BERRY-ESSEEN BOUNDS AND MODERATE DEVIATIONS FOR RANDOM WALKS ON GL d (R)

Hui Xiao, Ion Grama, Quansheng Liu

BERRY-ESSEEN BOUNDS AND MODERATE DEVIATIONS FOR RANDOM WALKS ON $GL_d(\mathbb{R})$

HUI XIAO, ION GRAMA, AND QUANSHENG LIU [1]

Abstract. Let $(g_n)_{n \geqslant 1}$ be a sequence of independent and identically distributed random elements of the general linear group $GL_d(\mathbb{R})$, with law μ. Consider the random walk $G_n := g_n \dots g_1$. Denote respectively by $\|G_n\|$ and $\rho(G_n)$ the operator norm and the spectral radius of G_n. For $\log \|G_n\|$ and $\log \rho(G_n)$, we prove moderate deviation principles under exponential moment and strong irreducibility conditions on μ; we also establish moderate deviation expansions in the normal range $[0, o(n^{1/6})]$ and Berry-Esseen bounds under the additional proximality condition on μ. Similar results are found for the couples $(X_n^x, \log \|G_n\|)$ and $(X_n^x, \log \rho(G_n))$ with target functions, where $X_n^x := G_n \cdot x$ is a Markov chain and x is a starting point on the projective space $\mathbb{P}(\mathbb{R}^d)$.

1. Introduction

1.1. **Background and previous results.** For any integer $d \geqslant 2$, denote by $\mathbb{G} = GL_d(\mathbb{R})$ the general linear group of real invertible $d \times d$ matrices. We equip the vector space $\mathbb{R}^d$ with the canonical Euclidean norm $|\cdot|$. Let $\mathbb{P}(\mathbb{R}^d)$ be the projective space in $\mathbb{R}^d$, which is defined as the set of elements $x = \mathbb{R}v$, where $v \in \mathbb{R}^d \setminus \{0\}$. For any $g \in \mathbb{G}$ and $v \in \mathbb{R}^d$, let gv be the multiplication of g by v. The action of a matrix $g \in \mathbb{G}$ on the direction $x = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^d)$ of a vector $v \in \mathbb{R}^d \setminus \{0\}$ is defined by $g \cdot x = \mathbb{R}gv$. For $g \in \mathbb{G}$, denote by $\|g\| = \sup_{v \in \mathbb{R}^{d-1} \setminus \{0\}} \frac{|gv|}{|v|}$ its operator norm, and by $\rho(g) = \lim_{k \to \infty} \|g^k\|^{1/k}$ its spectral radius. Let $(g_n)_{n \geqslant 1}$ be a sequence of independent and identically distributed (i.i.d.) random matrices with law μ on the group $\mathbb{G}$. Consider the left random walk $G_n = g_n \dots g_1$ on $\mathbb{G}$, and, for any starting point $X_0^x = x \in \mathbb{P}(\mathbb{R}^d)$, the Markov chain $X_n^x := G_n \cdot x$ on $\mathbb{P}(\mathbb{R}^d)$, where $n \geqslant 1$. The goal of this paper is to investigate Berry-Esseen type bounds and moderate deviation asymptotics for the operator norm

[1]Corresponding author, quansheng.liu@univ-ubs.fr

Date: September 1, 2021.

2010 *Mathematics Subject Classification.* Primary 60F10; Secondary 60B20, 60J05.

Key words and phrases. Berry-Esseen bounds; moderate deviation principles and expansions; random walks on groups; operator norm; spectral radius.

$\|G_n\|$ and the spectral radius $\rho(G_n)$, and more generally, for the couples $(X_n^x, \log \|G_n\|)$ and $(X_n^x, \log \rho(G_n))$ with target functions on X_n^x.

Let Γ_μ be the smallest closed subsemigroup of $\mathbb{G}$ generated by the support of μ. Denote $N(g) = \max\{\|g\|, \|g^{-1}\|\}$, where g^{-1} is the inverse matrix of $g \in \mathbb{G}$. Consider the following conditions.

A1 (Exponential moments). *There exists $\delta > 0$ such that $\mathbb{E}[N(g_1)^\delta] < \infty$.*

A2 (Strong irreducibility). *The support of μ acts strongly irreducibly on $\mathbb{R}^d$, i.e., no proper finite union of subspaces of $\mathbb{R}^d$ is invariant with respect to all elements of Γ_μ.*

A3 (Proximality). *Γ_μ contains at least one matrix with a unique eigenvalue of maximal modulus.*

Let us recall some basic results for the product G_n. Furstenberg and Kesten [14] first established the strong law of large numbers for the operator norm $\|G_n\|$: under the assumption that $\mathbb{E}[\max\{0, \log \|g_1\|\}] < \infty$, it holds that $\frac{1}{n} \log \|G_n\| \to \lambda$ almost surely as $n \to \infty$, where λ is a constant called top Lyapunov exponent of μ. This result turns out to be a consequence of Kingman's subadditive ergodic theorem [22] established later. The central limit theorem for $\|G_n\|$ is due to Le Page [23] (see also Bougerol and Lacroix [7]): if conditions **A1**, **A2** and **A3** hold, then $\frac{\log \|G_n\| - n\lambda}{\sigma\sqrt{n}}$ converges in law to the standard normal distribution, where $\sigma^2 > 0$ is the asymptotic variance of the random walk $(G_n)_{n \geqslant 1}$. Recently, using Gordin's martingale approximation method, Benoist and Quint [4] have relaxed the exponential moment condition **A1** to the optimal second moment condition $\mathbb{E}[\log^2 N(g_1)] < \infty$.

Similar law of large numbers and central limit theorem have been known for the spectral radius $\rho(G_n)$. Using the Hölder regularity of the invariant measure ν (see [17, 15]), Guivarc'h [15] has established the strong law of large numbers for $\rho(G_n)$: under conditions **A1**, **A2** and **A3**, $\frac{1}{n} \log \rho(G_n) \to \lambda$ almost surely as $n \to \infty$. Recently, under the same conditions, Benoist and Quint [5, Theorem 14.22] established the central limit theorem for $\rho(G_n)$: $\frac{\log \rho(G_n) - n\lambda}{\sigma\sqrt{n}}$ converges in law to the standard normal distribution. Further improvements have been done very recently: Aoun and Sert [2] proved the strong law of large numbers for $\rho(G_n)$ assuming only the second moment condition $\mathbb{E}[\log^2 N(g_1)] < \infty$, while Aoun [1] proved the central limit theorem for $\rho(G_n)$ under the second moment condition, the strong irreducibility condition **A2** and the unboundedness assumption of the semigroup Γ_μ.

Very little has been known about the Berry-Esseen bounds and moderate and large deviations, for the operator norm $\|G_n\|$ and the spectral radius $\rho(G_n)$. For Berry-Esseen type bounds, Cuny, Dedecker and Jan [12] (see also Cuny, Dedecker and Merlevède [13] in a more general setting) have recently established the following result about the rate of convergence in the central limit theorem for $\|G_n\|$: assuming $\mathbb{E}[\log^3 N(g_1)] < \infty$, conditions **A2** and **A3**, there exists a constant $C > 0$ such that for any $n \geqslant 2$,

$$\sup_{y\in\mathbb{R}} \left| \mathbb{P}\left(\frac{\log \|G_n\| - n\lambda}{\sigma\sqrt{n}} \leqslant y \right) - \Phi(y) \right| \leqslant \frac{C\sqrt{\log n}}{n^{1/4}}, \tag{1.1}$$

where Φ is the standard normal distribution function on $\mathbb{R}$. It is expected that the rate of convergence should be of the order $\frac{1}{\sqrt{n}}$, however, this is still an open problem even under stronger moment assumptions. The same question is also open for the spectral radius $\rho(G_n)$.

Moderate deviations have not yet been studied neither for $\|G_n\|$ nor for $\rho(G_n)$, to the best of our knowledge. For large deviations, the upper tail large deviation principle for $\|G_n\|$ has been established by Sert [24] and [25] under different conditions; it is conjectured in [24] that the usual large deviation principle would hold for $\rho(G_n)$.

1.2. **Objectives.** In this paper, we shall establish Berry-Esseen type bounds and moderate deviation results for both the operator norm $\|G_n\|$ and the spectral radius $\rho(G_n)$. Such kinds of results are important in applications because they give the rate of convergence in the central limit theorem and in the law of large numbers.

Our first objective is to establish the following Berry-Esseen type bound concerning the rate of convergence in the central limit theorem for the operator norm and for the spectral radius. We shall only give the results for the operator norm, since the results for the spectral radius are similar. Under conditions **A1**, **A2** and **A3**, for any Hölder continuous function φ on $\mathbb{P}(\mathbb{R}^d)$,

$$\sup_{x\in\mathbb{P}(\mathbb{R}^d)} \sup_{y\in\mathbb{R}} \left| \mathbb{E}\left[\varphi(X_n^x) \mathbb{1}_{\left\{ \frac{\log \|G_n\| - n\lambda}{\sigma\sqrt{n}} \leqslant y \right\}} \right] - \nu(\varphi)\Phi(y) \right| \leqslant \frac{C \log n}{\sqrt{n}}, \tag{1.2}$$

where ν is the unique invariant probability measure of the Markov chain $(X_n^x)_{n\geqslant 0}$. Under the stronger moment condition **A1**, our result improves on the bound (1.1) in two aspects. Firstly, we sharpen the rate of convergence by showing the rate $\frac{\log n}{\sqrt{n}}$. Secondly, we extend the validity of the bound for the couple $(X_n^x, \log \|G_n\|)$ with a target function on X_n^x.

Our second objective is to establish moderate deviation principles for $\log \|G_n\|$ and $\log \rho(G_n)$. We first deal with the couple $(X_n^x, \log \|G_n\|)$ under conditions **A1**, **A2** and **A3**. Namely, for any non-negative Hölder continuous function φ on $\mathbb{P}(\mathbb{R}^d)$ satisfying $\nu(\varphi) > 0$, any Borel set $B \subseteq \mathbb{R}$ and any sequence $(b_n)_{n\geqslant 1}$ of positive numbers satisfying $\frac{b_n}{\sqrt{n}} \to \infty$ and $\frac{b_n}{n} \to 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$
\begin{aligned}
- \inf_{y \in B^\circ} \frac{y^2}{2\sigma^2} &\leqslant \liminf_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{E} \left[\varphi(X_n^x) \mathbb{1}_{\left\{ \frac{\log \|G_n\| - n\lambda}{b_n} \in B \right\}} \right] \\
&\leqslant \limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{E} \left[\varphi(X_n^x) \mathbb{1}_{\left\{ \frac{\log \|G_n\| - n\lambda}{b_n} \in B \right\}} \right] \leqslant - \inf_{y \in \bar{B}} \frac{y^2}{2\sigma^2}, \quad (1.3)
\end{aligned}
$$

where B° and $\bar{B}$ are respectively the interior and the closure of B. It is also interesting to investigate the case when the proximality condition **A3** fails. It turns out that we are still able to prove the moderate deviation principle for $\|G_n\|$. We show (see Theorem 2.4) that there exists a constant $\sigma_0 > 0$ such that

$$
\begin{aligned}
- \inf_{y \in B^\circ} \frac{y^2}{2\sigma_0^2} &\leqslant \liminf_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P} \left(\frac{\log \|G_n\| - n\lambda}{b_n} \in B \right) \\
&\leqslant \limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P} \left(\frac{\log \|G_n\| - n\lambda}{b_n} \in B \right) \leqslant - \inf_{y \in \bar{B}} \frac{y^2}{2\sigma_0^2}. \quad (1.4)
\end{aligned}
$$

While the proximality condition **A3** ensures that the Markov chain $(X_n^x)_{n \geqslant 0}$ has a unique invariant probability measure ν on the projective space $\mathbb{P}(\mathbb{R}^d)$, in the opposite case there is no unique invariant probability measure. In this case a completely different approach is required; this is developed in Section 4.2. It is rather interesting to compare the moderate deviation result (1.4) with the large deviation asymptotic: when the proximality condition **A3** fails, the rate function in the large deviation principle is not known, moreover we do not even know whether the large deviation principle holds. We refer to Breuillard [10] and He, Lakrec and Lindenstrauss [18] for large deviation bounds for the operator norm. For the spectral radius we establish results similar to (1.3) and (1.4) under the same conditions. Very recent progress was made by Boulanger, Mathieu, Sert and Sisto [8], where the large deviation principle for the spectral radius has been established for the special case of simple linear algebraic groups of rank 1.

Our third objective is to establish a moderate deviation expansion. For the couple $(X_n^x, \log \|G_n\|)$ we show that under conditions **A1**, **A2** and **A3**, for any Hölder continuous function φ on $\mathbb{P}(\mathbb{R}^d)$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$ and $y \in [0, o(n^{1/6})]$,

$$
\frac{\mathbb{E} \left[\varphi(X_n^x) \mathbb{1}_{\left\{ \log \|G_n\| - n\lambda \geqslant \sqrt{n}\sigma y \right\}} \right]}{1 - \Phi(y)} = \nu(\varphi) + o(1). \quad (1.5)
$$

The expansion (1.5) is new even when $\varphi = \mathbf{1}$. A similar result is obtained for the spectral radius. These results are interesting since they give a precise asymptotic of moderate deviation probabilities in terms of the normal tail $1 - \Phi(y)$ in the range $y \in [0, o(n^{1/6})]$, which is known to be optimal.

1.3. **Proof outline.** In [12, 13], the proof of (1.1) consists of establishing the central limit theorem with rate of convergence in Wasserstein's distance utilising the martingale approximation method developed in [4]. With this approach, even if we obtain the best rate of convergence $\frac{1}{\sqrt{n}}$ in Wasserstein's distance, while passing to the Kolmogorov distance we can only get the rate $\frac{1}{n^{1/4}}$ in the Berry-Esseen bound. To get a better bound, a new approach is needed. Let $\sigma(\cdot, \cdot)$ be the norm cocycle defined by

$$\sigma(g, x) = \log \frac{|gv|}{|v|}, \tag{1.6}$$

where $g \in \mathbb{G}$ and $x = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^d)$ with $v \in \mathbb{R}^d \setminus \{0\}$. Our proof of (1.2) is based on the Berry-Esseen bound for the couple $(X_n^x, \sigma(G_n, x))$ recently established in [26], and on the following precise comparison between $\|G_n\|$ and $|G_n v|$ established in [5]: for any $a > 0$, there exist $c > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$ and $v \in \mathbb{R}^d \setminus \{0\}$,

$$\mathbb{P}\left(\left| \log \frac{\|G_n\|}{\|G_k\|} - \log \frac{|G_n v|}{|G_k v|} \right| \leqslant e^{-ak} \right) > 1 - e^{-ck}. \tag{1.7}$$

The basic idea to utilize this powerful inequality consists in carefully choosing certain integer k, taking the conditional expectation with respect to the σ-algebra $\sigma\{g_1, \ldots, g_k\}$ and using the large deviation bounds for $\log \|G_k\|$. This technique, in conjugation with limit theorems for the norm cocycle $\sigma(G_{n-k}, x)$, makes it possible to prove corresponding results for $\log \|G_n\|$; see [5] where a local limit theorem for $\log \|G_n\|$ has been established by taking $k = \lfloor \log^2 n \rfloor$, where $\lfloor a \rfloor$ denotes the integral part of a. In this paper, the proof of (1.2) is carried out by choosing $k = \lfloor C_1 \log n \rfloor$ with a sufficiently large constant $C_1 > 0$ and by using the Berry-Esseen bound for the couple $(X_n^x, \sigma(G_n, x))$ with a target function φ on X_n^x. In the same spirit, the moderate deviation principle (1.3) for the couple $(X_n^x, \log \|G_n\|)$ is established using the moderate deviation principle for the couple $(X_n^x, \sigma(G_n, x))$ proved in [26], together with the inequality (1.7) with $k = \lfloor C_1 \frac{b_n^2}{n} \rfloor$, where $C_1 > 0$ is a sufficiently large constant and the sequence $(b_n)_{n \geqslant 1}$ is given in (1.3).

As to the moderate deviation principle (1.4) for $\log \|G_n\|$ without assuming the proximality condition **A3**, its proof is more technical and delicate than that of (1.3). Indeed, when condition **A3** fails, the transfer operator of the Markov chain $(X_n^x)_{n \geqslant 0}$ has no spectral gap in general and it may happen that $(X_n^x)_{n \geqslant 0}$ possesses several invariant measures on the projective space $\mathbb{P}(\mathbb{R}^d)$. In this case, it becomes hopeless to prove a general form of (1.4) when a target function φ on X_n^x is taken into account. Nevertheless, the proof of (1.4) can be carried out by following the approach of Bougerol and Lacroix [7] (first announced in [6]), where central limit theorems and exponential large deviation bounds for $\log \|G_n\|$ and $\sigma(G_n, x)$ were established without giving the rate function. Specifically, employing this approach consists in finding the proximal dimension p of the semigroup Γ_μ generated by the matrix law μ and then applying Chevaley's algebraic irreducible representation [11] of the exterior powers $\wedge^p \mathbb{R}^d$, to show that the action of the semigroup Γ_μ is strongly irreducible and proximal on $\wedge^p \mathbb{R}^d$. Using this strategy together with (1.3) for $\varphi = \mathbf{1}$, we are able to establish (1.4).

For the proof of the Cramér type moderate deviation expansion (1.5) in the normal range $[0, o(n^{1/6})]$, when $y \in [0, \frac{1}{2}\sqrt{\log n}]$, we deduce the desired result from the Berry-Esseen type bound (1.2); when $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$, we make use of the moderate deviation expansion for the couple $(X_n^x, \sigma(G_n, x))$ recently established in [26] and the inequality (1.7) with $k = \lfloor C_1 y^2 \rfloor$, where $C_1 > 0$ is a sufficiently large constant.

All of the aforementioned results (1.2), (1.3), (1.4) and (1.5) for the operator norm $\|G_n\|$ turn out to be essential to establish analogous Berry-Esseen type bounds and moderate deviation results for the spectral radius $\rho(G_n)$. Another important ingredient in our proof is the precise comparison between $\rho(G_n)$ and $\|G_n\|$ established in [5]; see Lemma 3.2 below.

2. Main results

Let $\mathcal{C}(\mathbb{P}(\mathbb{R}^d))$ be the space of continuous complex-valued functions on the projective space $\mathbb{P}(\mathbb{R}^d)$ and $\mathbf{1}$ be the constant function with value 1 on $\mathbb{P}(\mathbb{R}^d)$. We equip the projective space $\mathbb{P}(\mathbb{R}^d)$ with the distance $\mathbf{d}$ defined by $\mathbf{d}(x, x') = \frac{|v \wedge v'|}{|v||v'|}$ for $x = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^d)$ and $x' = \mathbb{R}v' \in \mathbb{P}(\mathbb{R}^d)$, where $v \wedge v'$ denotes the exterior product of v and v' in $\mathbb{R}^d$. We assume that $\gamma > 0$ is a fixed small enough constant. Consider the Banach space $\mathcal{B}_\gamma := \{\varphi \in \mathcal{C}(\mathbb{P}(\mathbb{R}^d)) : \|\varphi\|_\gamma < +\infty\}$, where

$$\|\varphi\|_\gamma := \|\varphi\|_\infty + \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{\mathbf{d}^\gamma(x, y)} \quad \text{with} \quad \|\varphi\|_\infty := \sup_{x \in \mathbb{P}(\mathbb{R}^d)} |\varphi(x)|.$$

Recall that $g \cdot x$ denotes the action of the matrix $g \in \mathbb{G}$ on the element $x = \mathbb{R}v$ in the projective space $\mathbb{P}(\mathbb{R}^d)$, namely $g \cdot x = \mathbb{R}gv$. For any starting point $x \in \mathbb{P}(\mathbb{R}^d)$, the sequence $(X_n^x)_{n \geqslant 0}$ defined by

$$X_0^x = x, \qquad X_n^x = G_n \cdot x, \quad n \geqslant 1,$$

constitutes a Markov chain on the projective space $\mathbb{P}(\mathbb{R}^d)$. Under conditions **A1**, **A2** and **A3**, the chain $(X_n^x)_{n \geqslant 0}$ possesses a unique invariant probability measure ν on $\mathbb{P}(\mathbb{R}^d)$ such that $\mu * \nu = \nu$ (see [16]), where $\mu * \nu$ denotes the convolution of μ and ν. It is worth mentioning that if the proximality condition **A3** fails, then the invariant measure ν may not be unique (see [7, 5]). By [5, Proposition 14.17], the asymptotic variance σ^2 of the random walk $(G_n)_{n\geqslant 1}$ can be given by

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[(\log\|G_n\| - n\lambda)^2\right].$$

Throughout the paper, we denote by Φ the standard normal distribution function on $\mathbb{R}$. We write c, C for positive constants whose values may change from line to line.

2.1. **Berry-Esseen type bounds.** In this subsection, we present Berry-Esseen type bounds for the couples $(X_n^x, \log\|G_n\|)$ and $(X_n^x, \log\rho(G_n))$ with target functions on X_n^x. Recall that by Gelfand's formula, it holds that $\rho(g) = \lim_{k\to\infty}\|g^k\|^{1/k}$ for any $g \in \mathbb{G}$.

Theorem 2.1. *Assume conditions* ***A1***, ***A2*** *and* ***A3***. *Then there exists a constant $C > 0$ such that for all $n \geqslant 2$, $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in \mathbb{R}$ and $\varphi \in \mathcal{B}_\gamma$,*

$$\left|\mathbb{E}\left[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log\|G_n\|-n\lambda}{\sigma\sqrt{n}}\leqslant y\right\}}\right] - \nu(\varphi)\Phi(y)\right| \leqslant \frac{C\log n}{\sqrt{n}}\|\varphi\|_\gamma, \tag{2.1}$$

$$\left|\mathbb{E}\left[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log\rho(G_n)-n\lambda}{\sigma\sqrt{n}}\leqslant y\right\}}\right] - \nu(\varphi)\Phi(y)\right| \leqslant \frac{C\log n}{\sqrt{n}}\|\varphi\|_\gamma. \tag{2.2}$$

Using the fact that all matrix norms are equivalent, one can verify that in (2.1), the operator norm $\|\cdot\|$ can be replaced by any matrix norm.

Under the exponential moment condition **A1**, the Berry-Esseen type bound (2.1) with $\varphi = \mathbf{1}$ improves (1.1), which has been established recently by Cuny, Dedecker and Jan [12] (see also Cuny, Dedecker and Merlevède [13] in a more general setting) under the weaker third-order moment condition $\mathbb{E}[\log^3 N(g_1)] < \infty$.

The result with a general target function φ is worth some comments. On the one hand, it concerns the joint distribution of the couples $(X_n^x, \log\|G_n\|)$ and $(X_n^x, \log\rho(G_n))$, which give more information and can lead to interesting applications. On the other hand, the extension from the case $\varphi = \mathbf{1}$ to a general function φ is not trivial, for which a significant difficulty appears. The difficulty will be overcome by using the Berry-Esseen bound for the couple $(X_n^x, \sigma(G_n, x))$.

It is natural to make the conjecture that the optimal rate of convergence on the right hand sides of (2.1) and (2.2) should be $\frac{C}{\sqrt{n}}$ instead of $\frac{C\log n}{\sqrt{n}}$. For positive matrices, these optimal bounds have been proved in [27]. However, the proofs of the conjecture for invertible matrices seem to be rather delicate, for which new ideas and techniques are required. Nevertheless, we can prove the optimal bound $\frac{C}{\sqrt{n}}$ for large values of $|y|$, see the remark below.

Remark 2.2. Under the same conditions as in Theorem 2.1, if we consider $|y| > \sqrt{3\log\log n}$ instead of $y \in \mathbb{R}$, then the bound $\frac{C\log n}{\sqrt{n}}$ in (2.1) and (2.2) can be improved to be $\frac{C}{\sqrt{n}}$.

The proof of this remark will be given in the proof of Theorem 2.1.

2.2. **Moderate deviation principles.** We first state moderate deviation principles for the couples $(X_n^x, \log \|G_n\|)$ and $(X_n^x, \log \rho(G_n))$ with target functions on the Markov chain $(X_n^x)_{n\geqslant 0}$.

Theorem 2.3. *Assume conditions **A1**, **A2** and **A3**. Then, for any non-negative function $\varphi \in \mathcal{B}_\gamma$ satisfying $\nu(\varphi) > 0$, for any Borel set $B \subseteq \mathbb{R}$ and any sequence $(b_n)_{n\geqslant 1}$ of positive numbers satisfying $\frac{b_n}{\sqrt{n}} \to \infty$ and $\frac{b_n}{n} \to 0$, we have, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,*

$$
\begin{aligned}
- \inf_{y \in B^\circ} \frac{y^2}{2\sigma^2} &\leqslant \liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log\|G_n\| - n\lambda}{b_n} \in B\right\}}\Big] \\
&\leqslant \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log\|G_n\| - n\lambda}{b_n} \in B\right\}}\Big] \leqslant - \inf_{y\in \bar{B}} \frac{y^2}{2\sigma^2}, \quad (2.3)
\end{aligned}
$$

and

$$
\begin{aligned}
- \inf_{y \in B^\circ} \frac{y^2}{2\sigma^2} &\leqslant \liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log\rho(G_n) - n\lambda}{b_n} \in B\right\}}\Big] \\
&\leqslant \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log\rho(G_n) - n\lambda}{b_n} \in B\right\}}\Big] \leqslant - \inf_{y\in \bar{B}} \frac{y^2}{2\sigma^2}, \quad (2.4)
\end{aligned}
$$

where B° and $\bar{B}$ are respectively the interior and the closure of B.

Note that the target function φ in (2.3) and (2.4) is not necessarily strictly positive on $\mathbb{P}(\mathbb{R}^d)$. The moderate deviation principles (2.3) and (2.4) are new, even for $\varphi = \mathbf{1}$.

If we only consider the operator norm $\|G_n\|$ or the spectral radius $\rho(G_n)$, instead of the couples $(X_n^x, \log \|G_n\|)$ and $(X_n^x, \log \rho(G_n))$, we are still able to establish moderate deviation principles without assuming the proximality condition **A3**:

Theorem 2.4. *Assume conditions **A1**, **A2** and $\sigma^2 > 0$. Then, there exists a constant $\sigma_0 > 0$ such that for any Borel set $B \subseteq \mathbb{R}$ and any sequence $(b_n)_{n\geqslant 1}$ of positive numbers satisfying $\frac{b_n}{\sqrt{n}} \to \infty$ and $\frac{b_n}{n} \to 0$, we have,*

$$
\begin{aligned}
- \inf_{y \in B^\circ} \frac{y^2}{2\sigma_0^2} &\leqslant \liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log\|G_n\| - n\lambda}{b_n} \in B \right) \\
&\leqslant \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log\|G_n\| - n\lambda}{b_n} \in B \right) \leqslant - \inf_{y\in \bar{B}} \frac{y^2}{2\sigma_0^2}, \quad (2.5)
\end{aligned}
$$

and

$$-\inf_{y\in B^\circ}\frac{y^2}{2\sigma_0^2}\leqslant \liminf_{n\to\infty}\frac{n}{b_n^2}\log\mathbb{P}\left(\frac{\log\rho(G_n)-n\lambda}{b_n}\in B\right)$$
$$\leqslant \limsup_{n\to\infty}\frac{n}{b_n^2}\log\mathbb{P}\left(\frac{\log\rho(G_n)-n\lambda}{b_n}\in B\right)\leqslant -\inf_{y\in\bar B}\frac{y^2}{2\sigma_0^2}, \quad (2.6)$$

where B° and $\bar B$ are respectively the interior and the closure of B.

Remark 2.5. Assume conditions **A1** and **A2**. Let

$$\Gamma_{\mu,1}=\{|\det(g)|^{-1/d}g : g\in\Gamma_\mu\}$$

be the set of elements of Γ_μ normalized to have determinant 1.

(1) If $\Gamma_{\mu,1}$ is not contained in a compact subgroup of $\mathbb{G}$, then $\sigma>0$, as will be seen in the proof of Theorem 2.4 .

(2) If $\Gamma_{\mu,1}$ is contained in a compact subgroup of $\mathbb{G}$, then $c_1=\inf\{\|g\| : g\in\Gamma_{\mu,1}\}>0$ and $c_2=\sup\{\|g\| : g\in\Gamma_{\mu,1}\}<\infty$, so that

$$c_1^d\,|\det(g)|\leqslant\|g\|^d\leqslant c_2^d\,|\det(g)|\quad \forall g\in\Gamma_\mu. \quad (2.7)$$

Since $\log|\det(G_n)|=\sum_{i=1}^n\log|\det(g_i)|$ is a sum of i.i.d. real-valued random variables, from (2.7) (applied to $g=G_n$) it follows directly that the moderate deviation principle (2.5) holds with $\lambda=\frac{1}{d}\mathbb{E}\log|\det(g_1)|$ and $\sigma_0^2=\mathbb{E}[(\frac{1}{d}\log|\det(g_1)|-\lambda)^2]$ (which coincide with their original definitions), provided that $|\det(g_1)|$ is not a.s. a constant (which is equivalent to $\sigma_0^2>0$).

2.3. **Moderate deviation expansions.** In this subsection we formulate the Cramér type moderate deviation expansions in the normal range for the operator norm $\|G_n\|$ and the spectral radius $\rho(G_n)$. Our first result concerns the operator norm $\|G_n\|$.

Theorem 2.6. *Assume conditions **A1**, **A2** and **A3**. Then, we have, uniformly in $x\in\mathbb{P}(\mathbb{R}^d)$, $y\in[0,o(n^{1/6})]$ and $\varphi\in\mathcal{B}_\gamma$, as $n\to\infty$,*

$$\frac{\mathbb{E}\left[\varphi(X_n^x)\mathbb{1}_{\left\{\log\|G_n\|-n\lambda\geqslant\sqrt{n}\sigma y\right\}}\right]}{1-\Phi(y)}=\nu(\varphi)+\|\varphi\|_\gamma o(1), \quad (2.8)$$

$$\frac{\mathbb{E}\left[\varphi(X_n^x)\mathbb{1}_{\left\{\log\|G_n\|-n\lambda\leqslant-\sqrt{n}\sigma y\right\}}\right]}{\Phi(-y)}=\nu(\varphi)+\|\varphi\|_\gamma o(1). \quad (2.9)$$

When $\varphi=\mathbf{1}$, the expansions (2.8) and (2.9) are also new.

The proof of Theorem 2.6 is based on the Cramér type moderate deviation expansion for the couple $(X_n^x,\sigma(G_n,x))$ proved recently in [26], and on a fine comparison between the logarithm of the operator norm $\log\|G_n\|$ and the norm cocycle $\sigma(G_n,x)$ established in [5] (see Lemma 3.1 below). Note that

Theorem 2.6 covers the special case where $\nu(\varphi) = 0$; in this case the exact comparison with the normal tail remains open.

Our second result concerns the moderate deviation expansion for the spectral radius $\rho(G_n)$, also in the normal range.

Theorem 2.7. *Assume conditions* ***A1****,* ***A2*** *and* ***A3****. Then, we have, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in [0, o(n^{1/6})]$ and $\varphi \in \mathcal{B}_\gamma$, as $n \to \infty$,*

$$\frac{\mathbb{E}\left[\varphi(X_n^x)\mathbb{1}_{\left\{\log \rho(G_n) - n\lambda \geqslant \sqrt{n}\sigma y\right\}}\right]}{1 - \Phi(y)} = \nu(\varphi) + \|\varphi\|_\gamma o(1), \tag{2.10}$$

$$\frac{\mathbb{E}\left[\varphi(X_n^x)\mathbb{1}_{\left\{\log \rho(G_n) - n\lambda \leqslant -\sqrt{n}\sigma y\right\}}\right]}{\Phi(-y)} = \nu(\varphi) + \|\varphi\|_\gamma o(1). \tag{2.11}$$

The proof of Theorem 2.7 relies on Theorem 2.6 and on an estimate of the difference between spectral radius $\rho(G_n)$ and the operator norm $\|G_n\|$ established in [5] (see Lemma 3.2).

Like in Theorem 2.6, when $\varphi = \mathbf{1}$, the expansions (2.10) and (2.11) are also new; Theorem 2.7 also covers the case where $\nu(\varphi) = 0$, for which the exact comparaison with the normal tail is not known.

3. Berry-Esseen type bounds

The goal of this section is to prove Theorem 2.1 about Berry-Esseen type bounds for the operator norm $\|G_n\|$ and for the spectral radius $\rho(G_n)$.

We shall use the following result which is an interesting comparison theorem for $\log \|G_n\|$ and $\sigma(G_n, x)$.

Lemma 3.1. [5, Lemma 17.8] *Assume conditions* ***A1****,* ***A2*** *and* ***A3****. Then, for any $a > 0$, there exist $c > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$ and $v \in \mathbb{R}^d \setminus \{0\}$,*

$$\mathbb{P}\left(\left|\log \frac{\|G_n\|}{\|G_k\|} - \log \frac{|G_n v|}{|G_k v|}\right| \leqslant e^{-ak}\right) > 1 - e^{-ck}.$$

Proof of (2.1) *of Theorem 2.1.* Without loss of generality, we assume that the target function φ is non-negative (otherwise we can consider the positive and negative parts of φ). On the one hand, using the Berry-Esseen bound for the norm cocycle $\sigma(G_n, x)$ established in [26, Theorem 2.1] and the fact that $\log \|G_n\| \geqslant \sigma(G_n, x)$, we get the following upper bound: there exists a constant $C > 0$ such that for all $n \geqslant 1$, $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in \mathbb{R}$ and $\varphi \in \mathcal{B}_\gamma$,

$$I_n := \mathbb{E}\left[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log \|G_n\| - n\lambda}{\sigma\sqrt{n}} \leqslant y\right\}}\right] \leqslant \nu(\varphi)\Phi(y) + \frac{C}{\sqrt{n}}\|\varphi\|_\gamma. \tag{3.1}$$

On the other hand, applying Lemma 3.1, we deduce that for any $a > 0$, there exist $c > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$, it holds uniformly in $\varphi \in \mathcal{B}_\gamma$ and $x = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^d)$ with $v \in \mathbb{R}^d \setminus \{0\}$,

$$\begin{aligned} I_n &\geqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\frac{\log\|G_n\|-n\lambda}{\sigma\sqrt{n}}\leqslant y\big\}}\mathbb{1}_{\big\{\big|\log\|G_n\|-\log\frac{|G_nv|}{|G_kv|}-\log\|G_k\|\big|\leqslant e^{-ak}\big\}}\Big] \\ &\geqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\frac{\sigma(G_n,x)-\sigma(G_k,x)+\log\|G_k\|-n\lambda+e^{-ak}}{\sigma\sqrt{n}}\leqslant y\big\}}\Big] - e^{-ck}\|\varphi\|_\infty. \end{aligned} \tag{3.2}$$

For brevity, for any $n > k \geqslant 1$, we write $G_n = G_{n,k}G_k$ with

$$G_{n,k} = g_n \dots g_{k+1}, \qquad G_k = g_k \dots g_1.$$

From the large deviation bounds for $\log\|G_k\|$ (see [5] or [25]), we have that for any $q > \lambda$, there exist constants $c, C > 0$ such that for any $k \geqslant 1$,

$$\mathbb{P}(\log\|G_k\| > kq) \leqslant Ce^{-ck}. \tag{3.3}$$

Denote the σ-algebra $\mathscr{F}_k = \sigma(g_1, \dots, g_k)$. From (3.2), taking the conditional expectation with respect to the filtration $\mathscr{F}_k$, we derive that for any $q > \lambda$,

$$\begin{aligned} I_n &\geqslant \mathbb{E}\left\{\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\frac{\sigma(G_n,x)-\sigma(G_k,x)+\log\|G_k\|-n\lambda+e^{-ak}}{\sigma\sqrt{n}}\leqslant y\big\}}\Big|\mathscr{F}_k\Big]\right\} - e^{-ck}\|\varphi\|_\infty \\ &\geqslant \mathbb{E}\left\{\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\frac{\sigma(G_n,x)-\sigma(G_k,x)+\log\|G_k\|-n\lambda+e^{-ak}}{\sigma\sqrt{n}}\leqslant y\big\}}\mathbb{1}_{\{\log\|G_k\|\leqslant kq\}}\Big|\mathscr{F}_k\Big]\right\} \\ &\quad - e^{-ck}\|\varphi\|_\infty \\ &\geqslant \mathbb{E}\left\{\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\frac{\sigma(G_n,x)-\sigma(G_k,x)+kq-n\lambda+e^{-ak}}{\sigma\sqrt{n}}\leqslant y\big\}}\Big|\mathscr{F}_k\Big]\right\} - Ce^{-c_1k}\|\varphi\|_\infty, \end{aligned}$$

where in the last step we use the large deviation bound (3.3) and the constant $c_1 > 0$ is taken to be small enough. Since, for $x = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^d)$ with $v \in \mathbb{R}^d \setminus \{0\}$,

$$X_n^x = G_n \cdot x = G_{n,k} \cdot X_k^x \quad \text{and} \quad \sigma(G_n, x) - \sigma(G_k, x) = \sigma(G_{n,k}, X_k^x), \tag{3.4}$$

it follows that

$$I_n \geqslant \mathbb{E}\left\{\mathbb{E}\Big[\varphi(G_{n,k}\cdot X_k^x)\mathbb{1}_{\big\{\frac{\sigma(G_{n,k},X_k^x)+kq-n\lambda+e^{-ak}}{\sigma\sqrt{n}}\leqslant y\big\}}\Big|\mathscr{F}_k\Big]\right\} - Ce^{-c_1k}\|\varphi\|_\infty.$$

Using the Berry-Esseen bound for the norm cocycle $\sigma(G_n, x)$ (cf. [26, Theorem 2.1]), we obtain

$$I_n \geqslant \nu(\varphi)\Phi(y_1) - \frac{C}{\sqrt{n-k}}\|\varphi\|_\gamma - Ce^{-c_1k}\|\varphi\|_\infty,$$

where

$$y_1 = \frac{\sqrt{n}}{\sqrt{n-k}} y - \frac{k(q-\lambda) + e^{-ak}}{\sigma \sqrt{n-k}}.$$

Taking $k = \lfloor C_1 \log n \rfloor$ with $C_1 = \frac{1}{2c_1}$, we get that there exists a constant $C > 0$ such that $\frac{1}{\sqrt{n-k}} \leqslant \frac{C}{\sqrt{n}}$ and $e^{-c_1 k} \leqslant \frac{C}{\sqrt{n}}$. Note that

$$\Phi(y_1) = \Phi(y) - \frac{1}{\sqrt{2\pi}} \int_{y_1}^{y} e^{-\frac{t^2}{2}} dt.$$

By elementary calculations, there exists a constant $C_2 > 0$ such that

$$|y - y_1| \leqslant C_2 \left(\frac{\log n}{n} |y| + \frac{\log n}{\sqrt{n}} \right),$$

and for $n > k_0$ large enough,

$$\begin{aligned} e^{-\frac{y_1^2}{2}} &\leqslant \exp\Big\{ -\frac{1}{2} \frac{n}{n-k} y^2 + \sqrt{n} \frac{k(q-\lambda) + e^{-ak}}{\sigma(n-k)} y \Big\} \\ &\leqslant \exp\Big\{ -\frac{1}{2} y^2 + C_2 \frac{\log n}{\sqrt{n}} |y| \Big\}. \end{aligned}$$

Thus, there exists a constant $C > 0$ such that

$$\begin{aligned} \left| \int_{y_1}^{y} e^{-\frac{t^2}{2}} dt \right| &\leqslant |y - y_1| \max\Big\{ e^{-\frac{y^2}{2}}, e^{-\frac{y_1^2}{2}} \Big\} \\ &\leqslant C_2 \Big(\frac{\log n}{n} |y| + \frac{\log n}{\sqrt{n}} \Big) \exp\Big\{ -\frac{1}{2} y^2 + C_2 \frac{\log n}{\sqrt{n}} |y| \Big\} \\ &\leqslant \begin{cases} C \frac{\log n}{\sqrt{n}} & \forall y \in \mathbb{R}, \\ C \frac{1}{\sqrt{n}} & \text{if } |y| > \sqrt{3 \log\log n}. \end{cases} \end{aligned}$$

Consequently, we get the following lower bound for I_n: there exists a constant $C > 0$ such that for all $x \in \mathbb{P}(\mathbb{R}^d)$ and $\varphi \in \mathcal{B}_\gamma$,

$$I_n \geqslant \begin{cases} \nu(\varphi)\Phi(y) - \frac{C \log n}{\sqrt{n}} \|\varphi\|_\gamma & \forall y \in \mathbb{R}, \\ \nu(\varphi)\Phi(y) - \frac{C}{\sqrt{n}} \|\varphi\|_\gamma & \text{if } |y| > \sqrt{3 \log\log n}. \end{cases}$$

Together with the upper bound (3.1), this concludes the proof of (2.1) of Theorem 2.1 and the corresponding results in Remark 2.2. □

We now proceed to prove the Berry-Esseen type bound (2.2) for the spectral radius $\rho(G_n)$. The proof relies on the following comparison lemma between the operator norm $\|G_n\|$ and the spectral radius $\rho(G_n)$.

Lemma 3.2. [5, Lemma 14.13] *Assume conditions* ***A1*** *and* ***A2****. Then, for any* $\varepsilon > 0$*, there exist* $c > 0$ *and* $k_0 \in \mathbb{N}$*, such that for all* $n \geqslant k \geqslant k_0$*,*

$$\mathbb{P}\left(1 \geqslant \frac{\rho(G_n)}{\|G_n\|} > e^{-\varepsilon k}\right) \geqslant 1 - e^{-ck}.$$

Proof of (2.2) *of Theorem 2.1.* Without loss of generality, we assume that the target function φ is non-negative.

The lower bound is a direct consequence of (2.1) together with Remark 2.2 on it, and the inequality $\log \rho(G_n) \leqslant \log \|G_n\|$, from which we get that, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in \mathbb{R}$ and $\varphi \in \mathcal{B}_\gamma$,

$$\begin{aligned} I_n &:= \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log \rho(G_n) - n\lambda}{\sigma\sqrt{n}} \leqslant y\right\}}\Big] \\ &\geqslant \begin{cases} \nu(\varphi)\Phi(y) - \frac{C\log n}{\sqrt{n}}\|\varphi\|_\gamma & \forall y \in \mathbb{R}, \\ \nu(\varphi)\Phi(y) - \frac{C}{\sqrt{n}}\|\varphi\|_\gamma & \text{if } |y| > \sqrt{3\log\log n}. \end{cases} \end{aligned}$$

The upper bound is a consequence of (2.1) together with Remark 2.2 on it and Lemma 3.2. In fact, applying Lemma 3.2, we deduce that for any $\varepsilon > 0$, there exist $c_1 > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$,

$$\begin{aligned} I_n &\leqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log \rho(G_n) - n\lambda}{\sigma\sqrt{n}} \leqslant y\right\}}\mathbb{1}_{\left\{\log \rho(G_n) - \log\|G_n\| > -\varepsilon k\right\}}\Big] + e^{-c_1 k}\|\varphi\|_\infty \\ &\leqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log \|G_n\| - \varepsilon k - n\lambda}{\sigma\sqrt{n}} \leqslant y\right\}}\Big] + e^{-c_1 k}\|\varphi\|_\infty. \end{aligned}$$

Taking $k = \lfloor C_1 \log n \rfloor$ with $C_1 = \frac{1}{2c_1}$, we have $e^{-c_1 k} \leqslant \frac{C}{\sqrt{n}}$ for some constant $C > 0$. Using the bound (2.1) with y replaced by $y_1 := y + \frac{\varepsilon k}{\sigma\sqrt{n}}$, we obtain the following upper bound for I_n: there exists a constant $C > 0$ such that for all $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in \mathbb{R}$, $\varphi \in \mathcal{B}_\gamma$, and $n \geqslant k_0$ with k_0 large enough,

$$I_n \leqslant \begin{cases} \nu(\varphi)\Phi(y_1) + \frac{C\log n}{\sqrt{n}}\|\varphi\|_\gamma & \forall y \in \mathbb{R}, \\ \nu(\varphi)\Phi(y_1) + \frac{C}{\sqrt{n}}\|\varphi\|_\gamma & \text{if } |y| > \sqrt{3\log\log n}. \end{cases}$$

(Notice that $|y| > \sqrt{3\log\log n}$ implies $|y_1| > \sqrt{3\log\log n}$ for n large enough.) By an argument similar to that used in the proof of (2.1), it can be seen that

$$\Phi(y_1) \leqslant \begin{cases} \Phi(y) + \frac{C\log n}{\sqrt{n}} & \forall y \in \mathbb{R}, \\ \Phi(y) + \frac{C}{\sqrt{n}} & \text{if } |y| > \sqrt{3\log\log n}. \end{cases}$$

This concludes the proof of (2.2) and Remark 2.2 on it. $\square$

4. Moderate deviation principles

The goal of this section is to establish Theorems 2.3 and 2.4 about moderate deviation principles for the operator norm $\|G_n\|$ and the spectral radius $\rho(G_n)$. Notice that in the first theorem, we need the proximality condition, while in the second, we do not need it.

4.1. **Proof of Theorem 2.3.**

Proof of (2.3) *of Theorem 2.3.* Since the rate function $I(y) := \frac{y^2}{2\sigma^2}$, $y \in \mathbb{R}$, is strictly increasing on $[0, \infty)$ and strictly decreasing on $(-\infty, 0]$ with $I(0) = 0$, by Lemma 4.4 of [20], it suffices to prove the following moderate deviation asymptotics: for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log\|G_n\| - n\lambda}{b_n} \geqslant y\big\}}\Big] = -\frac{y^2}{2\sigma^2}, \tag{4.1}$$

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log\|G_n\| - n\lambda}{b_n} \leqslant -y\big\}}\Big] = -\frac{y^2}{2\sigma^2}. \tag{4.2}$$

We first prove (4.1). For the lower bound, using the moderate deviation principle [26, (2.10)] for the norm cocycle $\sigma(G_n, x)$, and the fact that $\sigma(G_n, x) \leqslant \log\|G_n\|$, we get that for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log\|G_n\| - n\lambda}{b_n} \geqslant y\big\}}\Big] \geqslant -\frac{y^2}{2\sigma^2}.$$

We now prove the upper bound. Denote by $(e_i)_{1\leqslant i\leqslant d}$ the standard orthonormal basis of $\mathbb{R}^d$. Since all matrix norms on $\mathbb{R}^d$ are equivalent, and both $g \mapsto \|g\|$ and $g \mapsto \max_{1\leqslant i\leqslant d} |g e_i|$ are matrix norms, there exists a positive constant c_1 such that $\log\|G_n\| \leqslant \max_{1\leqslant i\leqslant d} \log|G_n e_i| + c_1$. From this inequality, we derive that

$$\mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log\|G_n\| - n\lambda}{b_n} \geqslant y\big\}}\Big] \leqslant \sum_{i=1}^d \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log|G_n e_i| - n\lambda + c_1}{b_n} \geqslant y\big\}}\Big].$$

Since $b_n \to \infty$ as $n \to \infty$, we have that for any $\varepsilon > 0$, it holds that $\frac{c_1}{b_n} < \varepsilon$ for large enough n. Thus, using again the moderate deviation principle [26, (2.10)] for $\sigma(G_n, x)$, we obtain that for any $y > 0$,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log|G_n e_i| - n\lambda + c_1}{b_n} \geqslant y\big\}}\Big] \\
&\leqslant \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log|G_n e_i| - n\lambda}{b_n} \geqslant y - \varepsilon\big\}}\Big] = -\frac{(y-\varepsilon)^2}{2\sigma^2}.
\end{aligned}$$

Since $\varepsilon > 0$ can be arbitrary small, we get the desired upper bound. This concludes the proof of (4.1).

We next prove (4.2). The upper bound is straightforward: using again the moderate deviation principle [26, (2.10)] for $\sigma(G_n, x)$, and the fact that $\sigma(G_n, x) \leqslant \log \|G_n\|$, we get that for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log \|G_n\| - n\lambda}{b_n} \leqslant -y\big\}}\Big] \leqslant -\frac{y^2}{2\sigma^2}. \tag{4.3}$$

For the lower bound we need to use Lemma 3.1. For any $n \geqslant k$ and $v \in \mathbb{R}^d \setminus \{0\}$, consider the event

$$A_{n,k} = \Big\{ \Big| \log \|G_n\| - \log \frac{|G_n v|}{|G_k v|} - \log \|G_k\| \Big| \leqslant e^{-ak} \Big\},$$

and denote by $A_{n,k}^c$ its complement. By Lemma 3.1, for any $a > 0$, there exist $c_1 > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$ and $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\mathbb{P}(A_{n,k}^c) \leqslant e^{-c_1 k}. \tag{4.4}$$

By (4.4), we see that

$$\begin{aligned}
I_n :&= \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log \|G_n\| - n\lambda}{b_n} \leqslant -y\big\}}\Big] \\
&\geqslant \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\log \|G_n\| - n\lambda}{b_n} \leqslant -y\big\}} \mathbb{1}_{A_{n,k}}\Big] \\
&\geqslant \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\sigma(G_n,x) - \sigma(G_k,x) + \log \|G_k\| - n\lambda + e^{-ak}}{b_n} \leqslant -y\big\}} \mathbb{1}_{A_{n,k}}\Big] \\
&\geqslant \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\sigma(G_n,x) - \sigma(G_k,x) + \log \|G_k\| - n\lambda + e^{-ak}}{b_n} \leqslant -y\big\}}\Big] - e^{-c_1 k} \|\varphi\|_\infty. \qquad (4.5)
\end{aligned}$$

As in the proof of (2.1), for any $n \geqslant k \geqslant k_0$, we write $G_n = G_{n,k} G_k$ with $G_{n,k} = g_n \ldots g_{k+1}$ and $G_k = g_k \ldots g_1$. Taking the conditional expectation with respect to the filtration $\mathscr{F}_k = \sigma(g_1, \ldots, g_k)$, and using (3.3), we derive that for any $q > \lambda$, there exist constants $c_2, C > 0$ such that for any $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\begin{aligned}
I_n &\geqslant \mathbb{E}\Big\{ \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\sigma(G_n,x) - \sigma(G_k,x) + \log \|G_k\| - n\lambda + e^{-ak}}{b_n} \leqslant -y\big\}} \mathbb{1}_{\{\log \|G_k\| \leqslant kq\}} \Big| \mathscr{F}_k \Big] \Big\} \\
&\qquad - e^{-c_1 k} \|\varphi\|_\infty \\
&\geqslant \mathbb{E}\Big\{ \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\sigma(G_n,x) - \sigma(G_k,x) + kq - n\lambda + e^{-ak}}{b_n} \leqslant -y\big\}} \Big| \mathscr{F}_k \Big] \Big\} - C e^{-c_2 k} \|\varphi\|_\infty \\
&=: J_n - C e^{-c_2 k} \|\varphi\|_\infty. \qquad (4.6)
\end{aligned}$$

Using the moderate deviation principle [26, (2.10)] for $\sigma(G_n, x)$, we have that for any $y > 0$, $\epsilon > 0$ and sufficiently large n, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$e^{-\frac{b_n^2}{n}\big(\frac{y^2}{2\sigma^2} + \epsilon\big)} \leqslant \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\big\{\frac{\sigma(G_n,x) - n\lambda}{b_n} \leqslant -y\big\}}\Big] \leqslant e^{-\frac{b_n^2}{n}\big(\frac{y^2}{2\sigma^2} - \epsilon\big)}. \tag{4.7}$$

In the sequel, we take $k = \left\lfloor C_1 \frac{b_n^2}{n} \right\rfloor$, where $C_1 > 0$ is a constant to be chosen sufficiently large. If we denote

$$b_n' = b_n + \frac{k(q-\lambda) + e^{-ak}}{y},$$

then, using (3.4), the term J_n defined in (4.6) can be rewritten as

$$J_n = \mathbb{E}\left\{ \mathbb{E}\Big[\varphi(G_{n,k} \cdot X_k^x) \mathbb{1}_{\left\{\sigma(G_{n,k}, X_k^x) - (n-k)\lambda \leqslant -y b_n' \right\}} \Big| \mathscr{F}_k \Big] \right\}.$$

Since $\frac{b_n'}{\sqrt{n}} \to \infty$ and $\frac{b_n'}{n} \to 0$ as $n \to \infty$, applying (4.7) with n replaced by $n-k$, and with b_n replaced by b_n', we obtain that for any fixed $y > 0$ and $\epsilon > 0$ and for n large enough, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$e^{-\frac{(b_n')^2}{n-k}\left(\frac{y^2}{2\sigma^2} + \epsilon\right)} \leqslant J_n \leqslant e^{-\frac{(b_n')^2}{n-k}\left(\frac{y^2}{2\sigma^2} - \epsilon\right)}. \tag{4.8}$$

From (4.6) and (4.8), there exists a constant $c_3 > 0$ such that

$$\begin{aligned} \liminf_{n\to\infty} \frac{n}{b_n^2} \log I_n &\geqslant \liminf_{n\to\infty} \frac{n}{b_n^2} \log \left[e^{-\frac{(b_n')^2}{n-k}\left(\frac{y^2}{2\sigma^2} + \epsilon\right)} - C e^{-c_2 k} \right] \\ &\geqslant \liminf_{n\to\infty} \frac{n}{b_n^2} \log \left[e^{-\frac{(b_n')^2}{n-k}\left(\frac{y^2}{2\sigma^2} + \epsilon\right)} (1 - C e^{-c_3 k}) \right], \end{aligned}$$

where the last inequality holds due to the fact that as $n \to \infty$,

$$\frac{(b_n')^2}{k(n-k)} \Big(\frac{y^2}{2\sigma^2} + \epsilon \Big) \to \frac{1}{C_1} \Big(\frac{y^2}{2\sigma^2} + \epsilon \Big) < c_2$$

by choosing $C_1 > \left(\frac{y^2}{2\sigma^2} + \epsilon \right)/c_2$. Recalling that $k = \lfloor C_1 \frac{b_n^2}{n} \rfloor \to \infty$, we get $\lim_{n\to\infty} \frac{n}{b_n^2} \log(1 - C e^{-c_3 k}) = 0$ and

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log e^{-\frac{(b_n')^2}{n-k}\left(\frac{y^2}{2\sigma^2} + \epsilon\right)} = -\Big(\frac{y^2}{2\sigma^2} + \epsilon \Big).$$

Taking $\epsilon \to 0$, we get that for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{ \frac{\log \|G_n\| - n\lambda}{b_n} \leqslant -y \right\}} \Big] \geqslant -\frac{y^2}{2\sigma^2}.$$

Combining this with the upper bound (4.3), we obtain (4.2) and thus conclude the proof of (2.3). ☐

Using (2.3) and Lemma 3.2, we are now in a position to establish the moderate deviation principle (2.4) for the couple $(X_n^x, \log \rho(G_n))$.

Proof of (2.4) *of Theorem 2.3.* As explained in the proof of (2.3), it suffices to prove that, for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log \rho(G_n) - n\lambda}{b_n} \geqslant y\right\}}\Big] = -\frac{y^2}{2\sigma^2}, \tag{4.9}$$

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log \rho(G_n) - n\lambda}{b_n} \leqslant -y\right\}}\Big] = -\frac{y^2}{2\sigma^2}. \tag{4.10}$$

We first prove (4.9). On the one hand, since the function φ is non-negative, using (2.3) and the fact that $\rho(G_n) \leqslant \|G_n\|$, we get the upper bound: for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log \rho(G_n) - n\lambda}{b_n} \geqslant y\right\}}\Big] \leqslant -\frac{y^2}{2\sigma^2}. \tag{4.11}$$

On the other hand, by Lemma 3.2, we obtain that for any $\varepsilon > 0$, there exist $c_1 > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$,

$$\begin{aligned} I_n :&= \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log \rho(G_n) - n\lambda}{b_n} \geqslant y\right\}}\Big] \\ &\geqslant \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\{\log \|G_n\| - n\lambda \geqslant y b_n + \varepsilon k\}} \mathbb{1}_{\{\log \rho(G_n) - \log \|G_n\| \geqslant -\varepsilon k\}}\Big] \\ &\geqslant \mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\{\log \|G_n\| - n\lambda \geqslant y b_n + \varepsilon k\}}\Big] - e^{-c_1 k} \|\varphi\|_\infty. \end{aligned} \tag{4.12}$$

As in the proof of (2.3), we take $k = \lfloor C_1 \frac{b_n^2}{n} \rfloor$, where $C_1 > 0$ is a constant which will be chosen sufficiently large. By (2.3), for any $y > 0$ and $\eta > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$,

$$\mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\left\{\frac{\log \|G_n\| - n\lambda}{b_n} \geqslant y\right\}}\Big] \geqslant e^{-\frac{b_n^2}{n}\left(\frac{y^2}{2\sigma^2} + \eta\right)}. \tag{4.13}$$

Set $b_n' = b_n + \frac{\varepsilon k}{y}$. Since $\frac{b_n'}{\sqrt{n}} \to \infty$ and $\frac{b_n'}{n} \to 0$ as $n \to \infty$, using (4.13), we get that uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\mathbb{E}\Big[\varphi(X_n^x) \mathbb{1}_{\{\log \|G_n\| - n\lambda \geqslant y b_n + \varepsilon k\}}\Big] \geqslant e^{-\frac{(b_n')^2}{n}\left(\frac{y^2}{2\sigma^2} + \eta\right)}.$$

Implementing this bound into (4.12), we obtain

$$I_n \geqslant e^{-\frac{(b_n')^2}{n}\left(\frac{y^2}{2\sigma^2} + \eta\right)} \Big[1 - e^{-c_1 k + \frac{(b_n')^2}{n}\left(\frac{y^2}{2\sigma^2} + \eta\right)} \|\varphi\|_\infty\Big].$$

Choosing $C_1 > \frac{1}{c_1}\left(\frac{y^2}{2\sigma^2} + \eta\right)$, we have, as $n \to \infty$,

$$\frac{(b_n')^2}{kn} \left(\frac{y^2}{2\sigma^2} + \eta\right) \to \frac{1}{C_1} \left(\frac{y^2}{2\sigma^2} + \eta\right) < c_1.$$

Hence we get for some constant $c_2 > 0$,

$$I_n \geqslant e^{-\frac{(b_n')^2}{n}\left(\frac{y^2}{2\sigma^2}+\eta\right)}\Big[1 - e^{-c_2 k}\|\varphi\|_\infty\Big].$$

Therefore, using the fact that $k = \lfloor C_1 \frac{b_n^2}{n} \rfloor \to \infty$, we obtain

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log I_n \geqslant -\Big(\frac{y^2}{2\sigma^2} + \eta\Big).$$

Taking $\eta \to 0$, we obtain that for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log\rho(G_n)-n\lambda}{b_n}\geqslant y\right\}}\Big] \geqslant -\frac{y^2}{2\sigma^2}.$$

Together with the upper bound (4.11), this concludes the proof of (4.9).

We next prove (4.10). Using (4.2) and the fact that $\rho(G_n) \leqslant \|G_n\|$, we get the desired lower bound: for any $y > 0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log\rho(G_n)-n\lambda}{b_n}\leqslant -y\right\}}\Big] \geqslant -\frac{y^2}{2\sigma^2}. \tag{4.14}$$

For the upper bound, as before, set $k = \lfloor C_2 \frac{b_n^2}{n} \rfloor$, where $C_2 > 0$ is a constant to be chosen larger enough. By Lemma 3.2, for any $\varepsilon > 0$, there exist $c_3 > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$,

$$\begin{aligned} J_n :&= \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log\rho(G_n)-n\lambda}{b_n}\leqslant -y\right\}}\Big] \\ &\leqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\log\|G_n\|-n\lambda\leqslant -yb_n+\varepsilon k\right\}}\Big] + e^{-c_3 k}\|\varphi\|_\infty. \end{aligned}$$

By (4.2), for any $\eta > 0$, there exists $n_0 \in \mathbb{N}$ such that for any $n \geqslant n_0$,

$$\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\frac{\log\|G_n\|-n\lambda}{b_n}\leqslant -y\right\}}\Big] \leqslant e^{-\frac{b_n^2}{n}\left(\frac{y^2}{2\sigma^2}-\eta\right)}. \tag{4.15}$$

Let $b_n' = b_n - \frac{\varepsilon k}{y}$. We see that $\frac{b_n'}{\sqrt{n}} \to \infty$ and $\frac{b_n'}{n} \to 0$, as $n \to \infty$. From (4.15), it follows that uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$J_n \leqslant e^{-\frac{(b_n')^2}{n}\left(\frac{y^2}{2\sigma^2}-\eta\right)} + e^{-c_3 k}\|\varphi\|_\infty. \tag{4.16}$$

Since $b_n' = b_n - \frac{\varepsilon k}{y}$ and $k = \lfloor C_1 \frac{b_n^2}{n} \rfloor$, it holds that as $n \to \infty$, $\frac{b_n'}{b_n} \to 1$ and

$$\frac{(b_n')^2}{kn}\big(\frac{y^2}{2\sigma^2} - \eta\big) \to \frac{1}{C_1}\,\big(\frac{y^2}{2\sigma^2} - \eta\big) < c_3,$$

by choosing $C_1 > \frac{1}{c_3}\big(\frac{y^2}{2\sigma^2} - \eta\big)$. Thus,

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log J_n \leqslant -\lim_{n\to\infty}\Big(\frac{b_n'}{b_n}\Big)^2\Big(\frac{y^2}{2\sigma^2} - \eta\Big) = -\Big(\frac{y^2}{2\sigma^2} - \eta\Big).$$

Since $\eta > 0$ is arbitrary, we get the upper bound for J_n. Combining this with the lower bound (4.14), we conclude the proof of (4.10).

Putting together (4.9) and (4.10), we obtain (2.4). $\square$

4.2. **Proof of Theorem 2.4.** We now come to the proof of moderate deviation principles without assuming the proximality condition **A3**; see Theorem 2.4. The proof is based on Theorem 2.3 applied to $\| \wedge^p G_n\|$. In [7, Theorem V. 6.2], this approach is used to establish large deviation bounds for $\sigma(G_n, x)$ and $\log \|G_n\|$; it allows to relax the proximality condition **A3** for an exponential large deviation bound, but fails to give the rate function in the large deviation principle. For moderate deviations, the situation is different: with this approach we are able to get the rate function explicitly.

We need to introduce some additional notation. For any integer $1 \leqslant p \leqslant d$, the p-th exterior power $\wedge^p(\mathbb{R}^d)$ is a $\binom{d}{p}$-dimensional vector space with basis

$$\left\{e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p}, 1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant d\right\},$$

where $(e_i)_{1\leqslant i\leqslant d}$ is the standard orthonormal basis of $\mathbb{R}^d$; it is endowed with the standard norm still denoted by $|\cdot|$ as in the case of $\mathbb{R}^d$ (there should be no confusion in the context). For any $v_1, \dots, v_p \in \mathbb{R}^d$, the vector $v_1 \wedge \cdots \wedge v_p$ is nonzero if and only if $v_1, \dots, v_p$ are linearly independent in $\mathbb{R}^d$. We write $\wedge^p g$ for the image of $g \in \mathbb{G} = GL_d(\mathbb{R})$ under the representation $\wedge^p(\mathbb{R}^d)$; for any $v_1, \dots, v_p \in \mathbb{R}^d$, the action of the matrix $\wedge^p g$ on the vector $v_1 \wedge \cdots \wedge v_p$ is defined as follows:

$$\wedge^p g(v_1 \wedge \cdots \wedge v_p) = gv_1 \wedge \cdots \wedge gv_p.$$

The associated operator norm of $\wedge^p g$ is defined by

$$\| \wedge^p g\| = \sup\left\{|(\wedge^p g)v| : v \in \wedge^p(\mathbb{R}^d), |v| = 1\right\}.$$

Since $\wedge^p(gg') = (\wedge^p g)(\wedge^p g')$ for any $g, g' \in \mathbb{G}$, the submultiplicative property holds: $\| \wedge^p (gg')\| \leqslant \| \wedge^p g\|\| \wedge^p g'\|$. If the singular values of a matrix $g \in \mathbb{G}$ is given by $a_{11}, \dots, a_{dd}$ (arranged in decreasing order), then it holds that

$$\| \wedge^p g\| = a_{11} \dots a_{pp}. \tag{4.17}$$

As a consequence, we have $\| \wedge^p g\| \leqslant \|g\|^p$ and $\| \wedge^p g\|\| \wedge^{p+2} g\| \leqslant \| \wedge^{p+1} g\|^2$.

Let V be a subspace of $\wedge^p(\mathbb{R}^d)$. A set $S \subset \wedge^p(\mathbb{G}) := \{\wedge^p g : g \in \mathbb{G}\}$ is said to be *irreducible* on V if there is no proper subspace V_1 of V such that $gV_1 = V_1$ for all $g \in S$. A set $S \subset \wedge^p(\mathbb{G})$ is said to be *strongly irreducible* on V if there are no finite number of subspaces $V_1, \dots, V_m$ of V such that $g(V_1 \cup \ldots \cup V_m) = V_1 \cup \ldots \cup V_m$ for all $g \in S$. In particular, the strong irreducibility condition **A2** means that the Γ_μ (the smallest closed subsemigroup of $\mathbb{G}$ generated by the support of μ) is strongly irreducible on $\mathbb{R}^d$. Denote by G_μ the smallest closed subgroup of $\mathbb{G}$ generated by the support of μ. Then, condition **A2** is equivalent to saying that G_μ is strongly irreducible on $\mathbb{R}^d$. Indeed, the set $S = \{g \in \mathbb{G} : g(V_1 \cup \ldots \cup V_m) = V_1 \cup \ldots \cup V_m\}$ is a subgroup of $\mathbb{G}$, so that $\Gamma_\mu \subset S$ if and only if $G_\mu \subset S$, which means that $V_1 \cup \ldots \cup V_m$ is Γ_μ-invariant if and only if $V_1 \cup \ldots \cup V_m$ is G_μ-invariant. We refer to [7] for more details.

The following purely algebraic result is due to Chevalley [11]; see also Bougerol and Lacroix [7].

Lemma 4.1. *Let G be an irreducible subgroup of $GL_d(\mathbb{R})$. Then, for any integer $1 \leqslant p \leqslant d$, there exists a direct-sum decomposition of the p-th exterior power: $\wedge^p(\mathbb{R}^d) = V_1 \oplus \ldots \oplus V_k$ such that $(\wedge^p g)V_j = V_j$ for any $g \in G$ and $1 \leqslant j \leqslant k$. Moreover, $\wedge^p(G) := \{\wedge^p g : g \in G\}$ is irreducible on each subspace $V_j, j = 1, \cdots, k$.*

We say that an integer $1 \leqslant p \leqslant d$ is the *proximal dimension* of the semigroup Γ_μ, if p is the smallest integer with the following property: there exists a sequence of matrices $\{M_n\}_{n\geqslant 1} \subset \Gamma_\mu$ such that $\frac{M_n}{\|M_n\|}$ converges to a matrix with rank p. By definition, the proximality condition **A3** implies that the proximal dimension of Γ_μ is 1. The converse is also true if we assume that Γ_μ is irreducible, see [5] for the proof. Under the first moment condition $\mathbb{E}(\log N(g_1)) < \infty$, by Kingman's subadditive ergodic theorem [22], the Lyapunov exponents $(\lambda_p)_{1\leqslant p\leqslant d}$ of μ are defined recursively by

$$\lambda_1 + \ldots + \lambda_p = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}(\log \| \wedge^p G_n\|) = \lim_{n\to\infty} \frac{1}{n} \log \| \wedge^p G_n\|, \quad a.s..$$

This formula, together with the fact that $\| \wedge^{p-1} G_n\| \| \wedge^{p+1} G_n\| \leqslant \| \wedge^p G_n\|^2$, yields that $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_d$. The following fundamental result is due to Guivarc'h and Raugi [17] and gives a sufficient condition for ensuring that two successive Lyapunov exponents are distinct. It can also be found in [7, Proposition III. 6.2].

Lemma 4.2. *Assume condition **A2**. If $\mathbb{E}\log N(g_1) < \infty$ and the proximal dimension of the semigroup Γ_μ is p, then $\lambda_1 = \lambda_2 = \ldots = \lambda_p > \lambda_{p+1}$.*

The following result is from [7, Lemma III. 1.4].

Lemma 4.3. *Assume condition **A2**. If $\mathbb{E}\log N(g_1) < \infty$ and the proximal dimension of the semigroup Γ_μ is p, then there exists a constant $c > 0$ such that $c\|g\|^p \leqslant \| \wedge^p g\| \leqslant \|g\|^p$ for any $g \in \Gamma_\mu$.*

The following lemma was proved in [7, Proposition III. 1.7 and Remark III. 1.8]. Recall that $\Gamma_{\mu,1} = \{|\det(g)|^{-1/d} g : g \in \Gamma_\mu\}$.

Lemma 4.4. *(a) If the set $\Gamma_{\mu,1}$ is not contained in a compact subgroup of $\mathbb{G}$, then the proximal dimension p of the semigroup Γ_μ satisfies $1 \leqslant p \leqslant d-1$.*

(b) If the set $\Gamma_{\mu,1}$ is contained in a compact subgroup of $\mathbb{G}$, then there exists a scalar product on $\mathbb{R}^d$ for which all the matrices in $\Gamma_{\mu,1}$ are orthogonal.

In this case, $\log \|G_n\|$ *can be written as a sum of i.i.d. real-valued random variables.*

Now we are equipped to prove the moderate deviation principle (2.5) for the operator norm $\|G_n\|$ without assuming the proximality condition **A3**.

Proof of (2.5) *of Theorem 2.4.* We assume that $\Gamma_{\mu,1}$ is not contained in a compact subgroup of $\mathbb{G}$; the opposite case was already proved in Remark 2.5(2). Note that $\lambda = \lambda_1$. Without loss of generality, we assume that $\lambda_1 = 0$ since otherwise we can replace each matrix $g \in \Gamma_\mu$ by $e^{-\lambda_1} g$. As mentioned before, to prove (2.5), it is sufficient to show that for any $y > 0$,

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \|G_n\|}{b_n} \geqslant y \right) = -\frac{y^2}{2\sigma_0^2}, \tag{4.18}$$

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \|G_n\|}{b_n} \leqslant -y \right) = -\frac{y^2}{2\sigma_0^2}. \tag{4.19}$$

We first give a proof of (4.18). Let p be the proximal dimension of the semigroup Γ_μ. Since the set $\Gamma_{\mu,1}$ is not contained in a compact subgroup of $\mathbb{G}$, by Lemma 4.4 (a), we have $1 \leqslant p \leqslant d-1$. By Lemma 4.2, under condition **A2**, this implies that the Lyapunov exponents $(\lambda_p)_{1 \leqslant p \leqslant d}$ of μ satisfy

$$\lambda_1 = \ldots = \lambda_p = 0 > \lambda_{p+1}.$$

It follows that the two largest Lyapunov exponents of $\wedge^p G_n$ are given by $\lambda_1 + \cdots + \lambda_p = 0$ and $\lambda_2 + \cdots + \lambda_{p+1} = \lambda_{p+1} < 0$ (see [7, Proposition III. 1,2]). Applying Lemma 4.1 to $G = G_\mu$ (the smallest closed subgroup of $\mathbb{G}$ generated by the support of μ), we get the following direct-sum decomposition of the p-th exterior power $\wedge^p(\mathbb{R}^d)$:

$$\wedge^p(\mathbb{R}^d) = V_1 \oplus V_2 \oplus \ldots \oplus V_k,$$

where V_j are subspaces of $\wedge^p(\mathbb{R}^d)$ such that $(\wedge^p g)V_j = V_j$ for any $g \in G_\mu$ and $1 \leqslant j \leqslant k$, i.e. each V_j is invariant under $\wedge^p(G_\mu) := \{\wedge^p g : g \in G_\mu\}$. Moreover, $\wedge^p(G_\mu)$ is irreducible on each subspace V_j. Note that the set of all Lyapunov exponents of $\wedge^p G_n$ on the space $\wedge^p(\mathbb{R}^d)$ coincides with the union of all the Lyapunov exponents of $(\wedge^p G_n)$ restricted to each subspace V_j, $1 \leqslant j \leqslant k$. Hence we can choose V_1 in such a way that the restrictions of $\wedge^p G_n$ to V_1 and $V_2 \oplus \ldots \oplus V_k$, denoted respectively by G_n' and G_n'' (as usual we identify the linear transform with the corresponding matrice), satisfy:

$$\lim_{n\to\infty} \frac{1}{n} \log \|G_n'\| = \lambda_1 + \cdots + \lambda_p = 0 \quad \text{a.s.}, \tag{4.20}$$

$$\lim_{n\to\infty} \frac{1}{n} \log \|G_n''\| = \lambda_2 + \cdots + \lambda_{p+1} = \lambda_{p+1} < 0 \quad \text{a.s.}, \tag{4.21}$$

$$\| \wedge^p G_n \| = \max\{\|G_n'\|, \|G_n''\|\}. \tag{4.22}$$

Here, G_n' and G_n'' are products of i.i.d. invertible matrices of the form $G_n' = g_n' \cdots g_1'$ and $G_n'' = g_n'' \cdots g_1''$. We denote by μ_1 the law of the random matrix g_1, by d_1 the dimension of the vector space V_1, and by Γ_{μ_1} the smallest closed subsemigroup of $GL_{d_1}(\mathbb{R})$ generated by the support of μ_1. Then, following the same argument used in the proof of the central limit theorem for $\|G_n\|$ (see [7, Theorem V.5.4]), one can verify, under condition **A2** on μ, that the semigroup Γ_{μ_1} is strongly irreducible and proximal on $\mathbb{R}^{d_1}$. Therefore, μ_1 satisfies conditions **A2** and **A3**, so that we can apply the moderate deviation principle (2.3) with $\varphi = \mathbf{1}$ and G_n replaced by G_n', to get the following moderate deviation asymptotics: for any $y > 0$,

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \|G_n'\|}{b_n} \geqslant y \right) = -\frac{y^2}{2\sigma_1^2}, \tag{4.23}$$

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \|G_n'\|}{b_n} \leqslant -y \right) = -\frac{y^2}{2\sigma_1^2}, \tag{4.24}$$

where $\sigma_1^2 > 0$ is the asymptotic variance of the sequence $(G_n')_{n\geqslant 1}$ given by

$$\sigma_1^2 = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\Big[(\log \|G_n'\|)^2 \Big]. \tag{4.25}$$

From (4.22) and (4.23), we get the lower bound for $\| \wedge^p G_n\|$: for any $y > 0$,

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \| \wedge^p G_n\|}{b_n} \geqslant y \right) \geqslant -\frac{y^2}{2\sigma_1^2}. \tag{4.26}$$

On the other hand, since the upper Lyapunov exponent of the sequence $(G_n'')_{n\geqslant 1}$ is strictly less than 0 (see (4.21)), we have $\mathbb{E}(\log \|G_m''\|) < 0$ for sufficiently large integer $m \geqslant 1$. If we write $n = km + r$ with $k \geqslant 1$ and $0 \leqslant r < m$, then we have the identity

$$G_n'' = [G_n''(G_{km}'')^{-1}]\, [G_{km}''(G_{(k-1)m}'')^{-1}] \ldots [G_{2m}''(G_m'')^{-1}]\, G_m'',$$

and hence

$$\log \|G_n''\| \leqslant \log \|G_n''(G_{km}'')^{-1}\| + \log \|G_{km}''(G_{(k-1)m}'')^{-1}\| + \ldots + \log \|G_m''\|. \tag{4.27}$$

For fixed integer $m \geqslant 1$, we denote $u_m := -\mathbb{E}(\log \|G_m''\|) > 0$. Then,

$$\begin{aligned} \mathbb{P}(\log \|G_n''\| \geqslant 0) &\leqslant \mathbb{P}\Big(\log \|G_n''(G_{km}'')^{-1}\| \geqslant k\frac{u_m}{2} \Big) \\ &\quad + \mathbb{P}\Big(\log \|G_{km}''(G_{(k-1)m}'')^{-1}\| + \cdots + \log \|G_m''\| + k u_m \geqslant k\frac{u_m}{2} \Big). \end{aligned} \tag{4.28}$$

Using (4.22) and the fact that $\| \wedge^p g\| \leqslant \|g\|^p$ for any $g \in \Gamma_\mu$, we get that for constant $c > 0$ small enough,

$$\mathbb{E}(\|G_n''(G_{km}'')^{-1}\|^c) = \mathbb{E}(\|G_r''\|^c) \leqslant \mathbb{E}(\| \wedge^p G_r\|^c) \leqslant \mathbb{E}(\|G_r\|^{cp}) \leqslant \big[\mathbb{E}(\|g_1\|^{cp})\big]^r,$$

which is finite by condition **A1**. By Markov's inequality and the fact that $u_m > 0$ is a constant, it follows that there exist constants $c, C > 0$ such that

$$\mathbb{P}\Big(\log \|G''_n (G''_{km})^{-1}\| \geqslant k \frac{u_m}{2} \Big) \leqslant \mathbb{E}(\|G''_n (G''_{km})^{-1}\|^c) e^{-ck\frac{u_m}{2}} \leqslant C e^{-ck}. \tag{4.29}$$

Using the large deviation bounds for sums of i.i.d. real-valued random variables, the second term on the right hand side of (4.28) is dominated by Ce^{-ck}. Implementing this bound and (4.29) into (4.28), and taking into account $k \geqslant n/(m+1)$, we obtain

$$\mathbb{P}(\log \|G''_n\| \geqslant 0) \leqslant C e^{-cn}. \tag{4.30}$$

From (4.23) we derive that for any $y > 0$ and $\epsilon > 0$, there exists $n_0 \geqslant 1$ such that for any $n \geqslant n_0$,

$$\mathbb{P}\big(\log \|G'_n\| \geqslant y b_n \big) \leqslant \exp\Big\{ - \frac{b_n^2}{n} \Big(\frac{y^2}{2\sigma_1^2} - \epsilon \Big) \Big\}.$$

This, together with (4.30) and (4.22), yields the upper bound for $\| \wedge^p G_n \|$: for any $y > 0$ and $\epsilon > 0$,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \| \wedge^p G_n \|}{b_n} \geqslant y \right) \\
&\leqslant \limsup_{n\to\infty} \frac{n}{b_n^2} \log \left[\mathbb{P}\left(\frac{\log \|G'_n\|}{b_n} \geqslant y \right) + \mathbb{P}\Big(\frac{\log \|G''_n\|}{b_n} \geqslant y \Big) \right] \\
&\leqslant \limsup_{n\to\infty} \frac{n}{b_n^2} \log \left[e^{-\frac{b_n^2}{n}\left(\frac{y^2}{2\sigma_1^2} - \epsilon\right)} + C e^{-cn} \right] = -\Big(\frac{y^2}{2\sigma_1^2} - \epsilon \Big).
\end{aligned}$$

Since $\epsilon > 0$ can be arbitrary small, it follows that

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \| \wedge^p G_n \|}{b_n} \geqslant y \right) \leqslant -\frac{y^2}{2\sigma_1^2}. \tag{4.31}$$

Since the proximal dimension of the semigroup Γ_μ is p, by Lemma 4.3, the sequence $\{\log \| \wedge^p G_n\| - p \log \|G_n\|\}_{n \geqslant 1}$ is bounded by a constant from above and below. Combining this with (4.26) and (4.31), we get that for any $y > 0$,

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \|G_n\|}{b_n} \geqslant y \right) = -\frac{y^2}{2\sigma_0^2}, \tag{4.32}$$

where $\sigma_0^2 = (\sigma_1^2)/p^2 > 0$.

We next give a proof of (4.19). From (4.22) and (4.24), the upper bound follows: for any $y > 0$,

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log \| \wedge^p G_n \|}{b_n} \leqslant -y \right) \leqslant -\frac{y^2}{2\sigma_1^2}. \tag{4.33}$$

To prove the lower bound, observe that from (4.22) we have

$$\begin{aligned}\mathbb{P}\big(\log\|\wedge^p G_n\| \leqslant -yb_n\big) &= \mathbb{P}\big(\log\|G'_n\| \leqslant -yb_n, \log\|G''_n\| \leqslant -yb_n\big) \\ &\geqslant \mathbb{P}\big(\log\|G'_n\| \leqslant -yb_n\big) - \mathbb{P}(\log\|G''_n\| > -yb_n).\end{aligned}$$

Similarly to (4.30), with fixed integer $m \geqslant 1$ and $u_m = -\mathbb{E}(\log\|G''_m\|) > 0$, taking into account (4.27), we write

$$\begin{aligned}\mathbb{P}(\log\|G''_n\| > -yb_n) &\leqslant \mathbb{P}\Big(\log\|G''_n(G''_{km})^{-1}\| > k\frac{u_m}{2} - yb_n\Big) \\ &\quad + \mathbb{P}\Big(\log\|G''_{km}(G''_{(k-1)m})^{-1}\| + \cdots + \log\|G''_m\| + ku_m > k\frac{u_m}{2}\Big).\end{aligned}$$

In an analogous way as in the proof of (4.29), by Markov's inequality and the fact that $k = O(n)$ and $b_n = o(n)$, the first term on the right hand side of the above inequality is bounded by Ce^{-ck}. It has been shown in the proof of (4.30) that the second term is also bounded by Ce^{-ck}. Therefore, taking into account $k \geqslant n/(m+1)$, we get

$$\mathbb{P}(\log\|G''_n\| > -yb_n) \leqslant Ce^{-cn}.$$

Combining this bound with (4.24), we obtain

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log\|\wedge^p G_n\|}{b_n} \leqslant -y\right) \geqslant -\frac{y^2}{2\sigma_1^2}.$$

By Lemma 4.3, this implies

$$\liminf_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{\log\|G_n\|}{b_n} \leqslant -y\right) \geqslant -\frac{y^2}{2\sigma_0^2}, \tag{4.34}$$

where $\sigma_0^2 = (\sigma_1^2)/p^2 > 0$. Putting together (4.33) and (4.34), we conclude the proof of (4.19). $\square$

Proof of (2.6) *of Theorem 2.4.* Using Lemma 3.2, we can obtain (2.6) from (2.5) just as we obtained (2.4) from (2.3). The details are omitted. $\square$

5. Moderate deviation expansions

This section is devoted to proving Theorems 2.6 and 2.7 about Cramér type moderate deviation expansions in the normal range, for the operator norm $\|G_n\|$ and the spectral radius $\rho(G_n)$.

Proof of Theorem 2.6. By (2.1), there exists a constant $C > 0$ such that for all $n \geqslant 1$, $x \in \mathbb{P}(\mathbb{R}^d)$, $y > 0$ and $\varphi \in \mathcal{B}_\gamma$,

$$\left|\frac{\mathbb{E}\big[\varphi(X_n^x)\mathbb{1}_{\{\log\|G_n\| - n\lambda \leqslant -\sqrt{n}\sigma y\}}\big]}{\Phi(-y)} - \nu(\varphi)\right| \leqslant C\frac{\log n}{\sqrt{n}\Phi(-y)}\|\varphi\|_\gamma. \tag{5.1}$$

Using the elementary inequality

$$\frac{1}{\sqrt{2\pi}}\Big(\frac{1}{y}-\frac{1}{y^3}\Big)e^{-\frac{y^2}{2}} \leqslant \Phi(-y) \leqslant \frac{1}{\sqrt{2\pi}y}e^{-\frac{y^2}{2}} \quad \text{for } y>0, \tag{5.2}$$

it is easy to see that $\frac{\log n}{\sqrt{n}\Phi(-y)} = O(n^{-3/8}(\log n)^{3/2}) \to 0$, as $n \to \infty$, uniformly in $y \in [0, \frac{1}{2}\sqrt{\log n}]$. Therefore, from (5.1) we see that the expansion (2.9) holds uniformly in $y \in [0, \frac{1}{2}\sqrt{\log n}]$. In the same way, using (2.1) together with the fact that $|\mathbb{E}\varphi(X_n^x) - \nu(\varphi)| \leqslant Ce^{-cn}\|\varphi\|_\gamma$, one can also verify that the expansion (2.8) also holds uniformly in $y \in [0, \frac{1}{2}\sqrt{\log n}]$.

It remains to prove that the expansions (2.8) and (2.9) hold uniformly in $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$. We only give a proof of (2.9), since (2.8) can be established in a similar way. Without loss of generality we assume that the function φ is non-negative. For $x \in \mathbb{P}(\mathbb{R}^d)$ and $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$, denote

$$I_n := \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\log\|G_n\| - n\lambda \leqslant -\sqrt{n}\sigma y\big\}}\Big].$$

Using the moderate deviation expansion ([26, Theorem 2.3]) for the norm cocycle $\sigma(G_n, x)$, and the fact that $\sigma(G_n,x) \leqslant \log\|G_n\|$, the upper bound of I_n follows: there exists a constant $C>0$ such that for all $n \geqslant 1$, $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$ and $\varphi \in \mathcal{B}_\gamma$,

$$\frac{I_n}{\Phi(-y)} \leqslant \nu(\varphi) + C\|\varphi\|_\gamma \frac{y+1}{\sqrt{n}}. \tag{5.3}$$

For the lower bound of I_n, we shall use Lemma 3.1. For any $a>0$, $n > k \geqslant 1$ and $v \in \mathbb{R}^d \setminus \{0\}$, consider the event

$$A_{n,k} = \Big\{\Big|\log\|G_n\| - \log\frac{|G_n v|}{|G_k v|} - \log\|G_k\|\Big| \leqslant e^{-ak}\Big\},$$

and we write $A_{n,k}^c$ for its complement. By Lemma 3.1, for any $a>0$, there exist $c_1>0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\begin{aligned}
I_n &\geqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\log\|G_n\| - n\lambda \leqslant -\sqrt{n}\sigma y\big\}}\mathbb{1}_{A_{n,k}}\Big] \\
&\geqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\sigma(G_n,x)-\sigma(G_k,x)+\log\|G_k\| - n\lambda + e^{-ak} \leqslant -\sqrt{n}\sigma y\big\}}\mathbb{1}_{A_{n,k}}\Big] \\
&\geqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\sigma(G_n,x)-\sigma(G_k,x)+\log\|G_k\| - n\lambda + e^{-ak} \leqslant -\sqrt{n}\sigma y\big\}}\Big] - e^{-c_1 k}\|\varphi\|_\infty \\
&=: J_n - e^{-c_1 k}\|\varphi\|_\infty.
\end{aligned} \tag{5.4}$$

As before, for any $n > k \geqslant 1$, we write $G_n = G_{n,k}G_k$ with $G_{n,k} = g_n \dots g_{k+1}$ and $G_k = g_k \dots g_1$. We take the conditional expectation with respect to the filtration $\mathscr{F}_k = \sigma(g_1, \dots, g_k)$ and use (3.3) to obtain that, for any $q > \lambda$, there exists a constant $c_2 > 0$ such that for any $x \in \mathbb{P}(\mathbb{R}^d)$,

$$\begin{aligned}
J_n &\geqslant \mathbb{E}\left\{\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\sigma(G_n,x)-\sigma(G_k,x)+\log\|G_k\|-n\lambda+e^{-ak}\leqslant-\sqrt{n}\sigma y\right\}}\mathbb{1}_{\{\log\|G_k\|\leqslant kq\}}\Big|\mathscr{F}_k\Big]\right\} \\
&\geqslant \mathbb{E}\left\{\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\left\{\sigma(G_n,x)-\sigma(G_k,x)+kq-n\lambda+e^{-ak}\leqslant-\sqrt{n}\sigma y\right\}}\Big|\mathscr{F}_k\Big]\right\} - e^{-c_2 k}\|\varphi\|_\infty \\
&=: J_n' - e^{-c_2 k}\|\varphi\|_\infty. \qquad (5.5)
\end{aligned}$$

For brevity, we set

$$y_1 = y\sqrt{\frac{n}{n-k}} - \frac{k(q-\lambda)}{\sigma\sqrt{n-k}} - \frac{e^{-ak}}{\sigma\sqrt{n-k}}, \qquad (5.6)$$

then J_n' can be rewritten as

$$J_n' = \mathbb{E}\left\{\mathbb{E}\Big[\varphi(G_{n,k}\cdot X_k^x)\mathbb{1}_{\left\{\sigma(G_{n,k},X_k^x)-(n-k)\lambda\leqslant-\sqrt{n-k}\sigma y_1\right\}}\Big|\mathscr{F}_k\Big]\right\}.$$

For any $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$, we take $k = \lfloor C_1 y^2 \rfloor$, where $C_1 > 0$ is a constant to be chosen large enough. From (5.6), we see that $y \sim y_1 = o(n^{1/6})$ as $n \to \infty$. Hence, using the moderate deviation expansion ([26, Theorem 2.3]) for the norm cocycle $\sigma(G_n, x)$, we obtain that as $n \to \infty$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$ and $\varphi \in \mathcal{B}_\gamma$,

$$\frac{J_n'}{\Phi(-y_1)} = \nu(\varphi) + \|\varphi\|_\gamma O\Big(\frac{y_1+1}{\sqrt{n}}\Big). \qquad (5.7)$$

Using the asymptotic expansion $\sqrt{2\pi}\Phi(-y) = \frac{1}{y}e^{-\frac{y^2}{2}}[1 + O(\frac{1}{y^2})]$ as $y \to \infty$ (cf. (5.2)), we get that as $n \to \infty$, uniformly in $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$,

$$\frac{\Phi(-y)}{\Phi(-y_1)} = \frac{y_1}{y}e^{-\frac{y^2}{2}+\frac{y_1^2}{2}}(1+o(1)).$$

Taking into account (5.6) and the definition of k, one can find that, as $n \to \infty$, uniformly in $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$, we have $\frac{y_1}{y} = 1 + o(1)$ and $e^{-\frac{y^2}{2}+\frac{y_1^2}{2}} = 1 + o(1)$. Consequently, substituting the above estimates into (5.7), we get that, as $n \to \infty$, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$ and $\varphi \in \mathcal{B}_\gamma$,

$$\frac{J_n'}{\Phi(-y)} = \nu(\varphi) + \|\varphi\|_\gamma o(1).$$

This, together with (5.5), implies that

$$\frac{I_n}{\Phi(-y)} \geqslant \nu(\varphi) + \|\varphi\|_\gamma o(1) - \|\varphi\|_\infty \frac{2e^{-c_3\lfloor C_1 y^2\rfloor}}{\Phi(-y)}.$$

Using (5.2) and taking $C_1 > \frac{1}{c_3}$, it follows that, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$ and $\varphi \in \mathcal{B}_\gamma$,

$$\frac{I_n}{\Phi(-y)} \geqslant \nu(\varphi) + \|\varphi\|_\gamma o(1) - 4\|\varphi\|_\infty y e^{-c_3 \lfloor C_1 y^2 \rfloor + \frac{y^2}{2}} = \nu(\varphi) + \|\varphi\|_\gamma o(1).$$

Combining this with the upper bound (5.3) ends the proof of (2.9). □

We proceed to establish Theorem 2.7 based on Theorem 2.6, Lemma 3.2 and the Berry-Esseen type bound (2.2).

Proof of Theorem 2.7. We only prove the first expansion (2.10) since the proof of the second one (2.11) can be carried out in an analogous way.

We first remark that for the range of small values $y \in [0, \frac{1}{2}\sqrt{\log n}]$, the moderate deviation expansion (2.10) is a direct consequence of the Berry-Esseen type bound (2.2). Indeed, from (2.2) and the fact that $|\mathbb{E}\varphi(X_n^x) - \nu(\varphi)| \leqslant Ce^{-cn}\|\varphi\|_\gamma$, we derive that uniformly in $y > 0$,

$$\left| \frac{\mathbb{E}\big[\varphi(X_n^x)\mathbb{1}_{\{\log \rho(G_n) - n\lambda \geqslant \sqrt{n}\sigma y\}}\big]}{1 - \Phi(y)} - \nu(\varphi) \right| \leqslant C \frac{\log n}{\sqrt{n}(1 - \Phi(y))} \|\varphi\|_\gamma.$$

Using the inequality (5.2), Using the inequality (5.2), one can verify that $\frac{\log n}{\sqrt{n}(1-\Phi(y))} \to 0$, as $n \to \infty$, uniformly in $y \in [0, \frac{1}{2}\sqrt{\log n}]$. Hence the expansion (2.10) holds uniformly in $y \in [0, \frac{1}{2}\sqrt{\log n}]$.

Now we prove that (2.10) holds uniformly in $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$. Without loss of generality, we assume that the target function φ is non-negative. For brevity, we denote for $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$,

$$I_n := \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\{\log \rho(G_n) - n\lambda \geqslant \sqrt{n}\sigma y\}}\Big].$$

The proof consists of establishing upper and lower bounds.

For the upper bound, we use (2.8) and the inequlility $\rho(G_n) \leqslant \|G_n\|$, to get that, as $n \to \infty$, uniformly in $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$,

$$\frac{I_n}{1 - \Phi(y)} \leqslant \nu(\varphi) + \|\varphi\|_\gamma o(1). \tag{5.8}$$

For the lower bound, we shall apply Lemma 3.2 for a precise comparison between $\rho(G_n)$ and $\|G_n\|$. For any $\varepsilon > 0$ and $n > k \geqslant 1$, we denote

$$A_{n,k} = \Big\{ \log \rho(G_n) - \log \|G_n\| > -\varepsilon k \Big\}.$$

From Lemma 3.2 we know that for any $\varepsilon > 0$, there exist $c_1 > 0$ and $k_0 \in \mathbb{N}$ such that for all $n \geqslant k \geqslant k_0$, we have $\mathbb{P}(A_{n,k}) > 1 - e^{-c_1 k}$. Thus,

$$\begin{aligned} I_n &\geqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\log \rho(G_n) - n\lambda \geqslant \sqrt{n}\sigma y\big\}}\mathbb{1}_{A_{n,k}}\Big] \\ &\geqslant \mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\log \|G_n\| - n\lambda \geqslant \sqrt{n}\sigma y + \varepsilon k\big\}}\Big] - e^{-c_1 k}\|\varphi\|_\infty. \end{aligned} \tag{5.9}$$

By Theorem 2.6, we have, uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$ and $\varphi \in \mathcal{B}_\gamma$,

$$\frac{\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\log \|G_n\| - n\lambda \geqslant \sqrt{n}\sigma y + \varepsilon k\big\}}\Big]}{1 - \Phi(y + \frac{\varepsilon k}{\sqrt{n}\sigma})} = \nu(\varphi) + \|\varphi\|_\gamma o(1). \tag{5.10}$$

For $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$, we take $k = \lfloor C_1 y^2 \rfloor$, where $C_1 > 0$ is a constant to be chosen large enough. Since $\int_y^\infty e^{-\frac{t^2}{2}}dt = \frac{1}{y}e^{-\frac{y^2}{2}}[1 + O(\frac{1}{y^2})]$ as $y \to \infty$, we infer that as $n \to \infty$, uniformly in $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$,

$$\frac{1 - \Phi(y + \frac{\varepsilon k}{\sqrt{n}\sigma})}{1 - \Phi(y)} = \frac{y}{y + \frac{\varepsilon k}{\sqrt{n}\sigma}} \exp\Big\{ - y\frac{\varepsilon k}{\sqrt{n}\sigma} - \frac{\varepsilon^2 k^2}{2n\sigma^2}\Big\}(1 + o(1)).$$

Since $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$, taking into account that $k = \lfloor C_1 y^2 \rfloor$, we get

$$\frac{y}{y + \frac{\varepsilon k}{\sqrt{n}\sigma}} = 1 - \frac{\frac{\varepsilon k}{\sqrt{n}\sigma}}{y + \frac{\varepsilon k}{\sqrt{n}\sigma}} \geqslant 1 - \frac{\varepsilon k}{y\sqrt{n}\sigma} = 1 - \frac{\varepsilon\lfloor C_1 y^2\rfloor}{y\sqrt{n}\sigma} = 1 + o(1),$$

and

$$\exp\Big\{ - y\frac{\varepsilon k}{\sqrt{n}\sigma} - \frac{\varepsilon^2 k^2}{2n\sigma^2}\Big\} = \exp\Big\{ - y\frac{\varepsilon\lfloor C_1 y^2\rfloor}{\sqrt{n}\sigma} - \frac{\varepsilon^2\lfloor C_1 y^2\rfloor^2}{2n\sigma^2}\Big\} = 1 + o(1).$$

Hence, substituting the above estimates into (5.10), we get

$$\frac{\mathbb{E}\Big[\varphi(X_n^x)\mathbb{1}_{\big\{\log \|G_n\| - n\lambda \geqslant \sqrt{n}\sigma y + \varepsilon k\big\}}\Big]}{1 - \Phi(y)} \geqslant \nu(\varphi) + \|\varphi\|_\gamma o(1).$$

This, together with (5.9), implies the lower bound for I_n: uniformly in $x \in \mathbb{P}(\mathbb{R}^d)$, $y \in [\frac{1}{2}\sqrt{\log n}, o(n^{1/6})]$ and $\varphi \in \mathcal{B}_\gamma$,

$$\frac{I_n}{1 - \Phi(y)} \geqslant \nu(\varphi) + \|\varphi\|_\gamma o(1) - \frac{e^{-c_1 k}}{1 - \Phi(y)}\|\varphi\|_\infty \geqslant \nu(\varphi) + \|\varphi\|_\gamma o(1), \tag{5.11}$$

where in the last inequality we take $C_1 > \frac{1}{2c_1}$ and use (5.2).

Combining (5.11) with (5.8) finishes the proof of Theorem 2.7. $\square$

Acknowledgments

We would like to thank the referees and the editors for their helpful comments and remarks. The work was supported by the Centre Henri Lebesgue (CHL, ANR-11-LABX-0020-01), and the National Natural Science Foundation of China (Grants Nos. 11971063 and 11731012).

References

[1] Aoun R.: The central limit theorem for eigenvalues. *Proceedings of the American Mathematical Society*, 149(2): 859-873, 2021.

[2] Aoun, R., Sert, C.: Law of large numbers for the spectral radius of random matrix products. *American Journal of Mathematics*, 143(3): 995-1010, 2021.

[3] Benoist Y., Quint J. F.: Stationary measures and invariant subsets of homogeneous spaces (II). *Journal of the American Mathematical Society*, 26(3): 659-734, 2013.

[4] Benoist Y., Quint J. F.: Central limit theorem for linear groups. *The Annals of Probability*, 44(2): 1308-1340, 2016.

[5] Benoist Y., Quint J. F.: Random walks on reductive groups. *Springer International Publishing*, 2016.

[6] Bougerol P.: Stabilité en probabilité des équations différentielles stochastiques linéaires et convergence de produits de matrices aléatoires. *Comptes rendus des séances de l'Académie des sciences.* Série 1, Mathématique, 299(13): 631-634, 1984.

[7] Bougerol P., Lacroix J.: Products of random matrices with applications to Schrödinger operators. *Birkhäuser Boston*, 1985.

[8] Boulanger A., Mathieu P., Sert C., Sisto A.: Large deviations for random walks on hyperbolic spaces. *Annales Scientifiques de l'École Normale Supérieure*, to appear, see also arXiv:2008.02709, 2020.

[9] Buraczewski D., Damek E., Guivarc'h Y., Mentemeier S.: On multidimensional Mandelbrot cascades. *Journal of Difference Equations and Applications*, 20(11): 1523-1567, 2014.

[10] Breuillard E.: A non concentration estimate for random matrix products. Unpublished notes available at https://www. math. u-psud. fr/breuilla/RandomProducts2. pdf.

[11] Chevalley C.: Théorie des groupes de Lie. Tome II. Groupes algébriques. (French) Actualités Sci. Ind. no. 1152. Hermann, Paris, vii+189 pp, 1951.

[12] Cuny C., Dedecker J., Jan C.: Limit theorems for the left random walk on GL(d,$\mathbb{R}$). *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53(4): 1839-1865, 2017.

[13] Cuny C., Dedecker J., Merlevède F.: Rates of convergence in invariance principles for random walks on linear groups via martingale methods. *Transactions of the American Mathematical Society*, 374(1): 137-174, 2021.

[14] Furstenberg H., Kesten H.: Products of random matrices. *The Annals of Mathematical Statistics*, 31(2): 457-469, 1960.

[15] Guivarc'h Y.: Produits de matrices aléatoires et applications aux propriétés géométriques des sous-groupes du groupe linéaire. *Ergodic Theory and Dynamical Systems*, 10(3): 483-512, 1990.

[16] Guivarc'h Y., Le Page É.: Spectral gap properties for linear random walks and Pareto's asymptotics for affine stochastic recursions. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 52(2): 503-574, 2016.

[17] Guivarc'h Y., Raugi A.: Frontiere de Furstenberg, propriétés de contraction et théorèmes de convergence. *Probability Theory and Related Fields*, 69(2): 187-242, 1985.

[18] He W., Lakrec T., Lindenstrauss E.: Affine random walks on the torus. *International Mathematics Research Notices*, to appear, see also arXiv:2003.03743, 2020.

[19] Hennion H.: Limit theorems for products of positive random matrices. *The Annals of Probability*, 25(4): 1545-1587, 1997.

[20] Huang C., Liu Q.: Moments, moderate and large deviations for a branching process in a random environment. *Stochastic Processes and their Applications*, 122(2): 522-545, 2012.

[21] Kesten H.: Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131(1): 207-248, 1973.

[22] Kingman J. F. C.: Subadditive ergodic theory. *The Annals of Probability*, 1(6): 883-899, 1973.

[23] Le Page É.: Théorèmes limites pour les produits de matrices aléatoires. In Probability measures on groups. *Springer Berlin Heidelberg*, 258-303, 1982.

[24] Sert C.: Large deviation principle for random matrix products. *The Annals of Probability*, 47(3): 1335-1377, 2019.

[25] Xiao, H., Grama, I., Liu, Q.: Precise large deviation asymptotics for products of random matrices. *Stochastic Processes and their Applications*, 130(9): 5213-5242, 2020.

[26] Xiao, H., Grama, I., Liu, Q.: Berry-Esseen bound and precise moderate deviations for products of random matrices. *Journal of the European Mathematical Society*, to appear. See also *arXiv*:1907.02438, 2019.

[27] Xiao, H., Grama, I., Liu, Q.: Berry-Esseen bounds and moderate deviations for the norm, entries and spectral radius of products of positive random matrices. *arXiv*: 2010.00557, 2020.

Current address, Xiao, H.: Universität Hildesheim, Institut für Mathematik und Angewandte Informatik, Hildesheim, Germany.

Email address: `xiao@uni-hildesheim.de`

Current address, Grama, I.: Université de Bretagne-Sud, LMBA UMR CNRS 6205, Vannes, France

Email address: `ion.grama@univ-ubs.fr`

Current address, Liu, Q.: Université de Bretagne-Sud, LMBA UMR CNRS 6205, Vannes, France

Email address: `quansheng.liu@univ-ubs.fr`