-1} X_i \cdot X_j^T)| \\
&\quad + \mathbb{E}|\mathbb{E}_{n+m}(\sum_{m \leq i < j \leq n+m-1} X_i \cdot X_j^T) - \mathbb{E}(\sum_{m \leq i < j \leq n+m-1} X_i \cdot X_j^T)^T)| \\
&\leq \mathbb{E}|\mathbb{E}_{n+m}(\sum_{k=m}^{n+m-1} X_k \cdot X_k^T - \mathbb{E}(X_k \cdot X_k^T))| \\
&\quad + 2\mathbb{E}|\mathbb{E}_{n+m}(\sum_{m \leq i < j \leq n+m-1} X_i \cdot X_j^T - \mathbb{E}(X_i \cdot X_j^T))|. 
\end{align*}
\]

Estimate (4.3): by (A2”) and $\alpha \in (0, 1/2)$,

\[
(4.3) \leq \sum_{k=m}^{n+m-1} \mathbb{E}|\mathbb{E}_{n+m}[X_k \cdot X_k^T - \mathbb{E}(X_k \cdot X_k^T)]| \lesssim \sum_{k=m}^{n+m-1} (m + n - k)^{-1/\alpha} = O(1).
\]

Estimate (4.4): for any fixed $j \leq n + m - 1$:

\[
\mathbb{E}|\mathbb{E}_{j}[\sum_{m \leq i < j} X_i \cdot X_j^T - \mathbb{E}(X_i \cdot X_j^T)]| \leq \sum_{m \leq i < j} \mathbb{E}|\mathbb{E}_{j}(X_i \cdot X_j^T) - \mathbb{E}(X_i \cdot X_j^T)| \leq 2 \sum_{m \leq i < j} \mathbb{E}|\mathbb{E}_{j}(X_i \cdot X_j^T)|.
\]

By (A1”) and $\alpha \in (0, 1/2)$, the inequality above becomes:

\[
\lesssim \sum_{m \leq i < j} \mathbb{E}|\mathbb{E}_{j}X_i| \lesssim \sum_{m \leq i < j} (j - i)^{1-\alpha} = O(1).
\]

That is, for any fixed $j \leq n + m - 1$,

\[
\sum_{m \leq i < j} \mathbb{E}|\mathbb{E}_{j}[X_i \cdot X_j^T - \mathbb{E}(X_i \cdot X_j^T)]| = O(1). \tag{4.5}
\]
Let $\delta := \frac{\alpha}{1-\alpha} < 1$, then

\begin{equation}
\sum_{j=m+1}^{n+m-1} \sum_{m \leq i < j} \mathbb{E}[\mathbb{E}_{n+m} [X_i \cdot X_j^T - \mathbb{E}(X_i \cdot X_j^T)]]
\end{equation}

By (4.5) and (A3’), the inequality above becomes

\begin{align*}
&\lesssim [n^\delta] + \sum_{j=m+1}^{n+m-[n^\delta]} \sum_{m \leq i < j} \mathbb{E}[\mathbb{E}_{n+m} [X_i \cdot X_j^T - \mathbb{E}(X_i \cdot X_j^T)]] \\
&\lesssim [n^\delta] + \sum_{j=m+1}^{n+m-[n^\delta]} \sum_{m \leq i < j} (n + m - j)^{-1/\alpha} \\
&\leq [n^\delta] + \sum_{j=m+1}^{n+m-[n^\delta]} \frac{j - m}{(n + m - j)^{1/\alpha - 1}} = [n^\delta] + \sum_{j=1}^{n-[n^\delta]} \frac{j}{(n-j)^{1/\alpha - 1}} \\
&= [n^\delta] + \sum_{j=1}^{n-[n^\delta]} \frac{j/n}{(1 - j/n)^{1/\alpha - 1}} \cdot n^{-1} \cdot n^{3-1/\alpha} \\
&\lesssim [n^\delta] + \int_{0}^{n-[n^\delta]} \frac{x}{(1 - x)^{1/\alpha - 1}} dx \cdot n^{3-1/\alpha} = [n^\delta] + \int_{[n^\delta]/n}^{1} \frac{1 - x}{x^{1/\alpha - 1}} dx \cdot n^{3-1/\alpha} \\
&\lesssim \begin{cases} 
[n^\delta] + n^{1-\delta}, & 1/\alpha - 1 = 2 \\
n^{1+\delta(2-1/\alpha)} + n^\delta, & 1/\alpha - 1 \neq 2 
\end{cases}
\lesssim n^{1-\alpha}.
\end{align*}

All constants in $\lesssim$, $O(\cdot)$ do not depend on $m, n$. \qed

**Lemma 4.5.** If the conditions (A0’’)-(A3’’) are all satisfied, then for any $\epsilon \in (0, 1 - \frac{\alpha}{1-\alpha})$ there is a constant $C_\epsilon > 0$ such that for any $n, m \geq 1$

\[
\mathbb{E}\{[\mathbb{E}_{n+m} [(\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k)^T] - \mathbb{E}((\sum_{k=m}^{n+m-1} X_k) \cdot (\sum_{k=m}^{n+m-1} X_k)^T)]^{1+\epsilon}\} \leq C_\epsilon n^{1+\epsilon}.
\]
Proof. Let $\beta > \epsilon, \delta > 0$ (will be given later), and

$$
\Delta := \mathbb{E}_{n+m}( \sum_{k=m}^{n+m-1} X_k ) \cdot ( \sum_{k=m}^{n+m-1} X_k )^T - \mathbb{E}[ ( \sum_{k=m}^{n+m-1} X_k ) \cdot ( \sum_{k=m}^{n+m-1} X_k )^T ].
$$

Then

$$
\mathbb{E}( |\Delta|^{1+\epsilon} ) = \int_{|\Delta| > \delta} |\Delta|^{1+\epsilon} d\mu + \int_{|\Delta| \leq \delta} |\Delta|^{1+\epsilon} d\mu \leq \int_{|\Delta| > \delta} |\Delta|^{1+\beta} |\Delta|^{\epsilon-\beta} d\mu + \delta^\epsilon \cdot \mathbb{E}|\Delta| \\
\leq \delta^{\epsilon-\beta} \int |\Delta|^{1+\beta} d\mu + \delta^\epsilon \cdot \mathbb{E}|\Delta|.
$$

By the convexity of function $| \cdot |^{1+\beta}$ and the Hölder inequality, the inequality above becomes

$$
\leq \delta^{\epsilon-\beta} \int 2^\beta \cdot \left( \sum_{k=m}^{n+m-1} X_k \right) \cdot \left( \sum_{k=m}^{n+m-1} X_k \right)^T |1+\beta| d\mu + \delta^\epsilon \cdot \mathbb{E}|\Delta| \\
\leq 2^\beta \cdot \delta^{\epsilon-\beta} \cdot \int \left| \sum_{k=m}^{n+m-1} X_k \right| \cdot \left( \sum_{k=m}^{n+m-1} X_k \right)^T |1+\beta| d\mu + \delta^\epsilon \cdot \mathbb{E}|\Delta|.
$$

By the Minkowski’s inequality, Lemma 4.4 and (A0'''), the inequality above becomes

$$
\leq 2^\beta \cdot \delta^{\epsilon-\beta} \cdot \left( \sum_{k=m}^{n+m-1} ||X_k||_{L^{2+2\beta}} \right)^{2+2\beta} + \delta^\epsilon \cdot \mathbb{E}|\Delta| \leq 2^\beta \cdot \delta^{\epsilon-\beta} n^{2+2\beta} + \delta^\epsilon n^{1+\alpha}.
$$

Let $\delta = n^{\frac{2+2\beta-\frac{\alpha}{\beta}}{\beta}}$, then $\mathbb{E}( |\Delta|^{1+\epsilon} ) \leq 2^{\beta+1} \cdot n^{\frac{2+2\beta-\frac{\alpha}{\beta}}{\beta} + \frac{\alpha}{1-\alpha}}$. If $\epsilon \in (0, 1 - \frac{\alpha}{1-\alpha})$, we can choose a large $\beta$ such that $n^{\frac{2+2\beta-\frac{\alpha}{\beta}}{\beta} + \frac{\alpha}{1-\alpha}} \leq n^{1+\epsilon}$. Then $\mathbb{E}( |\Delta|^{1+\epsilon} ) \lesssim_{\epsilon} n^{1+\epsilon}$, and the constants in $\lesssim$ do not depend on $n, m$.

Lemma 4.6. If the conditions (A0''') and (A1'''') are satisfied, then there is a constant $C > 0$ such that for any $n, p, m, q \in \mathbb{N}

$$
|\mathbb{E}[ ( \sum_{k=m}^{n+m-1} X_k ) \cdot ( \sum_{k=q+m+n}^{q+n+m+p-1} X_k )^T ] | \leq C \max(n, p)^{\max(3-1/\alpha, 0)}.
$$

Proof. Let $\delta < 1$ (will be given later), $\bar{n} := \max(n, p)$, then for any $q \geq 0$,

$$
\mathbb{E}[ ( \sum_{k=m}^{n+m-1} X_k ) \cdot ( \sum_{k=q+m+n}^{q+n+m+p-1} X_k )^T ] = \sum_{k=m}^{n+m-1} \sum_{j=q+m+n}^{q+n+m+p-1} \mathbb{E}( X_k \cdot X_j^T ).
$$
By the conditions \((A0^*)\) and \((A1^*)\), the equality above becomes

\[
\begin{align*}
&\sum_{k=m}^{n+m-1} \sum_{j=q+n+m+1}^{q+n+m+p-1} \mathbb{E}[|X_k|] \lesssim \sum_{k=m}^{n+m-1} \sum_{j=q+m+n}^{n+m-1} (j - k)^{1-1/\alpha} \\
&= \sum_{k=m}^{n+m-1} \sum_{j=q+m+n}^{n+m-1} (j - (m + n) + (m + n) - k)^{1-1/\alpha} \\
&= \sum_{1 \leq k \leq n} \sum_{0 \leq j \leq p-1} (j + k + q)^{1-1/\alpha} \lesssim \sum_{1 \leq k \leq n} \sum_{1 \leq j \leq p} (j + k)^{1-1/\alpha} \\
&\lesssim \left[ \tilde{n}^{\delta} \right] + \sum_{[\tilde{n}^\delta] \leq k, j \leq \tilde{n}} (j/\tilde{n} + k/\tilde{n})^{1-1/\alpha} \tilde{n}^{-1} \cdot \tilde{n}^{-3-1/\alpha} \\
&\lesssim \left[ \tilde{n}^{\delta} \right] + \int_{[\tilde{n}^\delta]/\tilde{n}}^{1} \int_{[\tilde{n}^\delta]/\tilde{n}}^{1} (x + y)^{1-1/\alpha} dx dy \cdot \tilde{n}^{-3-1/\alpha} \\
&\lesssim \left[ \tilde{n}^{\delta} \right] + \tilde{n}^{3-1/\alpha} \cdot \int_{[\tilde{n}^\delta]/\tilde{n}}^{1} (1 + y)^{2-1/\alpha} - (y + [\tilde{n}^\delta]/\tilde{n})^{2-1/\alpha} dy \\
&\lesssim \left[ \tilde{n}^{\delta} \right] + \tilde{n}^{3-1/\alpha} \cdot \left[ 2^{3-1/\alpha} - (1 + [\tilde{n}^\delta]/\tilde{n})^{3-1/\alpha} - (1 + [\tilde{n}^\delta]/\tilde{n})^{3-1/\alpha} + (2[\tilde{n}^\delta]/\tilde{n})^{3-1/\alpha} \right] \\
&\lesssim \begin{cases} 
\tilde{n}^{\delta}, & 3 - 1/\alpha > 0 \\
\tilde{n}^{3-1/\alpha}, & 3 - 1/\alpha \leq 0 
\end{cases} \\
&= \tilde{n}^{\max(3-1/\alpha, 0)}.
\end{align*}
\]

All constants in \(\lesssim\) do not depend on \(m, n, p, q\).

**Lemma 4.7** (A maximal inequality).

If the conditions \((A0^*)-(A3^*)\) are satisfied, then for any \(\epsilon \in (0, \min\{1, 2 - \frac{2\alpha}{1-\alpha}\})\) there is a constant \(C_\epsilon > 0\) such that

\[
\mathbb{E}(\max_{m \leq k \leq m+n-1} \sum_{i=m}^{k} X_i^{2+\epsilon}) \leq C_\epsilon \cdot n^{1+\epsilon/2} \text{ for all } m, n \in \mathbb{N}.
\]

**Proof.** Similar to the martingale maximal inequality, Serfling [Ser68, Ser70] proved a maximal inequality for a random process (non-martingale) adapted to an increasing filtration. Although in different settings, we can still follow the idea of Theorem 3.1 in [Ser68], then apply Theorem B in [Ser70] to obtain the desired bound in our Lemma 4.7.

19
Note that if all \((X_k)_{k \in \mathbb{N}}\) satisfy the conditions (A0”)-(A3”), all coordinates of\((X_k)_{k \in \mathbb{N}}\) satisfy them too. Without loss of generality, we assume that all \((X_k)_{k \in \mathbb{N}}\) are random variables satisfying the conditions (A0”)-(A3”).

First we claim:
\[
\sup_{n,m \geq 1} \frac{\mathbb{E}(|\sum_{i=m}^{m+n-1} X_i|^{2+\epsilon})}{n^{2/3}} < \infty.
\]

Let \(A := \sum_{i=m}^{\lfloor n/2 \rfloor - 1} X_i, \quad B := \sum_{i=m+\lfloor n/2 \rfloor}^{m+n-1} X_i, \quad \epsilon \in (0, 1) \) (will be determined later),
\[
\mathbb{E}(|\sum_{i=m}^{m+n-1} X_i|^{2+\epsilon}) = \mathbb{E}(|A + B|^{2+\epsilon}) \leq \mathbb{E}(|A|^{2+\epsilon} + |B|^2 + 2|A| \cdot |B|)} \cdot (|A|^{\epsilon} + |B|^{\epsilon})
\]
\[
+ \mathbb{E}(|A|^{2+\epsilon} + |B|^{2+\epsilon} + 2|A| \cdot |B|)^{2+\epsilon} \cdot \mathbb{E}(|A|^{\epsilon} + A^2 \cdot |B|^{\epsilon}). \tag{4.6}
\]

Let \(s + t = 2 + \epsilon, s, \epsilon \in (0, 2], \epsilon/2 < 1 - \frac{\alpha}{1-n}\), by the Hölder inequality,
\[
\mathbb{E}(|A|^s) \cdot |B|^t = \mathbb{E}\{[\mathbb{E}_{m+[n/2]}(|A|^s)] \cdot |B|^t} \leq \mathbb{E}\{[\mathbb{E}_{m+[n/2]}(|A|^2)]^{s/2} \cdot |B|^t}\}
\]
\[
= \mathbb{E}\{[\mathbb{E}_{m+[n/2]}(|A|^2) - \mathbb{E}(|A|^2)]^{s/2} \cdot |B|^t} \}\leq \mathbb{E}\{\mathbb{E}_{m+[n/2]}(|A|^2) - \mathbb{E}(|A|^2)]^{s/2} \cdot |B|^t} + \mathbb{E}(\mathbb{E}_{m+[n/2]}(|A|^2) - \mathbb{E}(|A|^2)]^{s/2} \cdot |B|^t}\}
\]
\[
\leq \mathbb{E}(|B|^{2+\epsilon})^{2+\epsilon} \cdot \mathbb{E}\{[\mathbb{E}_{m+[n/2]}(|A|^2) - \mathbb{E}(|A|^2)]^{2+\epsilon} \cdot |n/2|^{\epsilon/2}
\]
\[
+ \mathbb{E}(|B|^{2+\epsilon})^{2+\epsilon} \cdot \mathbb{E}(\mathbb{E}_{m+[n/2]}(|A|^2) - \mathbb{E}(|A|^2)]^{2+\epsilon} \cdot |n/2|^{\epsilon/2}.
\]

By Lemma 4.2 and Lemma 4.5, there is a constant \(\bar{C}\) (does not depend on all \(n, m)\) such that the inequality above becomes
\[
\leq 2\mathbb{E}(|B|^{2+\epsilon})^{2+\epsilon} \cdot |n/2|^{\epsilon/2} \cdot \bar{C}.
\]

Apply the inequality above to (4.6) for \(s = 1, 1 + \epsilon, 2, \epsilon, 2, \) respectively, then
\[
\mathbb{E}(|A + B|^{2+\epsilon}) \leq \mathbb{E}(|A|^{2+\epsilon}) + \mathbb{E}(|B|^{2+\epsilon}) + 4 \cdot \bar{C} \cdot [\mathbb{E}(|B|^{2+\epsilon})]^{1+\epsilon} \cdot |n/2|^{1/2}
\]
\[
+ 4 \cdot \bar{C} \cdot [\mathbb{E}(|B|^{2+\epsilon})]^{1+\epsilon} \cdot |n/2|^{1/2} + 2 \cdot \bar{C} \cdot [\mathbb{E}(|B|^{2+\epsilon})]^{2+\epsilon} \cdot |n/2|^{\epsilon/2}
\]
\[
+ 2 \cdot \bar{C} \cdot [\mathbb{E}(|B|^{2+\epsilon})]^{2+\epsilon} \cdot |n/2|^{\epsilon/2}.
\]

Hence we have
\[
\frac{\mathbb{E}(|A + B|^{2+\epsilon})}{n^{2/3}} \leq \frac{\mathbb{E}(|A|^{2+\epsilon})}{n^{2/3}} \cdot \frac{|n/2|^{2+\epsilon}}{n^{2+\epsilon}} + \frac{\mathbb{E}(|B|^{2+\epsilon})}{n^{2+\epsilon}} \cdot \frac{(n - |n/2|)^{2+\epsilon}}{n^{2+\epsilon}}
\]
\[
+ 4 \cdot \bar{C} \cdot \left[\frac{\mathbb{E}(|B|^{2+\epsilon})}{(n - |n/2|)^{2+\epsilon}} \right]^{1+\epsilon} \cdot \frac{|n/2|^{1/2}}{n^{2+\epsilon}} \cdot \frac{(n - |n/2|)^{1+\epsilon}}{n^{2+\epsilon}}.
\]
Using \( \frac{\frac{n}{2}}{n^{\frac{1}{2}+\epsilon}} = [1/2 + o(1)]^{\frac{1}{2}+\epsilon} \), the inequality above becomes

\[
\frac{\mathbb{E}(|A|^{2+\epsilon})}{[n/2]^{\frac{1}{2}+\epsilon}} \cdot [1/2 + o(1)]^{\frac{1}{2}+\epsilon} + \frac{\mathbb{E}(|B|^{2+\epsilon})}{(n - [n/2])^{\frac{1}{2}+\epsilon}} \cdot [1/2 + o(1)]^{\frac{1}{2}+\epsilon} + 4 \cdot \bar{C} \cdot \left[ \frac{\mathbb{E}(|B|^{2+\epsilon})}{(n - [n/2])^{\frac{1}{2}+\epsilon}} \right]^{\frac{1}{2}+\epsilon} \cdot [1/2 + o(1)]^{\frac{1}{2}+\epsilon} \cdot [1/2 + o(1)]^{\frac{1}{2}+\epsilon} \]

\[
+ 2 \cdot \bar{C} \cdot \left[ \frac{\mathbb{E}(|B|^{2+\epsilon})}{(n - [n/2])^{\frac{1}{2}+\epsilon}} \right]^{\frac{2}{2+\epsilon}} \cdot [1/2 + o(1)]^{\frac{2}{2+\epsilon}} + 2 \cdot \bar{C} \cdot \left[ \frac{\mathbb{E}(|B|^{2+\epsilon})}{(n - [n/2])^{\frac{1}{2}+\epsilon}} \right]^{\frac{2}{2+\epsilon}} \cdot [1/2 + o(1)]^{\frac{2}{2+\epsilon}}.
\]

Let \( a_n := \max(\sup_{m \geq 1} \frac{\mathbb{E}(\sum_{i=m+1}^{n} X_i^{2+\epsilon})}{n^{\frac{1}{2}+\epsilon}}, \sup_{m \geq 1} \frac{\mathbb{E}(\sum_{i=m+1}^{n} X_i^{2+\epsilon})}{n^{\frac{1}{2}+\epsilon}}) \), the estimates above show that:

\[
a_n \leq [1/2 + o(1)]^{\frac{1}{2}+\epsilon} \cdot (2a_{[n/2]} + 4 \cdot \bar{C} \cdot a_{[n/2]}^{\frac{1}{2}+\epsilon} + 2 \cdot \bar{C} \cdot a_{[n/2]}^{\frac{1}{2}+\epsilon} + 2 \cdot \bar{C} \cdot a_{[n/2]}^{\frac{1}{2}+\epsilon}). \tag{4.7}
\]

Let \( g(x) := 2 + 4 \cdot \bar{C} \cdot x^{\frac{1}{2}+\epsilon} + 4 \cdot \bar{C} \cdot x^{\frac{1}{2}+\epsilon} + 2 \cdot \bar{C} \cdot x^{\frac{1}{2}+\epsilon} + 2 \cdot \bar{C} \cdot x^{\frac{1}{2}+\epsilon} \), then

\[
a_n \leq a_{[n/2]} \cdot [1/2 + o(1)]^{\frac{1}{2}+\epsilon} \cdot g(a_{[n/2]}).
\]

There is \( x_0 \) such that for all \( x \geq x_0 \), \( g(x) \approx 2 \).

Since \( o(1) \to 0 \) as \( n \to \infty \), then there is \( N \) such that for all \( n \geq N \),

\[
[1/2 + o(1)]^{\frac{1}{2}+\epsilon} \approx (1/2)^{\frac{1}{2}+\epsilon} < 1/2.
\]

Then we can choose large \( x_0, N \) such that for any \( n \geq N, x \geq x_0 \),

\[
[1/2 + o(1)]^{\frac{1}{2}+\epsilon} \cdot g(x) < 1.
\]

Let \( b_n = \max(a_n, x_0) \), then for any \( n \geq N \), (4.7) becomes:

\[
a_n \leq [1/2 + o(1)]^{\frac{1}{2}+\epsilon} \cdot (2b_{[n/2]} + 4 \cdot \bar{C} \cdot b_{[n/2]}^{\frac{1}{2}+\epsilon} + 2 \cdot \bar{C} \cdot b_{[n/2]}^{\frac{1}{2}+\epsilon} + 2 \cdot \bar{C} \cdot b_{[n/2]}^{\frac{1}{2}+\epsilon}) < b_{[n/2]}.
\]

Therefore, \( b_n \leq b_{[n/2]} \) for any \( n \geq N \). Furthermore, for any \( n \geq 1 \),

\[
\sup_{m \geq 1} \frac{\mathbb{E}(\sum_{i=m+1}^{n} X_i^{2+\epsilon})}{n^{\frac{1}{2}+\epsilon}} \leq a_n \leq b_n \leq \max(b_1, b_2, \cdots, b_N) < \infty.
\]
Second, apply the inequality above to the Theorem B in [Ser70], that is, suppose \((X_i)_{i \geq 1}\) have finite variances, zero means, and

\[
\sup_m \mathbb{E}(\sum_{i=m+1}^{m+n} |X_i|^{2+\epsilon}) \lesssim n^{\frac{2+\epsilon}{2}},
\]

then

\[
\mathbb{E}(\max_{1 \leq k \leq n} \sum_{i=m+1}^{m+k} |X_i|^{2+\epsilon}) \lesssim n^{\frac{2+\epsilon}{2}} \text{ for all } m, n \in \mathbb{N},
\]

where the constant in \(\lesssim\) does not depend on \(m, n\). So we obtain our desired maximal inequality.

To find the desired Gaussian vectors in the definition of the VASIP, Berkes and Philipp [BP79] gave a criterion:

**Theorem 4.8** (See [BP79]).

Given a probability space \((X, \mathcal{B}, \mu)\), let \((Y_k)_{k \geq 1}\) be a sequence of random vectors in \(\mathbb{R}^d\), adapted to an increasing filtration \((\mathcal{G}_k)_{k \geq 1}\), that is, \(Y_k\) is \(\mathcal{G}_k\)-measurable. Let \((H_k)_{k \geq 1}\) be a family of positive semi-definite \(d \times d\) matrices. Assume that \(\mu_k\) is a Gaussian distribution with a characteristic function \(e^{-\frac{1}{2}(u, H_k u)}\). Suppose that there are some non-negative numbers \(T_k \geq 10^8d, \lambda_k, \delta_k\) such that for any \(u \in \mathbb{R}^d\) with \(|u| \leq T_k\):

\[
\mathbb{E}[\mathbb{E}[\exp(i \langle u, Y_k \rangle) | \mathcal{G}_{k-1}] - \exp(-\frac{u^T H_k u}{2})] \leq \lambda_k,
\]

\[
\mu_k\{u : |u| \geq T_k/4\} \leq \delta_k.
\]

Then without changing its distribution we can define \((Y_k)_{k \geq 1}\) on a richer probability space together with a family of independent Gaussian vectors \((G_k)_{k \geq 1}\) whose distributions are \((\mu_k)_{k \geq 1}\) and

\[
\tilde{P}(|Y_k - G_k| \geq \alpha_k) \leq \alpha_k,
\]

where \(\alpha_1 = 1\), \(\alpha_k := 16d \cdot \frac{\log T_k}{T_k} + 4 \lambda_k^{1/2} \cdot T_k + \delta_k, k \geq 2\), \(\tilde{P}\) is the probability w.r.t. the richer probability space.

In particularly, if \(\sum_{k \geq 1} \alpha_k < \infty\), then almost surely,

\[
\sum_{k \geq 1} |Y_k - G_k| < \infty.
\]

Berkes and Philipp constructed Gaussian vectors inductively, which relies heavily on the increasing filtration \((\mathcal{G}_k)_{k \geq 1}\). However, our filtration \((\mathcal{E}_k)_{k \geq 1}\) is decreasing. We will derive the following lemma for this case, which plays a crucial role in our proof of the Theorem 4.1.
**Lemma 4.9 (A VASIP criterion).**

Given a probability space \((X, \mathcal{B}, \mu)\), let \((Y_k)_{k \geq 1}\) be a sequence of random vectors in \(\mathbb{R}^d\), \((\mathcal{F}_k)_{k \geq 1}\) be a decreasing filtration, \(Y_k\) be \(\mathcal{F}_k\)-measurable, \((H_k)_{k \geq 1}\) be a family of positive semi-definite \(d \times d\) matrices. Assume that \(\mu_k\) is a Gaussian distribution with a characteristic function \(e^{-\frac{1}{2}u^T H_k u}\). Suppose that there are some non-negative numbers \(T_k \geq 10^8d, \lambda_k, \delta_k\), such that for any \(u \in \mathbb{R}^d\) with \(|u| \leq T_k\):

\[
\mathbb{E}[\mathbb{E}[\exp(iu^T \cdot Y_k)|\mathcal{F}_{k+1}] - \exp(-\frac{u^T H_k u}{2})] \leq \lambda_k, \\
\mu_k\{u : |u| \geq T_k/4\} \leq \delta_k.
\]

Then without changing its distribution we can define \((Y_k)_{k \geq 1}\) on a richer probability space together with a family of independent Gaussian vectors \((G_k)_{k \geq 1}\) whose distributions are \((\mu_k)_{k \geq 1}\) and

\[
\tilde{P}(|Y_k - G_k| \geq \alpha_k) \leq \alpha_k,
\]

where \(\alpha_k := 16d \cdot \frac{\log T_k}{T_k} + 4\lambda_k^{1/2} \cdot T_k^d + \delta_k\) for all \(k \geq 1\), \(\tilde{P}\) is the probability w.r.t. the richer probability space.

In particular, if \(\sum_{k \geq 1} \alpha_k < \infty\), then almost surely,

\[
\sum_{k \geq 1} |Y_k - G_k| < \infty.
\]

**Proof.** Before proving this lemma, let’s recall the procedure of how to construct Gaussian vectors in [BP79]: \(G_1\) is constructed with the distribution \(\mu_1\), extend the probability space \(\Omega\) to \(\Omega \times I\) by multiplying an unit interval \(I\) endowed with the Lebesgue measure if the original probability space has atoms. Inductively, assume that \(G_1, G_2, \ldots, G_{k-1}\) have been constructed, partition the extended probability space such that it is a union of countably many \(\sigma(G_1, \ldots, G_{k-1})\)-measurable sets. Locally, on each of these sets, construct \(G_k\), and extend the extended probability space by multiplying a new unit interval. Obtain global \(G_k\) by gluing all local \(G_k\). The final extended probability space is \(\Omega \times I^N\).

To prove our result, let \(I_n = [n, n + 1], n \in \mathbb{Z}\), we will construct a triangular array of Gaussian vectors \((G_{ik})_{1 \leq k \leq n, n \geq 1}\) together with extended probability spaces \((\Omega_n)_{n \geq 1}\):

For the 1-st row of the array, let \(G^1_1 := G_1, \mu_1 = L(G^1_1), \Omega_1 := \Omega \times I_1\) (we denote its probability by \(\tilde{P}\)) as in [BP79]. Assume that the previous \((n - 1)\) rows of the array are done: the extended probability space \(\Omega_{n-1}\) (still denote its probability by \(\tilde{P}\)) and \((G^m_{k-1})_{1 \leq k \leq n-1}\) are constructed.

For the \(n\)-th row of the array, consider the increasing filtration \((\mathcal{F}_{n+2-k})_{1 \leq k \leq n+1}\). By Theorem 4.8, we can construct \(G^n_{n+1}, G^n_1, \ldots, G^n_n\) and a probability space \(\Omega_{n+1} \times I^N_n\) (still denote its probability by \(\tilde{P}\)) such that

\[
\tilde{P}(|Y_k - G^n_k| \geq \alpha_k) \leq \alpha_k, \mu_k = L(G^n_k), 1 \leq k \leq n,
\]

\[
\mathbb{E}\mathbb{E}[\exp(iu^T \cdot Y_k)|\mathcal{F}_{k+1}] - \exp(-\frac{u^T H_k u}{2})] \leq \lambda_k, \\
\mu_k\{u : |u| \geq T_k/4\} \leq \delta_k.
\]
where $\alpha_{n+1} = 1$, $\alpha_k = 16d \cdot \log \frac{N_k}{k} + 4\lambda_k^{1/2} \cdot T_k \cdot \delta_k$, $1 \leq k \leq n$.

Since $G_{n+1}^m$ and $\alpha_{n+1}$ do not make contributions, drop them. Then we have $G_n^m$, $\cdots$, $G_1^m$ and a probability space $\Omega_{n-1} \times I^n_1$ such that

$$\tilde{P}(|Y_k - G_k^m| \geq \alpha_k) \leq \alpha_k,$$

where $\alpha_k = 16d \cdot \log \frac{N_k}{k} + 4\lambda_k^{1/2} \cdot T_k \cdot \delta_k$, $1 \leq k \leq n$.

This procedure ends up with a large extended probability space $\Omega \times \prod_{i \geq 1} I_i^n$ (still denote its probability by $\tilde{P}$) and a triangular array of Gaussian vectors $(G_k^m)_{1 \leq k \leq n, n \geq 1}$ such that

$$\tilde{P}(|Y_k - G_k^m| \geq \alpha_k) \leq \alpha_k,$$

$$\mu_k = L(G_k^m)$$

for all $k, n, 1 \leq k \leq n$.

Consider a new triangular array $(Y_k, G_k^m)_{1 \leq k \leq n, n \geq 1}$, we will construct the desired $(G_k)_{k \geq 1}$ inductively:

Start with the 1-st step, since $(G_1^m)_{n \geq 1}$ have the same distribution, hence $(Y_1, G_1^m)_{n \geq 1}$ is tight. Then along a subsequence, there is a weak limit $(Y_1', G_1')$ such that

$$(Y_1, G_1^m) \to_d (Y_1', G_1'),$$

$$\tilde{P}(|Y_1' - G_1'| \geq \alpha_1) \leq \alpha_1.$$

Assume that the $(m-1)$-th step is done, that is, $(Y_1, G_1^m, Y_2, G_2^m, \cdots, Y_{m-1}, G_{m-1}^m)_{n \geq m-1}$ has a subsequence with a weak limit $(Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}')$ and an extended probability space $\prod_{m-1 \leq i \leq 1} I_i \times \Omega \times \prod_{i \geq 1} I_i^n$ (still denote its probability by $\tilde{P}$) such that

$$(Y_1, G_1^m, Y_2, G_2^m, \cdots, Y_{m-1}, G_{m-1}^m) \to_d (Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}'),$$

$$\tilde{P}(|Y_k' - G_k'| \geq \alpha_k) \leq \alpha_k$$

for any $k \leq m - 1$.

For the $m$-th step, since $(Y_1, G_1^m, Y_2, G_2^m, \cdots, Y_m, G_m^m)_{n \geq m}$ is tight, then along a subsequence of the subsequence in the $(m-1)$-th step, there is a weak limit

$$(Y_1, G_1^m, Y_2, G_2^m, \cdots, Y_m, G_m^m) \to_d (Y_1', G_1', \cdots, \hat{Y}_k', \cdots, \hat{Y}_{m-1}', G_{m-1}', G_m').$$

Compare with the weak limit in the $(m-1)$-th step, we have

$$(Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}') = (Y_1', G_1', \cdots, \hat{Y}_k', \cdots, \hat{Y}_{m-1}', G_{m-1}').$$

By Lemma 8.3, there is $(Y_m', G_m')$ such that

$$(Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}', Y_m', G_m') = (Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}', \hat{Y}_m', \hat{G}_m').$$
Meanwhile, we have an extended probability space \( \prod_{-m \leq i \leq -1} I_i \times \Omega \times \prod_{i \geq 1} I_i^N \) (still denote its probability by \( \tilde{P} \)).

Therefore, in this \( m \)-th step, we have a weak convergence along a subsequence:

\[
(Y_1, G_1^m, Y_2, G_2^m, \ldots, Y_m, G_m^m) \rightarrow_d (Y_1', G_1', Y_2', G_2', \ldots, Y_m', G_m').
\]

Then by the diagonal argument, there is a subsequence (does not depend on \( m \)),
such that for any \( m \geq 1 \),

\[
(Y_1, G_1^m, Y_2, G_2^m, \ldots, Y_m, G_m^m) \rightarrow_d (Y_1', G_1', Y_2', G_2', \ldots, Y_m', G_m'),
\]

\[
\tilde{P}(|Y_k' - G_k'| \geq \alpha_k) \leq \alpha_k \text{ for any } k \leq m.
\]

Therefore, for any \( k \geq 1 \),
\[
L(Y_1', \ldots, Y_k') = L(Y_1, \ldots, Y_k),
\]
\[
L(G_1', \ldots, G_k') = \bigotimes_{1 \leq i \leq k} \mu_i.
\]

They imply that \((G_i')_{i \geq 1}\) are independent and \((Y_i')_{i \geq 1} \overset{d}{=} (Y_i)_{i \geq 1}\). The extended probability space becomes \( \prod_{i \leq -1} I_i \times \Omega \times \prod_{i \geq 1} I_i^N \) (still denote its probability by \( \tilde{P} \)).

Using Lemma 8.3 again, there are \((G_i)_{i \geq 1}\) and an extended probability space \( \prod_{i \leq -1} I_i \times I_0 \times \Omega \times \prod_{i \geq 1} I_i^N \) (still denote its probability by \( \tilde{P} \)) such that

\[
((Y_i')_{i \geq 1}, (G_i')_{i \geq 1}) \overset{d}{=} ((Y_i)_{i \geq 1}, (G_i)_{i \geq 1}).
\]

Therefore, for any \( k \geq 1 \),
\[
\tilde{P}(|Y_k' - G_k'| \geq \alpha_k) \leq \alpha_k,
\]
\[
L(G_1, \ldots, G_k) = \bigotimes_{1 \leq i \leq k} \mu_i,
\]

where \( \alpha_k := 16d \cdot \log T_k \cdot T_k^{1/2} + 4\lambda_k^{1/2} \cdot T_k^d + \delta_k, k \geq 1 \).

With all lemmas above, we are ready to prove Theorem 4.1.

### 4.2 Proof of Theorem 4.1

**Introduction of the blocks**

We will construct consecutive blocks \( \{I_n, n \geq 1\} \) in \( \mathbb{N} \) without gaps between them:

let \( I_n \) be the interval in \( \mathbb{N} \) such that \( |I_n| = [n^c], c > 0 \). So \( \bigcup_{i \geq 1} I_i = \mathbb{N} \). Let \( a \in (1/2, 1), c_n := [n^{c(1-a)}] \). Construct consecutive blocks \( \{I_{n,i}, 1 \leq i \leq c_n + 1\} \) in
We are going to apply Lemma 4.9 to $Y_n, \mathcal{F}_n, H_n, T_n, \mu_n, \delta_n$ and estimate

$$E|E[\exp(iu^T \cdot Y_n)|\mathcal{F}_{n+1}] - \exp\left(-\frac{u^T \cdot H_n \cdot u}{2}\right)|$$

We are going to apply Lemma 4.9 to $Y_n, \mathcal{F}_n, H_n, T_n, \mu_n, \delta_n$ and estimate

$$E|E[\exp(iu^T \cdot Y_n)|\mathcal{F}_{n+1}] - \exp\left(-\frac{u^T \cdot H_n \cdot u}{2}\right)|$$

We have

$$= E|E_{a_{n+1}} \exp(iu^T \cdot \frac{X_n}{\sqrt{b_n}}) - \exp\left(-\frac{u^T \cdot \left[\frac{E_n \cdot X_n^T}{b_n}\right]}{2b_n}\right)|.$$ (4.8)

First, we note that

$$E(X_n \cdot X_n^T) = \sum_{i=1}^{c_{n+1}} E(X_{n,i} \cdot X_{n,i}^T) + \sum_{1 \leq i < j \leq c_n+1} E(X_{n,i} \cdot X_{n,j}^T) + \sum_{1 \leq i < j \leq c_n+1} E(X_{n,i} \cdot X_{n,j})^T$$

$$= \sum_{i=1}^{c_{n+1}} E(X_{n,i} \cdot X_{n,i}^T) + \sum_{i=1}^{c_n} E\left[\sum_{i<j \leq c_n+1} X_{n,j}^T\right] + \sum_{i=1}^{c_n} E\left[\sum_{i<j \leq c_n+1} X_{n,i} \cdot X_{n,j}^T\right].$$

By Lemma 4.6, the equality above becomes

$$= \sum_{i=1}^{c_{n+1}} E(X_{n,i} \cdot X_{n,i}^T) + \sum_{i=1}^{c_n} O(|I_n|^\max(3-1/\alpha, 0)) = \sum_{i=1}^{c_{n+1}} E(X_{n,i} \cdot X_{n,i}^T) + c_n \cdot O(|I_n|^\max(3-1/\alpha, 0))$$

$$= \sum_{i=1}^{c_{n+1}} E(X_{n,i} \cdot X_{n,i}^T) + O(n^{c(1-\alpha)+c \max(3-1/\alpha, 0)}).$$

Then we have the following estimate:

$$E|E_{a_{n+1}} \exp(iu^T \cdot \frac{X_n}{\sqrt{b_n}}) - \exp\left(-u^T \cdot \frac{E_n \cdot X_n^T}{2b_n}\right) \cdot u)|$$
\[
\begin{align*}
&\leq \mathbb{E}[\mathcal{E}_{n+1} \exp(iu^T \cdot \frac{X_n}{\sqrt{b_n}}) - \exp(-u^T \cdot \frac{\sum_{i=1}^{c_n+1} \mathbb{E}(\bar{X}_{n,i} \cdot X_{n,i}^T)}{2b_n} u)] \\
&\quad + |\exp(-u^T \cdot \frac{\sum_{i=1}^{c_n+1} \mathbb{E}(\bar{X}_{n,i} \cdot X_{n,i}^T)}{2b_n} u) - \exp(-u^T \cdot \frac{\mathbb{E}(X \cdot \bar{X}^T)}{2b_n} u)| \\
&\leq \mathbb{E}[\mathcal{E}_{n+1} \exp(iu^T \cdot \frac{X_n}{\sqrt{b_n}}) - \exp(-u^T \cdot \frac{\sum_{i=1}^{c_n+1} \mathbb{E}(\bar{X}_{n,i} \cdot X_{n,i}^T)}{2b_n} u)] \\
&\quad + \frac{O(|u|^2 \cdot n^{c(1-a)+c \max(3-1/a,0)})}{b_n}. \tag{4.9}
\end{align*}
\]

Since
\[
\begin{align*}
&\exp(iu^T \cdot \frac{X_n}{\sqrt{b_n}}) - \exp(-u^T \cdot \frac{\sum_{i=1}^{c_n+1} \mathbb{E}(\bar{X}_{n,i} \cdot X_{n,i}^T)}{2b_n} u) \\
&= \sum_{k=0}^{c_n} \left[ \exp(-\frac{\sum_{i=1}^{k} \mathbb{E}(u^T \cdot \bar{X}_{n,i})^2}{2b_n}) \cdot \exp\left(i \frac{\sum_{i=k+1}^{c_n} u^T \cdot \bar{X}_{n,i}}{\sqrt{b_n}}\right) \\
&\quad - \exp(-\frac{\sum_{i=1}^{k} \mathbb{E}(u^T \cdot \bar{X}_{n,i})^2}{2b_n}) \cdot \exp\left(i \frac{\sum_{i=2+k}^{c_n} u^T \cdot \bar{X}_{n,i}}{\sqrt{b_n}}\right) \right] \\
&= \sum_{k=0}^{c_n} \left\{ \exp(-\frac{\mathbb{E}(u^T \cdot \bar{X}_{n,k+1})^2}{2b_n}) \cdot \left[ \exp\left(i \frac{\sum_{i=k+2}^{c_n} u^T \cdot \bar{X}_{n,i}}{\sqrt{b_n}}\right) \right] \\
&\quad - \exp\left(-\frac{\mathbb{E}(u^T \cdot \bar{X}_{n,k+1})^2}{2b_n}\right) \right\},
\end{align*}
\]

then (4.9) becomes
\[
\begin{align*}
&= \mathbb{E}[\mathcal{E}_{n+1} \left\{ \sum_{k=0}^{c_n} \exp(-\frac{\sum_{i=1}^{k} \mathbb{E}(u^T \cdot \bar{X}_{n,i})^2}{2b_n}) \cdot \left[ \exp\left(i \frac{\sum_{i=k+2}^{c_n} u^T \cdot \bar{X}_{n,i}}{\sqrt{b_n}}\right) \right] \\
&\quad - \exp\left(-\frac{\mathbb{E}(u^T \cdot \bar{X}_{n,k+1})^2}{2b_n}\right) \right\}] + \frac{O(|u|^2 \cdot n^{c(1-a)+c \max(3-1/a,0)})}{b_n}.
\end{align*}
\]

Since \(\exp(i \frac{\sum_{i=k+2}^{c_n} u^T \cdot \bar{X}_{n,i}}{\sqrt{b_n}})\) is \(\mathcal{F}_{n,k+2}\)-measurable, then the equality above becomes
\[
\begin{align*}
&= \mathbb{E}[\mathcal{E}_{n+1} \left\{ \sum_{k=0}^{c_n} \exp(-\frac{\sum_{i=1}^{k} \mathbb{E}(u^T \cdot \bar{X}_{n,i})^2}{2b_n}) \cdot \exp\left(i \frac{\sum_{i=k+2}^{c_n} u^T \cdot \bar{X}_{n,i}}{\sqrt{b_n}}\right) \right\} \\
&\times \mathbb{E}[\exp\left(i \frac{u^T \cdot \bar{X}_{n,k+1}}{\sqrt{b_n}}\right) - \exp\left(-\frac{\mathbb{E}(u^T \cdot \bar{X}_{n,k+1})^2}{2b_n}\right) | \mathcal{F}_{n,k+2}] \right\}] + \frac{O(|u|^2 \cdot n^{c(1-a)+c \max(3-1/a,0)})}{b_n} \\
&\leq \sum_{k=0}^{c_n} \mathbb{E}[\mathbb{E}[\exp\left(i \frac{u^T \cdot \bar{X}_{n,k+1}}{\sqrt{b_n}}\right) - \exp\left(-\frac{\mathbb{E}(u^T \cdot \bar{X}_{n,k+1})^2}{2b_n}\right) | \mathcal{F}_{n,k+2}] \right\}] + \frac{O(|u|^2 \cdot n^{c(1-a)+c \max(3-1/a,0)})}{b_n}.
\end{align*}
\]
Using the Taylor expansion: for any $\epsilon_0 \in (0, \min(1, 2 - \frac{2\alpha}{1-\alpha}))$,

$$e^{-x} = 1 - x + O(x^2),$$

$$e^{ix} = 1 + ix - x^2/2 + x^2 \cdot O(\min(|x|, 1)) = 1 + ix - x^2/2 + O(|x|^{2+\epsilon_0}),$$

the inequality above becomes:

$$= \sum_{k=0}^{\epsilon_0} E^2 \{ [1 + i \frac{u^T \cdot X_{n,k+1}}{2b_n} - (\frac{u^T \cdot X_{n,k+1}}{2b_n})^2] + O((\frac{u^T \cdot X_{n,k+1}}{2b_n})^{2+\epsilon_0}) | F_{n,k+2} \}$$

$$- \sum_{0 \leq k \leq c_n} \{ |u| \sqrt{b_n} + \frac{|u|^2}{2b_n} \cdot E \{ |X_{n,k+1} \cdot X_{n,k+1}^T - E(X_{n,k+1} \cdot X_{n,k+1}^T) | F_{n,k+2} \}$$

$$+ \frac{|u|^2 \cdot E(\frac{|X_{n,k+1} \cdot X_{n,k+1}^T|^2}{b_n})^2}{b_n^{(2+\epsilon_0)/2}} + \frac{|u|^{2+\epsilon_0} \cdot E(\frac{|X_{n,k+1} |^{2+\epsilon_0}}{b_n^{(2+\epsilon_0)/2}})}{b_n} + O(\frac{|u|^2 \cdot n^c(1-a) + c \max(3-1/a, 0)}{b_n})$$

Let $|u| \leq T_n = n^\kappa$ and apply Lemmas 4.2, 4.3, 4.7 to $X_{n,k+1}$, Lemma 4.4 to $X_{n,k+1}$ and $F_{n,k+2}$; the inequality above becomes

$$\lesssim \frac{n^{\kappa+c(1-a)}}{n^{(1+\epsilon_0)(1+c)}} + \frac{n^{2\kappa+c(1-a)+c\alpha_n/1-\alpha}}{n^{(1+c)/2}} + \frac{n^{c(1-a)+4\kappa+2c\alpha}}{n^{(1+c)}} + \frac{n^{\kappa(2+\epsilon_0)+c(1-a)+c\alpha_2}}{n^{(1+c)(2+\epsilon_0)}}$$

$$+ \frac{n^{\kappa+e(1-a)+c \max(3-1/a, 0)}}{n^{(1+c)}} =: n^{-\kappa_1} + n^{-\kappa_2} + n^{-\kappa_3} + n^{-\kappa_4} + n^{-\kappa_5}.$$

Here $\epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha})$. Let $v := \min\{\kappa_1, \kappa_2, \kappa_3, \kappa_4, \kappa_5\}$. Then when $|u| \leq n^\kappa$,

$$E \{ E[\exp(iu^T \cdot Y_n)] | F_{n+1} \} - \exp(-\frac{u^T \cdot H_n \cdot u}{2}) \lesssim n^{-v}.$$  

**Gaussian approximations for $(X_n)_{n \geq 1}$**

From now on, we choose appropriate blocks $\{I_n : n \geq 1\}$ and $\{I_{n,i} : 1 \leq i \leq c_n + 1\}$ (that is, appropriate constants $c, \gamma, \kappa, v, a$ and $\epsilon_0$) such that they satisfy

$$\min(\kappa, v/2 - d\kappa) > 1, \quad (4.10)$$

$$\gamma(c + 1)/2 > 1 + (c + 1)/2 - \min(\kappa, v/2 - d\kappa), \quad (4.11)$$

$$c - \gamma(c + 1) < 0, \quad (4.12)$$

$$1 + (c + 1)(\max\{3 - 1/\alpha, 0\} - \gamma) < 0 \quad (4.13)$$

28
\[
\gamma(c + 1)(2 + \epsilon_0)/2 - c(1 + \epsilon_0/2) > 1. \tag{4.14}
\]

There are many choices for these constants, for example,

\[
\kappa = 2, c = \max \left\{ \frac{\epsilon_0 a + (2 + \epsilon_0)(8d + 12)}{\epsilon_0(1 - a)}, \frac{\max(3 - 1/\alpha, 0) + \epsilon_0/(2 + \epsilon_0)}{1 - \max(3 - 1/\alpha, 0)} \right\},
\]

\[
a = \max \left\{ \frac{\epsilon_0 + 2\alpha}{(1 - \alpha)(2\epsilon_0 + 2)}, \frac{2 + 2\epsilon_0}{3\epsilon_0 + 4}, \frac{(2 + \epsilon_0)(1 + \max(3 - 1/\alpha, 0)) - 2}{2 + 2\epsilon_0} \right\},
\]

\[
\epsilon_0 \text{ is any number in } (0, \min\{1, 2 - \frac{2\alpha}{1 - \alpha}\}),
\]

\[
\gamma \text{ is any number in } \left( \frac{c}{c + 1} + \frac{2}{(c + 1)(2 + \epsilon_0)}, 1 \right).
\]

With these blocks, we apply Lemma 4.9 and have the following approximation.

**Lemma 4.10.** There are independent Gaussian vectors \(G''_n\) with variance matrices \(E(X_n \cdot X_n^\top)\) such that

\[
\sum_{i=1}^n X_i - \sum_{i=1}^n G''_i \preceq \lambda(\sigma^2_{aa})(1 - \epsilon')/2 \text{ a.s.}
\]

for a sufficiently small \(\epsilon' > 0\) (depends on \(\gamma, c, \kappa, v\)).

**Proof.** First, by Lemma 4.2, \(|H_n| = |E(X_n \cdot X_n^\top)| \preceq n^{c - \gamma(c+1)}\). Together with (4.12), we have

\[
\delta_n = \mu_n\{u : |u| \geq T_n/4\} \preceq \int_{|t| > T_n/4} \det(H_n)^{-1/2} \exp(- \langle t, H_n^{-1}t \rangle/2) dt \preceq \int_{|t| > (n^{\gamma(c+1)-c}T_n)/4} e^{-|t|^2/2} dt
\]

decays exponentially. So \(\delta_n = \mu_n\{u : |u| \geq T_n/4\} \preceq n^{-v}\).

Then

\[
\alpha_n := 16d \cdot \frac{\log T_n}{T_n} + 4\lambda_n^{1/2} \cdot T_n^d + \delta_n \preceq n^{-\kappa} + n^{d\kappa - v/2} \preceq n^{-\min(\kappa, v/2 - d\kappa)},
\]

which is summable (i.e. \(\sum_{n \geq 1} \alpha_n < \infty\)) due to (4.10). By Lemma 4.9, there are Gaussian vectors \(G''_n\) with variance matrices \(E(X_n \cdot X_n^\top)\) such that

\[
|Y_n - \frac{G''_n}{\sqrt{b_n}}| = \frac{|X_n}{\sqrt{b_n}} - \frac{G''_n}{\sqrt{b_n}} < \alpha_n \text{ i.o.}
\]
Then almost surely,
\[
\sum_{i=1}^{n} X_i - G_i'' \lesssim \sum_{i=1}^{n} \alpha_i \sqrt{b_i} \lesssim \sum_{i=1}^{n} i^{-\min(\kappa,v/2-\delta \kappa)} \cdot i^{s+1} \lesssim n^{1+\frac{s+1}{2}-\min(\kappa,v/2-\delta \kappa)}.
\]

By (4.11), we can choose a small \( \epsilon' \) such that
\[
\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} G_i'' \lesssim n^{1+\frac{s+1}{2}-\min(\kappa,v/2-\delta \kappa)} \lesssim \lambda\left(\sigma_{a_n}^2\right)^{\frac{1}{\epsilon'+2}} \text{ a.s.} \]

**Comparisons between \((\overline{X}_n)_{n \geq 1}\) and \((X_m)_{m \geq 1}\)**

For any \( m \), there is \( n \in \mathbb{N} \) such that \( a_n \leq m < a_{n+1} \) and we have

**Lemma 4.11.** \( \lambda(\sigma_m^2) \approx \lambda(\sigma_{a_n}^2) \).

**Proof.** By Lemma 4.2 and Lemma 4.6,
\[
\mathbb{E}\left[\left(\sum_{i=1}^{m} X_i\right) \cdot \left(\sum_{i=1}^{m} X_i\right)^T\right] = \mathbb{E}\left[\left(\sum_{i=1}^{n} \overline{X}_i\right) \cdot \left(\sum_{i=1}^{n} \overline{X}_i\right)^T\right] + \mathbb{E}\left[\left(\sum_{i=1}^{n} \overline{X}_i\right) \cdot \left(\sum_{i=a_n+1}^{m} X_i\right)^T\right]
\]
\[
+ \mathbb{E}\left[\left(\sum_{i=a_n+1}^{m} X_i\right) \cdot \left(\sum_{i=1}^{n} \overline{X}_i\right)^T\right] + \mathbb{E}\left[\left(\sum_{i=a_n+1}^{m} X_i\right) \cdot \left(\sum_{i=a_n+1}^{m} X_i\right)^T\right]
\]
\[
\lesssim \mathbb{E}\left[\left(\sum_{i=1}^{n} \overline{X}_i\right) \cdot \left(\sum_{i=1}^{n} \overline{X}_i\right)^T\right] + n\alpha_{a_n+1}^{\max(3-1/\alpha,0)} + n\alpha_{a_n+1}^{\max(3-1/\alpha,0)} + a_{n+1} - a_n.
\]

Using \( a_n \approx a_{n+1} \) and \( \mathbb{E}\left[\left(\sum_{i=1}^{n} \overline{X}_i\right) \cdot \left(\sum_{i=1}^{n} \overline{X}_i\right)^T\right] \lesssim a_n^{\gamma} \), the inequality above becomes
\[
\lesssim \left\{ \mathbb{E}\left[\left(\sum_{i=1}^{n} \overline{X}_i\right) \cdot \left(\sum_{i=1}^{n} \overline{X}_i\right)^T\right] \right\} \cdot \left[ 1 + n\alpha_{a_n}^{\max(3-1/\alpha,0)-\gamma} + n\alpha_{a_n}^{\max(3-1/\alpha,0)-\gamma} + a_{a_n}^{-\gamma}(a_{n+1} - a_n) \right]
\]
\[
\lesssim \left\{ \mathbb{E}\left[\left(\sum_{i=1}^{n} \overline{X}_i\right) \cdot \left(\sum_{i=1}^{n} \overline{X}_i\right)^T\right] \right\} \cdot \left[ 1 + n^{1+(c+1)(\max(3-1/\alpha,0)-\gamma)} + n^{c-\gamma(c+1)} \right].
\]

By (4.12) and (4.13), we have
\[
\lambda(\sigma_m^2) \lesssim \inf_{|u|=1} u^T \cdot \mathbb{E}\left[\left(\sum_{i=1}^{n} \overline{X}_i\right) \cdot \left(\sum_{i=1}^{n} \overline{X}_i\right)^T\right] \cdot u = \lambda(\sigma_{a_n}^2).
\]

Similarly, by Lemmas 4.2 and 4.6,
\[
\sigma_{a_n}^2 \lesssim \mathbb{E}\left[\left(\sum_{i=1}^{m} X_i\right) \cdot \left(\sum_{i=1}^{m} X_i\right)^T\right] + n\alpha_{a_n}^{\max(3-1/\alpha,0)} + n\alpha_{a_n}^{\max(3-1/\alpha,0)} + a_{n+1} - a_n.
\]
$$\sum_{i=1}^{m} X_i \cdot \left( \sum_{i=1}^{m} X_i \right)^T + m^{1/\alpha + \max(3-1/\alpha, 0)} + m^{c-\gamma}.$$ 

Since \( \mathbb{E}\left[ \left( \sum_{i=1}^{m} X_i \right) \cdot \left( \sum_{i=1}^{m} X_i \right)^T \right] \gtrsim m^\gamma \), then the inequality above becomes

$$\lesssim \left\{ \mathbb{E}\left[ \left( \sum_{i=1}^{m} X_i \right) \cdot \left( \sum_{i=1}^{m} X_i \right)^T \right] \right\} \cdot (1 + m^{1/\alpha + \max(3-1/\alpha, 0)} - \gamma + m^{c-\gamma}).$$

By (4.12) and (4.13) again, we have \( \lambda(\sigma_{a_n}^2) \lesssim \lambda(\sigma_m^2) \).

Lemma 4.12.

$$\sup_{a_n \leq m \leq a_{n+1}} \left| \sum_{i=a_n+1}^{m} X_i \right| \lesssim \lambda(\sigma_{a_n}^2)^{1/2-\epsilon'} \text{ a.s.}$$

for a sufficiently small \( \epsilon' > 0 \) (depends on \( \gamma, c, \epsilon_0 \)).

Proof. By Lemma 4.7,

$$\mathbb{E}\left( \frac{\sup_{a_n \leq m \leq a_{n+1}} \left| \sum_{i=a_n+1}^{m} X_i \right|}{\lambda(\sigma_{a_n}^2)^{1/2-\epsilon'}} \right)^{2+\epsilon_0} \lesssim \frac{n^c(1+\epsilon_0/2)}{n^{\gamma(c+1)(2+\epsilon_0)(1/2-\epsilon')}}.$$ 

From (4.14), there is a small \( \epsilon' > 0 \) such that \( \gamma(c+1)(2+\epsilon_0)(1/2-\epsilon') - c(1+\epsilon_0/2) > 1 \). By the Borel-Cantelli Lemma, we have

$$\sup_{a_n \leq m \leq a_{n+1}} \left| \sum_{i=a_n+1}^{m} X_i \right| \lesssim \lambda(\sigma_{a_n}^2)^{1/2-\epsilon'} \text{ a.s.}$$

Gaussian approximations for \( (X_m)_{m \geq 1} \)

It is always possible to find nonzero independent Gaussian vectors \( \{G_k, k \geq 1\} \) such that for each \( n \in \mathbb{N} \), \( \sum_{k \in I_n} G_k = G_n'' \) where \( G_n'' \) are in the Lemma 4.10. We claim that \( \sum_{i \leq m} G_i \) matches \( \sum_{i \leq m} X_i \) for any \( m \in \mathbb{N} \) in the sense of (4.1) and (4.2):

Verify (4.2): for any \( m \), there is \( n \) such that \( a_n \leq m < a_{n+1} \). Recall

$$\mathbb{E}(G_i'' \cdot G_i''^T) = \sum_{j \in I_i} \mathbb{E}[(G_j) \cdot (G_j)^T] = \mathbb{E}(\overline{X}_i \cdot \overline{X}_i^T),$$

where \( \mathbb{E} \) is the expectation of the probability of the extended probability space. Then, by (4.15), we have

$$\mathbb{E}\left[ \left( \sum_{i=1}^{m} X_i \right) \cdot \left( \sum_{i=1}^{m} X_i \right)^T \right] - \sum_{i \leq m} \mathbb{E}(G_i \cdot G_i^T)$$
By (4.15), Lemma 4.2 and Lemma 4.6, the equality above becomes

\[ \lesssim na_n(3-1/\alpha,0) + na_n(3-1/\alpha,0) + 2(a_{n+1} - a_n) \lesssim n^{1+(c+1)\max(3-1/\alpha,0)} + n^c \]

\[ \lesssim \lambda(\sigma_{an}^2)\frac{1+(c+1)\max(3-1/\alpha,0)}{\gamma(c+1)} + \lambda(\sigma_{an}^2)\gamma(c+1). \]

By (4.12), (4.13), there is a small \( \epsilon' > 0 \) such that

\[ \frac{1 + (c + 1) \max(3 - 1/\alpha, 0)}{\gamma(c + 1)} < 1 - \epsilon', \quad \frac{c}{\gamma(c + 1)} < 1 - \epsilon'. \]

Therefore, by Lemma 4.11, we have

\[ |\mathbb{E}[(\sum_{i\leq m} X_i) \cdot (\sum_{i \leq m} X_i)^T] - \sum_{i \leq m} \mathbb{E}(G_i \cdot G_i^T)| \lesssim \lambda(\sigma_{an}^2) \approx \lambda(\sigma_m^2) \]

Verify (4.1): by Lemma 4.11 and Lemma 4.12, we have

\[ \sum_{i=1}^{m} X_i - \sum_{i=1}^{m} G_i = \sum_{i=1}^{n} \tilde{X}_i - n \sum_{i=a_n+1}^{m} G''_i + \sum_{i=a_n+1}^{m} X_i - \sum_{i=a_n+1}^{m} G_i. \quad (4.16) \]

Note that

\[ \mathbb{E}(\sup_{a_n \leq m \leq a_{n+1}} |\sum_{i=a_n+1}^{m} G_i|^{1/2-\epsilon'}) \lesssim (a_{n+1} - a_n)^{1+\epsilon_0/2} \frac{n^{c(1+\epsilon_0/2)}}{n^{\gamma(c+1)(2+\epsilon_0)(1/2-\epsilon')}} \lesssim \frac{n^{c(1+\epsilon_0/2)}}{n^{\gamma(c+1)(2+\epsilon_0)(1/2-\epsilon')}} ; \]

so by (4.14), there is a small \( \epsilon' > 0 \) such that

\[ \sup_{a_n \leq m \leq a_{n+1}} |\sum_{i=a_n+1}^{m} G_i| \lesssim \lambda(\sigma_{an}^2) \approx \lambda(\sigma_m^2) \text{ a.s.} \]

Hence, by Lemma 4.10, (4.16) can be estimated as follows

\[ \lesssim \lambda(\sigma_{an}^2) \frac{1}{a_n} + \sup_{a_n \leq m \leq a_{n+1}} |\sum_{i=a_n+1}^{m} X_i| \lesssim \lambda(\sigma_{an}^2) \frac{1}{a_n} \approx \lambda(\sigma_m^2) \frac{1}{a_n} \text{ a.s.} \]

Therefore (4.1) and (4.2) hold. So we finish the proof of the Theorem 4.1. \( \square \)
5 Proof of Theorem 2.6

Proof. We are going to apply our Theorem 4.1 to prove the Theorem 2.6. Let $X_k := \phi_k \circ T^k$, $E_k := T^{-k}B$. Clearly, (A0”) holds.

Verify (A1”): by (A1),

$$\begin{align*}
\sup_{i \geq 1} \mathbb{E}[E_{n+i}X_i]d\mu &= \sup_{i \geq 1, ||\psi|| \leq 1} \int \psi \circ T^{n+i} \cdot \mathbb{E}_{n+i} \phi_i \circ T^i d\mu \\
&= \sup_{i \geq 1, ||\psi|| \leq 1} \int \psi \circ T^{n+i} \cdot \phi_i \circ T^i d\mu = \sup_{i \geq 1, ||\psi|| \leq 1} \int \psi \circ T^{n+i} \cdot \phi_i \cdot P^i 1 d\mu \\
&= \sup_{i \geq 1} \int |P^{n+i}(\phi_i \cdot P^i 1)| d\mu \lesssim n^{1-\alpha}.
\end{align*}$$

Verify (A2”): by (A2),

$$\begin{align*}
\sup_{i \geq 1} \mathbb{E}[\mathbb{E}_{n+i}[X_i \cdot X_i^T - \mathbb{E}(X_i \cdot X_i^T)]] &= \sup_{i \geq 1, ||\psi|| \leq 1} \int \psi \circ T^{n+i} \cdot [\phi_i \circ T^i \cdot \phi_i^T \circ T^i - \int \phi_i \circ T^i \cdot \phi_i^T \circ T^i d\mu] d\mu \\
&= \sup_{i \geq 1, ||\psi|| \leq 1} \int \psi \circ T^{n+i} \cdot [(\phi_i \cdot \phi_i^T - \int \phi_i \circ T^i \cdot \phi_i^T \circ T^i d\mu) \cdot P^i 1] d\mu \\
&= \sup_{i \geq 1} \int |P^{n+i}[(\phi_i \cdot \phi_i^T - \int \phi_i \circ T^i \cdot \phi_i^T \circ T^i d\mu) \cdot P^i 1]| d\mu \lesssim n^{1-\alpha}.
\end{align*}$$

Verify (A3”): by (A3),

$$\begin{align*}
\mathbb{E}[\mathbb{E}_{n+i+j}[X_i \cdot X_i^{T+j} - \mathbb{E}(X_i \cdot X_i^{T+j})]] &= \sup_{||\psi|| \leq 1} \mathbb{E}[\psi \circ T^{n+i+j} \cdot [X_i \cdot X_i^{T+j} - \mathbb{E}(X_i \cdot X_i^{T+j})]] \\
&= \sup_{||\psi|| \leq 1} \int \psi \circ T^{n+i+j} \cdot [\phi_i \cdot \phi_i^T \circ T_i^{T+j} - \mathbb{E}(\phi_i \cdot \phi_i^T \circ T_i^{T+j} \cdot P^i 1)] \cdot P^i 1 d\mu \\
&= \sup_{||\psi|| \leq 1} \int \psi \circ T^{i+1+j+n} \cdot \{P^{i+1}(\phi_i \cdot P^i 1) \cdot \phi_i^{T+j} - P^{i+j} 1 \cdot \int P^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^{T+j} d\mu\} d\mu \\
&= \int |P^{i+j+n}(\phi_i \cdot P^i 1) \cdot \phi_i^{T+j} - P^{i+j} 1 \cdot \int P^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^{T+j} d\mu)| d\mu \lesssim n^{1-\alpha},
\end{align*}$$

where the constant in $\lesssim$ does not depend on $i, j, n$.

By Theorem 4.1, there is $\gamma \in (0, 1)$ which depends on $d, \alpha$ only ($\gamma$ will be given in Appendix, Lemma 8.1), such that if $\lambda(\sigma_n^2) \gtrsim n^\gamma$, then $(\phi_k \circ T^k)_{k \geq 1}$ satisfies the VASIP. Therefore we finish the proof of Theorem 2.6. \(\square\)
6 Proof of Theorem 2.9

Proof. For almost every \( \omega \in \Omega \), we will apply our Theorem 4.1 to the probability space \((X, \mathcal{B}, \mu_\omega)\) and the maps \((T_{\nu} \circ \omega)_{k \geq 0}\). Let \( X_k := \phi_{\sigma^k} \circ T_k \) and \( \mathcal{E}_k := (T_k)^{-1} \mathcal{B} \), define:

\[
E^\omega(\cdot) = \int (\cdot) d\mu_\omega, \quad E_n^\omega(\cdot) := E^\omega(\cdot)(T_n^m)^{-1} \mathcal{B}.
\]

Clearly, (A0*) holds.

Verify (A1*): by (A1'),

\[
sup_{i \geq 1} E_n^\omega(\sup_{1 \leq \|\psi\| \leq 1} \int \psi \circ T_{n+i} d\mu_\omega) = \sup_{i \geq 1} E_n^\omega(\sup_{1 \leq \|\psi\| \leq 1} \int \psi \circ T_{\sigma_i} d\mu_\sigma_\omega) \leq \sup_{i \geq 0} \int P_n^\omega(\phi_{\sigma_i} \circ h_{\sigma_i}) d\mu \lesssim n^{1-1/\alpha}.
\]

Verify (A2*): by (A2')

\[
sup_{k \geq 1} E_n^\omega(\sup_{k \geq 1} E_{n+k}^\omega \|X_k \cdot X_k^T - \mathbb{E}(X_k \cdot X_k^T)\|)
\]

\[
= \sup_{k \geq 1} E_n^\omega(\sup_{k \geq 1} \|\psi \circ T_k d\mu_\omega\|)
\]

\[
\lesssim \sup_{k \geq 0} \int P_n^\omega(\{\phi_{\alpha_k} \circ \phi_{\alpha_k}^T - \mathbb{E}\phi_{\alpha_k} \circ \phi_{\alpha_k}^T\} \cdot h_{\alpha_k}) d\mu \lesssim n^{1-1/\alpha}.
\]

Verify (A3*): by (A3')

\[
E_n^\omega(\sup_{k \geq 1} E_{n+i+j}^\omega \|X_i \cdot X_i^T \cdot X_i \cdot X_i^T\|)
\]

\[
= \sup_{\|\psi\| \leq 1} \int \psi \circ T_{\sigma_i}^{n+j} d\mu_\sigma_\omega \left(\phi_{\sigma_i} \circ \phi_{\sigma_i+j} \circ T_{\sigma_i}^j - \mathbb{E}\phi_{\sigma_i} \circ \phi_{\sigma_i+j} \circ T_{\sigma_i}^j\right) d\mu_\sigma_\omega
\]

\[
= \sup_{\|\psi\| \leq 1} \int \psi \circ T_{\sigma_i}^{n+j} d\mu_\sigma_\omega \left(\phi_{\sigma_i} \circ \phi_{\sigma_i+j} \circ h_{\sigma_i} \circ \phi_{\sigma_i+j}^T - h_{\sigma_i} \circ \phi_{\sigma_i+j}^T\right) d\mu_\sigma_\omega
\]

\[
= \int P_{\sigma_i}^{n+j} d\mu_\sigma_\omega \left(\phi_{\sigma_i} \circ \phi_{\sigma_i+j} \circ h_{\sigma_i} \circ \phi_{\sigma_i+j}^T - h_{\sigma_i} \circ \phi_{\sigma_i+j}^T\right) d\mu_\sigma_\omega \lesssim n^{1-1/\alpha},
\]

where the constant in \( \lesssim \) does not depend on \( i, j, n, \omega \).

Therefore for this \( \omega \in \Omega \), by our Theorem 4.1, (4.1) and (4.2) hold under the condition \( \sigma_n^2(\omega) \gtrsim n^7 \). In other words, if \( \sigma_n^2(\omega) \gtrsim n^7, (2.5) \) and (2.6) hold.
Next we claim the quenched VASIP or the coboundary in Theorem 2.9 by verifying the variance growth (2.7).

The verification of whether (2.7) holds or not follows the argument of Lemma 12 in [DFGTV18a] except the following:

1. in our case, the last inequality of page 2270 in [DFGTV18a] becomes
   \[ \leq \bar{K} \cdot \sum_{i \geq 1} i^{-1/\alpha} < \infty, \]

2. in our case, the inequality in the middle of page 2271 becomes
   \[ \leq \bar{K} \sum_{i=0}^{n-1} \sum_{k \geq n-i} k^{-1/\alpha} = \sum_{k=1}^{n-1} k \cdot k^{-1/\alpha} + n \sum_{k \geq n} k^{-1/\alpha} \lesssim n^{3-1/\alpha} \int_{1/n}^1 x^{2-1/\alpha} dx + n \sum_{k \geq n} k^{-1/\alpha}. \]

Then (35) in [DFGTV18a] becomes \( \leq n^{-1} \cdot (n^{3-1/\alpha} + n \sum_{k \geq n} k^{-1/\alpha}) \to 0. \)

Therefore there is a \( d \times d \) positive semi-definite matrix \( \Sigma^2 \) such that almost every \( \omega \in \Omega, \)
   \[ \lim_{n \to \infty} n^{-1} \int (\sum_{k=0}^{n-1} \phi_{\sigma^k(\omega)} \circ T^k_\omega) \cdot (\sum_{k=0}^{n-1} \phi_{\sigma^k(\omega)} \circ T^k_\omega)^T d\mu_\omega = \Sigma^2. \]

If \( \Sigma^2 > 0 \), then \( \sigma^2_\omega(\omega) \) grows linearly for a.e. \( \omega \in \Omega \), that is, (2.7) holds. By Theorem 4.1, the quenched VASIP for \( (\phi_{\sigma^k} \circ T^k_\omega)_{k \geq 1, \omega \in \Omega} \) holds.

If \( \det(\Sigma^2) = 0 \), without loss of generality, we assume that \( \Sigma^2 = \begin{bmatrix} I_{d_1 \times d_1} & 0 \\ 0 & 0_{d_2 \times d_2} \end{bmatrix} \).

If \( d_1 = 0 \), we claim that the coboundary holds:

Without loss of generality, we assume that all \( (\phi_{\omega})_{\omega \in \Omega} \) are scalar functions, and denote \( \bar{\phi}(\omega, x) := \phi_{\omega}(x) \). Similar to the computations of Lemma 12 (36) in [DFGTV18a], we have:

\[ 0 = \Sigma^2 = \int \bar{\phi}^2(\omega, x) d\mu_\omega dP + 2 \sum_{i \geq 1} \int \bar{\phi}(\omega, x) \cdot \bar{\phi} \circ \tau^i(\omega, x) d\mu_\omega dP, \]

where \( \tau(\omega, x) := (\sigma^\omega, f^\omega(x)) \) (here \( (\Omega \times X, \tau, d\mu_\omega dP) \) is a stationary dynamical system).

For the stationary dynamical system \( (\Omega \times X, \tau, d\mu_\omega dP) \) with the observable \( \tilde{\phi} \in L^\infty(\Omega \times X) \), we denote the transfer operator of \( \tau \) by \( \tau^* \). We will verify conditions (1) and (2) of Theorem 1.1 in [Liv96]: by (A1'),

\[ \sum_{n \geq 0} \left| \int \tilde{\phi} \cdot \bar{\phi} \circ \tau^n d\mu_\omega dP \right| \lesssim \sum_{n \geq 0} \int |P^n(\phi_{\omega} \cdot h_\omega)| d\mu dP \lesssim \sum_{n \geq 1} n^{1-1/\alpha} < \infty, \]
\[
\sum_{n \geq 0} \int |\tau^n \phi| d\mu_\omega d\mathbb{P} = \sum_{n \geq 0} \sup_{|\xi|, |\eta| \leq 1} \int \xi \circ \tau^n \cdot \phi d\mu d\mathbb{P} \lesssim \sum_{n \geq 0} \int |P^n_\omega (\phi \cdot h_\omega)| d\mu d\mathbb{P} \lesssim \sum_{n \geq 0} n^{1-1/\alpha} < \infty.
\]

Therefore, by Theorem 1.1 of [Liv96], there is \( \psi \in L^1(\Omega \times X) \) such that:

\[
\phi_{\sigma \omega}(T_\omega x) = \psi(\sigma(\omega), T_\omega(x)) - \psi(\omega, x) \text{ a.e. } (\omega, x).
\]

If \( d_1 > 0, d_2 > 0 \), then \( \mathbb{R}^d = \mathbb{R}^{d_1} \oplus \mathbb{R}^{d_2} \) with projections \( \pi_1 : \mathbb{R}^{d_1} \oplus \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}, \pi_2 : \mathbb{R}^{d_1} \oplus \mathbb{R}^{d_2} \to \mathbb{R}^{d_2} \). Then \( (\pi_1 \circ \phi_{\sigma^k(\omega)} \circ T^k_\omega)_{k \geq 1, \omega \in \Omega} \) has the quenched VASIP by the argument of \(" \Sigma^2 > 0 \)" above. For \( (\pi_2 \circ \phi_{\sigma^k(\omega)} \circ T^k_\omega)_{k \geq 1, \omega \in \Omega} \), we follow the argument of \(" d_1 = 0 \)" above, so there is \( \psi \in L^1(\Omega \times X; \mathbb{R}^{d_2}) \) such that:

\[
\pi_2 \circ \phi_{\sigma \omega}(T_\omega x) = \psi(\sigma(\omega), T_\omega(x)) - \psi(\omega, x) \text{ a.e. } (\omega, x).
\]

So we finish the proof of Theorem 2.9. \(\square\)

### 7 Proof of Corollaries

**Proof of Corollary 3.1.** Define \( X(x) := x, x \in [0, 1] \), for a sufficiently large \( a_0 > 1 \), consider a cone \( C_{a_0} \subset L^1[0, 1] \):

\[
C_{a_0} := \{ f \in \text{Lip}_{\text{loc}}(0, 1) : f \geq 0, f \text{ decreasing, } X^{\alpha+1} \cdot f \text{ increasing}, f(x) \leq a_0 \cdot x^{-\alpha} \cdot \int f dm \}.
\]

To prove the Corollary 3.1, we need the following lemma:

**Lemma 7.1** (See also [AHN+15, LSV99, NTV18]).

Assume that \( K, M > 0, \phi_i \in \text{Lip}[0, 1] \) and \( h_k \in C_{a_0} \) with \( \|\phi_i\|_{\text{Lip}} \leq K, \|h_k\|_{L^1} \leq M \) for all \( i, k \geq 1 \). Then for a sufficiently large \( a_0 > 1 \) (does not depend on \( K, M \)), there are constants \( \lambda, v, \delta \) (only depends on \( K, M, \alpha, a_0 \)) such that the following holds:

\[
h_{i,k} := (\phi_i + \lambda \cdot X + v)h_k + \delta, h_{i,k}^2 := (\lambda \cdot X + v)h_k + \delta + \int \phi_i \cdot h_k dm \in C_{a_0}, \ (7.1)
\]

\[
\phi_i \cdot h_k - \int \phi_i \cdot h_k dm = h_{i,k}^1 - h_{i,k}^2 \in C_{a_0} - C_{a_0},
\]

\[
\int h_{i,k}^1 dm = \int h_{i,k}^2 dm,
\]

\( \mathbf{1} \in C_{a_0}, C_{a_0} \) is preserved by all \( T_k \)'s transfer operators \( P_k \).
Furthermore, there are constants $C_{K,M,a,0}, C_{a,0}$ such that for all $m, n \in \mathbb{N}, h \in C_{a,0}$:

$$||P_{m+1}^{n+m}(\phi_k \cdot h_k - \int \phi_k \cdot h_k dm)||_{L^1} \leq C_{K,M,a,0} \cdot n^{1-1/\alpha},$$  (7.2)

$$||P_{m+1}^{n+m}(h - \int h dm)||_{L^1} \leq C_{a,0} \cdot ||h||_{L^1} \cdot n^{1-1/\alpha}. \quad (7.3)$$

**Proof.** [AHN+15, NTV18] proved these properties for the cone $C_{a,0} \cap C^1(0,1)$. However, the $C^1$ conditions are not used in their proofs, so the decay of correlation (7.3) still holds for our $C_{a,0}$, and $C_{a,0}$ is still a $P_k$-invariant cone. To prove (7.1), the argument is replacing $|\phi_k'|_{\infty}$ with its Lipschitz constant $\text{Lip}(\phi_k)$ in Lemma 2.4 of [NTV18]. Then (7.2) holds by applying (7.3) and (7.1).

With this lemma, we can prove the VASIP for $(\phi_k)_{k \in \mathbb{N}} \subset \text{Lip}[0,1]$ in Corollary 3.1 now:

Since $\sup_i ||\phi_i||_{\text{Lip}} < \infty, \sup_i ||\phi_i \cdot \phi_i^T||_{\text{Lip}} \leq 2 \sup_i ||\phi_i||^2_{\text{Lip}} < \infty, P^k 1 \in C_{a,0}, ||P^k 1||_{L^1} = 1$, so the conditions (A1) and (A2) are easily verified by (7.2). Now we verify (A3):

$$\int |P_{i+1}^{i+j+n}(P_{i+1}^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^T) - \int P_{i+1}^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^T dm| dm \lesssim n^{1-1/\alpha}.$$  

For the fixed $i, j$ above, by (7.1), there are $h_1, h'_1, h''_1, h'''_1, h_2, h''_2, h'''_2 \in C_{a,0}$ and the following decompositions:

$h_1 - h_2 = \phi_i \cdot P^i 1 \in C_{a,0} - C_{a,0}, h'_1 := P_{i+1}^{i+j} h_1 \in C_{a,0}, h'_2 := P_{i+1}^{i+j} h_2 \in C_{a,0}, h''_1 - h''_2 = h''_1 \cdot \phi_i^T dm \in C_{a,0} - C_{a,0}, h''_1 - h''_2 = h'_2 \cdot \phi_i^T dm - h''_2 \cdot \phi_i^T dm \in C_{a,0} - C_{a,0}.$

So $P_{i+1}^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^T dm = h'_1 \cdot \phi_i^T dm - h''_2 \cdot \phi_i^T dm$.

By (7.1), (7.3),

$$\int |P_{i+1}^{i+j+n}(P_{i+1}^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^T) - \int P_{i+1}^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^T dm| dm = \int |P_{i+1}^{i+j+n}(h''_1 - h''_2 - \int h''_1 - h''_2 dm) \lesssim C_{a,0} \cdot (||h''_1||_{L^1} + ||h''_1||_{L^1}) \cdot n^{1-1/\alpha} \lesssim C_{a,0} \cdot C_{\text{Lip}} ||\phi_k||_{L^1} \sup_{i} ||h_i||_{L_1}||h_i||_{L_1} \cdot n^{1-1/\alpha} = C_{a,0} \cdot C_{\text{Lip}} ||\phi_k||_{L^1} \sup_{i} ||h_i||_{L_1}||h_i||_{L_1} \cdot n^{1-1/\alpha}.$$  

By (7.1), $||h_1||_{L^1}, ||h_2||_{L^1}$ are bounded by a constant $C_{\text{Lip}} ||\phi_k||_{L^1}$. Therefore,

$$\int |P_{i+1}^{i+j+n}(P_{i+1}^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^T) - \int P_{i+1}^{i+j}(\phi_i \cdot P^i 1) \cdot \phi_i^T dm| dm \leq C_{\text{Lip}} ||\phi_k||_{L^1} \cdot n^{1-1/\alpha},$$

for some constant $C_{\text{Lip}} ||\phi_k||_{L^1} > 0$.

Therefore the VASIP holds for $(\phi_i \circ T^i)_{i \geq 1}$ provided $\lambda(\sigma^2) \gtrsim n^\gamma$.  

37
For the self-norming CLT (2.8), we will give a similar but simpler proof: let $(\phi_k)_{k \in \mathbb{N}} \subset \text{Lip}([0, 1]; \mathbb{R})$, $I_n = [1, n]$. Let $a \in (1/2, 1)$, $c_n := \lfloor n^{1-a} \rfloor$. Construct consecutive blocks $I_{n,i}$ in $I_n$ such that: $|I_{n,i}| = \lfloor n^a \rfloor$, $1 \leq i \leq c_n$, the first block $I_{n,1}$ contains the least number of $I_n$, the last block $I_{n,c_n+1} := I_n \setminus \bigcup_{1 \leq i \leq c_n} I_{n,i}$ contains the largest number of $I_n$. So $|I_{n,c_n+1}| \leq 2\lfloor n^a \rfloor$ and $\bigcup_{1 \leq i \leq c_n} I_{n,i} = I_n$. Similar to the proof of Theorem 4.1, let $X_n := \sum_{i \leq n} \phi_i \circ T^i$, $b_n := \lambda(\sigma_n^2) = \sigma_n^2 \asymp n^{\gamma_1}$ ($\gamma_1$ will be given in Appendix, Lemma 8.2), fix any $u \in \mathbb{R}$:

$$|\mathbb{E}[\exp(iu \cdot \frac{X_n}{\sqrt{b_n}})] - \exp(-u^2/2)| \leq \mathbb{E}[:\mathbb{E}_{n+1}[\exp(iu \cdot \frac{X_n}{\sqrt{b_n}})] - \exp(-u^2/2)].$$

Using the same method which estimates (4.8), the inequality above becomes

$$\sum_{k=0}^{c_n} \mathbb{E}[\mathbb{E}[:1 + iu X_{n,k+1} - (u \cdot X_{n,k+1})^2 \mathbb{E}[: \mathbb{E}[:1 + iu X_{n,k+1} - (u \cdot X_{n,k+1})^2 + O(\frac{u \cdot X_{n,k+1}}{\sqrt{b_n}}) | F_{n,k+2}]|^2]] + O(n^{c(1-a)+c\max(3-1/a,0)} \lambda^2 n^{1-a} + c\max(3-1/a,0)))$$

By Lemmas 4.2, 4.3, 4.4, 4.7, the inequality above becomes

$$\sum_{k=0}^{c_n} \mathbb{E}[\mathbb{E}[:|u|^2 | X_{n,k+1}^2 + O(\frac{|u|^2 \cdot n^{(1-a)+c\max(3-1/a,0)}}{\sqrt{b_n}})]]$$

In order to have the self-norming CLT (2.8), we will choose appropriate constants $a, \gamma_1, \epsilon_0$ such that $\kappa_1, \kappa_2, \kappa_3, \kappa_4, \kappa_5$ are negative. One possible choice is,

$$a = \max\left(\frac{\epsilon_0 + (2 + \epsilon_0) \max(0, 3 - 1/a)}{2 + 2\epsilon_0}, \frac{\epsilon_0}{2 + \epsilon_0}, \frac{2 + \epsilon_0}{4 + 3\epsilon_0}\right),$$

where $\epsilon_0$ is any number in $(0, \min\{1, 2 - \frac{2a}{1-a}\})$, $\gamma_1$ is any number in $(\frac{2a + \epsilon_0}{2 + \epsilon_0}, 1)$. □

**Proof of Corollary 3.3.** It is not hard to show that under the assumptions of Corollary 3.3, there is $\lambda \in (0, 1)$ such that for any $f \in \mathcal{V}$,

$$\sup_m ||P_m^m(f - \int f d\mu)||_{\mathcal{V}} \lesssim \lambda^n ||f - \int f d\mu||_{\mathcal{V}}.$$
Note that $\mathcal{V}$ is a Banach algebra, so the conditions (A1)-(A3) are all satisfied. By Theorem 2.6, Corollary 3.3 holds.

**Proof of Corollary 3.6.** For the existence and uniqueness of the quasi-invariant probabilities, the proofs are the same as in Proposition 1 of [DFGT18a]. The conditions (A1’)-(A3’) can be verified by the methods used in the proof of Corollary 3.3. So by Theorem 2.9, we have the desired result but $\psi \in L^1(\Omega \times X, d\mu_\omega d\mathbb{P})$ in the case of the coboundary. To prove $\psi \in L^2(\Omega \times X, d\mu_\omega d\mathbb{P})$, it is exactly the same as in the Proposition 3 of [DFGT18a].

**Proof of Corollary 3.8.** The existence of the quasi-invariant probabilities is constructed similar to [DFGT18a]: consider the Banach space

$$Y = \{ v : \Omega \times [0, 1] \to \mathbb{R} : v_\omega := v(\omega, \cdot) \in L^1([0, 1], m), \sup_\omega \| v_\omega \|_{L^1} < \infty \}$$

with norm $\| v \| := \sup_\omega \| v_\omega \|_{L^1}$.

Define an operator $\mathcal{L} : Y \to Y$: $\mathcal{L}(v)_\omega := P_{\sigma^{-1}} \omega v_{\sigma^{-1}} \omega$. So $\| \mathcal{L} v \| \leq \| v \|$. Consider $(\mathcal{L}^n 1)_{n \geq 1}$. We claim this is a Cauchy sequence:

By Lemma 7.1, since $P_\omega 1 \in C_{a_0}$ for any $\omega \in \Omega$, then for any $n < m$,

$$\| \mathcal{L}^n 1 - \mathcal{L}^m 1 \| = \sup_\omega \| P_{\sigma^{-n}} \omega (1 - P_{\sigma^{-m}} \omega 1) \|_{L^1} \lesssim n^{1/\alpha}.$$ 

Then there is $h \in Y$ such that $\mathcal{L} h = h$, that is, $P_{\sigma^{-1}} h_{\sigma^{-1}} \omega = h_\omega$ for a.e. $\omega \in \Omega$. So $h_\omega$, as the limit of $P_{\sigma^{-n}} 1 \in C_{a_0}$, satisfies all conditions of $C_{a_0}$ except $h_\omega \in \text{Lip}_{\text{loc}}(0, 1)$. To prove it, the method is the same as in the Lemma 2.3 of [LSV99] (the last two lines of page 674 in [LSV99]). Therefore $h_\omega \in C_{a_0}$ for a.e. $\omega \in \Omega$. Define the quasi-invariant probability $d\mu_\omega := h_\omega dm$, so $(T_\omega)_* \mu_\omega = \mu_{\sigma \omega}$ for a.e. $\omega \in \Omega$. The verification of the conditions (A1’)-(A3’) is the same as in the Corollary 3.1. By Theorem 2.9, this corollary holds.

**Proof of Corollary 3.10.** First, we will show: if (A4) holds, then there is a $d \times d$ positive semi-definite matrix $\Sigma^2$ and $\epsilon \in (0, 1)$ such that

$$\mathbb{E}[\left(\sum_{i=1}^n \phi \circ T^i\right) \cdot \left(\sum_{i=1}^n \phi \circ T^i\right)^T] = n \cdot \Sigma^2 + o(n^{1-\epsilon}). \quad (7.4)$$

Note that, by (A4),

$$\sum_{i \geq 1} \mathbb{E}(\phi \cdot \phi^T \circ T^i) \lesssim \sum_{i \geq 1} i^{1-1/\alpha} < \infty$$

absolutely converges.
Let $\Sigma^2 := \mathbb{E}(\phi \cdot \phi^T) + \sum_{i \geq 1} \mathbb{E}(\phi \cdot \phi^T \circ T^i) + \sum_{i \geq 1} \mathbb{E}(\phi \cdot \phi^T \circ T^i)^T$, then

$$\mathbb{E}[\left(\sum_{i=1}^{n} \phi \circ T^i\right) \cdot \left(\sum_{i=1}^{n} \phi \circ T^i\right)^T] - n \cdot \Sigma^2$$

$$= \sum_{i=1}^{n} \mathbb{E}(\phi \circ T^i \cdot \phi^T \circ T^i) + \sum_{1 \leq i < j \leq n} \mathbb{E}[(\phi \circ T^i \cdot \phi^T \circ T^j) + (\phi \circ T^i \cdot \phi^T \circ T^j)^T] - n \cdot \Sigma^2$$

$$= n \cdot \mathbb{E}(\phi \cdot \phi^T) + \sum_{1 \leq i < j \leq n} \mathbb{E}(\phi \cdot \phi^T \circ T^j) + \sum_{1 \leq i < j \leq n} \mathbb{E}(\phi \cdot \phi^T \circ T^j) - n \cdot \Sigma^2$$

$$= n \cdot \mathbb{E}(\phi \cdot \phi^T) + \sum_{i=1}^{n} \sum_{j=1}^{n-i} \mathbb{E}(\phi \cdot \phi^T \circ T^j) + \sum_{i=1}^{n} \sum_{j=1}^{n-i} \mathbb{E}(\phi \cdot \phi^T \circ T^j) - n \cdot \sum_{i \geq 1} \mathbb{E}[(\phi \cdot \phi^T \circ T^i) + (\phi \cdot \phi^T \circ T^i)^T].$$

Then we just need to estimate:

$$\sum_{i=1}^{n} \sum_{j=1}^{n-i} \mathbb{E}(\phi \cdot \phi^T \circ T^j) - n \cdot \sum_{i \geq 1} \mathbb{E}(\phi \cdot \phi^T \circ T^i) = \sum_{i=1}^{n} \sum_{j=1}^{n-i} \mathbb{E}(\phi \cdot \phi^T \circ T^j) - \sum_{i=1}^{n} \sum_{j \geq 1} \mathbb{E}(\phi \cdot \phi^T \circ T^j)$$

$$= \sum_{i=1}^{n} \sum_{j=n-i+1}^{n} \mathbb{E}(\phi \cdot \phi^T \circ T^j) = \sum_{i=1}^{n} \sum_{j=n-i+1}^{n} \mathbb{E}(\phi \cdot \phi^T \circ T^j) + n \cdot \sum_{j>n} \mathbb{E}(\phi \cdot \phi^T \circ T^j).$$

By (A4) again, the equality above becomes

$$\leq \sum_{i=1}^{n} \sum_{j=n-i+1}^{n} j^{-1/\alpha} + n \cdot \sum_{j>n} j^{-1/\alpha} \leq \sum_{i=1}^{n} i^{-1/\alpha} + n \cdot \int_{n}^{\infty} x^{-1/\alpha} \, dx$$

$$= n^{3-1/\alpha} + n \cdot n^{2-1/\alpha} \leq n^{3-1/\alpha}.$$

Since $3 - 1/\alpha < 1$, then there is $\epsilon \in (0, 1)$ such that

$$\sigma^2_n - n \cdot \Sigma^2 = \mathbb{E}[(\sum_{i=1}^{n} \phi \circ T^i) \cdot (\sum_{i=1}^{n} \phi \circ T^i)^T] - n \cdot \Sigma^2 \lessapprox n^{3-1/\alpha} = o(n^{1-\epsilon}).$$

If $\det(\Sigma^2) > 0$, then $\sigma^2_n \gtrapprox n$. So, by Theorem 2.6, the VASIP holds if conditions (A5),(A6) are satisfied as well. Moreover, by Lemma 8.4, the Gaussian vectors are i.i.d. with variances $\Sigma^2$.

If $\det(\Sigma^2) = 0$, without loss of generality, assume that $\Sigma^2 = \begin{bmatrix} I_{d_1 \times d_1} & 0 \\ 0 & 0_{d_2 \times d_2} \end{bmatrix}_{d \times d}$. The argument in this case is exactly the same as in Theorem 2.9, we will not repeat it here.
To prove the VASIP for the Young tower $\Delta$, Young [You99] proved the first order decay of correlations (A4) and (A5) already, so we just need to verify the second order decay of correlation (A6):

$$\sup_j \int |P^n[P^j(\phi) \cdot \phi^T - P^j(\phi) \cdot \phi^T \, dv]| \, dv \lesssim n^{1-1/\alpha},$$

where $dv = \frac{dm}{dm}$, both $\frac{dm}{dm}$ and $\phi$ are in $L^\infty(\Delta) \cap C^\beta(\Delta)$, $\inf \frac{dm}{dm} > 0$ where $C^\beta(\Delta)$ is the same as in [You99]. Indeed, we just need to show $P^j\phi$ is also a Lipschitz function with a Lipschitz exponent which does not depend on $j$, then (A6) holds by using (A4):

Without loss of generality, we assume that $\phi$ is a scalar function with a Lipschitz exponent $C\phi$; for any $(a, m_a) \in F^{-j}\Delta_{m,i}$, the orbit $\{F^0(a, m_a), \ldots, F^j(a, m_a)\}$ touches $\Delta_{a,m}$ for $q_a$ times $(0 \leq q_a \leq j)$, $a \in \Delta_{0,1,a} \cap (F^R)^{-q_a-1} \Delta_{0,1,a} \cap \cdots \cap (F^R)^{-q_a-1} \Delta_{0,1,a}$. We denote $P_a := ((F^R)^{-q_a} \Delta_{0,1,a} \cap (F^R)^{-q_a-1} \Delta_{0,1,a} \cap \cdots \cap (F^R)^{-q_a-1} \Delta_{0,1,a}) \times m_a$. Therefore $F^j(P_a) = \Delta_{m,i}$. For different $P_a$, they are either exactly the same or disjoint.

Now we show that $P^j\phi$ are Lipschitz and Lipschitz exponents are uniformly bounded.

1. $P^j\phi$ is locally Lipschitz:

For any $z_1 := (x_1, m), z_2 := (x_2, m) \in \Delta_{m,i}$, for any $a$ discussed above, there are $y_a^1 \in P_a, y_a^2 \in P_a$ such that $F^jy_a^1 = (x_1, m), F^jy_a^2 = (x_2, m)$.

$$P^j(\phi)(x_1, m) = \sum_{F^j(y) = (x_1, m)} \frac{\phi(y) \frac{dm}{dm}(y)}{J^F(y) \frac{dm}{dm}(x_1, m)} = \sum_a \frac{\phi(y_a^1) \frac{dm}{dm}(y_a^1)}{J^F(y_a^1) \frac{dm}{dm}(x_1, m)}.$$

$$P^j(\phi)(x_2, m) = \sum_{F^j(y) = (x_2, m)} \frac{\phi(y) \frac{dm}{dm}(y)}{J^F(y) \frac{dm}{dm}(x_2, m)} = \sum_a \frac{\phi(y_a^2) \frac{dm}{dm}(y_a^2)}{J^F(y_a^2) \frac{dm}{dm}(x_2, m)}.$$

Then

$$|P^j(\phi)(x_1, m) - P^j(\phi)(x_2, m)|$$

$$\leq \frac{1}{\frac{dm}{dm}(x_1, m)} \left| \sum_a \frac{\phi(y_a^1) \frac{dm}{dm}(y_a^1)}{J^F(y_a^1)} - \sum_a \frac{\phi(y_a^2) \frac{dm}{dm}(y_a^2)}{J^F(y_a^2)} \right|$$

$$+ \frac{1}{\frac{dm}{dm}(x_1, m)} - \frac{1}{\frac{dm}{dm}(x_2, m)} \left| \sum_a \frac{\phi(y_a^2) \frac{dm}{dm}(y_a^2)}{J^F(y_a^2)} \right|$$

$$\lesssim \left| \sum_a \frac{\phi(y_a^1) \frac{dm}{dm}(y_a^1) - \phi(y_a^2) \frac{dm}{dm}(y_a^2)}{J^F(y_a^2)} \right|$$

$$+ \left| \sum_a \frac{\phi(y_a^2) \frac{dm}{dm}(y_a^2)}{J^F(y_a^2)} (1 - \frac{J^F(y_a^2)}{J^F(y_a^2)}) \right|.$$
+ \sum_a \phi(y_a^2) \frac{dv}{m(y_a^2)} \cdot \beta_s(x_1, x_2),

where \( y_a^1, y_a^2 \in P_a \). Using the distortion (3.2), \( F^{j+R_i-m}P_a = F^{R_i-m}\Delta_{m,i} = \Delta_0 \) and \( JF^{R_i-m}|_{\Delta_{m,i}} = 1 \), the inequality above becomes

\[ \lesssim C_\phi \cdot C_v \cdot \sum_a \frac{d(y_a^1, y_a^2)}{JF^j(y_a)} + C_\phi \cdot C_v \cdot \sum_a \frac{\beta_s(x_1, x_2)}{JF^j(y_a^2)}. \]

\[ \lesssim \sum_a C_v \cdot C_\phi \cdot \frac{m(P_a)}{m(\Delta_0)} \cdot d(z_1, z_2) \leq C_v \cdot C_\phi \cdot \frac{m(\Delta)}{m(\Delta_0)} \cdot d(z_1, z_2). \]

2. \( P^j\phi \) is bounded:

\[ |P^j(\phi)(x_1, m)| \leq P^j(1)(x_1, m) \cdot ||\phi||_\infty = ||\phi||_\infty. \]

Therefore, \( P^j(\phi) \) is globally Lipschitz, that is, \( P^j(\phi) \in C_\beta(\Delta) \): for any \( z_1, z_2 \in \Delta \),

\[ |P^j(\phi)(z_1) - P^j(\phi)(z_2)| \lesssim ||\phi||_\infty \cdot C_v \cdot C_\phi \cdot d(z_1, z_2). \]

The Lipschitz exponents of \( P^j(\phi) \), as shown above, are uniformly bounded. \( \square \)

**Acknowledgments**

The author thanks his Ph.D. advisor Prof. Andrew Török for giving him one of the problems in this paper and University of Houston for a good place to study dynamical systems. The author thanks Prof. Leonid Bunimovich and Prof. Ian Melbourne for helpful comments and the anonymous reviewer for valuable remarks and comments.

8 **Appendix**

**Lemma 8.1 (Computations of the range of \( \gamma \)).**

The parameter \( \gamma \) in Theorems 4.1 and 2.6 can be any number in \( \left( \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon_0)}, 1 \right) \), where

\[ c = \max(\frac{\epsilon_0 a + (2 + \epsilon_0)(8d + 12)}{\epsilon_0(1-a)}, \frac{1 - \frac{2}{2+\epsilon_0} + \max(3 - \frac{1}{a}, 0)}{1 - \max(3 - \frac{1}{a}, 0)}), \]

\[ a = \max(\frac{\epsilon_0 + 2\alpha}{(1-a)(2\epsilon_0 + 2)}, \frac{2 + 2\epsilon_0}{3\epsilon_0 + 4}, \frac{(2 + \epsilon_0)(1 + \max(3 - \frac{1}{a}, 0)) - 2}{2 + 2\epsilon_0}), \]

\[ \epsilon_0 = \min(1, 2 - \frac{2\alpha}{1-\alpha}). \]

d, \alpha are the parameters in Theorems 4.1 and 2.6.
Proof. From the subsection 4.2, we know that \( \gamma \) can be found by solving (4.12)-(4.14), so we summarize them here:

1. \( \min(\kappa, \frac{v}{2} - d\kappa) > 1, \) where \( \kappa > 1, v := \min(\frac{\gamma(1+c)}{2} - \kappa - c(1-a), \gamma(1+c) - 2\kappa - c(1-a) - ca \frac{\alpha}{1-\alpha}, 2\gamma(1+c) - c(1-a) - 4\kappa - 2ca, \frac{\gamma(1+c)(2+\epsilon_0)}{2} - \kappa(2+\epsilon_0) - c(1-a) - ca \frac{2+\epsilon_0}{2}, \gamma(1+c) - 2\kappa - c(1-a) - c \max(3 - \frac{1}{\alpha}, 0)), a \in (\frac{1}{2}, 1), \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha}), c > 1. \)

2. \( \gamma \frac{c+1}{2} > 1 + \frac{c+1}{2} - \min(\kappa, \frac{v}{2} - d\kappa). \)

3. \( c - \gamma(c + 1) < 0. \)

4. \( 1 + (c + 1)(\max(3 - \frac{1}{\alpha}, 0) - \gamma) < 0. \)

5. \( \frac{c}{\gamma(c+1)} < 1. \)

6. \( \frac{1}{2} \gamma(c+1)(2+\epsilon) - c(1 + \frac{v}{2}) > 1, \epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha}). \)

If \( v > 2(d + 1)\kappa, \) then \( \min(\kappa, \frac{v}{2} - \kappa d) = \kappa. \) So we use this to simplify 1,2 above to 1,2 below. Note that 3 and 5 above are the same, the inequality 6 above is \( \gamma > \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}, \epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha}). \)

Use the definition of \( v \) to change \( v > 2(d + 1)\kappa \) to 1,2,3,4,5 below, and move 2,3,4 above to 6,7,8 below, then we have

1. \( \frac{\gamma(1+c)}{2} - \kappa - c(1-a) > 2(d + 1)\kappa, a \in (\frac{1}{2}, 1), c > 1. \)

2. \( \gamma(1+c) - 2\kappa - c(1-a) - ca \frac{\alpha}{1-\alpha} > 2(d + 1)\kappa. \)

3. \( 2\gamma(1+c) - c(1-a) - 4\kappa - 2ca > 2(d + 1)\kappa. \)

4. \( \frac{\gamma(1+c)(2+\epsilon_0)}{2} - \kappa(2+\epsilon_0) - c(1-a) - ca \frac{2+\epsilon_0}{2} > 2(d + 1)\kappa, \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha}). \)

5. \( \gamma(1+c) - 2\kappa - c(1-a) - c \max(3 - \frac{1}{\alpha}, 0) > 2(d + 1)\kappa. \)

6. \( \gamma > 1 - \frac{2}{c+1} (\kappa - 1). \)
7. $\gamma > \frac{1}{c+1} + \max(3 - \frac{1}{\alpha}, 0)$.

8. $\gamma > \frac{c}{c+1} + \frac{2}{c+1} (1-a), a \in (\frac{1}{2}, 1), c > 1.$

We rewrite the inequalities above to represent the range of $\gamma$:

1. $\gamma > \frac{(4d+6)\kappa}{c+1} + \frac{2c}{c+1} (1-a), a \in (\frac{1}{2}, 1), c > 1.$

2. $\gamma > \frac{(2d+4)\kappa}{c+1} + \frac{c\alpha}{c+1} + \frac{c}{c+1}(1-a)$.

3. $\gamma > \frac{(d+3)\kappa}{c+1} + \frac{c(a+1)}{2(c+1)}.$

4. $\gamma > \frac{2(2d+4+\epsilon_0)}{(1+\epsilon)(2+\epsilon_0)} \kappa + \frac{2c+c\epsilon_0}{(c+1)(2+\epsilon_0)}, \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha}).$

5. $\gamma > \frac{2d+4}{1+c} \kappa + \frac{c(1-a)}{c+1} + \frac{c}{c+1} \max(3 - \frac{1}{\alpha}, 0)$.

6. $\gamma > 1 - \frac{2}{c+1} (\kappa - 1)$.

7. $\gamma > \frac{1}{c+1} + \max(3 - \frac{1}{\alpha}, 0)$.

8. $\gamma > \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}, \epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha})$.

Since $\frac{c}{c+1} < 1, \frac{2}{2+\epsilon} < 1, \kappa < 1, \epsilon_0 < 1$, we can change the range of $\gamma$ 1-5 above to 1-5 below:

1. $\gamma > \frac{(4d+6)\kappa}{c+1} + 2(1-a)$.

2. $\gamma > \frac{(2d+4)\kappa}{c+1} + \frac{\alpha}{1-\alpha} + (1-a)$.

3. $\gamma > \frac{(d+3)\kappa}{c+1} + \frac{(a+1)}{2}$.

4. $\gamma > \frac{2(2d+4+\epsilon_0)}{(1+\epsilon)(2+\epsilon_0)} \kappa + \frac{2c+c\epsilon_0}{(c+1)(2+\epsilon_0)}.$

5. $\gamma > \frac{2d+4}{1+c} \kappa + (1-a) + \max(3 - \frac{1}{\alpha}, 0), a \in (\frac{1}{2}, 1), \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha})$.

6. $\gamma > 1 - \frac{2}{c+1} (\kappa - 1)$.

7. $\gamma > \frac{1}{c+1} + \max(3 - \frac{1}{\alpha}, 0)$.

8. $\gamma > \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}, \epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha})$.

Let $\kappa := 2$, since $2d + 5 > d + 3, \frac{a+1}{2} < \frac{2+c\epsilon_0}{2+\epsilon_0}$, we can simplify inequalities 3,4 above to 3 below. Since $1 - \frac{2}{c+1} < \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}$, we can simplify the inequalities 6,8 above to 6 below:
1. $\gamma > \frac{2(4d+6)}{c+1} + 2(1-a)$.

2. $\gamma > \frac{2(2d+4)}{c+1} + \frac{a}{1-\alpha} + (1-a)$.

3. $\gamma > \frac{2(2d+5)}{1+c} + \frac{2+\alpha\epsilon_0}{2+\epsilon_0}$.

4. $\gamma > \frac{2(2d+4)}{1+c} + (1-a) + \max(3 - \frac{1}{\alpha}, 0), a \in (\frac{1}{2}, 1), \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha})$.

5. $\gamma > \frac{1}{c+1} + \max(3 - \frac{1}{\alpha}, 0)$.

6. $\gamma > \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}, \epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha})$.

Note that if $a > \frac{2+2\epsilon_0}{3\epsilon_0+4}$, then $\frac{2+\alpha\epsilon_0}{2+\epsilon_0} > 2(1-a)$.

If $a > \frac{\epsilon_0+2\alpha}{(1-\alpha)(2\epsilon_0+2)}$, then $\frac{2+\alpha\epsilon_0}{2+\epsilon_0} > \frac{\alpha}{1-\alpha} + 1 - a$.

If $a > \frac{(2+\epsilon_0)(1+\max(3 - \frac{1}{\alpha}, 0)) - 2}{2+2\epsilon_0}$, then $\frac{2+\alpha\epsilon_0}{2+\epsilon_0} > 1 - a + \max(0, 3 - \frac{1}{\alpha})$.

Therefore, use $4d+6 > \max\{2d+4, 2d+5\}$, when

$$a > \max\left(\frac{\epsilon_0 + 2\alpha}{(1-\alpha)(2\epsilon_0 + 2)}, \frac{2 + 2\epsilon_0}{3\epsilon_0 + 4}, \frac{(2 + \epsilon_0)(1 + \max(3 - \frac{1}{\alpha}, 0)) - 2}{2 + 2\epsilon_0}\right),$$

$$\epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha}),$$

we can simplify inequalities 1,2,3,4 above to 1 below and move 5,6 above to 2,3 below:

1. $\gamma > \frac{2(4d+6)}{1+c} + \frac{2+\alpha\epsilon_0}{2+\epsilon_0}$.

2. $\gamma > \frac{1}{c+1} + \max(3 - \frac{1}{\alpha}, 0)$.

3. $\gamma > \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}, \epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha})$.

Note that if $c > \frac{1-\frac{2}{2+\epsilon} + \max(3 - \frac{1}{\alpha}, 0)}{1 - \min(3 - \frac{1}{\alpha}, 0)}$, $\epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha})$, then

$$\frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)} > \frac{1}{c+1} + \max(3 - \frac{1}{\alpha}, 0).$$

If $c > \frac{2+\epsilon_0 a + (2+\epsilon_0)(8d+12) - \frac{2(2+\epsilon_0)}{2+\epsilon}}{\epsilon_0(1-\alpha)}$, then

$$\frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)} > \frac{2(4d+6)}{1+c} + \frac{2+\alpha\epsilon_0}{2+\epsilon_0}.$$

Therefore, if

$$c > \max\left(\frac{2 + \epsilon_0 a + (2 + \epsilon_0)(8d + 12) - \frac{2(2+\epsilon_0)}{2+\epsilon}}{\epsilon_0(1-\alpha)}, \frac{1-\frac{2}{2+\epsilon} + \max(3 - \frac{1}{\alpha}, 0)}{1 - \min(3 - \frac{1}{\alpha}, 0)}\right),$$

45
\[ a > \max \left( \frac{\epsilon_0 + 2\alpha}{(1 - \alpha)(2\epsilon_0 + 2)}, \frac{2 + 2\epsilon_0}{3\epsilon_0 + 4} \cdot \frac{(2 + \epsilon_0)(1 + \max(3 - \frac{1}{\alpha}, 0)) - 2}{2 + 2\epsilon_0} \right), \]

\[ \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1 - \alpha}), \]

the inequalities 1, 2, 3 above can be combined as follows:

\[ \gamma > \frac{c}{c + 1} + \frac{2}{(c + 1)(2 + \epsilon)}. \]

Let \( \epsilon_0 = \epsilon \), then

\[ \gamma > \frac{c}{c + 1} + \frac{2}{(c + 1)(2 + \epsilon_0)}. \]

Since this is a strict inequality for \( \gamma \), so \( c, a, \epsilon_0 \) can take the supremums or infinums, respectively, that is,

\[ c = \max \left( \frac{2 + \epsilon_0 a + (2 + \epsilon_0)(8d + 12) - 2}{\epsilon_0(1 - a)}, \frac{1 - \frac{2}{2 + \epsilon_0} + \max(3 - \frac{1}{\alpha}, 0)}{1 - \max(3 - \frac{1}{\alpha}, 0)} \right), \]

\[ a = \max \left( \frac{\epsilon_0 + 2\alpha}{(1 - \alpha)(2\epsilon_0 + 2)}, \frac{2 + 2\epsilon_0}{3\epsilon_0 + 4} \cdot \frac{(2 + \epsilon_0)(1 + \max(3 - \frac{1}{\alpha}, 0)) - 2}{2 + 2\epsilon_0} \right), \]

\[ \epsilon_0 = \min(1, 2 - \frac{2\alpha}{1 - \alpha}). \]

\[ \square \]

**Lemma 8.2** (Computations of the range of \( \gamma_1 \)).

From the proof of Corollary 3.1, \( \gamma_1 \) satisfies the following inequalities:

1. \( \frac{\gamma_1}{2} > 1 - a \).
2. \( \gamma_1 > 1 - a + a \frac{\alpha}{1 - \alpha} \).
3. \( 2\gamma_1 > 1 + a \).
4. \( \gamma_1 > (1 + a \frac{\alpha}{2}) \cdot \frac{2}{2 + \epsilon_0}, \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1 - \alpha}) \).
5. \( \gamma_1 > 1 - a + \max(0, 3 - \frac{1}{\alpha}) \).

Then \( \gamma_1 \) can be any number in \((\frac{2 + \alpha\epsilon_0}{2 + \epsilon_0}, 1)\), where

\[ a = \max \left( \frac{\epsilon_0 + (2 + \epsilon_0) \max(0, 3 - \frac{1}{\alpha})}{2 + 2\epsilon_0}, \frac{\epsilon_0}{2 + \epsilon_0 + \frac{1 - 2\alpha}{1 - \alpha}}, \frac{2 + 2\epsilon_0}{4 + 5\epsilon_0} \right), \epsilon_0 = \min(1, 2 - \frac{2\alpha}{1 - \alpha}). \]
Proof. Since \( \frac{1+a}{2} < \frac{2+\alpha a}{2+\epsilon_0} \), then \( \gamma_1 > (1 + \epsilon_0 \cdot \frac{2}{2+\epsilon_0} > \frac{1+a}{2} \).

If \( a > \frac{2+2a}{4+\epsilon_0} \), then \( (1 + \epsilon_0 \cdot \frac{2}{2+\epsilon_0} > 2(1 - a) \).

If \( a > \frac{\epsilon_0}{2+\epsilon_0} \cdot \frac{1-a}{1-a} \), then \( (1 + \epsilon_0 \cdot \frac{2}{2+\epsilon_0} > 1 - a + \frac{\alpha}{1-a} \).

If \( a > \frac{\epsilon_0 + (2+\epsilon_0) \max(0,3 - \frac{1}{\alpha})}{2+\epsilon_0} \), then \( (1 + \epsilon_0 \cdot \frac{2}{2+\epsilon_0} > 1 - a + \max(0,3 - \frac{1}{\alpha}) \).

So when \( a > \max(\frac{\epsilon_0 + (2+\epsilon_0) \max(0,3 - \frac{1}{\alpha})}{2+\epsilon_0}, \frac{\epsilon_0}{2+\epsilon_0 + \frac{1-2\alpha}{1-\alpha}}, 2+2\epsilon_0, \frac{2+2\epsilon_0}{4+5\epsilon_0} \) \) and \( \epsilon_0 < \min(1, 2 - \frac{2\alpha}{1-\alpha}) \),

\[ \gamma_1 > \frac{2 + \alpha \epsilon_0}{2 + \epsilon_0}. \]

Since this is a strict inequality for \( \gamma_1 \), so \( a, \epsilon_0 \) can take the supremums or infimums, respectively, that is, \( \gamma_1 \) can be any number in \( (\frac{2+\alpha \epsilon_0}{2+\epsilon_0}, 1) \), where

\[ a = \max(\frac{\epsilon_0 + (2+\epsilon_0) \max(0,3 - \frac{1}{\alpha})}{2+\epsilon_0}, \frac{\epsilon_0}{2+\epsilon_0 + \frac{1-2\alpha}{1-\alpha}}, 2+2\epsilon_0, \frac{2+2\epsilon_0}{4+5\epsilon_0}), \epsilon_0 = \min(1, 2 - \frac{2\alpha}{1-\alpha}) \].

\[ \square \]

Lemma 8.3 (Transfer, see [Kal02] Theorem 6.10).

For any measurable space \( S \) and Borel space \( T \), let \( \xi \stackrel{d}{=} \xi' \) and \( \eta \) be random elements in \( S \) and \( T \), respectively (that is, \( \xi \) and \( \eta \) are defined on the same probability space, \( \xi \) and \( \xi' \) have the same distribution but are not necessarily defined on the same probability space). Then there exists a random element \( \eta' \) in \( T \) with

\[ (\eta, \xi) \stackrel{d}{=} (\eta', \xi'). \]

More precisely, there exists a measurable function \( f : S \times [0,1] \rightarrow T \) such that \( \eta' = f(\xi', U) \) where \( U \sim U(0,1) \) and \( \xi' \) are independent.

Indeed, to guarantee the independence above, we can simply extend the probability space by multiplying an interval \( (I, \text{Leb}) \).

Lemma 8.4 (Embedding in a \( d \)-dimensional Brownian motion).

If \( (\phi_k \circ T^k)_{k \geq 1} \) satisfies the VASIP, and there is a constant \( \epsilon \in (0,1/2) \) and a positive definite \( d \times d \) matrix \( \Sigma^2 \) such that \( \sigma_n^2 = n \cdot \Sigma^2 + o(n^{1-\epsilon}) \), then there is a constant \( \bar{\epsilon} \in (0,1/2) \) and a standard \( d \)-dimensional Brownian motion \( B_t \) such that

\[ \sum_{k=1}^{n} \phi_k \circ T^k - \Sigma \cdot B_n = o(n^{1/2-\bar{\epsilon}}) \text{ a.s.} \]

Proof. Since \( d = 1 \) is trivial, we assume that \( d > 1 \). By the Definition 2.3, we have:

\[ \sum_{k=1}^{n} \phi_k \circ T^k - \sum_{k=1}^{n} G_k = o(n^{1/2-\epsilon}) \text{ a.s.}, \]

47
\[
\sigma_n^2 = \int \left( \sum_{k=1}^{n} \phi_k \circ T^k \right) \cdot \left( \sum_{k=1}^{n} \phi_k \circ T^k \right)^T d\mu = \sum_{k=1}^{n} \tilde{E}(G_k \cdot G_k^T) + o(n^{1-\epsilon}),
\]

where \( \tilde{E} \) is the expectation of the probability \( \tilde{P} \) of the extended probability space \((X, \mathcal{B}, \mu)\). Then
\[
\sum_{k \leq n} \tilde{E}(G_k \cdot G_k^T) = n \cdot \Sigma^2 + o(n^{1-\epsilon}).
\]

Without loss of generality, we assume that \( \Sigma^2 = I_{d \times d} \).

Let \( c \in \mathbb{N} \) (will be given later), then
\[
\sum_{k=n^c+1}^{(1+n)^c} \tilde{E}(G_k \cdot G_k^T) = [(1 + n)^c - n^c] \cdot I_{d \times d} + o((n + 1)^c(1-\epsilon)). \tag{8.1}
\]

If \( c \) is large enough such that \( c - 1 > c(1 - \epsilon) \), then
\[
\frac{\sum_{k=n^c+1}^{(1+n)^c} \tilde{E}(G_k \cdot G_k^T)}{(1 + n)^c - n^c} - I_{d \times d} = \frac{o((n + 1)^c(1-\epsilon))}{n^{c-1}} = o(n^{1-c\epsilon}).
\]

We denote
\[
A := \sum_{k=n^c+1}^{(1+n)^c} \tilde{E}(G_k \cdot G_k^T) = Q_n \cdot \begin{bmatrix} 
\lambda_1^n & 0 & \cdots & 0 \\
0 & \lambda_2^n & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_d^n
\end{bmatrix} \cdot Q_n^T, \tag{8.2}
\]

where \( \lambda_1^n \leq \lambda_2^n \leq \cdots \leq \lambda_d^n \) are eigenvalues, \( Q_n \) is an orthogonal matrix, and denote
\[
A_1 := Q_n \cdot \begin{bmatrix}
\min(\lambda_1^n, (1+n)^c-n^c) & 0 & \cdots & 0 \\
0 & \min(\lambda_2^n, (1+n)^c-n^c) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \min(\lambda_d^n, (1+n)^c-n^c)
\end{bmatrix} \cdot Q_n^T,
\]
\[
A_2 := A - A_1,
\]
\[
A_3 := ((1+n)^c-n^c) \cdot I_{d \times d} - A_1.
\]

For each \( n \), pick arbitrary independent Gaussian vectors \( \tilde{g}_1^{n+1}, \tilde{g}_2^{n+1}, \tilde{g}_3^{n+1} \) such that
\[
\tilde{E}[\tilde{g}_1^{n+1} \cdot (\tilde{g}_1^{n+1})^T] = A_1, \quad \tilde{E}[\tilde{g}_2^{n+1} \cdot (\tilde{g}_2^{n+1})^T] = A_2, \quad \tilde{E}[\tilde{g}_3^{n+1} \cdot (\tilde{g}_3^{n+1})^T] = A_3.
\]
Therefore \( g_1^{n+1} + g_2^{n+1} \overset{d}{=} \sum_{k=n^c+1}^{(1+n)^c} G_k \). By Lemma 8.3, there are zero-mean independent Gaussian vectors \( g_1^{n+1}, g_2^{n+1}, g_3^{n+1} \) (extend the probability space if necessary, still denote its probability by \( P \) and its expectation by \( \mathbb{E} \)) such that

\[
(g_1^{n+1} + g_2^{n+1}, g_1^{n+1}, g_2^{n+1}, g_3^{n+1}) \overset{d}{=} (\sum_{k=n^c+1}^{(1+n)^c} G_k, g_1^{n+1}, g_2^{n+1}, g_3^{n+1}).
\]

Therefore,

\[
\mathbb{E}[g_1^{n+1} \cdot (g_1^{n+1})^T] = A_1, \mathbb{E}[g_2^{n+1} \cdot (g_2^{n+1})^T] = A_2, \mathbb{E}[g_3^{n+1} \cdot (g_3^{n+1})^T] = A_3,
\]

\[
\sum_{k=n^c+1}^{(1+n)^c} G_k = g_1^{n+1} + g_2^{n+1} \text{ a.s.},
\]

\[
\mathbb{E}[(g_1^{n+1} + g_3^{n+1}) \cdot (g_1^{n+1} + g_3^{n+1})^T] = [(1+n)^c - n^c] \cdot I_{d \times d}.
\]

Furthermore, since \( A_2 \) and \( A_3 \), after being diagonalized by \( Q_n \), have nonzero numbers on disjoint entries of the diagonal line. Therefore by (8.1),

\[
\mathbb{E}[g_2^{n+1} \cdot (g_2^{n+1})^T] = o((n+1)^c(1-\epsilon)),
\]

\[
\mathbb{E}[g_3^{n+1} \cdot (g_3^{n+1})^T] = o((n+1)^c(1-\epsilon)).
\]

By Lemma 8.3, we know \( g_i^n \) depends on \( \sum_{n^c+1}^{(1+n)^c} G_k \). Since for any \( n_1 \neq n_2 \in \mathbb{N} \), \( \sum_{k=n_1^c+1}^{(1+n_1)^c} G_k \) and \( \sum_{k=n_2^c+1}^{(1+n_2)^c} G_k \) are independent, then

\[
(g_1^{n_1+1}, g_2^{n_1+1}, g_3^{n_1+1}) \text{ and } (g_1^{n_2+1}, g_2^{n_2+1}, g_3^{n_2+1}) \text{ are independent.}
\]

So there is a Brownian motion \( B_t \) such that for each \( n \in \mathbb{N} \),

\[
g_1^{n+1} + g_3^{n+1} = B_{(1+n)^c} - B_{n^c}.
\]

Therefore

\[
\sum_{k=n^c+1}^{(1+n)^c} G_k - (B_{(1+n)^c} - B_{n^c}) = g_2^{n+1} - g_3^{n+1} \text{ a.s.},
\]

\[
\sum_{k=1}^{n^c} G_k - B_{n^c} = \sum_{i=1}^{n} (g_2^i - g_3^i) \text{ a.s.}
\]

For any \( m \in \mathbb{N} \), there is \( n \) such that \( n^c < m \leq (n+1)^c \) and

\[
\sum_{k=1}^{m} G_k - B_m = \sum_{k=1}^{n^c} G_k - B_{n^c} + \sum_{k=1+n^c}^{m} G_k - (B_m - B_{n^c})
\]
\begin{align*}
&= \sum_{i=1}^{n} (g_{i}^2 - g_{3}^2) + \sum_{k=n+1}^{m} G_{k} - (B_{m} - B_{n}) \\
&\leq \left| \sum_{i=1}^{n} (g_{i}^2 - g_{3}^2) \right| + \sup_{n^{c} < m \leq (n+1)^{c}} \left| \sum_{k=n+1}^{m} G_{k} \right| + \sup_{n^{c} < m \leq (n+1)^{c}} \left| B_{m} - B_{n} \right|.
\end{align*}

To estimate the last two terms, without loss of generality, we assume that these two terms are Gaussian random variables. Then if \(2c \bar{\epsilon} < 1, \bar{\epsilon} < \epsilon,\)

\[
\tilde{P}(\sup_{n^{c} < m \leq (n+1)^{c}} \left| \sum_{k=n+1}^{m} G_{k} \right| > n^{c(1/2-\bar{\epsilon})}) \leq \tilde{P}(\sup_{n^{c} < m \leq (n+1)^{c}} \left| B_{m} - B_{n} \right| > n^{c(1/2-\bar{\epsilon})})
\]

\[
\lesssim \int_{|t| > \frac{n^{c(1/2-\epsilon)}}{\sqrt{\text{Var}(\sum_{k=n+1}^{m} G_{k})}}} e^{-t^{2}/2} dt \lesssim \int_{|t| > \frac{n^{c(1/2-\bar{\epsilon})}}{\sqrt{n^{c(1-\epsilon)/2}}}} e^{-t^{2}/2} dt \lesssim n^{-2},
\]

\[
\tilde{P}(\sup_{n^{c} < m \leq (n+1)^{c}} \left| B_{m} - B_{n} \right| > n^{c(1/2-\bar{\epsilon})}) \leq \tilde{P}(\sup_{n^{c} < m \leq (n+1)^{c}} \left| B_{(n+1)^{c}} - B_{n} \right| > n^{c(1/2-\bar{\epsilon})})
\]

\[
\lesssim \int_{|t| > \frac{n^{c(1/2-\bar{\epsilon})}}{\sqrt{(n+1)^{c} - n^{c}}}} e^{-t^{2}/2} dt \lesssim n^{-2}.
\]

The estimates of \(g_{2}^i\) and \(g_{3}^i\) are the same, so we just estimate \(g_{2}^i\):

\[
\tilde{P}(\sum_{i=1}^{n} g_{i}^2 > n^{c(1/2-\bar{\epsilon})}) \lesssim \int_{|t| > \frac{n^{c(1/2-\bar{\epsilon})}}{\sqrt{\text{Var}(\sum_{i=1}^{n} g_{i}^2)}}} e^{-t^{2}/2} dt \lesssim \int_{|t| > \frac{n^{c(1/2-\bar{\epsilon})}}{\sqrt{n^{c(1-\bar{\epsilon})}/2}}} e^{-t^{2}/2} dt \lesssim n^{-2}.
\]

By the Borel-Cantelli Lemma,

\[
\sum_{k=1}^{m} G_{k} - B_{m} \leq \left| \sum_{i=1}^{n} (g_{2} - g_{3}^i) \right| + \sup_{n^{c} < m \leq (n+1)^{c}} \left| \sum_{k=n+1}^{m} G_{k} \right| + \sup_{n^{c} < m \leq (n+1)^{c}} \left| B_{m} - B_{n} \right|
= o(n^{c(1/2-\bar{\epsilon})}) = o(m^{1/2-\bar{\epsilon}}) \text{ a.s.}
\]

\section*{References}

[AHN+15] Romain Aimino, Huyi Hu, Matthew Nicol, Andrei Török, and Sandro Vaienti, \textit{Polynomial loss of memory for maps of the interval with a neutral fixed point}, Discrete Contin. Dyn. Syst. 35 (2015), no. 3, 793–806. MR 3277171

[BB84] Daniel Berend and Vitaly Bergelson, \textit{Ergodic and mixing sequences of transformations}, Ergodic Theory Dynam. Systems 4 (1984), no. 3, 353–366. MR 776873
István Berkes and Walter Philipp, *Approximation theorems for independent and weakly dependent random vectors*, Ann. Probab. 7 (1979), no. 1, 29–54. MR 515811

C. Cuny, J. Dedecker, A. Korepanov, and F. Merlevède, *Rates in almost sure invariance principle for quickly mixing dynamical systems*, Stoch. Dyn. 20 (2020), no. 1, 2050002, 28. MR 4066797

C. Cuny, J. Dedecker, A. Korepanov, and F. Merlevède, *Rates in almost sure invariance principle for slowly mixing dynamical systems*, Ergodic Theory Dynam. Systems 40 (2020), no. 9, 2317–2348. MR 4130806

Jean-Pierre Conze and Albert Raugi, *Limit theorems for sequential expanding dynamical systems on [0,1]*, Ergodic theory and related fields, Contemp. Math., vol. 430, Amer. Math. Soc., Providence, RI, 2007, pp. 89–121. MR 2331327

D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti, *Almost sure invariance principle for random piecewise expanding maps*, Nonlinearity 31 (2018), no. 5, 2252–2280. MR 3816673

D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti, *A spectral approach for quenched limit theorems for random expanding dynamical systems*, Comm. Math. Phys. 360 (2018), no. 3, 1121–1187. MR 3803820

Sébastien Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probab. Theory Related Fields 128 (2004), no. 1, 82–122. MR 2027296

Sébastien Gouëzel, *Almost sure invariance principle for dynamical systems by spectral methods*, Ann. Probab. 38 (2010), no. 4, 1639–1671. MR 2663640

Olli Hella and Juho Leppänen, *Central limit theorems with a rate of convergence for time-dependent intermittent maps*, arXiv e-prints (2018), arXiv:1811.11170.

Nicolai Haydn, Matthew Nicol, Andrew Török, and Sandro Vaienti, *Almost sure invariance principle for sequential and non-stationary dynamical systems*, Trans. Amer. Math. Soc. 369 (2017), no. 8, 5293–5316. MR 3646763

Olli Hella and Mikko Stenlund, *Quenched normal approximation for random sequences of transformations*, arXiv e-prints (2018), arXiv:1810.10760.
[Kal02] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169

[Kor18] Alexey Korepanov, *Rates in almost sure invariance principle for dynamical systems with some hyperbolicity*, Comm. Math. Phys. 363 (2018), no. 1, 173–190. MR 3849987

[KP80] J. Kuelbs and Walter Philipp, *Almost sure invariance principles for partial sums of mixing B-valued random variables*, Ann. Probab. 8 (1980), no. 6, 1003–1036. MR 602377

[Liv96] Carlangelo Liverani, *Central limit theorem for deterministic systems*, International Conference on Dynamical Systems (Montevideo, 1995), Pitman Res. Notes Math. Ser., vol. 362, Longman, Harlow, 1996, pp. 56–75. MR 1460797

[LM13] Stefano Luzzatto and Ian Melbourne, *Statistical properties and decay of correlations for interval maps with critical points and singularities*, Comm. Math. Phys. 320 (2013), no. 1, 21–35. MR 3046988

[LSV99] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems 19 (1999), no. 3, 671–685. MR 1695915

[MN05] Ian Melbourne and Matthew Nicol, *Almost sure invariance principle for nonuniformly hyperbolic systems*, Comm. Math. Phys. 260 (2005), no. 1, 131–146. MR 2175992

[MN09] ———, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, Ann. Probab. 37 (2009), no. 2, 478–505. MR 2510014

[MT04] Ian Melbourne and Andrei Török, *Statistical limit theorems for suspension flows*, Israel J. Math. 144 (2004), 191–209. MR 2121540

[NPT19] Matthew Nicol, Felipe Perez Pereira, and Andrew Torok, *Large deviations and central limit theorems for sequential and random systems of intermittent maps*, to appear in Ergodic Theory Dynam. Systems (2019), arXiv:1909.07435.

[NTV18] Matthew Nicol, Andrew Török, and Sandro Vaienti, *Central limit theorems for sequential and random intermittent dynamical systems*, Ergodic Theory Dynam. Systems 38 (2018), no. 3, 1127–1153. MR 3784257

52
[OSY09] William Ott, Mikko Stenlund, and Lai-Sang Young, *Memory loss for time-dependent dynamical systems*, Math. Res. Lett. **16** (2009), no. 3, 463–475. MR 2511626

[Ryc83] Marek Rychlik, *Bounded variation and invariant measures*, Studia Math. **76** (1983), no. 1, 69–80. MR 728198

[Ser68] R. J. Serfling, *Contributions to central limit theory for dependent variables*, Ann. Math. Statist. **39** (1968), 1158–1175. MR 0228053

[Ser70] ______, *Moment inequalities for the maximum cumulative sum*, Ann. Math. Statist. **41** (1970), 1227–1234. MR 0268938

[Su19a] Yaofeng Su, *Almost surely invariance principle for non-stationary and random intermittent dynamical systems*, Discrete Contin. Dyn. Syst. **39** (2019), no. 11, 6585–6597. MR 4026993

[Su19b] Yaofeng Su, *Random Young Towers and Quenched Limit Laws*, arXiv e-prints (2019), arXiv:1907.12199.

[You99] Lai-Sang Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188. MR 1750438