LIMIT THEOREMS FOR NON-STATIONARY AND RANDOM DYNAMICAL SYSTEMS

by
Yaofeng Su

A dissertation submitted to the Department of Mathematics,

College of Natural Sciences and Mathematics

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Mathematics

Chair of Committee: Andrew Török

Committee Member: Vaughn Climenhaga

Committee Member: Gemunu Gunaratne

Committee Member: Matthew Nicol

University of Houston
May 2020

ACKNOWLEDGMENTS

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study. First, my utmost gratitude to my advisor Dr. Andrew Török for the support and help he provided me throughout my graduate education; my committee members, Dr. Vaughn Climenhaga, Dr. Gemunu Gunaratne, and Dr. Matthew Nicol for their participation and their thoughtful comments; all professors in the dynamical system research group of University of Houston, Dr. Vaughn Climenhaga, Dr. Alan Haynes, Dr. Matthew Nicol, Dr. William Ott, and Dr. Andrew Török for organizing the wonderful "Houston Summer School on Dynamical Systems", where I learned a lot about dynamical systems, ergodic theory, number theory, and quantum mechanics.

ABSTRACT

We study the limit behavior of the non-stationary/random chaotic dynamical systems and prove a strong statistical limit theorem: (vector-valued) almost sure invariance principle for the non-stationary dynamical systems and quenched (vector-valued) almost sure invariance principle for the random dynamical systems. It is a matching of the trajectories of the dynamical system with a Brownian motion in such a way that the error is negligible in comparison with the Birkhoff sum. We develop a method called "reverse Gaussian approximation" and apply it to the classical block construction. We apply our results to the stationary chaotic systems which can be described by the Young tower, and the (non)uniformly expanding non-stationary/random dynamical systems with intermittency or uniform spectral gap. Our results imply that the systems we study have many limit results that are satisfied by Brownian motion.

TABLE OF CONTENTS

1 Introduction

1.1 History review

Dynamical systems is a math subject describing the long term evolution of systems for which the 'infinitesimal' evolution rule is known. Examples and applications arise from all branches of science and technology, like physics, chemistry, economics, ecology, communications, biology, computer science, and meteorology, to mention just a few.

These systems have in common the fact that each possible state may be described by a finite (or infinite) number of observable quantities. Usually, there are some constraints between these quantities. So the space of states X (called phase space) is often a manifold.

For the continuous time systems, the evolution rule can be a differential equation: each state $x \in X$ associates the speed and direction in which the system is going to evolve. Even when the real phenomenon is supposed to evolve in continuous time, it is convenient to consider a discrete time model-discrete time systems. In this case, the evolution rule is a transformation $T : X \to X$, assigning to the present state $x \in X$ the next state $T(x)$ after one unit of time. In both cases, the main problem of dynamical system is describing the behavior as time goes to infinity for the majority of orbit.

The term non-stationary dynamical systems, introduced by Berend and Bergelson [7], refers to a (non-stationary) system in which a sequence of concatenation of maps $T_k \circ T_{k-1} \circ \cdots \circ T_1$ acts on the state space X, where the maps $T_i : X \to X$ are allowed to vary with i. For more discussions about its behavior as time goes to infinity for the majority of orbit, see [10, 40].

One interesting limit behavior is called almost sure invariance principle (ASIP): it is a matching of the trajectories of the dynamical system with a Brownian motion in such a way that the error is negligible in comparison with the Birkhoff sum. Limit theorems such as the Central Limit Theorem (CLT), the Functional Central Limit Theorem (FCLT), the Law of the Iterated Logarithm (LIL) etc. transfer from the Brownian motion to time-series generated by observables on the dynamical system. These kinds of results for the stationary dynamical systems have a lot of consequences (see, e.g., Melbourne, Nicol, and Stenlund [34, 35, 47]).

Such results were given by [11, 12, 20, 23, 37, 48, 49] for one-dimensional processes under the stationary or non-stationary dynamical systems and by [16,52] for independent higher-dimensional processes; however, for higher-dimensional dependent processes, difficulties arise since the techniques relying on Skorokhod embedding [42] do not work efficiently. In this direction, an approximation argument was introduced by Berkes, Philipp, and Stout [8, 41] and was generalized by Kuelbs and Philipp [26]. Using Kuelbs and Philipp's result, Melbourne and Nicol [35] obtained the vector-valued almost sure invariance principle (VASIP) for the (non)uniformly hyperbolic/expanding stationary dynamical systems by choosing a suitable filtration. Their proof relies on the Young tower and the tower technique developed by Melbourne and Török in [36]; hence it works very well for the stationary dynamical systems when they have some Markovian behavior and sufficient hyperbolicity.

Unfortunately, it is quite common to encounter the stationary dynamical systems for which there is no natural well-behaved filtration. Gouëzel [19] showed that a sufficient control condition on the characteristic functions of a process implies the VASIP. This condition is easy to check for large classes of dynamical systems or Markov chains using strong or weak spectral perturbation arguments. His method relies on the good spectrum of the quasi-compact transfer operator acting on a suitable Banach space (also known as spectral gap). The advantage of his result is that the invariant density is not required to be bounded away from zero. This helps Luzzatto and Melbourne [33] to obtain the VASIP successfully for a large class of (non)uniformly expanding interval maps with singularities.

However, if the dynamical system is non-stationary, the tower technique in [34–36] no longer works. If the transfer operator of the dynamic has no spectral gap on the space of functions that are of interest, Gouëzel's approach fails to work. Such examples and their related statistical properties are provided in the following papers [1, 10, 13–15, 22, 27, 28, 38, 39].

Conze and Raugi [10] considered the composition of a family of uniformly expanding interval maps, extended the spectral theory of single transfer operator to the case of a sequence of transfer operators, and developed a martingale technique based on Gordin's approach [18] (decompose Birkhoff sum as reverse martingale differences plus an error term) to prove CLT. Dragičević, Froyland, González-Tokman, and Vaienti [14, 15] and Haydn, Nicol, Török, and Vaienti [20] used this technique to prove scalar ASIP for the non-stationary/random dynamical systems with exponential decay of correlation under a variance growth assumption.

For the case without spectral gap, Aimino, Hu, Nicol, Török, Vaienti [1] considered the composition of a family of Pomeau-Manneville like maps, obtained by perturbing the slope at the indifferent fixed point 0. They got polynomial decay of correlations for C^1-observables. Nicol, Török, Vaienti [39] considered the same system, used the martingale technique in [10] and Sprindžuk's Theorem [45] to prove self-norming CLT under the assumptions that the system has sufficiently fast decay of correlation and the variance grows at a certain rate. Moreover, they proved self-norming CLT for nearby maps in this family and quenched CLT for random compositions of finitely many maps in this family under an assumption of fast decay of correlation.

However, this martingale technique causes a new problem: the estimate of the error terms requires a sufficiently fast decay of correlation, as [20, 39] did, so the limit laws are not known if the decay of correlation of the dynamical system is too slow. Therefore, in this dissertation, martingales are not used; instead, we obtain the VASIP for the non-stationary dynamical systems by setting up several dynamical inequalities and a new approximation method (modify [8] to work for arbitrary decreasing filtration, in particular, dynamical systems). The three assumptions on the decay of correlations (A4)-(A6) in our Theorem 2.8 are quite natural under the setting of dynamical systems. As applications, we apply our results to a large class of the non-stationary dynamical systems in [10, 20, 39] and the random systems in [14, 15, 39]. In particular, we obtain the optimal range ($\alpha < \frac{1}{2}$) for the Theorem 3.1 in [39]. We also recover the classical stationary results in [19, 34, 35, 51]. See section 3, Applications.

Finally, we want to compare our results with Gouëzel [19], Melbourne and Nicol [34, 35] in the stationary dynamics setting:

- Advantage: compared to assumption (H) in [19], our condition is simpler, automatically satisfied by the stationary dynamical systems in [19]. Same as [19], our VASIP result does not require invariant density to be bounded away zero. Unlike [19, 34, 35], our result shows that we do not need to deal with different types of systems (Young tower or spectral gap) separately, that is, we can prove the VASIP for the stationary systems only by verifying the three assumptions (A4)-(A6) in Theorem 2.8. In section 3 and section 6, we will show that the Young tower and spectral gap satisfy assumptions (A4)-(A6).

- Disadvantage: we do not have an explicit formula or a good estimate for the VASIP error rate (the difference between the Birkhoff sum and Brownian motion), while [35] obtains an explicit formula for the Young towers and [19] obtains $O(n^{\frac{1}{4}})$ for the dynamics with spectral gap, which is independent of the dimension of observables. The reason for this drawback is that we consider the systems under several assumptions on the slow decay of correlations (especially for the non-uniformly expanding non-stationary dynamical systems in [1]). The parameters in our Theorem 2.8, variance growth rate (the γ in Lemma 7.1) and the VASIP error rate, are far from optimal.

1.2 Outline of this dissertation

- In section 2, we will give our main theorems.
- In section 3, several corollaries of our main theorems are given as applications.
- In section 4, we obtain several technical lemmas which will be used in the proofs of the main theorems.
- In section 5, we give the proofs of our main theorems.
- Section 6 contains the proofs of corollaries in section 3.
- Section 7, Appendix, mainly focuses on the computation of the parameters in Theorem 2.8.

2 Definitions, notations, main theorems

2.1 Definitions and notations

Consider a probability space $(X, \mathcal{B}, \mu)$ with μ as a reference probability and a family of non-singular (w.r.t. μ) maps $T_k : X \to X, k \geq 1$. For any $n, m, k \in \mathbb{N}$, denote:

$$T_m^{n+m} := T_{n+m} \circ T_{n+m-1} \circ \cdots \circ T_m,$$

$$T^n := T_1^n = T_n \circ T_{n-1} \circ \cdots \circ T_1.$$

The transfer operator (Perron-Frobenius operator) P_k associated to T_k is defined by the duality relation:

$$\int g \cdot P_k f d\mu = \int g \circ T_k \cdot f d\mu \text{ for all } f \in L^1, g \in L^\infty.$$

Similar to T_m^{n+m} and T^n, denote:

$$P_m^{n+m} := P_{n+m} \circ P_{n+m-1} \circ \cdots \circ P_m,$$

$$P^n := P_1^n = P_n \circ P_{n-1} \circ \cdots \circ P_1.$$

Notation 2.1

1. *$a_n \approx b_n$ (resp. "$a_n \precsim b_n$") means there is a constant $C \geq 1$ such that $C^{-1} \cdot b_n \leq a_n \leq C \cdot b_n$ for all n (resp. $a_n \leq C \cdot b_n$ for all n).*
2. *C_a denotes a constant that depends only on a.*
3. ***1** denotes the constant function **1** on X.*
4. *For any $m \in \mathbb{N}$, scalar function f and L^1-matrix $\big[f_{ij}\big]$ (i.e., $f_{ij} \in L^1(X)$ for all $i, j \geq 1$), define:*

$$f \cdot P_m(\big[f_{ij}\big]) = P_m(\big[f_{ij}\big]) \cdot f := \big[f \cdot P_m(f_{ij})\big].$$

5. *From now on, let* $\{\phi_k \in L^\infty(X, \mu; \mathbb{R}^d) : k \in \mathbb{N}\}$ *be the observables satisfying*

$$\sup_k ||\phi_k||_{L^\infty} < \infty,$$

$$\int \phi_k \circ T^k d\mu = 0.$$

Definition 2.2 *For* $(T^n)_{n\geq 1} := (T_1^n)_{n\geq 1}$ *defined above, we have a decreasing filtration* $(T^{-n}\mathcal{B})_{n\geq 1}$. *Denote conditional expectation w.r.t.* $T^{-n}\mathcal{B}$ *and* μ *by:*

$$\mathbb{E}_n(\cdot) := \mathbb{E}(\cdot|T^{-n}\mathcal{B}).$$

In particular, the expectation, that is, conditional expectation w.r.t. $\{\emptyset, X\}$ *and* μ*, is denoted by:*

$$\mathbb{E}(\cdot) := \int (\cdot) d\mu.$$

Definition 2.3 (Non-stationary, stationary and random dynamical systems)

$(X, \mathcal{B}, (T_k)_{k\geq 1}, \mu)$ *is called a non-stationary dynamical system, where* $(T_k)_{k\geq 1}$ *are non-singular maps on* $(X, \mathcal{B}, \mu)$ *as stated above. In contrast, a stationary dynamical system means that* $T_k = T_1$ *for all* $k \geq 1$ *and* $(T_1)_*\mu = \mu$.

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

1. $\sigma : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Omega, \mathcal{F}, \mathbb{P})$ *is an invertible ergodic transformation preserving the probability* $\mathbb{P}$.

2. $(T_\omega)_{\omega\in\Omega}$ *are non-singular (w.r.t.* μ*) maps on* $(X, \mathcal{B}, \mu)$.

3. *The probability* μ_ω *on* X *is absolutely continuous w.r.t.* μ.

4. $(T_\omega)_*\mu_\omega = \mu_{\sigma\omega}$ *a.e.* $\omega \in \Omega$.

Definition 2.4 (VASIP for non-stationary dynamical system, see [20])

For the non-stationary dynamical system $(X, \mathcal{B}, (T_k)_{k\geq 1}, \mu)$ *with the observables* $(\phi_k)_{k\in\mathbb{N}}$*, denote* *the $d \times d$ variance matrix by*

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

Denote the least eigenvalue of σ_n^2 by

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

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

$$\sum_{k \leq n} \phi_k \circ T^k - \sum_{k \leq n} G_k = o(\lambda(\sigma_n^2)^{\frac{1-\epsilon}{2}}) \text{ a.s.}, \tag{2.1}$$

$$\int (\sum_{k \leq n} \phi_k \circ T^k) \cdot (\sum_{k \leq n} \phi_k \circ T^k)^T d\mu = \sum_{k \leq n} \tilde{\mathbb{E}}(G_k \cdot G_k^T) + o(\lambda(\sigma_n^2)^{1-\epsilon}), \tag{2.2}$$

$$\lambda(\sigma_n^2) \to \infty, \tag{2.3}$$

where $\tilde{\mathbb{E}}(\cdot)$ in (2.2) means the expectation w.r.t. the probability $\tilde{P}$ of the extended probability space of $(X, \mathcal{B}, \mu)$.

Remark 2.5

1. *Some G_k can be zero, since the zero Gaussian random vector is independent of any random vectors.*

2. *In general, $\sum_{k \leq n} G_k$ does not form a d-dimensional Brownian motion due to the non-stationary process $(\phi_k \circ T^k)_{k \geq 1}$. However, by Lemma 7.4, if there is a constant $\epsilon \in (0,1)$ and a positive definite $d \times d$ matrix $\sigma^2 > 0$, such that $\sigma_n^2 = n \cdot \sigma^2 + o(n^{1-\epsilon})$, then $(G_k)_{k \geq 1}$ can be modified as i.i.d. Gaussian vectors with covariance σ^2, and $\sum_{k \leq n} G_k$ would be replaced by d-dimensional Browian motion stopped at time n. So the VASIP, in this case, coincides with the classical*

VASIP for the stationary dynamical system. This remark says that (2.2) is an equivalence relation: the Gaussian vectors in the definition of the VASIP rely on the behavior of variance growth (2.2) and covariance σ^2 only.

3. *If $d = 1$, G_k can be embedded into a Brownian motion, and Definition 2.4 becomes scalar ASIP, that is, there is a matching of the Birkhoff sum $\sum_{k\leq n} \phi_k \circ T^k$ with a standard Brownian motion B_t observed at time $\tilde{\mathbb{E}}[(\sum_{k\leq n} G_k)^2]$ so that $\sum_{k\leq n} \phi_k \circ T^k$ equals $B_{\tilde{\mathbb{E}}[(\sum_{k\leq n} G_k)^2]}$ plus negligible error almost surely, which equals $B_{\sigma_n^2}$ plus a negligible error almost surely. Then $(\phi_k \circ T^k)_{k\geq 1}$ also satisfies self-norming CLT and LIL:*

$$\frac{\sum_{k\leq n} \phi_k \circ T^k}{\sigma_n} \xrightarrow{d} N(0,1),$$

$$\limsup_{n\to\infty} \frac{\sum_{k\leq n} \phi_k \circ T^k}{\sqrt{\sigma_n^2 \log\log \sigma_n^2}} = 1,$$

$$\liminf_{n\to\infty} \frac{\sum_{k\leq n} \phi_k \circ T^k}{\sqrt{\sigma_n^2 \log\log \sigma_n^2}} = -1.$$

4. *If $d = 1$, scalar ASIP implies scalar self-norming FCLT:*

   Let $\rho_n^2 := \tilde{\mathbb{E}}[(\sum_{k\leq n} G_k)^2]$, where $(G_k)_{k\geq 1}$ are the Gaussian variables in the definition of the VASIP and $S_n := \sum_{k\leq n} \phi_k \circ T^k$. For any $n \geq 1$, define the piecewise continuous function S^n w.r.t. time variable $t \in [0,1]$:

$$S_t^n := \frac{S_i}{\rho_n} + \frac{t - \frac{\rho_i^2}{\rho_n^2}}{\frac{\rho_{i+1}^2}{\rho_n^2} - \frac{\rho_i^2}{\rho_n^2}} \cdot \frac{S_{i+1} - S_i}{\rho_n}, t \in [\frac{\rho_i^2}{\rho_n^2}, \frac{\rho_{i+1}^2}{\rho_n^2}],$$

$$\textit{where } 0 \leq i \leq n-1, S_0 := 0, \frac{0}{0} := 0.$$

Then $S^n \xrightarrow{d} B$ on $C[0,1]$, where B is a standard Brownian motion.

Definition 2.6 (Assumptions on decays of correlations)

There is $\alpha < \frac{1}{2}$ such that for any $i, j, n \in \mathbb{N}$ (the constants indicated in $\precsim$ below are uniform *over all $\phi_i, i, j, n \in \mathbb{N}$), the following holds:*

$$\int |P_{i+1}^{n+i}(\phi_i \cdot P^i \boldsymbol{1})| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}, \tag{A1}$$

$$\int |P_{i+1}^{n+i}[(\phi_i \cdot \phi_i^T - \int \phi_i \circ T^i \cdot \phi_i^T \circ T^i d\mu) \cdot P^i \boldsymbol{1}]| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}, \tag{A2}$$

$$\int |P_{i+j+1}^{i+j+n}\{[P_{i+1}^{i+j}(\phi_i \cdot P^i \boldsymbol{1}) \cdot \phi_{i+j}^T - P^{i+j} \boldsymbol{1} \cdot \int P_{i+1}^{i+j}(\phi_i \cdot P^i \boldsymbol{1}) \cdot \phi_{i+j}^T d\mu]\}| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}. \tag{A3}$$

Remark 2.7

1. *We assume $\alpha < \frac{1}{2}$ throughout this dissertation.*

2. *For the stationary dynamical system, that is, $\phi_k := \phi, T_k = T$ for all $k \geq 1$, $(T)_*\mu = \mu$ and $\int \phi d\mu = 0$. We denote the transfer operator of T by P. Then $P^i \boldsymbol{1} = \boldsymbol{1}$ a.s. for any $i \geq 1$ and the assumptions (A1)-(A3) become:*

$$\int |P^n(\phi)| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}, \tag{A4}$$

$$\int |P^n(\phi \cdot \phi^T - \int \phi \cdot \phi^T d\mu)| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}, \tag{A5}$$

$$\int |P^n[P^j(\phi) \cdot \phi^T - \int P^j(\phi) \cdot \phi^T d\mu]| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}. \tag{A6}$$

(A4), (A5) are well-known to be decay of correlations if ϕ has certain regularity. In this dissertation, they are called first order decay of correlation for the stationary dynamical system, (A6) is called second order decay of correlation for the stationary dynamical system.

2.2 Main theorems

Theorem 2.8 (VASIP)

Assume the non-stationary dynamical system $(X, \mathcal{B}, (T_k)_{k\geq 1}, \mu)$ with the observables $(\phi_k)_{k\in\mathbb{N}}$ satisfies (A1)-(A3). Then there is $\gamma \in (0,1)$ depending on d, α only (will be given in Appendix, Lemma 7.1), such that if $\lambda(\sigma_n^2) \succsim n^\gamma$, then $(\phi_k \circ T^k)_{k\geq 1}$ satisfies the VASIP.

Theorem 2.9 (Quenched VASIP)

Consider the random dynamical system $(X, \mathcal{B}, (T_\omega)_{\omega\in\Omega}, \mu, \Omega, \mathcal{F}, \mathbb{P}, \sigma, (\mu_\omega)_{\omega\in\Omega})$. Denote P_ω as the transfer operator of T_ω w.r.t. μ. Define $\tau : \Omega \times X \to \Omega \times X$ by $\tau(\omega, x) := (\sigma\omega, T_\omega(x))$. Define the random composition of transformations and transfer operators: $T^k_\omega := T_{\sigma^{k-1}(\omega)} \circ T_{\sigma^{k-2}(\omega)} \circ \cdots \circ T_\omega$, $P^k_\omega := P_{\sigma^{k-1}(\omega)} \circ P_{\sigma^{k-2}(\omega)} \circ \cdots \circ P_\omega$. Assume for a.e. $\omega \in \Omega$, $d\mu_\omega := h_\omega d\mu$ for some density function $h_\omega \in L^1(\mu)$ (this implies $P_\omega h_\omega = h_{\sigma\omega}$). Assume the observables $\{\phi_\omega \in L^\infty(X, \mu; \mathbb{R}^d) : \omega \in \Omega\}$ satisfy

$$\sup_{\omega\in\Omega} ||\phi_\omega||_{L^\infty(X,\mu)} < \infty, \int \phi_\omega d\mu_\omega = 0$$

and (A1')-(A3') below, where $\alpha < \frac{1}{2}$ and the constants indicated in $\precsim$ are uniform over all $i, j, n \in \mathbb{N}$ and $(\phi_\omega)_{\omega\in\Omega}$:

$$\int |P^n_{\sigma^i\omega}(\phi_{\sigma^i\omega} \cdot h_{\sigma^i\omega})| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}, \tag{A1'}$$

$$\int |P^n_{\sigma^i\omega}[(\phi_{\sigma^i\omega} \cdot \phi^T_{\sigma^i\omega} - \int \phi_{\sigma^i\omega} \cdot \phi^T_{\sigma^i\omega} d\mu_{\sigma^i\omega}) \cdot h_{\sigma^i\omega}]| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}, \tag{A2'}$$

$$\int |P^n_{\sigma^{i+j}\omega}\{[P^j_{\sigma^i\omega}(\phi_{\sigma^i\omega} \cdot h_{\sigma^i\omega}) \cdot \phi^T_{\sigma^{i+j}\omega} - h_{\sigma^{i+j}\omega} \int P^j_{\sigma^i\omega}(\phi_{\sigma^i\omega} \cdot h_{\sigma^i\omega}) \cdot \phi^T_{\sigma^{i+j}\omega} d\mu]\}| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}. \tag{A3'}$$

Then there are two linear subspaces (independent of ω): $W_1, W_2 \subset \mathbb{R}^d$, $\mathbb{R}^d = W_1 \bigoplus W_2$ with projections $\pi_1 : W_1 \bigoplus W_2 \to W_1, \pi_2 : W_1 \bigoplus W_2 \to W_2$ such that

- *Quenched VASIP: the dynamical system $(\pi_1 \circ \phi_{\sigma^k\omega} \circ T^k_\omega)_{k\geq 1}$ satisfies the VASIP w.r.t. μ_ω for a.e. $\omega \in \Omega$.*
- *Coboundary: the dynamical system $(\pi_2 \circ \phi_{\sigma^k\omega} \circ T^k_\omega)_{k\geq 1}$ is a coboundary: there is $\psi \in L^1(\Omega \times X, d\mu_\omega d\mathbb{P})$ such that:*

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

Remark 2.10

1. *Conditions (A3), (A3'), (A6) can be easily verified by the invariant cone and tower extension methods, as shown in our Corollaries 3.1 and 3.8.*

2. The quasi-invariant density h_ω is not required to be bounded away from zero.

3 Applications

3.1 Applications to non-stationary dynamical systems

Corollary 3.1 (Polynomially mixing non-stationary system)

Consider the non-stationary dynamical system $([0,1],\mathcal{B},(T_{\beta_k})_{k\geq 1},m)$ *in [39], where* m *is the Lebesgue measure,* $T_k := T_{\beta_k}$ *are Pomeau-Manneville like maps with* $0<\beta_k<\alpha<\frac{1}{2}$*:*

$$T_{\beta_k}(x) = \begin{cases} x + 2^{\beta_k}x^{1+\beta_k}, & 0 \leq x \leq \frac{1}{2} \\ 2x-1, & \frac{1}{2} < x \leq 1 \end{cases}. \tag{3.1}$$

Consider the observables $(\phi_k)_{k\in\mathbb{N}} \subset \mathrm{Lip}([0,1];\mathbb{R}^d)$ *with* $\sup_k ||\phi_k||_{\mathrm{Lip}} < \infty$*. Then there is* $\gamma \in (0,1)$ *(same as* γ *in Theorem 2.8), such that if* $\lambda(\sigma_n^2) \gtrsim n^\gamma$*, then* $(\phi_i \circ T^i)_{i\geq 1}$ *satisfies the VASIP.*

Consider the observables $(\phi_k)_{k\in\mathbb{N}} \subset \mathrm{Lip}([0,1];\mathbb{R})$ *with* $\sup_k ||\phi_k||_{\mathrm{Lip}} < \infty$*. Then there is* $\gamma_1 < 1$ *depending on* α *only (simpler than* γ*, will be given in Appendix, Lemma 7.2), such that if* $\lambda(\sigma_n^2) \gtrsim n^{\gamma_1}$*, then* $(\phi_i \circ T^i)_{i\geq 1}$ *satisfies self-norming CLT.*

Remark 3.2 *In [39], the martingale method was used to prove CLT holds when* $\alpha < \frac{1}{8}$*. Our method proves CLT holds for the slower mixing system (*$\alpha < \frac{1}{2}$*), which coincides with the sharp result in the stationary case, see [51].*

Corollary 3.3 (Exponentially mixing non-stationary system)

Consider the non-stationary dynamical system $(X,\mathcal{B},(T_k)_{k\geq 1},\mu)$*, assume* $(\mathcal{V},||\cdot||_{\mathcal{V}})$ *is a* $(P_k)_{k\geq 1}$*-invariant Banach sub-algebra of* $(L^1,||\cdot||_{L^1})$ *satisfying the following assumptions:*

1. $\mathbf{1} \in \mathcal{V}$.
2. $||\cdot||_{L^1} \lesssim ||\cdot||_{\mathcal{V}}$.
3. *There is a constant* A*, such that for any* $n,m \in \mathbb{N}$*, any* $v \in \mathcal{V}$*,*

$$||P_{m+1}^{n+m} v||_{\mathcal{V}} \leq A \cdot ||v||_{\mathcal{V}}.$$

4. *There is $\rho \in (0,1)$ and a constant B such that for any $n, m \in \mathbb{N}$, any $v \in \mathcal{V}_0 := \{v \in \mathcal{V} : \int v d\mu = 0\}$, we have*

$$||P_{m+1}^{n+m} v||_{\mathcal{V}} \leq B \cdot \rho^n \cdot ||v||_{\mathcal{V}}.$$

Consider the observables $(\phi_k)_{k\in\mathbb{N}} \subset \mathcal{V} \cap L^\infty(X, \mu; \mathbb{R}^d)$ with $\sup_k\{||\phi_k||_{\mathcal{V}}, ||\phi_k||_{L^\infty}\} < \infty$. Then there is $\gamma < 1$ (same as γ in Theorem 3.1), 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

1. *As applications, we can use this result for the dynamical systems considered in [3–6, 10, 17, 20, 21, 24, 29, 30, 46]:*

   - *Shrinking target problem for an expanding map in Theorem 5.1 of [20]: β-transformation, smooth expanding map, the Gauss map, and mixing Rychlik-type map. $\mathcal{V} := BV, T_k := T$ with $T_*\mu = \mu$, $\phi_k := 1_{A_k} - \mu(A_k)$ with $\sup_k ||\phi_k||_{BV} < \infty$ and $\mu(A_k) \gtrsim n^{\gamma-1}$ where γ is the one in our Theorem 2.8.*

   - *Non-stationary observations on an Axiom A dynamical system in Corollary 6.2 of [20]: $\mathcal{V} := C^{0,\alpha}$ (α-Hölder space), $T_k := T$ with $T_*\mu = \mu$, $\sup_k ||\phi_k||_{C^{0,\alpha}} < \infty$ with $\sigma_n^2 \gtrsim n^{\max(\gamma, \frac{\sqrt{17}-1}{4})}$ where γ is the one in our Theorem 2.8.*

   - *The systems in section 7 of [20] are essentially the same, so we just choose 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 SRB for T, $\phi_k := \phi - \int \phi \circ T_1^k d\mu$ where $\phi \in \mathcal{V}$ is not a coboundary for T and $\int \phi d\mu = 0$. By our Theorem 2.8 and Lemma 7.1 of [20], $(\phi_k \circ T_1^k := \phi \circ T_1^k - \int \phi \circ T_1^k d\mu)_{k\geq 1}$ has the VASIP. Since for any $n \geq 1$,*

$$\sum_{k\leq n} \int \phi \circ T_1^k d\mu = \sum_{k\leq n} \int \phi \cdot P_1^k(h) dm = \sum_{k\leq n} \int \phi \cdot [P_1^k(h) - P^k(h)] dm$$

$$\lesssim \sum_{k\leq n} ||\phi||_{\mathcal{V}} \cdot ||P_1^k(h) - P^k(h)||_{L^1}$$

where P_1^k and P are the transfer operators of T_1^k and T respectively, then by Lemma

2.13 in [10],

$$\sup_n |\sum_{k\geq 1} \int \phi \circ T_1^k d\mu| = O(1).$$

So we have the same statement of the VASIP for $(\phi \circ T_1^k)_{k\geq 1}$ as Theorem 7.4 in [20].

2. *Conditions (Min) in [10] and (LB) in [20] are not required here. This Corollary works for all non-stationary dynamical systems whose transfer operators have uniform spectral gap in the sense of [10]. Gouëzel in [19] obtained the following similar result: if the transfer operator of the stationary dynamical system has a spectral gap, then the VASIP holds, without assuming the (Min) or (LB) conditions. Using Gouëzel's result, Luzzatto and Melbourne in [33] obtained the VASIP for interval maps with singularities.*

3. *$\mathcal{V}$ is usually chosen to be anisotropic Banach spaces or simpler Banach spaces such as Hölder functions, Lipschitz functions or bounded variation functions. This implies the decay of correlation w.r.t. the norm $||\cdot||_{\mathcal{V}}$ is exponential, see for instance Proposition 3.1 in [2].*

3.2 Applications to random dynamical systems

Corollary 3.5 (Exponentially mixing random system)

Consider the same system as in [14]: $(X, \mathcal{B}, (T_\omega)_{\omega\in\Omega}, m)$ with a notion of a variation $\mathrm{var}: L^1(X,m) \to [0,\infty]$ *and a probability space $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ where $\sigma : \Omega \to \Omega$ is an invertible $\mathbb{P}$-preserving ergodic transformation. Define random composition of transformations T_ω^k and random composition of transfer operators P_ω^k as in Theorem 2.9. Assume:*

1. $\mathrm{var}(th) = |t|\,\mathrm{var}(h), t \in \mathbb{R}$.

2. $\mathrm{var}(g+h) \leq \mathrm{var}(g) + \mathrm{var}(h)$.

3. $||h||_{L^\infty} \leq C_{\mathrm{var}} \cdot (||h||_{L^1} + \mathrm{var}(h))$ *for some constant* $1 \leq C_{\mathrm{var}} < \infty$.

4. *For any $C > 0$, the set $\{h : X \to \mathbb{R} : ||h||_{L^1} + \mathrm{var}(h) \leq C\}$ is $L^1(X,m)$-compact.*

5. $\mathrm{var}(\boldsymbol{1}) < \infty$.

6. $\{h : X \to \mathbb{R} : ||h||_{L^1} = 1$ *and* $\mathrm{var}(h) < \infty\}$ *is $L^1(X,m)$-dense in* $\{h : X \to \mathbb{R} : ||h||_{L^1} = 1\}$.

7. *There is* $K_{\text{var}} < \infty$ *such that for every* $g, h \in B$,

$$\text{var}(gh) + ||gh||_{L^1} \le K_{\text{var}} \cdot (\text{var}(h) + ||h||_{L^1}) \cdot (\text{var}(g) + ||g||_{L^1}),$$

where

$$B := BV(X, m) := \{h \in L^1(X, m) : \text{var}(h) < \infty\}$$

with norm

$$||h||_B := \text{var}(h) + ||h||_{L^\infty}.$$

8. *For any* $g \in L^1(X, m)$ *such that* $\text{ess inf}\, g > 0$, *we have* $\text{var}(\frac{1}{g}) \le \frac{\text{var}(g)}{(\text{ess inf}\, g)^2}$.

9. *The map* $(\omega, x) \to (P_\omega H(\omega, \cdot))(x)$ *is* $\mathbb{P} \times m$*-measurable, that is, measurable on the space* $(\Omega \times X, \mathcal{F} \otimes \mathcal{B})$ *for every* $\mathbb{P} \times m$*-measurable function* $H : \Omega \times X \to \mathbb{R}^d$ *such that* $H(\omega, \cdot) \in L^1(X, m)$ *for a.e.* $\omega \in \Omega$.

10. *There is* $C > 0$ *such that for a.e.* $\omega \in \Omega$ *and any* $\phi \in B$,

$$||P_\omega \phi||_B \le C \cdot ||\phi||_B.$$

11. *There are constants* $K, \lambda > 0$ *such that for a.e.* $\omega \in \Omega$, $n \ge 0$ *and* $\phi \in B$ *with* $\int \phi dm = 0$, *we have*

$$||P_\omega^n \phi||_B \le K \cdot e^{-\lambda n} ||\phi||_B.$$

Then there is an unique quasi-invariant probability $d\mu_\omega := h_\omega dm$ *such that for a.e.* $\omega \in \Omega$, $P_\omega h_\omega = h_{\sigma\omega}, \sup_\omega ||h_\omega||_B < \infty$. *Moreover, for the observables* $(\phi_\omega)_{\omega \in \Omega} \subset L^\infty(X, m; \mathbb{R}^d) \cap B$ *with* $\int \phi_\omega d\mu_\omega = 0$ *and* $\sup_{\omega \in \Omega} ||\phi_\omega||_B < \infty$. *Then there are two linear sub-spaces (independent of* ω*):* $W_1, W_2 \subset \mathbb{R}^d$, $\mathbb{R}^d = W_1 \bigoplus W_2$ *with projections* $\pi_1 : W_1 \bigoplus W_2 \to W_1, \pi_2 : W_1 \bigoplus W_2 \to W_2$ *such that*

- *Quenched VASIP: the dynamical system* $(\pi_1 \circ \phi_{\sigma^k \omega} \circ T_\omega^k)_{k \ge 1}$ *satisfies the VASIP w.r.t.* μ_ω *for a.e.* $\omega \in \Omega$.

- *Coboundary: the dynamical system* $(\pi_2 \circ \phi_{\sigma^k \omega} \circ T_\omega^k)_{k \geq 1}$ *is a coboundary: there is* $\psi \in L^2(\Omega \times X, d\mu_\omega d\mathbb{P})$ *such that:*

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

Remark 3.6

1. *[14] and [15] considered the same random dynamical systems, but present different results for limit theorems.*

2. *Note that our assumptions for the random dynamical system are (V1)-(V8) and (H1), (H2), (H5) in [14]. (H3), (H4) in [14] and (C4) in [15] are not required here. Our Corollary 3.5 works for the random dynamical systems in [14, 15]:*

    - *Random piecewise expanding maps in higher dimensions.*

    - *Random Lasota-Yorke maps.*

Corollary 3.7 (Polynomially mixing random system)

Consider the system $(X, \mathcal{B}, (T_\omega)_{\omega \in [0,\alpha]^{\mathbb{Z}}}, m)$ *where* $T_\omega := T_{\omega_0}$ *is the Pomeau-Manneville like map which is picked from* $\{T_\beta : \beta \in [0, \alpha], \alpha < \frac{1}{2}\}$*. Define* $(\Omega, \mathcal{F}, \mathbb{P}, \sigma) := ([0, \alpha]^{\mathbb{Z}}, \mathcal{F}, \mathbb{P}, \sigma)$ *where* $\sigma : \Omega \to \Omega$ *is an invertible ergodic left shift preserving the probability* $\mathbb{P}$*. Define random composition of transformations* T_ω^k *and random composition of transfer operators* P_ω^k *as in Theorem 2.9.*

Then there is a quasi-invariant probability $d\mu_\omega := h_\omega dm$ *such that* $P_\omega h_\omega = h_{\sigma\omega}$ *for a.e.* $\omega \in \Omega$*; moreover, consider the observable* $(\phi_\omega)_{\omega \in \Omega} \subset \mathrm{Lip}([0, 1]; \mathbb{R}^d)$ *with* $\sup_\omega ||\phi_\omega||_{\mathrm{Lip}} < \infty$ *and* $\int \phi_\omega d\mu_\omega = 0$*. Then there are two linear subspaces (independent of* ω*):* $W_1, W_2 \subset \mathbb{R}^d$*,* $\mathbb{R}^d = W_1 \bigoplus W_2$ *with projections* $\pi_1 : W_1 \bigoplus W_2 \to W_1, \pi_2 : W_1 \bigoplus W_2 \to W_2$ *such that*

- *Quenched VASIP: the dynamical system* $(\pi_1 \circ \phi_{\sigma^k \omega} \circ T_\omega^k)_{k \geq 1}$ *satisfies the VASIP w.r.t.* μ_ω *for a.e.* $\omega \in \Omega$*.*

- *Coboundary: the dynamical system* $(\pi_2 \circ \phi_{\sigma^k \omega} \circ T_\omega^k)_{k \geq 1}$ *is a coboundary: there is* $\psi \in L^1(\Omega \times$

$X, d\mu_\omega d\mathbb{P}$) such that:

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

3.3 Applications to stationary dynamical systems

Corollary 3.8 (Stationary dynamical system)

Consider the stationary dynamical system $(X, \mathcal{B}, T, \mu)$ (that is, $T_*\mu = \mu$) with mean zero observable $\phi : X \to \mathbb{R}^d$ satisfying (A4), then there is a $d \times d$ positive semi-definite matrix σ^2 and $\epsilon \in (0,1)$, such that $\sigma_n^2 = n \cdot \sigma^2 + o(n^{1-\epsilon})$. If (A4)-(A6) are all satisfied, then there are two linear subspaces: $W_1, W_2 \subset \mathbb{R}^d$ such that $\mathbb{R}^d = W_1 \bigoplus W_2$ with projection $\pi_1 : W_1 \bigoplus W_2 \to W_1, \pi_2 : W_1 \bigoplus W_2 \to W_2$, such that:

- *The dynamical system $(\pi_1 \circ \phi \circ T^k)_{k \geq 1}$ satisfies the VASIP.*
- *The dynamical system $(\pi_2 \circ \phi \circ T^k)_{k \geq 1}$ 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 the dynamical system can be described by a Young tower Δ [50, 51], that is, $(\Delta, \mathcal{B}, F, v)$ with $v \circ F^{-1} = v$, $dv = \frac{dv}{dm} dm$ is exact, dm is the reference measure on Δ, return map R is defined on the base of the the tower: $\Delta_0 = \bigsqcup_{i \geq 1} \Delta_{0,i}$ such that $R|_{\Delta_{0,i}} \equiv R_i \in \mathbb{N}$, $\int_{\Delta_0} R dm < \infty$ 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

$$|\frac{JF^R(x)}{JF^R(y)} - 1| \precsim \beta^{s(F^R(x), F^R(x))} \tag{3.2}$$

where J is the Jacobian w.r.t. dm, $\beta \in (0,1)$, $s(x,y)$ 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 Δ: 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 dynamical system $(\Delta, \mathcal{B}, F, v)$ and any mean zero observable $\phi \in \mathrm{Lip}(\Delta)$, (A4)-(A6) are all satisfied. On the other hand, the discrete stationary dynamical systems such as Pomeau-Manneville maps, Viana maps, etc. considered in [34], [35] can be described by a Young tower, so we recover the VASIP for those systems.

4 Dynamical inequalities

In this section, we will obtain several inequalities based on (A1)-(A3), which are needed in section 5 to prove Theorem 2.8.

Lemma 4.1 *If (A1) is satisfied, then for all $n, m \in \mathbb{N}$, the following holds:*

$$\int \Big(\sum_{m \le k \le n+m-1} \phi_k \circ T^k \Big) \cdot \Big(\sum_{m \le k \le m+n-1} \phi_k \circ T^k \Big)^T d\mu = O(n),$$

where the constant indicated in $O(\cdot)$ is uniform over all ϕ_k, m, n.

Proof

$$\int \Big(\sum_{m \le k \le n+m-1} \phi_k \circ T^k \Big) \cdot \Big(\sum_{m \le k \le m+n-1} \phi_k \circ T^k \Big)^T d\mu = \int \sum_{m \le k \le n+m-1} \phi_k \circ T^k \cdot \phi_k^T \circ T^k$$

$$+ \sum_{m \le i < j \le n+m-1} \phi_i \circ T^i \cdot \phi_j^T \circ T^j + \Big(\sum_{m \le i < j \le n+m-1} \phi_i \circ T^i \cdot \phi_j^T \circ T^j \Big)^T d\mu.$$

By $\sup_k ||\phi_k||_{L^\infty} < \infty$, the above equality becomes

$$= O(n) + \sum_{m \le i < j \le n+m-1} \int \phi_i \cdot \phi_j^T \circ T_{i+1}^j \cdot P^i \mathbf{1} + (\phi_i \cdot \phi_j^T \circ T_{i+1}^j \cdot P^i \mathbf{1})^T d\mu$$

$$\precsim O(n) + \sum_{m \le i < j \le n+m-1} \int |P_{i+1}^j (\phi_i \cdot P^i \mathbf{1})| d\mu.$$

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

$$\precsim O(n) + \sum_{m \le i < j \le n+m} \frac{1}{(j-i)^{\frac{1}{\alpha}-1}} = O(n)$$

$$+ \sum_{0 \le i < j \le n} \frac{1}{(j-i)^{\frac{1}{\alpha}-1}} \precsim O(n) + \sum_{0 < j \le n} \sum_{0 \le i < j} \frac{1}{(j-i)^{\frac{1}{\alpha}-1}} = O(n).$$

All constants indicated in $\precsim$, $O(\cdot)$ are uniform over all ϕ_k, m, n. ∎

Lemma 4.2 *If (A1) is satisfied, then for all $n, m \in \mathbb{N}$, the following holds:*

$$\mathbb{E}|\mathbb{E}_{n+m} \sum_{m \le k \le n+m-1} \phi_k \circ T^k| = O(1),$$

where $O(1)$ is uniform over all ϕ_k, n, m.

Proof

$$\mathbb{E}|\mathbb{E}_{n+m} \sum_{m \le k \le n+m-1} \phi_k \circ T^k| = \sup_{||\psi||_{L^\infty(X;\mathbb{R})} \le 1} \int \psi \circ T^{n+m} \cdot \sum_{m \le k \le n+m-1} \phi_k \circ T^k d\mu.$$

By $||\psi||_{L^\infty} \le 1$ and (A1), the above equality becomes

$$\le \sum_{m \le k \le n+m-1} \int |P_{k+1}^{n+m}(\phi_k \cdot P^k \mathbf{1})| d\mu \precsim \sum_{m \le k \le n+m-1} \frac{1}{(m+n-k)^{\frac{1}{\alpha}-1}} = O(1).$$

All constants indicated in $\precsim$, $O(\cdot)$ are uniform over all ϕ_k, m, n. The last equality holds because of $\frac{1}{\alpha} - 1 > 1$.

∎

Lemma 4.3 *If (A1)-(A3) are satisfied, then for all $n, m \in \mathbb{N}$, the following holds:*

$$\mathbb{E}|\mathbb{E}_{n+m}[(\sum_{k=m}^{n+m-1} \phi_k \circ T^k) \cdot (\sum_{k=m}^{n+m-1} \phi_k \circ T^k)^T] - \mathbb{E}[(\sum_{k=m}^{n+m-1} \phi_k \circ T^k) \cdot (\sum_{k=m}^{n+m-1} \phi_k \circ T^k)^T]| = O(n^{\frac{\alpha}{1-\alpha}}),$$

where the constant indicated in $O(\cdot)$ is uniform over all ϕ_k, m, n.

Proof

$$\mathbb{E}|\mathbb{E}_{n+m}[(\sum_{k=m}^{n+m-1} \phi_k \circ T^k) \cdot (\sum_{k=m}^{n+m-1} \phi_k \circ T^k)^T] - \mathbb{E}[(\sum_{k=m}^{n+m-1} \phi_k \circ T^k) \cdot (\sum_{k=m}^{n+m-1} \phi_k \circ T^k)^T]|$$

$$\le \mathbb{E}|\mathbb{E}_{n+m}(\sum_{m \le k \le n+m-1} \phi_k \circ T^k \cdot \phi_k^T \circ T^k) - \mathbb{E}(\sum_{m \le k \le n+m-1} \phi_k \circ T^k \cdot \phi_k^T \circ T^k)|$$

$$+\mathbb{E}|\mathbb{E}_{n+m}(\sum_{m\le i<j\le n+m-1}\phi_i\circ T^i\cdot\phi_j^T\circ T^j)-\mathbb{E}(\sum_{m\le i<j\le n+m-1}\phi_i\circ T^i\cdot\phi_j^T\circ T^j)|$$

$$+\mathbb{E}|[\mathbb{E}_{n+m}(\sum_{m\le i<j\le n+m-1}\phi_i\circ T^i\cdot\phi_j^T\circ T^j)-\mathbb{E}(\sum_{m\le i<j\le n+m-1}\phi_i\circ T^i\cdot\phi_j^T\circ T^j)]^T|$$

$$\le\mathbb{E}|\mathbb{E}_{n+m}[\sum_{m\le k\le n+m-1}\phi_k\circ T^k\cdot\phi_k^T\circ T^k-\mathbb{E}(\phi_k\circ T^k\cdot\phi_k^T\circ T^k)]| \tag{4.1}$$

$$+2\mathbb{E}|\mathbb{E}_{n+m}[\sum_{m\le i<j\le n+m-1}\phi_i\circ T^i\cdot\phi_j^T\circ T^j-\mathbb{E}(\phi_i\circ T^i\cdot\phi_j^T\circ T^j)]| \tag{4.2}$$

Estimate (4.1):

$$(4.1)=\sup_{||\psi||_{L^\infty(X;\mathbb{R})}\le 1}\int\psi\circ T^{n+m}\cdot[\sum_{m\le k\le n+m-1}\phi_k\circ T^k\cdot\phi_k^T\circ T^k-\mathbb{E}(\phi_k\circ T^k\cdot\phi_k^T\circ T^k)]d\mu$$

$$\precsim\sum_{k=m}^{n+m-1}\int|P_{k+1}^{n+m}\{[\phi_k\cdot\phi_k^T-\mathbb{E}(\phi_k\circ T^k\cdot\phi_k^T\circ T^k)]\cdot P^k\mathbf{1}\}|d\mu.$$

By (A2) and $\alpha\in(0,\frac{1}{2})$, the above inequality becomes

$$\precsim\sum_{k=m}^{n+m-1}\frac{1}{(m+n-k)^{\frac{1}{\alpha}-1}}=O(1).$$

To estimate (4.2), for any fixed $j\le n+m-1$:

$$\mathbb{E}|\mathbb{E}_j[\sum_{m\le i<j}\phi_i\circ T^i\cdot\phi_j^T\circ T^j-\mathbb{E}(\phi_i\circ T^i\cdot\phi_j^T\circ T^j)]|\le\sum_{m\le i<j}\mathbb{E}|\mathbb{E}_j(\phi_i\circ T^i\cdot\phi_j^T\circ T^j)-\mathbb{E}(\phi_i\circ T^i\cdot\phi_j^T\circ T^j)|$$

$$\le 2\sum_{m\le i<j}\mathbb{E}|\mathbb{E}_j(\phi_i\circ T^i\cdot\phi_j^T\circ T^j)|\le 2\sum_{m\le i<j}\sup_{||\psi||_{L^\infty(X;\mathbb{R})}\le 1}\int\psi\circ T^j\cdot\phi_i\circ T^i\cdot\phi_j^T\circ T^jd\mu.$$

By (A1), $||\psi||_{L^\infty}\le 1,\sup_j||\phi_j||_{L^\infty}<\infty$ and $\alpha\in(0,\frac{1}{2})$, the above inequality becomes:

$$\precsim 2\sum_{m\le i<j}\int|P_{i+1}^j(\phi_i\cdot P^i\mathbf{1})|d\mu\precsim\sum_{m\le i<j}\frac{1}{(j-i)^{\frac{1}{\alpha}-1}}=O(1).$$

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

$$\sum_{m \le i < j} \mathbb{E}|\mathbb{E}_j[\phi_i \circ T^i \cdot \phi_j^T \circ T^j - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]| = O(1). \tag{4.3}$$

Let $\delta := \frac{\alpha}{1-\alpha} < 1$, then

$$\begin{aligned}
(4.2) \precsim & \sum_{m \le i < j \le n+m-1} \mathbb{E}|\mathbb{E}_{n+m}[\phi_i \circ T^i \cdot \phi_j^T \circ T^j - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]| \\
= & \sum_{m < j \le n+m-1} \sum_{m \le i < j} \mathbb{E}|\mathbb{E}_{n+m}[\phi_i \circ T^i \cdot \phi_j^T \circ T^j - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]| \\
= & \sum_{m < j \le n+m-1} \sum_{m \le i < j} \mathbb{E}|\mathbb{E}_{n+m}\mathbb{E}_j[\phi_i \circ T^i \cdot \phi_j^T \circ T^j - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]| \\
\le & \sum_{n+m-\lfloor n^\delta \rfloor < j \le n+m-1} \sum_{m \le i < j} \mathbb{E}|\mathbb{E}_j[\phi_i \circ T^i \cdot \phi_j^T \circ T^j - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]| \\
+ & \sum_{m < j \le n+m-\lfloor n^\delta \rfloor} \sum_{m \le i < j} \mathbb{E}|\mathbb{E}_{n+m}[\phi_i \circ T^i \cdot \phi_j^T \circ T^j - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]|.
\end{aligned}$$

By (4.3), the above inequality becomes

$$\precsim \lfloor n^\delta \rfloor + \sum_{m < j \le n+m-\lfloor n^\delta \rfloor} \sum_{m \le i < j} \mathbb{E}|\mathbb{E}_{n+m}[\phi_i \circ T^i \cdot \phi_j^T \circ T^j - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]|$$

$$\precsim \lfloor n^\delta \rfloor + \sum_{j=m+1}^{n+m-\lfloor n^\delta \rfloor} \sum_{m \le i < j} \sup_{||\psi||_{L^\infty(X;\mathbb{R})} \le 1} \int \psi \circ T^{n+m} \cdot [(\phi_i \circ T^i \cdot \phi_j^T \circ T^j) - \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)] d\mu$$

$$= \lfloor n^\delta \rfloor + \sum_{j=m+1}^{n+m-\lfloor n^\delta \rfloor} \sum_{m \le i < j} \sup_{||\psi||_{L^\infty(X;\mathbb{R})} \le 1} \int \psi \circ T_{j+1}^{n+m} \cdot [P_{i+1}^j(\phi_i \cdot P^i \mathbf{1}) \cdot \phi_j^T - P^j \mathbf{1} \cdot \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)] d\mu$$

$$\le \lfloor n^\delta \rfloor + \sum_{j=m+1}^{n+m-\lfloor n^\delta \rfloor} \sum_{m \le i < j} \int |P_{j+1}^{n+m}[P_{i+1}^j(\phi_i \cdot P^i \mathbf{1}) \cdot \phi_j^T - P^j \mathbf{1} \cdot \mathbb{E}(\phi_i \circ T^i \cdot \phi_j^T \circ T^j)]| d\mu.$$

By (A3), the above inequality becomes

$$\precsim \lfloor n^\delta \rfloor + \sum_{j=m+1}^{n+m-\lfloor n^\delta \rfloor} \sum_{m\le i<j} \frac{1}{(n+m-j)^{\frac{1}{\alpha}-1}} \le \lfloor n^\delta \rfloor + \sum_{m<j\le n+m-\lfloor n^\delta \rfloor} \frac{j-m}{(n+m-j)^{\frac{1}{\alpha}-1}}$$

$$= \lfloor n^\delta \rfloor + \sum_{0<j\le n-\lfloor n^\delta \rfloor} \frac{j}{(n-j)^{\frac{1}{\alpha}-1}} = \lfloor n^\delta \rfloor + \sum_{0<j\le n-\lfloor n^\delta \rfloor} \frac{\frac{j}{n}}{(1-\frac{j}{n})^{\frac{1}{\alpha}-1}} \cdot \frac{1}{n} \cdot \frac{1}{n^{\frac{1}{\alpha}-3}}$$

$$\precsim \lfloor n^\delta \rfloor + \int_0^{\frac{n-\lfloor n^\delta \rfloor}{n}} \frac{x}{(1-x)^{\frac{1}{\alpha}-1}} dx \cdot \frac{1}{n^{\frac{1}{\alpha}-3}}$$

$$= \lfloor n^\delta \rfloor + \int_{\frac{\lfloor n^\delta \rfloor}{n}}^1 \frac{1-x}{x^{\frac{1}{\alpha}-1}} dx \cdot \frac{1}{n^{\frac{1}{\alpha}-3}} \precsim \begin{cases} n^\delta + n^{1-\delta}, & \frac{1}{\alpha}-1=2 \\ n^{1+\delta(2-\frac{1}{\alpha})} + n^\delta, & \frac{1}{\alpha}-1 \neq 2 \end{cases} \precsim n^{\frac{\alpha}{1-\alpha}}. \tag{4.4}$$

All constants indicated in $\precsim$, $O(\cdot)$ are uniform over all ϕ_k, m, n.

∎

Lemma 4.4 *If (A1)-(A3) are satisfied, then for any $\epsilon \in (0, 1-\frac{\alpha}{1-\alpha})$, there is constant C_ϵ such that for all $n, m \in \mathbb{N}$, the following holds:*

$$\mathbb{E}\{|\mathbb{E}_{n+m}[(\sum_{k=m}^{n+m-1} \phi_k \circ T^k) \cdot (\sum_{k=m}^{n+m-1} \phi_k \circ T^k)^T] - \mathbb{E}[(\sum_{k=m}^{n+m-1} \phi_k \circ T^k) \cdot (\sum_{k=m}^{n+m-1} \phi_k \circ T^k)^T]|^{1+\epsilon}\}$$

$\precsim C_\epsilon \cdot n^{1+\epsilon}$, *where the constant indicated in $\precsim$ is uniform over all ϕ_k, m, n, ϵ.*

Proof

Let $\beta > \epsilon, \delta > 0$ (will be given later), and

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

Then

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

$$\le \delta^{\epsilon-\beta}\int |\Delta|^{1+\beta} d\mu + \delta^{\epsilon}\cdot\mathbb{E}|\Delta|.$$

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

$$\le \delta^{\epsilon-\beta}\int 2^{\beta}\cdot\mathbb{E}[|(\sum_{k=m}^{n+m-1}\phi_k\circ T^k)\cdot(\sum_{k=m}^{n+m-1}\phi_k\circ T^k)^T|^{1+\beta}]d\mu$$

$$+\delta^{\epsilon}\cdot\mathbb{E}|\Delta| \le 2^{\beta}\cdot\delta^{\epsilon-\beta}\cdot\int |(\sum_{m\le k\le n+m-1}\phi_k\circ T^k)\cdot(\sum_{m\le k\le n+m-1}\phi_k\circ T^k)^T|^{1+\beta} d\mu + \delta^{\epsilon}\cdot\mathbb{E}|\Delta|.$$

By Minkowski's inequality, Lemma 4.3 and $\sup_i ||\phi_i||_{L^\infty} < \infty$, the above inequality becomes

$$\le 2^{\beta}\cdot\delta^{\epsilon-\beta}\cdot(\sum_{m\le k\le n+m-1}||\phi_k\circ T^k||_{L^{2+2\beta}})^{2+2\beta} + \delta^{\epsilon}\cdot\mathbb{E}|\Delta| \le 2^{\beta}\cdot\delta^{\epsilon-\beta}n^{2+2\beta} + \delta^{\epsilon}n^{\frac{\alpha}{1-\alpha}}.$$

Take $\delta = n^{\frac{2+2\beta-\frac{\alpha}{1-\alpha}}{\beta}}$, then $\mathbb{E}(|\Delta|^{1+\epsilon}) \le 2^{\beta+1}\cdot n^{\epsilon\cdot\frac{2+2\beta-\frac{\alpha}{1-\alpha}}{\beta}+\frac{\alpha}{1-\alpha}}$. If $\epsilon\in(0,1-\frac{\alpha}{1-\alpha})$, we can choose big β such that $n^{\epsilon\cdot\frac{2+2\beta-\frac{\alpha}{1-\alpha}}{\beta}+\frac{\alpha}{1-\alpha}} \le n^{1+\epsilon}$. Then $\mathbb{E}(|\Delta|^{1+\epsilon}) \precsim n^{1+\epsilon}$, and the constant indicated in $\precsim$ is uniform over ϕ_k, m, n. ∎

Lemma 4.5 *If (A1) is satisfied, for any $m,n,p,q\in\mathbb{N}$, the following holds:*

$$\mathbb{E}[(\sum_{k=m}^{n+m-1}\phi_k\circ T^k)\cdot(\sum_{k=q+m+n}^{q+n+m+p-1}\phi_k\circ T^k)^T] = O(\max(n,p)^{\max(3-\frac{1}{\alpha},0)}),$$

where the constant indicated in $O(\cdot)$ is uniform over all ϕ_k, m, n, p.

Proof

Let $\delta<1$ (will be given later), $\bar{n} := \max(n,p)$, then

$$\mathbb{E}[(\sum_{k=m}^{n+m-1}\phi_k\circ T^k)\cdot(\sum_{k=q+m+n}^{q+n+m+p-1}\phi_k\circ T^k)^T] = \sum_{k=m}^{n+m-1}\sum_{j=q+m+n}^{q+n+m+p-1}\mathbb{E}(\phi_k\circ T^k\cdot\phi_j^T\circ T^j).$$

By (A1) and $\sup_j ||\phi_j||_{L^\infty} < \infty$, the above equality becomes

$$\precsim \sum_{k=m}^{n+m-1} \sum_{j=q+m+n}^{q+n+m+p-1} \int |P_{k+1}^{j}(\phi_k \cdot P^k \mathbf{1})| d\mu \precsim \sum_{k=m}^{n+m-1} \sum_{j=q+m+n}^{q+n+m+p-1} \frac{1}{(j-k)^{\frac{1}{\alpha}-1}}$$

$$= \sum_{k=m}^{n+m-1} \sum_{j=q+m+n}^{q+n+m+p-1} \frac{1}{(j-(m+n)+(m+n)-k)^{\frac{1}{\alpha}-1}} = \sum_{1\le k\le n} \sum_{0\le j\le p-1} \frac{1}{(j+k+q)^{\frac{1}{\alpha}-1}}$$

$$\precsim \sum_{1\le k\le n} \sum_{1\le j\le p} \frac{1}{(j+k)^{\frac{1}{\alpha}-1}} \le \sum_{1\le k\le \bar{n}} \sum_{1\le j\le \bar{n}} \frac{1}{(j+k)^{\frac{1}{\alpha}-1}} \precsim 2\lfloor \bar{n}^{\delta} \rfloor + \sum_{\lfloor \bar{n}^{\delta}\rfloor \le k, j\le \bar{n}} \frac{1}{(j+k)^{\frac{1}{\alpha}-1}}$$

$$\precsim 2\lfloor \bar{n}^{\delta} \rfloor + \sum_{\lfloor \bar{n}^{\delta}\rfloor \le k, j\le \bar{n}} \frac{1}{(\frac{j}{\bar{n}}+\frac{k}{\bar{n}})^{\frac{1}{\alpha}-1}} \frac{1}{\bar{n}} \cdot \frac{1}{\bar{n}} \cdot \frac{1}{\bar{n}^{\frac{1}{\alpha}-3}} \precsim 2\lfloor \bar{n}^{\delta} \rfloor + \int_{\frac{\lfloor \bar{n}^{\delta}\rfloor}{\bar{n}}}^{1} \int_{\frac{\lfloor \bar{n}^{\delta}\rfloor}{\bar{n}}}^{1} \frac{1}{(x+y)^{\frac{1}{\alpha}-1}} dxdy \cdot \frac{1}{\bar{n}^{\frac{1}{\alpha}-3}}$$

$$\precsim 2\lfloor \bar{n}^{\delta} \rfloor + \frac{1}{\bar{n}^{\frac{1}{\alpha}-3}} \cdot \left(\int_{\frac{\lfloor \bar{n}^{\delta}\rfloor}{\bar{n}}}^{1} (1+y)^{2-\frac{1}{\alpha}} - (y + \frac{\lfloor \bar{n}^{\delta}\rfloor}{\bar{n}})^{2-\frac{1}{\alpha}} dy\right)$$

$$\precsim \lfloor \bar{n}^{\delta} \rfloor + \frac{1}{\bar{n}^{\frac{1}{\alpha}-3}} \cdot \left(2^{3-\frac{1}{\alpha}} - (1+\frac{\lfloor \bar{n}^{\delta}\rfloor}{\bar{n}})^{3-\frac{1}{\alpha}} - (1+\frac{\lfloor \bar{n}^{\delta}\rfloor}{\bar{n}})^{3-\frac{1}{\alpha}} + (\frac{2\lfloor \bar{n}^{\delta}\rfloor}{\bar{n}})^{3-\frac{1}{\alpha}}\right)$$

$$\precsim \begin{cases} \bar{n}^{\delta} + \bar{n}^{3-\frac{1}{\alpha}}, & 3-\frac{1}{\alpha} > 0 \\ \bar{n}^{\delta}, & 3-\frac{1}{\alpha} \le 0 \end{cases} \precsim \begin{cases} \bar{n}^{3-\frac{1}{\alpha}}, & 3-\frac{1}{\alpha} > 0, \delta = 3-\frac{1}{\alpha} \\ O(1), & 3-\frac{1}{\alpha} \le 0, \delta = 0 \end{cases} \precsim \bar{n}^{\max(3-\frac{1}{\alpha}, 0)}. \tag{4.5}$$

All constants indicated in $\precsim$, $O(\cdot)$ are uniform over all ϕ_k, m, n, p, q.

∎

Lemma 4.6 (Maximal inequality)

If (A1)-(A3) are satisfied, then for any $\epsilon \in (0, \min(1, 2-\frac{2\alpha}{1-\alpha}))$, we have

$$\mathbb{E}(\max_{m\le k\le m+n-1} |\sum_{m\le i\le k} \phi_i \circ T^i|^{2+\epsilon}) \precsim C_\epsilon \cdot n^{1+\frac{\epsilon}{2}},$$

where the constant indicated in $\precsim$ is uniform over all ϕ_k, m, n.

Proof Similar to martingale maximal inequality, Serfling in [43,44] proved maximal inequality for some random processes (non-martingale) adapted to a increasing filtration. Although in different settings, we can still follow the idea of Theorem 3.1 in [43], then apply Theorem B in [44] to obtain the desired bound of our Lemma 4.6.

Note that if all $(\phi_k)_{k\in\mathbb{N}}$ satisfy (A1)-(A3), all coordinates of $(\phi_k)_{k\in\mathbb{N}}$ satisfy them too. Without loss of generality, we assume all $(\phi_k)_{k\in\mathbb{N}}$ are scalar functions satisfying (A1)-(A3).

First we claim:

$$\sup_{n,m\geq 1}\frac{\mathbb{E}(|\sum_{m\leq i\leq m+n-1}\phi_i\circ T^i|^{2+\epsilon})}{n^{\frac{2+\epsilon}{2}}}<\infty.$$

Let $A=\sum_{m\leq i<m+\lfloor\frac{n}{2}\rfloor}\phi_i\circ T^i$, $B=\sum_{m+\lfloor\frac{n}{2}\rfloor\leq i\leq m+n-1}\phi_i\circ T^i$, $\epsilon\in(0,1)$ (will be determined later),

$$\mathbb{E}(|\sum_{m\leq i\leq m+n-1}\phi_i\circ T^i|^{2+\epsilon})=\mathbb{E}(|\sum_{m\leq i<m+\lfloor\frac{n}{2}\rfloor}\phi_i\circ T^i+\sum_{m+\lfloor\frac{n}{2}\rfloor\leq i\leq m+n-1}\phi_i\circ T^i|^{2+\epsilon})$$

$$\leq\mathbb{E}[(|A|+|B|)^2\cdot(|A|^\epsilon+|B|^\epsilon)]=\mathbb{E}[(A^2+B^2+2|A|\cdot|B|)\cdot(|A|^\epsilon+|B|^\epsilon)]$$

$$=\mathbb{E}(|A|^{2+\epsilon}+|B|^{2+\epsilon}+2|A|\cdot|B|^{1+\epsilon}+2|B|\cdot|A|^{1+\epsilon}+B^2\cdot|A|^\epsilon+A^2\cdot|B|^\epsilon).\tag{4.6}$$

Let $s+t=2+\epsilon, s\in(0,2]$, $\frac{\epsilon}{2}<1-\frac{\alpha}{1-\alpha}$, by Hölder inequality,

$$\mathbb{E}[|A|^s\cdot|B|^t]=\mathbb{E}\{[\mathbb{E}_{m+\lfloor\frac{n}{2}\rfloor}(|A|^s)]\cdot|B|^t\}\leq\mathbb{E}\{[\mathbb{E}_{m+\lfloor\frac{n}{2}\rfloor}(|A|^2)]^{\frac{s}{2}}\cdot|B|^t\}$$

$$=\mathbb{E}\{[\mathbb{E}_{m+\lfloor\frac{n}{2}\rfloor}(|A|^2)-\mathbb{E}(|A|^2)+\mathbb{E}(|A|^2)]^{\frac{s}{2}}\cdot|B|^t\}\leq\mathbb{E}[|\mathbb{E}_{m+\lfloor\frac{n}{2}\rfloor}(|A|^2)-\mathbb{E}(|A|^2)|^{\frac{s}{2}}\cdot|B|^t]$$

$$+[\mathbb{E}(|A|^2)]^{\frac{s}{2}}\cdot\mathbb{E}(|B|^t)\leq[\mathbb{E}(|B|^{2+\epsilon})]^{\frac{t}{2+\epsilon}}\cdot\{\mathbb{E}[|\mathbb{E}_{m+\lfloor\frac{n}{2}\rfloor}(|A|^2)-\mathbb{E}(|A|^2)|^{\frac{2+\epsilon}{2}}]\}^{\frac{s}{2+\epsilon}}$$

$$+[\mathbb{E}(|B|^{2+\epsilon})]^{\frac{t}{2+\epsilon}}\cdot[\mathbb{E}(|A|^2)]^{\frac{s}{2}}.$$

By Lemma 4.1, Lemma 4.4, there is a constant $\bar{C}$ which is uniform over all ϕ_k,n,m such that the above inequality becomes

$$\leq 2[\mathbb{E}(|B|^{2+\epsilon})]^{\frac{t}{2+\epsilon}}\cdot\lfloor\frac{n}{2}\rfloor^{\frac{s}{2}}\cdot\bar{C}.$$

Apply the above inequality to (4.6) for $s=1,1+\epsilon,\epsilon,2$, respectively, then

$$\mathbb{E}(|A+B|^{2+\epsilon}) \le \mathbb{E}(|A|^{2+\epsilon}) + \mathbb{E}(|B|^{2+\epsilon})$$

$$+4\cdot\bar{C}\cdot[\mathbb{E}(|B|^{2+\epsilon})]^{\frac{1+\epsilon}{2+\epsilon}}\cdot\lfloor\frac{n}{2}\rfloor^{\frac{1}{2}} + 4\cdot\bar{C}\cdot[\mathbb{E}(|B|^{2+\epsilon})]^{\frac{1}{2+\epsilon}}\cdot\lfloor\frac{n}{2}\rfloor^{\frac{1+\epsilon}{2}}$$

$$+2\cdot\bar{C}\cdot[\mathbb{E}(|B|^{2+\epsilon})]^{\frac{2}{2+\epsilon}}\cdot\lfloor\frac{n}{2}\rfloor^{\frac{\epsilon}{2}} + 2\cdot\bar{C}\cdot[\mathbb{E}(|B|^{2+\epsilon})]^{\frac{\epsilon}{2+\epsilon}}\cdot\lfloor\frac{n}{2}\rfloor.$$

Then

$$\frac{\mathbb{E}(|A+B|^{2+\epsilon})}{n^{\frac{2+\epsilon}{2}}} \le \frac{\mathbb{E}(|A|^{2+\epsilon})}{\lfloor\frac{n}{2}\rfloor^{\frac{2+\epsilon}{2}}}\cdot\frac{\lfloor\frac{n}{2}\rfloor^{\frac{2+\epsilon}{2}}}{n^{\frac{2+\epsilon}{2}}} + \frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}\cdot\frac{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}{n^{\frac{2+\epsilon}{2}}}$$

$$+4\cdot\bar{C}\cdot[\frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}]^{\frac{1+\epsilon}{2+\epsilon}}\cdot\frac{\lfloor\frac{n}{2}\rfloor^{\frac{1}{2}}\cdot(n-\lfloor\frac{n}{2}\rfloor)^{\frac{1+\epsilon}{2}}}{n^{\frac{2+\epsilon}{2}}} + 4\cdot\bar{C}\cdot[\frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}]^{\frac{1}{2+\epsilon}}\cdot\frac{\lfloor\frac{n}{2}\rfloor^{\frac{1+\epsilon}{2}}\cdot(n-\lfloor\frac{n}{2}\rfloor)^{\frac{1}{2}}}{n^{\frac{2+\epsilon}{2}}}$$

$$+2\cdot\bar{C}\cdot[\frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}]^{\frac{2}{2+\epsilon}}\cdot\frac{\lfloor\frac{n}{2}\rfloor^{\frac{\epsilon}{2}}\cdot(n-\lfloor\frac{n}{2}\rfloor)}{n^{\frac{2+\epsilon}{2}}} + 2\cdot\bar{C}\cdot[\frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}]^{\frac{\epsilon}{2+\epsilon}}\cdot\frac{\lfloor\frac{n}{2}\rfloor\cdot(n-\lfloor\frac{n}{2}\rfloor)^{\frac{\epsilon}{2}}}{n^{\frac{2+\epsilon}{2}}}$$

$$= \frac{\mathbb{E}(|A|^{2+\epsilon})}{\lfloor\frac{n}{2}\rfloor^{\frac{2+\epsilon}{2}}}\cdot[\frac{1}{2}+o(1)]^{\frac{2+\epsilon}{2}} + \frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}\cdot[\frac{1}{2}+o(1)]^{\frac{2+\epsilon}{2}}$$

$$+4\cdot\bar{C}\cdot[\frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}]^{\frac{1+\epsilon}{2+\epsilon}}\cdot[\frac{1}{2}+o(1)]^{\frac{2+\epsilon}{2}} + 4\cdot\bar{C}\cdot[\frac{\mathbb{E}(|B|^{2+\epsilon})}{(n-\lfloor\frac{n}{2}\rfloor)^{\frac{2+\epsilon}{2}}}]^{\frac{1}{2+\epsilon}}\cdot[\frac{1}{2}+o(1)]^{\frac{2+\epsilon}{2}}$$

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

Let $a_n := \max(\sup_{m\ge 1}\frac{\mathbb{E}(|\sum_{m\le i\le m+n-1}\phi_i\circ T^i|^{2+\epsilon})}{n^{\frac{2+\epsilon}{2}}}, \sup_{m\ge 1}\frac{\mathbb{E}(|\sum_{m\le i\le m+n}\phi_i\circ T^i|^{2+\epsilon})}{(n+1)^{\frac{2+\epsilon}{2}}})$, the estimates above shows that:

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

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

$$a_n \le a_{\lfloor\frac{n}{2}\rfloor}\cdot[\frac{1}{2}+o(1)]^{\frac{2+\epsilon}{2}}\cdot g(a_{\lfloor\frac{n}{2}\rfloor}).$$

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

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

$$[\frac{1}{2} + o(1)]^{\frac{2+\epsilon}{2}} \text{ is close to } (\frac{1}{2})^{\frac{2+\epsilon}{2}} < \frac{1}{2}.$$

Then we can choose big x_0, N such that for any $n \geq N, x \geq x_0$, $[\frac{1}{2} + o(1)]^{\frac{2+\epsilon}{2}} \cdot g(x) < 1$.

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

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

Therefore, $b_n \leq b_{\lfloor \frac{n}{2} \rfloor}$ for any $n \geq N$; furthermore, for any $n \geq 1$,

$$\sup_{m \geq 1} \frac{\mathbb{E}(|\sum_{m \leq i \leq m+n-1} \phi_i \circ T^i|^{2+\epsilon})}{n^{\frac{2+\epsilon}{2}}} \leq a_n \leq b_n \leq \max(b_1, b_2, \cdots, b_N) < \infty.$$

Second, apply the above inequality to Theorem B in [44]:

Lemma 4.7 (Theorem B in [44])

Suppose X_i has finite variance, zero mean, and

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

then

$$\sup_m \mathbb{E}(\max_{1 \leq k \leq n} |\sum_{i=m+1}^{m+k} X_i|^{2+\epsilon}) \precsim n^{\frac{2+\epsilon}{2}},$$

where the constant indicated in $\precsim$ is independent of n.

Let $X_i := \phi_i \circ T^i$, we obtain the desired maximal inequality.

∎

To find the desired Gaussian vectors in the definition of the VASIP, Berkes and Philipp [8] gave a criterion:

Theorem 4.8 (see [8])

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

$$\mathbb{E}|\mathbb{E}[\exp(iu^T\cdot X_k)|\mathcal{G}_{k-1}]-\exp(-\frac{1}{2}u^T\cdot H_k\cdot u)|\leq\lambda_k,$$

$$\mu_k\{u:|u|\geq\frac{1}{4}T_k\}\leq\delta_k.$$

Then without changing its distribution we can define $(X_k)_{k\geq 1}$ 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

$$\tilde{P}(|X_k-G_k|\geq\alpha_k)\leq\alpha_k,$$

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

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

$$\sum_{k\geq 1}|X_k-G_k|<\infty.$$

Berkes and Philipp constructed Gaussian vectors inductively, which relies heavily on the increasing filtration $(\mathcal{G}_k)_{k\geq 1}$. However, our filtration $(T^{-k}\mathcal{B})_{k\geq 1}$ is decreasing. One way to overcome this difficulty is to construct increasing σ-algebra from T_k with certain Markovian behavior (see [34,35]). Since we do not have too much information on T_k so far, we will keep using decreasing filtration $(T^{-k}\mathcal{B})_{k\geq 1}$ and derive the following lemma, which plays a crucial role in our proof:

Lemma 4.9 (VASIP criteria)

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. Let $(H_k)_{k\geq 1}$ be a family of positive semi-definite $d\times d$ matrices. Assume μ_k is Gaussian distribution with characteristic function $e^{-\frac{1}{2}u^T\cdot H_k\cdot u}$. Suppose *that there are some non-negative numbers $T_k \geq 10^8 d, \lambda_k, \delta_k$, such that for any $u \in \mathbb{R}^d$ with $|u| \leq T_k$:*

$$\mathbb{E}|\mathbb{E}[\exp(iu^T \cdot Y_k)|\mathcal{F}_{k+1}] - \exp(-\frac{1}{2}u^T \cdot H_k \cdot u)| \leq \lambda_k,$$

$$\mu_k\{u : |u| \geq \frac{1}{4}T_k\} \leq \delta_k.$$

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

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

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

Proof Before proving this lemma, let's recall the procedure of how to construct Gaussian vectors in [8]: G_1 is constructed with distribution μ_1, extending the probability space Ω to $\Omega \times I$ by multiplying an unit interval I with Lebesgue measure if the original probability space has atoms. Inductively, assume $G_1, G_2, \cdots, G_{k-1}$ have been constructed, partition the extended probability space as the union of countably many $\sigma(G_1, \cdots, G_{k-1})$-measurable elements. Locally, on each such element, construct G_k, and extend the extended probability space by multiplying another unit interval. Obtain global G_k by gluing the local G_k. The final extended probability space is $\Omega \times I^{\mathbb{N}}$.

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

For the 1-st row of the array, let $G_1^1 := G_1, \mu_1 = L(G_1^1), \Omega_1 := \Omega \times I_1$ (denote its probability by P) as in [8]. Assume the previous $(n-1)$ rows of the array are done: the extended probability space Ω_{n-1} (still denote its probability by P) and $(G_k^{n-1})_{k\leq n-1}$ are constructed.

For the n-th row of the array, consider increasing filtration $(\mathcal{F}_{n+2-k})_{1\leq k\leq n+1}$. By Theorem 4.8, we can construct $G_{n+1}^n, G_1^n, \cdots, G_n^n$ and probability space $\Omega_{n-1} \times I_n^{\mathbb{N}}$ (still denote its probability by

P) such that

$$P(|Y_k - G_k^n| \geq \alpha_k) \leq \alpha_k, \mu_k = L(G_k^n), 1 \leq k \leq n,$$

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

Since G_{n+1}^n and α_{n+1} do not make contribution, discard them. Then we have $G_n^n, \cdots, G_1^n$ and probability space $\Omega_{n-1} \times I_n^{\mathbb{N}}$ such that

$$P(|Y_k - G_k^n| \geq \alpha_k) \leq \alpha_k,$$

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

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

$$P(|Y_k - G_k^n| \geq \alpha_k) \leq \alpha_k,$$

$$\mu_k = L(G_k^n) \text{ for all } k, n \geq 1,$$

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

Consider a new triangular array $(Y_k, G_k^n)_{1\leq k\leq n, n\geq 1}$, we will use induction to construct the $(G_k)_{k\geq 1}$ as indicated in our lemma:

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

$$(Y_1, G_1^n) \xrightarrow[subsequence]{d} (Y_1', G_1'),$$

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

Assume the $(m-1)$-th step is done, that is, $(Y_1, G_1^n, Y_2, G_2^n, \cdots, Y_{m-1}, G_{m-1}^n)_{n\geq m-1}$ has a subsequence with weak limit $(Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}')$, and an extended probability space

$\prod_{-(m-1)\le i\le -1} I_i \times \Omega \times \prod_{i\ge 1} I_i^{\mathbb{N}}$ (still denote its probability by P) such that

$$(Y_1, G_1^n, Y_2, G_2^n, \cdots, Y_{m-1}, G_{m-1}^n) \xrightarrow[subsequence]{d} (Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}'),$$

$$P(|Y_k' - G_k'| \ge \alpha_k) \le \alpha_k, \text{ for any } k \le m-1.$$

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

$$(Y_1, G_1^n, Y_2, G_2^n, \cdots, Y_m, G_m^n) \xrightarrow[subsequence]{d} (\bar{Y}_1', \bar{G}_1', \cdots, \bar{Y}_m', \bar{G}_m').$$

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

$$(Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}') \stackrel{d}{=} (\bar{Y}_1', \bar{G}_1', \bar{Y}_2', \bar{G}_2', \cdots, \bar{Y}_{m-1}', \bar{G}_{m-1}').$$

By Lemma 7.3, there is (Y_m', G_m') such that

$$(Y_1', G_1', Y_2', G_2', \cdots, Y_{m-1}', G_{m-1}', Y_m', G_m') \stackrel{d}{=} (\bar{Y}_1', \bar{G}_1', \bar{Y}_2', \bar{G}_2', \cdots, \bar{Y}_{m-1}', \bar{G}_{m-1}', \bar{Y}_m', \bar{G}_m').$$

Meanwhile, we have an extended probability space $\prod_{-m\le i\le -1} I_i \times \Omega \times \prod_{i\ge 1} I_i^{\mathbb{N}}$ (still denote its probability by P).

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

$$(Y_1, G_1^n, Y_2, G_2^n, \cdots, Y_m, G_m^n) \xrightarrow[subsequence]{d} (Y_1', G_1', Y_2', G_2', \cdots, Y_m', G_m').$$

Then by the diagonal argument, there is a subsequence (independent of m), such that for any $m \ge 1$,

$$(Y_1, G_1^n, Y_2, G_2^n, \cdots, Y_m, G_m^n) \xrightarrow[subsequence]{d} (Y_1', G_1', Y_2', G_2', \cdots, Y_m', G_m'),$$

$$P(|Y_k' - G_k'| \ge \alpha_k) \le \alpha_k, \text{ for any } k \le m.$$

Therefore, for any $k \geq 1$,

$$L(Y_1', \cdots, Y_k') = L(Y_1, \cdots, Y_k),$$

$$L(G_1', \cdots, G_k') = \bigotimes_{1 \leq i \leq k} \mu_i.$$

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

Use Lemma 7.3 again in a similar way, there is $(G_i)_{i \geq 1}$ and a final extended probability space $\prod_{i \leq -1} I_i \times I_0 \times \Omega \times \prod_{i \geq 1} I_i^{\mathbb{N}}$ (still denote its probability by $\tilde{P}$) such that

$$((Y_i')_{i \geq 1}, (G_i')_{i \geq 1}) \stackrel{d}{=} ((Y_i)_{i \geq 1}, (G_i)_{i \geq 1}).$$

Therefore, for any $k \geq 1$,

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

$$L(G_1, \cdots, G_k) = \bigotimes_{1 \leq i \leq k} \mu_i,$$

where $\alpha_k := 16d \cdot \frac{\log T_k}{T_k} + 4\lambda_k^{\frac{1}{2}} \cdot T_k^d + \delta_k, k \geq 1$. ∎

With all lemmas above, we are ready to prove Theorem 2.8.

5 Proofs of main theorems

5.1 Proof of Theorem 2.8

Blocks construction

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 $|I_n| = \lfloor n^c \rfloor, c > 0$. So $\bigcup_{i\geq 1} I_i = \mathbb{N}$. Let $a \in (\frac{1}{2}, 1), c_n := \lfloor n^{c(1-a)} \rfloor$. Construct consecutive blocks $\{I_{n,i}, 1 \leq i \leq c_n + 1\}$ in I_n such that: $|I_{n,i}| = \lfloor n^{ca} \rfloor, 1 \leq i \leq c_n$, 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 $|I_{n,c_n+1}| \leq 2\lfloor n^{ca} \rfloor$ and $\bigcup_{1\leq i\leq c_n+1} I_{n,i} = I_n$. Let $a_n := \sum_{i\leq n} |I_i| \approx n^{c+1}$ and

$$X_n := \sum_{i\in I_n} \phi_i \circ T^i, Y_n := \frac{X_n}{\sqrt{b_n}}, H_n := \frac{\mathbb{E}(X_n \cdot X_n^T)}{b_n}, \mathcal{F}_n := T^{-a_{n-1}-1}\mathcal{B},$$

$$\text{where } b_n := \lambda(\sigma_{a_n}^2) \gtrsim n^{\gamma(c+1)}, X_n, Y_n \text{ are } \mathcal{F}_n\text{-measurable},$$

$$X_{n,i} := \sum_{i\in I_{n,i}} \phi_i \circ T^i, \mathcal{F}_{n,i} := T^{-\sum_{k\leq i-1}|I_{n,k}|-1}\mathcal{B}, X_{n,i} \text{ is } \mathcal{F}_{n,i}\text{-measurable},$$

$$T_n := n^\kappa, \delta_n := \mu_n\{u : |u| \geq \frac{1}{4}T_n\},$$

where μ_n is mean zero Gaussian distribution with variance H_n, and κ, c, a will be given in Appendix, Lemma 7.1 and its proof.

Estimate of $\mathbb{E}|\mathbb{E}[\exp(i\boldsymbol{u}^{\boldsymbol{T}} \cdot \boldsymbol{Y_n})|\boldsymbol{\mathcal{F}_{n+1}}] - \exp(-\frac{1}{2}\boldsymbol{u}^{\boldsymbol{T}} \cdot \boldsymbol{H_n} \cdot \boldsymbol{u})|$

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

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

$$= \mathbb{E}|\mathbb{E}_{a_n+1} \exp(iu^T \cdot \frac{X_n}{\sqrt{b_n}}) - \exp(-\frac{1}{2}u^T \cdot \frac{\mathbb{E}(X_n \cdot X_n^T)}{b_n} \cdot u).| \tag{5.1}$$

First, we note that

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

$$+\sum_{1\le i<j\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,j}^T)^T=\sum_{1\le i\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,i}^T)+\sum_{1\le i<c_n+1}\mathbb{E}[X_{n,i}\cdot(\sum_{i<j\le c_n+1}X_{n,j}^T)]$$

$$+\sum_{1\le i<c_n+1}\mathbb{E}[X_{n,i}\cdot(\sum_{i<j\le c_n+1}X_{n,j}^T)]^T.$$

By Lemma 4.5, the above equality becomes

$$=\sum_{1\le i\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,i}^T)+2\sum_{1\le i<c_n+1}O(|I_n|^{\max(3-\frac{1}{\alpha},0)})$$

$$=\sum_{1\le i\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,i}^T)+c_n\cdot O(|I_n|^{\max(3-\frac{1}{\alpha},0)})=\sum_{1\le i\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,i}^T)+O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)}).$$

Therefore, inspired by the estimate method in [9] and use the above equality, we have the following:

$$\mathbb{E}|\mathbb{E}_{a_n+1}\exp(iu^T\cdot\frac{X_n}{\sqrt{b_n}})-\exp(-\frac{1}{2}u^T\cdot\frac{\mathbb{E}(X_n\cdot X_n^T)}{b_n}\cdot u)|$$

$$\le\mathbb{E}|\mathbb{E}_{a_n+1}\exp(iu^T\cdot\frac{X_n}{\sqrt{b_n}})-\exp(-\frac{1}{2}u^T\cdot\frac{\sum_{1\le i\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,i}^T)}{b_n}\cdot u)|$$

$$+|\exp(-\frac{1}{2}u^T\cdot\frac{\sum_{1\le i\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,i}^T)}{b_n}\cdot u)-\exp(-\frac{1}{2}u^T\cdot\frac{\mathbb{E}(X_n\cdot X_n^T)}{b_n}\cdot u)|$$

$$\precsim\mathbb{E}|\mathbb{E}_{a_n+1}\exp(iu^T\cdot\frac{X_n}{\sqrt{b_n}})-\exp(-u^T\cdot\frac{\sum_{1\le i\le c_n+1}\mathbb{E}(X_{n,i}\cdot X_{n,i}^T)}{2b_n}\cdot u)|+|u|^2\cdot\frac{O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)})}{b_n}$$

$$=\mathbb{E}|\mathbb{E}_{a_n+1}[\sum_{0\le k\le c_n}\exp(-\frac{\sum_{0<i\le k}\mathbb{E}(u^T\cdot X_{n,i})^2}{2b_n})\cdot\exp(i\frac{\sum_{k<i\le c_n+1}u^T\cdot X_{n,i}}{\sqrt{b_n}})$$

$$-\exp(-\frac{\sum_{0<i\le k+1}\mathbb{E}(u^T\cdot X_{n,i})^2}{2b_n})\cdot\exp(i\cdot\frac{\sum_{1+k<i\le c_n+1}u^T\cdot X_{n,i}}{\sqrt{b_n}})]|+|u|^2\cdot\frac{O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)})}{b_n}$$

$$=\mathbb{E}|\mathbb{E}_{a_n+1}\{\sum_{0\le k\le c_n}\exp(-\frac{\sum_{0<i\le k}\mathbb{E}(u^T\cdot X_{n,i})^2}{2b_n})\cdot[\exp(i\frac{u^T\cdot X_{n,k+1}}{\sqrt{b_n}})$$

$$-\exp(-\frac{\mathbb{E}(u^T\cdot X_{n,k+1})^2}{2b_n})]\cdot\exp(i\frac{\sum_{k+1<i\leq c_n+1}u^T\cdot X_{n,i}}{\sqrt{b_n}})\}|+|u|^2\cdot\frac{O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)})}{b_n}$$

$$=\mathbb{E}|\mathbb{E}_{a_n+1}\{\sum_{0\leq k\leq c_n}\exp(-\frac{\sum_{0<i\leq k}\mathbb{E}(u^T\cdot X_{n,i})^2}{2b_n})\cdot\exp(i\cdot\frac{\sum_{k+1<i\leq c_n+1}u^T\cdot X_{n,i}}{\sqrt{b_n}})\times$$

$$\mathbb{E}[\exp(i\frac{u^T\cdot X_{n,k+1}}{\sqrt{b_n}})-\exp(-\frac{\mathbb{E}(u^T\cdot X_{n,k+1})^2}{2b_n})|\mathcal{F}_{n,k+2}]\}|+|u|^2\cdot\frac{O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)})}{b_n}$$

$$\leq\sum_{0\leq k\leq c_n}\mathbb{E}|\mathbb{E}[\exp(i\frac{u^T\cdot X_{n,k+1}}{\sqrt{b_n}})-\exp(-\frac{\mathbb{E}(u^T\cdot X_{n,k+1})^2}{2b_n})|\mathcal{F}_{n,k+2}]|+|u|^2\cdot\frac{O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)})}{b_n}.$$

Let $|u|\leq T_n=n^\kappa$, by Lemma 4.1, 4.2, 4.3, 4.6, and Taylor expansion:

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

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

for any $\epsilon_0\in(0,\min(1,2-\frac{2\alpha}{1-\alpha}))$, the above inequality becomes:

$$=\sum_{0\leq k\leq c_n}\mathbb{E}|\mathbb{E}\{[1+i\frac{u^T\cdot X_{n,k+1}}{\sqrt{b_n}}-\frac{1}{2}(\frac{u^T\cdot X_{n,k+1}}{\sqrt{b_n}})^2+O(|\frac{u^T\cdot X_{n,k+1}}{\sqrt{b_n}}|^{2+\epsilon_0})]|\mathcal{F}_{n,k+2}\}$$

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

$$\leq\sum_{0\leq k\leq c_n}\{\frac{|u|}{\sqrt{b_n}}+\frac{1}{2}\frac{|u|^2}{b_n}\cdot\mathbb{E}|\mathbb{E}\{[X_{n,k+1}\cdot X_{n,k+1}^T-\mathbb{E}(X_{n,k+1}\cdot X_{n,k+1}^T)]|\mathcal{F}_{n,k+2}\}|$$

$$+|u|^4\cdot\frac{|\mathbb{E}(X_{n,k+1}\cdot X_{n,k+1}^T)|^2}{b_n^2}+|u|^{2+\epsilon_0}\cdot\frac{\mathbb{E}(|X_{n,k+1}|^{2+\epsilon_0})}{b_n^{\frac{2+\epsilon_0}{2}}}\}+|u|^2\cdot\frac{O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)})}{b_n}$$

$$\precsim\frac{n^{\kappa+c(1-a)}}{n^{\frac{\gamma(1+c)}{2}}}+\frac{n^{2\kappa+c(1-a)+ca\frac{\alpha}{1-\alpha}}}{n^{\gamma(1+c)}}+\frac{n^{c(1-a)+4\kappa+2ca}}{n^{2\gamma(1+c)}}$$

$$+\frac{n^{\kappa(2+\epsilon_0)+c(1-a)+ca\frac{2+\epsilon_0}{2}}}{n^{\frac{\gamma(1+c)(2+\epsilon_0)}{2}}}+\frac{n^{2\kappa+c(1-a)+c\max(3-\frac{1}{\alpha},0)}}{n^{\gamma(1+c)}}\precsim\frac{1}{n^v},$$

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

Equations (5.2)-(5.7) for the range of γ

First, by Lemma 4.1,

$$|H_n| = |\frac{\mathbb{E}(X_n \cdot X_n^T)}{b_n}| \precsim \frac{n^c}{n^{\gamma(c+1)}}.$$

So if

$$c - \gamma(c+1) < 0, \tag{5.2}$$

denote the probability for distribution μ_n by P, then

$$\delta_n = \mu_n\{u : |u| \geq \frac{1}{4}T_n\} = P(|N(0, H_n)| > \frac{1}{4}T_n) \precsim P(|N(0, \frac{n^c}{n^{\gamma(c+1)}})| > \frac{1}{4}T_n)$$

is exponential decay, so

$$\mu_n\{u : |u| \geq \frac{1}{4}T_n\} = \delta_n \precsim \frac{1}{n^v}.$$

Then

$$\alpha_n = 16d \cdot \frac{\log T_n}{T_n} + 4\lambda_n^{\frac{1}{2}} \cdot T_n^d + \delta_n \precsim \frac{1}{n^\kappa} + \frac{1}{n^{\frac{v}{2}-d\kappa}} \precsim \frac{1}{n^{\min(\kappa, \frac{v}{2}-d\kappa)}}.$$

Note that $\frac{\alpha}{1-\alpha} < 1, 3 - \frac{1}{\alpha} < 1$. If $\kappa > 1$, choose γ, a closed to 1 carefully (γ will be given in Appendix, Lemma 7.1 and its proof) such that

$$\min(\kappa, \frac{v}{2} - d\kappa) > 1. \tag{5.3}$$

So $\sum_{n\geq 1} \alpha_n < \infty$. By Lemma 4.9, there are Gaussian vectors G''_n with covariance matrix $\mathbb{E}(X_n \cdot X_n^T)$ 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,

$$\sum_{i\le n} X_i - G_i'' \precsim \sum_{i\le n} \alpha_i \cdot \sqrt{b_i} \precsim \sum_{i\le n} \frac{1}{i^{\min(\kappa, \frac{v}{2}-d\kappa)}} \cdot i^{\frac{c+1}{2}} \precsim n^{1+\frac{c+1}{2}-\min(\kappa,\frac{v}{2}-d\kappa)}.$$

Choosing γ, κ, a, c carefully (will be given in Appendix Lemma 7.1) such that

$$\gamma\frac{c+1}{2} > 1 + \frac{c+1}{2} - \min(\kappa, \frac{v}{2} - d\kappa). \tag{5.4}$$

Therefore, there is a small ϵ' such that

$$\sum_{i\le n} X_i - \sum_{i\le n} G_i'' \precsim n^{1+\frac{c+1}{2}-\min(\kappa,\frac{v}{2}-d\kappa)} \precsim \lambda(\sigma_{a_n}^2)^{\frac{1-\epsilon'}{2}} \text{ a.s..}$$

From block length a_n to general $m \in \mathbb{N}$:

For any m, there is n such that $a_n \le m < a_{n+1}$, then we have the following lemmas:

Lemma 5.1

If

$$c - \gamma(c+1) < 0, \tag{5.5}$$

$$1 + (c+1)(\max\{3 - \frac{1}{\alpha}, 0\} - \gamma) < 0 \tag{5.6}$$

(these relations are possible, see Appendix, Lemma 7.1), then

$$\lambda(\sigma_m^2) \approx \lambda(\sigma_{a_n}^2).$$

Proof

By Lemma 4.1 and Lemma 4.5,

$$\mathbb{E}[(\sum_{i\le m} \phi_i \circ T^i) \cdot (\sum_{i\le m} \phi_i \circ T^i)^T] = \mathbb{E}[(\sum_{i\le n} X_i) \cdot (\sum_{i\le n} X_i)^T] + \mathbb{E}[(\sum_{i\le n} X_i) \cdot (\sum_{a_n<i\le m} \phi_i \circ T^i)^T]$$

$$+\mathbb{E}[(\sum_{i\le n} X_i) \cdot (\sum_{a_n<i\le m} \phi_i \circ T^i)^T]^T + \mathbb{E}[(\sum_{a_n<i\le m} \phi_i \circ T^i) \cdot (\sum_{a_n<i\le m} \phi_i \circ T^i)^T]$$

$$\precsim \mathbb{E}[(\sum_{i\le n} X_i)\cdot(\sum_{i\le n} X_i)^T] + n\cdot a_{n+1}^{\max(3-\frac{1}{\alpha},0)} + n\cdot a_{n+1}^{\max(3-\frac{1}{\alpha},0)} + a_{n+1} - a_n.$$

Since $a_n \approx a_{n+1}$, the above inequality becomes

$$\precsim \mathbb{E}[(\sum_{i\le n} X_i)\cdot(\sum_{i\le n} X_i)^T] + a_n^{\gamma}\cdot n\cdot a_n^{\max(3-\frac{1}{\alpha},0)-\gamma} + a_n^{\gamma}\cdot n\cdot a_n^{\max(3-\frac{1}{\alpha},0)-\gamma} + a_n^{\gamma}\cdot(\frac{a_{n+1}-a_n}{a_n^{\gamma}}).$$

Since $\mathbb{E}[(\sum_{i\le n} X_i)\cdot(\sum_{i\le n} X_i)^T] \succsim a_n^{\gamma}$, then the above inequality becomes

$$\precsim \{\mathbb{E}[(\sum_{i\le n} X_i)\cdot(\sum_{i\le n} X_i)^T]\}\cdot[1 + n\cdot a_n^{\max(3-\frac{1}{\alpha},0)-\gamma} + n\cdot a_n^{\max(3-\frac{1}{\alpha},0)-\gamma} + (\frac{a_{n+1}-a_n}{a_n^{\gamma}})].$$

$$\precsim \{\mathbb{E}[(\sum_{i\le n} X_i)\cdot(\sum_{i\le n} X_i)^T]\}\cdot(1 + n^{1+(c+1)(\max(3-\frac{1}{\alpha},0)-\gamma)} + n^{c-\gamma(c+1)}).$$

By (5.5), (5.6), we have

$$\lambda(\sigma_m^2) \precsim 3\cdot\inf_{|u|=1} u^T\cdot\mathbb{E}[(\sum_{i\le n} X_i)\cdot(\sum_{i\le n} X_i)^T]\cdot u = 3\lambda(\sigma_{a_n}^2).$$

Similarly, by Lemma 4.1 and Lemma 4.5,

$$\mathbb{E}[(\sum_{i\le n} X_i)\cdot(\sum_{i\le n} X_i)^T] \precsim \mathbb{E}[(\sum_{i\le m}\phi_i\circ T^i)\cdot(\sum_{i\le m}\phi_i\circ T^i)^T] + n\cdot a_n^{\max(3-\frac{1}{\alpha},0)} + n\cdot a_n^{\max(3-\frac{1}{\alpha},0)} + a_{n+1} - a_n$$

$$\precsim \mathbb{E}[(\sum_{i\le m}\phi_i\circ T^i)\cdot(\sum_{i\le m}\phi_i\circ T^i)^T] + m^{\frac{1}{c+1}}\cdot m^{\max(3-\frac{1}{\alpha},0)} + m^{\frac{1}{c+1}}\cdot m^{\max(3-\frac{1}{\alpha},0)} + m^{\frac{c}{c+1}}.$$

Since $\mathbb{E}[(\sum_{i\le m}\phi_i\circ T^i)\cdot(\sum_{i\le m}\phi_i\circ T^i)^T] \succsim m^{\gamma}$, then the above inequality becomes

$$\precsim \{\mathbb{E}[(\sum_{i\le m}\phi_i\circ T^i)\cdot(\sum_{i\le m}\phi_i\circ T^i)^T]\}\cdot(1 + m^{\frac{1}{c+1}}\cdot m^{\max(3-\frac{1}{\alpha},0)-\gamma} + m^{\frac{c}{c+1}-\gamma}).$$

By (5.5), (5.6) again, we have

$$\lambda(\sigma_{a_n}^2) \precsim 3\lambda(\sigma_m^2).$$

∎

Lemma 5.2

If

$$\frac{1}{2}\gamma(c+1)(2+\epsilon)-c(1+\frac{\epsilon}{2})>1, \tag{5.7}$$

where $\epsilon<\min(1,2-\frac{2\alpha}{1-\alpha})$*, then there is a small* $\epsilon'>0$ *such that*

$$\sup_{a_n\le m\le a_{n+1}}|\sum_{a_n<i\le m}\phi_i\circ T^i|\precsim\lambda(\sigma_{a_n}^2)^{\frac{1}{2}-\epsilon'}\ \textit{a.s..}$$

Proof

By Lemma 4.6,

$$\mathbb{E}(|\frac{\sup_{a_n\le m\le a_{n+1}}|\sum_{a_n<i\le m}\phi_i\circ T^i|}{\lambda(\sigma_{a_n}^2)^{\frac{1}{2}-\epsilon'}}|^{2+\epsilon})\precsim\frac{n^{c(1+\frac{\epsilon}{2})}}{n^{\gamma(c+1)(2+\epsilon)(\frac{1}{2}-\epsilon')}}=\frac{1}{n^{\gamma(c+1)(2+\epsilon)(\frac{1}{2}-\epsilon')-c(1+\frac{\epsilon}{2})}}.$$

From (5.7), there is a small $\epsilon'>0$ such that

$$\gamma(c+1)(2+\epsilon)(\frac{1}{2}-\epsilon')-c(1+\frac{\epsilon}{2})>1.$$

By the Borel-Cantelli Lemma, we have

$$\sup_{a_n\le m\le a_{n+1}}|\sum_{a_n<i\le m}\phi_i\circ T^i|\precsim\lambda(\sigma_{a_n}^2)^{\frac{1}{2}-\epsilon'}\ \text{a.s..}$$

∎

Find the independent Gaussian vectors in the definition of VASIP

For any $k\in\mathbb{N}$, define

$$G_k:=\begin{cases}G_i'', & \text{if } k=a_i\\ 0, & \text{if } k\neq a_i\end{cases}. \tag{5.8}$$

We claim $\sum_{i\le m}G_i$ matches $\sum_{i\le m}\phi_i\circ T^i$ in the sense of (2.2) and (2.1):

Verify (2.2): for any m, there is n such that $a_n \leq m < a_{n+1}$. Recall $\tilde{\mathbb{E}}(G_i'' \cdot G_i''^T) = \mathbb{E}(X_i \cdot X_i^T)$, where $\tilde{\mathbb{E}}(\cdot)$ is the expectation of the probability of the extended probability space, then we have

$$\mathbb{E}[(\sum_{i\leq m} \phi_i \circ T^i) \cdot (\sum_{i\leq m} \phi_i \circ T^i)^T] - \sum_{i\leq m} \tilde{\mathbb{E}}(G_i \cdot {G_i}^T)$$

$$= \mathbb{E}[(\sum_{i\leq m} \phi_i \circ T^i) \cdot (\sum_{i\leq m} \phi_i \circ T^i)^T] - \sum_{i\leq n} \tilde{\mathbb{E}}(G_i'' \cdot G_i''^T)$$

$$= \sum_{1\leq i<j\leq n} \mathbb{E}(X_i \cdot X_j^T) + \sum_{1\leq i<j\leq n} \mathbb{E}(X_i \cdot X_j^T)^T + \mathbb{E}[(\sum_{i\leq n} X_i) \cdot (\sum_{a_n<i\leq m} \phi_i \circ T^i)^T]$$

$$+\mathbb{E}[(\sum_{i\leq n} X_i) \cdot (\sum_{a_n<i\leq m} \phi_i \circ T^i)^T]^T + \mathbb{E}[(\sum_{a_n<i\leq m} \phi_i \circ T^i) \cdot (\sum_{a_n<i\leq m} \phi_i \circ T^i)^T].$$

By Lemma 4.1 and Lemma 4.5, the above equality becomes

$$\precsim n \cdot a_n^{\max(3-\frac{1}{\alpha},0)} + n \cdot a_{n+1}^{\max(3-\frac{1}{\alpha},0)} + a_{n+1} - a_n \precsim n^{1+(c+1)\max(3-\frac{1}{\alpha},0)} + n^c$$

$$\precsim \lambda(\sigma_{a_n}^2)^{\frac{1+(c+1)\max(3-\frac{1}{\alpha},0)}{\gamma(c+1)}} + \lambda(\sigma_{a_n}^2)^{\frac{c}{\gamma(c+1)}}.$$

By (5.5), (5.6), there is a small $\epsilon' > 0$ such that

$$\frac{1+(c+1)\max(3-\frac{1}{\alpha},0)}{\gamma(c+1)} < 1-\epsilon',$$

$$\frac{c}{\gamma(c+1)} < 1-\epsilon'.$$

Therefore, by Lemma 5.1,

$$\mathbb{E}[(\sum_{i\leq m} \phi_i \circ T^i) \cdot (\sum_{i\leq m} \phi_i \circ T^i)^T] - \sum_{i\leq n} \tilde{\mathbb{E}}(G_i'' \cdot G_i''^T) \precsim \lambda(\sigma_{a_n}^2)^{1-\epsilon'} \precsim \lambda(\sigma_m^2)^{1-\epsilon'}.$$

Verify (2.1): By Lemma 5.1 and Lemma 5.2, we have

$$\sum_{i\leq m} \phi_i \circ T^i - \sum_{i\leq m} G_i = \sum_{i\leq n} X_i - \sum_{i\leq n} G_i'' + \sum_{a_n<i\leq m} \phi_i \circ T^i$$

$$\precsim \lambda(\sigma_{a_n}^2)^{\frac{1-\epsilon'}{2}} + \sup_{a_n \le m \le a_{n+1}} |\sum_{a_n < i \le m} \phi_i \circ T^i| \precsim \lambda(\sigma_{a_n}^2)^{\frac{1-\epsilon'}{2}} \approx \lambda(\sigma_m^2)^{\frac{1-\epsilon'}{2}} \text{ a.s..}$$

In sum, (2.1) and (2.2) hold if (5.2)-(5.7) are all satisfied, then the VASIP holds. The range for γ will be derived from (5.2)-(5.7), see the computation in Appendix, Lemma 7.1.

∎

5.2 Proof of Theorem 2.9

Proof Fix $\omega \in \Omega$, we will apply our Theorem 2.8 to the non-stationary dynamical system $(X, \mathcal{B}, (T_{\sigma^k\omega})_{k\ge0}, \mu_\omega)$ to prove the VASIP w.r.t. μ_ω. From the proof of Theorem 2.8, we note that the VASIP holds via proving Lemma 4.1-4.6. So, to prove Theorem 2.9, we will prove the random versions of Lemma 4.1-4.6. Since Lemma 4.4 and Lemma 4.6 are deduced from the other four, and Lemma 4.5 is deduced from Lemma 4.1 and Lemma 4.3, so we just need to give the random versions of Lemma 4.1, Lemma 4.2, Lemma 4.3. For a.e. $\omega \in \Omega$, fix it and take μ_ω as the reference probability, define:

$$\mathbb{E}^\omega(\cdot) = \int (\cdot) d\mu_\omega, \mathbb{E}_n^\omega(\cdot) := \mathbb{E}^\omega[(\cdot)|(T_\omega^n)^{-1}\mathcal{B}].$$

The proof of random version of Lemma 4.1 is similar to Lemma 4.1: by (A1'),

$$\int (\sum_{m\le k\le n+m-1} \phi_{\sigma^k\omega} \circ T_\omega^k) \cdot (\sum_{m\le k\le m+n-1} \phi_{\sigma^k\omega} \circ T_\omega^k)^T d\mu_\omega$$

$$\precsim O(n) + \sum_{m\le i<j\le n+m-1} \int |P_{\sigma^i\omega}^{j-i}(\phi_{\sigma^i\omega} \cdot h_{\sigma^i\omega})| d\mu \precsim O(n) + \sum_{m\le i<j\le n+m} \frac{1}{(j-i)^{\frac{1}{\alpha}-1}} = O(n).$$

The proof of random version of Lemma 4.2 is similar to Lemma 4.2: by (A1'),

$$\mathbb{E}^\omega|\mathbb{E}_{n+m}^\omega \sum_{m\le k\le n+m-1} \phi_{\sigma^k\omega} \circ T_\omega^k| = \sup_{||\psi||_{L^\infty(X;\mathbb{R})}\le 1} \int \psi \circ T_\omega^{n+m} \cdot \sum_{m\le k\le n+m-1} \phi_{\sigma^k\omega} \circ T_\omega^k d\mu_\omega$$

$$\leq \sum_{m\leq k\leq n+m-1} \int |P_{\sigma^k\omega}^{n+m-k}(\phi_{\sigma^k\omega}\cdot h_{\sigma^k\omega})|d\mu \precsim \sum_{m\leq k\leq n+m-1} \frac{1}{(m+n-k)^{\frac{1}{\alpha}-1}} = O(1).$$

The proof of random version of Lemma 4.3 is similar to the estimate of Lemma 4.3, we will just outline the key parts:

$$\mathbb{E}^\omega|\mathbb{E}_{n+m}^\omega[(\sum_{k=m}^{n+m-1} \phi_{\sigma^k\omega}\circ T_\omega^k)\cdot(\sum_{k=m}^{n+m-1} \phi_{\sigma^k\omega}\circ T_\omega^k)^T] - \mathbb{E}^\omega[(\sum_{k=m}^{n+m-1} \phi_{\sigma^k\omega}\circ T_\omega^k)\cdot(\sum_{k=m}^{n+m-1} \phi_{\sigma^k\omega}\circ T_\omega^k)^T]|$$

$$\leq \mathbb{E}^\omega|\mathbb{E}_{n+m}^\omega[\sum_{m\leq k\leq n+m-1} \phi_{\sigma^k\omega}\circ T_\omega^k\cdot\phi_{\sigma^k\omega}^T\circ T_\omega^k - \mathbb{E}^\omega(\phi_{\sigma^k\omega}\circ T_\omega^k\cdot\phi_{\sigma^k\omega}^T\circ T_\omega^k)]| \tag{5.9}$$

$$+2\mathbb{E}^\omega|\mathbb{E}_{n+m}^\omega[\sum_{m\leq i<j\leq n+m-1} \phi_{\sigma^i\omega}\circ T_\omega^i\cdot\phi_{\sigma^j\omega}^T\circ T_\omega^j - \mathbb{E}^\omega(\phi_{\sigma^i\omega}\circ T_\omega^i\cdot\phi_{\sigma^j\omega}^T\circ T_\omega^j)]|. \tag{5.10}$$

By (A2'), (5.9) becomes

$$\mathbb{E}^\omega|\mathbb{E}_{n+m}^\omega[\sum_{m\leq k\leq n+m-1} \phi_{\sigma^k\omega}\circ T_\omega^k\cdot\phi_{\sigma^k\omega}^T\circ T_\omega^k - \mathbb{E}^\omega(\phi_{\sigma^k\omega}\circ T_\omega^k\cdot\phi_{\sigma^k\omega}^T\circ T_\omega^k)]|$$

$$\precsim \sum_{k=m}^{n+m-1} \int |P_{\sigma^k\omega}^{n+m-k}\{[\phi_{\sigma^k\omega}\cdot\phi_{\sigma^k\omega}^T - \mathbb{E}^\omega(\phi_{\sigma^k\omega}\circ T_\omega^k\cdot\phi_{\sigma^k\omega}^T\circ T_\omega^k)]\cdot h_{\sigma^k\omega}\}|d\mu$$

$$\precsim \sum_{k=m}^{n+m-1} \frac{1}{(m+n-k)^{\frac{1}{\alpha}-1}} = O(1).$$

To estimate (5.10), for any fixed $j \leq n+m-1$, by (A1'):

$$\mathbb{E}^\omega|\mathbb{E}_j^\omega[\sum_{m\leq i<j} \phi_{\sigma^i\omega}\circ T_\omega^i\cdot\phi_{\sigma^j\omega}^T\circ T_\omega^j - \mathbb{E}^\omega(\phi_{\sigma^i\omega}\circ T_\omega^i\cdot\phi_{\sigma^j\omega}^T\circ T_\omega^j)]| \precsim \sum_{m\leq i<j} \mathbb{E}^\omega|\mathbb{E}_j^\omega(\phi_{\sigma^i\omega}\circ T_\omega^i\cdot\phi_{\sigma^j\omega}^T\circ T_\omega^j)|$$

$$\precsim \sum_{m\leq i<j} \int |P_{\sigma^i\omega}^{j-i}(\phi_{\sigma^i\omega}\cdot h_{\sigma^i\omega})|d\mu \precsim \sum_{m\leq i<j} \frac{1}{(j-i)^{\frac{1}{\alpha}-1}} = O(1).$$

Let $\delta = \frac{\alpha}{1-\alpha}$, by (A3') and the above inequality:

$$(5.10) \precsim \sum_{n+m-\lfloor n^\delta\rfloor<j\leq n+m-1} \sum_{m\leq i<j} \mathbb{E}^\omega|\mathbb{E}_j^\omega[\phi_{\sigma^i\omega}\circ T_\omega^i\cdot\phi_{\sigma^j\omega}^T\circ T_\omega^j - \mathbb{E}^\omega(\phi_{\sigma^i\omega}\circ T_\omega^i\cdot\phi_{\sigma^j\omega}^T\circ T_\omega^j)]|$$

$$+\sum_{m<j\leq n+m-\lfloor n^\delta\rfloor}\sum_{m\leq i<j}\mathbb{E}^\omega|\mathbb{E}^\omega_{n+m}[\phi_{\sigma^i\omega}\circ T^i_\omega\cdot\phi^T_{\sigma^j\omega}\circ T^j_\omega-\mathbb{E}^\omega(\phi_{\sigma^i\omega}\circ T^i_\omega\cdot\phi^T_{\sigma^j\omega}\circ T^j_\omega)]|$$

$$\precsim\lfloor n^\delta\rfloor+\sum_{j=m+1}^{n+m-\lfloor n^\delta\rfloor}\sum_{m\leq i<j}\int|P^{n+m-j}_{\sigma^j\omega}[P^{j-i}_{\sigma^i\omega}(\phi_{\sigma^i\omega}\cdot h_{\sigma^i\omega})\cdot\phi^T_{\sigma^j\omega}-h_{\sigma^j\omega}\cdot\mathbb{E}^\omega(\phi_{\sigma^i\omega}\circ T^i_\omega\cdot\phi^T_{\sigma^j\omega}\circ T^j_\omega)]|d\mu$$

$$\precsim\lfloor n^\delta\rfloor+\sum_{j=m+1}^{n+m-\lfloor n^\delta\rfloor}\sum_{m\leq i<j}\frac{1}{(n+m-j)^{\frac{1}{\alpha}-1}}\precsim n^{\frac{\alpha}{1-\alpha}},$$

where all constants indicated in $\precsim$, $O(\cdot)$ are uniform over all ϕ_ω, m, n.

The proof of the VASIP w.r.t. μ_ω only relies on the random version of Lemma 4.1-4.6, which were proved above. So, by Theorem 2.8, the VASIP holds for a.e. $\omega\in\Omega$ under the condition of variance growth.

Next we claim the VASIP w.r.t. μ_ω or the coboundary based on variance growth: the proof is exactly the same as Lemma 12 in [14] except the following:

1. the last inequality of page 2270 in [14] becomes:

$$\leq\bar{K}\cdot\sum_{i\geq 1}\frac{1}{i^{\frac{1}{\alpha}-1}}<\infty.$$

2. the inequality in the middle of page 2271 becomes:

$$\leq\bar{K}\sum_{i\leq n-1}\sum_{k\geq n-i}\frac{1}{k^{\frac{1}{\alpha}-1}}=\sum_{k\leq n-1}\frac{k}{k^{\frac{1}{\alpha}-1}}+n\sum_{k\geq n}\frac{1}{k^{\frac{1}{\alpha}-1}}\precsim n^{3-\frac{1}{\alpha}}\int_{\frac{1}{n}}^1\frac{1}{x^{\frac{1}{\alpha}-2}}dx+n\sum_{k\geq n}\frac{1}{k^{\frac{1}{\alpha}-1}}.$$

Then (35) in [14] becomes $\leq\frac{1}{n}\cdot(n^{3-\frac{1}{\alpha}}+n\sum_{k\geq n}\frac{1}{k^{\frac{1}{\alpha}-1}})\to 0$.

Therefore there is a $d\times d$ positive semi-definite matrix $\sigma^2\geq 0$ such that almost surely,

$$\lim_{n\to\infty}\frac{1}{n}\int(\sum_{0\leq k<n}\phi_{\sigma^k(\omega)}\circ T^k_\omega)\cdot(\sum_{0\leq k<n}\phi_{\sigma^k(\omega)}\circ T^k_\omega)^T d\mu_\omega=\sigma^2.$$

If $\sigma^2>0$, then variance grows linearly for a.e. $\omega\in\Omega$. By Theorem 2.8, the VASIP w.r.t. μ_ω holds for a.e. $\omega\in\Omega$.

If $\det(\sigma^2) = 0$, without loss of generality, assume $\sigma^2 = \begin{bmatrix} I_{d_1 \times d_1} & 0 \\ 0 & \mathbf{0}_{d_2 \times d_2} \end{bmatrix}_{d \times d}$.

If $d_1 = 0$, we claim it has the coboundary:

Without loss of generality, assume all $(\phi_\omega)_{\omega \in \Omega}$ are scalar functions, denote $\bar{\phi}(\omega, x) := \phi_\omega(x)$, similar to the computation of Lemma 12 (36) in [14], we have:

$$0 = \sigma^2 = \int \bar{\phi}^2(\omega, x) d\mu_\omega d\mathbb{P} + 2 \sum_{i \geq 1} \int \bar{\phi}(\omega, x) \cdot \bar{\phi} \circ \tau^i(\omega, x) d\mu_\omega d\mathbb{P},$$

where $\tau(\omega, x) := (\sigma\omega, f_\omega(x))$.

For the stationary dynamical system $(\Omega \times X, \tau, d\mu_\omega d\mathbb{P})$ with observable $\bar{\phi} \in L^\infty(\Omega \times X)$, denote the transfer operator of τ by τ^*. We will verify conditions (1) and (2) of Theorem 1.1 in [31]: by (A1'),

$$\sum_{n \geq 0} \left| \int \bar{\phi} \cdot \bar{\phi} \circ \tau^n d\mu_\omega d\mathbb{P} \right| \precsim \sum_{n \geq 0} \int |P_\omega^n(\phi_\omega h_\omega)| d\mu d\mathbb{P} \precsim \sum_{n \geq 1} \frac{1}{n^{\frac{1}{\alpha} - 1}} < \infty,$$

$$\sum_{n \geq 0} \int |\tau^{*n} \bar{\phi}| d\mu_\omega d\mathbb{P} = \sum_{n \geq 0} \sup_{||\xi||_{L^\infty(\Omega \times X)} \leq 1} \int \xi \circ \tau^n \cdot \bar{\phi} d\mu_\omega d\mathbb{P}$$

$$\precsim \sum_{n \geq 0} \int |P_\omega^n(\phi_\omega h_\omega)| d\mu d\mathbb{P} \precsim \sum_{n \geq 1} \frac{1}{n^{\frac{1}{\alpha} - 1}} < \infty.$$

Therefore, by Theorem 1.1 of [31], 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 $d_1 > 0, d_2 > 0$, we will follow the argument of [19]: $\mathbb{R}^d = \mathbb{R}^{d_1} \bigoplus \mathbb{R}^{d_2}$ with projections $\pi_1 : \mathbb{R}^{d_1} \bigoplus \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}, \pi_2 : \mathbb{R}^{d_1} \bigoplus \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$. The dynamical system $(\pi_1 \circ \phi_{\sigma^k(\omega)} \circ T_\omega^k)_{k \geq 1}$ has the VASIP w.r.t. μ_ω by the argument of "$\sigma^2 > 0$" above. For the dynamical system $(\pi_2 \circ \phi_{\sigma^k(\omega)} \circ T_\omega^k)_{k \geq 1}$, 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:

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

∎

6 Proofs of corollaries

6.1 Proof of Corollary 3.1

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

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

Lemma 6.1 (see also [1, 32, 39])

Assume $K > 0$, $\phi_i \in \mathrm{Lip}[0,1]$ and $h_k \in C_{a_0}$ with $||\phi_i||_{\mathrm{Lip}} \leq K, ||h_k||_{L^1} \leq M$ for all $i, k \geq 1$. Then for a sufficiently large $a_0 > 1$ independent of K, M, there are constants λ, v, δ (only depends on K, M, α, a_0) such that the following holds:

$$h^1_{i,k} := (\phi_i + \lambda \cdot X + v)h_k + \delta, h^2_{i,k} := (\lambda \cdot X + v)h_k + \delta + \int \phi_i \cdot h_k dm \in C_{a_0}, \tag{6.1}$$

$$\phi_i \cdot h_k - \int \phi_i \cdot h_k dm = h^1_{i,k} - h^2_{i,k} \in C_{a_0} - C_{a_0},$$

$$\int h^1_{i,k} dm = \int h^2_{i,k} dm,$$

$\mathbf{1} \in C_{a_0}, C_{a_0}$ is preserved by all T_k's transfer operators P_k.

Furthermore, there are constants $C_{K,M,\alpha,a_0}, C_{\alpha,a_0}$ such that for all $m, n \in \mathbb{N}$, $h \in C_{a_0}$:

$$||P^{n+m}_{m+1}(\phi_k \cdot h_k - \int \phi_k \cdot h_k dm)||_{L^1} \precsim C_{K,M,\alpha,a_0} \cdot \frac{1}{n^{\frac{1}{\alpha}-1}}, \tag{6.2}$$

$$||P^{n+m}_{m+1}(h - \int h dm)||_{L^1} \precsim C_{\alpha,a_0} \cdot ||h||_{L^1} \cdot \frac{1}{n^{\frac{1}{\alpha}-1}}. \tag{6.3}$$

Proof *[1, 39] proved these properties for the cone $C_{a_0} \cap C^1(0,1]$. However, since the C^1 properties are not used in their proofs, so the decay of correlation (6.3) still holds for our C_{a_0}, and C_{a_0} is still an P_k-invariant cone. To prove (6.1), the argument is replacing $|\phi_k'|_\infty$ with its Lipschitz constant* $\mathrm{Lip}(\phi_k)$ *in Lemma 2.4 in [39]. Then (6.2) holds by applying (6.3) and (6.1).*

∎

With this lemma, we can prove our corollary for $(\phi_k)_{k\in\mathbb{N}} \subset \mathrm{Lip}[0,1]$ now:

Since $\sup_i ||\phi_i||_{\mathrm{Lip}} < \infty, \sup_i ||\phi_i \cdot \phi_i^T||_{\mathrm{Lip}} \leq 2\sup_i ||\phi_i||^2_{\mathrm{Lip}} < \infty, ||P^k\mathbf{1}||_{L^1} = 1$, so (A1) and (A2) are easily verified by (6.2). Now we verify (A3):

$$\int |P_{i+j+1}^{i+j+n}\{[P_{i+1}^{i+j}(\phi_i \cdot P^i\mathbf{1}) \cdot \phi_{i+j}^T - \int P_{i+1}^{i+j}(\phi_i \cdot P^i\mathbf{1}) \cdot \phi_{i+j}^T d\mu]\}| d\mu \precsim \frac{1}{n^{\frac{1}{\alpha}-1}}.$$

For the fixed i, j above, by (6.1), there are $h_1, h_1', h_1''', h_1'''', h_2, h_2', h_2''', h_2'''' \in C_{a_0}$ and the following decompositions:

$h_1 - h_2 = \phi_i \cdot P^i\mathbf{1} \in C_{a_0} - C_{a_0}, h_1' := P_{i+1}^{i+j}h_1 \in C_{a_0}, h_2' := P_{i+1}^{i+j}h_2 \in C_{a_0}, h_1''' - h_2''' = h_1' \cdot \phi_{i+j}^T - \int h_1' \cdot \phi_{i+j}^T dm \in C_{a_0} - C_{a_0}, h_1'''' - h_2'''' = h_2' \cdot \phi_{i+j}^T - \int h_2' \cdot \phi_{i+j}^T dm \in C_{a_0} - C_{a_0}$.

So $\int P_{i+1}^{i+j}(\phi_i \cdot P^i\mathbf{1}) \cdot \phi_{i+j}^T d\mu))dm = \int h_1' \cdot \phi_{i+j}^T dm - \int h_2' \cdot \phi_{i+j}^T dm$.

By (6.1), (6.3),

$$\int |P_{i+j+1}^{i+j+n}\{[P_{i+1}^{i+j}(\phi_i \cdot P^i\mathbf{1}) \cdot \phi_{i+j}^T - \int P_{i+1}^{i+j}(\phi_i \cdot P^i\mathbf{1}) \cdot \phi_{i+j}^T d\mu]\}| dm$$

$$= \int |P_{i+j+1}^{i+j+n}[h_1''' - h_2''' - (h_1'''' - h_2'''')]| dm \precsim C_{a_0,\alpha}(||h_1'''||_{L^1} + ||h_1''''||_{L^1})\frac{1}{n^{\frac{1}{\alpha}-1}}$$

$$\precsim C_{a_0,\alpha} \cdot C_{\sup_k ||\phi_k||_{\mathrm{Lip}}, ||h_1'||_{L^1}, ||h_2'||_{L^1}} \cdot \frac{1}{n^{\frac{1}{\alpha}-1}} = C_{a_0,\alpha} \cdot C_{\sup_k ||\phi_k||_{\mathrm{Lip}}, ||h_1||_{L^1}, ||h_2||_{L^1}} \cdot \frac{1}{n^{\frac{1}{\alpha}-1}}.$$

By (6.1), $||h_1||_{L^1}, ||h_2||_{L^1}$ are bounded by a constant $C_{\sup_k ||\phi_k||_{\mathrm{Lip}}}$. Therefore,

$$\int |P_{i+j+1}^{i+j+n}\{[P_{i+1}^{i+j}(\phi_i \cdot P^i\mathbf{1}) \cdot \phi_{i+j}^T - \int P_{i+1}^{i+j}(\phi_i \cdot P^i\mathbf{1}) \cdot \phi_{i+j}^T d\mu]\}| dm \precsim C_{\sup_k ||\phi_k||_{\mathrm{Lip}}, \alpha, a_0} \cdot \frac{1}{n^{\frac{1}{\alpha}-1}}.$$

Therefore the VASIP holds for this non-stationary dynamical system.

For scalar self-norming CLT, we will give a similar but simpler proof than VASIP's:

Let $(\phi_k)_{k\in\mathbb{N}} \subset \mathrm{Lip}([0,1];\mathbb{R})$, $I_n = [1, n]$. Let $a \in (\frac{1}{2}, 1), c_n := \lfloor n^{(1-a)} \rfloor$ (this a is different from the one in Lemma 6.1). 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 $J_{n,c_n+1} := I_n \setminus \bigcup_{1\leq i\leq c_n} I_{n,i}$ contains the largest number of I_n. So $|I_{n,c_n+1}| \leq 2\lfloor n^a \rfloor$ and $\bigcup_{1\leq i\leq c_n+1} I_{n,i} = I_n$. Similar to the proof of Theorem 2.8, let $X_n := \sum_{i\leq n} \phi_i \circ T^i$, $b_n := \lambda(\sigma_n^2) \gtrsim n^{\gamma_1}$, fix any $u \in \mathbb{R}$:

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

With the same estimates as in (5.1), the above inequality becomes

$$\sum_{0\leq k\leq c_n} \mathbb{E}|\mathbb{E}\{[1 + i\frac{u \cdot X_{n,k+1}}{\sqrt{b_n}} - \frac{1}{2}(\frac{u \cdot X_{n,k+1}}{\sqrt{b_n}})^2 + O(|\frac{u \cdot X_{n,k+1}}{\sqrt{b_n}}|^{2+\epsilon_0})]|\mathcal{F}_{n,k+2}\}$$

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

$$\leq \sum_{0\leq k\leq c_n} \{\frac{|u|}{\sqrt{b_n}} + \frac{1}{2}\frac{|u|^2}{b_n} \cdot \mathbb{E}|\mathbb{E}\{[X_{n,k+1}^2 - \mathbb{E}(X_{n,k+1}^2)]|\mathcal{F}_{n,k+2}\}|$$

$$+|u|^4 \cdot \frac{|\mathbb{E}(X_{n,k+1}^2)|^2}{b_n^2} + |u|^{2+\epsilon_0} \cdot \frac{\mathbb{E}(|X_{n,k+1}|^{2+\epsilon_0})}{b_n^{\frac{2+\epsilon_0}{2}}}\} + |u|^2 \cdot \frac{O(n^{c(1-a)+c\max(3-\frac{1}{\alpha},0)})}{b_n}$$

$$\lesssim \frac{n^{(1-a)}}{n^{\frac{\gamma_1}{2}}}|u| + \frac{n^{(1-a)+a\frac{\alpha}{1-\alpha}}}{n^{\gamma_1}}|u|^2 + \frac{n^{(1-a)+2a}}{n^{2\gamma_1}}|u|^4 + \frac{n^{(1-a)+a\frac{2+\epsilon_0}{2}}}{n^{\frac{\gamma_1(2+\epsilon_0)}{2}}}|u|^{2+\epsilon_0} + \frac{n^{(1-a)+\max(3-\frac{1}{\alpha},0)}}{n^{\gamma_1}}|u|^2.$$

To let it go to zero, we need the following conditions:

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

So when γ_1 is any number in $(\frac{2+a\epsilon_0}{2+\epsilon_0}, 1]$, self-norming CLT holds, where

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

For the computation of γ_1, see Appendix, Lemma 7.2. ∎

Proof of Corollary 3.3

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

$$||P_m^{n+m}(f - \int f d\mu)||_{\mathcal{V}} \precsim \lambda^n ||f - \int f d\mu||_{\mathcal{V}},$$

$$\sup_i ||P^i \mathbf{1}||_{\mathcal{V}} < \infty.$$

Note that $\mathcal{V}$ is a Banach algebra, so (A1)-(A3) are all satisfied. By Theorem 2.8, Corollary 3.3 holds.

∎

6.2 Proof of Corollary 3.5

Proof For the existence and uniqueness of the quasi-invariant probability, the proofs are given in [14]. (A1')-(A3') can be verified similar to Corollary 3.3. So by Theorem 2.9, we have the desired result with $\psi \in L^1(\Omega \times X, d\mu_\omega d\mathbb{P})$. To prove $\psi \in L^2(\Omega \times X, d\mu_\omega d\mathbb{P})$, it is exactly the same as Proposition 3 in [14].

∎

6.3 Proof of Corollary 3.7

Proof The existence of the quasi-invariant probability is constructed similar to [14]: Consider the Banach space

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

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 $||\mathcal{L}v|| \le ||v||$. Consider $(\mathcal{L}^n \mathbf{1})_{n \ge 1}$. We claim this is a Cauchy sequence:

By Lemma 6.1, since $P_\omega \mathbf{1} \in C_{a_0}$ for any $\omega \in \Omega$, then for any $n < m$,

$$||\mathcal{L}^n \mathbf{1} - \mathcal{L}^m \mathbf{1}|| = \sup_\omega ||P^n_{\sigma^{-n}\omega}(\mathbf{1} - P^{m-n}_{\sigma^{-m}\omega}\mathbf{1})||_{L^1} \le K \cdot \frac{1}{n^{\frac{1}{\alpha}-1}}.$$

Then there is $h \in Y$ such that $\mathcal{L}h = h$, that is, $P_{\sigma^{-1}\omega} h_{\sigma^{-1}\omega} = h_\omega$ for a.e.-$\omega \in \Omega$. So h_ω satisfies all conditions of C_{a_0} except its regularity. To prove $h_\omega \in \mathrm{Lip}_{loc}(0,1]$, the method is the same as Lemma 2.3 in [32]. 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 (A1')-(A3') is the same as Corollary 3.1. By Theorem 2.9, this corollary holds.

∎

6.4 Proof of Corollary 3.8

Proof First, we will show: there is a $d \times d$ positive semi-definite matrix $\sigma^2 \ge 0$ and $\epsilon \in (0,1)$ such that

$$\mathbb{E}[(\sum_{i \le n} \phi \circ T^i) \cdot (\sum_{i \le n} \phi \circ T^i)^T] = n \cdot \sigma^2 + o(n^{1-\epsilon}). \tag{6.4}$$

Note that, by (A4),

$$\sum_{i \ge 1} \mathbb{E}(\phi \cdot \phi^T \circ T^i) \precsim \sum_{i \ge 1} \frac{1}{i^{\frac{1}{\alpha}-1}} < \infty \text{ absolutely converges.}$$

Let $\sigma^2 := \mathbb{E}(\phi \cdot \phi^T) + \sum_{i \ge 1} \mathbb{E}(\phi \cdot \phi^T \circ T^i) + \sum_{i \ge 1} \mathbb{E}(\phi \cdot \phi^T \circ T^i)^T$, then

$$\mathbb{E}[(\sum_{i \le n} \phi \circ T^i) \cdot (\sum_{i \le n} \phi \circ T^i)^T] - n \cdot \sigma^2$$

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

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

$$= n \cdot \mathbb{E}(\phi \cdot \phi^T) + \sum_{1 \le i \le n} \sum_{0 < j \le n-i} \mathbb{E}(\phi \cdot \phi^T \circ T^j) + \sum_{1 \le i \le n} \sum_{0 < j \le n-i} \mathbb{E}(\phi \cdot \phi^T \circ T^j)^T - n \cdot \sigma^2$$

$$= \sum_{1\le i\le n}\sum_{0<j\le n-i}\mathbb{E}(\phi\cdot\phi^T\circ T^j)+\sum_{1\le i\le n}\sum_{0<j\le n-i}\mathbb{E}(\phi\cdot\phi^T\circ T^j)^T-n\cdot\sum_{i\ge 1}\mathbb{E}(\phi\cdot\phi^T\circ T^i)-n\cdot\sum_{i\ge 1}\mathbb{E}(\phi\cdot\phi^T\circ T^i)^T.$$

Then we just need to estimate:

$$\sum_{1\le i\le n}\sum_{0<j\le n-i}\mathbb{E}(\phi\cdot\phi^T\circ T^j)-n\cdot\sum_{i\ge 1}\mathbb{E}(\phi\cdot\phi^T\circ T^i)$$

$$=\sum_{1\le i\le n}\sum_{0<j\le n-i}\mathbb{E}(\phi\cdot\phi^T\circ T^j)-\sum_{1\le i\le n}\sum_{j\ge 1}\mathbb{E}(\phi\cdot\phi^T\circ T^j)=\sum_{1\le i\le n}\sum_{j>n-i}\mathbb{E}(\phi\cdot\phi^T\circ T^j)$$

$$=\sum_{1\le i\le n}\sum_{n-i<j\le n}\mathbb{E}(\phi\cdot\phi^T\circ T^j)+n\cdot\sum_{j>n}\mathbb{E}(\phi\cdot\phi^T\circ T^j)\precsim\sum_{1\le i\le n}\sum_{n-i<j\le n}\frac{1}{j^{\frac{1}{\alpha}-1}}+n\cdot\sum_{j>n}\frac{1}{j^{\frac{1}{\alpha}-1}}$$

$$\precsim\sum_{i\le n} i\cdot\frac{1}{i^{\frac{1}{\alpha}-1}}+n\cdot\int_n^\infty\frac{1}{x^{\frac{1}{\alpha}-1}}dx=n^{3-\frac{1}{\alpha}}+n\cdot n^{2-\frac{1}{\alpha}}\precsim n^{3-\frac{1}{\alpha}}.$$

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

$$\sigma_n^2-n\cdot\sigma^2=\mathbb{E}[(\sum_{i\le n}\phi\circ T^i)\cdot(\sum_{i\le n}\phi\circ T^i)^T]-n\cdot\sigma^2\precsim n^{3-\frac{1}{\alpha}}=o(n^{1-\epsilon}).$$

If $\det(\sigma^2)>0$, then $\sigma_n^2\succsim n$. So, by Theorem 2.8, the VASIP holds if (A5),(A6) are satisfied as well; moreover, by Lemma 7.4, the Gaussian vectors are i.i.d. with covariance σ^2.

If $\det(\sigma^2)=0$, without loss of generality, assume $\sigma^2=\begin{bmatrix} I_{d_1\times d_1} & 0 \\ 0 & \mathbf{0}_{d_2\times d_2}\end{bmatrix}_{d\times d}$.

The argument in this case is exactly same as in Theorem 2.9, we will not repeat it here.

To prove the VASIP for the Young tower Δ, Young [51] proved the first order decay of correlation (A4) and (A5) already, so we just need to verify the second order decay of correlation (A6):

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

where $dv=\frac{dv}{dm}dm$, both $\frac{dv}{dm}$ and ϕ are in $L^\infty(\Delta)\cap C_\beta(\Delta),\inf\frac{dv}{dm}>0$ where $C_\beta(\Delta)$ is the same as in [51]. However, we just need to show $P^j\phi$ is also Lipschitz function with uniform Lipschitz exponent independent of j, then (A6) holds by using (A4):

Without loss of generality, assume ϕ is scalar function with Lipschitz exponent C_ϕ: for any

$(a, m_a) \in F^{-j}\Delta_{m,i}$, the orbit $\{F^0(a, m_a), \cdots, F^j(a, m_a)\}$ touches Δ_0 for q_a times $(0 \le q_a \le j)$, $a \in \Delta_{0,i_{0,a}} \cap (F^R)^{-1}\Delta_{0,i_{1,a}} \cap \cdots \cap (F^R)^{-(q_a-1)}\Delta_{0,i_{q_a-1,a}}$. Denote $P_a := ((F^R)^{-q_a}\Delta_{0,i} \cap \Delta_{0,i_{0,a}} \cap (F^R)^{-1}\Delta_{0,i_{1,a}} \cap \cdots \cap (F^R)^{-(q_a-1)}\Delta_{0,i_{q_a-1,a}}) \times m_a$. Therefore $F^j(P_a) = \Delta_{m,i}$. For different P_a, they are either exactly the same, or disjoint.

- $P^j\phi$ is locally Lipschitz:

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

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

$$P^j(\phi)(x_2, m) = \frac{1}{\frac{dv}{dm}(x_2, m)} \cdot \sum_{F^j(y)=(x_2,m)} \frac{\phi(y)\frac{dv}{dm}(y)}{JF^j(y)} = \frac{1}{\frac{dv}{dm}(x_2, m)} \cdot \sum_a \frac{\phi(y_a^2)\frac{dv}{dm}(y_a^2)}{JF^j(y_a^2)}.$$

$$|P^j(\phi)(x_1, m) - P^j(\phi)(x_2, m)| \le \frac{1}{\frac{dv}{dm}(x_1, m)} |\sum_a \frac{\phi(y_a^1)\frac{dv}{dm}(y_a^1)}{JF^j(y_a^1)} - \sum_a \frac{\phi(y_a^2)\frac{dv}{dm}(y_a^2)}{JF^j(y_a^2)}|$$

$$+|\frac{1}{\frac{dv}{dm}(x_1, m)} - \frac{1}{\frac{dv}{dm}(x_2, m)}| \cdot |\sum_a \frac{\phi(y_a^2)\frac{dv}{dm}(y_a^2)}{JF^j(y_a^2)}| \precsim |\sum_a \frac{\phi(y_a^1)\frac{dv}{dm}(y_a^1) - \phi(y_a^2)\frac{dv}{dm}(y_a^2)}{JF^j(y_a^1)}|$$

$$+|\sum_a \frac{\phi(y_a^2)\frac{dv}{dm}(y_a^2)}{JF^j(y_a^1)}(1 - \frac{JF^j(y_a^1)}{JF^j(y_a^2)})| + |\sum_a \frac{\phi(y_a^2)\frac{dv}{dm}(y_a^2)}{JF^j(y_a^2)}| \cdot \beta^{s(x_1,x_2)}$$

  where $y_a^1, y_a^2 \in P_a$. Use 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 above inequality becomes

$$\precsim C_\phi \cdot C_v \cdot \sum_a \frac{1}{JF^j(y_a^1)} \cdot d(y_a^1, y_a^2) + C_\phi \cdot C_v \cdot \sum_a \frac{1}{JF^j(y_a^2)} \cdot \beta^{s(x_1,x_2)}$$

$$\precsim \sum_a C_v \cdot C_\phi \cdot \frac{m(P_a)}{m(\Delta_0)} \cdot d(z_1, z_2) \precsim C_v \cdot C_\phi \cdot \frac{m(\Delta)}{m(\Delta_0)} \cdot d(z_1, z_2).$$

- $P^j\phi$ is bounded:

$$|P^j(\phi)(x_1, m)| \le P^j(\mathbf{1})(x_1, m) \cdot ||\phi||_{L^\infty} = ||\phi||_{L^\infty}.$$

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

$$|P^j(\phi)(z_1) - P^j(\phi)(z_2)| \precsim 2||\phi||_{L^\infty} \cdot C_v \cdot C_\phi \cdot d(z_1, z_2)$$

where its Lipschitz exponent, as shown above, is independent of j. ∎

7 Appendix

Lemma 7.1 (Computation of the range of γ)

The parameter γ in Theorem 2.8 can be any number in $(\frac{c}{c+1}+\frac{2}{(c+1)(2+\epsilon_0)},1]$, *where*

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

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

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

d,α *are the parameters in Theorem* 2.8.

Proof To find the range of γ, we summarize (5.2)-(5.7) 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. $\frac{1}{2}\gamma(c+1)(2+\epsilon)-c(1+\frac{\epsilon}{2})>1,\epsilon<\min(1,2-\frac{2\alpha}{1-\alpha})$.

If $v>2(d+1)\kappa$, then $\min(\kappa,\frac{v}{2}-\kappa d)=\kappa$. So we can use this to simplify 1,2 above as 1,2 below. Note that 3 and 5 above are the same, the inequality 6 above is $\gamma>\frac{c}{c+1}+\frac{2}{(c+1)(2+\epsilon)}$ which implies 3 and 5 above: $\gamma>\frac{c}{c+1}$. So we can combine 3,5,6 above as 4 below. Therefore the above inequalities can be rewritten as:

1. $v>2(d+1)\kappa$, $\kappa>1$.

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

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

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

Use the definition of v to expand $v > 2(d+1)\kappa$ as 1,2,3,4,5 below, and copy 2,3,4 above as 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. $\gamma > \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}, \epsilon < \min(1, 2 - \frac{2\alpha}{1-\alpha})$.

We transform the above inequalities to represent the range of γ:

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

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

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

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

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

Use $\frac{c}{c+1} < 1, \frac{2}{2+\epsilon_0} < 1, a < 1, \epsilon_0 < 1$, we can narrow the range of γ 1-5 above as 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. $\gamma > \frac{(2d+5)}{(1+c)}\kappa + \frac{2+a\epsilon_0}{(2+\epsilon_0)}$.

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 $\kappa := 2$, use $2d+5 > d+3, \frac{a+1}{2} < \frac{2+\epsilon_0 a}{2+\epsilon_0}$, we can combine inequalities 3,4 above as 3 below. Use $1 - \frac{2}{c+1} < \frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)}$, we can combine the inequalities 6,8 above as 6 below:

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

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

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

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

If $a > \frac{\epsilon_0 + 2\alpha}{(1-\alpha)(2\epsilon_0+2)}$, then $\frac{2+a\epsilon_0}{(2+\epsilon_0)} > \frac{\alpha}{1-\alpha} + 1 - a$.
Note that if $a > \frac{2+2\epsilon_0}{3\epsilon_0+4}$, then $\frac{2+a\epsilon_0}{(2+\epsilon_0)} > 2(1-a)$.
If $a > \frac{(2+\epsilon_0)(1+\max(3-\frac{1}{\alpha},0))-2}{2+2\epsilon_0}$, then $\frac{2+a\epsilon_0}{(2+\epsilon_0)} > 1 - a + \max(0, 3 - \frac{1}{\alpha})$.

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

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

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

we can combine inequalities 1,2,3,4 above as 1 below and copy 5,6 above as 2,3 below:

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

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

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

If $c > \frac{2+\epsilon_0 a+(2+\epsilon_0)(8d+12)-\frac{2(2+\epsilon_0)}{2+\epsilon}}{\epsilon_0(1-a)}$, then $\frac{c}{c+1} + \frac{2}{(c+1)(2+\epsilon)} > \frac{2(4d+6)}{(1+c)} + \frac{2+a\epsilon_0}{(2+\epsilon_0)}$.

Therefore, if

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

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

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

the inequalities 1,2,3 above can be combined as the following one inequality:

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

Let $\epsilon_0 = \epsilon$, then

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

Since this is a strict inequality for γ, so c, a, ϵ_0 can take their infimums/supremum, respectively, that is,

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

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

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

∎

Lemma 7.2 (Computation of the range of γ_1)

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

Then γ_1 can be any number in $(\frac{2+a\epsilon_0}{2+\epsilon_0}, 1]$, where

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

Proof Since $\frac{1+a}{2} < \frac{2+a\epsilon_0}{2+\epsilon_0}$, then $\gamma_1 > (1 + a\frac{\epsilon_0}{2}) \cdot \frac{2}{2+\epsilon_0} > \frac{1+a}{2}$.

If $a > \frac{2+2\epsilon_0}{4+5\epsilon_0}$, then $(1 + a\frac{\epsilon_0}{2}) \cdot \frac{2}{2+\epsilon_0} > 2(1-a)$.

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

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

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

$$\gamma_1 > \frac{2+a\epsilon_0}{2+\epsilon_0}.$$

Since this is a strict inequality for γ_1, so a, ϵ_0 can take their infimum/supremum, respectively, that is, γ_1 can be any number in $(\frac{2+a\epsilon_0}{2+\epsilon_0}, 1]$, where

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

∎

Lemma 7.3 (Transfer, see [25] Theorem 6.10)

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

$$(\eta, \xi) \overset{d}{=} (\eta', \xi').$$

More precisely, there exists a measurable function $f : S \times [0,1] \to T$ such that we may take $\eta' = f(\xi', U)$ where $U \sim U(0,1)$ is independent of ξ and f.

Indeed, to guarantee the independence, we may simply extend the probability space where ξ' lives in, by multiplying an interval (I, Leb).

Lemma 7.4 (Embedding in a d-dimensional Brownian motion)

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

$$\sum_{k\leq n} \phi_k \circ T^k - \sigma B_n = o(n^{\frac{1}{2}-\bar{\epsilon}}) \text{ a.s.}.$$

Proof Since $d = 1$ is trivial, we assume $d > 1$. By the VASIP Definition 2.4, we have:

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

$$\sigma_n^2 = \int (\sum_{k\leq n} \phi_k \circ T^k) \cdot (\sum_{k\leq n} \phi_k \circ T^k)^T d\mu = \sum_{k\leq n} \tilde{\mathbb{E}}(G_k \cdot G_k^T) + o(n^{1-\epsilon}),$$

where $\tilde{\mathbb{E}}(\cdot)$ is the expectation of the probability $\tilde{\mathbb{P}}$ of the extended probability space $(X, \mathcal{B}, \mu)$. Then

$$\sum_{k \leq n} \tilde{\mathbb{E}}(G_k \cdot G_k^T) = n \cdot \sigma^2 + o(n^{1-\epsilon}).$$

Without loss of generality, we can assume $\sigma^2 = I_{d\times d}$.

Let $c \in \mathbb{N}$ (will be given later), then

$$\sum_{n^c < k \leq (1+n)^c} \tilde{\mathbb{E}}(G_k \cdot G_k^T) = [(1+n)^c - n^c] \cdot I_{d\times d} + o((n+1)^{c(1-\epsilon)}). \tag{7.1}$$

If c is big enough such that $c - 1 > c(1-\epsilon)$, then

$$\frac{\sum_{n^c < k \leq (1+n)^c} \tilde{\mathbb{E}}(G_k \cdot G_k^T)}{(1+n)^c - n^c} - I_{d\times d} = \frac{o((n+1)^{c(1-\epsilon)})}{n^{c-1}} = o(n^{1-c\epsilon}).$$

Denote

$$A := \sum_{n^c < k \leq (1+n)^c} \tilde{\mathbb{E}}(G_k \cdot G_k^T) = Q_n \cdot \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d^n \end{bmatrix} \cdot Q_n^T, \tag{7.2}$$

where $\lambda_1^n \leq \lambda_2^n \leq \cdots \leq \lambda_d^n$ are eigenvalues, Q_n is an orthogonal matrix. 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 $\bar{g}_1^{n+1}, \bar{g}_2^{n+1}, \bar{g}_3^{n+1}$ such that

$$\tilde{\mathbb{E}}[\bar{g}_1^{n+1} \cdot (\bar{g}_1^{n+1})^T] = A_1, \tilde{\mathbb{E}}[\bar{g}_2^{n+1} \cdot (\bar{g}_2^{n+1})^T] = A_2, \tilde{\mathbb{E}}[\bar{g}_3^{n+1} \cdot (\bar{g}_3^{n+1})^T] = A_3.$$

Therefore $\bar{g}_1^{n+1} + \bar{g}_2^{n+1} \overset{d}{=} \sum_{n^c < k \leq (1+n)^c} G_k$. By Lemma 7.3, there are independent mean zero 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 $\tilde{P}$ and its expectation by $\tilde{E}(\cdot)$) such that

$$(\bar{g}_1^{n+1} + \bar{g}_2^{n+1}, \bar{g}_1^{n+1}, \bar{g}_2^{n+1}, \bar{g}_3^{n+1}) \overset{d}{=} (\sum_{n^c < k \leq (1+n)^c} G_k, g_1^{n+1}, g_2^{n+1}, g_3^{n+1}).$$

Therefore,

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

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

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

Furthermore, since A_2 and A_3, after being diagonalized by Q_n, have nonzero entries on disjoint positions of diagonal line. Therefore by (7.1),

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

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

By Lemma 7.3, we know g_i^n depends on $\sum_{n^c < k \leq (1+n)^c} G_k$. Since for any $n_1 \neq n_2 \in \mathbb{N}$, $\sum_{n_1^c < k \leq (1+n_1)^c} G_k$ is independent of $\sum_{n_2^c < k \leq (1+n_2)^c} G_k$, then

$$(g_1^{n_1+1}, g_2^{n_1+1}, g_3^{n_1+1}) \text{ is independent of } (g_1^{n_2+1}, g_2^{n_2+1}, g_3^{n_2+1}).$$

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

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

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

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

$$\sum_{k\leq m} G_k-B_m=\sum_{k\leq n^c} G_k-B_{n^c}+\sum_{n^c<k\leq m} G_k-(B_m-B_{n^c})=\sum_{i\leq n}(g_2^i-g_3^i)+$$

$$\sum_{n^c<k\leq m} G_k-(B_m-B_{n^c})\leq|\sum_{i\leq n}(g_2^i-g_3^i)|+\sup_{n^c<m\leq(n+1)^c}|\sum_{n^c<k\leq m} G_k|+\sup_{n^c<m\leq(n+1)^c}|B_m-B_{n^c}|.$$

To estimate the last two terms, we just need to estimate each coordinate of them. So without loss of generality, we assume the last two terms are scalar Gaussian random elements. Then if $2c\bar{\epsilon}<1,\bar{\epsilon}<\epsilon$,

$$\tilde{P}(\sup_{n^c<m\leq(n+1)^c}|\sum_{n^c<k\leq m} G_k|>n^{c(\frac{1}{2}-\bar{\epsilon})})\leq\tilde{P}(|N(0,\tilde{\mathbb{E}}[(\sum_{n^c<k\leq(n+1)^c} G_k)^2]|>n^{c(\frac{1}{2}-\bar{\epsilon})})$$

$$=\tilde{P}(|N(0,1)|>\frac{n^{c(\frac{1}{2}-\bar{\epsilon})}}{\sqrt{\tilde{\mathbb{E}}[(\sum_{n^c<k\leq(n+1)^c} G_k)^2]}})\precsim\tilde{P}(|N(0,1)|>\frac{n^{c(\frac{1}{2}-\bar{\epsilon})}}{n^{\frac{c-1}{2}}})\leq e^{-n^{1-2c\bar{\epsilon}}}\precsim\frac{1}{n^2},$$

$$\tilde{P}(\sup_{n^c<m\leq(n+1)^c}|B_m-B_{n^c}|>n^{c(\frac{1}{2}-\bar{\epsilon})})\leq\tilde{P}(|N(0,(n+1)^c-n^c)|>n^{c(\frac{1}{2}-\bar{\epsilon})})$$

$$=\tilde{P}(|N(0,1)|>\frac{n^{c(\frac{1}{2}-\bar{\epsilon})}}{\sqrt{(n+1)^c-n^c}})\precsim\tilde{P}(|N(0,1)|>\frac{n^{c(\frac{1}{2}-\bar{\epsilon})}}{n^{\frac{c-1}{2}}})\leq e^{-n^{1-2c\bar{\epsilon}}}\precsim\frac{1}{n^2}.$$

The estimate of g_2^i and g_3^i are the same, so we just estimate g_2^i:

$$\tilde{P}(|\sum_{i\leq n} g_2^i|>n^{c(\frac{1}{2}-\bar{\epsilon})})=\tilde{P}(|N(0,\tilde{\mathbb{E}}[(\sum_{i\leq n} g_2^i)^2]])|>n^{c(\frac{1}{2}-\bar{\epsilon})})$$

$$=\tilde{P}(|N(0,1)|>\frac{n^{c(\frac{1}{2}-\bar{\epsilon})}}{\sqrt{\tilde{\mathbb{E}}[(\sum_{i\leq n} g_2^i)^2]}})\precsim\tilde{P}(|N(0,1)|>\frac{n^{c(\frac{1}{2}-\bar{\epsilon})}}{n^{\frac{1}{2}[1+c(1-\epsilon)]}})\precsim e^{-n^{c(\epsilon-2\bar{\epsilon})-1}}\precsim\frac{1}{n^2}.$$

By the Borel-Cantelli Lemma,

$$\sum_{k\leq m} G_k - B_m \leq |\sum_{i\leq n}(g_2^i - g_3^i)| + \sup_{n^c<m\leq(n+1)^c} |\sum_{n^c<k\leq m} G_k|$$

$$+ \sup_{n^c<m\leq(n+1)^c} |B_m - B_{n^c}| = o(n^{c(\frac{1}{2}-\bar{\epsilon})}) = o(m^{\frac{1}{2}-\bar{\epsilon}}) \text{ a.s..}$$

∎

Bibliography

[1] Aimino, R., Hu, H., Nicol, M., Török, A., and Vaienti, S. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete Contin. Dyn. Syst. 35*, 3 (2015), 793–806.

[2] Aimino, R., Nicol, M., and Vaienti, S. Annealed and quenched limit theorems for random expanding dynamical systems. *Probab. Theory Related Fields 162*, 1-2 (2015), 233–274.

[3] Aimino, R., and Rousseau, J. Concentration inequalities for sequential dynamical systems of the unit interval. *Ergodic Theory Dynam. Systems 36*, 8 (2016), 2384–2407.

[4] Alves, J. F., Freitas, J. M., Luzzatto, S., and Vaienti, S. From rates of mixing to recurrence times via large deviations. *Adv. Math. 228*, 2 (2011), 1203–1236.

[5] Aytaç, H., Freitas, J. M., and Vaienti, S. Laws of rare events for deterministic and random dynamical systems. *Trans. Amer. Math. Soc. 367*, 11 (2015), 8229–8278.

[6] Bahsoun, W., and Vaienti, S. Escape rates formulae and metastability for randomly perturbed maps. *Nonlinearity 26*, 5 (2013), 1415–1438.

[7] Berend, D., and Bergelson, V. Ergodic and mixing sequences of transformations. *Ergodic Theory Dynam. Systems 4*, 3 (1984), 353–366.

[8] Berkes, I., and Philipp, W. Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab. 7*, 1 (1979), 29–54.

[9] Brown, B. M. Martingale central limit theorems. *Ann. Math. Statist. 42* (1971), 59–66.

[10] Conze, J.-P., and Raugi, A. Limit theorems for sequential expanding dynamical systems on $[0, 1]$. In *Ergodic theory and related fields*, vol. 430 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 2007, pp. 89–121.

[11] Cuny, C., and Merlevède, F. Strong invariance principles with rate for "reverse" martingale differences and applications. *J. Theoret. Probab. 28*, 1 (2015), 137–183.

[12] Denker, M., and Philipp, W. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergodic Theory Dynam. Systems 4*, 4 (1984), 541–552.

[13] Dragičević, D., Froyland, G., González-Tokman, C., and Vaienti, S. A spectral approach for quenched limit theorems for random hyperbolic dynamical systems. *arXiv e-prints* (Dec 2018), arXiv:1812.07340.

[14] Dragičević, D., Froyland, G., González-Tokman, C., and Vaienti, S. Almost sure invariance principle for random piecewise expanding maps. *Nonlinearity 31*, 5 (2018), 2252–2280.

[15] Dragičević, D., Froyland, G., González-Tokman, C., and Vaienti, S. A spectral approach for quenched limit theorems for random expanding dynamical systems. *Comm. Math. Phys. 360*, 3 (2018), 1121–1187.

[16] Einmahl, U. Extensions of results of Komlós, Major, and Tusnády to the multivariate case. *J. Multivariate Anal. 28*, 1 (1989), 20–68.

[17] Field, M., Melbourne, I., and Török, A. Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions. *Ergodic Theory Dynam. Systems 23*, 1 (2003), 87–110.

[18] Gordin, M. I. The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR 188* (1969), 739–741.

[19] Gouëzel, S. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab. 38*, 4 (2010), 1639–1671.

[20] Haydn, N., Nicol, M., Török, A., and Vaienti, S. Almost sure invariance principle for sequential and non-stationary dynamical systems. *Trans. Amer. Math. Soc. 369*, 8 (2017), 5293–5316.

[21] Haydn, N., Nicol, M., Vaienti, S., and Zhang, L. Central limit theorems for the shrinking target problem. *J. Stat. Phys. 153*, 5 (2013), 864–887.

[22] Hella, O., and Leppänen, J. Central limit theorems with a rate of convergence for time-dependent intermittent maps. *arXiv e-prints* (Nov. 2018), arXiv:1811.11170.

[23] Hofbauer, F., and Keller, G. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z. 180*, 1 (1982), 119–140.

[24] Hu, H., and Vaienti, S. Absolutely continuous invariant measures for non-uniformly expanding maps. *Ergodic Theory Dynam. Systems 29*, 4 (2009), 1185–1215.

[25] Kallenberg, O. *Foundations of modern probability*, second ed. Probability and its Applications (New York). Springer-Verlag, New York, 2002.

[26] Kuelbs, J., and Philipp, W. Almost sure invariance principles for partial sums of mixing B-valued random variables. *Ann. Probab. 8*, 6 (1980), 1003–1036.

[27] Leppänen, J. Intermittent quasistatic dynamical systems: weak convergence of fluctuations. *Nonauton. Dyn. Syst. 5*, 1 (2018), 8–34.

[28] Leppänen, J., and Stenlund, M. Quasistatic dynamics with intermittency. *Math. Phys. Anal. Geom. 19*, 2 (2016), Art. 8, 23.

[29] Liverani, C. Decay of correlations. *Ann. of Math. (2) 142*, 2 (1995), 239–301.

[30] Liverani, C. Decay of correlations for piecewise expanding maps. *J. Statist. Phys. 78*, 3-4 (1995), 1111–1129.

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

[32] Liverani, C., Saussol, B., and Vaienti, S. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems 19*, 3 (1999), 671–685.

[33] Luzzatto, S., and Melbourne, I. Statistical properties and decay of correlations for interval maps with critical points and singularities. *Comm. Math. Phys. 320*, 1 (2013), 21–35.

[34] Melbourne, I., and Nicol, M. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys. 260*, 1 (2005), 131–146.

[35] Melbourne, I., and Nicol, M. A vector-valued almost sure invariance principle for hyperbolic dynamical systems. *Ann. Probab. 37*, 2 (2009), 478–505.

[36] Melbourne, I., and Török, A. Statistical limit theorems for suspension flows. *Israel J. Math. 144* (2004), 191–209.

[37] Merlevède, F., and Rio, E. Strong approximation of partial sums under dependence conditions with application to dynamical systems. *Stochastic Process. Appl. 122*, 1 (2012), 386–417.

[38] Nándori, P., Szász, D., and Varjú, T. A central limit theorem for time-dependent dynamical systems. *J. Stat. Phys. 146*, 6 (2012), 1213–1220.

[39] Nicol, M., Török, A., and Vaienti, S. Central limit theorems for sequential and random intermittent dynamical systems. *Ergodic Theory Dynam. Systems 38*, 3 (2018), 1127–1153.

[40] Ott, W., Stenlund, M., and Young, L.-S. Memory loss for time-dependent dynamical systems. *Math. Res. Lett. 16*, 3 (2009), 463–475.

[41] Philipp, W., and Stout, W. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc. 2*, 161,, 161 (1975), iv+140.

[42] Scott, D. J., and Huggins, R. M. On the embedding of processes in Brownian motion and the law of the iterated logarithm for reverse martingales. *Bull. Austral. Math. Soc. 27*, 3 (1983), 443–459.

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

[44] Serfling, R. J. Moment inequalities for the maximum cumulative sum. *Ann. Math. Statist. 41* (1970), 1227–1234.

[45] Sprindžuk, V. G. *Metric theory of Diophantine approximations.* V. H. Winston & Sons, Washington, D.C.; A Halsted Press Book, John Wiley & Sons, New York-Toronto, Ont.-London, 1979. Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics.

[46] Stenlund, M. Non-stationary compositions of Anosov diffeomorphisms. *Nonlinearity 24*, 10 (2011), 2991–3018.

[47] Stenlund, M. A vector-valued almost sure invariance principle for Sinai billiards with random scatterers. *Comm. Math. Phys. 325*, 3 (2014), 879–916.

[48] Su, Y. Almost sure invariance principle for non-stationary and random intermittent dynamical systems. *arXiv e-prints* (Mar 2019), arXiv:1903.09758.

[49] Su, Y. Random Young towers and quenched limit laws. *arXiv e-prints* (July 2019), arXiv:1907.12199.

[50] Young, L.-S. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2) 147*, 3 (1998), 585–650.

[51] Young, L.-S. Recurrence times and rates of mixing. *Israel J. Math. 110* (1999), 153–188.

[52] Zaitsev, A. Y. Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments. *ESAIM Probab. Statist. 2* (1998), 41–108.