

Dmitry Dolgopyat and Omri Sarig

# Local limit theorems for inhomogeneous Markov chains



# Contents

<b>Notation</b>	3
<b>Preface</b>	7
<b>1 Additive functionals on Markov arrays</b>	19
1.1 The basic setup	19
1.1.1 Inhomogeneous Markov chains	19
1.1.2 Inhomogeneous Markov arrays	22
1.1.3 Additive functionals	23
1.2 Uniform ellipticity	24
1.2.1 The definition	24
1.2.2 Contraction estimates and exponential mixing	25
1.2.3 Hitting probabilities and bridge probabilities	29
1.3 Structure constants	30
1.3.1 Hexagons, balance, and structure constants	30
1.3.2 The ladder process	34
1.3.3 $\gamma$ -step ellipticity conditions	36
1.3.4 Uniform ellipticity and strong mixing conditions	37
1.4 Notes and references	39
<b>2 Variance growth, center-tightness, and the central limit theorem</b>	41
2.1 Main results	41
2.2 Proofs	44
2.2.1 The Gradient Lemma	44
2.2.2 The estimate for $\text{Var}(S_N)$	47
2.2.3 Characterization of center-tight additive functionals	49
2.2.4 McLeish's martingale central limit theorem	49
2.2.5 Proof of Dobrushin's central limit theorem	54
2.2.6 Almost sure convergence for sums of functionals with summable variance	57
2.2.7 Convergence of moments	59
2.3 Notes and references	62
<b>3 The essential range and irreducibility</b>	63
3.1 Definitions and motivation	63
3.2 Main results	64
3.2.1 Results for Markov chains	64
3.2.2 Results for Markov arrays	65
3.2.3 Hereditary arrays	66
3.3 Proofs	68
3.3.1 Reduction lemmas	68
3.3.2 The possible values of the co-range	78
3.3.3 Calculation of the essential range	79
3.3.4 Existence of irreducible reductions	80

3.3.5	Proofs of results on hereditary arrays . . . . .	81
3.4	Notes and references . . . . .	83
<b>4</b>	<b>The local limit theorem in the irreducible case . . . . .</b>	<b>85</b>
4.1	Main results . . . . .	85
4.1.1	Local limit theorems for Markov chains . . . . .	85
4.1.2	Local limit theorems for Markov arrays . . . . .	86
4.1.3	Mixing local limit theorems . . . . .	87
4.2	Proofs . . . . .	87
4.2.1	Characteristic functions . . . . .	88
4.2.2	The LLT in the irreducible non-lattice case . . . . .	92
4.2.3	The LLT for the irreducible lattice case . . . . .	95
4.2.4	The mixing LLT . . . . .	96
4.3	Notes and references . . . . .	98
<b>5</b>	<b>The local limit theorem in the reducible case . . . . .</b>	<b>99</b>
5.1	Main results . . . . .	99
5.1.1	Heuristics and warm up examples . . . . .	99
5.1.2	The LLT in the reducible case . . . . .	101
5.1.3	Irreducibility as a necessary condition for the mixing LLT . . . . .	103
5.1.4	Universal bounds for $\mathbb{P}_x[S_N - z_N \in (a, b)]$ . . . . .	103
5.2	Proofs . . . . .	104
5.2.1	Characteristic functions . . . . .	104
5.2.2	Proof of the LLT in the reducible case . . . . .	111
5.2.3	Necessity of the irreducibility assumption . . . . .	116
5.2.4	Universal bounds for Markov chains . . . . .	120
5.2.5	Universal bounds for Markov arrays . . . . .	123
5.3	Notes and references . . . . .	127
<b>6</b>	<b>Local limit theorems for large and moderate deviations . . . . .</b>	<b>129</b>
6.1	The moderate deviations and large deviations regimes . . . . .	129
6.2	Local limit theorems for large deviations . . . . .	130
6.2.1	The log moment generating functions . . . . .	130
6.2.2	The rate functions . . . . .	132
6.2.3	The LLT for moderate deviations . . . . .	133
6.2.4	The LLT for large deviations . . . . .	133
6.3	Proofs . . . . .	136
6.3.1	Strategy of proof . . . . .	136
6.3.2	A parameterized family of changes of measure . . . . .	137
6.3.3	Choosing the parameters . . . . .	144
6.3.4	The asymptotic behavior of $\tilde{V}^{\xi_N}(S_N)$ . . . . .	150
6.3.5	Asymptotics of the log moment generating functions . . . . .	155
6.3.6	Asymptotics of the rate functions . . . . .	157
6.3.7	The local limit theorem for large deviations . . . . .	159
6.3.8	Rough bounds in the reducible case . . . . .	162
6.4	Large deviations threshold . . . . .	162
6.5	Notes and references . . . . .	167
<b>7</b>	<b>Miscellaneous examples and special cases . . . . .</b>	<b>169</b>
7.1	Homogenous Markov chains . . . . .	169
7.2	Perturbations of homogeneous chains . . . . .	174
7.3	Small additive functionals . . . . .	176
7.4	Equicontinuous additive functionals . . . . .	177
7.4.1	Range . . . . .	177
7.4.2	Large deviation threshold . . . . .	178
7.5	Notes and references . . . . .	180

Contents	3
<b>8 LLT for Markov chains in random environment</b> .....	183
8.1 Markov chains in random environment .....	183
8.2 Main results .....	186
8.3 Proofs .....	189
8.3.1 The essential range is a.s. constant .....	189
8.3.2 Variance growth .....	190
8.3.3 The local limit theorem .....	193
8.3.4 Log-moment generating functions and rate functions .....	197
8.4 Notes and references .....	202
References .....	204
<b>Index</b> .....	209





## Notation

$\nabla a$	the additive functional $\{a_{n+1}^{(N)}(X_{n+1}^{(N)}) - a_n^{(N)}(X_n^{(N)})\}$ (a gradient)
$\mathcal{B}(\mathfrak{S})$	the Borel $\sigma$ -algebra of a separable complete metric space $\mathfrak{S}$
$c^-, c^+$	large deviations threshold, see §6.4
$C_c(\mathbb{R})$	the space of continuous $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ with compact support
$C_{mix}$	the mixing constant from Proposition 1.11
$\text{Cov}$	the covariance
$\text{CVar}$	the circular variance, see §3.3.1
$d_n(\xi), d_n^{(N)}(\xi)$	structure constants, see §1.3
$D_N(\xi)$	structure constants, see §1.3
$\delta(\pi)$	the contraction coefficient of a Markov operator $\pi$ , see §1.2.2
$\delta(f)$	the graininess constant of $f$ , see chapter 3
$\varepsilon_0$	(usually) the uniform ellipticity constant, see §1.2.1
$\mathbb{E}, \mathbb{E}_x$	the expectation operator. $\mathbb{E}_x := \mathbb{E}(\cdot   X_1 = x)$
$\text{ess sup}$	the essential supremum, see chapter 1
$f, g, h$	additive functionals
$f_n, f_n^{(N)}$	an entry of an additive functional $f$ of a Markov chain or array
$\mathcal{F}_N(\xi)$	the normalized log-moment generating function, see chapter 6
$G_{alg}(X, f)$	the algebraic range, see chapter 3
$G_{ess}(X, f)$	the essential range, see chapter 3
$\Gamma$	the balance (of a hexagon), see §1.3.1
$H(X, f)$	the co-range, see chapter 3
$\text{Hex}(N, n)$	the space of level $N$ hexagons at position $n$ , see §1.3.1
$\mathcal{I}_N(\eta)$	the rate function, see chapter 6
$k_N$	(usually) the length of the $N$ -th row of an array, minus one
$\mu(dx)$	a measure with its integration variable
$\log$	the natural logarithm (same as $\ln$ )
$\mathbb{N}$	$\{1, 2, 3, \dots\}$
$\text{Osc}$	the oscillation, see §1.2.2
$(\Omega, \mathcal{F}, \mu, T)$	a measurable map $T : \Omega \rightarrow \Omega$ on a measure space $(\Omega, \mathcal{F}, \mu)$
$\mathbb{P}(A), \mathbb{P}_x(A)$	the probability of the event $A$ . $\mathbb{P}_x(A) := \mathbb{P}(A   X_1 = x)$
$\pi_{n,n+1}(x, dy)$	the $n$ -th transition kernel of a Markov chain
$p_n(x, y)$	(usually) the density of $\pi_{n,n+1}(x, dy)$
$\Phi_N(x, \xi)$	characteristic functions, see §4.2.1
$S_N$	$\sum_{i=1}^N f_i(X_i, X_{i+1})$ (chains), or $\sum_{i=1}^{k_N} f_i^{(N)}(X_i^{(N)}, X_{i+1}^{(N)})$ (arrays)
$\text{sgn}(x)$	the sign of $x$ : $(+1)$ when $x > 0$ , $(-1)$ when $x < 0$ , and 0 for $x = 0$
$\mathfrak{S}_n, \mathfrak{S}_n^{(N)}$	the state space of $X_n$ (chains) or of $X_n^{(N)}$ (arrays)
$u_n, u_n^{(N)}, U_N$	structure constants, see §1.3
$\text{Var}$	the variance
$V_N$	the variance of $S_N$
$X_n, X_n^{(N)}$	an entry of a Markov chain, or the $N$ -th row of a Markov array
$X$	a Markov chain or a Markov array
$z_N$	(usually) a real number not too far from $\mathbb{E}(S_N)$

a.e.	almost everywhere
a.s.	almost surely
TFAE	the following are equivalent
s.t.	such that
w.l.o.g.	without loss of generality
$\emptyset$	the empty set
$\because$	because
$\therefore$	therefore
$1_E$	the indicator function of the set $E$ , equal to 1 on $E$ and to zero elsewhere
$a \pm \varepsilon$	a quantity inside $[a - \varepsilon, a + \varepsilon]$
$e^{\pm \varepsilon} a$	a quantity in $[e^{-\varepsilon} a, e^{\varepsilon} a]$
$\sim$	$a_n \sim b_n \Leftrightarrow a_n/b_n \xrightarrow{n \rightarrow \infty} 1$
$\asymp$	$a_n \asymp b_n \Leftrightarrow 0 < \liminf(a_n/b_n) \leq \limsup(a_n/b_n) < \infty$
$\lesssim$	$a_n \lesssim b_n \Leftrightarrow \limsup(a_n/b_n) < \infty$
$\ll$	for measures: $\mu \ll \nu$ means " $\nu(E) = 0 \Rightarrow \mu(E) = 0$ for all measurable $E$ ; For numbers: non-rigorous shorthand for "much smaller than"
$\approx$	non-rigorous shorthand for "approximately equal"
$\stackrel{!}{=}$	is defined to be equal to
$\stackrel{!}{=}$	an equality that will be justified later
$\stackrel{?}{=}$	a possibly false equality that requires checking

$X_n \xrightarrow[n \rightarrow \infty]{prob} Y$  convergence in probability

$X_n \xrightarrow[n \rightarrow \infty]{dist} Y$  convergence in distribution

$X_n \xrightarrow[n \rightarrow \infty]{L^p} Y$  convergence in  $L^p$

$[S_N > t]$	conditions in brackets indicate the events that the conditions happen. For example, if $\varphi : \mathfrak{S} \rightarrow \mathbb{R}$ , then $[\varphi(\omega) > t] := \{\omega \in \mathfrak{S} : \varphi(\omega) > t\}$
$\lfloor x \rfloor, \lceil x \rceil$	$\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}, \lceil x \rceil := \min\{n \in \mathbb{Z} : n \geq x\}$
$\{x\}, \langle x \rangle$	$\{x\} := x - \lfloor x \rfloor; \langle x \rangle$ is the unique number in $[-\pi, \pi]$ s.t. $x - \langle x \rangle \in 2\pi\mathbb{Z}$
$\{x\}_{t\mathbb{Z}}, [x]_{t\mathbb{Z}}$	$\{x\}_{t\mathbb{Z}} := t\{x/t\}, [x]_{t\mathbb{Z}} := x - \{x\}_{t\mathbb{Z}}$ , so $[x]_{t\mathbb{Z}} \in t\mathbb{Z}$ and $\{x\}_{t\mathbb{Z}} \in [0, t)$

The *Fourier transform* of an  $L^1$  function  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is

$$\widehat{\phi}(\xi) := \int_{\mathbb{R}} e^{-i\xi u} \phi(u) du.$$

The *Legendre-Fenchel transform* of a convex real-valued function  $\phi$  on  $\mathbb{R}$  is

$$\phi^*(\eta) := \sup_{\xi} [\xi \eta - \phi(\xi)].$$



## Preface

### Setup and aim

Our aim is to provide asymptotic formulas for the probabilities  $\mathbb{P}_x[S_N - z_N \in (a, b)]$ , where  $X_n$  is a Markov chain,  $x$  is some initial state,  $\mathbb{P}_x := \mathbb{P}[\cdot | X_1 = x]$ ,

$$S_N = \sum_{n=1}^N f_n(X_n, X_{n+1}), \text{ and } z_N \text{ are real numbers not too far from } \mathbb{E}(S_N).$$

Such results are called **local limit theorems (LLT)**,<sup>1</sup> and they have a long history, see the end of this chapter. The novelty of this work is that we allow the Markov chain to be **inhomogeneous**. This means that we allow the set of states, the transition probabilities, and the summands  $f_n$  to depend on  $n$ .

We will always assume that  $f_n$  are uniformly bounded real-valued functions, and that  $\{X_n\}$  is uniformly elliptic, a technical condition which will be stated in chapter 1, and which implies uniform exponential mixing.

These assumptions place us in the Gaussian domain of attraction. The analogy with classical results for sums of independent identically distributed (iid) random variables suggests that in the best of all situations, we should expect the following (in what follows  $V_N = \text{Var}(S_N)$  and  $A_N \sim B_N \Leftrightarrow A_N/B_N \xrightarrow[N \rightarrow \infty]{} 1$ ):

(1) **Local deviations:** If  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , then  $\mathbb{P}_x[S_N - z_N \in (a, b)] \sim \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} |a - b|$ .

(2) **Moderate deviations:** If  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then

$$\mathbb{P}_x[S_N - z_N \in (a, b)] \sim \frac{e^{-\frac{1+o(1)}{2} \left( \frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \right)^2}}{\sqrt{2\pi V_N}} |a - b|.$$

(3) **Large deviations:** If  $\left| \frac{z_N - \mathbb{E}(S_N)}{V_N} \right| < c$  with  $c > 0$  sufficiently small, then for every  $x$  in the state space of  $X_1$ ,

$$\mathbb{P}_x[S_N - z_N \in (a, b)] \sim \frac{e^{-V_N \mathcal{J}_N(\frac{z_N}{V_N})}}{\sqrt{2\pi V_N}} \cdot \int_a^b e^{-t \xi_N \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right)} dt \cdot \rho_N \left( x, \frac{z_N - \mathbb{E}(S_N)}{V_N} \right), \text{ where}$$

- $\mathcal{J}_N(\cdot)$  are the Legendre transforms of  $\mathcal{F}_N(\xi) := \frac{1}{V_N} \log \mathbb{E}(e^{\xi S_N})$ .

---

<sup>1</sup> By contrast, *central* limit theorems describe  $\mathbb{P}[S_N - z_N \in (a\sqrt{\text{Var}(S_N)}, b\sqrt{\text{Var}(S_N)})]$ .

- $\xi_N : (-c, c) \rightarrow \mathbb{R}$  satisfy  $\xi_N(0) = 0$ ,  $\text{sgn}(\xi_N(\eta)) = \text{sgn}(\eta)$ , and  $\exists C > 0$  independent of  $N$  s.t.  $C^{-1}|\eta| \leq |\xi_N(\eta)| \leq C|\eta|$  for all  $N \in \mathbb{N}$ ,  $|\eta| < c$ .
- $\rho_N(x, t) \xrightarrow[t \rightarrow 0]{} 1$  uniformly in  $N, x$ , and  $\rho_n(\cdot, \cdot)$  are uniformly bounded away from  $0, \infty$  on  $\mathfrak{S}_1 \times [-c, c]$  where  $\mathfrak{S}_1$  is the state space of  $X_1$ .
- $c, \xi_N, \rho_N$  depend on the Markov chain, but not on  $z_N$  or on  $(a, b)$ .

(The asymptotic results in the large deviation regime are more precise than in the moderate deviation case, but less universal. See Chapter 6 for more details.)

Although the asymptotic formulas (1)–(3) above are true in many cases, they do sometime fail — even when  $S_N$  is a sum of iid's. The aim of this work is to give general sufficient conditions for (1)–(3), and to provide the necessary asymptotic corrections when some of these conditions fail. To do this we first identify all the obstructions to (1)–(3), and then we analyze  $S_N$  when these obstructions happen.

## The obstructions to the local limit theorems

The *algebraic range* is the smallest closed additive subgroup  $G \leq \mathbb{R}$  for which there are  $c_n \in \mathbb{R}$  so that  $f_n(X_n, X_{n+1}) - c_n \in G$  almost surely for all  $n$ . We show that the following list is a complete set of obstructions to (1)–(3):

- (I) **Lattice behavior:** The algebraic range is  $t\mathbb{Z}$  with  $t \in \mathbb{R}$ .
- (II) **Center-tightness:**  $\text{Var}(S_N)$  does not tend to infinity. In chapter 2 we will see that in this case  $\text{Var}(S_N)$  must be bounded.
- (III) **Reducibility:**  $f_n(X_n, X_{n+1}) = g_n(X_n, X_{n+1}) + c_n(X_n, X_{n+1})$  where  $\{c_n(X_n, X_{n+1})\}$  is center-tight, and the algebraic range of  $\{g_n(X_n, X_{n+1})\}$  is strictly smaller than the algebraic range of  $\{f_n(X_n, X_{n+1})\}$ .

One of our main results is that (1)–(3) hold whenever (I), (II), (III) fail.

## How to show that the obstructions do not occur

While it is usually easy to rule out the lattice obstruction (I), it is often not clear how to rule out (II) and (III). What is needed is a tool that determines from the data of  $f_n$  and  $X_n$  whether  $\{f_n(X_n, X_{n+1})\}$  is center-tight or reducible.

In chapter 1, we introduce numerical constants  $d_n(\xi)$  ( $n \geq 3, \xi \in \mathbb{R}$ ) which are defined purely in terms of the transition probabilities  $\pi_{n,n+1}(x, E) := \mathbb{P}(X_{n+1} \in E | X_n = x)$  and the functions  $f_n(x, y)$ , and which can be used to determine which obstructions occur and which vanish:

- If  $\sum d_n^2(\xi) = \infty$  for all  $\xi \neq 0$ , then the obstructions (I),(II),(III) do not occur, and the asymptotic expansions (1)–(3) hold.
- If  $\sum d_n^2(\xi) < \infty$  for all  $\xi \neq 0$ , then  $\text{Var}(S_N)$  is bounded (obstruction II).
- If  $\sum d_n^2(\xi) = \infty$  for some but not all  $\xi \neq 0$ , then  $\text{Var}(S_N) \rightarrow \infty$  but we are either lattice or reducible: (II) fails, but at least one of (I),(III) occurs.

We call  $d_n(\xi)$  the *structure constants* of  $X = \{X_n\}$  and  $f = \{f_n\}$ .

## What happens when the obstructions do occur

### (I) The lattice case

The lattice obstruction (I) already happens for sums of iid's, and the classical approach how to adjust (1)–(3) to this setup extends without much difficulty to the inhomogeneous Markov case.

Suppose the algebraic range is  $t\mathbb{Z}$  with  $t \neq 0$ , i.e. there are constants  $c_n$  such that  $f_n(X_n, X_{n+1}) - c_n \in t\mathbb{Z}$  almost surely for all  $n$ . Assume further that  $t\mathbb{Z}$  is the smallest group with this property. In this case

$$S_N \in \gamma_N + t\mathbb{Z} \text{ a.s. for all } N,$$

where  $\gamma_N = \sum_{i=1}^N c_i \pmod{t\mathbb{Z}}$ . Instead of analyzing  $\mathbb{P}_x[S_N - z_N \in (a, b)]$ , which might be equal to zero, we study  $\mathbb{P}_x[S_N - z_N = kt]$ , with  $k \in \mathbb{Z}$  fixed and  $z_N \in \gamma_N + t\mathbb{Z}$ .

We show that in case (I), if the algebraic range is  $t\mathbb{Z}$ , and obstructions (II) and (III) do not occur, then (as in the case of iid's):

(1') If  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ ,  $z_N \in \gamma_N + t\mathbb{Z}$  and  $k \in \mathbb{Z}$ , then  $\mathbb{P}_x[S_N - z_N = kt] \sim \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} |t|$ .

(2') If  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ ,  $z_N \in \gamma_N + t\mathbb{Z}$  and  $k \in \mathbb{Z}$ , then

$$\mathbb{P}_x[S_N - z_N = kt] \sim \frac{1}{\sqrt{2\pi V_N}} e^{-\frac{1+o(1)}{2} \left( \frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \right)^2} |t|.$$

(3') If  $\left| \frac{z_N - \mathbb{E}(S_N)}{V_N} \right| < c$  with  $c > 0$  sufficiently small,  $z_N \in \gamma_N + t\mathbb{Z}$  and  $(a, b) \cap t\mathbb{Z} \neq \emptyset$ , then for every  $x$  in the state space of  $X_1$ ,

$$\mathbb{P}_x[S_N - z_N \in (a, b)] \sim \frac{e^{-V_N \mathcal{I}_N(\frac{z_N}{V_N})}}{\sqrt{2\pi V_N}} \cdot \sum_{\tau \in (a, b) \cap t\mathbb{Z}} |t| e^{-\tau \xi_N(\frac{z_N - \mathbb{E}(S_N)}{V_N})} \cdot \rho_N \left( x, \frac{z_N - \mathbb{E}(S_N)}{V_N} \right)$$

where  $\mathcal{I}_N(\cdot)$ ,  $\rho_N$  and  $\xi_N$  have the properties listed in the non-lattice case (3).

The previous results hold for lattice valued, irreducible, non-center tight additive functionals, that is, when (I) holds and (II),(III) fail. Here is an equivalent condition in terms of the data of  $X_n$  and  $f_n$ :

$$\exists t \neq 0 \text{ s.t. } \sum d_n^2(\xi) < \infty \text{ exactly when } \xi \in \frac{2\pi}{t} \mathbb{Z}.$$

Under this condition, (1')–(3') hold with parameter  $|t|$ .

### (II) The center-tight case

We show that obstruction (II) happens iff  $f_n(X_n, X_{n+1})$  can be put in the form

$$f_n(X_n, X_{n+1}) = a_{n+1}(X_{n+1}) - a_n(X_n) + h_n(X_n, X_{n+1}) + c_n \quad (*)$$

where  $a_n(X_n)$  are uniformly bounded,  $c_n$  are constants,  $h_n(X_n, X_{n+1})$  have mean zero, and  $\sum \text{Var}[h_n(X_n, X_{n+1})] < \infty$ .

The freedom in choosing  $a_n(X_n)$  is too great to allow general statements on the asymptotic behavior of  $\mathbb{P}_x[S_N - z_N \in (a, b)]$ , see Example 2.2.<sup>2</sup> But as we shall see in chapter 2, (\*) does provide us with some almost sure control:

$$S_N = a_{N+1}(X_{N+1}) - a_1(X_1) + \sum_{n=1}^N h_n(X_n, X_{n+1}) + \gamma_N,$$

where  $\gamma_N = \sum_{i=1}^N c_i$ , and  $\sum_{n=1}^{\infty} h_n(X_n, X_{n+1})$  converges almost surely. This means that in the center-tight scenario,  $S_N - \mathbb{E}(S_N)$  can be decomposed into the sum of two terms: A bounded oscillatory term which only depends on  $X_1, X_{N+1}$ , and a term which depends on the entire past  $X_1, \dots, X_{N+1}$  and which converges almost surely.

### (III) The reducible case

In the reducible case, we can decompose

$$f_n(X_n, X_{n+1}) = g_n(X_n, X_{n+1}) + c_n(X_n, X_{n+1}) \quad (**)$$

where  $\{c_n(X_n, X_{n+1})\}$  is center-tight, and the algebraic range of  $\{g_n(X_n, X_{n+1})\}$  is strictly smaller than the algebraic range of  $\{f_n(X_n, X_{n+1})\}$ .

In principle, it is possible that  $\{g_n(X_n, X_{n+1})\}$  is reducible too, but in chapter 5 we show that one can find an “optimal” decomposition (\*\*) where  $\{g_n(X_n, X_{n+1})\}$  is not reducible, and cannot be decomposed further. The algebraic range of the “optimal”  $\{g_n(X_n, X_{n+1})\}$  is the “infimum” of all possible reduced ranges:

$$G_{ess} := \bigcap \left\{ G : \begin{array}{l} G \text{ is the algebraic range of some } \{g_n(X_n, X_{n+1})\} \\ \text{which satisfies (**) with } \{c_n(X_n, X_{n+1})\} \text{ center-tight} \end{array} \right\}.$$

We call  $G_{ess}$  the *essential range* of  $\{f_n\}$ . It can be calculated explicitly from the data of  $f_n$  and  $X_n$  in terms of the structure constants, see Theorem 3.2.

It follows from the definitions that  $G_{ess}$  is a proper closed subgroup of  $\mathbb{R}$ , so  $G_{ess} = \{0\}$  or  $t\mathbb{Z}$  or  $\mathbb{R}$ . In the reducible case,  $G_{ess} = \{0\}$  or  $t\mathbb{Z}$ , because if  $G_{ess} = \mathbb{R}$ , then the algebraic range (which contains  $G_{ess}$ ) is also equal to  $\mathbb{R}$ .

If  $G_{ess} = \{0\}$ , then the optimal  $\{g_n\}$  has algebraic range  $\{0\}$ , and  $g_n$  are constant functions. In this case  $f_n$  is center-tight, and we are back in case (II).

If  $G_{ess} = t\mathbb{Z}$  with  $t \neq 0$ , then  $\{g_n(X_n, X_{n+1})\}$  is lattice, non-center-tight, and irreducible. Therefore

$$S_N = \underbrace{\sum_{n=1}^N g_n(X_n, X_{n+1})}_{S_N(g)} + \underbrace{\sum_{n=1}^N c_n(X_n, X_{n+1})}_{S_N(c)} \quad (\dagger)$$

where  $S_N(g)$  satisfies the lattice local limit theorems (1’)-(3’) with parameter  $t$ , and  $\text{Var}[S_N(c)] = O(1)$ . Trading constants between  $g$  and  $c$ , we can also arrange  $\mathbb{E}(S_N(c)) = O(1)$ .

Unfortunately even though  $\text{Var}[S_N(f)] \rightarrow \infty$  and  $\text{Var}[S_N(c)] = O(1)$ , examples show that  $S_N(c)$  is still powerful enough to disrupt the local limit theorem for  $S_N$ , lattice or non-lattice (example

<sup>2</sup> Throughout this work, Example X.Y is example number Y in chapter X. Similarly for Theorems, Propositions etc.

**5.1).** Heuristically, what happens is that the mass of  $S_N(g)$  concentrates on cosets of  $t\mathbb{Z}$  according to (1')–(3'), but  $S_N(c)$  smudges this mass to a neighborhood of the lattice in a non-universal manner.

This suggests that (1)–(3) should be approximately true for intervals  $(a, b)$  of length  $|a - b| \gg |t|$ , but false for intervals of length  $|a - b| \ll |t|$ . In chapter 5 we prove results in this direction.

For intervals with size  $|a - b| > 2|t|$ , we show that for all  $z_N \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , for all  $N$  large enough

$$\frac{1}{3} \left( \frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}} \right) \leq \mathbb{P}_x[S_N - z_N \in (a, b)] \leq 3 \left( \frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}} \right).$$

If  $|a - b| > L > |t|$ , we can replace 3 by a constant  $C(L)$  such that  $C(L) \xrightarrow[L/|t| \rightarrow \infty]{} 1$ .

For general intervals, possibly with length less than  $|t|$ , we show the following: There are uniformly bounded functions  $b_N(x_1, x_{N+1})$  and a random variable  $\mathfrak{H} = \mathfrak{H}(X_1, X_2, X_3, \dots)$  so that for every  $z_N \in t\mathbb{Z}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , for every  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  continuous with compact support,

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \mathbb{E}_x[\phi(S_N - z_N - b_N(X_1, X_{N+1}))] = \frac{e^{-z^2/2}|t|}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \mathbb{E}_x[\phi(mt + \mathfrak{H})]. \quad (\ddagger)$$

For  $\phi \approx 1_{[a,b]}$  with  $|a - b| \gg |t|$ , the right-hand-side of  $(\ddagger)$  is approximately equal to  $\frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi}}$ , in accordance with (1), see Lemma 5.7. But for  $|a - b| \ll |t|$ , the right-hand-side depends on the essential range  $t\mathbb{Z}$  and on the detailed structure of  $\{c_n(X_n, X_{n+1})\}$  through  $t$ ,  $b_N(X_1, X_{N+1})$  and  $\mathfrak{H}$ .

What are  $b_N(X_N, X_{N+1})$  and  $\mathfrak{H}$ ? Recall that the term  $c_n(X_n, X_{n+1})$  on the right-hand-side of  $(\ddagger)$  is center-tight. As such, it can be put in the form

$$c_n(X_n, X_{n+1}) = a_{n+1}(X_{n+1}) - a_n(X_n) + h_n(X_n, X_{n+1}) + c_n^*,$$

where  $\sup_n(\text{ess sup}|a_n|) < \infty$ ,  $c_n^*$  are constants,  $\mathbb{E}(h_n(X_n, X_{n+1})) = 0$ , and  $\sum h_n$  converges almost surely. Let  $\gamma_N := \sum_{n=1}^N c_n^* = \mathbb{E}(S_N(c)) + O(1) = O(1)$ . The proof of  $(\ddagger)$  shows that

- $b_N = a_{N+1}(X_{N+1}) - a_1(X_1) + \{\gamma_N\}_{t\mathbb{Z}}$ , where  $\{x\}_{t\mathbb{Z}} = |t|\{x/|t|\} = x \pmod{t\mathbb{Z}}$ ;
- $\mathfrak{H} = \sum_{n=1}^{\infty} h_n(X_n, X_{n+1})$ .<sup>3</sup>

This works as follows. Let  $z_N^* := z_N - [\gamma_N]_{t\mathbb{Z}}$ , where  $[x]_{t\mathbb{Z}} := x - \{x\}_{t\mathbb{Z}} \in t\mathbb{Z}$ . Then  $z_N^* \in t\mathbb{Z}$ ,  $\frac{z_N^* - \mathbb{E}(S_N)}{\sqrt{V_N}} = \frac{z_N - \mathbb{E}(S_N) + O(1)}{\sqrt{V_N}} \rightarrow z$ , and

$$S_N - b_N - z_N = [S_N(g) - z_N^*] + S_N(h).$$

By subtracting  $b_N$  from  $S_N$ , we are shifting the distribution of  $S_N$  to the distribution of the sum of two terms: The first,  $S_N(g)$ , is an *irreducible*  $t\mathbb{Z}$ -valued additive functional; and the second,  $S_N(h)$ , converges almost surely to  $\mathfrak{H}$ .

Suppose for the sake of discussion that  $S_N(g), S_N(h)$  were independent, then the lattice LLT for  $S_N(g)$  and the definition of  $\mathfrak{H}$  would imply that

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \mathbb{E}_x[\phi(S_N - b_N - z_N)] = \int_{\mathbb{R}} \phi(x) m(dx),$$

<sup>3</sup> It is possible to replace  $\mathfrak{H}$  by a different random variable  $\mathfrak{F}$  which is bounded, see chapter 5.

where  $m := \frac{e^{-z^2/2}}{\sqrt{2\pi}} m_{t\mathbb{Z}} * m_{\mathfrak{H}}$ , and  $m_{\mathfrak{H}}(E) := \mathbb{P}[\mathfrak{H} \in E]$ ,  $m_{t\mathbb{Z}} := |t| \cdot$ counting measure of  $t\mathbb{Z}$ . Calculating, we find that  $\int_{\mathbb{R}} \phi dm =$ right-hand-side of (‡).

In general,  $S_N(g)$  and  $S_N(h)$  are not independent, and the problem of proving (‡) reduces to the problem of proving that  $S_N(g)$  and  $S_N(h)$  are *asymptotically* independent. This is done in chapter 5.

For further consequences of (‡), including an interpretation in terms of the asymptotic distributional behavior of  $S_N$  modulo  $t\mathbb{Z}$ , see chapter 5.

### Final words on the setup of this work

Before we end the preface, we would like to comment on a choice we made when we wrote this work, specifically, our focus on additive functionals of the form  $f_n = f_n(X_n, X_{n+1})$ .

This choice is somewhat unorthodox: The theory of Markov processes is mostly concerned with the case  $f_n = f_n(X_n)$  (see e.g. [45, 109, 139]), and the theory of stochastic processes is mostly concerned with the case  $f_n = f_n(X_n, X_{n+1}, \dots)$ , under assumptions of weak dependence of  $X_k, X_n$  when  $|k - n| \gg 1$  (see e.g. [74, 130]). We decided to study  $f_n = f_n(X_n, X_{n+1})$  for the following reasons:

- The case  $f_n = f_n(X_n, X_{n+1})$  is richer than the case  $f_n = f_n(X_n)$  because it contains gradients  $a_{n+1}(X_{n+1}) - a_n(X_n)$ . Two additive functionals which differ by a gradient with uniformly bounded  $\text{ess sup} |a_n|$  will have the same CLT behavior, but they may have different LLT behavior, because their algebraic ranges can be different. This leads to an interesting reduction theory which we would have missed had we only considered the case  $f_n = f_n(X_n)$ .<sup>4</sup>
- The case  $f_n(X_n, \dots, X_{n+m})$  with  $m > 1$  can be deduced from the case  $f_n(X_n, X_{n+1})$ , and does not require new ideas, see Example 1.3 and the discussion in §1.3.3. We decided to keep  $m = 1$  and leave the extension to  $m > 1$  to the reader.
- The case  $f_n = f_n(X_n, X_{n+1}, \dots)$  is of great interest, and we hope to address it in the future, but at the moment our results do not cover it.

We hope to stimulate research into the local limit theorem of additive functionals of general non-stationary stochastic processes with mixing conditions. Such work will have applications outside the theory of stochastic processes, such as the theory of dynamical systems. Our work here is a step in this direction.

### Notes and references

**Local limit theorems for sums of iid's.** The first LLT is of course the celebrated de Moivre–Laplace Theorem. De Moivre, in his 1738 book [33], gave approximations for  $\mathbb{P}[a \leq S_n \leq b]$  when  $S_n = X_1 + \dots + X_n$ , and  $X_i$  are iid, equal to zero or one with equal probabilities. Laplace extended de Moivre's results to the case when  $X_i$  are equal to zero or one with non-equal probabilities [89, 90]. Pólya, in 1921, extended these results to the vector valued iid which generate the simple random walk on  $\mathbb{Z}^d$ , and deduced his famous criterion for the recurrence of simple random walks [121].

<sup>4</sup> We cannot reduce the case  $f_n(X_n, X_{n+1})$  to the case  $f_n(Y_n)$  by working with the Markov chain  $Y_n = (X_n, X_{n+1})$  because  $\{Y_n\}$  may no longer satisfy some of our standing assumptions, specifically the uniformly ellipticity condition (see chapter 1).

The next historical landmark is Gnedenko's 1948 work [59, 60] which initiated the study of the LLT for sums of iid with *general* lattice distributions. He asked for the weakest possible assumptions on the distribution of iid's  $X_i$  which lead to a LLT with Gaussian or stable limit. Khinchin popularized the problem by emphasizing its importance to the foundations of quantum statistical physics [77], and it was studied intensively by the Russian school, with important contributions by Linnik, Ibragimov, Prohorov, Richter, Saulis, Petrov and others. We will comment on some of these contributions in later chapters. For the moment, we refer the reader to the excellent books by Gnedenko & Kolmogorov [61], Ibragimov & Linnik [74], and Petrov [117] and the many references they contain.

The early works on the local limit theorem all focused on the lattice case. The Gnedenko–Kolmogorov book [61] contains the first result we are aware of which could be considered to be a non-lattice local limit theorem. The authors assume that each of the iid's  $X_i$  have a probability density function  $p(x) \in L^r$  with finite variance  $\sigma^2$ , and show that the density function  $p_n(x)$  of  $X_1 + \dots + X_n$  satisfies

$$\sigma\sqrt{n}p_n(\sigma\sqrt{n}x) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

There could be non-lattice iid's without density functions, for example the iid's  $X_i$  equal to  $(-1)$ ,  $0$ , or  $\sqrt{2}$  with equal probabilities (the algebraic range is  $\mathbb{R}$ , because the group generated by  $(-1)$  and  $\sqrt{2}$  is dense). Shepp [141] was the first to consider non-lattice LLT in such cases. His approach was to provide asymptotic formulas for  $\mathbb{P}[a \leq S_n - \mathbb{E}(S_N) \leq b]$  for arbitrary intervals  $[a, b]$ , or for

$$\sqrt{2\pi\text{Var}(S_N)}\mathbb{E}[\phi(S_N - \mathbb{E}(S_N))]$$

for all test functions  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  which are continuous with compact support. In this monograph, we use a slight modification of Shepp's formulation of the LLT. Instead of working with  $S_N - \mathbb{E}(S_N)$ , we work with  $S_N - z_N$  subject to the assumptions that  $z_N$  is “not too far” from  $\mathbb{E}(S_N)$ , and that  $S_N - z_N \in$  algebraic range.

Stone proved non-lattice LLT in Shepp's sense for sums of vector valued iid in [147], extending earlier work of Rvačeva [132] who treated the lattice case. These works are important not only because of the intrinsic interest in the vector valued case, but also because of technical innovations which became tools of the trade, see e.g. [17].

**Local limit theorems for stationary stochastic processes.** The earliest local limit theorem for non-iid sequences  $\{X_i\}$  is due to Kolmogorov [83]. He considered stationary homogeneous Markov chains  $\{X_i\}$  with a finite set of states  $\mathfrak{S} = \{a_1, \dots, a_n\}$ , and proved a local limit theorem for the occupation times

$$S_N = \sum_{i=1}^N \vec{f}(X_i), \text{ where } \vec{f}(x) = (1_{a_1}(x), \dots, 1_{a_n}(x)).$$

Following further developments for finite state Markov chains by Sirazhdinov [142], Nagaev [109] was able to obtain a very general local limit theorems for  $S_N = \sum_{i=1}^N f(X_i)$  for a large class of stationary homogeneous countable Markov chains  $\{X_i\}$  and for a variety of unbounded functions  $f$ , both in the gaussian and stable cases.

Nagaev's paper introduced the method of characteristic function operators, which is also applicable outside the context of Markov chains. This opened the way for proving LLT for other weakly dependent stationary stochastic processes, and in particular to time series of probability preserving dynamical systems. Guivarc'h & Hardy [65] proved gaussian local limit theorems for Birkhoff sums  $S_N = \sum_{i=1}^N f(T^i x)$  for Anosov diffeomorphisms  $T : X \rightarrow X$  with an invari-

ant Gibbs measure, and Hölder continuous functions  $f$ . Rousseau-Egele [127] and Broise [19] proved such theorems for piecewise expanding interval map possessing an absolutely continuous invariant measure,  $X = [0, 1]$ , and  $f \in BV$ . Aaronson & Denker [4] gave general LLT for stationary processes generated by Gibbs-Markov maps both in the gaussian and in the non-gaussian domain of attraction. These results have found many applications in infinite ergodic theory, dynamical systems and hyperbolic geometry, see for example [1], [3], [5]. The influence of Nagaev's method can also be recognized in other works on other asymptotic problems in dynamics and geometry, see for example [9], [10], [69], [76], [87], [88], [91], [92], [119], [120], [140]. For the connection between the LLT and the behavior of local times for stationary stochastic processes, see [38, 51].

**Local limit theorems for non-stationary stochastic processes.** The interest in limit theorems for sums of *non-identically* distributed, independent, random variables goes back to the works of Chebyshev [148], Lyapunov [99], and Lindeberg [95] who considered the central limit theorem for such sums.

The study of LLT for sums of non-identically distributed random variables started later, in the works of Prohorov [122] and Rozanov [128]. A common theme in these works and those that followed them is to assume an asymptotic for  $\mathbb{P}[a \leq \frac{S_N - A_N}{B_N} \leq b]$  for suitable normalizing constants  $A_N, B_N$ , and then ask what extra conditions imply an asymptotic for  $\mathbb{P}[a \leq S_N - A_N \leq b]$ .

An important counterexample by Gamerklidze [58] pointed the way towards the phenomenon that the distribution of  $S_N$  may lie close to a proper sub-group of its algebraic range without actually charging it, and a variety of sufficient conditions which rule this out were developed over the years. We mention especially Rozanov's condition in the lattice case [128] (see the end of chapter 3), the Mineka-Silverman condition in the non-lattice case [104], and Statulevicius's condition [146], and conditions motivated by additive number theory such as those appearing in [106] and [107]. For a discussion of these conditions, see [108].

Dolgopyat proved a LLT for sums of non-identically distributed, independent random variables which also applies to the reducible case [49].

Dobrushin proved a general central limit theorem for inhomogeneous Markov chains in [45] (see chapter 2). Local limit theorems for inhomogeneous Markov chains are considered in [145].

Merlevède, M. Peligrad and C. Peligrad proved local limit theorems for sums  $\sum_{i=1}^N f_i(X_i)$  where  $\{X_i\}$  is a  $\psi$ -mixing inhomogeneous Markov chain, under the irreducibility condition of Mineka & Silverman [114]. Hafouta obtained local limit theorems for a class of inhomogeneous Markov chains in [67]. In a different direction, central limit theorems for time-series of inhomogeneous sequences of Anosov diffeomorphisms are proved in [12] and [27].

An important source of examples of inhomogeneous Markov chains is a Markov chain in random environment, when considered for a specific ("quenched") realizations of the environment (see chapter 8). Hafouta & Kifer proved local limit theorems for non-conventional ergodic sums in [68], and local limit theorems for random dynamical systems including Markov chains in random environment in [69]. Demers, Péne & Zhang [36] prove a LLT for an integer valued observable for a random dynamical system.

Comparing the theory of inhomogeneous Markov chains to theory of Markov chains in random environment studied in [69], we note the following differences:

- (a) The theory of inhomogeneous Markov chains applies to fixed realizations of noise and not just to almost every realization of noise;

- (b) In the random environment setup, a center-tight additive functional must be a coboundary, while in the general case it can also have a component with summable variances;
- (c) In the non center-tight random environment setup, the variance grows linearly for a.e. realization of noise. But for a general inhomogeneous Markov chain it can grow arbitrarily slowly.

**The contribution of this work.** The novelty of this work is in providing optimal sufficient conditions for the classical asymptotic formulas for  $\mathbb{P}[S_N - z_N \in (a, b)]$ , and in the analysis of  $\mathbb{P}[S_N - z_N \in (a, b)]$  when these conditions fail.

In particular, we derive a new asymptotic formula for  $\mathbb{P}[S_N - z_N \in (a, b)]$  in the reducible case, subject to assumption that  $V_N := \text{Var}(S_N) \rightarrow \infty$ , and we prove a structure theorem for  $S_N$  in case  $V_N \not\rightarrow \infty$ .

Unlike previous works, our analysis does not require any assumptions on the rate of growth of  $V_N$ , beyond convergence to infinity.

**Acknowledgements:** The work on this monograph was partially supported by the BSF grant 201610. The authors thank the staff of Weizmann Institute for excellent working conditions. O.S. was also partially supported by ISF grant 1149/18. D.D. was partially supported by NSF grant DMS 1665046. The authors are indebted to Manfred Denker, Yuri Kifer, and Ofer Zeitouni for useful discussions and suggestions. The authors are in particularly indebted to Yeor Hafouta for many useful and penetrating comments on the first draft of this work.



# Chapter 1

## Additive functionals on Markov arrays

*This chapter discusses the setup and standing assumptions used in this work.*

### 1.1 The basic setup

#### 1.1.1 Inhomogeneous Markov chains

A **Markov chain** is given by the following data:

- **State spaces:** Borel spaces  $(\mathfrak{S}_n, \mathcal{B}(\mathfrak{S}_n))$  ( $n \geq 1$ ), where  $\mathfrak{S}_n$  is a complete separable metric space, and  $\mathcal{B}(\mathfrak{S}_n)$  is the Borel  $\sigma$ -algebra of  $\mathfrak{S}_n$ .  $\mathfrak{S}_n$  is the set of “the possible states of the Markov chain at time  $n$ .”
- **Transition probabilities:** Borel probability measures  $\pi_{n,n+1}^{(N)}(x, dy)$  on  $\mathfrak{S}_{n+1}$  ( $x \in \mathfrak{S}_n, n \geq 1$ ), so that for every Borel  $E \subset \mathfrak{S}_{n+1}$ , the function  $x \mapsto \pi_{n,n+1}^{(N)}(x, E)$  is measurable. The measure  $\pi_n(x, E)$  is “the probability of event  $E$  at time  $n+1$ , given that the state at time  $n$  was  $x$ .”
- **Initial distribution:**  $\pi(dx)$ , a Borel probability measure on  $\mathfrak{S}_1$ .  $\pi(E)$  is “the probability that the state  $x$  at time 1 satisfies  $x \in E$ .”

The **Markov chain** associated with this data is the Markov process  $\mathbf{X} := \{X_n\}_{n \geq 1}$  such that  $X_n \in \mathfrak{S}_n$  for all  $n$ , and so that for all Borel  $E_i \subset \mathfrak{S}_i$ ,

$$\mathbb{P}(X_1 \in E_1) = \pi(E_1), \quad \mathbb{P}(X_{n+1} \in E_{n+1} | X_n = x_n) = \pi_{n,n+1}(x_n, E_{n+1}).$$

$\mathbf{X}$  is uniquely defined, with joint distribution

$$\begin{aligned} \mathbb{P}(X_1 \in E_1, \dots, X_n \in E_n) := \\ \int_{E_{n-1}} \int_{E_{n-2}} \dots \int_{E_1} \pi_{n-1,n}(x_{n-1}, E_n) \pi(dx_1) \pi_{1,2}(x_1, dx_2) \dots \pi_{n-2,n-1}(x_{n-2}, dx_{n-1}). \end{aligned} \tag{1.1.1}$$

$\mathbf{X}$  satisfies the following important **Markov property**:

$$\mathbb{P}(X_{n+1} \in E | X_n, X_{n-1}, \dots, X_1) = \mathbb{P}(X_{n+1} \in E | X_n) = \pi_{n,n+1}(X_n, E). \tag{1.1.2}$$

See, for instance, [17, Ch. 7].

In what follows  $\mathbb{P}, \mathbb{E}$  and  $\text{Var}$  denote the probability, expectation, and variance calculated using this joint distribution. In the special case when  $\pi$  is the point mass at  $x$ , we write  $\mathbb{P}_x, \mathbb{E}_x$  and  $\text{Var}_x$ .

If the state spaces and the transition probabilities do not depend on  $n$ , i.e.,  $\mathfrak{S}_n = \mathfrak{S}_1$  and  $\pi_{n,n+1}(x, dy) = \pi_{1,2}(x, dy)$  for all  $n$ , then we call  $X$  a **homogeneous** Markov chain. Otherwise,  $X$  is called an **inhomogeneous** Markov chain. In this work, we are mainly interested in the inhomogeneous case.

**Example 1.1 (Markov chain with finite state spaces).** *These are Markov chains  $X$  with state spaces  $\mathfrak{S}_n = \{1, \dots, d_n\}$ ,  $\mathcal{B}(\mathfrak{S}_n) = \{\text{subsets of } \mathfrak{S}_n\}$ .*

In this case the transition probabilities are completely characterized by the rectangular stochastic matrices with entries

$$\pi_{xy}^n := \pi_{n,n+1}(x, \{y\}) \quad (x = 1, \dots, d_n; y = 1, \dots, d_{n+1}),$$

and the initial distribution is completely characterized by the probability vector

$$\pi_x := \pi(\{x\}) \quad (x = 1, \dots, d_n).$$

The joint distribution of  $\{X_n\}$  is given by

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \pi_{x_1} \pi_{x_1 x_2}^1 \pi_{x_2 x_3}^2 \dots \pi_{x_{n-1} x_n}^{n-1},$$

and this leads to the following discrete version of (1.1.1):

$$\mathbb{P}(X_1 \in E_1, \dots, X_n \in E_n) = \sum_{x_{n-1} \in E_{n-1}} \sum_{x_{n-2} \in E_{n-2}} \dots \sum_{x_1 \in E_1} \pi_{x_1} \pi_{x_1 x_2}^1 \pi_{x_2 x_3}^2 \dots \pi_{x_{n-1} x_n}^{n-1}.$$

**Example 1.2 (Markov chains in random environment).**

Let  $X$  denote a homogeneous Markov chain with state space  $\mathfrak{S}$ , transition probability  $\pi(x, dy)$ , and initial distribution concentrated at a point  $x_1$ . It is possible to view  $X$  as a model for the motion of a particle on  $\mathfrak{S}$  as follows. At time 1, the particle is located at  $x_1$ , and a particle at position  $x$  will jump after one time step to a random location  $y$ , distributed like  $\pi(x, dy)$ :  $\mathbb{P}(y \in E) = \pi(x, E)$ . With this interpretation,

$$X_n = \text{the position of the particle at time } n.$$

The homogeneity of  $X$  is reflected in the fact that the law of motion which governs the jumps does not change in time.

Let us now refine the model by adding a dependence of the transition probabilities on an external parameter  $\omega$ , which we think of as “the environment.” For example,  $\omega$  can represent a external force field which affects the likelihood of various movements, and which can be modified by God or some other experimentalist. The transition probabilities become  $\pi(x, \omega, dy)$ .

Suppose the environment  $\omega$  changes in time according to some deterministic rule. This is modeled by a map  $T : \Omega \rightarrow \Omega$ , where  $\Omega$  is the collection of all possible states of the environment, and  $T$  is a deterministic law of motion which says that an environment at state  $\omega$  will evolve after one unit of time to the state  $T(\omega)$ . Iterating we see that if the initial state of the environment at time zero was  $\omega$ , then its state at time  $n$  will be  $\omega_n = T^{n-1}(\omega) = (T \circ \dots \circ T)(\omega)$ .

Returning to our particle, we see that if the initial condition of the environment at time one is  $\omega$ , then the transition probabilities at time  $n$  are

$$\pi_{n,n+1}^\omega(x, dy) = \pi(x, T^{n-1}(\omega), dy).$$

Thus each  $\omega \in \Omega$  gives rise to an inhomogeneous Markov chain  $X^\omega$ , which describes the Markovian dynamics of a particle, coupled to a changing environment, and corresponding to the initial condition that at time one, the particle is at position  $x_1$  and the environment is at state  $\omega$ .

If  $T(\omega) = \omega$ , the environment stays fixed, and the Markov chain is homogeneous, otherwise the Markov chain is inhomogeneous. We will return to Markov chains in random environment in chapter 8.

**Example 1.3 (Markov chains with finite memory).**

We can weaken the Markov property (1.1.2) by specifying that for some fixed  $k_0 \geq 1$ , for all  $E \in \mathcal{B}(\mathfrak{S}_{n+1})$ ,

$$\mathbb{P}(X_{n+1} \in E | X_n, \dots, X_1) = \begin{cases} \mathbb{P}(X_{n+1} \in E | X_n, \dots, X_{n-k_0+1}) & n > k_0; \\ \mathbb{P}(X_{n+1} \in E | X_n, \dots, X_1) & n \leq k_0. \end{cases}$$

Stochastic processes like that are called ‘‘Markov chains with finite memory’’ (of length  $k_0$ ). Markov chains with memory of length 1 are ordinary Markov chains. Markov chains with memory of length  $k_0 > 1$  can be recast as ordinary Markov chains by considering the stochastic process  $\tilde{X} = \{(X_n, \dots, X_{n+k_0-1})\}_{n \geq 1}$  with its natural state spaces, initial distribution, and transition kernels.

**Example 1.4 (A non-example).** *Every inhomogeneous Markov chain  $X$  can be presented as a homogeneous Markov chain  $Y$ , but this is not very useful.*

Let  $\mathfrak{S}_i$  denote the state spaces of  $X$ . These are complete separable metric spaces, and therefore they are Borel isomorphic to  $\mathbb{R}$ , or to  $\mathbb{Z}$ , or to a finite set (see e.g. [143], §3). So we can construct Borel bi-measurable injections  $\varphi_i : \mathfrak{S}_i \hookrightarrow \mathbb{R}$ . Let

$$Y_n = (\varphi_n(X_n), n).$$

We claim that  $Y = \{Y_n\}_{n \geq 1}$  is a homogeneous Markov chain. Let  $\delta_\xi$  denote the Dirac measure at  $\xi$ , defined by  $\delta_\xi(E) := 1$  when  $E \ni \xi$  and  $\delta_\xi(E) := 0$  otherwise. Let  $\mathfrak{S}_n, \pi_{n,n+1}$  and  $\pi$  denote the states spaces, transition probabilities, and initial distribution of  $X$ . Define a *homogeneous* Markov chain  $Z$  with

- state space  $\mathfrak{S} := \mathbb{R} \times \mathbb{N}$
- initial distribution  $\hat{\pi} := (\pi \circ \varphi_1^{-1}) \times \delta_1$ , a measure on  $\mathfrak{S}_1 \times \{1\}$
- transition probabilities

$$\hat{\pi}((x, n), A \times B) := \begin{cases} \pi_{n,n+1}(\varphi_n^{-1}(x), \varphi_{n+1}^{-1}(A)) \delta_{n+1}(B) & x \in \varphi_n(\mathfrak{S}_n) \\ \delta_0(A) \delta_1(B) & \text{otherwise.} \end{cases}$$

A direct calculation shows that the joint distribution  $Z$  is equal to the joint distribution of  $Y = \{(\varphi_n(X_n), n)\}_{n \geq 1}$ . So  $Y$  is a homogeneous Markov chain.

Such presentations will not be useful to us, because they destroy useful structures which are essential for our work on the local limit theorem. For example, they destroy the uniform ellipticity property in section 1.2 below.

### 1.1.2 Inhomogeneous Markov arrays

For technical reasons that we will explain later, it is useful to consider a generalization of a Markov chain, called a **Markov array**. To define a Markov array, we need the following data:

- **Row lengths:**  $k_N + 1$  where  $k_N \geq 1$  and  $(k_N)_{N \geq 1}$  is strictly increasing.
- **State spaces:**  $(\mathfrak{S}_n^{(N)}, \mathcal{B}(\mathfrak{S}_n^{(N)}))$ ,  $(1 \leq n \leq k_N + 1)$ , where  $\mathfrak{S}_n^{(N)}$  is a complete separable metric space with more than one point, and  $\mathcal{B}(\mathfrak{S}_n^{(N)})$  is its Borel  $\sigma$ -algebra.
- **Transition probabilities:**  $\{\pi_{n,n+1}^{(N)}(x, dy)\}_{x \in \mathfrak{S}_n^{(N)}}$  ( $1 \leq n \leq k_N$ ) where  $\pi_{n,n+1}^{(N)}(x, dy)$  are Borel probability measures on  $\mathfrak{S}_{n+1}^{(N)}$ , so that for every Borel  $E \subset \mathfrak{S}_{n+1}^{(N)}$ , the function  $x \mapsto \pi_{n,n+1}^{(N)}(x, E)$  is measurable, and for all  $x$ , and  $\pi_{n,n+1}^{(N)}(x, \cdot)$  is not carried by a single atom.
- **Initial distributions:** Borel probability measures  $\pi^{(N)}(dx)$  on  $\mathfrak{S}_1^{(N)}$ .

For each  $N \geq 1$ , this data determines a finite Markov chain of length  $k_N + 1$   $\mathbf{X}^{(N)} = (X_1^{(N)}, X_2^{(N)}, \dots, X_{k_N+1}^{(N)})$ , called the  **$N$ -th row** of the array. We will continue to denote the joint probability distribution, expectation, and variance of  $\mathbf{X}^{(N)}$  by  $\mathbb{P}$ ,  $\mathbb{E}$ , and  $\text{Var}$ . These objects depend on  $N$ , but the index  $N$  will always be obvious from the context, and can be suppressed. As always, in cases when we wish to condition on the initial state  $X_1^{(N)} = x$ , we will write  $\mathbb{P}_x$  and  $\mathbb{E}_x$ .

The rows  $\mathbf{X}^{(N)} = (X_1^{(N)}, X_2^{(N)}, \dots, X_{k_N+1}^{(N)})$  can be arranged in an array of random variables

$$\mathbf{X} = \begin{cases} X_1^{(1)}, \dots, X_{k_1+1}^{(1)} \\ X_1^{(2)}, \dots, X_{k_1+1}^{(2)}, \dots, X_{k_2+1}^{(2)} \\ X_1^{(3)}, \dots, X_{k_1+1}^{(3)}, \dots, X_{k_2+1}^{(3)}, \dots, X_{k_3+1}^{(3)} \\ \dots \end{cases}$$

Each horizontal row  $\mathbf{X}^{(N)} = (X_1^{(N)}, X_2^{(N)}, \dots, X_{k_N+1}^{(N)})$  comes equipped with a joint distribution, which depends on  $N$ . But *no joint distribution on elements of different rows is specified*.

#### Example 1.5 (Markov chains as Markov arrays).

Every Markov chain  $\{X_n\}$  gives rise to a Markov array with row lengths  $k_N = N + 1$  and rows  $\mathbf{X}^{(N)} = (X_1, \dots, X_{N+1})$ . In this case  $\mathfrak{S}_n^{(N)} = \mathfrak{S}_n$ ,  $\pi_{n,n+1}^{(N)} = \pi_{n,n+1}$ , and  $\pi^{(N)} = \pi$ .

Conversely, any Markov array so that  $\mathfrak{S}_n^{(N)} = \mathfrak{S}_n$ ,  $\pi_{n,n+1}^{(N)} = \pi_{n,n+1}$ , and  $\pi^{(N)} = \pi$  determines a Markov chain with state spaces  $\mathfrak{S}_n$ , transition probabilities  $\pi_{n,n+1}^{(N)} = \pi_{n,n+1}$ , and initial distributions  $\pi^{(N)} = \pi$ .

**Example 1.6 (Change of measure).** Suppose  $\{X_n\}_{n \geq 1}$  is a Markov chain with data  $\mathfrak{S}_n, \pi_{n,n+1}, \pi$ , and let  $\varphi_n^{(N)}(x, y)$  be a family of positive measurable functions on  $\mathfrak{S}_n \times \mathfrak{S}_{n+1}$ . Define new transition probabilities by

$$\pi_{n,n+1}^{(N)}(x, dy) := \frac{\varphi_{n,n+1}^{(N)}(x, y)}{\int \varphi_{n,n+1}^{(N)}(x, y) \pi_{n,n+1}(x, dy)} \pi_{n,n+1}(x, dy).$$

Then the data  $k_N = N + 1$ ,  $\mathfrak{S}_n^{(N)} := \mathfrak{S}_n$ ,  $\pi^{(N)} := \pi$  and  $\pi_{n,n+1}^{(N)}$  determines a Markov array called the **change of measure** of  $\{X_n\}$  with **weights**  $\varphi_n^{(N)}$ .

Why study Markov arrays? There are several reasons, and the one most relevant to this work is the following: The theory of large deviations for Markov *chains*, relies on a change of measure which results in Markov *arrays*. Thus, readers who are only interested in local limit theorems for Markov chains in the **local regime**  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}} \rightarrow z$ , may ignore the theory of arrays and limit their attention to Markov chains. But those who are also interested in the **large deviations regime**, where  $|\frac{z_N - \mathbb{E}(S_N)}{\text{Var}(S_N)}|$  is of order 1, will need the theory for Markov arrays.

### 1.1.3 Additive functionals

An **additive functional** of a Markov chain is a sequence  $f = \{f_n\}_{n \geq 1}$  of measurable functions  $f_n : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \rightarrow \mathbb{R}$ , where  $\mathfrak{S}_n$  are the states spaces of the Markov chain. The pair  $\mathbf{X} = \{X_n\}, f = \{f_n\}$  determines a stochastic process

$$S_N = f_1(X_1, X_2) + f_2(X_2, X_3) + \cdots + f_N(X_N, X_{N+1}) \quad (N \geq 1).$$

We will often abuse terminology and call  $(\mathbf{X}, f)$  and  $\{S_N\}_{N \geq 1}$  “additive functionals.”

An **additive functional** of a Markov array  $\mathbf{X}$  with row lengths  $k_N + 1$  and state spaces  $\mathfrak{S}_n^{(N)}$  is an array of measurable functions  $f_n^{(N)} : \mathfrak{S}_n^{(N)} \times \mathfrak{S}_{n+1}^{(N)} \rightarrow \mathbb{R}$  with row lengths  $k_N$ :

$$f = \begin{cases} f_1^{(1)}, \dots, f_{k_1}^{(1)} \\ f_1^{(2)}, \dots, f_{k_1}^{(2)}, \dots, f_{k_2}^{(2)} \\ f_1^{(3)}, \dots, f_{k_1}^{(3)}, \dots, f_{k_2}^{(3)}, \dots, f_{k_3}^{(3)} \\ \dots \dots \dots \dots \dots \dots \end{cases}$$

Again, this determines a sequence of random variables

$$S_N = f_1^{(N)}(X_1^{(N)}, X_2^{(N)}) + f_2^{(N)}(X_2^{(N)}, X_3^{(N)}) + \cdots + f_{k_N}^{(N)}(X_{k_N}^{(N)}, X_{k_N+1}^{(N)}) \quad (N \geq 1),$$

which we also refer to as “additive functional.” But be careful! *This is not a stochastic process*, because no joint distribution of  $S_1, S_2, \dots$  is specified.

Suppose  $f, g$  are two additive functionals on  $\mathbf{X}$ . If  $\mathbf{X}$  is a Markov chain,

$$f + g := \{f_n + g_n\}, \quad cf := \{cf_n\}, \quad |f| := \sup_n \left( \sup_{x,y} |f_n(x,y)| \right)$$

and  $\text{ess sup}_n |f| := \sup_n (\text{ess sup} |f_n(X_n, X_{n+1})|)$ .

Similarly, if  $\mathbf{X}$  is a Markov array with row lengths  $k_N + 1$ , then

$$f + g := \{f_n^{(N)} + g_n^{(N)}\}, \quad cf := \{cf_n^{(N)}\}, \quad |f| := \sup_N \sup_{1 \leq n \leq k_N} \left( \sup_{x,y} |f_n^{(N)}(x,y)| \right),$$

and

$$\text{ess sup}_N |f| := \sup_N \sup_{1 \leq n \leq k_N} \left( \text{ess sup} |f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})| \right).$$

The notation  $|f| \leq K$  a.s. will mean that  $\text{ess sup } |f| \leq K$  (“a.s.” stands for “almost surely”). An additive functional is called **uniformly bounded** if there is a constant  $K$  such that  $|f| \leq K$ , and **uniformly bounded a.s.** if  $\exists K$  such that  $|f| \leq K$  a.s.

## 1.2 Uniform ellipticity

### 1.2.1 The definition

A Markov chain  $X$  with state spaces  $\mathfrak{S}_n$  and transition probabilities  $\pi_{n,n+1}(x, dy)$  is called **uniformly elliptic**, if there exists a Borel probability measure  $\mu_n$  on  $\mathfrak{S}_n$ , Borel measurable functions  $p_n : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \rightarrow [0, \infty)$ , and a constant  $0 < \varepsilon_0 < 1$  called the **ellipticity constant** such that for all  $n \geq 1$ ,

- (a)  $\pi_{n,n+1}(x, dy) = p_n(x, y) \mu_{n+1}(dy)$ ;
- (b)  $0 \leq p_n \leq 1/\varepsilon_0$ ;
- (c)  $\int_{\mathfrak{S}_{n+1}} p_n(x, y) p_{n+1}(y, z) \mu_{n+1}(dy) > \varepsilon_0$ .

We will see in Proposition 1.12 below that one can always assume without loss of generality that  $\mu_n$  are the measures  $\mu_n(E) = \mathbb{P}(X_n \in E)$ .

The integral in (c) is the two-step transition probability  $\mathbb{P}(X_{n+2} = z | X_n = x)$ , and we will sometime call (c) a **two-step ellipticity condition**. For more general  $\gamma$ -step ellipticity conditions, see §1.3.3.

**Example 1.7 (Doeblin chains)** Suppose  $X$  has finite state spaces  $\mathfrak{S}_n$  s.t  $|\mathfrak{S}_n| \leq M < \infty$  for all  $n$ , and  $\pi_{xy}^n := \pi_{n,n+1}(x, \{y\})$  satisfy

- (1)  $\exists \varepsilon'_0 > 0$  s.t. for all  $n \geq 1$  and  $(x, y) \in \mathfrak{S}_n \times \mathfrak{S}_{n+1}$ , either  $\pi_{xy}^n = 0$  or  $\pi_{xy}^n > \varepsilon'_0$ ;
- (2) for all  $n$ , for all  $(x, z) \in \mathfrak{S}_n \times \mathfrak{S}_{n+2}$ , there exists  $y \in \mathfrak{S}_{n+1}$  such that  $\pi_{xy}^n \pi_{yz}^{n+1} > 0$ .

Doeblin chains are uniformly elliptic: Take  $\mu_n$  to be the uniform measure on  $\mathfrak{S}_n$  and  $p_n(x, y) := \pi_{xy}^n / |\mathfrak{S}_{n+1}|$ . Then (a) is clear, (b) holds with any  $\varepsilon_0 < 1/M$ , and (c) holds with  $\varepsilon_0 := (\varepsilon'_0/M)^2$ . Doeblin chains are named after W. Doeblin, who studied homogeneous countable Markov chains satisfying similar conditions.

Here is the formulation of the uniform ellipticity conditions for Markov arrays. A Markov array  $X$  with state spaces  $\mathfrak{S}_n^{(N)}$ , transition probabilities  $\pi_{n,n+1}^{(N)}(x, dy)$ , and row lengths  $k_N + 1$  is called **uniformly elliptic**, if there exist Borel probability measures  $\mu_n^{(N)}$  on  $\mathfrak{S}_n^{(N)}$ , Borel measurable functions  $p_n^{(N)} : \mathfrak{S}_n^{(N)} \times \mathfrak{S}_{n+1}^{(N)} \rightarrow [0, \infty)$ , and a constant  $0 < \varepsilon_0 < 1$  as follows: For all  $N \geq 1$  and  $1 \leq n \leq k_N$ ,

- (a)  $\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y) \mu_{n+1}^{(N)}(dy)$ ;
- (b)  $0 \leq p_n^{(N)} \leq 1/\varepsilon_0$ ;
- (c)  $\int_{\mathfrak{S}_{n+1}^{(N)}} p_n^{(N)}(x, y) p_{n+1}^{(N)}(y, z) \mu_{n+1}^{(N)}(dy) > \varepsilon_0$ .

**Example 1.8** Suppose  $X$  is a uniformly elliptic Markov chain and suppose  $Y$  is a Markov array obtained from  $X$  by the change of measure construction described in Example 1.6. If the weights  $\varphi_n^{(N)}(x, y)$  are uniformly bounded away from zero and infinity, then  $Y$  is uniformly elliptic.

### 1.2.2 Contraction estimates and exponential mixing

Suppose  $\mathfrak{X}, \mathfrak{Y}$  are complete and separable metric spaces. A **transition kernel** from  $\mathfrak{X}$  to  $\mathfrak{Y}$  is a family  $\{\pi(x, dy)\}_{x \in \mathfrak{X}}$  of Borel probability measures on  $\mathfrak{Y}$  so that  $x \mapsto \pi(x, E)$  is measurable for all  $E \subset \mathfrak{Y}$  Borel. A transition kernel  $\{\pi(x, dy)\}_{x \in \mathfrak{X}}$  determines two **Markov operators**, one acting on measures and the other acting on functions. The action on measures takes a probability measure  $\mu$  on  $\mathfrak{X}$  and maps it to a probability measure on  $\mathfrak{Y}$  via

$$\pi(\mu)(E) := \int_{\mathfrak{X}} \pi(x, E) \mu(dx).$$

The action on functions takes a bounded Borel function  $u : \mathfrak{Y} \rightarrow \mathbb{R}$  and maps it to a bounded Borel function on  $\mathfrak{X}$  via

$$\pi(u)(x) = \int_{\mathfrak{Y}} u(y) \pi(x, dy).$$

The two operators are dual:  $\int u(y) \pi(\mu)(dy) = \int \pi(u)(x) \mu(dx)$ .

These operators are contractions in the following sense. Define the **oscillation** of a function  $u : \mathfrak{Y} \rightarrow \mathbb{R}$  to be

$$\text{Osc}(u) := \sup_{y_1, y_2 \in \mathfrak{Y}} |u(y_1) - u(y_2)|.$$

The **contraction coefficient** of  $\{\pi(x, dy)\}_{x \in \mathfrak{X}}$  is

$$\delta(\pi) := \sup\{|\pi(x_1, E) - \pi(x_2, E)| : x_1, x_2 \in \mathfrak{X}, E \in \mathcal{B}(\mathfrak{Y})\}.$$

The **total variation distance** between two probability measures  $\mu_1, \mu_2$  on  $\mathfrak{X}$  is

$$\begin{aligned} \|\mu_1 - \mu_2\|_{\text{Var}} &:= \sup\{|\mu_1(A) - \mu_2(A)| : A \subset \mathfrak{X} \text{ is measurable}\} \\ &\equiv \frac{1}{2} \sup \left\{ \int w(x) (\mu_1 - \mu_2)(dx) \mid w : \mathfrak{X} \rightarrow [-1, 1] \text{ is measurable} \right\}. \end{aligned}$$

Caution!  $\|\mu_1 - \mu_2\|_{\text{Var}}$  is actually *one half* of the total variation of  $\mu_1 - \mu_2$ , because it is equal to  $(\mu_1 - \mu_2)^+(\mathfrak{X})$  and to  $(\mu_1 - \mu_2)^-(\mathfrak{X})$ , but not to

$$|\mu|(\mathfrak{X}) = (\mu_1 - \mu_2)^+(\mathfrak{X}) + (\mu_1 - \mu_2)^-(\mathfrak{X}).$$

**Lemma 1.9 ([139])** Suppose  $\mathfrak{X}, \mathfrak{Y}$  are complete and separable metric spaces, and  $\{\pi(x, dy)\}_{x \in \mathfrak{X}}$  is a transition kernel from  $\mathfrak{X}$  to  $\mathfrak{Y}$ . Then:

- (a)  $0 \leq \delta(\pi) \leq 1$ .
- (b)  $\delta(\pi) = \sup\{\text{Osc}[\pi(u)] \mid u : \mathfrak{Y} \rightarrow \mathbb{R} \text{ measurable, and } \text{Osc}(u) \leq 1\}$ .
- (c) If  $\mathfrak{Z}$  is a complete separable metric space,  $\pi_1$  is a transition kernel from  $\mathfrak{X}$  to  $\mathfrak{Y}$ , and  $\pi_2$  is a transition kernel from  $\mathfrak{Y}$  to  $\mathfrak{Z}$ , then  $\delta(\pi_1 \circ \pi_2) \leq \delta(\pi_1) \delta(\pi_2)$ .
- (d)  $\text{Osc}[\pi(u)] \leq \delta(\pi) \text{Osc}(u)$  for every  $u : \mathfrak{Y} \rightarrow \mathbb{R}$  bounded and measurable.
- (e)  $\|\pi(\mu_1) - \pi(\mu_2)\|_{\text{Var}} \leq \delta(\pi) \|\mu_1 - \mu_2\|_{\text{Var}}$  for all Borel probability measures  $\mu_1, \mu_2$  on  $\mathfrak{X}$ .
- (f) Suppose  $\lambda$  is a probability measure on  $\mathfrak{X} \times \mathfrak{Y}$  with marginals  $\mu_{\mathfrak{X}}, \mu_{\mathfrak{Y}}$ , and transition kernel  $\{\pi(x, dy)\}$ , i.e.  $\lambda(E \times \mathfrak{Y}) = \mu_{\mathfrak{X}}(E)$ ,  $\lambda(\mathfrak{X} \times E) = \mu_{\mathfrak{Y}}(E)$ , and

$$\lambda(dx, dy) = \int_{\mathfrak{X}} \pi(x, dy) \mu_{\mathfrak{X}}(dx).$$

Let  $f \in L^2(\mu_{\mathfrak{X}}), g \in L^2(\mu_{\mathfrak{Y}})$  be two elements with zero integral. Then

$$\left| \int_{\mathfrak{X} \times \mathfrak{Y}} f(x)g(y)\lambda(dx,dy) \right| \leq \sqrt{\delta(\pi)} \|f\|_{L^2(\mu_{\mathfrak{X}})} \|g\|_{L^2(\mu_{\mathfrak{Y}})}.$$

*Proof.* (a) is trivial.

The inequality  $\leq$  in (b) is because for every  $E \in \mathcal{B}(\mathfrak{Y})$ ,  $u := 1_E$  satisfies  $\text{Osc}(u) \leq 1$ . To see  $\geq$ , fix some  $u : \mathfrak{Y} \rightarrow \mathbb{R}$  measurable such that  $\text{Osc}(u) \leq 1$ . Suppose first that  $u$  is a simple function (a measurable function with finitely many values), then we can write  $u = c + \sum_{i=1}^m \alpha_i 1_{A_i}$  where  $c \in \mathbb{R}$ ,  $|\alpha_i| \leq \frac{1}{2} \text{Osc}(u)$ , and  $A_i$  measurable and pairwise disjoint. For every pair of points  $x_1, x_2 \in \mathfrak{X}$ ,

$$\begin{aligned} |\pi(u)(x_1) - \pi(u)(x_2)| &= \left| \sum_{i=1}^m \alpha_i [\pi(x_1, A_i) - \pi(x_2, A_i)] \right| \\ &\leq \left| \sum_{\pi(x_1, A_i) > \pi(x_2, A_i)} \alpha_i [\pi(x_1, A_i) - \pi(x_2, A_i)] \right| + \left| \sum_{\pi(x_1, A_i) < \pi(x_2, A_i)} \alpha_i [\pi(x_1, A_i) - \pi(x_2, A_i)] \right| \\ &\leq \frac{1}{2} \text{Osc}(u) \delta(\pi) + \frac{1}{2} \text{Osc}(u) \delta(\pi) = \delta(\pi) \text{Osc}(u) = \delta(\pi). \end{aligned}$$

So  $\text{Osc}[\pi(u)] \leq \delta(\pi)$  for all simple functions  $u$  with  $\text{Osc}(u) \leq 1$ . A standard approximation argument now shows that  $\text{Osc}[\pi(u)] \leq \delta(\pi)$  for all measurable  $u$  s.t.  $\text{Osc}(u) \leq 1$ . This proves (b). Part (c) and (d) immediately follow.

To see (e), we restrict to the non-trivial case  $\mu_1 \neq \mu_2$ . Let  $\mu := \mu_1 - \mu_2$ , and decompose  $\mu = \mu^+ - \mu^-$  where  $\mu^\pm$  are singular positive measures (this is the Jordan decomposition). Since  $\mu(\mathfrak{X}) = 0$ ,  $\mu^+, \mu^-$  has equal total mass, and

$$\mu^\pm(\mathfrak{X}) = \frac{1}{2}(\mu^+(\mathfrak{X}) + \mu^-(\mathfrak{X})) = \frac{1}{2}|\mu|(\mathfrak{X}) \equiv \|\mu_1 - \mu_2\|_{\text{Var}}.$$

Let

$$\hat{\mu}_1 := \mu^+ / \|\mu_1 - \mu_2\|_{\text{Var}}, \quad \hat{\mu}_2 := \mu^- / \|\mu_1 - \mu_2\|_{\text{Var}}, \quad \hat{\mu} := \hat{\mu}_1 - \hat{\mu}_2 = \frac{\mu_1 - \mu_2}{\|\mu_1 - \mu_2\|_{\text{Var}}}.$$

Note that  $\hat{\mu}_1$  and  $\hat{\mu}_2$  are probability measures.

For every non-constant measurable function  $w : \mathfrak{Y} \rightarrow [-1, 1]$ ,

$$\begin{aligned} \frac{\frac{1}{2} \int_{\mathfrak{Y}} w(y) \pi(\mu)(dy)}{\|\mu_1 - \mu_2\|_{\text{Var}}} &= \frac{1}{2} \int_{\mathfrak{Y}} w(y_1) \pi(\hat{\mu}_1)(dy_1) - \int_{\mathfrak{Y}} w(y_2) \pi(\hat{\mu}_2)(dy_2) \\ &= \frac{1}{2} \int_{\mathfrak{X}} \pi(w)(x_1) \hat{\mu}_1(dx_1) - \int_{\mathfrak{X}} \pi(w)(x_2) \hat{\mu}_2(dx_2) \\ &= \frac{1}{2} \int_{\mathfrak{X}} \int_{\mathfrak{X}} [\pi(w)(x_1) - \pi(w)(x_2)] \hat{\mu}_1(dx_1) \hat{\mu}_2(dx_2), \text{ because } \hat{\mu}_i(\mathfrak{X}) = 1, \\ &\leq \frac{1}{2} \delta(\pi) \text{Osc}(w) \leq \delta(\pi), \text{ by (b) and because } \text{Osc}(w) \leq 2\|w\|_\infty \leq 2. \end{aligned}$$

Passing to the supremum over all  $w(y)$  gives part (e).

Part (f) is the content of Lemma 4.1 in [139, Lemma 4.1], and we reproduce the proof given there. Consider the  $\sigma$ -algebra  $\mathcal{G} := \{\mathfrak{X} \times E : E \subset \mathfrak{Y} \text{ is measurable}\}$ , which represents the information on the  $\mathfrak{Y}$ -coordinate of  $(x, y) \in \mathfrak{X} \times \mathfrak{Y}$ .

Let  $\tilde{\pi}_y$  be a measurable family of conditional probabilities given  $\mathcal{G}$ , i.e.  $\tilde{\pi}_y$  is a probability measure on  $\mathfrak{X} \times \{y\}$ ,  $y \mapsto \int f d\tilde{\pi}_y$  is Borel for every Borel function  $f : \mathfrak{X} \times \mathfrak{Y} \rightarrow [0, 1]$ ,  $\lambda = \int_{\mathfrak{X} \times \mathfrak{Y}} \tilde{\pi}_y d\lambda$ , and for every  $\lambda$ -absolutely integrable  $f(x, y)$ ,

$$\mathbb{E}_\lambda(f(x, y) | \mathcal{G})(y) = \int_{\mathfrak{X}} f d\tilde{\pi}_y \text{ } \lambda\text{-a.e.}$$

We may identify  $\tilde{\pi}_y$  with a probability measure  $\hat{\pi}(y, dx)$  on  $\mathfrak{X}$  defined by

$$\hat{\pi}(y, E) = \tilde{\pi}_y(E \times \{y\}) \quad (E \subset \mathfrak{X} \text{ Borel}).$$

It is useful to think of  $\hat{\pi}(y, dx)$  as the transition kernel “which goes the opposite way” to  $\pi(x, dy)$ . Indeed, if  $\pi(x, dy)$  is the transition probability of a Markov chain  $\{X_n\}$  from  $n$  to  $n+1$ , and  $\lambda$  is the joint distribution of  $(X_n, X_{n+1})$ , then  $\hat{\pi}(y, dx)$  is the transition probability from  $n+1$  to  $n$ , i.e.  $\hat{\pi}(y, E) = \mathbb{P}(X_n \in E | X_{n+1} = y)$ .

The operators  $\pi : L^2(\mu_{\mathfrak{Y}}) \rightarrow L^2(\mu_{\mathfrak{X}})$  and  $\hat{\pi} : L^2(\mu_{\mathfrak{X}}) \rightarrow L^2(\mu_{\mathfrak{Y}})$  are dual to one another, because  $\int_{\mathfrak{X}} f(x) \pi(g)(x) d\mu_{\mathfrak{X}}(x)$  and  $\int_{\mathfrak{Y}} \hat{\pi}(f)(y) g(y) d\mu_{\mathfrak{Y}}(y)$  are both equal to  $\int f(x) g(y) \lambda(dx, dy)$ .

**CLAIM:**  $Q := \pi \circ \hat{\pi} : L^2(\mu_{\mathfrak{X}}) \rightarrow L^2(\mu_{\mathfrak{X}})$  is self-adjoint,  $Q$  preserves the linear subspace  $L_0^2(\mu_{\mathfrak{X}}) := \{f \in L^2(\mu_{\mathfrak{X}}) : \int f d\mu_{\mathfrak{X}} = 0\}$ , and the spectral radius of  $Q : L_0^2 \rightarrow L_0^2$  is at most  $\delta(Q)$ .

*Proof of the claim:*  $Q$  is self adjoint, because  $Q^* = (\pi \hat{\pi})^* = \hat{\pi}^* \pi^* = \pi \hat{\pi}$ .

It is useful to notice that  $Q$  is given by  $(Qf)(x) = \int_{\mathfrak{X}} f(x') Q(x, dx')$  where  $Q(x, E)$  is the probability measure on  $\mathfrak{X}$  given by  $Q(x, E) = \int \hat{\pi}(y, E) \pi(x, dy)$ .  $Q(x, dx')$  is a transition probability from  $\mathfrak{X}$  to  $\mathfrak{X}$ . Notice that  $Q(\mu_{\mathfrak{X}}) = \mu_{\mathfrak{X}}$ :

$$\begin{aligned} (Q\mu_{\mathfrak{X}})(E) &= \int_{\mathfrak{X}} Q(x, E) \mu_{\mathfrak{X}}(dx) = \int_{\mathfrak{X}} \int_{\mathfrak{Y}} \mu_{\mathfrak{X}}(dx) \pi(x, dy) \tilde{\pi}_y(E \times \{y\}) \\ &= \int_{\mathfrak{X} \times \mathfrak{Y}} \tilde{\pi}_y(E \times \{y\}) \lambda(dx, dy) = \int_{\mathfrak{X} \times \mathfrak{Y}} \tilde{\pi}_y(E \times \mathfrak{Y}) d\lambda = \lambda(E \times \mathfrak{Y}) = \mu_{\mathfrak{X}}(E). \end{aligned}$$

Thus, for all  $f \in L^2(\mu_{\mathfrak{X}})$ ,  $\int Q f d\mu_{\mathfrak{X}} = \int f d(Q\mu_{\mathfrak{X}}) = \int f d\mu_{\mathfrak{X}}$ . It follows that  $Q : L^2(\mu_{\mathfrak{X}}) \rightarrow L^2(\mu_{\mathfrak{X}})$  preserves the linear space  $L_0^2$ .

For every  $\varphi \in L_0^2 \cap L^\infty$ ,  $\|\varphi\|_\infty \leq \text{Osc}(\varphi)$ . Since  $Q$  preserves  $L_0^2 \cap L^\infty$ , for every  $f$  in this space, we have by parts (c) and (d) that

$$\|Q^n f\|_2 \leq \|Q^n f\|_\infty \leq \text{Osc}(Q^n f) \leq \delta(Q)^n \text{Osc}(f). \quad (1.2.1)$$

This implies that the spectral radius of  $Q : L_0^2 \rightarrow L_0^2$  is less than or equal to  $\delta(Q)$ . Otherwise there is an  $L_0^2$ -function, part of whose spectral decomposition corresponds to the part of the spectrum outside  $\{\lambda \in \mathbb{R} : |\lambda| \leq \delta(Q) + \varepsilon\}$  (self-adjoint operators have real spectrum). Any sufficiently close  $L_0^2 \cap L^\infty$ -function would have components with similar properties; but the existence of such components is inconsistent with (1.2.1). The proof of the claim is complete.

We are ready for the proof of (f). Since  $Q : L_0^2 \rightarrow L_0^2$  is a self-adjoint operator on a Hilbert space with spectral radius  $\leq \delta(Q)$ ,  $\langle Q(f), f \rangle_{L_0^2} \leq \delta(Q) \|f\|_{L_0^2}^2$  for all  $f \in L_0^2(\mu_{\mathfrak{X}})$ . It follows that

$$\|\hat{\pi}(f)\|_{L_0^2(\mu_{\mathfrak{Y}})}^2 = \langle \hat{\pi}(f), \hat{\pi}(f) \rangle_{L_0^2(\mu_{\mathfrak{Y}})} = \langle Q(f), f \rangle_{L_0^2(\mu_{\mathfrak{X}})} \leq \delta(Q) \|f\|_{L_0^2(\mu_{\mathfrak{X}})}^2.$$

So every  $f \in L_0^2(\mu_{\mathfrak{X}})$ ,  $g \in L_0^2(\mu_{\mathfrak{Y}})$

$$\begin{aligned} \left| \int_{\mathfrak{X} \times \mathfrak{Y}} f(x)g(y)\lambda(dx,dy) \right| &= \left| \int_{\mathfrak{Y}} \mu_{\mathfrak{Y}}(dy) \int_{\mathfrak{X}} \widehat{\pi}(y,dx)f(x)g(y) \right| = \langle \widehat{\pi}(f), g \rangle_{L^2(\mu_{\mathfrak{Y}})} \\ &\leq \|\widehat{\pi}(f)\|_2 \|g\|_2 \leq \sqrt{\delta(Q)} \|f\|_2 \|g\|_2, \text{ as required.} \end{aligned} \quad \square$$

We now return to the setup of Markov arrays  $\mathsf{X} = \{X_n^{(N)} : 1 \leq n \leq k_N + 1, N \geq 1\}$  and consider the following *two-step transition probabilities*

$$\pi_{n,n+2}^{(N)}(x, E) := \int \pi_{n+1,n+2}^{(N)}(y, E) \pi_{n,n+1}^{(N)}(x, dy)$$

defined for  $1 \leq n < N < \infty$ ,  $x \in \mathfrak{S}_n^{(N)}$ , and  $E \in \mathcal{B}(\mathfrak{S}_{n+2}^{(N)})$ . The uniform ellipticity condition gives the following uniform bound for  $\delta(\pi_{n,n+2}^{(N)})$ :

**Lemma 1.10** *Let  $\mathsf{X}$  be a uniformly elliptic Markov array with ellipticity coefficient  $\varepsilon_0$ . Then  $\sup_N \sup_{1 \leq n \leq k_N} \delta(\pi_{n,n+2}^{(N)}) \leq 1 - \varepsilon_0$ . Similarly for Markov chains.*

*Proof.* We fix  $N$  and drop the superscripts  $^{(N)}$ .

Uniform ellipticity implies that  $\pi_{n,n+2}(x, E) \ll \mu_{n+2}$  and that the Radon-Nikodym density is bounded from below by  $\varepsilon_0$ . This allows us to write

$$\pi_{n,n+2}(x, dy) = \varepsilon_0 \mu_{n+2}(dy) + (1 - \varepsilon_0) \widehat{\pi}_{n,n+2}(x, dy). \quad (1.2.2)$$

Note that the first term does not depend on  $x$ .

Let  $u : \mathfrak{S}_{n+2} \rightarrow \mathbb{R}$  be a measurable function with  $\text{Osc}(u) \leq 1$ , then we can write  $u(\cdot) = c + w(\cdot)$  where  $c$  is a constant and  $\|w\|_{\infty} \leq \frac{1}{2}$ . A direct calculation shows that

$$\begin{aligned} &\left| \int_{\mathfrak{S}_n} u(z) \pi_{n,n+2}(x_1, dz) - \int_{\mathfrak{S}_n} u(z) \pi_{n,n+2}(x_2, dz) \right| \\ &= \left| \int_{\mathfrak{S}_n} w(z) \pi_{n,n+2}(x_1, dz) - \int_{\mathfrak{S}_n} w(z) \pi_{n,n+2}(x_2, dz) \right| \\ &= (1 - \varepsilon_0) \left| \int_{\mathfrak{S}_n} w(z) \widehat{\pi}_{n,n+2}(x_1, dz) - \int_{\mathfrak{S}_n} w(z) \widehat{\pi}_{n,n+2}(x_2, dz) \right| \\ &\leq (1 - \varepsilon_0) \|w\|_{\infty} [\pi_{n,n+2}(x_1, \mathfrak{S}_{n+2}) + \pi_{n,n+2}(x_2, \mathfrak{S}_{n+2})] \leq 1 - \varepsilon_0, \end{aligned}$$

where the last inequality holds since  $\|w\|_{\infty} \leq \frac{1}{2}$ .  $\square$

**Proposition 1.11** *If  $\mathsf{X}$  is uniformly elliptic, then there exist  $\theta \in (0, 1)$  and  $C_{\text{mix}} > 0$ , which only depend on the ellipticity constant  $\varepsilon_0$  as follows. Suppose  $h_n^{(N)}(x, y)$  are measurable functions on  $\mathfrak{S}_n^{(N)} \times \mathfrak{S}_{n+1}^{(N)}$ , and let  $h_n^{(N)} := h_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})$ , then*

(1) *If  $h_n^{(N)}$  is bounded and  $\mathbb{E}(h_n^{(N)}) = 0$ , then for all  $1 \leq m < n \leq k_N$*

$$\|\mathbb{E}(h_n^{(N)} | X_m^{(N)})\|_{\infty} \leq C_{\text{mix}} \theta^{n-m} \|h_n^{(N)}\|_{\infty}. \quad (1.2.3)$$

(2) *If  $\text{Var}(h_n^{(N)}), \text{Var}(h_m^{(N)}) < \infty$  and  $\mathbb{E}(h_n^{(N)}), \mathbb{E}(h_m^{(N)}) = 0$ , then for all  $1 \leq m < n \leq k_N$*

$$\|\mathbb{E}(h_n^{(N)} | X_m^{(N)})\|_2 \leq C_{\text{mix}} \theta^{n-m} \|h_n^{(N)}\|_2. \quad (1.2.4)$$

$$|\mathbb{E}(h_m^{(N)} h_n^{(N)})| \leq C_{\text{mix}} \theta^{n-m} \|h_m^{(N)}\|_2 \|h_n^{(N)}\|_2. \quad (1.2.5)$$

The analogous statements hold for Markov chains.

*Proof.* We fix  $N$  and let  $\pi_{n,n+1} := \pi_{n,n+1}^{(N)}, X_n = X_n^{(N)}, h_n := h_n^{(N)}$ . Define for  $k \leq n$

$$w_{n,k}(X_k) := \mathbb{E}(h_n|X_k),$$

then  $w_{n,n}(X_n) := \mathbb{E}(h_n|X_n) = \int h_n(X_n, y) \pi_{n,n+1}(X_n, dy) = \pi_{n,n+1}[h_n(X_n, \cdot)]$ . By the Markov property,  $w_{n,n}(X_n) = \mathbb{E}(h_n|X_n, X_{n-1}, \dots, X_1)$ , and this allows us to write  $\pi_{n-1,n}(w_{n,n})(X_{n-1}) \equiv \mathbb{E}(w_{n,n}(X_n)|X_{n-1}) = \mathbb{E}(\mathbb{E}(h_n|X_n, \dots, X_1)|X_{n-1}) = \mathbb{E}(h_n|X_{n-1})$ . So  $\pi_{n-1,n}(w_{n,n})(X_{n-1}) = w_{n,n-1}(X_{n-1})$ .

Applying the Markov operator  $\pi_{n-2,n-1}$  on both sides gives in a similar way  $(\pi_{n-2,n-1} \circ \pi_{n-1,n})(w_{n,n})(X_{n-2}) = w_{n,n-2}(X_{n-2})$ .

Continuing in this way we arrive eventually to the identity

$$w_{n,m}(X_m) := \mathbb{E}(h_n|X_m) = (\pi_{m,m+1} \circ \dots \circ \pi_{n-1,n})(w_{n,n})(X_m).$$

By the previous lemmas  $\text{Osc}[w_{n,m}] \leq (1 - \varepsilon_0)^{\lfloor \frac{n-m}{2} \rfloor} \text{Osc}[w_{n,n}]$ .

Notice that for every bounded measurable function  $v$ ,  $\|v\|_\infty \leq |\mathbb{E}(v)| + \text{Osc}(v)$ . Since by assumption  $\mathbb{E}(w_{n,m}(X_m)) = \mathbb{E}(h_n) = 0$ ,

$$\|w_{n,m}(X_m)\|_\infty \leq (1 - \varepsilon_0)^{\lfloor \frac{n-m}{2} \rfloor} \text{Osc}[w_{n,n}].$$

$\text{Osc}[w_{n,n}] \leq 2\|w_{n,n}\|_\infty \leq 2\|h_n\|_\infty$ , and part 1 follows.

Part 2 is proved in a similar way, using Lemma 1.9(f).  $\square$

### 1.2.3 Hitting probabilities and bridge probabilities

Throughout this section, let  $\mathbf{X}$  be an inhomogeneous Markov array with row lengths  $k_N$ , and data  $\mathfrak{S}_n^{(N)}, \pi_{n,n+1}^{(N)}, \pi^{(N)}$ . Suppose  $\mathbf{X}$  is uniformly elliptic:

$$\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y) \mu_{n+1}(dy)$$

where  $0 \leq p_n^{(N)} \leq 1/\varepsilon_0$  and  $\int_{\mathfrak{S}_{n+1}} p_n^{(N)}(x, y) p_{n+1}^{(N)}(y, z) \mu_{n+1}(dy) > \varepsilon_0$ .

The following proposition estimates  $\mathbb{P}(X_n^{(N)} \in E)$  in terms of  $\mu_n^{(N)}$ :

**Proposition 1.12** *Under the above assumptions, for every  $3 \leq n \leq k_N + 1 < \infty$  and every Borel set  $E \subset \mathfrak{S}_N^{(N)}$ ,  $\varepsilon_0 \leq \frac{\mathbb{P}(X_n^{(N)} \in E)}{\mu_n^{(N)}(E)} \leq \varepsilon_0^{-1}$ . Similarly for Markov chains.*

*Proof.* We fix a row  $N$ , and drop the superscripts  $^{(N)}$ . Define a probability measure on  $\mathfrak{S}_n$  by  $P_n(E) = \mathbb{P}(X_n \in E)$ , then for every  $1 \leq n < k_N$ , for every bounded measurable  $\varphi : \mathfrak{S}_{n+2} \rightarrow \mathbb{R}$ ,

$$\begin{aligned} \int \varphi dP_{n+2} &= \mathbb{E}(\varphi(X_{n+2})) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{E}(\varphi(X_{n+2})|X_{n+1}, X_n) \middle| X_n\right)\right) \\ &= \mathbb{E}\left(\mathbb{E}(\mathbb{E}(\varphi(X_{n+2})|X_{n+1}) \middle| X_n)\right) \quad (\because \text{Markov property}) \\ &= \int \int \int \varphi(z) \pi_{n+1,n+2}(y, dz) \pi_{n,n+1}(x, dy) P_n(dx) \\ &= \int \int \int \varphi(z) p_{n+1}(y, z) p_n(x, y) \mu_{n+2}(dz) \mu_{n+1}(dy) P_n(dx) \end{aligned}$$

$$= \int \varphi(z) \left[ \int \left( \int p_{n+1}(y, z) p_n(x, y) \mu_{n+1}(dy) \right) P_n(dx) \right] \mu_{n+2}(dz)$$

The quantity in the square brackets is bounded below by  $\varepsilon_0$  and bounded above by  $\varepsilon_0^{-1}$ . So the measures  $P_{n+2}, \mu_{n+2}$  are equivalent, and  $\varepsilon_0 \leq \frac{dP_{n+2}}{d\mu_{n+2}} \leq \varepsilon_0^{-1}$ .  $\square$

Notice that in checking the uniform ellipticity condition, we are free to modify  $\mu_n^{(N)}$  by a density bounded away from zero and infinity. Thus, proposition 1.12 allows us to assume without loss of generality that  $\mu_n^{(N)}(E) = \mathbb{P}(X_n^{(N)} \in E)$  for  $3 \leq n \leq k_N$ .

The ellipticity property implies that for all  $x \in \mathfrak{S}_n^{(N)}, z \in \mathfrak{S}_{n+2}^{(N)}$ ,

$$Z_n^{(N)}(x, z) := \int_{\mathfrak{S}_{n+1}} p_n^{(N)}(x, y) p_{n+1}^{(N)}(y, z) \mu_{n+1}^{(N)}(dy) \neq 0.$$

This allows us to make the following definition: The **bridge distribution** of  $X_{n+1}^{(N)}$  given that  $X_n^{(N)} = x$  and  $X_{n+2}^{(N)} = z$  is the measure on  $\mathfrak{S}_{n+1}^{(N)}$  which assigns to a Borel set  $E \subset \mathfrak{S}_{n+1}^{(N)}$  the probability

$$\mathbb{P} \left( E \mid \begin{array}{l} X_n^{(N)} = x \\ X_{n+2}^{(N)} = z \end{array} \right) := \frac{1}{Z_n^{(N)}(x, z)} \int_E p_n^{(N)}(x, y) p_{n+1}^{(N)}(y, z) \mu_{n+1}^{(N)}(dy). \quad (1.2.6)$$

The definition makes sense because  $Z_n^{(N)}(x, z) \neq 0$ . The following lemma explains why the formula (1.2.6) is reasonable:

**Lemma 1.13** *Let  $\psi_E(x, z) :=$  right hand side of (1.2.6), then*

$$\psi_E(X_n^{(N)}, X_{n+2}^{(N)}) = \mathbb{P} \left( X_{n+1}^{(N)} \in E \mid X_n^{(N)}, X_{n+2}^{(N)} \right) \text{ } \mathbb{P}\text{-almost everywhere.}$$

We omit the proof, which is routine. The lemma does not “prove” (1.2.6): Conditional probabilities are only defined almost everywhere, and are by their very nature non-canonical. But (1.2.6) makes sense everywhere. It is a definition, not a theorem.

### 1.3 Structure constants

Throughout this section we assume that  $f$  is an additive functional on a uniformly elliptic Markov array  $X$  with row lengths  $k_N + 1$ , state spaces  $\mathfrak{S}_n^{(N)}$ , and transition probabilities as in the ellipticity condition:  $\pi_{n,n+1}^{(N)}(x, dy) = p_n(x, y) \mu_n(dy)$ , where  $\mu_n^{(N)}(E) = \mathbb{P}(X_n^{(N)} \in E)$ . See §1.2.3 why we may assume this on  $\mu_n^{(N)}$ .

#### 1.3.1 Hexagons, balance, and structure constants

A **Level  $N$  hexagon at position  $3 \leq n \leq k_N$**  is a configuration

$$P_n^{(N)} := \left( x_{n-2}; \frac{x_{n-1}}{y_{n-1}}; \frac{x_n}{y_n}; y_{n+1} \right)$$

where  $x_i, y_i \in \mathfrak{S}_i^{(N)}$ . A hexagon is called **admissible** if

$$\begin{aligned} p_{n-2}^{(N)}(x_{n-2}, x_{n-1}) p_{n-1}^{(N)}(x_{n-1}, x_n) p_n^{(N)}(x_n, y_{n+1}) &\neq 0 \\ p_{n-2}^{(N)}(x_{n-2}, y_{n-1}) p_{n-1}^{(N)}(y_{n-1}, y_n) p_n^{(N)}(y_n, y_{n+1}) &\neq 0 \end{aligned}$$

Admissible hexagons exist because of uniform ellipticity.

The space of level  $N$  admissible hexagons at position  $n$  will be denoted by  $\text{Hex}(N, n)$ .

One can put a natural probability measure on  $\text{Hex}(N, n)$  by taking  $\{Y_n^{(N)}\}$  to be an independent copy of  $\{X_n^{(N)}\}$ , and looking at the distribution of

$$\left( X_{n-2}^{(N)}, X_{n-1}^{(N)}; X_n^{(N)}; Y_{n+1}^{(N)} \right) \text{ conditioned on } \begin{cases} X_{n-2}^{(N)} = Y_{n-2}^{(N)} \\ X_{n+1}^{(N)} = Y_{n+1}^{(N)} \end{cases}$$

Writing the measure explicitly is possible, but cumbersome. It is better to think of it as the result of the following sampling procedure for  $\left( x_{n-2}; \frac{x_{n-1}}{y_{n-1}}; \frac{x_n}{y_n}; y_{n+1} \right)$ :

- $(x_{n-2}, x_{n-1})$  is sampled from the distribution of  $(X_{n-2}^{(N)}, X_{n-1}^{(N)})$ ;
- $(y_n, y_{n+1})$  is sampled from the distribution of  $(Y_n^{(N)}, Y_{n+1}^{(N)})$  (so it is independent of  $(x_n, x_{n+1})$ );
- $x_n$  and  $y_{n-1}$  are conditionally independent given the previous choices, and are sampled using the bridge distributions

$$\begin{aligned} \mathbb{P}(x_n \in E | x_{n-1}, y_{n+1}) &= \mathbb{P}\left(X_n^{(N)} \in E \middle| \begin{array}{l} X_{n-1}^{(N)} = x_{n-1} \\ X_{n+1}^{(N)} = y_{n+1} \end{array}\right) \\ \mathbb{P}(y_{n-1} \in E | x_{n-2}, y_n) &= \mathbb{P}\left(Y_{n-1}^{(N)} \in E \middle| \begin{array}{l} Y_{n-2}^{(N)} = x_{n-2} \\ Y_n^{(N)} = y_n \end{array}\right). \end{aligned}$$

We call the resulting measure the **hexagon measure** on  $\text{Hex}(N, n)$ .

The **balance** of a hexagon  $P_n^{(N)} := \left( x_{n-2}; \frac{x_{n-1}}{y_{n-1}}; \frac{x_n}{y_n}; y_{n+1} \right)$  is

$$\begin{aligned} \Gamma(P_n^{(N)}) &:= f_{n-2}^{(N)}(x_{n-2}, x_{n-1}) + f_{n-1}^{(N)}(x_{n-1}, x_n) + f_n^{(N)}(x_n, y_{n+1}) \\ &\quad - f_{n-2}^{(N)}(x_{n-2}, y_{n-1}) - f_{n-1}^{(N)}(y_{n-1}, y_n) - f_n^{(N)}(y_n, y_{n+1}). \end{aligned} \tag{1.3.1}$$

**Definition 1.14** The structure constants of  $\mathfrak{f} = \{f_n^{(N)}\}$  are

$$\begin{aligned} u_n^{(N)} &:= u_n^{(N)}(\mathfrak{f}) := \mathbb{E}((\Gamma(P_n^{(N)}))^2)^{1/2} \text{ (expectation on } \text{Hex}(N, n) \text{)} \\ d_n^{(N)}(\xi) &:= d_n^{(N)}(\xi, \mathfrak{f}) := \mathbb{E}(|e^{i\xi \Gamma(P_n^{(N)})} - 1|^2)^{1/2} \text{ (expectation on } \text{Hex}(N, n) \text{)} \\ U_N &:= U_N(\mathfrak{f}) := \sum_{n=3}^{k_N} (u_n^{(N)})^2, \quad D_N(\xi) := \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2. \end{aligned} \tag{1.3.2}$$

If  $\mathbf{X}$  is a Markov chain, we write  $u_n = u_n^{(N)}$ ,  $d_n(\xi) = d_n^{(N)}(\xi)$ .

The significance of the structure constants will become clear in later chapters. At this point we can only hint and say that the behavior of  $U_N$  determines if  $\text{Var}(S_N) \rightarrow \infty$ , and the behavior

of  $D_N(\xi)$  determines “how close”  $f$  is to an additive functional whose values all belong to the lattice  $(2\pi/\xi)\mathbb{Z}$ .

**Lemma 1.15** *Suppose  $f, g$  are two additive functionals on a uniformly elliptic Markov array  $X$ , then*

- (a)  $d_n^{(N)}(\xi + \eta, f)^2 \leq 8(d_n^{(N)}(\xi, f)^2 + d_n^{(N)}(\eta, f)^2)$ ;
- (b)  $d_n^{(N)}(\xi, f + g)^2 \leq 8(d_n^{(N)}(\xi, f)^2 + d_n^{(N)}(\xi, g)^2)$ ;
- (c)  $d_n^{(N)}(\xi, f) \leq |\xi| u_n^{(N)}(f)$ ;
- (d)  $u_n^{(N)}(f + g)^2 \leq 2[u_n^{(N)}(f)^2 + u_n^{(N)}(g)^2]$ .

*Proof.* For any  $z, w \in \mathbb{C}$  such that  $|z|, |w| \leq 2$ , we have <sup>1</sup>

$$|zw + z + w|^2 \leq 8(|z|^2 + |w|^2).$$

So if  $P$  is a level  $N$  hexagon  $P$  at position  $n$ , and  $\xi_P := \xi\Gamma(P)$ ,  $\eta_P := \eta\Gamma(P)$ , then

$$\begin{aligned} |e^{i(\xi_P + \eta_P)} - 1|^2 &= |(e^{i\xi_P} - 1)(e^{i\eta_P} - 1) + (e^{i\xi_P} - 1) + (e^{i\eta_P} - 1)|^2 \\ &\leq 8(|e^{i\xi_P} - 1|^2 + |e^{i\eta_P} - 1|^2). \end{aligned} \quad (1.3.3)$$

Part (a) follows by integrating over all  $P \in \text{Hex}(n, N)$ . Part (b) has a similar proof which we omit. Part (c) follows from the inequality  $|e^{i\theta} - 1|^2 = 4\sin^2 \frac{\theta}{2} \leq |\theta|^2$ . Part (d) follows from Minkowski’s inequality and  $|ab| \leq \frac{1}{2}(a^2 + b^2)$ .  $\square$

**Example 1.16 (Gradients)** *Suppose  $f_n(x, y) = a_{n+1}(y) - a_n(x) + c_n$  for all  $n$ , then the balance of each hexagon is zero and  $u_n, d_n(\xi)$  are all zero. For a converse statement, see §2.2.1.*

*Suppose  $f_n(x, y) = a_{n+1}(y) - a_n(x) + c_n \pmod{\frac{2\pi}{\xi}\mathbb{Z}}$  for all  $n$ . Then  $e^{i\xi\Gamma(P)} = 1$  for all hexagons  $P$ , and  $d_n(\xi)$  are all zero. For a converse statement, see §3.3.1.*

**Example 1.17 (Sums of independent random variables)** *Let  $S_N = X_1 + \dots + X_N$ , where  $X_i$  are independent real valued random variables with non-zero variance. Let us see what  $u_n$  and  $d_n(\xi)$  measure in this case.*

**Proposition 1.18**  $u_n^2 = 2(\text{Var}(X_{n-1}) + \text{Var}(X_n))$  and  $\sum_{n=3}^N u_n^2 \asymp \text{Var}(S_N)$  (i.e.  $\exists N_0$  such that the ratio of the two sides is uniformly bounded for  $N \geq N_0$ ).

*Proof.* Let  $\{Y_n\}$  be an independent copy of  $\{X_n\}$ , and let  $X_i^* := X_i - Y_i$  (the symmetrization of  $X_i$ ). A simple calculation shows that the balance of a position  $n$  hexagon is equal in distribution to  $X_{n-1}^* + X_n^*$ . Clearly  $\mathbb{E}[X_i^*] = 0$  and  $\mathbb{E}[(X_i^*)^2] = 2\text{Var}(X_i)$ . Consequently,

$$u_n^2(\xi) = \mathbb{E}[(X_{n-1}^*)^2 + (X_n^*)^2] = 2\text{Var}(X_{n-1}) + 2\text{Var}(X_n).$$

Summing over  $n$  we obtain  $\sum_{n=3}^N u_n^2 \asymp \text{Var}(S_N)$ .  $\square$

We remark that the proposition also holds for Markov arrays satisfying the *one-step ellipticity* condition (see §1.3.3).

Next we relate  $d_n^2(\xi)$  to the distance of  $X_i$  from a coset of  $\frac{2\pi}{\xi}\mathbb{Z}$ . The distance of a random variable  $X$  from a coset  $\frac{2\pi}{\xi}\mathbb{Z}$  is measured by the following quantity:

<sup>1</sup>  $(zw + z + w)^2 = z^2w^2 + z^2 + w^2 + 2(z^2w + zw^2 + zw)$ , and  $|z^2w^2| \leq 4|zw| \leq 2|z|^2 + 2|w|^2$ ,  $|z^2w| \leq 2|z|^2$ ,  $2|zw| \leq |z|^2 + |w|^2$ ,  $|zw^2| \leq 2|w|^2$ .

$$\mathfrak{D}(X, \xi) := \min_{\theta \in \mathbb{R}} \mathbb{E} \left[ \text{dist}^2 \left( X, \theta + \frac{2\pi}{\xi} \mathbb{Z} \right) \right]^{1/2}.$$

The minimum exists because the quantity we are minimizing is a periodic and continuous function of  $\theta$ .

**Proposition 1.19** *For every  $\xi \neq 0$   $d_n(\xi) = 0$  iff  $X_i \in \text{coset of } \frac{2\pi}{\xi} \mathbb{Z}$  a.s. ( $i = n-1, n$ ). In addition, there exists  $C(\xi) > 1$  such that if  $d_n(\xi) \neq 0$  then*

$$C(\xi)^{-1} \leq \frac{d_n^2(\xi)}{\mathfrak{D}(X_{n-1}, \xi)^2 + \mathfrak{D}(X_n, \xi)^2} \leq C(\xi).$$

*Proof.* Choose  $\theta_i \in [0, \frac{2\pi}{\xi}]$  s.t.  $\mathfrak{D}(X_i, \xi) = \mathbb{E}[\text{dist}^2(X_i, \theta_i + \frac{2\pi}{\xi} \mathbb{Z})]$ . There is no loss of generality in assuming that  $\theta_i = 0$ , because the structure constants of  $f_i(x) = x$  and  $g_i(x) = x - \theta_i$  are the same. Henceforth we assume that

$$\mathfrak{D}(X_i, \xi) = \mathbb{E}[\text{dist}^2(X_i, \frac{2\pi}{\xi} \mathbb{Z})]. \quad (1.3.4)$$

As in the proof of the previous proposition, the balance of a position  $n$  hexagon is equal in distribution to  $X_{n-1}^* + X_n^*$ , where  $X_i^* := X_i - Y_i$  and  $\{Y_i\}$  is an independent copy of  $\{X_i\}$ . So  $d_n^2(\xi) = \mathbb{E}(|e^{i(X_{n-1}^* + X_n^*)} - 1|^2)$ .

We need the following elementary facts:

$$|e^{i(x+y)} - 1|^2 = 4 \sin^2 \frac{x+y}{2} = 4(\sin \frac{x}{2} \cos \frac{y}{2} + \sin \frac{y}{2} \cos \frac{x}{2})^2 \quad (x, y \in \mathbb{R}) \quad (1.3.5)$$

$$\frac{4}{\pi^2} \text{dist}^2(t, \pi \mathbb{Z}) \leq \sin^2 t \leq \text{dist}^2(t, \pi \mathbb{Z}) \quad (t \in \mathbb{R}) \quad (1.3.6)$$

$$\mathbb{P}[X_i^* \in [0, \frac{\pi}{2\xi}] + \frac{2\pi}{\xi} \mathbb{Z}] \geq \frac{1}{4} \quad (i \geq 1) \quad (1.3.7)$$

(1.3.5) is trivial; (1.3.6) is because of the inequality  $2t/\pi \leq \sin t \leq t$  on  $[0, \frac{\pi}{2}]$ , which the reader may verify by drawing the graphs. To see (1.3.7) note that  $\mathbb{R} = ([0, \frac{\pi}{2\xi}] + \frac{\pi}{\xi} \mathbb{Z}) \uplus ([0, \frac{\pi}{2\xi}] + \frac{\pi}{2\xi} + \frac{\pi}{\xi} \mathbb{Z})$ , and therefore there exists  $k = 0, 1$  such that  $\mathbb{P}[X_i \in [0, \frac{\pi}{2\xi}] + \frac{k\pi}{2\xi} + \frac{\pi}{\xi} \mathbb{Z}] \geq \frac{1}{2}$ . Since  $Y_i$  is an independent copy of  $X_i$ ,  $\mathbb{P}[X_i, Y_i \in [0, \frac{\pi}{2\xi}] + \frac{k\pi}{2\xi} + \frac{\pi}{\xi} \mathbb{Z}] \geq \frac{1}{4}$ . This event is a subset of  $[X_i^* \in [0, \frac{\pi}{2\xi}] + \frac{2\pi}{\xi} \mathbb{Z}]$ .

Returning to the identity  $d_n^2(\xi) = \mathbb{E}(|e^{i(X_{n-1}^* + X_n^*)} - 1|^2)$ , we see that by (1.3.5)

$$\begin{aligned} d_n^2(\xi) &= \mathbb{E}(|e^{i\xi(X_{n-1}^* + X_n^*)} - 1|^2) \\ &= 4\mathbb{E} \left( \sin^2 \frac{\xi X_{n-1}^*}{2} \cos^2 \frac{\xi X_n^*}{2} + \sin^2 \frac{\xi X_n^*}{2} \cos^2 \frac{\xi X_{n-1}^*}{2} + \frac{1}{2} \sin(\xi X_{n-1}^*) \sin(\xi X_n^*) \right) \\ &= 4\mathbb{E} \left( \sin^2 \frac{\xi X_{n-1}^*}{2} \right) \mathbb{E} \left( \cos^2 \frac{\xi X_n^*}{2} \right) + 4\mathbb{E} \left( \sin^2 \frac{\xi X_n^*}{2} \right) \mathbb{E} \left( \cos^2 \frac{\xi X_{n-1}^*}{2} \right) \end{aligned} \quad (1.3.8)$$

where we used the symmetry of the distribution of  $X_i^*$  to see that  $\mathbb{E}[\sin(\xi X_i^*)] = 0$ . By (1.3.7),  $\mathbb{E} \left( \cos^2 \frac{\xi X_i^*}{2} \right) \geq \cos^2(\frac{\pi}{4}) \mathbb{P}[X_i^* \in [0, \frac{\pi}{2\xi}] + \frac{2\pi}{\xi} \mathbb{Z}] \geq \frac{1}{8}$ , and therefore there exists  $C_n \in [\frac{1}{8}, 4]$  such that

$$d_n^2(\xi) = C_n \left[ \mathbb{E} \left( \sin^2 \frac{\xi X_{n-1}^*}{2} \right) + \mathbb{E} \left( \sin^2 \frac{\xi X_n^*}{2} \right) \right]. \quad (1.3.9)$$

It remains to bound  $\mathbb{E} \left( \sin^2 \frac{\xi X_{n-1}^*}{2} \right)$  in terms of  $\mathfrak{D}(X_i, \xi)$ .

Recall that  $X_i^* = X_i - Y_i$  where  $Y_i$  is an independent copy of  $X_i$ , and use (1.3.5) and independence to find that

$$\begin{aligned} \mathbb{E} \left( \sin^2 \frac{\xi X_i^*}{2} \right) &= \mathbb{E} \left[ \left( \sin \frac{\xi X_i}{2} \cos \frac{\xi Y_i}{2} - \sin \frac{\xi Y_i}{2} \cos \frac{\xi X_i}{2} \right)^2 \right] \\ &= 2\mathbb{E}(\sin^2 \frac{\xi X_i}{2})\mathbb{E}(\cos^2 \frac{\xi X_i}{2}) - \frac{1}{2}\mathbb{E}(\sin(\xi X_i))^2 \leq 2\mathbb{E}(\sin^2 \frac{\xi X_i}{2}) \\ &\leq 2\mathbb{E}(\text{dist}^2(\frac{\xi X_i}{2}, \pi\mathbb{Z})) \equiv \frac{\xi^2}{2}\mathbb{E}(\text{dist}^2(X_i, \frac{2\pi}{\xi}\mathbb{Z})) = \frac{\xi^2}{2}\mathfrak{D}(X_i, \xi), \text{ by (1.3.4), (1.3.6).} \end{aligned}$$

Next by (1.3.6) and the definition of  $\mathfrak{D}(X_i, \xi)$ ,

$$\begin{aligned} \mathbb{E} \left( \sin^2 \frac{\xi X_i^*}{2} \right) &\geq \frac{4}{\pi^2}\mathbb{E} \left( \text{dist}^2(\frac{\xi X_i^*}{2}, \pi\mathbb{Z}) \right) = \frac{\xi^2}{\pi^2}\mathbb{E} \left( \text{dist}^2(X_i - Y_i, \frac{2\pi}{\xi}\mathbb{Z}) \right) \\ &= \frac{\xi^2}{\pi^2}\mathbb{E}_{Y_i} \left[ \mathbb{E}_{X_i} \left( \text{dist}^2(X_i, Y_i + \frac{2\pi}{\xi}\mathbb{Z}) \right) \right] \geq \frac{\xi^2}{\pi^2}\mathbb{E}_{Y_i} [\mathfrak{D}(X_i, \xi)] = \frac{\xi^2}{\pi^2}\mathfrak{D}(X_i, \xi). \end{aligned}$$

The proposition follows from (1.3.9).  $\square$

### 1.3.2 The ladder process

The material of this section is needed for the proofs of the gradient lemma and the reduction lemma in chapters 2 and 3, but will not be used elsewhere.

Suppose  $\mathsf{X} = \{X_i^{(N)}\}$  is a Markov array with row lengths  $k_N + 1$ , state spaces  $\mathfrak{S}_n^{(N)}$ , and transition probabilities  $\pi_{n,n+1}^{(N)}(x, dy)$ . Let  $\mu_n^{(N)}(E) := \mathbb{P}(X_n^{(N)} \in E)$ . Suppose  $\mathsf{X}$  is uniformly elliptic. In particular,

$$\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y)\mu_{n+1}^{(N)}(dy),$$

with  $p_n^{(N)}(x, y)$  as in the uniform ellipticity condition.

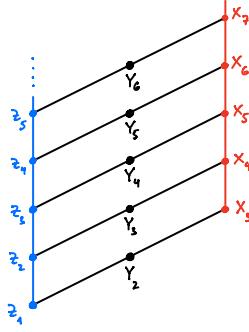
We would like to define a new Markov array  $\mathsf{L}$ , called the **Ladder process**, with the following structure (figure 1.3.2):

- (a) Each row has entries  $\underline{L}_n^{(N)} = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)})$  ( $3 \leq n \leq k_N + 1$ ),
- (b)  $\{Z_i^{(N)}\}$  is an independent copy of  $\mathsf{X}$ ,
- (c)  $Y_{n-1}^{(N)} \in \mathfrak{S}_{n-1}^{(N)}$  are independent given  $\{X_i^{(N)}\}, \{Z_i^{(N)}\}$ , and
- (d)  $\mathbb{P} \left( Y_{n-1}^{(N)} \in E \middle| \{X_i^{(N)}\} = \{x_i\}, \{Z_i^{(N)}\} = \{z_i\} \right) = \mathbb{P} \left( X_{n-1}^{(N)} \in E \middle| \begin{array}{l} X_{n-2}^{(N)} = z_{n-2} \\ X_n^{(N)} = x_n \end{array} \right)$ ,

see the discussion of bridge probabilities above.

Let  $\underline{L}_n^{(N)} = (z_{n-2}, y_{n-1}, x_n)$ . Define the probability measures

$$m_n^{(N)}(d\underline{L}_n^{(N)}) := \frac{p_{n-2}^{(N)}(z_{n-2}, y_{n-1})p_{n-1}^{(N)}(y_{n-1}, x_n)}{\int_{\mathfrak{S}_{n-1}^{(N)}} p_{n-2}^{(N)}(z_{n-2}, \eta)p_{n-1}^{(N)}(\eta, x_n)\mu_{n-1}^{(N)}(d\eta)} \mu_{n-2}^{(N)}(dz_{n-2})\mu_{n-1}^{(N)}(dy_{n-1})\mu_n^{(N)}(dx_n).$$



**Fig. 1.1** The ladder process.  $\{Z_i^{(N)}\}$ ,  $\{X_i^{(N)}\}$  are independent copies.  $Y_n^{(N)}$  are conditionally independent given  $\{X_i^{(N)}\}, \{Z_i^{(N)}\}$ .

**Lemma 1.20**  $\mathsf{L}$  exists, is Markov, and is uniformly elliptic with ellipticity constant  $\varepsilon_0^2$  (with respect to the background measure  $m_n$ ), where  $\varepsilon_0$  is the ellipticity constant of  $\mathsf{X}$ . For every  $N$ ,

- (1)  $\{X_n^{(N)}\}_{n=3}^{k_N+1}$ ,  $\{Z_n^{(N)}\}_{n=1}^{k_N-1}$  are independent, and distributed like the corresponding pieces of the  $N$ -th rows of  $\mathsf{X}$ .
- (2)  $Y_n^{(N)}$  are conditionally independent given  $\{X_i^{(N)}\}$ ,  $\{Z_i^{(N)}\}$ .
- (3)  $P_n^{(N)} := \left( Z_{n-2}^{(N)}, \frac{Z_{n-1}^{(N)}}{Y_{n-1}^{(N)}}, \frac{Y_n^{(N)}}{X_n^{(N)}}, X_{n+1}^{(N)} \right)$  is distributed like the level  $N$ , position  $n$ , random hexagon.

*Proof.* Let  $\mathbb{P} \left( dy_n \middle| \begin{array}{l} X_{n-1}^{(N)} = z_{n-1} \\ X_{n+1}^{(N)} = x_{n+1} \end{array} \right)$  denote the bridge measure on  $\mathfrak{S}_n^{(N)}$  with boundary conditions  $X_{n-1}^{(N)} = z_{n-1}$ ,  $X_{n+1}^{(N)} = x_{n+1}$ . Define the Markov array  $\mathsf{L}$  with

- Rows  $\underline{L}_n^{(N)} = (z_{n-2}, y_{n-1}, x_n)$  ( $3 \leq n \leq k_N + 1, N \geq 1$ )
- State spaces:  $\underline{\mathfrak{S}}_n^{(N)} := \mathfrak{S}_{n-2}^{(N)} \times \mathfrak{S}_{n-1}^{(N)} \times \mathfrak{S}_n^{(N)}$  ( $3 \leq n \leq k_N + 1$ ).
- Initial distribution:  $\pi^{(N)}(dz_1, dy_2, dx_3) = \int_{\mathfrak{S}_1^{(N)} \times \mathfrak{S}_3^{(N)}} \mu_1^{(N)}(dz) \mu_3^{(N)}(dx) \mathbb{P} \left( dy \middle| \begin{array}{l} X_1^{(N)} = z \\ X_3^{(N)} = x \end{array} \right)$
- Transition probabilities  $\pi_n^{(N)}((z_{n-2}, y_{n-1}, x_n), E_{n-1} \times E_n \times E_{n+1}) =$

$$= \int_{E_{n-1} \times E_n \times E_{n+1}} p_{n-2}^{(N)}(z_{n-2}, z_{n-1}) p_n^{(N)}(x_n, x_{n+1}) \mathbb{P} \left( dy_n \middle| \begin{array}{l} X_{n-1}^{(N)} = z_{n-1} \\ X_{n+1}^{(N)} = x_{n+1} \end{array} \right).$$

(We evolve  $z_{n-2} \rightarrow z_{n-1}$  and  $x_n \rightarrow x_{n+1}$  independently according to  $\pi_{n-2}^{(N)}(z_{n-2}, dz)$ ,  $\pi_n^{(N)}(x_n, dx)$ , and then sample  $y_n$  using the relevant bridge distribution.)

It is routine to check that  $\mathsf{L}$  has the structure described at the beginning of the section, and that it satisfies the properties listed in the lemma.

Here for example is the proof of uniform ellipticity. In what follows we fix  $N$ , suppose  $x_i, y_i, z_i \in \mathfrak{S}_i$ , and write  $p_n^{(N)} = p$  whenever the subscript is clear from the variables.

Then  $\pi_n^{(N)}(\underline{L}_n, d\underline{L}_{n+1}) = P(\underline{L}_n, \underline{L}_{n+1}) m_{n+1}(d\underline{L}_{n+1})$ , where

$$P(\underline{L}_n, \underline{L}_{n+1}) := p(z_{n-2}, z_{n-1})p(x_n, x_{n+1}).$$

If  $X$  has ellipticity constant  $\varepsilon_0$ , then  $P(\underline{L}_n, \underline{L}_{n+1}) \leq \varepsilon_0^{-2}$ , and

$$\begin{aligned} & \int P(\underline{L}_n, \underline{L}_{n+1})P(\underline{L}_{n+1}, \underline{L}_{n+2})m_{n+1}(d\underline{L}_{n+1}) \\ & \geq \iiint p(z_{n-2}, z_{n-1})p(x_n, x_{n+1})p(z_{n-1}, z_n)p(x_{n+1}, x_{n+2}) \times \\ & \quad \times \frac{p(z_{n-1}, y_n)p(y_n, x_{n+1})}{\int p(z_{n-1}, \eta)p(\eta, x_{n+1})\mu_n(d\eta)}\mu_{n-1}(dz_{n-1})\mu_n(dy_n)\mu_{n+1}(dx_{n+1}) \\ & = \iint p(z_{n-2}, z_{n-1})p(x_n, x_{n+1})p(z_{n-1}, z_n)p(x_{n+1}, x_{n+2})\mu_{n-1}(dz_{n-1})\mu_{n+1}(dx_{n+1}) \\ & = \int p(z_{n-2}, z_{n-1})p(z_{n-1}, z_n)\mu_{n-1}(dz_{n-1}) \int p(x_n, x_{n+1})p(x_{n+1}, x_{n+2})\mu_{n+1}(dx_{n+1}) \\ & \geq \varepsilon_0^{-2}. \end{aligned}$$

So the ladder process is uniformly elliptic with ellipticity constant  $\varepsilon_0^2$ .  $\square$

### 1.3.3 $\gamma$ -step ellipticity conditions

We mention a few possible variants of the uniform ellipticity condition discussed in this chapter. Suppose  $X$  is a Markov array with row lengths  $k_N + 1$  and transition probabilities taking the form  $\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y)\mu_{n+1}^{(N)}(dy)$ .

The **one-step ellipticity condition** is that for some  $\varepsilon_0 > 0$ , for all  $N \geq 1$ ,  $1 \leq n \leq k_N$ , and for every  $x \in \mathfrak{S}_n^{(N)}, y \in \mathfrak{S}_{n+1}^{(N+1)}$ ,

$$\varepsilon_0 < p_n^{(N)}(x, y) \leq \varepsilon_0^{-1}.$$

Notice that this implies that all transitions  $x \rightarrow y$  have positive probability.

The  **$\gamma$ -step ellipticity condition** ( $\gamma = 2, 3, \dots$ ) is that for some  $\varepsilon_0 > 0$ , for all  $N \geq 1, n \leq k_N$ ,

$$0 \leq p_n^{(N)} \leq 1/\varepsilon_0$$

and for all  $n \leq k_N - \gamma + 1$ , and every  $x \in \mathfrak{S}_n^{(N)}, z \in \mathfrak{S}_{n+\gamma}^{(N)}$ , the iterated integral

$$\int_{\mathfrak{S}_{n+1}^{(N)}} \cdots \int_{\mathfrak{S}_{n+\gamma}^{(N)}} p_n^{(N)}(x, y_1) \prod_{i=1}^{\gamma-2} p_{n+i}^{(N)}(y_i, y_{i+1}) p_{n+\gamma-1}^{(N)}(y_{\gamma-1}, z) \mu_{n+1}(dy_1) \cdots \mu_{n+\gamma}(dy_{\gamma-1})$$

is bigger than  $\varepsilon_0$  (with the convention that  $\prod_{i=1}^0 := 1$ ).

The ellipticity condition we use in this work corresponds to  $\gamma = 2$ . This is weaker than the one-step condition, but stronger than the  $\gamma$ -step condition for  $\gamma \geq 3$ .

The results of this work could in principle be reproduced assuming only a  $\gamma$ -step condition with  $\gamma \geq 2$ . To do this, one needs to replace the space of hexagons by the space of  $2(\gamma+1)$ -gons  $\left( x_{n-\gamma}; \begin{matrix} x_{n-\gamma+1} & \dots & x_n \\ y_{n-\gamma+1} & \dots & y_n \end{matrix}; y_{n+1} \right)$  with its associated structure constants, and its associated  $\gamma$ -ladder

process  $\underline{L}_n^{(N)} = (Z_{n-\gamma-1}^{(N)}, Y_{n-\gamma}^{(N)}, \dots, Y_{n-1}^{(N)}, X_n^{(N)})$ . Since no new ideas are needed, and since our notation is already heavy enough as it is, we will only treat the case  $\gamma = 2$  in this work.

### 1.3.4 Uniform ellipticity and strong mixing conditions

The contents of this section are not used elsewhere in this work.

Suppose  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space, and let  $\mathcal{A}, \mathcal{B}$  be two sub  $\sigma$ -algebras of  $\mathcal{F}$ . There are several standard measures for the dependence between  $\mathcal{A}$  and  $\mathcal{B}$ :

$$\begin{aligned}\alpha(\mathcal{A}, \mathcal{B}) &:= \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{A}, B \in \mathcal{B}\}; \\ \rho(\mathcal{A}, \mathcal{B}) &:= \sup \left\{ |\mathbb{E}(fg) - \mathbb{E}(f)\mathbb{E}(g)| : \begin{array}{l} f \in L^2(\mathcal{A}), g \in L^2(\mathcal{B}); \\ \|f - \mathbb{E}(f)\|_2 = 1, \|g - \mathbb{E}(g)\|_2 = 1 \end{array} \right\}; \\ \phi(\mathcal{A}, \mathcal{B}) &:= \sup \{ |\mathbb{P}(B|A) - \mathbb{P}(B)| : A \in \mathcal{A}, B \in \mathcal{B}, \mathbb{P}(A) \neq 0 \}; \\ \psi(\mathcal{A}, \mathcal{B}) &:= \sup \left\{ \left| \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} - 1 \right| : A \in \mathcal{A}, B \in \mathcal{B} \text{ with non-zero probabilities} \right\}.\end{aligned}$$

If one of these quantities vanishes then they all vanish, and this happens iff  $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$  for all  $A \in \mathcal{A}, B \in \mathcal{B}$ . In this case we say that  $\mathcal{A}, \mathcal{B}$  are **independent**. In the dependent case,  $\alpha, \rho, \phi, \psi$  can be used to bound the covariance between (certain)  $\mathcal{A}$ -measurable and  $\mathcal{B}$ -measurable random variables:

**Theorem 1.1.** Suppose  $X$  is  $\mathcal{A}$ -measurable,  $Y$  is  $\mathcal{B}$ -measurable, then

- (1)  $|\text{Cov}(X, Y)| \leq 8\alpha(\mathcal{A}, \mathcal{B})^{1-\frac{1}{p}-\frac{1}{q}} \|X\|_p \|Y\|_q$  whenever  $p \in (1, \infty]$ ,  $q \in (1, \infty]$ ,  
 $\frac{1}{p} + \frac{1}{q} < 1$ ,  $X \in L^p$ ,  $Y \in L^q$ .
- (2)  $|\text{Cov}(X, Y)| \leq \rho(\mathcal{A}, \mathcal{B}) \|X - \mathbb{E}Y\|_2 \|Y - \mathbb{E}Y\|_2$  whenever  $X, Y \in L^2$ .
- (3)  $|\text{Cov}(X, Y)| \leq 2\phi(\mathcal{A}, \mathcal{B}) \|X\|_1 \|Y\|_\infty$  whenever  $X \in L^1$ ,  $Y \in L^\infty$ .
- (4)  $|\text{Cov}(X, Y)| \leq \psi(\mathcal{A}, \mathcal{B}) \|X\|_1 \|Y\|_1$  whenever  $X \in L^1$ ,  $Y \in L^\infty$ .

For proof and references, see [16, vol 1, ch. 3].

**Definition 1.21** Let  $X := \{X_n\}_{n \geq 1}$  be a general stochastic process, not necessarily stationary or Markov. Let  $\mathcal{F}_1^n$  denote the  $\sigma$ -algebra generated by  $X_1, \dots, X_n$ , and let  $\mathcal{F}_m^\infty$  denote the  $\sigma$ -algebra generated by  $X_k$  for  $k \geq m$ .

- (1)  $X$  is called  **$\alpha$ -mixing**, if  $\alpha(n) := \sup_{k \geq 1} \alpha(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty) \xrightarrow{n \rightarrow \infty} 0$ .
- (2)  $X$  is called  **$\rho$ -mixing**, if  $\rho(n) := \sup_{k \geq 1} \rho(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty) \xrightarrow{n \rightarrow \infty} 0$ .
- (3)  $X$  is called  **$\phi$ -mixing**, if  $\phi(n) := \sup_{k \geq 1} \phi(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty) \xrightarrow{n \rightarrow \infty} 0$ .
- (4)  $X$  is called  **$\psi$ -mixing**, if  $\psi(n) := \sup_{k \geq 1} \psi(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty) \xrightarrow{n \rightarrow \infty} 0$ .

**Theorem 1.2.** If  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space, and  $\mathcal{A}, \mathcal{B}$  are sub- $\sigma$ -algebras of  $\mathcal{F}$ , then  $\alpha := \alpha(\mathcal{A}, \mathcal{B})$ ,  $\rho := \rho(\mathcal{A}, \mathcal{B})$ ,  $\phi := \phi(\mathcal{A}, \mathcal{B})$ ,  $\psi := \psi(\mathcal{A}, \mathcal{B})$  satisfy the inequalities

$$2\alpha \leq \phi \leq \frac{1}{2}\psi, \quad \rho \leq 2\sqrt{\phi}. \quad (1.3.10)$$

For the proof, see [16, vol 1, Prop. 3.11]). It follows that

$$\psi\text{-mixing} \Rightarrow \phi\text{-mixing} \Rightarrow \rho\text{-mixing} \Rightarrow \alpha\text{-mixing}.$$

These implications are strict, see [16, vol 1 §5.23].

Let us see what is the connection of  $\phi$ -mixing to uniform ellipticity. First we'll show that uniform ellipticity implies exponential  $\psi$ -mixing, and then we'll give a weak converse of this statement for finite state Markov chains.

**Proposition 1.22** *Let  $X$  be a uniformly elliptic Markov chain, then for every  $x \in \mathfrak{S}_1$ ,  $X$  conditioned on  $X_1 = x$  is  $\psi$ -mixing. Moreover,  $\alpha(n), \rho(n), \phi(n), \psi(n) \xrightarrow[n \rightarrow \infty]{} 0$  exponentially fast, uniformly in  $x$ .*

*Proof.* We will need the following fact:

**CLAIM.** *There exists a constant  $K$  which only depends on the ellipticity constant of  $X$  as follows. For every  $x \in \mathfrak{S}_1$ ,  $k \geq 2$ , and for every bounded measurable function  $h_k : \mathfrak{S}_k \rightarrow \mathbb{R}$ , we have the inequality  $\|\mathbb{E}_x(h_k(X_k)|X_{k-2})\|_\infty \leq K\mathbb{E}_x(|h_k(X_k)|)$ .*

*Proof of the claim.* By the uniform ellipticity of  $X$ , the transition kernels of  $X$  can be put in the form  $\pi_{n,n+1}(x, dy) = p_n(x, y)\mu_{n+1}(dy)$ , where  $0 \leq p_n \leq \varepsilon_0^{-1}$  and  $\int p_n(x, y)p_{n+1}(y, z)\mu_{n+1}(dy) > \varepsilon_0$ . In addition, Prop. 1.12 tells us that the Radon-Nikodym derivative of  $\mu_{n+1}$  with respect to the measure  $\mathbb{P}_x(X_{n+1} \in E)$  is almost everywhere in  $[\varepsilon_0, \varepsilon_0^{-1}]$ . It follows that for all  $\xi$ ,

$$\begin{aligned} |\mathbb{E}_x(h_{k+2}(X_{k+2})|X_k = \xi)| &\leq \iint p_k(\xi, y)p_{k+1}(y, z)|h_{k+2}(z)|\mu_{k+1}(dy)\mu_{k+2}(dz) \\ &\leq \varepsilon_0^{-2} \int |h_{k+2}(z)|\mu_{k+2}(dz) \leq \varepsilon_0^{-3}\mathbb{E}_x(|h_{k+2}(X_{k+2})|). \end{aligned}$$

We now prove the proposition. Fix  $x \in \mathfrak{S}_1$ , and let  $\psi_x$  denote the  $\psi$  measure of dependence for  $X$  conditioned on  $X_1 = x$ . Let  $\mathcal{F}_k$  denote the  $\sigma$ -algebra generated by  $X_k$ . Using the Markov property, it is not difficult to see that

$$\psi_x(n) = \sup_{k \geq 1} \psi_x(\mathcal{F}_k, \mathcal{F}_{k+n}),$$

see [16, vol 1, pp. 206–7].

Suppose now that  $n > 2$ , and fix some  $A \in \mathcal{F}_k, B \in \mathcal{F}_{k+n}$  with positive  $\mathbb{P}_x$ -measure. Let  $h_k := 1_A$  and  $h_{k+n} := 1_B - \mathbb{P}_x(B)$ . Then

$$\begin{aligned} |\mathbb{P}_x(A \cap B) - \mathbb{P}_x(A)\mathbb{P}_x(B)| &= |\mathbb{E}_x(h_k h_{k+n})| = |\mathbb{E}_x(\mathbb{E}_x(h_k h_{k+n}|\mathcal{F}_k))| \\ &= |\mathbb{E}_x(h_k \mathbb{E}_x(h_{k+n}|X_k))| \leq \mathbb{E}_x(|h_k|)\|\mathbb{E}_x(h_{k+n}|X_k)\|_\infty \\ &= \mathbb{P}_x(A)\|\mathbb{E}_x(\mathbb{E}_x(h_{k+n}|X_{k+n-2})|X_k)\|_\infty \\ &\leq \mathbb{P}_x(A) \cdot C_{mix} \theta^{n-2} \|\mathbb{E}_x(h_{k+n}|X_{k+n-2})\|_\infty, \text{ by uniform ellipticity and (1.2.3)} \\ &\leq \mathbb{P}_x(A) \cdot C_{mix} \theta^{n-2} \cdot K\mathbb{E}_x(|h_{k+n}|), \text{ by the claim} \\ &\leq 2K C_{mix} \theta^{n-2} \mathbb{P}_x(A)\mathbb{P}_x(B). \end{aligned}$$

Dividing by  $\mathbb{P}_x(A)\mathbb{P}_x(B)$  and passing to the supremum over  $A \in \mathcal{F}_k, B \in \mathcal{F}_{k+n}$ , gives  $\psi_x(n) \leq 2K C_{mix} \theta^{n-2}$ .

Recall from Proposition 1.11 that  $C_{mix}, \theta$  depend on the ellipticity constant of  $X$ , but not on  $x$ . So  $\psi_x(n) \rightarrow 0$  exponentially fast, uniformly in  $x$ . By (1.3.10),  $\alpha_x(n), \rho_x(n), \phi_x(n) \rightarrow 0$  exponentially fast, uniformly in  $n$ .  $\square$

**Proposition 1.23** *Let  $X$  be a Markov chain such that*

- (1)  $\exists \kappa > 0$  s.t.  $\mathbb{P}(X_n = x) > \kappa$  for every  $n \geq 1$ ,  $x \in \mathfrak{S}_n$  (in particular,  $|\mathfrak{S}_n| < 1/\kappa$ ).
- (2)  $\phi(n) \xrightarrow[n \rightarrow \infty]{} 0$ .

*Then  $X$  satisfies the  $\gamma$ -step ellipticity condition for all  $\gamma$  large enough.*

*Proof.* By (1), all state spaces are finite sets. Define a measure on  $\mathfrak{S}_n$  by  $\mu_n(E) = \mathbb{P}(X_n \in E)$ , and let  $p_n(x, y) := \frac{\mathbb{P}(X_{n+1} = y | X_n = x)}{\mathbb{P}(X_{n+1} = y)}$ . This is well-defined by (1), and:

- (a) By construction,  $\pi_{n,n+1}(x, dy) = p_n(x, y)\mu_{n+1}(dy)$ .
- (b) By (1),  $p_n(x, y) \leq 1/\mathbb{P}(X_{n+1} = y) \leq \kappa^{-1}$ .
- (c) By (2), for all  $\gamma$  large enough,  $\phi(\gamma) < \frac{1}{2}\kappa$ . For such  $\gamma$ ,

$$\begin{aligned} & \int_{\mathfrak{S}_{n+1}} \cdots \int_{\mathfrak{S}_{n+\gamma}} p_n(x, y_1) \prod_{i=1}^{\gamma-2} p_{n+i}(y_i, y_{i+1}) p_{n+\gamma-1}(y_{\gamma-1}, z) \mu_{n+1}(dy_1) \cdots \mu_{n+\gamma}(dy_{\gamma-1}) \\ &= \mathbb{P}(X_{n+\gamma} = z | X_n = x) \geq \mathbb{P}(X_{n+\gamma} = z) - \phi(\mathcal{F}_n, \mathcal{F}_{n+\gamma}) \geq \kappa - \phi(\gamma) > \frac{1}{2}\kappa. \end{aligned}$$

We obtain the  $\gamma$ -ellipticity condition with ellipticity constant  $\frac{1}{2}\kappa$ . □

## 1.4 Notes and references

For a comprehensive treatment of inhomogeneous Markov chains on general state spaces, see Doob's book [52]. The uniform ellipticity condition is one of a plethora of contraction conditions for Markov operators, which were developed over the years as sufficient conditions for results such as Propositions 1.11 and 1.12. We mention in particular the works of Markov [100], Doeblin [46, 47], Hajnal [70], Doob [52], and Dobrushin [45] (see also Seneta [137] and Sethuraman & Varadhan [139]).

The contraction coefficient mentioned in section 1.2.2 is also called an “ergodicity coefficient,” and it plays a major role in Dobrushin's proof of the CLT for inhomogeneous Markov chains [45]. Our treatment of contraction coefficients follows closely [139]. In particular, Lemma 1.9 and the proof of part (f) of that lemma is taken from there.

Proposition 1.12 is similar in spirit to Doeblin's estimates for the stationary probability vector of a Markov chain satisfying Doeblin's condition in terms of the stochastic matrix of the chain [46, 47].

For a discussion of the “change of measure” construction see chapter 6. The quantities  $\mathfrak{D}(X, \xi)$  were introduced by Mukhin for the purpose of studying local limit theorem for sums of independent random variables. See [108] and references therein.

For a comprehensive account of measures of dependence and mixing conditions, see [16].



## Chapter 2

# Variance growth, center-tightness, and the central limit theorem

In this chapter we analyze the variance of  $S_N = f_1(X_1, X_2) + \dots + f_N(X_N, X_{N+1})$  as  $N \rightarrow \infty$ , characterize the additive functionals for which  $\text{Var}(S_N) \not\rightarrow \infty$ , and prove Dobrushin's Theorem: If  $V_N \rightarrow \infty$  then the central limit theorem holds.

### 2.1 Main results

Let  $X$  be a Markov array with row lengths  $k_N + 1$ , let  $f$  an additive functional on  $X$ , and define  $S_N = \sum_{i=1}^{k_N} f_i^{(N)}(X_i^{(N)}, X_{i+1}^{(N)})$ .

**Definition 2.1**  $f$  is called **center-tight** if there are constants  $m_N$  s.t. for every  $\varepsilon > 0$ , there exists  $M$  s.t.  $\mathbb{P}[|S_N - m_N| > M] < \varepsilon$  for all  $N$ .

Center-tightness is an obstruction to the local limit theorem. We shall see below (Theorem 2.2) that  $f$  is center-tight iff  $\text{Var}(S_N) \not\rightarrow \infty$ . Obviously, in such a situation the right hand side in  $\mathbb{P}[S_N - z_N \in (a, b)] \stackrel{?}{\sim} \frac{e^{-z^2/2} |a-b|}{\sqrt{2\pi V_N}}$  can be made bigger than one by choosing  $|a-b|$  sufficiently big, and the asymptotic relation fails. One could hope for a different universal asymptotic behavior, but as the following class of examples shows, this is hopeless:

**Example 2.2 (Non-universality in the LLT for center-tight functionals):**

Let  $X = \{X_n\}_{n \geq 1}$  be a sequence of identically distributed independent random variables with uniform distribution on  $[0, 1]$ . Choose an *arbitrary* sequence of random variables  $\{Z_n\}_{n \geq 1}$  taking values in  $[0, 1]$ . By the isomorphism theorem for Lebesgue spaces, there are measurable functions  $g_n : [0, 1] \rightarrow [0, 1]$  such that

$$g_0 \equiv 0, \quad g_n(X_n) = Z_n \text{ in distribution.}$$

Let  $f = \{f_n\}_{n \geq 1}$  with  $f_n(X_n, X_{n+1}) := g_{n+1}(X_{n+1}) - g_n(X_n)$ . Then  $S_N = Z_{N+1}$  in distribution, whence  $\mathbb{P}(S_N \in (a, b)) = \mathbb{P}(Z_{N+1} \in (a, b))$  is completely arbitrary.

Every Markov array admits center-tight additive functionals. Here are three constructions which lead to such examples (in the uniformly bounded, uniformly elliptic case, *all* center-tight additive functional arise this way, see Theorem 2.2 below):

**Example 2.3 (Gradients):** Gradients on Markov chains are additive functionals of the form

$$f_n(x, y) := (\nabla a)_n(x, y) := a_{n+1}(y) - a_n(x).$$

where  $a_n : \mathfrak{S}_n \rightarrow \mathbb{R}$  is measurable, and  $\mathbf{a} = \{a_n\}$  is a.s. uniformly bounded.

Gradients on Markov arrays are defined similarly by the formula  $f_n^{(N)}(x, y) := a_{n+1}^{(N)}(y) - a_n^{(N)}(x)$ , where  $a_n^{(N)} : \mathfrak{S}_n^{(N)} \rightarrow \mathbb{R}$  is measurable, and  $\mathbf{a} = \{a_n^{(N)}\}$  is a.s. uniformly bounded. We write  $\mathbf{f} = \nabla \mathbf{a}$ , and say that  $\mathbf{f}$  is the **gradient** of  $\mathbf{a}$  and  $\mathbf{a}$  is the **potential** of  $\mathbf{f}$ .<sup>1</sup>

The gradient of an a.s. uniformly bounded potential is center-tight because if  $|\mathbf{a}| \leq K$ , then  $|S_N| = |a_{k_N+1}^{(N)}(X_{N+1}) - a_1^{(N)}(X_1)| \leq 2K$ .

**Example 2.4 (Summable variance):** We say that an additive functional  $\mathbf{f}$  on a Markov chain  $\mathbf{X}$  has **summable variance** if it is a.s. uniformly bounded, and

$$V_\infty := \sum_{n=1}^{\infty} \text{Var}[f_n(X_n, X_{n+1})] < \infty.$$

The definition of summable variance for additive functionals on arrays is similar, except that now  $V_\infty$  is defined by  $V_\infty := \sup_N \sum_{n=1}^{k_N} \text{Var}[f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})] < \infty$ .

If  $\mathbf{X}$  is uniformly elliptic and  $|\mathbf{f}| \leq K$  a.s., then summable variance implies center-tightness. This follows from Chebyshev's inequality and the following lemma:

**Lemma 2.5** Let  $\mathbf{f}$  be a uniformly bounded functional of the uniformly elliptic Markov array. Then  $V_N \leq \bar{V}_N \left(1 + \frac{2C_{\text{mix}}}{1-\theta}\right)$  where  $\bar{V}_N := \sum_{n=1}^{k_N} \text{Var}(f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}))$ , and  $C_{\text{mix}}$  and  $0 < \theta < 1$  are as in Prop. 1.11.

*Proof.* We give the proof for Markov chains (the proof for arrays is identical):

$$\begin{aligned} \text{Var}(S_N) &= \sum_{n=1}^N \text{Var}(f_n) + 2 \sum_{n=1}^{N-1} \sum_{m=n+1}^N \text{Cov}(f_n, f_m) \\ &\leq \bar{V}_N + 2C_{\text{mix}} \sum_{n=1}^{N-1} \sum_{m=n+1}^N \theta^{m-n} \sqrt{\text{Var}(f_n) \text{Var}(f_m)}, \text{ with } C_{\text{mix}}, \theta \text{ as in (1.2.5)} \\ &\leq \bar{V}_N + 2C_{\text{mix}} \sum_{j=1}^{N-1} \theta^j \sum_{n=1}^{N-j} \sqrt{\text{Var}(f_n) \text{Var}(f_{n+j})} < \bar{V}_N + \frac{2C_{\text{mix}} \bar{V}_N}{1-\theta} \end{aligned}$$

by the Cauchy-Schwarz inequality.  $\square$

**Example 2.6** Suppose  $\mathbf{X}$  is uniformly elliptic. Then every additive functional of the form  $\mathbf{f} = \mathbf{g} + \mathbf{h}$  where  $\mathbf{g}$  is a gradient and  $\mathbf{h}$  has summable variance is center-tight.

We will now state the main results of this chapter. We assume throughout that

- (E)  $\mathbf{X} = \{X_n^{(N)}\}$  is a **uniformly elliptic** inhomogeneous Markov array with row lengths  $k_N + 1$ , state spaces  $\mathfrak{S}_n^{(N)}$ , transition probabilities  $\pi_{n,n+1}^{(N)}$ , initial distributions  $\pi^{(N)}$ , and ellipticity constant  $\varepsilon_0$ .
- (B)  $\mathbf{f} = \{f_n^{(N)}\}$  is an **a.s. uniformly bounded** additive functional on  $\mathbf{X}$ , satisfying the bound  $|\mathbf{f}| \leq K$  almost surely.

<sup>1</sup> In the ergodic theoretic literature,  $\mathbf{f}$  is called a **coboundary** and  $\mathbf{a}$  is called a **transfer function**.

Let  $V_N := \text{Var}(S_N)$ , and  $U_N := \sum_{n=3}^{k_N} (u_n^{(N)})^2$  where  $u_n^{(N)}$  are as in (1.3.2).

**Theorem 2.1.** *There are constants  $C_1, C_2 > 0$  which only depend on  $\varepsilon_0, K$  s.t. for every uniformly elliptic array with ellipticity constant  $\varepsilon_0$  and every additive functional  $f$  on  $X$  s.t.  $|f| \leq K$  a.s.,*

$$C_1^{-1}U_N - C_2 \leq \text{Var}(S_N) \leq C_1U_N + C_2 \text{ for all } N.$$

**Corollary 2.7** *Suppose  $X$  is a Markov chain. Either  $\text{Var}(S_N) \rightarrow \infty$  or  $\text{Var}(S_N) = O(1)$ . Moreover,  $\text{Var}(S_N) \asymp \sum_{n=3}^N u_n^2$  where  $u_n$  are the structure constants from (1.3.2).*

(The corollary is clearly false for arrays.) Returning to arrays, we'll show:

**Theorem 2.2.**  *$\text{Var}(S_N)$  is bounded iff  $f$  is center-tight iff  $f = \nabla a + h$  where  $a$  is a uniformly bounded potential, and  $h$  has summable variance.*

**Corollary 2.8**  *$f$  is center-tight iff  $\sup_N U_N < \infty$ .*

Theorem 2.1 is a statement on the localization of cancellations. In general, if the variance of an additive functional of a stochastic process does not tend to infinity, then there must be some strong cancellations in  $S_N$ . A priori, these cancellations may involve summands located far apart from one another. Theorem 2.1 says that strong cancellations must already occur among three consecutive terms  $f_{n-2}^{(N)} + f_{n-1}^{(N)} + f_n^{(N)}$ : This is what  $U_N$  measures.

If  $f$  depends only on one variable  $f_n(x, y) = f_n(x)$ , and we have the one-step ellipticity condition  $p_N(x, y) \geq \varepsilon_0$  one can define the ladder process using quadrilaterals

$$Q_n^N = \begin{pmatrix} X_{n-1}^N & X_n^N & Y_{n+1}^N \\ & Y_n^N & \end{pmatrix}$$

instead of hexagons. As a result  $u_n$  is replaced by

$$(\bar{u}_n^{(N)})^2 \asymp \iint |f_n^{(N)}(y_1) - f_n^{(N)}(y_2)|^2 d\mu_n(y_1) d\mu_n(y_2) = 2\text{Var}(f_n). \quad (2.1.1)$$

Repeating the arguments from the proof of Theorem 2.1 we obtain that there are constants  $\widehat{C}_1, \widehat{C}_2$  such that

$$\widehat{C}_1^{-1} \sum_n \text{Var}(f_n(X_n)) - \widehat{C}_2 \leq V_N \leq \widehat{C}_1 \left( \sum_n \text{Var}(f_n(X_n)) \right) + \widehat{C}_2.$$

This estimate has been previously obtained in [45, 139] under weaker ellipticity assumptions. A similar estimate does *not* hold in case  $f_n^{(N)}$  depends on two variables. Indeed if  $f_n^{(N)}$  is a

gradient, then  $V_N$  is bounded while  $\sum_{n=1}^N \text{Var}(f_n(X_n, X_{n+1}))$  can be arbitrarily large.

We end the chapter with the reproduction of the proofs of the following two known well-known results.

**Theorem 2.3 (Dobrushin).** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$ . If  $\text{Var}(S_N) \rightarrow \infty$ , then for every interval,*

$$\mathbb{P} \left[ \frac{S_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}} \in (a, b) \right] \xrightarrow{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt.$$

The proof we give, which is due to Sethuraman & Varadhan, is based on McLeish's martingale central limit theorem. For the convenience of the reader we prove the martingale CLT in section 2.2.4.

The next result reduces in the case of identically distributed independent random variables to Khintchin-Kolmogorov's Two-Series Theorem. The result is stated for Markov chains, and not Markov arrays, because it relates to the properties of  $S_N$  as a stochastic process.

**Theorem 2.4.** *Let  $f = \{f_n\}$  be an a.e. uniformly bounded additive functional of a uniformly elliptic inhomogeneous Markov chain  $X = \{X_n\}$ . If  $\sum_{n=1}^{\infty} \text{Var}[f_n(X_n, X_{n+1})]$  is finite, then*

$$\sum_{n=1}^{\infty} [f_n(X_n, X_{n+1}) - \mathbb{E}(f_n(X_n, X_{n+1}))] \text{ converges almost surely.}$$

## 2.2 Proofs

### 2.2.1 The Gradient Lemma

**Lemma 2.9 (Gradient Lemma)** *Suppose  $f$  is an additive functional on a uniformly elliptic Markov array  $X$ , and assume  $|f| \leq K$  almost surely. Then we can write*

$$f = \tilde{f} + \nabla a + c,$$

where  $\tilde{f}, a, c$  are additive functionals on  $X$  with the following properties:

- (a)  $|a| \leq 2K$  and  $a_n^{(N)}(x)$  are measurable functions on  $\mathfrak{S}_n^{(N)}$ .
- (b)  $|c| \leq K$  and  $c_n^{(N)}$  are constant functions.
- (c)  $|\tilde{f}| \leq 6K$  and  $\tilde{f}_n^{(N)}(x, y)$  satisfy  $\|\tilde{f}_n^{(N)}\|_2 \leq u_n^{(N)}$  for all  $3 \leq n \leq k_N + 1$ .

If  $X$  is a Markov chain, we can choose  $f_n^{(N)} = f_n$ ,  $a_n^{(N)} = a_n$ ,  $c_n^{(N)} = c_n$ .

**Proof for Doeblin chains:** Before proving the lemma in full generality, we consider the important special case of Doeblin chains (Example 1.7), for which the proof is particularly simple.

Recall that a Doeblin chain is a Markov chain  $X$  with *finite* state spaces  $\mathfrak{S}_n$  of *uniformly bounded cardinality*, and whose associated transition matrices  $\pi_{xy}^n := \pi_{n,n+1}(x, \{y\})$  satisfy the following properties:

- (E1)  $\exists \epsilon'_0 > 0$  s.t. for all  $n \geq 1$  and  $(x, y) \in \mathfrak{S}_n \times \mathfrak{S}_{n+1}$ , either  $\pi_{xy}^n = 0$  or  $\pi_{xy}^n > \epsilon'_0$ ;
- (E2) for all  $n$ , for all  $(x, z) \in \mathfrak{S}_n \times \mathfrak{S}_{n+2}$ ,  $\exists y \in \mathfrak{S}_{n+1}$  such that  $\pi_{xy}^n \pi_{yz}^{n+1} > 0$ .

We saw in example 1.7 that  $X$  is uniformly elliptic.

We re-label the states in  $\mathfrak{S}_n$  so that  $\mathfrak{S}_n = \{1, \dots, d_n\}$  where  $d_n \leq d$ , and in such a way that  $\pi_{11}^n > 0$  for all  $n$ . Assumption (E2) guarantees that for every  $n \geq 3$  and every  $x \in \mathfrak{S}_n$  there exists a state  $\xi_{n-1}(x) \in \mathfrak{S}_{n-1}$  s.t.  $\pi_{1, \xi_{n-1}(x)}^{n-2} \pi_{\xi_{n-1}(x), x}^{n-1} > 0$ . Let

$$a_0 \equiv 0, \quad a_1 \equiv 0, \quad \text{and } a_n(x) := f_{n-2}(1, \xi_{n-1}(x)) + f_{n-1}(\xi_{n-1}(x), x) \text{ for } n \geq 3$$

$$c_0 := 0, \quad c_1 := 0, \quad \text{and } c_n := f_{n-2}(1, 1) \text{ for } n \geq 3$$

$$\tilde{f} := f - \nabla a - c.$$

We claim that  $\tilde{f}, a, c$  satisfy our requirements. To explain why and to motivate the construction, consider the special case  $u_n = 0$ . In this  $\|\tilde{f}\|_2 = 0$  and the lemma reduces to constructing functions  $b_n : \mathfrak{S}_n \rightarrow \mathbb{R}$  s.t.  $f = \nabla b + c$ . We first try to solve  $f = \nabla b$  with  $c = 0$ . Any solution must satisfy

$$f_n(x, y) = b_{n+1}(y) - b_n(x). \quad (2.2.1)$$

Necessarily,  $b_n(y) = b_2(x_2) + f_2(x_2, x_3) + \cdots + f_{n-2}(x_{n-2}, x_{n-1}) + f_{n-1}(x_{n-1}, y)$  for all paths  $(x_2, \dots, x_{n-1}, y)$  with positive probability. The path  $x_2 = \cdots = x_{n-2} = 1, x_{n-1} = \xi_{n-1}(y)$  suggests to define

$$b_2 \equiv 0, \quad b_n(y) := \sum_{k=2}^{n-3} f_k(1, 1) + f_{n-2}(1, \xi_{n-1}(y)) + f_{n-1}(\xi_{n-1}(y), y)$$

This works: for every  $n \geq 3$ , if  $\pi_{xy}^n > 0$  then

$$\begin{aligned} b_{n+1}(y) - b_n(x) &= [f_{n-2}(1, 1) + f_{n-1}(1, \xi_n(y)) + f_n(\xi_n(y), y) \\ &\quad - f_{n-2}(1, \xi_{n-1}(x)) - f_{n-1}(\xi_{n-1}(x), x) - f_n(x, y)] + f_n(x, y) \\ \therefore b_{n+1}(y) - b_n(x) &= \Gamma_n \left( 1 \quad \begin{matrix} 1 & \xi_n(y) \\ \xi_{n-1}(x) & x \end{matrix} \quad y \right) + f_n(x, y) \stackrel{!}{=} f_n(x, y). \end{aligned} \quad (2.2.2)$$

Here is the justification of  $\stackrel{!}{=}$ . In the setup we consider, the natural measure on the level  $n$  hexagons is atomic, and every admissible hexagon has positive mass. So  $u_n = 0$  implies that  $\Gamma_n(P) = 0$  for every admissible hexagon, and  $\stackrel{!}{=}$  follows.

We proved (2.2.1), but we are not yet done because  $b$  is not necessarily uniformly bounded. To fix this decompose  $b_n(y) = a_n(y) + \sum_{k=2}^{n-3} f_k(1, 1)$ . Then  $|a| \leq 2K$ , and a direct calculation shows that  $f_n(x, y) = a_{n+1}(y) - a_n(x) + f_{n-2}(1, 1)$ , whence  $f = \nabla a + c$  as we claimed.

This proves the lemma in case  $u_n = 0$ . The general case  $u_n \geq 0$  is done in exactly the same way, except that now the identity (2.2.2) gives for  $\tilde{f} := f - \nabla a - c$

$$\tilde{f}_n(x, y) = f_n(x, y) - (a_{n+1}(y) - a_n(x)) - c_n = -\Gamma_n \left( 1 \quad \begin{matrix} 1 & \xi_n(y) \\ \xi_{n-1}(x) & x \end{matrix} \quad y \right).$$

If  $|f| \leq K$ , then  $|\Gamma_n| \leq 6K$ , whence  $|\tilde{f}| \leq 6K$ . Next,

$$\|\tilde{f}_n\|_2^2 \leq \mathbb{E} \left[ \Gamma_n \left( 1 \quad \begin{matrix} 1 & \xi_n(X_{n+1}) \\ \xi_{n-1}(X_n) & X_n \end{matrix} \quad X_{n+1} \right)^2 \right].$$

In the scenario we consider the space of admissible hexagons has a finite number of elements, and each has probability uniformly bounded below. So there is a global constant  $C$  which only depends on  $\sup |\mathfrak{S}_n|$  and on  $\varepsilon'_0$  in (E2) such that

$$\mathbb{E} \left[ \Gamma_n \left( 1 \quad \begin{matrix} 1 & \xi_n(X_{n+1}) \\ \xi_{n-1}(X_n) & X_n \end{matrix} \quad X_{n+1} \right)^2 \right] \leq C \mathbb{E}[\Gamma(P)^2],$$

where the last expectation is over all position  $n$  hexagons. So  $\|\tilde{f}\|_2 \leq \sqrt{C} \cdot u_n^2$ .

(The gradient lemma says that we can choose  $a$  and  $c$  so that  $C = 1$ . The argument we gave does not quite give this, but the value of the constant is not important for the applications we have in mind.)

**The proof of the gradient lemma in the general case:** Recall the ladder process  $L = \{\underline{L}_n^{(N)}\}$ ,  $\underline{L}_n^{(N)} = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)})$  from §1.3.2. In what follows we omit the superscripts  $(N)$  on the right hand side of identities. Define

$$\begin{aligned} F_n^{(N)}(\underline{L}_n^{(N)}) &:= F_n(\underline{L}_n) = f_{n-2}(Z_{n-2}, Y_{n-1}) + f_{n-1}(Y_{n-1}, X_n) \\ F_n^{(N)}(\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)}) &:= F_n(\underline{L}_n, \underline{L}_{n+1}) = \Gamma \left( Z_{n-2} \frac{Z_{n-1}}{Y_{n-1}} \frac{Y_n}{X_n} X_{n+1} \right), \text{ see (1.3.1).} \end{aligned}$$

Then we have the following identity:

$$f_n^{(N)}(X_n, X_{n+1}) = F_{n+1}(\underline{L}_{n+1}) - F_n(\underline{L}_n) + f_{n-2}(Z_{n-2}, Z_{n-1}) - \Gamma_n(\underline{L}_n, \underline{L}_{n+1}). \quad (2.2.3)$$

Next define  $a_n^{(N)} : \mathfrak{S}_n^{(N)} \rightarrow \mathbb{R}$  and  $c_n^{(N)} \in \mathbb{R}$  by

$$a_n^{(N)}(\xi) := \mathbb{E} \left( \mathbb{E}(F_n(\underline{L}_n) | X_n = \xi) \right) \quad (3 \leq n \leq k_N) \quad (2.2.4)$$

$$c_n^{(N)} := \mathbb{E}[f_{n-2}(Z_{n-2}, Z_{n-1})]. \quad (2.2.5)$$

We will show that the lemma holds with  $a, c$  and  $\tilde{f} := f - \nabla a - c$ .

Since  $|f| \leq K$  by assumption, it is clear that  $|a| \leq 2K$  and  $|c| \leq K$ . It remains to bound  $\tilde{f}$  in  $L^\infty$  and  $L^2$ .

**CLAIM:** For every  $(\xi, \eta) \in \mathfrak{S}_n \times \mathfrak{S}_{n+1}$ ,

$$\begin{aligned} c_n^{(N)} &= \mathbb{E} \left[ \mathbb{E} \left( f_{n-2}(Z_{n-2}, Z_{n-1}) \middle| \begin{array}{l} X_{n+1} = \eta \\ X_n = \xi \end{array} \right) \right], \\ a_n^{(N)}(\xi) &= \mathbb{E} \left( F_n(\underline{L}_n) \middle| \begin{array}{l} X_{n+1} = \eta \\ X_n = \xi \end{array} \right) \\ a_{n+1}^{(N)}(\eta) &= \mathbb{E} \left( F_{n+1}(\underline{L}_{n+1}) \middle| \begin{array}{l} X_{n+1} = \eta \\ X_n = \xi \end{array} \right) \end{aligned}$$

*Proof of the claim.* The proof is based on Lemma 1.20. The first identity is because  $\{Z_n\}$  is independent from  $\{X_n\}$ . The second identity is because conditioned on  $X_n$ ,  $\underline{L}_n$  is independent of  $X_{n+1}$ . The third identity is because conditioned on  $X_{n+1}$ ,  $\underline{L}_{n+1}$  is independent of  $X_n$ .

With the claim proved, we can proceed to bound  $\tilde{f}$ . Taking the conditional expectation  $\mathbb{E}(\cdot | X_{n+1}^{(N)} = \eta, X_n^{(N)} = \xi)$  on both sides of (2.2.3), we find that

$$f_n^{(N)}(\xi, \eta) = a_{n+1}(\eta) - a_n(\xi) + c_n - \mathbb{E} \left( \Gamma_n(\underline{L}_n, \underline{L}_{n+1}) \middle| \begin{array}{l} X_{n+1} = \eta \\ X_n = \xi \end{array} \right),$$

whence  $\tilde{f}_n(\xi, \eta) := -\mathbb{E} \left( \Gamma_n(\underline{L}_n, \underline{L}_{n+1}) \middle| \begin{array}{l} X_{n+1} = \eta \\ X_n = \xi \end{array} \right)$ .

Clearly  $|\tilde{f}| \leq 6K$ . To bound the  $L^2$  norm we recall that the marginal distribution of  $\{X_n\}$  with respect to the distribution of the ladder process is precisely the distribution of our original array. Therefore

$$\begin{aligned}\|\tilde{f}_n^{(N)}\|_2^2 &\equiv \mathbb{E} \left[ \tilde{f}_n^{(N)}(X_n, X_{n+1})^2 \right] = \mathbb{E} \left[ \mathbb{E} \left( \Gamma_n(\underline{L}_n, \underline{L}_{n+1}) \middle| X_{n+1} X_n \right)^2 \right] \\ &\leq \mathbb{E} [\mathbb{E} (\Gamma_n(\underline{L}_n, \underline{L}_{n+1})^2)]\end{aligned}$$

because conditional expectations contract  $L^2$ -norms.

Next we use Lemma 1.20(3) to see that  $\Gamma_n^{(N)}(\underline{L}_n, \underline{L}_{n+1})$  is equal in distribution to the balance of a random level  $N$  hexagon at position  $n$ , whence  $\mathbb{E}(\Gamma_n^{(N)}) = (u_n^{(N)})^2$ .  $\square$

The gradient lemma splits an additive functional into a gradient term, and a term with controlled variance. The next lemma estimates the covariances between the two terms.

**Lemma 2.10** *Suppose  $f$  is a uniformly bounded functional of a uniformly elliptic Markov array. There is a constant  $C$  s.t. if  $h_{\ell_N}^{(N)}$  are uniformly bounded measurable functions on  $\mathfrak{S}_{\ell_N}^{(N)} \times \mathfrak{S}_{\ell_{N+1}}^{(N)}$ , and  $\text{ess sup } |f_n^{(N)}| \leq K$ ,  $\text{ess sup } |h_{\ell_N}^{(N)}| \leq L$ , then*

$$\text{Cov} \left( S_N, h_{\ell_N}^{(N)}(X_{\ell_N}^{(N)}, X_{\ell_{N+1}}^{(N)}) \right) \leq CKL.$$

*Proof.* This follows from the decomposition  $\text{Cov}(S_N, h_{\ell_N}^{(N)}) = \sum_{n=1}^{k_N} \text{Cov}(f_n^{(N)}, h_{\ell_N}^{(N)})$  and the exponential mixing of  $X$  (Proposition 1.11).  $\square$

### 2.2.2 The estimate for $\text{Var}(S_N)$

We prove Theorem 2.1. Let  $f = \{f_n^{(N)}\}$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X = \{X_n^{(N)}\}$  with row lengths  $k_N + 1$ . Our aim is to bound  $\text{Var}(S_N)$  above and below by affine functions of the structure constants  $U_N = \sum_{n=3}^{k_N} (u_n^{(N)})^2$ . Assume  $|f| \leq K$  almost surely.

Throughout the proof, we fix  $N$  and drop the superscripts  $(N)$ . So  $X_n^{(N)} = X_n$ ,  $f_n^{(N)} = f_n$ ,  $u_n^{(N)} = u_n$  etc.

**Lower bound for the variance.** Let's split  $U_N = \sum_{n=3}^{k_N} u_n^2$  into three sums:

$$U_N = \sum_{\gamma=0,1,2} U_N(\gamma), \text{ where } U_N(\gamma) := \left( \sum_{n=3}^{k_N} u_n^2 1_{[n=\gamma \bmod 3]}(n) \right).$$

For every  $N$  there is at least one  $\gamma_N \in \{0, 1, 2\}$  such that  $U_N(\gamma_N) \geq \frac{1}{3} U_N$ . Let

$$\alpha_N := \gamma_N + 1.$$

and define  $\beta_N$  by

$$k_N - \beta_N + 1 = \max \{n \leq k_N : n = \alpha_N \bmod 3\}.$$

With these choices,  $\alpha_N, \beta_N \in \{1, 2, 3\}$ , and  $k_N - \beta_N + 1 = \alpha_N \bmod 3$ .

We begin by bounding from below the variance of  $S'_N := \sum_{k=\alpha_N}^{k_N-\beta_N} f_j(X_j, X_{j+1})$ . Write  $k_N - \beta_N + 1 = 3M_N + \alpha_N$ , with  $M_N \in \mathbb{N}$ , then

$$S'_N = F_0 + \cdots + F_{M_N-1}, \text{ where } F_k := f_{3k+\alpha_N} + f_{3k+\alpha_N+1} + f_{3k+\alpha_N+2}.$$

Observe that  $S'_N$  is a function of the following variables:

$$\boxed{X_{\alpha_N}}, X_{\alpha_N+1}, X_{\alpha_N+2}, \boxed{X_{\alpha_N+3}}, X_{\alpha_N+4}, X_{\alpha_N+5}, \dots, \boxed{X_{k_N-\beta_N+1}},$$

where we have boxed the terms with indices congruent to  $\alpha_N \pmod{3}$ . Let  $\mathcal{F}_N$  denote the  $\sigma$ -algebra generated by the boxed random variables. Conditioned on  $\mathcal{F}_N$ ,  $F_k$  are independent. Therefore,

$$\text{Var}(S'_N | \mathcal{F}_N) = \sum_{k=0}^{M_N-1} \text{Var}(F_k | \mathcal{F}_N) = \sum_{k=0}^{M_N-1} \text{Var}(F_k | X_{3k+\alpha_N}, X_{3(k+1)+\alpha_N})$$

Taking the expectation on both sides, and using the general inequality  $\text{Var}(S'_N) \geq \mathbb{E}(\text{Var}(S'_N | \mathcal{F}_N))$ , we obtain

$$\text{Var}(S'_N) \geq \sum_{k=0}^{M_N-1} \mathbb{E} \left( \text{Var}(F_k | X_{3k+\alpha_N}, X_{3(k+1)+\alpha_N}) \right).$$

To estimate the summands, we recall that for every random variable  $W$ ,  $\text{Var}(W) = \frac{1}{2}\mathbb{E}[(W' - W'')^2]$  where  $W', W''$  are two independent copies of  $W$ . Thus

$$\begin{aligned} & \text{Var}(F_k | X_{3k+\alpha_N} = a, X_{3(k+1)+\alpha_N} = b) \\ &= \frac{1}{2} \mathbb{E} \left[ \Gamma \left( \begin{matrix} X_{3k+\alpha_N+1} & X_{3k+\alpha_N+2} & Y_{3(k+1)+\alpha_N} \\ X_{3k+\alpha_N} & Y_{3k+\alpha_N+1} & Y_{3k+\alpha_N+2} \end{matrix} \right)^2 \middle| \begin{matrix} X_{3k+\alpha_N} = Y_{3k+\alpha_N} = a \\ X_{3(k+1)+\alpha_N} = Y_{3(k+1)+\alpha_N} = b \end{matrix} \right], \end{aligned}$$

whence  $\mathbb{E}(\text{Var}(F_k | X_{3k+\alpha_N}, X_{3(k+1)+\alpha_N})) \equiv \mathbb{E}(\Gamma(P)^2) \equiv (u_{3k+\alpha_N+2}^{(N)})^2$  where  $\Gamma(P)$  is the balance of a random hexagon  $P \in \text{Hex}(N, 3k + \alpha_N + 2)$ . So

$$\begin{aligned} \text{Var}(S'_N) &\geq \frac{1}{2} \sum_{k=0}^{M_N-1} (u_{3k+\alpha_N+2}^{(N)})^2 = \frac{1}{2} \sum_{k=0}^{M_N-1} (u_{3(k+1)+\gamma_N}^{(N)})^2 \quad (\because \alpha_N = \gamma_N + 1) \\ &\geq \frac{1}{2} \sum_{n=3}^{k_N} u_n^2 \mathbf{1}_{[n=\gamma_n \pmod{3}]}(n) - 2 \sup\{u_j^2\} \geq \frac{1}{2} U_N(\gamma_N) - 2 \cdot (6K)^2 \\ &> \frac{1}{6} U_N - 100K^2, \text{ by choice of } \gamma_N. \end{aligned}$$

Now we claim that  $|\text{Var}(S_N) - \text{Var}(S'_N)|$  is uniformly bounded from below. To see this, let  $f_j^* := f_j - \mathbb{E}(f_j)$ , and let  $A_N := \{j \in \mathbb{N} : 1 \leq j \leq \alpha \text{ or } k_N - \beta \leq j \leq k_N\}$ . Then  $S_N = S'_N + \sum_{j \in A_N} f_j$ , whence

$$\text{Var}(S_N) = \text{Var}(S'_N) + \text{Var} \left( \sum_{j \in A_N} f_j \right) + 2 \sum_{j \in A_N} \text{Cov}(S'_N, f_j).$$

Since  $|f| \leq K$  and  $|A_N| \leq 6$ , the second term bounded by  $4K^2|A_N| \leq 24K^2$ . Next by uniform ellipticity and (1.2.5), there are mixing constants  $\theta \in (0, 1)$  and  $C_{mix} > 0$  which only depend on

$\varepsilon_0$ , the ellipticity constant of  $X$ , so that

$$\text{Cov}(S'_N, f_j) \leq C_{\text{mix}} \sum_{n=1}^{k_N} \|f_n^*\|_2 \|f_j^*\|_2 \theta^{|n-j|} \leq \frac{2C_{\text{mix}}K^2}{1-\theta}.$$

It follows that  $\text{Var}(S_N) \geq \text{Var}(S'_N) - \text{const} \geq \text{const}U_N - \text{const}$ , where the constants depends only on  $K$  and the ellipticity constant  $\varepsilon_0$ .

**Upper bound for the variance.** Write  $f = \tilde{f} + \nabla a + c$  as in the gradient lemma. In particular,  $\text{Var}(\tilde{f}_n(X_{n-1}, X_n)) \leq u_n^2$ . Then

$$\text{Var}\left(\sum_{n=1}^{k_N} f_n\right) = \text{Var}\left(\sum_{n=1}^{k_N} \tilde{f}_n\right) + \text{Var}(a_{N+1} - a_1) + 2\text{Cov}\left(\sum_{n=1}^{k_N} \tilde{f}_n, a_{N+1} - a_1\right).$$

The first term is smaller than  $C_1 U_N + C'_2$  due to the gradient Lemma and Lemma 2.5 the second term is smaller than  $C''_2$  due to Lemma 2.10.  $\square$

### 2.2.3 Characterization of center-tight additive functionals

We prove Theorem 2.2. Suppose  $f$  is an a.s. uniformly bounded functional on a uniformly elliptic array  $X$ . We will show that the following conditions are equivalent:

- (a)  $\text{Var}(S_N) = O(1)$ ;
- (b)  $f$  is the sum of a gradient and an additive functional with summable variance;
- (c)  $f$  is center tight.

**(a)  $\Rightarrow$  (b):** By the gradient lemma  $f = \nabla a + (\tilde{f} + c)$ , where  $a_n^{(N)}(x)$  are measurable functions on  $\mathfrak{S}_n^{(N)}$  with uniformly bounded  $L^\infty$  norm,  $c_n^{(N)}$  are uniformly bounded constants, and  $\|\tilde{f}_n\|_2 \leq u_n^{(N)}$ .

By Theorem 2.1,  $\sup_N \sum_{n=3}^{k_N} (u_n^{(N)})^2 < \infty$ , so  $\tilde{f} + c$  has summable variance, proving (b).

**(b)  $\Rightarrow$  (c):** We already saw that gradients and functionals with summable variance are center-tight. Since the sum of center-tight functionals is center-tight, (c) is proved.

**(c)  $\Rightarrow$  (a):** Assume by way of contradiction that  $\exists N_i \uparrow \infty$  such that  $V_{N_i} = \text{Var}(S_{N_i}) \rightarrow \infty$ . By Dobrushin's CLT (see [45], [139] and §2.2.5),  $\frac{S_{N_i} - \mathbb{E}(S_{N_i})}{\sqrt{V_{N_i}}}$  converges in distribution to a standard

Gaussian distribution. But center-tightness implies that there are constants  $\mu'_N$  s.t.  $\frac{S_N - \mu'_N}{\sqrt{V_N}}$  converges in distribution to the deterministic random variable  $W \equiv 0$ , and both statements cannot be true simultaneously.  $\square$

### 2.2.4 McLeish's martingale central limit theorem

A **martingale difference array** with row lengths  $k_N$  is a (possibly non-Markov) array  $\Delta$  of random variables

$$\Delta = \{\Delta_j^{(N)} : N \geq 1, 1 \leq j \leq k_N\}$$

together with an array of  $\sigma$ -algebras  $\{\mathcal{F}_j^{(N)} : N \geq 1, 1 \leq j \leq k_N\}$ , so that:

- (1) For each  $N$ ,  $\Delta_1^{(N)}, \dots, \Delta_{k_N}^{(N)}$  are random variables on the same probability space  $(\mathfrak{S}_N, \mathcal{F}_N, \mu_N)$ .
- (2)  $\mathcal{F}_1^{(N)} \subset \mathcal{F}_2^{(N)} \subset \mathcal{F}_3^{(N)} \subset \dots \subset \mathcal{F}_{k_N}^{(N)}$  are sub  $\sigma$ -algebras of  $\mathcal{F}_N$ .
- (3)  $\Delta_j^{(N)}$  is  $\mathcal{F}_j^{(N)}$ -measurable,  $\mathbb{E}(|\Delta_j^{(N)}|) < \infty$ , and  $\mathbb{E}(\Delta_{j+1}^{(N)} | \mathcal{F}_j^{(N)}) = 0$ .

We say that  $\Delta$  has **finite variance**, if every  $\Delta_j^{(N)}$  has finite variance. Notice that  $\mathbb{E}(\Delta_j^{(N)}) = 0$  for all  $j = 2, \dots, k_{N+1}$ . If in addition  $\mathbb{E}(\Delta_1^{(N)}) = 0$  for all  $N$ , then we say that  $\Delta$  has **zero mean**.

**Example 2.11** Suppose  $\{S_n\}$  is a martingale relative to  $\{\mathcal{F}_n\}$ , then

$$\Delta_1^{(N)} := S_1, \Delta_j^{(N)} := S_j - S_{j-1}, \mathcal{F}_j^{(N)} := \mathcal{F}_j, j = 1, \dots, N$$

is a martingale difference array.

The following basic observation on martingale difference arrays is a key to many of their properties:

**Lemma 2.12** Suppose  $\Delta$  is a martingale difference array with finite variance, then for each  $N$   $\Delta_1^{(N)}, \dots, \Delta_{k_N}^{(N)}$  are uncorrelated, and if  $\Delta$  has zero mean, then

$$\text{Var}\left(\sum_{n=1}^{k_N} \Delta_n^{(N)}\right) = \sum_{n=1}^{k_N} \mathbb{E}[(\Delta_n^{(N)})^2].$$

*Proof.* Fix  $N$  and write  $\Delta_j^{(N)} = \Delta_j, \mathcal{F}_j^{(N)} = \mathcal{F}_j$ .

If  $i < j$ , then  $\mathbb{E}(\Delta_j \Delta_i) = \mathbb{E}[\mathbb{E}(\Delta_j \Delta_i | \mathcal{F}_{j-1})] = \mathbb{E}[\mathbb{E}(\Delta_i \mathbb{E}(\Delta_j | \mathcal{F}_{j-1}))] = \mathbb{E}(\Delta_i \cdot 0) = 0$ . The identity for the variance immediately follows.  $\square$

**Theorem 2.5 (McLeish's Martingale Central Limit Theorem).** Let  $\Delta = \{\Delta_j^{(N)}\}$  be a martingale difference array with row lengths  $k_N$ , zero mean, and finite variance, and let  $V_N := \sum_{j=1}^{k_N} \mathbb{E}[(\Delta_j^{(N)})^2]$ . Suppose:

- (1)  $\max_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}}$  has uniformly bounded  $L^2$  norm;
- (2)  $\max_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}} \xrightarrow[N \rightarrow \infty]{} 0$  in probability; and
- (3)  $\frac{1}{\sqrt{V_N}} \sum_{j=1}^{k_N} (\Delta_j^{(N)})^2 \xrightarrow[N \rightarrow \infty]{} 1$  in probability.

Then for all intervals  $(a, b)$ ,  $\mathbb{P}\left[\frac{1}{\sqrt{V_N}} \sum_{j=1}^{k_N} \Delta_j^{(N)} \in (a, b)\right] \xrightarrow[N \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$ .

We prepare the ground for the proof.

A sequence of random variables  $\{Y_n\}$  on  $(\Omega, \mathcal{F}, \mu)$  is called **uniformly integrable** if for every  $\varepsilon$ ,  $\exists K$  s.t.  $\mathbb{E}(|Y_n| \mathbf{1}_{[|Y_n| > K]}) < \varepsilon$  for all  $n$ . This is strictly stronger than tightness (there are tight non-integrable random variables).

**Example 2.13** If  $M_p := \sup_n \|Y_n\|_p < \infty$  for some  $p > 1$ , then  $\{Y_n\}$  is uniformly integrable.

Indeed, by Chebyshev's inequality,  $\mu[|Y_n| > K] \leq \frac{1}{K^p} M_p^p$ , and by Hölder's inequality  $\mathbb{E}(|Y_n| \mathbf{1}_{[|Y_n| > K]}) \leq M_p \mu[|Y_n| > K]^{1/q} = O(K^{-p/q})$  for the  $q$  s.t.  $\frac{1}{p} + \frac{1}{q} = 1$ .

**Lemma 2.14** Suppose  $Y_n, Y \in L^1(\Omega, \mathcal{F}, \mu)$ , then  $Y_n \xrightarrow[n \rightarrow \infty]{L^1} Y$  iff  $\{Y_n\}$  are uniformly integrable and  $Y_n \xrightarrow[n \rightarrow \infty]{\text{in probability}} Y$ . In this case  $\mathbb{E}(Y_n) \xrightarrow[n \rightarrow \infty]{\text{in probability}} \mathbb{E}(Y)$ .

*Proof.* We include the well-known, standard proof for completeness.

Proof of  $(\Rightarrow)$ : Since  $Y \in L^1$ , it follows (for example from the Dominated Convergence Theorem) that  $\lim_{K \rightarrow \infty} \mathbb{E}(|Y|1_{|Y| \geq K}) = 0$ . Given  $\varepsilon$  take  $K$  so that  $\mathbb{E}(|Y|1_{|Y| \geq K}) < \varepsilon$ . Let  $\delta = \mathbb{P}(|Y| \leq K)$  then it is easy to see that

$$\mathbb{E}(|Y|1_F) < \varepsilon \text{ for all measurable sets } F \text{ s.t. } \mu(F) < \delta. \quad (2.2.6)$$

Fix  $\varepsilon > 0$ , and choose  $\delta$  as in (2.2.6).

Suppose  $\|Y_n - Y\|_1 \rightarrow 0$ . By Markov's inequality  $\mathbb{P}[|Y_n - Y| > \varepsilon] \leq \|Y_n - Y\|_1 / \varepsilon \rightarrow 0$ , and  $Y_n \rightarrow Y$  in probability.

Markov's inequality also implies that  $\mathbb{P}[|Y_n| > K] \leq K^{-1} \sup \|Y_n\|_1 = O(K^{-1})$ , so there exists  $K$  s.t.  $\mathbb{P}[|Y_n| > K] < \delta$  for all  $n$ . By the choice of  $\delta$ ,

$$\int_{[|Y_n| > K]} |Y_n| d\mu \leq \int_{[|Y_n| > K]} |Y| d\mu + \int_{[|Y_n| > K]} |Y_n - Y| d\mu \leq \varepsilon + \|Y_n - Y\|_1 \xrightarrow[n \rightarrow \infty]{\text{in probability}} \varepsilon.$$

Uniform integrability follows.

Proof of  $(\Leftarrow)$ : Given a random variable  $Z$ , let  $Z^K := Z1_{[|Z| \leq K]}$ . Since  $\{Y_n\}$  is uniformly integrable, for every  $\varepsilon$  there is a  $K > 1$  s.t.  $\|Y_n^K - Y_n\|_1 < \varepsilon$  for all  $n$ . Similarly,  $\|Y^K - Y\|_1 < \varepsilon$  for all  $K$  large enough. Thus for all  $n$ ,

$$\begin{aligned} \|Y_n - Y\|_1 &\leq \|Y_n^K - Y^K\|_1 + 2\varepsilon \leq \varepsilon \mu[|Y_n^K - Y^K| \leq \varepsilon] + 2K\mu[|Y_n^K - Y^K| > \varepsilon] + 2\varepsilon \\ &\leq 3\varepsilon + 2K \left( \mu[|Y_n - Y| > \varepsilon] + \mu[|Y_n| > K] + \mu[|Y| > K] \right) \\ &\leq 3\varepsilon + 2K\mu[|Y_n - Y| > \varepsilon] + 2\mathbb{E}(|Y_n|1_{[|Y_n| > K]}) + 2\mathbb{E}(|Y|1_{[|Y| > K]}) \\ &\therefore \limsup_{n \rightarrow \infty} \|Y_n - Y\|_1 \leq 3\varepsilon + 2\sup_n \mathbb{E}(|Y_n|1_{[|Y_n| > K]}) + 2\mathbb{E}(|Y|1_{[|Y| > K]}), \end{aligned}$$

where we have used the assumption that  $Y_n \rightarrow Y$  in probability. The last expression can be made arbitrarily small, by choosing  $\varepsilon$  sufficiently small,  $K$  sufficiently large, and appealing to the uniform integrability of  $Y_n$ .  $\square$

**Lemma 2.15 (McLeish)** Let  $\{W_j^{(N)} : 1 \leq j \leq k_N\}$  be a triangular array of random variables<sup>2</sup>, where  $W_1^{(N)}, \dots, W_{k_N}^{(N)}$  are defined on the same probability space. Fix  $t \in \mathbb{R}$  and let  $T_N(t) := \prod_{j=1}^{k_N} (1 + itW_j^{(N)})$ . Suppose

- (1)  $\{T_N(t)\}$  is uniformly integrable and  $\mathbb{E}(T_N) \xrightarrow[N \rightarrow \infty]{\text{in probability}} 1$ ,
- (2)  $\sum_{j=1}^{k_N} (W_j^{(N)})^2 \xrightarrow[N \rightarrow \infty]{\text{in probability}} 1$ ,
- (3)  $\max_{1 \leq j \leq k_N} |W_j^{(N)}| \xrightarrow[N \rightarrow \infty]{\text{in probability}} 0$ .

Then  $\mathbb{E}(e^{it(W_1^{(N)} + \dots + W_{k_N}^{(N)})}) \xrightarrow[N \rightarrow \infty]{\text{in probability}} e^{-\frac{1}{2}t^2}$ .

<sup>2</sup> Not necessarily a martingale difference array or a Markov array.

*Proof.* Define a function  $r(x)$  on  $[-1, 1]$  by the identity  $e^{ix} = (1 + ix)e^{-\frac{1}{2}x^2 + r(x)}$ , then  $r(x) = -\log(1 + ix) + ix + \frac{1}{2}x^2 = O(|x|^3)$ . Fix  $C$  s.t.  $|r(x)| \leq C|x|^3$  for  $|x| < 1$ .

Substituting  $S_N := W_1^{(N)} + \dots + W_{k_N}^{(N)}$  in  $e^{ix} = (1 + ix)e^{-\frac{1}{2}x^2 + r(x)}$  gives (in what follows we drop the superscripts  $(N)$  and abbreviate  $T_n := T_n(t)$ ):

$$\begin{aligned}\mathbb{E}(e^{itS_N}) &= \mathbb{E}\left(\prod_{j=1}^{k_N} e^{itW_j}\right) = \mathbb{E}(T_N e^{-\frac{1}{2}\sum_{j=1}^{k_N} t^2 W_j^2 + r(tW_j)}) \\ &= \mathbb{E}(T_N U_N), \text{ where } U_N := \exp\left[-\frac{1}{2} \sum_{j=1}^{k_N} t^2 (W_j^{(N)})^2 + r(tW_j^{(N)})\right].\end{aligned}$$

$T_N$  and  $U_N$  have the following properties:

(a)  $\mathbb{E}(T_N) \xrightarrow[N \rightarrow \infty]{} 1$ , by assumption.

(b)  $\{T_N\}$  is uniformly integrable by assumption, and  $|T_N U_N| = |e^{itS_N}| = 1$ .

(c)  $U_N \xrightarrow[N \rightarrow \infty]{\text{prob}} \exp\left(-\frac{1}{2}t^2\right)$ , because

o  $\sum_{j=1}^{k_N} (W_j^{(N)})^2 \xrightarrow[N \rightarrow \infty]{\text{prob}} 1$ , by assumption,

o  $\max_{1 \leq j \leq k_N} |W_j^{(N)}| \xrightarrow[n \rightarrow \infty]{\text{prob}} 0$  by assumption, so with asymptotic probability one,

$$\left| \sum_{i=1}^{k_N} r(tW_j^{(N)}) \right| \leq C|t|^3 \max_{1 \leq j \leq k_N} |W_j^{(N)}| \sum_{j=1}^{k_N} (W_j^{(N)})^2 \xrightarrow[N \rightarrow \infty]{\text{prob}} 0.$$

We claim that this implies that  $\mathbb{E}(e^{itS_N}) = \mathbb{E}(T_N U_N) \xrightarrow[N \rightarrow \infty]{} e^{-\frac{1}{2}t^2}$ . Let  $L := e^{-\frac{1}{2}t^2}$ . Since  $|\mathbb{E}(T_N U_N) - L| \leq |\mathbb{E}(T_N(U_N - L))| + L|\mathbb{E}(T_N) - 1|$ , (a) tells us that

$$|\mathbb{E}(T_N U_N) - L| \leq |\mathbb{E}(T_N(U_N - L))| + o(1). \quad (2.2.7)$$

Next, for every  $K, \varepsilon, \mu$ ,  $\mu[|T_N(U_N - L)| > \varepsilon] \leq \mu[|T_N| > K] + \mu[|U_N - L| > \varepsilon/K]$ . Therefore by (b) and (c),

$$T_N(U_N - L) \xrightarrow[N \rightarrow \infty]{} 0 \text{ in probability.} \quad (2.2.8)$$

Finally,  $|T_N(U_N - L)| \leq 1 + L|T_N|$ , so  $T_N(U_N - L)$  is uniformly integrable by (b). By Lemma 2.14,  $\mathbb{E}(T_N(U_N - L)) \rightarrow 0$ , and by (2.2.7),  $\mathbb{E}(e^{itS_N}) = \mathbb{E}(T_N U_N) \rightarrow e^{-\frac{1}{2}t^2}$ .  $\square$

**Proof of the Martingale CLT [103]:** Let  $\Delta = \{\Delta_j^{(N)}\}$  be a martingale difference array with row lengths  $k_N$ , which satisfies the assumptions of Theorem 2.5, and let

$$S_N := \sum_{j=1}^{k_N} \Delta_j^{(N)} \text{ and } V_N := \text{Var}(S_N) \equiv \sum_{j=1}^{k_N} \mathbb{E}[(\Delta_j^{(N)})^2] \text{ (see Lemma 2.12).}$$

It is tempting to apply McLeish's Lemma to the normalized array  $\Delta_j^{(N)} / \sqrt{V_N}$ , but to do this we need to check the uniform integrability of  $\prod_{j=1}^n (1 + it\Delta_j^{(N)} / \sqrt{V_N})$  and this is difficult. It is

easier to work with the following array of truncations:

$$W_1^{(N)} := \frac{1}{\sqrt{V_N}} \Delta_1^{(N)}, \quad W_n^{(N)} := \frac{1}{\sqrt{V_N}} \Delta_n^{(N)} \mathbf{1}_{[\sum_{k=1}^{n-1} (\Delta_k^{(N)})^2 \leq 2V_N]}.$$

It is easy to check that  $\{W_n^{(N)}\}$  is a martingale difference array relative to  $\mathcal{F}_n^{(N)}$ , and that  $\{W_n^{(N)}\}$  has zero mean, and finite variance.

In addition,  $S_N^* := \sum_{n=1}^{k_N} W_n^{(N)}$  are close to  $S_N / \sqrt{V_N}$  in probability:

$$\mu[S_N^* \neq \frac{S_N}{\sqrt{V_N}}] \leq \mu \left[ \exists 1 \leq j \leq k_N \text{ s.t. } \sum_{k=1}^{j-1} (\Delta_k^{(N)})^2 > 2V_N \right] \leq \mu \left[ \sum_{j=1}^{k_N} (\Delta_j^{(N)})^2 > 2V_N \right] \xrightarrow[N \rightarrow \infty]{} 0$$

because  $\frac{1}{V_N} \sum_{j=1}^{k_N} (\Delta_j^{(N)})^2 \xrightarrow[N \rightarrow \infty]{\text{prob}} 1$  by assumption.

Thus to prove the theorem, it is enough to show that  $S_N^*$  converges in distribution to the standard Gaussian distribution. To do this, we check that  $\{W_n^{(N)}\}$  satisfies the conditions of McLeish's Lemma.

Fix  $t \in \mathbb{R}$ , and let  $T_N = T_N(t) := \prod_{j=1}^{k_N} (1 + it W_j^{(N)})$ . Let  $J_N := \max \{2 \leq j \leq k_N : \sum_{k=1}^{j-1} (\Delta_k^{(N)})^2 \leq 2V_N\}$  (or  $J_N = 1$  if the maximum is over the empty set). Writing  $W_j = W_j^{(N)}$  and  $\Delta_j = \Delta_j^{(N)}$ , we obtain

$$\begin{aligned} |T_N| &= \prod_{j=1}^{k_N} (1 + t^2 W_j^2)^{1/2} = \prod_{j=1}^{J_N} \left( 1 + \frac{t^2 \Delta_j^2}{V_N} \right)^{1/2} \\ &= \left( \prod_{j=1}^{J_N-1} \left( 1 + \frac{t^2 \Delta_j^2}{V_N} \right) \right)^{1/2} \cdot \left( 1 + \frac{t^2 \Delta_{J_N}^2}{V_N} \right)^{1/2}, \text{ where } \prod_{j=1}^0 (\dots) := 1 \\ &\leq \exp \left( \frac{t^2}{2V_N} \sum_{j=0}^{J_N-1} \Delta_j^2 \right) \left( 1 + \frac{t^2}{V_N} \Delta_{J_N}^2 \right)^{1/2} \leq e^{t^2} \left( 1 + |t| \max_{1 \leq j \leq k_N} \left| \frac{\Delta_j^{(N)}}{\sqrt{V_N}} \right| \right). \end{aligned}$$

Thus

$$\|T_N(t)\|_2^2 \leq e^{2t^2} \left( 1 + |t| \mathbb{E} \left( \max_{1 \leq j \leq k_N} \left| \frac{\Delta_j^{(N)}}{\sqrt{V_N}} \right| \right) \right)^2.$$

By the first assumption of the theorem, the last quantity is uniformly bounded for each  $t$ . It follows that  $\{T_N(t)\}_{N \geq 1}$  is uniformly integrable for each  $t$ . Next, successive conditioning shows that  $\mathbb{E}(T_N) = 1 + it \mathbb{E}(\Delta_1^{(N)}) = 1$ . The first condition of McLeish's Lemma is verified.

The second condition of McLeish's Lemma follows from the assumption  $\frac{1}{V_N} \sum_{n=1}^{k_N} (\Delta_n^{(N)})^2 \rightarrow 0$  in probability, and the estimate

$$\mu \left[ \sum_{n=1}^{k_N} (W_n^{(N)})^2 \neq \sum_{n=1}^{k_N} \left( \frac{\Delta_n^{(N)}}{\sqrt{V_N}} \right)^2 \right] \leq \mu \left[ \exists 1 \leq n \leq k_N \text{ s.t. } \sum_{j=1}^n (\Delta_j^{(N)})^2 > 2V_N \right] \leq$$

$$\leq \mu \left[ \sum_{n=1}^{k_N} (\Delta_n^{(N)})^2 > 2V_N \right] \xrightarrow[N \rightarrow \infty]{} 0, \text{ because } \frac{1}{V_n} \sum_{j=1}^{k_N} (\Delta_j^{(N)})^2 \rightarrow 1 \text{ in probability.}$$

The third condition of McLeish's Lemma follows from the assumption that

$$\max_{1 \leq j \leq k_N} |W_j^{(N)}| \rightarrow 0 \text{ in probability, for similar reasons.}$$

So McLeish's lemma applies to  $\{W_n^{(N)}\}$ , and  $\mathbb{E}(e^{itS_N^*}) \rightarrow e^{-\frac{1}{2}t^2}$  for all  $t \in \mathbb{R}$ . By Lévy's continuity theorem, this implies that  $S_N^* \xrightarrow[N \rightarrow \infty]{\text{dist}} N(0, 1)$ .

As explained above, this implies that  $\frac{S_N}{\sqrt{V_N}} \xrightarrow[N \rightarrow \infty]{\text{dist}} N(0, 1)$ .  $\square$

### 2.2.5 Proof of Dobrushin's central limit theorem

Let  $\mathbb{X} = \{X_n^{(N)}\}$  be a uniformly elliptic Markov array with row lengths  $k_N + 1$ , and let  $\mathbf{f} = \{f_n^{(N)}\}$  be an a.s. uniformly bounded additive functional on  $\mathbb{X}$ . Define as before  $S_N = \sum_{n=1}^{k_N} f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})$ ,  $V_N := \text{Var}(S_N)$ . Without loss of generality,

$$\mathbb{E}[f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})] = 0 \text{ and } |f_n^{(N)}| \leq K \text{ for all } n, N.$$

Define  $\mathcal{F}_n^{(N)} := \sigma(X_1^{(N)}, \dots, X_{n+1}^{(N)})$  for  $n \geq 1$ , and  $\mathcal{F}_0^{(N)} :=$  trivial  $\sigma$ -algebra. Fix  $N$  and write  $f_k = f_k^{(N)}(X_k^{(N)}, X_{k+1}^{(N)})$  and  $\mathcal{F}_k = \mathcal{F}_k^{(N)}$ , then  $\mathbb{E}(f_k | \mathcal{F}_k) = f_k$ ,  $\mathbb{E}(f_k | \mathcal{F}_0) = \mathbb{E}(f_k) = 0$ , and therefore

$$\begin{aligned} S_N &= \sum_{k=1}^{k_N} f_k = \sum_{k=1}^{k_N} (\mathbb{E}(f_k | \mathcal{F}_k) - \mathbb{E}(f_k | \mathcal{F}_0)) = \sum_{k=1}^{k_N} \sum_{n=1}^k (\mathbb{E}(f_k | \mathcal{F}_n) - \mathbb{E}(f_k | \mathcal{F}_{n-1})) \\ &= \sum_{n=1}^{k_N} \sum_{k=n}^{k_N} (\mathbb{E}(f_k | \mathcal{F}_n) - \mathbb{E}(f_k | \mathcal{F}_{n-1})) \\ &= \sum_{n=1}^{k_N} \Delta_n^{(N)}, \text{ where } \Delta_n^{(N)} := \sum_{k=n}^{k_N} (\mathbb{E}(f_k^{(N)} | \mathcal{F}_n^{(N)}) - \mathbb{E}(f_k^{(N)} | \mathcal{F}_{n-1}^{(N)})). \end{aligned}$$

The array  $\{\Delta_n^{(N)} : 1 \leq n \leq k_N; N \geq 1\}$  is a martingale difference array relative to the filtrations  $\mathcal{F}_n^{(N)}$ , with zero mean and finite variances. To prove the theorem, it suffices to check that  $\{\Delta_n^{(N)}\}$  satisfies the conditions of the martingale CLT.

STEP 1:  $\max_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}}$  has uniformly bounded  $L^2$  norm, and  $\max_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}} \xrightarrow[N \rightarrow \infty]{\text{prob}} 0$ .

*Proof.* The proof is based on the exponential mixing of uniformly elliptic Markov arrays (Proposition 1.11): Let  $K := \text{ess sup } |\mathbf{f}|$ , then there are constants  $C_{\text{mix}} > 1$  and  $0 < \theta < 1$  such that for all  $k \geq n$ ,

$$\|\mathbb{E}(f_k^{(N)} | \mathcal{F}_n^{(N)})\|_\infty \leq C_{\text{mix}} K \theta^{k-n-1}.$$

It follows that  $|\Delta_j^{(N)}| < 2C_{mix}K\sum_{\ell=-1}^{\infty} \theta^\ell = \frac{2C_{mix}K\theta^{-2}}{1-\theta}$ . The step follows from the assumption that  $V_N \rightarrow \infty$ .

STEP 2:  $\frac{1}{V_N} \sum_{n=1}^{k_N} (\Delta_n^{(N)})^2 \xrightarrow[N \rightarrow \infty]{} 1$  in probability.

*Proof.* We follow [139] closely.

Let  $Y_i^{(N)} := (\Delta_i^{(N)})^2/V_N$ . We will show that  $\left\| \sum_{i=1}^{k_N} Y_i^{(N)} - 1 \right\|_2^2 \xrightarrow[N \rightarrow \infty]{} 0$ , and use the general fact that  $L^2$ -convergence implies convergence in probability (by Chebyshev's inequality).

Notice that  $\mathbb{E} \left( \sum_{i=1}^{k_N} Y_i^{(N)} \right) = 1$ , because by Lemma 2.12, this expectation equals

$$\frac{1}{V_N} \times \text{Var} \left( \sum_{n=1}^{k_N} \Delta_n^{(N)} \right) = \frac{1}{V_N} \text{Var}(S_N) = 1.$$

So

$$\begin{aligned} \left\| \sum_{i=1}^{k_N} Y_i^{(N)} - 1 \right\|_2^2 &= \mathbb{E} \left[ \left( \sum_{i=1}^{k_N} Y_i^{(N)} \right)^2 \right] - 2\mathbb{E} \left[ \sum_{i=1}^{k_N} Y_i^{(N)} \right] + 1 \\ &= \mathbb{E} \left[ \sum_{i=1}^{k_N} (Y_i^{(N)})^2 \right] + 2\mathbb{E} \left[ \sum_{i < j} Y_i^{(N)} Y_j^{(N)} \right] - 2 + 1 \\ &= O \left( \max_{1 \leq \ell \leq k_N} \|Y_\ell^{(N)}\|_\infty \right) \cdot \mathbb{E} \left[ \sum_{\ell=1}^{k_N} Y_\ell^{(N)} \right] + 2\mathbb{E} \left[ \sum_{i < j} Y_i^{(N)} Y_j^{(N)} \right] - 1. \end{aligned}$$

We saw in the proof of step 1 that  $\|\Delta_j^{(N)}\|_\infty$  are uniformly bounded. Thus  $\max_{1 \leq \ell \leq k_N} \|Y_\ell^{(N)}\|_\infty = O(1/V_N)$ , so  $\left\| \sum_{i=1}^{k_N} Y_i^{(N)} - 1 \right\|_2^2 = 2\mathbb{E} \left[ \sum_{i < j} Y_i^{(N)} Y_j^{(N)} \right] - 1 + o(1)$ . It remains to show that

$$2\mathbb{E} \left[ \sum_{i < j} Y_i^{(N)} Y_j^{(N)} \right] \xrightarrow[N \rightarrow \infty]{} 1. \quad (2.2.9)$$

The proof of (2.2.9) is based on the following fact:

$$\text{Osc}(N) := \max_{1 \leq i \leq k_N} \text{Osc} \left( \mathbb{E} \left( \sum_{j=i+1}^{k_N} Y_j^{(N)} \middle| \mathcal{F}_i^{(N)} \right) \right) \xrightarrow[N \rightarrow \infty]{} 0. \quad (2.2.10)$$

Here  $\text{Osc}$  is the oscillation, which was defined in §1.2.1. Before proving this, we explain why (2.2.10) implies (2.2.9). Write  $x = y \pm \varepsilon$  whenever  $y - \varepsilon \leq x \leq y + \varepsilon$ . Every bounded function  $\varphi$  satisfies  $\varphi = \mathbb{E}(\varphi) \pm \text{Osc}(\varphi)$ . So

$$\begin{aligned} 2\mathbb{E} \left[ \sum_{i < j} Y_i^{(N)} Y_j^{(N)} \right] &= 2\mathbb{E} \left[ \sum_{i=1}^{k_N} Y_i^{(N)} \sum_{j=i+1}^{k_N} Y_j^{(N)} \right] = 2\mathbb{E} \left[ \sum_{i=1}^{k_N} Y_i^{(N)} \mathbb{E} \left( \sum_{j=i+1}^{k_N} Y_j^{(N)} \middle| \mathcal{F}_i^{(N)} \right) \right] \\ &= 2\mathbb{E} \left[ \sum_{i=1}^{k_N} Y_i^{(N)} \mathbb{E} \left( \sum_{j=i+1}^{k_N} Y_j^{(N)} \right) \right] \pm 2\mathbb{E} \left[ \sum_{i=1}^{k_N} Y_i^{(N)} \right] \text{Osc}(N) \end{aligned}$$

$$\begin{aligned}
&= 2 \sum_{i=1}^{k_N} \mathbb{E}(Y_i^{(N)}) \sum_{j=i+1}^{k_N} \mathbb{E}(Y_j^{(N)}) \pm 2 \text{Osc}(N) \quad (\because \sum_{i=1}^{k_N} \mathbb{E}(Y_i^{(N)}) = 1) \\
&= \left( \sum_{i=1}^{k_N} \mathbb{E}(Y_i^{(N)}) \right)^2 - \sum_{i=1}^{k_N} \mathbb{E}(Y_i^{(N)})^2 \pm 2 \text{Osc}(N) \\
&= 1 + O\left( \max_{1 \leq i \leq k_N} \|Y_i^{(N)}\|_\infty \right) \pm 2 \text{Osc}(N), \quad \because \sum \mathbb{E}(Y_i^{(N)})^2 \leq \underbrace{\sum \mathbb{E}(Y_i^{(N)})}_{=1} \max \|Y_i^{(N)}\|_\infty \\
&= 1 + O(V_N^{-1}) + O(\text{Osc}(N)).
\end{aligned}$$

So (2.2.10) implies (2.2.9), and with it the step.

We turn to the proof of (2.2.10). Henceforth we fix  $N$  and drop all the  $^{(N)}$  superscripts. First we note that a routine modification of the proof of Lemma 2.12 shows that for all  $j, k > i$ ,  $\mathbb{E}(\Delta_j \Delta_k | \mathcal{F}_i) = 0$ . It follows that

$$\begin{aligned}
\mathbb{E}\left(\sum_{j=i+1}^{k_N} Y_j \middle| \mathcal{F}_i\right) &\equiv \frac{1}{V_N} \mathbb{E}\left(\sum_{j=i+1}^{k_N} \Delta_j^2 \middle| \mathcal{F}_i\right) = \frac{1}{V_N} \mathbb{E}\left(\left(\sum_{n=i+1}^{k_N} \Delta_n\right)^2 \middle| \mathcal{F}_i\right) \\
&= \frac{1}{V_N} \mathbb{E}\left(\left(\sum_{n=i+1}^{k_N} \sum_{k=n}^{k_N} [\mathbb{E}(f_k | \mathcal{F}_n) - \mathbb{E}(f_k | \mathcal{F}_{n-1})]\right)^2 \middle| \mathcal{F}_i\right) \\
&= \frac{1}{V_N} \mathbb{E}\left(\left(\sum_{k=i+1}^{k_N} \sum_{n=i+1}^k [\mathbb{E}(f_k | \mathcal{F}_n) - \mathbb{E}(f_k | \mathcal{F}_{n-1})]\right)^2 \middle| \mathcal{F}_i\right) \\
&= \frac{1}{V_N} \mathbb{E}\left(\left(\sum_{k=i+1}^{k_N} [f_k - \mathbb{E}(f_k | \mathcal{F}_i)]\right)^2 \middle| \mathcal{F}_i\right) \\
&= \frac{1}{V_N} \sum_{k, \ell=i+1}^{k_N} \mathbb{E}\left([(f_k - \mathbb{E}(f_k | \mathcal{F}_i))(f_\ell - \mathbb{E}(f_\ell | \mathcal{F}_i))] \middle| \mathcal{F}_i\right) \\
&= \frac{1}{V_N} \sum_{k, \ell=i+1}^{k_N} \mathbb{E}\left[f_k f_\ell + \mathbb{E}(f_k | \mathcal{F}_i) \mathbb{E}(f_\ell | \mathcal{F}_i) - f_k \mathbb{E}(f_\ell | \mathcal{F}_i) - f_\ell \mathbb{E}(f_k | \mathcal{F}_i) \middle| \mathcal{F}_i\right] \\
&= \frac{1}{V_N} \sum_{k, \ell=i+1}^{k_N} [\mathbb{E}[f_k f_\ell | \mathcal{F}_i] - \mathbb{E}(f_\ell | \mathcal{F}_i) \mathbb{E}(f_k | \mathcal{F}_i)] \tag{2.2.11}
\end{aligned}$$

The oscillation of the summands can be estimated as follows. By Lemma 1.9(d)

$$\text{Osc}\left(\mathbb{E}(u(X_k^{(N)}, X_{k+1}^{(N)}) | X_j^{(N)})\right) \leq \delta(\pi_{j,k}^{(N)}) \text{Osc}(u),$$

where  $\delta(\pi_{j,k}^{(N)})$  is the contraction coefficient of the  $(k-j)$ -step Markov operator  $\pi_{j,k}^{(N)}$ . In the uniformly elliptic case, by Lemma 1.10,  $\delta(\pi_{j,j+2}^{(N)}) \leq 1 - \varepsilon_0$ , where  $\varepsilon_0 > 0$  is the ellipticity constant of  $\mathbf{X}$ . Iterating Lemma 1.9(c) we conclude that there exists  $C_0 > 0$  and  $0 < \theta < 1$  such that for all  $k > i + 1$ , and for every bounded function  $u : \mathfrak{S}_k^{(N)} \times \mathfrak{S}_{k+1}^{(N)} \rightarrow \mathbb{R}$ ,

$$\text{Osc}\left(\mathbb{E}(u(X_k^{(N)}, X_{k+1}^{(N)}) | \mathcal{F}_i^{(N)})\right) \leq C_0 \theta^{k-i} \text{Osc}(u).$$

This, (1.2.3), and the inequalities  $|f_j| \leq K$ ,  $\text{Osc}(u) \leq 2\|u\|_\infty$  and  $\text{Osc}(uv) \leq \|u\|_\infty \text{Osc}(v) + \|v\|_\infty \text{Osc}(u)$  imply the existence of constants  $C_1 > 0$  and  $0 < \theta < 1$  such that for every  $N \geq 1$  and  $i+2 \leq k \leq \ell \leq k_N$ ,

$$\begin{aligned} & \text{Osc}(\mathbb{E}(f_\ell|\mathcal{F}_i)\mathbb{E}(f_k|\mathcal{F}_i)) \\ & \leq \text{Osc}(\mathbb{E}(f_\ell|\mathcal{F}_i))\|\mathbb{E}(f_k|\mathcal{F}_i)\|_\infty + \|\mathbb{E}(f_\ell|\mathcal{F}_i)\|_\infty \text{Osc}(\mathbb{E}(f_k|\mathcal{F}_i)) \leq C_1 \theta^{k-i} \theta^{\ell-i}. \\ & \text{Osc}\left(\mathbb{E}[f_k f_\ell|\mathcal{F}_i]\right) = \text{Osc}\left(\mathbb{E}[f_k \mathbb{E}(f_\ell|\mathcal{F}_k)|\mathcal{F}_i]\right) \\ & \leq C_0 \theta^{k-i} \text{Osc}(f_k \mathbb{E}(f_\ell|\mathcal{F}_k)) \leq C_0 \theta^{k-i} [K \cdot \text{Osc}(\mathbb{E}(f_\ell|\mathcal{F}_k)) + \text{Osc}(f_k) \|\mathbb{E}(f_\ell|\mathcal{F}_k)\|_\infty] \\ & \leq C_1 \theta^{k-i} \theta^{\ell-k}. \end{aligned}$$

We have stated these bounds for  $k, \ell \geq i+2$ , but in fact they remain valid for  $k = i+2$  or  $\ell = i+2$ , if we increase  $C_1$  to guarantee that  $C_1 \theta^2 > 2K^2$ .

Substituting these bounds in (2.2.11), we find that

$$\text{Osc}(N) \leq \frac{2C_1}{V_N} \sum_{k, \ell=i+1}^{\infty} \theta^{k-i} \theta^{\ell-k} \leq \frac{2C_1}{V_N} \left( \frac{\theta}{1-\theta} \right)^2 \xrightarrow{N \rightarrow \infty} 0.$$

This proves (2.2.10), and completes the proof of step 2.

Steps 1 and 2 verify the conditions of the martingale CLT. So  $\frac{1}{\sqrt{V_N}} \sum_{n=1}^{k_N} \Delta_n^{(N)}$  converges in distribution to the standard Gaussian distribution. By construction,  $\frac{1}{\sqrt{V_N}} S_N \equiv \frac{1}{\sqrt{V_N}} \sum_{n=1}^{k_N} \Delta_n^{(N)}$ , and the theorem is proved.  $\square$

### 2.2.6 Almost sure convergence for sums of functionals with summable variance

We prove Proposition 2.4. Let  $f_0^* := 0$ ,  $f_n^* := f_n(X_n, X_{n+1}) - \mathbb{E}f_n(X_n, X_{n+1})$ , let  $\mathcal{F}_0$  denote the trivial  $\sigma$ -algebra, and let  $\mathcal{F}_n$  denote the  $\sigma$ -algebra generated by  $X_1, \dots, X_n$ . Then  $f_k^*$  is  $\mathcal{F}_{k+1}$ -measurable, so

$$f_k^* = \mathbb{E}(f_k^*|\mathcal{F}_{k+1}) - \mathbb{E}(f_k^*|\mathcal{F}_0) = \sum_{n=0}^k \mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n).$$

Therefore (numbered equalities are justified below):

$$\begin{aligned} \sum_{k=1}^N f_k^* &= \sum_{k=1}^N \sum_{n=0}^k [\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)] = \sum_{n=0}^N \sum_{k=n}^N [\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)] \\ &\stackrel{(1)}{=} \sum_{n=0}^N \sum_{k=n}^{\infty} (\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)) - \sum_{n=0}^N \sum_{k=N+1}^{\infty} (\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)) \\ &\stackrel{(2)}{=} \sum_{n=0}^N \sum_{k=n}^{\infty} (\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)) - \sum_{k=N+1}^{\infty} \sum_{n=0}^N (\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)) \\ &\stackrel{(3)}{=} \sum_{n=0}^N \sum_{k=n}^{\infty} (\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)) - \sum_{k=N+1}^{\infty} \mathbb{E}(f_k^*|\mathcal{F}_{N+1}). \end{aligned}$$

To justify the numbered inequalities almost surely, we need to establish the convergence of the series which they involve.

By (1.2.4),  $\|\mathbb{E}(f_k^*|\mathcal{F}_{n+1})\|_2 + \|\mathbb{E}(f_k^*|\mathcal{F}_n)\|_2 \leq 2C_{mix}\sqrt{\text{Var}(f_k)}\theta^{k-n+1}$ , so by the Cauchy-Schwarz inequality and the assumption  $\sum \text{Var}(f_n) < \infty$ ,

$$\sum_{n=0}^N \sum_{k=n}^{\infty} \|\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)\|_2 < \infty.$$

This justifies  $\stackrel{(1)}{=}$  and  $\stackrel{(2)}{=}$ .

Next by assumption,  $|f| \leq K$  a.s. for some constant  $K$ . By (1.2.3),  $\|\mathbb{E}(f_k^*|\mathcal{F}_0)\|_{\infty} + \|\mathbb{E}(f_k^*|\mathcal{F}_n)\|_{\infty} \leq 4KC_{mix}\theta^{n-k}$  so  $\sum_{k=N+1}^{\infty} |\mathbb{E}(f_k^*|\mathcal{F}_{N+1})| < \infty$ . This justifies  $\stackrel{(3)}{=}$ .

In summary,  $\sum_{k=1}^N f_k^* = \sum_{n=0}^N \Delta_n - Z_N$ , where

$$\Delta_n := \sum_{k=n}^{\infty} (\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)) , Z_N := \sum_{k=N+1}^{\infty} \mathbb{E}(f_k^*|\mathcal{F}_{N+1}).$$

To finish the proof, we show that  $\sum_{n=0}^{\infty} \Delta_n$  and  $\lim_{N \rightarrow \infty} Z_N$  exist a.s.

CLAIM 1.  $M_N := \sum_{n=0}^{N-1} \Delta_n$  is a martingale relative to  $\{\mathcal{F}_N\}$ , and  $\sup \|M_N\|_2 < \infty$ . Consequently,  $\lim M_N$  exists almost surely.

*Proof.*  $\mathbb{E}(M_{N+1} - M_N|\mathcal{F}_N) = \mathbb{E}(\Delta_N|\mathcal{F}_N) \stackrel{!}{=} \sum_{k=N}^{\infty} \mathbb{E}(\mathbb{E}(f_k^*|\mathcal{F}_{N+1})|\mathcal{F}_N) - \mathbb{E}(\mathbb{E}(f_k^*|\mathcal{F}_N)|\mathcal{F}_N) = 0$ .

To justify  $\stackrel{!}{=}$  we note that the series

$$\Delta_N = \sum_{k=N}^{\infty} [\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)]$$

converges in  $L^2$ , because  $\|\mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n)\|_{\infty} = O(\theta^{k-n})$ , so its conditional expectation can be calculated term-by-term.

Next we show that  $\|M_N\|_2$  is uniformly bounded:

$$\begin{aligned} \|M_{N+1}\|_2 &\leq \left\| \sum_{n=0}^N \sum_{k=n}^{\infty} \mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n) \right\|_2 \\ &\leq \left\| \sum_{k=0}^{\infty} \sum_{n=0}^{k \wedge N} \mathbb{E}(f_k^*|\mathcal{F}_{n+1}) - \mathbb{E}(f_k^*|\mathcal{F}_n) \right\|_2 = \left\| \sum_{k=0}^{\infty} \mathbb{E}(f_k^*|\mathcal{F}_{(k \wedge N)+1}) \right\|_2 \\ &\leq \left\| \sum_{k=0}^N f_k^* \right\|_2 + \left\| \sum_{k=N+1}^{\infty} \mathbb{E}(f_k^*|\mathcal{F}_{N+1}) \right\|_2 \\ &\leq \sqrt{\sum_{k=0}^N \|f_k^*\|_2^2 + 2 \sum_{0 \leq k < \ell \leq N} \text{Cov}(f_k^*, f_{\ell}^*) + \sum_{k=N+1}^{\infty} \|\mathbb{E}(f_k^*|\mathcal{F}_{N+1})\|_{\infty}^2} \end{aligned}$$

$$\leq \sqrt{\sum_{k=0}^{\infty} \|f_k^*\|_2^2 + 2C_{mix} \sum_{0 \leq k < \ell \leq \infty} \theta^{\ell-k} \|f_k^*\|_2 \|f_{\ell}^*\|_2 + C_{mix} \sum_{k=N+1}^{\infty} \|f_k^*\|_{\infty} \theta^{k-N}}.$$

The last expression is uniformly bounded, because  $\sum \text{Var}(f_k) < \infty$  and

$$\begin{aligned} \sum_{0 \leq k < \ell < \infty} \theta^{\ell-k} \|f_k^*\|_2 \|f_{\ell}^*\|_2 &\leq \sum_{r=1}^{\infty} \theta^r \sum_{k=0}^{\infty} \|f_k^*\|_2 \|f_{k+r}^*\|_2 \leq \frac{1}{1-\theta} \sum_{k=0}^{\infty} \|f_k^*\|_2^2 \\ \sum_{k=N+1}^{\infty} \|f_k^*\|_{\infty} \theta^{k-N} &= \frac{1}{1-\theta} \sup_k \|f_k^*\|_{\infty}. \end{aligned}$$

CLAIM 2.  $Z_N \xrightarrow[N \rightarrow \infty]{} 0$  almost surely.

*Proof.* It is enough to prove that  $\sum \|Z_N\|_2^2 < \infty$ , because this implies using Chebyshev's inequality that  $\sum \mathbb{P}[|Z_N| > \varepsilon] \leq \frac{1}{\varepsilon^2} \sum \|Z_N\|_2^2 < \infty$  for all  $\varepsilon > 0$ , whence, by the Borel-Cantelli Lemma,  $\limsup |Z_N| \leq \varepsilon$  a.s. for all  $\varepsilon$ . Equivalently,  $\lim Z_N = 0$  a.s.

Here is the proof that  $\sum \|Z_N\|_2^2 < \infty$ :

$$\begin{aligned} \frac{1}{2} \sum_{N=1}^{\infty} \|Z_N\|_2^2 &= \sum_{N=1}^{\infty} \sum_{k_2 \geq k_1 > N} \mathbb{E} \left[ \mathbb{E}(f_{k_1}^* | \mathcal{F}_{N+1}) \mathbb{E}(f_{k_2}^* | \mathcal{F}_{N+1}) \right] \\ &= \sum_{N=1}^{\infty} \sum_{k_2 \geq k_1 > N} \mathbb{E} \left[ f_{k_2}^* \mathbb{E}(f_{k_1}^* | \mathcal{F}_{N+1}) \right] \\ &\leq C_{mix} \sum_{N=1}^{\infty} \sum_{k_2 \geq k_1 > N} \theta^{k_2-N+1} \|f_{k_2}^*\|_2 \|\mathbb{E}(f_{k_1}^* | \mathcal{F}_{N+1})\|_2 \quad \text{by (1.2.5)} \\ &\leq C_{mix}^2 \sum_{N=1}^{\infty} \sum_{k_2 \geq k_1 > N} \theta^{k_2-N+1} \|f_{k_2}^*\|_2 \cdot \theta^{k_1-N+1} \|f_{k_1}^*\|_2 \quad \text{by (1.2.4)} \\ &= C_{mix}^2 \sum_{j \geq 0} \theta^j \sum_{k > 0} \theta^{2k} \sum_{N=1}^{\infty} \|f_{k+N+j}^*\|_2 \|f_{k+N}^*\|_2 \\ &\quad (\text{after changing indices } j = k_2 - k_1, k = k_1 - N + 1) \\ &\leq C_{mix}^2 \sum_{j \geq 0} \theta^j \sum_{k > 0} \theta^{2k} \sqrt{\sum_{N=1}^{\infty} \|f_{k+N+j}^*\|_2^2 \sum_{N=1}^{\infty} \|f_{k+N}^*\|_2^2} \\ &\leq \frac{C_{mix}^2}{1-\theta} \sum_{k > 0} \theta^{2k} \sum_{N=k}^{\infty} \|f_N^*\|_2^2 = \frac{C_{mix}^2}{1-\theta} \sum_{N=1}^{\infty} \|f_N^*\|_2^2 \sum_{k=1}^N \theta^{2k} < \infty, \end{aligned}$$

because  $0 < \theta < 1$  and  $\sum \|f_k^*\|_2^2 < \infty$ . □

### 2.2.7 Convergence of moments.

Dobrushin's CLT (Theorem 2.3) shows that if  $V_N \rightarrow \infty$  then for any bounded continuous function  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  we have

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \phi \left( \frac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \right) \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(z) e^{-z^2/2} dz. \quad (2.2.12)$$

In applications, one often need to have convergence of expectations for unbounded functions, such as polynomials. This problem is addressed in the present section.

**Lemma 2.16** *Let  $f$  be a centered bounded additive functional of a uniformly elliptic Markov chain such that  $V_N \rightarrow \infty$ . Then for each  $r \in \mathbb{N}$  there is a constant  $C_r$  such that for all  $N$ ,*

$$|\mathbb{E}[S_N^r]| \leq C_r V_N^{\lfloor r/2 \rfloor}.$$

**Corollary 2.17** *Under the assumptions of Lemma 2.16*

$$\lim_{N \rightarrow \infty} \frac{\mathbb{E}[S_N^r]}{V_N^{r/2}} = \begin{cases} 0 & r \text{ is odd,} \\ (r-1)!! = \prod_{k=0}^{(r/2)-1} (r-2k-1) & r \text{ is even.} \end{cases}$$

The corollary follows from Dobrushin's CLT (Theorem 2.3), using the fact that by Lemma 2.16 and the de la Vallée-Poussin Lemma,  $(S_N/\sqrt{V_N})^r$  is uniformly integrable for all  $r > 1$  even, and therefore  $\lim \mathbb{E}[(S_N/\sqrt{V_N})^r] = \mathbb{E}[N^r]$ , where  $N$  is a Gaussian random variable with mean zero and variance one.

The proof of Lemma 2.16 proceeds by expanding  $S_N^r$  into a sum of  $r$ -tuples  $f_{n_1} \cdots f_{n_r}$  ( $n_1 \leq \cdots \leq n_r$ ), and by estimating the expectation of each tuple. (Here and throughout,  $f_n = f_n(X_n, X_{n+1})$ .) In view of the gradient lemma it is sufficient to prove Lemma 2.16 under the assumption that there is some constant  $C > 0$  such that  $\tilde{u}_n := \|f_n\|_{L^2}$  satisfy  $\sum_n \tilde{u}_n^2 \leq CV_N$ .

Consider an  $r$  tuple  $f_{n_1} \cdots f_{n_r}$  where  $n_1 \leq n_2 \leq \cdots \leq n_r$ . Segments of the form  $[n_j, n_{j+1}]$  will be called *edges*. The vertices belonging to an edge are called *bound*, the other vertices are called *free*.

A *marking* is a non-empty collection of edges satisfying the following two conditions. Firstly, each vertex  $n_j$  belongs to at most one edge. Secondly, for every free vertex  $n_l$ , either

- (i) there exists a minimal  $f(l) > l$  such that  $n_{f(l)}$  is bound, and for all  $l \leq i < f(l)$ ,  $n_{i+1} - n_i \leq n_{f(l)+1} - n_{f(l)}$ ; or
- (ii) there exists a maximal  $p(l) < l$  such that  $n_{p(l)}$  is bound, and for all  $p(l) < i \leq l$ ,  $n_i - n_{i-1} \leq n_{p(l)} - n_{p(l)-1}$ .

If (i) holds we will say that  $n_l$  is associated to the edge  $[n_{f(l)}, n_{f(l)+1}]$  otherwise it is associated to  $[n_{p(l)-1}, n_{p(l)}]$ .

**Lemma 2.18** *There are constants  $L = L(r) > 0$  and  $0 < \theta < 1$  such that*

$$\left| \mathbb{E} \left[ \prod_{i=1}^r f_{n_i} \right] \right| \leq L \sum_{\text{markings}} \prod_{[n_j, n_{j+1}] \text{ is an edge}} \left( \theta^{(n_{j+1} - n_j)} \tilde{u}_{n_j} \tilde{u}_{n_{j+1}} \right).$$

*Proof.* If  $r = 1$  then the result holds since  $\mathbb{E}[f_n] = 0$  (in this case there are no markings, and we let the empty sum be equal to zero).

If  $r = 2$  then the lemma says that  $|\mathbb{E}[f_{n_1} f_{n_2}]| \leq K \theta^{n_2 - n_1} \|f_{n_1}\|_{L^2} \|f_{n_2}\|_{L^2}$  which is true due to Proposition 1.11(2).

For  $r \geq 3$  we use induction. Take  $j$  such that  $n_{j+1} - n_j$  is the largest. Then

$$\mathbb{E} \left[ \prod_{i=1}^r f_{n_i} \right] = \mathbb{E} \left[ \prod_{i=1}^j f_{n_i} \right] \mathbb{E} \left[ \prod_{i=j+1}^r f_{n_i} \right] + O \left( \theta^{(n_{j+1} - n_j)} \left\| \prod_{i=1}^j f_{n_i} \right\|_{L^2} \left\| \prod_{i=j+1}^r f_{n_i} \right\|_{L^2} \right).$$

Let  $K := \text{ess sup} |\mathbf{f}|$ , then the second term is smaller than  $\theta^{(n_{j+1}-n_j)} \tilde{u}_{n_j} \tilde{u}_{n_{j+1}} K^{r-2}$ . Thus this term is controlled by the marking with only one marked edge  $[n_j, n_{j+1}]$ . Applying the inductive assumption to each factor in the first term we obtain the result.  $\square$

**Lemma 2.19** *There exists  $\bar{C}_r > 0$  s.t. for every set  $\mathcal{C}$  of  $r$  tuples  $1 \leq n_1 \leq \dots \leq n_r \leq N$ ,*

$$\Gamma_{\mathcal{C}} := \sum_{(n_1, \dots, n_r) \in \mathcal{C}} \left| \mathbb{E} \left[ \prod_{i=1}^r f_{n_i} \right] \right| \leq \bar{C}_r V_N^{\lfloor r/2 \rfloor}.$$

Lemma 2.19 implies Lemma 2.16 since

$$\mathbb{E}[S_N^r] = \sum_{s=1}^r \sum_{k_1 + \dots + k_s = r} \frac{r!}{k_1! \dots k_s!} \sum_{1 \leq n_1 < \dots < n_s \leq N} \mathbb{E} \left[ \prod_{j=1}^s f_{n_j}^{k_j} \right].$$

Therefore it suffices to prove Lemma 2.19.

*Proof.* By Lemma 2.18

$$\Gamma_{\mathcal{C}} \leq L \sum_{(n_1, \dots, n_r) \in \mathcal{C}} \sum_{\substack{\text{markings } (e_1, \dots, e_s) \\ \text{of } (n_1, \dots, n_r)}} \prod_{j=1}^s \left( \tilde{u}_{e_j^-} \tilde{u}_{e_j^+} \theta^{(e_j^+ - e_j^-)} \right)$$

where the marked edges are  $e_j = [e_j^-, e_j^+]$ ,  $j = 1, \dots, s$ . Collecting all terms with a fixed set of marked edges  $(e_1, \dots, e_s)$  we obtain

$$\Gamma_{\mathcal{C}} \leq C(r) \sum_s \sum_{(e_1, \dots, e_s)} \prod_{j=1}^s \left( \tilde{u}_{e_j^-} \tilde{u}_{e_j^+} \theta^{(e_j^+ - e_j^-)} (e_j^+ - e_j^-)^{r-2} \right) \quad (2.2.13)$$

where  $C(r) \prod_j (e_j^+ - e_j^-)^r$  accounts for all tuples which admit a marking  $(e_1, \dots, e_s)$ . Indeed, for every edge  $e = [e^-, e^+]$  there are at most  $0 \leq j \leq r-2$  vertices which may be associated to  $e$  and the positions of those vertices are located inside

$$[e^- - (r-2)(e^+ - e^-), e^-] \cup (e^+, e^+ + (r-2)(e^+ - e^-)].$$

It follows that there are at most  $2(r-2)(e^+ - e^-)$  choices to place each vertex associated to a given edge. This gives

$$\prod_e \left( \sum_{j=0}^{r-2} [2(r-2)(e^+ - e^-)]^j \right) \leq C(r) \prod_e (e^+ - e^-)^{r-2}$$

possibilities for tuples with marking  $(e_1, \dots, e_s)$  proving (2.2.13).

The sum over  $(e_1, \dots, e_s)$  in (2.2.13) can be estimated by

$$\left( \sum_{n=1}^{N-1} \sum_{m=1}^{N-n} \tilde{u}_n \tilde{u}_{n+m} \theta^m m^{r-2} \right)^s.$$

For each  $m$ ,  $\sum_n \tilde{u}_n \tilde{u}_{n+m} = O(V_N)$  due to the Cauchy-Schwartz inequality and because  $\sum_{n=1}^N \tilde{u}_n^2 \leq CV_N$  by assumption. Summing over  $m$  gives  $\Gamma_{\mathcal{C}} \leq \text{const} \sum_{2s \leq r} V_N^s$  where the condition  $2s \leq r$  ap-

pears because each edge involves two distinct vertices, and no vertex belongs to more than one edge. The result follows.  $\square$

### 2.3 Notes and references

The connection between the non-growth of variance and representation in terms of gradients is well-known for stationary stochastic processes. The first result in this direction we are aware of is Leonov's Theorem [94]. He showed that the asymptotic variance of a homogeneous additive functional of a stationary homogeneous Markov chain is zero iff the additive functional is the sum of a gradient and a constant. Rousseau-Egele [127] and Guivarc'h & Hardy [65] extended this to the context of dynamical systems preserving an invariant Gibbs measure. Kifer [80], Conze & Raugi [28], Dragičević, Froyland & González-Tokman [53] have proved versions of Leonov's theorem for random and/or sequential dynamical systems.

The connection between center-tightness and gradients is a central feature of the theory of cocycles over ergodic transformations. Suppose  $T : X \rightarrow X$  is an ergodic probability preserving transformation on a non-atomic probability space. For every measurable  $f : X \rightarrow \mathbb{R}$ ,  $\{f \circ T^n\}$  is a stationary stochastic process, and

$$S_N = f + f \circ T + \cdots + f \circ T^{N-1}$$

are called the “ergodic sums of the cocycle  $f$ .” A “coboundary” is a function of the form  $f = g - g \circ T$  with  $g$  measurable. Schmidt characterized cocycles with center-tight  $S_N$  as those arising from coboundaries [136, page 181]. These results extend to cocycles taking values in locally compact groups, see Moore & Schmidt [105] and Aaronson & Weiss [7]. For more on this, see Aaronson [2, chapter 8], and Bradley [16, chapters 8,19]. We also refer to [64] for an analogous result in the continuous setting.

Notice that inhomogeneous theory is different from the stationary theory in that there is another cause for center-tightness: Having summable variance. This cannot happen in the stationary homogeneous world (unless all  $f_i$  are constant).

Theorem 2.3 is a special case of a more general result due to Dobrushin, which can be found in [45]. The conditions for Dobrushin's full result are more general than uniform boundedness or uniform ellipticity. Our proof follows the paper of Sethuraman & Varadhan [139], except for some changes we needed to make to deal with additive functionals of the form  $f_k(X_k, X_{k+1})$ , and not just  $f_k(X_k)$  as in [139]. McLeish's Lemma, the martingale CLT, and their proofs are due to McLeish [103]. We refer the reader to Hall & Heyde [71] for the history of this result, further extensions, and references.

Theorem 2.4 extends the Kolmogorov-Khintchin “Two-Series Theorem” [84]. There are other extensions to sums of dependent random variables. We mention for example a version for martingales (Hall & Heyde [71, chapter 2]), for sums of negatively dependent random variables (Matuła, [101]) and for expanding maps ([28]).

The proofs of theorems 2.3 and 2.4 use Gordin's “martingale-coboundary decomposition” [62], see also [71],[86].

## Chapter 3

# The essential range and irreducibility

In this chapter we discuss the following question: How small can we make the range of an additive functional, by subtracting from it a center-tight functional?

### 3.1 Definitions and motivation

Let  $f = \{f_n\}$  be an additive functional of a Markov chain  $X := \{X_n\}$ . The **algebraic range** of  $(X, f)$  is the intersection  $G_{alg}(X, f)$  of all closed groups  $G$  s.t.,

$$\exists c_n \in \mathbb{R} \text{ s.t. } \mathbb{P}[f_n(X_n, X_{n+1}) - c_n \in G] = 1 \text{ for all } n \geq 1. \quad (3.1.1)$$

We will see later (Lemma 3.9) that  $G_{alg}(X, f)$  itself satisfies (3.1.1), therefore  $G_{alg}(X, f)$  is the smallest closed group satisfying (3.1.1).

**Example 3.1 (The simple random walk).** Suppose  $\{X_n\}$  are independent random variables such that  $\mathbb{P}(X_n = \pm 1) = \frac{1}{2}$ , and let  $f_n(x, y) = x$ . Then  $S_n = X_1 + \dots + X_n$  is the simple random walk on  $\mathbb{Z}$ . The algebraic range in this case is  $2\mathbb{Z}$ .

*Proof:*  $G_{alg} \subset 2\mathbb{Z}$ , because we can take  $c_n := -1$ . Assume by contradiction that  $G_{alg} \subsetneq 2\mathbb{Z}$ , then  $G_{alg} = t\mathbb{Z}$  for  $t \geq 4$ , and the supports of  $S_n$  are cosets of  $t\mathbb{Z}$ .

But this is false, because  $\exists a_1, a_2$  s.t.  $|a_1 - a_2| < t$  and  $\mathbb{P}(S_n = a_i) \neq 0$ : For  $n$  even take  $a_i = (-1)^i$ , and for  $n$  odd take  $a_i = 1 + (-1)^i$ .  $\square$

The **lattice case** is the case when  $G_{alg}(X, f) = t\mathbb{Z}$  for some  $t \geq 0$ . The **non-lattice case** is the case when  $G_{alg}(X, f) = \mathbb{R}$ . The distinction is important for the following reason. If  $G_{alg}(X, f) = t\mathbb{Z}$  and  $\gamma_N := c_1 + \dots + c_N$ , then

$$\mathbb{P}(S_N \in \gamma_N + t\mathbb{Z}) = 1 \text{ for all } N.$$

In this case it is not true that  $\mathbb{P}(S_N - z_N \in (a, b)) \sim \frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}$  whenever  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , because  $\mathbb{P}(S_N - z_N \in (a, b)) = 0$  whenever  $|a - b| < t$  and  $z_N + (a, b)$  falls inside the gaps of  $\gamma_N + t\mathbb{Z}$ . This is the **lattice obstruction** to the local limit theorem.

There is a related, but more subtle, obstruction. An additive functional  $f$  is called **reducible** on  $X$ , if there is another additive functional  $g$  on  $X$  such that  $f - g$  is center-tight, and

$$G_{alg}(X, g) \subsetneq G_{alg}(X, f).$$

In this case we say that  $g$  is a **reduction** of  $f$ , and call the algebraic range of  $g$  a **reduced range** of  $f$ .

**Example 3.2 (Simple random walk with continuous first step):** Suppose  $\{X_n\}_{n \geq 1}$  are independent real valued random variables such that  $X_1$  has continuous non-uniform distribution  $\mathfrak{F}$  with compact support, and  $X_2, X_3, \dots$  are equal to  $\pm 1$  with equal probabilities. Let  $f_n(x, y) = x$ , then  $S_n = X_1 + X_2 + \dots + X_n$ .

Because of the continuously distributed first step,  $G_{alg}(f) = \mathbb{R}$ . But if we subtract from  $f$  the center-tight functional  $c$  with components

$$c_n(x, y) = x \text{ when } n = 1 \text{ and } c_n(x, y) \equiv 0 \text{ when } n > 1,$$

then the result  $g := f - c$  has algebraic range  $2\mathbb{Z}$ . So  $f$  is reducible.

The reduction  $g$  satisfies the lattice local limit theorem (see the preface), because it generates the (delayed) simple random walk. But by the assumptions on  $\mathfrak{F}$ , the original functional  $f = g + c$  does not satisfy the LLT, lattice or non-lattice. This can be seen by direct calculation from the observation that the distribution of  $S_n$  is the convolution of  $\mathfrak{F}$  and the centered binomial distribution. See chapter 5 for details.

Here we see an instance of the **reducibility obstruction** to the local limit theorem: A situation when the LLT fails because the additive functional is a sum of a lattice term which satisfies the lattice LLT and a non-lattice center-tight term which spoils it. The reducibility obstruction to the LLT raises the following questions:

1. Given an additive functional  $f$ , how small can we make its algebraic range by subtracting from it a center-tight term?
2. Is there an “optimal” center-tight functional  $c$  such that the algebraic range of  $f - c$  cannot be reduced further?

Motivated by these questions, we introduce the following definitions. The **essential range** of  $f$  is

$$G_{ess}(X, f) := \bigcap \{G_{alg}(X, g) : f - g \text{ is center tight}\}.$$

This is a closed sub-group of  $G_{alg}(X, f)$ .

An additive functional without reductions is called **irreducible**. Equivalently,  $f$  is irreducible iff  $G_{ess}(X, f) = G_{alg}(X, f)$ .

In this terminology questions 1 and 2 call for the calculation of  $G_{ess}(X, f)$  and ask for an irreducible reduction of  $f$ .

## 3.2 Main results

### 3.2.1 Results for Markov chains

The questions raised at the end of the last section can be answered using the structure constants  $d_n(\xi)$  introduced in (1.3.2). Define the **co-range** of  $f$  to be the set

$$H(X, f) := \{\xi \in \mathbb{R} : \sum_{n=3}^{\infty} d_n(\xi)^2 < \infty\}.$$

**Theorem 3.1.** Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . If  $f$  is center-tight then  $H(X, f) = \mathbb{R}$ , and if not then either  $H(X, f) = \{0\}$ , or  $H(X, f) = t\mathbb{Z}$  for some  $t \geq \pi/(6\text{ess sup } |f|)$ .

**Theorem 3.2.** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ , then*

- (a) *If  $H(X, f) = 0$ , then  $G_{ess}(X, f) = \mathbb{R}$ .*
- (b) *If  $H(X, f) = t\mathbb{Z}$  with  $t \neq 0$ , then  $G_{ess}(X, f) = \frac{2\pi}{t}\mathbb{Z}$ .*
- (c) *If  $H(X, f) = \mathbb{R}$ , then  $G_{ess}(X, f) = \{0\}$ .*

**Theorem 3.3.** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . Then there exists an irreducible uniformly bounded additive functional  $g$  such that  $f - g$  is center-tight, and*

$$G_{alg}(X, g) = G_{ess}(X, g) = G_{ess}(X, f).$$

**Corollary 3.3** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . If  $G_{ess}(X, f) = t\mathbb{Z}$  with  $t \neq 0$ , then  $|t| \leq 12\text{ess sup}|f|$ .*

The corollary follows directly from Theorems 3.1 and 3.2(b).

### 3.2.2 Results for Markov arrays

The previous discussion applies to Markov arrays. Let  $f$  be an additive functional on a Markov array  $X$  with row lengths  $k_N + 1$ :

- (1) The **algebraic range**  $G_{alg}(X, f)$  is the intersection of all closed subgroups  $G$  of  $\mathbb{R}$  such that for all  $1 \leq k \leq k_N, N \geq 1$

$$\exists c_k^{(N)} \in \mathbb{R} \text{ s.t. } \mathbb{P}[f_k^{(N)}(X_k^{(N)}, X_{k+1}^{(N)}) - c_k^{(N)} \in G] = 1.$$

- (2) The **essential range**  $G_{ess}(X, f)$  is the intersection of the algebraic ranges of all additive functionals of the form  $f - h$  where  $h$  is center-tight.

- (3) The **co-range** is  $H(X, f) := \{\xi \in \mathbb{R} : \sup_N \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty\}$ .

- (4) An additive functional  $f$  is called **irreducible** if  $G_{ess}(X, f) = G_{alg}(X, f)$ .

This is consistent with the definitions for Markov chains, see Corollary 3.4 below.

**Theorem 3.4.** *The results of Theorems 3.1, 3.2 3.3 and of Corollary 3.3 hold for all a.s. uniformly bounded additive functionals on uniformly elliptic Markov arrays.*

**Corollary 3.4** *Suppose  $f = \{f_n\}$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X = \{X_n\}$ . Let  $\tilde{f} = \{f_n^{(N)}\}$  be an additive functional on a Markov array  $\tilde{X} = \{\tilde{X}_n^{(N)}\}$  s.t.  $f_n^{(N)} = f_n$  and  $X_n^{(N)} = X_n$ . Then*

$$G_{alg}(\tilde{X}, \tilde{f}) = G_{alg}(X, f), \quad G_{ess}(\tilde{X}, \tilde{f}) = G_{ess}(X, f), \quad H(\tilde{X}, \tilde{f}) = H(X, f).$$

*Proof.* The equality of the algebraic ranges and co-ranges is trivial, but the equality of the essential ranges requires justification, because some center-tight functionals of  $\{X_n^{(N)}\}$  are not of the form  $h_n^{(N)} = h_n$ .

However, since the co-ranges agree, the essential ranges must also agree, by the version of Theorem 3.2 for arrays.  $\square$

### 3.2.3 Hereditary arrays

Some results for Markov chains do not extend to general Markov arrays. Of particular importance is the following fact, which we need for the proof of the LLT (see the proof of Theorem 4.1, claim 2). Recall the definition of  $D_N(\xi)$  from (1.3.2).

**Theorem 3.5.** *Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ , then*

$$D_N(\xi) \xrightarrow[N \rightarrow \infty]{} \infty \text{ uniformly on compact subsets of } \mathbb{R} \setminus H(X, f). \quad (3.2.1)$$

*Proof.* Suppose  $\xi \in \mathbb{R} \setminus H(X, f)$ , then  $\sup_N D_N(\xi) = \infty$ , whence

$$D_N(\xi) = \sum_{k=3}^N d_k^{(N)}(\xi)^2 \xrightarrow[N \rightarrow \infty]{} \sum_{k=3}^{\infty} d_k^{(N)}(\xi)^2 \equiv \sup_N D_N(\xi) = \infty.$$

Since  $D_N(\xi)$  is non-decreasing and  $\xi \mapsto D_N(\xi)$  are continuous, the convergence is uniform on compact subsets of  $\mathbb{R} \setminus H(X, f)$ .  $\square$

The following two examples show that Theorem 3.5 fails for some arrays:

**Example 3.5** *Let  $X_n$  be a sequence of independent uniform random variables with zero mean and variance equal to one. Form an array by setting*

$$X_k^{(N)} = \begin{cases} X_k & 1 \leq k \leq N+1, N \text{ odd} \\ 0 & 1 \leq k \leq N+1, N \text{ even} \end{cases} \quad (k = 1, \dots, N)$$

and let  $f_k^{(N)}(x, y) := x$ . Then for every  $0 \neq \xi \in \mathbb{R} \setminus H(X, f)$ ,  $D_N(\xi) \not\rightarrow \infty$ .

*Proof.* We claim that  $\sup_N D_{2N+1}(\xi) = \infty$  for every  $\xi \neq 0$ .

To see this, suppose  $P = \begin{pmatrix} X_{n-2} & X_{n-1} & X_n \\ Y_{n-2} & Y_{n-1} & Y_n \end{pmatrix}$  is a random level  $2N+1$  hexagon at position  $n$ , then  $\Gamma(P) = X_{n-1} + X_n - Y_{n-1} - Y_n$  where  $X_i, Y_j$  are independent random variables each having uniform distribution with mean zero and unit variance. So  $\Gamma(P)$  is a *non lattice random variable* and for every  $\xi \neq 0$ ,  $d_n^{(2N+1)}(\xi)^2 = \mathbb{E}(|e^{i\xi\Gamma(P)} - 1|^2) = c(\xi)$ , where  $c(\xi)$  is a positive constant independent of  $n$ . So

$$D_{2N+1}(\xi) = (2N-1)c(\xi) \xrightarrow[N \rightarrow \infty]{} \infty.$$

Thus  $H(X, f) = \{0\}$ . But  $D_N(\xi) \not\rightarrow \infty$  for  $\xi \neq 0$ , because  $D_{2N}(\xi) = 0$ .  $\square$

**Example 3.6** *Suppose  $X_n$  are a sequence of independent identically distributed random variables, equal to  $\pm 1$  with probability  $\frac{1}{2}$ . Form an array with row lengths  $N+1$  by setting  $X_n^{(N)} = X_n$ , and let*

$$f_n^{(N)}(X_n, X_{n+1}) := \frac{1}{2} \left( 1 + \frac{1}{\sqrt[3]{N}} \right) X_n \quad (1 \leq n \leq N+1).$$

*Then  $D_N(\xi) \rightarrow \infty$  for all  $\xi \notin H(X, f)$ , but the convergence is not uniform on compact subsets of  $\mathbb{R} \setminus H(f)$ .*

*Proof.*  $\Gamma \left( +1 \begin{smallmatrix} +1 & +1 \\ +1 & -1 \end{smallmatrix} +1 \right) = 1 + N^{-1/3}$ . Since  $\text{Hex}(N, n)$  consists of  $2^6$  hexagons, the hexagon  $\left( +1 \begin{smallmatrix} +1 & +1 \\ +1 & -1 \end{smallmatrix} +1 \right)$  has probability  $2^{-6}$ . It follows that

$$d_n^{(N)}(\xi) \geq 2^{-6} |e^{i\xi(1+N^{-1/3})} - 1|^2 = \frac{1}{16} \sin^2 \frac{\xi(1+N^{-1/3})}{2}$$

$$D_N(\xi) \geq \frac{N-2}{16} \sin^2 \frac{\xi(1+N^{-1/3})}{2} \sim \begin{cases} 16^{-1} N \sin^2 \frac{\xi}{2} & \xi \notin 2\pi\mathbb{Z} \\ 16^{-1} \sqrt[3]{N} & \xi \in 2\pi\mathbb{Z}. \end{cases}$$

We see that  $D_N(\xi) \rightarrow \infty$  for all  $\xi \neq 0$ , whence  $H(\mathbf{X}, \mathbf{f}) = \{0\}$ , and  $D_N(\xi) \rightarrow \infty$  for all  $\xi \notin H(\mathbf{X}, \mathbf{f})$ . But the convergence is not uniform on any compact neighborhood of  $2\pi k$ ,  $k \neq 0$ , because  $D_N(\xi_N) \equiv 0$  for  $\xi_N = 2\pi k(1+N^{-1/3})^{-1} \rightarrow 2\pi k$ .  $\square$

Because of the importance of property (3.2.1) to the proof of the LLT, we would like to characterize the additive functionals on Markov arrays which satisfy it. Examples 1 and 2 point the way.

Let  $\mathbf{X}$  be a Markov array with row lengths  $k_N$ . A **sub-array** of  $\mathbf{X}$  is an array  $\mathbf{X}'$  of the form  $\{X_k^{(N_\ell)} : 1 \leq k \leq k_{N_\ell} + 1, \ell \geq 1\}$  where  $N_\ell \uparrow \infty$ . The **restriction** of  $\mathbf{f}$  to  $\mathbf{X}'$  is

$$\mathbf{f}|_{\mathbf{X}'} = \{f_k^{(N_\ell)} : 1 \leq k \leq k_{N_\ell}, \ell \geq 1\}.$$

$(\mathbf{X}, \mathbf{f})$  is called **hereditary**, if  $G_{ess}(\mathbf{X}', \mathbf{f}|_{\mathbf{X}'}) = G_{ess}(\mathbf{X}, \mathbf{f})$  for all sub-arrays  $\mathbf{X}'$ , and **stably hereditary** if  $(\mathbf{X}, \mathbf{g})$  is hereditary whenever  $\mathbf{g} = \{(1 + \varepsilon_N) f_k^{(N)}\}$  with  $\varepsilon_N \rightarrow 0$ .

**Theorem 3.6.** *Let  $\mathbf{f}$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $\mathbf{X}$ , then the following conditions are equivalent:*

(1)  $\mathbf{f}$  is hereditary;

(2) for all  $\xi$ ,  $\liminf_{N \rightarrow \infty} \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty \Rightarrow \limsup_{N \rightarrow \infty} \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty$ ;

(3) for all  $\xi \notin H(\mathbf{X}, \mathbf{f})$ ,  $D_N(\xi) \xrightarrow[N \rightarrow \infty]{} \infty$ ;

(4)  $H(\mathbf{X}', \mathbf{f}|_{\mathbf{X}'}) = H(\mathbf{X}, \mathbf{f})$  for every sub-array  $\mathbf{X}'$  of  $\mathbf{X}$ .

In addition,  $\mathbf{f}$  is stably hereditary iff the convergence in (3) is uniform on compact subsets of  $\mathbb{R} \setminus H(\mathbf{X}, \mathbf{f})$ .

**Example 3.7 (Markov chains):** Suppose  $\mathbf{f}$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $\mathbf{X}$ . If  $f_n^{(N)} = f_n$  and  $X_n^{(N)} = X_n$ , then  $\mathbf{f}$  is stably hereditary.

*Proof.* This follows from Theorems 3.5 and 3.6.  $\square$

**Example 3.8 (“Change of measure”):** Let  $\mathbf{Y}$  be an array obtained from a Markov chain  $\mathbf{X}$  using the change of measure construction (example 1.6). Let  $\varphi_n^{(N)}$  denote the weights of the change of measure. If  $\exists C > 0$  s.t.

$$C^{-1} < \varphi_n^{(N)} < C \text{ for all } n, N,$$

then for every a.s. uniformly bounded additive functional  $\mathbf{f}$  on  $\mathbf{X}$ , the additive functional  $f_n^{(N)} := f_n$  is stably hereditary on  $\mathbf{Y}$ .

*Proof.* If  $d_n(\xi, X)$  are the structure constants of  $f$  on  $X$ , and  $d_n^{(N)}(\xi, Y)$  are the structure constants of  $f$  on  $Y$ , then  $C^{-6}d_n(\xi, X) \leq d_n^{(N)}(\xi, Y) \leq C^6d_n(\xi, X)$ . So  $H(Y, f) = H(X, f)$ .

Theorem 3.5 says that  $D_N(\xi, X) \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(X, f)$ . Since  $D_N(\xi, Y) \geq C^{-6}D_N(\xi, X)$ ,  $D_N(\xi, X) \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(Y, f)$ .  $\square$

Sometimes (but not always, Example 3.6), every hereditary functional is stably hereditary:

**Theorem 3.7.** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$ .*

- (a) *Suppose  $G_{ess}(X, f) = t\mathbb{Z}$  or  $\{0\}$ . If  $f$  is hereditary then  $f$  is stably hereditary.*
- (b) *Suppose  $f$  is integer valued and not center-tight, and  $|f| \leq K$ , then  $G_{ess}(X, f) = \frac{k}{2\pi}\mathbb{Z}$  for some  $0 < k \leq 12K$ , and if  $f$  is hereditary then  $f$  is stably hereditary.*

### 3.3 Proofs

#### 3.3.1 Reduction lemmas

**Lemma 3.9** *Let  $f$  be an additive functional on a Markov array  $X$  with row lengths  $k_N + 1$ . For every  $N \geq 1$  and  $1 \leq n \leq k_N$ , there exists  $c_k^{(N)}$  s.t.  $f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) - c_n^{(N)} \in G_{alg}(X, f)$  almost surely.*

*Proof.*  $G_{alg}(X, f)$  is the intersection of all closed subgroups  $G$  such that

$$\exists c_k^{(N)} \text{ s.t. } f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) - c_n^{(N)} \in G \text{ almost surely.} \quad (3.3.1)$$

This is a closed subgroup of  $\mathbb{R}$ . The lemma is trivial when  $G_{alg}(X, f) = \mathbb{R}$  (take  $c_n^{(N)} \equiv 0$ ), so we focus on the case  $G_{alg}(X, f) \neq \mathbb{R}$ .

In this case (3.3.1) holds with some  $G = t\mathbb{Z}$  with  $t \geq 0$ , and  $f_n^{(N)}(X_n, X_{n+1})$  must be a discrete random variable. Let  $A_n^{(N)}$  denote the set of values attained by  $f_n^{(N)}(X_n, X_{n+1})$  with positive probability. Since  $G = t\mathbb{Z}$  satisfies (3.3.1),  $A_n^{(N)} \subset \text{coset of } t\mathbb{Z}$ , and  $D_n^{(N)} := A_n^{(N)} - A_n^{(N)} \subset t\mathbb{Z}$ . Let  $G_0$  denote the group generated by  $\bigcup_{N \geq 1} \bigcup_{1 \leq n \leq k_N} D_n^{(N)}$ . Then  $G_0$  is a subgroup of  $t\mathbb{Z}$ . In particular,  $G_0$  is closed.

By the previous paragraph,  $G_0 \subset t\mathbb{Z}$  for any group  $t\mathbb{Z}$  which satisfies (3.3.1). So  $G_0 \subseteq G_{alg}(X, f)$ . Next, we fix  $n, N$  and observe that all the values of  $f_n^{(N)}(X_n, X_{n+1})$  belong to the same translate of  $A_n^{(N)} - A_n^{(N)}$ , and therefore to the same coset of  $G_0$ . So  $G_0$  satisfies (3.3.1), and  $G_0 \supset G_{alg}(X, f)$ . So  $G_{alg}(X, f) = G_0$ . Since  $G_0$  satisfies (3.3.1),  $G_{alg}(X, f)$  satisfies (3.3.1).  $\square$

**Lemma 3.10 (Reduction Lemma)** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$ . If  $\xi \neq 0$  and  $\sup_N \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty$ , then there exists a uniformly bounded additive functional  $g$  on  $X$  s.t.*

$$f - g \text{ is center-tight, and } G_{alg}(g) \subset \frac{2\pi}{\xi}\mathbb{Z}.$$

*If  $X_n^{(N)} = X_n$  and  $f_n^{(N)} = f_n$  (as in the case additive functionals of Markov chains), then we can take  $g$  such that  $g_n^{(N)} = g_n$ .*

**Proof for Doeblin chains:** As in the case of the gradient lemma, the reduction lemma has a particularly simple proof in the important special case of Doeblin Markov chains (Example 1.7). Recall that Doeblin chains have finite state spaces  $\mathfrak{S}_n$ . Let  $\pi_{xy}^n := \pi_{n,n+1}(x, \{y\})$ , and re-label the states  $\mathfrak{S}_n = \{1, \dots, d_n\}$  in such a way that  $\pi_{11}^n = \pi_{n,n+1}(1, \{1\}) \neq 0$  for all  $n$ . The Doeblin condition guarantees that for every  $x \in \mathfrak{S}_n$ , there exists a state  $\xi_n(x) \in \mathfrak{S}_{n+1}$  such that  $\pi_{1,\xi_n(x)}^{n-1} \pi_{\xi_n(x),1}^n > 0$ .

Define as in the proof of the gradient lemma,

$$\begin{aligned} a_0 &\equiv 0, \quad a_1 \equiv 0, \quad \text{and } a_n(x) := f_{n-2}(1, \xi_{n-1}(x)) + f_{n-1}(\xi_{n-1}(x), x) \text{ for } n \geq 3 \\ c_0 &:= 0, \quad c_1 := 0, \quad \text{and } c_n := f_{n-2}(1, 1) \text{ for } n \geq 3 \\ \tilde{f} &:= f - \nabla a - c. \end{aligned}$$

Then  $\tilde{f}_n(x, y) = f_n(x, y) - (a_{n+1}(y) - a_n(x)) - c_n = -\Gamma_n \begin{pmatrix} 1 & 1 & \xi_n(y) \\ & \xi_{n-1}(x) & x \\ & & y \end{pmatrix}$ , where  $\Gamma_n$  denotes the balance of a hexagon, see (1.3.1).

For Doeblin chains, there are finitely many admissible hexagons at position  $n$ , and the hexagon measure assigns each of them a mass which is uniformly bounded from below. Let  $C^{-1}$  be a uniform lower bound for this mass, then

$$|e^{i\xi \tilde{f}_n(x,y)} - 1|^2 \leq C\mathbb{E}(|e^{i\xi \Gamma_n} - 1|^2) = Cd_n^2(\xi).$$

Decompose  $\tilde{f}_n(x, y) = g_n(x, y) + h_n(x, y)$  where  $g_n(x, y) \in \frac{2\pi}{\xi}\mathbb{Z}$  and  $h_n(x, y) \in [-\frac{\pi}{\xi}, \frac{\pi}{\xi}]$ . Clearly  $|g| \leq |f| + |\nabla a| + |c| + |h| \leq 6|f| + \pi/\xi$ , and  $G_{alg}(\mathbf{X}, \mathbf{h}) \subset \frac{2\pi}{\xi}\mathbb{Z}$ .

We show that  $f - g$  is center tight. We need the following inequality:<sup>1</sup>

$$\frac{4x^2}{\pi^2} \leq |e^{ix} - 1|^2 \leq x^2 \text{ for all } |x| \leq \pi. \quad (3.3.2)$$

By (3.3.2),  $|h_n(x, y)|^2 \leq \frac{\pi^2}{4\xi^2} |e^{i\xi h_n(x,y)} - 1|^2 = \frac{\pi^2}{4\xi^2} |e^{i\xi \tilde{f}_n(x,y)} - 1|^2 \leq C \frac{\pi^2}{4\xi^2} d_n^2(\xi)$ , whence

$$\sum_{n=3}^{\infty} \text{Var}(h_n(X_n, X_{n+1}) + c_n) = \sum_{n=3}^{\infty} \text{Var}(h_n(X_n, X_{n+1})) \leq \frac{C\pi^2}{4\xi^2} \sum_{n=3}^{\infty} d_n^2(\xi) < \infty.$$

So  $h + c$  has summable variance. Therefore  $f - g = \nabla a + (h + c)$  is center tight.  $\square$

### Preparations for the proof in the general case.

**Lemma 3.11** Suppose  $E_1, \dots, E_N$  are measurable events, and let  $W$  denote the random variable which counts how many of  $E_i$  occur simultaneously, then

$$\mathbb{P}(W \geq t) \leq \frac{1}{t} \sum_{k=1}^N \mathbb{P}(E_k).$$

*Proof.* Apply Markov's inequality to  $W = \sum 1_{E_k}$ .  $\square$

Suppose  $W$  is a real-valued random variable. A **circular mean** of  $W$  is a real number  $\theta \in [-\pi, \pi]$  which minimizes the quantity  $\mathbb{E}(|e^{i(W-\theta)} - 1|^2)$ . Such numbers always exist, because

<sup>1</sup> Proof of (3.3.2): Since  $y = \sin x$  is concave on  $[0, \frac{\pi}{2}]$ , its graph lies above the chord  $y = 2x/\pi$  and below the tangent  $y = x$ . So  $2x/\pi \leq \sin x \leq x$  on  $[0, \frac{\pi}{2}]$ . Now use the identity  $|e^{ix} - 1|^2 = 2(1 - \cos x) = 4\sin^2 \frac{x}{2}$ .

$\theta \mapsto \mathbb{E}(|e^{i(W-\theta)} - 1|^2)$  is continuous and  $2\pi$ -periodic. But circular means are not unique: If, for example,  $W$  is uniformly distributed on  $[-\pi, \pi]$ , then every  $\theta \in [-\pi, \pi]$  is a circular mean.

The **circular variance** of a real random variable  $W$  is defined to be

$$\text{CVar}(W) := \min_{\theta \in [-\pi, \pi)} \mathbb{E}(|e^{i(W-\theta)} - 1|^2) \equiv \min_{\theta \in [-\pi, \pi)} 4\mathbb{E}(\sin^2 \frac{W-\theta}{2}).$$

For every  $x \in \mathbb{R}$ , let

$$\langle x \rangle := \text{unique element of } [-\pi, \pi) \text{ s.t. } x - \langle x \rangle \in 2\pi\mathbb{Z}. \quad (3.3.3)$$

It is not difficult to see, using (3.3.2), that for every circular mean  $\theta$

$$\frac{4}{\pi^2} \text{Var}(\langle W - \theta \rangle) \leq \text{CVar}(W) \leq \text{Var}(W). \quad (3.3.4)$$

**Lemma 3.12** *For every real-valued random variable  $W$ , we can write  $W = W_1 + W_2$  where  $W_1 \in 2\pi\mathbb{Z}$  almost surely, and  $\text{Var}(W_2) \leq \frac{\pi^2}{4} \text{CVar}(W)$ .*

*Proof.*  $W_1 := (W - \theta) - \langle W - \theta \rangle$ ,  $W_2 := \langle W - \theta \rangle + \theta$ ,  $\theta :=$  a circular mean.  $\square$

**Proof of the Reduction Lemma in the general case:** Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$ , with row lengths  $k_N$ , and fix  $\xi \neq 0$  such that

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 < \infty.$$

Let  $L$  denote the ladder process associated to  $X$  (see section 1.3.2). We remind the reader that this is a Markov array with entries  $\underline{L}_n^{(N)} = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)})$  ( $3 \leq n \leq k_N$ ), and for every  $N$ : (a)  $\{X_n^{(N)}\}$ ,  $\{Z_n^{(N)}\}$  are two independent copies of  $X^{(N)}$ ; (b)  $Y_n^{(N)}$  are conditionally independent given  $\{X_i^{(N)}\}$  and  $\{Z_i^{(N)}\}$ ; and (c) the conditional distribution of  $Y_n^{(N)}$  given  $\{Z_i^{(N)}\}$  and  $\{X_i^{(N)}\}$  is given by

$$\mathbb{P} \left( Y_{n-1}^{(N)} \in E \middle| \begin{array}{l} \{Z_i^{(N)}\} = \{\zeta_i^{(N)}\} \\ \{X_i^{(N)}\} = \{\xi_i^{(N)}\} \end{array} \right) = \begin{array}{l} \text{bridge probability for } X \text{ that } X_{n-1}^{(N)} \in E \\ \text{given that } X_{n-2}^{(N)} = \zeta_{n-2}^{(N)} \text{ and } X_n^{(N)} = \xi_n^{(N)}. \\ (\text{see } \S 1.2.3) \end{array}$$

Let  $F, H$  be the additive functionals on  $L$  with entries

$$\begin{aligned} F^{(N)}(\underline{L}_n) &:= f_{n-2}^{(N)}(Z_{n-2}^{(N)}, Y_{n-1}^{(N)}) + f_{n-1}^{(N)}(Y_{n-1}^{(N)}, X_n^{(N)}) \\ H_n^{(N)}(\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)}) &:= \left\langle \xi \Gamma \begin{pmatrix} Z_{n-2}^{(N)} & Z_{n-1}^{(N)} & Y_n^{(N)} & X_{n+1}^{(N)} \\ Y_{n-1}^{(N)} & X_n^{(N)} & & \end{pmatrix} \right\rangle \quad (3 \leq n \leq k_N, N \geq 1) \end{aligned}$$

(see (1.3.1) and (3.3.3)). Clearly  $\text{ess sup } |F| \leq 2\text{ess sup } |f|$  and  $|H| \leq \pi$ .

STEP 1:  $|\mathbb{E}(H_n^{(N)})| \leq \frac{\pi}{4} d_n^{(N)}(\xi)^2$ ,  $\mathbb{E}[(H_n^{(N)})^2] \leq \frac{\pi^2}{4} d_n^{(N)}(\xi)^2$ , and

$$\sup_N \mathbb{E}[(H_3^{(N)} + \dots + H_{k_N}^{(N)})^2] < \infty.$$

PROOF OF STEP 1. We fix  $N$  and drop the superscripts  $(N)$ .

The map  $\iota : \left( Z_{n-2} \frac{Z_{n-1}}{Y_{n-1}} \frac{Y_n}{X_n}, X_{n+1} \right) \mapsto \left( Z_{n-2} \frac{Y_{n-1}}{Z_{n-1}} \frac{X_n}{Y_n}, X_{n+1} \right)$  preserves the natural measure on the space of hexagons, and is an involution:  $\iota^2 = id$ . Clearly

$$\Gamma \circ \iota = -\Gamma.$$

Using the partial symmetry  $\langle -x \rangle = -\langle x \rangle$  for all  $x \notin -\pi + 2\pi\mathbb{Z}$ , we find that  $H_n \circ \iota = -H_n$  on  $[H_n \neq -\pi]$ . So  $\mathbb{E}(H_n 1_{[H_n \neq -\pi]}) = 0$ , and therefore

$$|\mathbb{E}(H_n)| = \pi \mathbb{P}(H_n = -\pi) \leq \frac{\pi}{4} \mathbb{E}(|e^{iH_n} - 1|^2) = \frac{\pi}{4} \mathbb{E}(|e^{i\xi\Gamma} - 1|^2) = \frac{\pi}{4} d_n(\xi)^2,$$

which is the first statement we needed to show.

Next we observe from (3.3.2) that  $\mathbb{E}(H_n^2) \leq \frac{\pi^2}{4} \mathbb{E}(|e^{i\xi\Gamma} - 1|^2) = \frac{\pi^2}{4} d_n^{(N)}(\xi)^2$ , which is the second statement we had to prove.

The two statements already proven and the boundedness of  $d_n$  show that there is a constant  $C$  s. t.  $\text{Var}(H_n) \leq C d_n^2(\xi)^2$ . Now the third statement follows from Lemma 2.5. The proof of step 1 is complete.

From now on, fix a constant  $D$  such that

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 + \sup_N \mathbb{E} \left[ \left( \sum_{n=3}^{k_N} H_n^{(N)} \right)^2 \right] < D.$$

STEP 2: For every  $N \geq 1$  there exists  $\underline{\zeta}^{(N)} = (\zeta_1^{(N)}, \dots, \zeta_{k_N+1}^{(N)}) \in \prod_{i=1}^{k_N+1} \mathfrak{S}_i^{(N)}$  s.t.

$$\sum_{n=3}^{k_N} \mathbb{E} \left( H_n^{(N)} (\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)})^2 \middle| \{Z_i^{(N)}\} = \underline{\zeta}^{(N)} \right) < \pi^2 D,$$

$$\mathbb{E} \left[ \left( \sum_{n=3}^{k_N} H_n^{(N)} (\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)}) \right)^2 \middle| \{Z_i^{(N)}\} = \underline{\zeta}^{(N)} \right] < \pi^2 D,$$

$$\mathbb{E}_X \left[ \sum_{n=3}^{k_N} \text{CVar} \left( \xi F_n^{(N)} (\underline{L}_n^{(N)}) \middle| \{Z_i^{(N)}\} = \underline{\zeta}^{(N)}, X_n^{(N)} \right) \right] < \pi^2 D \text{ and}$$

$$|f_n^{(N)}(\zeta_n^{(N)}, \zeta_{n+1}^{(N)})| \leq \text{ess sup } |f| \text{ for all } 3 \leq n \leq k_N.$$

Here and throughout  $\underline{L}_n = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)})$ , and  $\mathbb{E}_X$  indicates averaging on  $\{X_i^{(N)}\}$ .

PROOF OF STEP 2. We fix  $N$  and drop the  $^{(N)}$  superscripts.

Let  $\Omega_1 := \left\{ \underline{\zeta} : \sum_{n=3}^{k_N} \mathbb{E}(H_n^2 | \{Z_i\} = \underline{\zeta}) \leq \pi^2 D \right\}$ . By step 1,

$$\mathbb{E}_Z \left[ \mathbb{E} \left( \sum_{n=3}^{k_N} H_n^2 \middle| \{Z_i\} = \underline{\zeta} \right) \right] = \sum_{n=3}^{k_N} \mathbb{E}(H_n^2) \leq \frac{\pi^2}{4} \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 \leq \frac{\pi^2}{4} D,$$

where  $\mathbb{E}_Z$  = integration over  $\underline{\zeta}$  with respect to the distribution of  $\{Z_i^{(N)}\}$  (recall that  $\{Z_i^{(N)}\} \stackrel{dist}{=} \{X_i^{(N)}\}$ ). By Markov's inequality,  $\mathbb{P}[\{Z_i^{(N)}\} \in \Omega_1] > \frac{3}{4}$ .

Let  $\Omega_2 := \{\underline{\zeta} : \mathbb{E}[(\sum_{n=3}^{k_N} H_n(\underline{L}_n, \underline{L}_{n+1}))^2 | \{Z_i\} = \underline{\zeta}] < \pi^2 D\}$ . As before, by Markov's inequality,  $\mathbb{P}[\{Z_i^{(N)}\} \in \Omega_2] \geq 1 - \frac{1}{\pi^2}$ .

Let  $\Omega_3 := \left\{ \underline{\zeta} : \mathbb{E}_X \left[ \sum_{n=3}^{k_N} \text{CVar}(\xi F(\underline{L}_{n+1}) | \{Z_i\} = \underline{\zeta}, X_{n+1}) \right] < \pi^2 D \right\}$ ,

$$\theta^*(\underline{L}_n, X_{n+1}, Z_{n-1}) := -\xi f_{n-2}(Z_{n-2}, Z_{n-1}) + \xi F(\underline{L}_n) + \xi f_n(X_n, X_{n+1}).$$

Then  $\exp[iH_n(\underline{L}_n, \underline{L}_{n+1})] = \exp[i\xi F(\underline{L}_{n+1}) - i\theta^*(\underline{L}_{n+1}, X_{n+1}, Z_{n-1})]$ .

Given  $X_{n+1}$  and  $\{Z_i\}$ ,  $\underline{L}_{n+1}$  is conditionally independent from  $\underline{L}_n$ ,  $\{X_i\}_{i \neq n+1}$ . So

$$\begin{aligned} \mathbb{E}_{Z,X} \left( \text{CVar} \left( \xi F(\underline{L}_{n+1}) | \{Z_i\}, X_{n+1} \right) \right) &= \mathbb{E} \left( \text{CVar} \left( \xi F(\underline{L}_{n+1}) | \underline{L}_n, \{Z_i\}, \{X_i\} \right) \right) \\ &\stackrel{!}{\leq} \mathbb{E} \left( \mathbb{E} (|e^{i\xi F(\underline{L}_{n+1}) - i\theta^*(\underline{L}_n, X_{n+1}, Z_{n-1})} - 1|^2 | \underline{L}_n, \{X_i\}, \{Z_i\}) \right) \\ &= \mathbb{E} (|e^{i(\xi F(\underline{L}_{n+1}) - \theta^*)} - 1|^2) \equiv \mathbb{E} (|e^{iH_n} - 1|^2) = \mathbb{E} (|e^{i\xi \Gamma} - 1|^2) = d_n(\xi)^2, \end{aligned}$$

where  $\stackrel{!}{\leq}$  is because  $\theta^*$  is conditionally constant. So

$$\mathbb{E}_Z \left[ \mathbb{E}_X \left( \sum_{n=3}^{k_N} \text{CVar}(\xi F(\underline{L}_{n+1}) | \{Z_i\}, X_{n+1}) \right) \right] < D.$$

By Markov's inequality,  $\mathbb{P}(\{Z_i^{(N)}\} \in \Omega_3) \geq 1 - \frac{1}{\pi^2}$ .

Finally, let  $\Omega_4 := \{\underline{\zeta} : |f_n(\zeta_n, \zeta_{n+1})| \leq \text{ess sup } |\mathbf{f}| \}$ , then  $\mathbb{P}(\{Z_i^{(N)}\} \in \Omega_4) = 1$ .

In summary  $\mathbb{P} \left[ \bigcup_{1 \leq i \leq 4} \Omega_i^c \right] \leq \frac{2}{\pi^2} + \frac{1}{4} < 1$ . Necessarily  $\Omega_1 \cap \Omega_2 \cap \Omega_3 \cap \Omega_4 \neq \emptyset$ . Any  $\underline{\zeta} = \underline{\zeta}^{(N)}$

in the intersection satisfies the requirements of step 2.

STEP 3: *There exist measurable functions  $\theta_n^{(N)} : \mathfrak{S}_n^{(N)} \rightarrow [-\pi, \pi]$  s.t.*

$$\sum_{n=3}^{k_N} \mathbb{E} \left( |e^{i\xi F_n^{(N)}(\underline{L}_n^{(N)}) - i\theta_n^{(N)}(X_n)} - 1|^2 | \{Z_i^{(N)}\} = \underline{\zeta}^{(N)} \right) < 2\pi^2 D.$$

*Proof.* We fix  $N$  and drop the  $^{(N)}$  superscripts.

Clearly,  $\theta \mapsto \mathbb{E}(|e^{i(W-\theta)} - 1|^2)$  is continuous for every random variable  $W$ . So  $\text{CVar}(W) = \inf_{q \in \mathbb{Q}} \mathbb{E}(|e^{i(W-q)} - 1|^2)$ , an infimum over a countable set, whence  $\text{CVar}(\xi F_n | \{Z_i\} = \underline{\zeta}, X_n) = \inf_{q \in \mathbb{Q}} \mathbb{E}(|e^{i\xi F(\underline{L}_n) - iq} - 1|^2 | \{Z_i\} = \underline{\zeta}, X_n = \xi_n)$ .

The expectation can be expressed explicitly using integrals with respect to the bridge distributions, and this expression shows that

$$\eta \mapsto \text{CVar}(\xi F_n | \{Z_i\} = \underline{\zeta}, X_n = \eta)$$

is measurable on  $\mathfrak{S}_n^{(N)}$ .

Fix  $N$  and  $\underline{\zeta} = \underline{\zeta}^{(N)}$ . We say that  $(\eta, q) \in S_n^{(N)} \times \mathbb{R}$  have “property  $P_n(\eta, q)$ ”, if the following condition holds:

$$\begin{aligned} & \mathbb{E}(|e^{i\xi F_n(\underline{L}_n) - iq} - 1|^2 | \{Z_n\} = \underline{\zeta}, X_n = \eta) \\ & \leq \text{CVar}(\xi F_n(\underline{L}_n) | \{Z_n\} = \underline{\zeta}, X_n = \eta) + \frac{D}{n^2} \end{aligned} \quad (P_n(\xi, q))$$

By the previous paragraph,  $\{\eta : P_n(\eta, q) \text{ holds}\}$  is measurable, and for every  $\eta$  there exists  $q \in \mathbb{Q} \cap (-\pi, \pi)$  such that  $P_n(\eta, q)$  holds. Let

$$\theta_n(\eta) = \theta_n^{(N)}(\eta) := \inf \{q : q \in \mathbb{Q} \cap (-\pi, \pi) \text{ s.t. } P_n(\eta, q) \text{ holds}\}.$$

Again, this is a measurable function, and since for fixed  $\eta$ ,  $P_n(\eta, q)$  is a closed property of  $q$ ,  $\theta_n^{(N)}(\eta)$  itself satisfies property  $P_n(\eta, \theta_n^{(N)}(\xi))$ . So

$$\begin{aligned} & \mathbb{E}_X \left[ \sum_{n=3}^{k_N} \mathbb{E}(|e^{i\xi F_n(\underline{L}_n) - i\theta_n^{(N)}(X_n)} - 1|^2 | \{Z_n\} = \underline{\zeta}, X_n) \right] \\ & \leq \mathbb{E}_X \left[ \sum_{n=3}^{k_N} \text{CVar}(\xi F_n(\underline{L}_n) | \{Z_n\} = \underline{\zeta}, X_n) \right] + \frac{\pi^2}{6} D \\ & < 2\pi^2 D, \text{ by choice of } \underline{\zeta}. \end{aligned}$$

STEP 4 (THE REDUCTION). Let  $\underline{\zeta} = \underline{\zeta}^{(N)}$ ,  $\theta_n = \theta_n^{(N)}$ ,  $f_n = f_n^{(N)}$ ,  $F_n = F_n^{(N)}$ ,  $X_n = X_n^{(N)}$ ,  $Z_n = Z_n^{(N)}$ . Define

$$\begin{aligned} c_n^{(N)} &:= f_n(\zeta_{n-2}, \zeta_{n-1}) \\ a_n^{(N)}(x) &:= \frac{1}{\xi} \left[ \theta_n(X_n) + \mathbb{E}(\langle \xi F_n(\underline{L}_n) - \theta_n(X_n) \rangle | \{Z_i\} = \underline{\zeta}, X_n = x) \right] \quad (x \in \mathfrak{S}_n^{(N)}) \\ \tilde{f} &:= \frac{1}{\xi} \left\langle \xi(f - \nabla a - c) \right\rangle \\ g &:= f - \nabla a - c - \tilde{f}. \end{aligned}$$

Then  $a, c, \tilde{f}, g$  are uniformly bounded, and  $G_{alg}(g) \subset \frac{2\pi}{\xi} \mathbb{Z}$ .

*Proof.* By choice of  $\underline{\zeta}^{(N)}$ ,  $|c| \leq \text{ess sup}|f|$ , and by the definition of  $\theta^{(N)}$  and  $\langle \cdot \rangle$ ,  $|a| \leq 2\pi/|\xi|$  and  $|\tilde{f}| \leq \pi/|\xi|$ . It follows that  $|g| \leq 2\text{ess sup}|f| + 3\pi/|\xi|$ . Next,

$$g \equiv \frac{1}{\xi} \left( \xi(f - \nabla a - c) - \langle \xi(f - \nabla a - c) \rangle \right).$$

The term in the brackets belongs to  $2\pi\mathbb{Z}$  by the definition of  $\langle \cdot \rangle$ , so  $G_{alg}(g) \subset \frac{2\pi}{\xi} \mathbb{Z}$ , and the proof of step 4 is complete.

Notice that  $f - g = \nabla a + c + \tilde{f}$ . Gradients and constant functionals are center tight. So to complete the proof of the reduction lemma, it suffices to show:

STEP 5:  $\tilde{f}$  is center-tight.

*Proof.* We fix  $N$  and drop the  $^{(N)}$  superscripts.

We begin with a few identities. Suppose  $\{Z_i^{(N)}\} = \{\zeta_i^{(N)}\}$ , and consider the hexagon  $P_n := \begin{pmatrix} Z_{n-2} & Z_{n-1} & Y_n & X_n & X_{n+1} \\ Y_{n-2} & Y_{n-1} & X_n & X_n & X_{n+1} \end{pmatrix} = \begin{pmatrix} \zeta_{n-2} & \zeta_{n-1} & Y_n & X_n & X_{n+1} \\ Y_{n-1} & X_n & X_n & X_{n+1} & X_{n+1} \end{pmatrix}$ , then

$$-\Gamma(P_n) = -f_{n-2}(Z_{n-2}, Z_{n-1}) - F_{n+1}(\underline{L}_{n-1}) + F_n(\underline{L}_n) + f_n(X_n, X_{n+1}),$$

whence

$$\begin{aligned} \xi \tilde{f}_n(X_n, X_{n+1}) &= \left\langle \xi \left( -\Gamma(P_n) + a_n(X_n) - F(\underline{L}_n) + F(\underline{L}_{n+1}) - a_{n+1}(X_{n+1}) \right) \right\rangle \\ &= \left\langle -H_n(\underline{L}_n, \underline{L}_{n+1}) + \xi (a_n(X_n) - F(\underline{L}_n)) + \xi (F(\underline{L}_{n+1}) - a_{n+1}(X_{n+1})) \right\rangle. \end{aligned}$$

Define a new functional  $W$  of the ladder process  $\{\underline{L}_n\}$  with entries

$$W(\underline{L}_n) := \langle \xi F(\underline{L}_n) - \theta_n(X_n) \rangle - \mathbb{E} \left( \langle \xi F(\underline{L}_n) - \theta_n(X_n) \rangle \mid \{Z_i\} = \underline{\zeta}^{(N)}, X_n \right).$$

Notice that  $W(\underline{L}_n) = \xi(F(\underline{L}_n) - a_n(X_n)) \pmod{2\pi\mathbb{Z}}$ . Therefore

$$\xi \tilde{f}_n(X_n, X_{n+1}) = \left\langle W(\underline{L}_{n+1}) - W(\underline{L}_n) - H_n(\underline{L}_n, \underline{L}_{n+1}) \right\rangle. \quad (3.3.5)$$

CLAIM. Given  $\delta > 0$ , let  $T_\delta = 11\pi^2 D/\delta$ . Then there exists a measurable set  $\Omega_X$  of  $\{X_i\}$  such that  $\mathbb{P}(\Omega_X) > 1 - \delta$  and such that for all  $\underline{\xi} \in \Omega_X$ ,

- (1)  $\sum_{n=3}^{k_N} \mathbb{P} \left( |W(\underline{L}_n)| > \frac{\pi}{3} \mid \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) < T_\delta,$
- (2)  $\sum_{n=3}^{k_N} \mathbb{P} \left( |H_n(\underline{L}_n, \underline{L}_{n+1})| > \frac{\pi}{3} \mid \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) < T_\delta,$
- (3)  $\mathbb{E} \left( \left| \sum_{n=3}^{k_N} H_n(\underline{L}_n, \underline{L}_{n+1}) \right| \mid \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) < T_\delta.$

*Proof of the claim.*  $\underline{L}_n$  is conditionally independent of  $\{X_i\}_{i \neq n}$  given  $\{Z_i\}, X_n$ . So

$$\begin{aligned} &\sum_{n=3}^{k_N} \mathbb{P} \left( |W(\underline{L}_n)| \geq \frac{\pi}{4} \mid \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) \\ &= \sum_{n=3}^{k_N} \mathbb{P} \left( |W(\underline{L}_n)| \geq \frac{\pi}{4} \mid \{Z_i\} = \underline{\zeta}, X_n = \xi_n \right). \end{aligned}$$

Since  $\mathbb{E}(W(\underline{L}_n) \mid \{Z_i\} = \underline{\zeta}, X_n) = 0$ , we can use the Chebyshev inequality to bound the sum of probabilities from above by

$$\begin{aligned} &\leq \frac{16}{\pi^2} \sum_{n=3}^{k_N} \text{Var}(\langle \xi F(\underline{L}_n) - \theta_n(X_n) \rangle \mid \{Z_i\} = \underline{\zeta}, X_n) \\ &\leq 4 \sum_{n=3}^{k_N} \mathbb{E}(|e^{i\xi F(\underline{L}_n) - i\theta_n(X_n)} - 1|^2 \mid \{Z_i\} = \underline{\zeta}, X_n), \text{ see (3.3.2).} \end{aligned}$$

Integrating over  $\{X_i\}$  we have by the choice of  $\theta_n^{(N)}(X_n)$  (step 3) that

$$\mathbb{E}_X \left[ \sum_{n=3}^{k_N} \mathbb{P} \left( |W(\underline{L}_n)| \geq \frac{\pi}{4} \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) \right] \leq 8\pi^2 D.$$

By Markov's inequality, the set

$$\Omega_X^1(T) := \left\{ \underline{\xi} : \sum_{n=3}^{k_N} \mathbb{P} \left( |W(\underline{L}_n)| > \frac{\pi}{3} \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) \leq T \right\}$$

has probability  $\mathbb{P}[\Omega_X^1(T)] \geq 1 - 8\pi^2 D/T$ .

Similarly, by Markov's inequality

$$\mathbb{P} \left( |H_n| \geq \frac{\pi}{4} \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) \leq \frac{16}{\pi^2} \mathbb{E} \left( H_n^2 \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right).$$

By the choice of  $\underline{\zeta}$ ,  $\mathbb{E}_X \left[ \sum_{n=3}^{k_N} \mathbb{P} \left( |H_n| \geq \frac{\pi}{4} \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} \right) \right] \leq 16D$ . So the set

$$\Omega_X^2(T) := \left\{ \underline{\xi} : \sum_{n=3}^{k_N} \mathbb{P} \left( |H_n(\underline{L}_n, \underline{L}_{n+1})| > \frac{\pi}{3} \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) \leq T \right\}$$

has probability  $\mathbb{P}[\Omega_X^2(T)] \geq 1 - 16D/T > 1 - 2\pi^2 D/T$ .

Finally, since conditional expectations contract  $L^2$ -norms,

$$\begin{aligned} & \mathbb{E}_X \left[ \mathbb{E} \left( \left| \sum_{n=3}^{k_N} H_n(\underline{L}_n, \underline{L}_{n+1}) \right| \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right)^2 \right] \\ & \leq \mathbb{E} \left[ \left( \sum_{n=3}^{k_N} H_n(\underline{L}_n, \underline{L}_{n+1}) \right)^2 \middle| \{Z_i\} = \underline{\zeta} \right] \leq \pi^2 D. \end{aligned}$$

So  $\Omega_X^3(T) := \left\{ \underline{\xi} : \mathbb{E} \left( \left| \sum_{n=3}^{k_N} H_n(\underline{L}_n, \underline{L}_{n+1}) \right| \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) \leq T \right\}$  has probability

$$\mathbb{P}[\Omega_X^3(T)] > 1 - \pi^2 D/T^2.$$

We see that if  $T > 1$ , then  $\mathbb{P}[\Omega_X^1(T) \cap \Omega_X^2(T) \cap \Omega_X^3(T)] > 1 - \frac{11\pi^2 D}{T}$ . The claim follows.  $\square$

We can now complete the proof of the step 5 (and the reduction lemma) and show that  $\tilde{f}$  is center-tight.

Fix  $\delta > 0$  and  $\Omega_X, T_\delta$  as in the claim. Fix  $N$  and define the random set

$$A_N(\{\underline{L}_n^{(N)}\}) := \{3 \leq n \leq k_N : |W(\underline{L}_n)| \geq \frac{\pi}{3} \text{ or } |H_n(\underline{L}_n, \underline{L}_{n+1})| \geq \frac{\pi}{3}\}.$$

For all  $\underline{\xi} \in \Omega_X$ , we have the following bound (Lemma 3.11):

$$\mathbb{P} \left( |A_N| > 4T_\delta \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) < \frac{1}{2}.$$

Similarly, for all  $\underline{\xi} \in \Omega_X$ ,  $\mathbb{P} \left( \left| \sum_{n=3}^{k_N} H_n \right| > 4T_\delta \middle| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi} \right) \leq \frac{1}{4}$ .

Since the probabilities of these events add up to less than one, the intersection of their complements is non-empty. So for every  $\underline{\xi} \in \Omega_X$  we can find  $\{Y_i^{(N)}(\underline{\xi})\}_{i=2}^{k_N-1}$  such that  $\underline{L}_n^* := \underline{L}_n^*(\underline{\xi}) = (\zeta_{n-2}^{(N)}, Y_{n-1}^{(N)}(\underline{\xi}), \xi_n)$  has the following two properties:

$$\left| \sum_{n=3}^{k_N} H_n(\underline{L}_n^*, \underline{L}_{n+1}^*) \right| \leq 4T_\delta, \text{ and}$$

$$M := \#\left\{3 \leq n \leq k_N : |W(\underline{L}_n^*)| \geq \frac{\pi}{3} \text{ or } |H_n(\underline{L}_n^*, \underline{L}_{n+1}^*)| \geq \frac{\pi}{3}\right\} \leq 4T_\delta.$$

Let  $n_1 < \dots < n_M$  be an enumeration of the indices  $n$  where  $|W(\underline{L}_n^*)| \geq \frac{\pi}{3}$  or  $|H_n(\underline{L}_n^*, \underline{L}_{n+1}^*)| \geq \frac{\pi}{3}$ . By (3.3.5), if  $n_i < n < n_{i+1} - 1$ ,

$$\xi \tilde{f}_n(\xi_n, \xi_{n+1}) = W(\underline{L}_{n+1}^*) - W(\underline{L}_n^*) - H_n(\underline{L}_n^*, \underline{L}_{n+1}^*),$$

because  $\langle x + y + z \rangle = x + y + z$  whenever  $|x|, |y|, |z| < \frac{\pi}{3}$ . So

$$- \sum_{n=n_i}^{n_{i+1}-1} \xi \tilde{f}_n(\xi_n, \xi_{n+1}) = \sum_{n=n_i+1}^{n_{i+1}-1} H_n(\underline{L}_n^*, \underline{L}_{n+1}^*) \pm 6\pi$$

where we have used the bounds  $|W| \leq 2\pi$  and  $|H_n| \leq \pi$ . Summing over  $i$  we find that for every  $\underline{\xi} \in \Omega_X$ ,

$$\left| \xi \sum_{n=3}^{k_N} \tilde{f}_n(\xi_n, \xi_{n+1}) \right| \leq \left| \sum_{n=3}^{k_N} H_n(\underline{L}_n^*, \underline{L}_{n+1}^*) \right| + 10M\pi \leq 4T_\delta + 40T_\delta\pi < 42\pi T_\delta.$$

Setting  $C_\delta := 42\pi T_\delta / \xi$ , we find that  $\mathbb{P}(|\sum_{n=3}^{k_N} \tilde{f}_n^{(N)}| \geq C_\delta) < \delta$  for all  $N$ , whence the (center-)tightness of  $\tilde{f}$ .  $\square$

In chapter 5 we will need the following variant of the reduction lemma for integer valued  $f$ .

**Lemma 3.13 (Integer Reduction Lemma)** *Let  $X$  be a uniformly elliptic Markov chain, and  $f$  an integer valued additive functional on  $X$  s.t.  $|f| \leq K$  a.s. For every  $N$ ,  $f_n(x, y) = g_n^{(N)}(x, y) + a_n^{(N)}(x) - a_{n+1}^{(N)}(y) + c_n^{(N)}$  ( $n = 1, \dots, N$ ) where*

- (1)  $c_n^{(N)}$  are integers such that  $|c_n^{(N)}| \leq K$ ,
- (2)  $a_n^{(N)}$  are measurable integer valued functions on  $\mathfrak{S}_n$  s.t.  $|a_n^{(N)}| \leq 2K$ ,
- (3)  $g_n^{(N)}$  are measurable, integer valued, and  $\sum_{n=3}^N \mathbb{E}[g_n^{(N)}(X_n, X_{n+1})^2] \leq 10^3 K^4 \sum_{n=3}^N u_n^2$ , with  $u_n$  the structure constants of  $f$ .

*Proof.* Let  $\left( Z_{n-2} \frac{Z_{n-1}}{Y_{n-1}} \frac{Y_n}{X_n} X_{n+1} \right)$  be a random hexagon. By the definition of the structure constants,

$$\mathbb{E} \left[ \sum_{n=3}^N \mathbb{E} \left( \Gamma \left( Z_{n-2} \frac{Z_{n-1}}{Y_{n-1}} \frac{Y_n}{X_n} X_{n+1} \right)^2 \middle| Z_{n-2}, Z_{n-1} \right) \right] = \sum_{n=3}^N u_n^2.$$

Therefore, for every  $N$  there exists  $z_n = z_n(N) \in \mathfrak{S}_n$  ( $n = 1, \dots, N-2$ ) such that

$$\sum_{n=3}^N \mathbb{E} \left[ \mathbb{E} \left( \Gamma \left( Z_{n-2} \frac{Z_{n-1} Y_n}{Y_{n-1} X_n} X_{n+1} \right)^2 \middle| Z_{n-2} = z_{n-2}, Z_{n-1} = z_{n-1} \right) \right] \leq \sum_{n=3}^N u_n^2.$$

We emphasize that  $z_n$  depends on  $N$ .

Let  $c_n^{(N)} := f_{n-2}(z_{n-2}, z_{n-1})$ , and let  $a_n^{(N)}(x_n)$  be the (smallest) most likely value of

$$f_{n-2}(z_{n-2}, Y) + f_{n-1}(Y, x_n),$$

where  $Y$  has the bridge distribution of  $X_{n-1}$  conditioned on  $X_{n-2} = z_{n-2}$  and  $X_n = x_n$ . The most likely value exists, and has probability bigger than  $\delta_K := \frac{1}{5K}$ , because  $f_{n-2}(z_{n-2}, Y) + f_{n-1}(Y, x_n) \in [-2K, 2K] \cap \mathbb{Z}$ .

Set  $g_n^{(N)}(x_n, x_{n+1}) := f_n(x_n, x_{n+1}) + a_n^{(N)}(x_n) - a_{n+1}^{(N)}(x_{n+1}) - c_n^{(N)}$ . Equivalently,  $g_n^{(N)}(x_n, x_{n+1}) = -\Gamma \left( Z_{n-2} \frac{z_{n-1} y_n}{y_{n-1} x_n} x_{n+1} \right)$  for the  $y_k$  which maximize the likelihood of the value  $f_{k-1}(z_{k-1}, Y) + f_k(Y, x_{k+1})$  when  $Y$  has the bridge distribution of  $X_k$  given  $X_{k-1} = z_{k-1}, X_{k+1} = x_{k+1}$ .

Our task is to estimate  $\sum_{n=3}^N \mathbb{E}[g_n^{(N)}(X_n, X_{n+1})^2]$ . Define for this purpose the functions  $h_n^{(N)} : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \rightarrow \mathbb{R}$ ,

$$h_n^{(N)}(x_n, x_{n+1}) := \mathbb{E} \left( \Gamma \left( Z_{n-2} \frac{Z_{n-1} Y_n}{Y_{n-1} X_n} X_{n+1} \right)^2 \middle| \begin{array}{ll} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{array} \right)^{1/2},$$

Our plan is to show the following:

- (a)  $\sum_{n=3}^N \mathbb{E}(h_n^{(N)}(X_n, X_{n+1})^2) \leq \sum_{n=3}^N u_n^2$
- (b) If  $h_n^{(N)}(x_n, x_{n+1}) < \delta_K$ , then  $g_n^{(N)}(x_n, x_{n+1}) = 0$ .
- (c)  $\mathbb{E}(g_n^{(N)}(X_n, X_{n+1})^2) \leq (6K)^2 \mathbb{P}[h_n^{(K)} \geq \delta_K] \leq 36K^2 \delta_K^{-2} \mathbb{E}[h_n^{(N)}(X_n, X_{n+1})^2]$ .

Part (a) is because of the choice of  $z_n$ . To see part (b), note that since  $f$  is integer valued, either the balance of a hexagon is zero, or it has absolute value  $\geq 1$ . Therefore, if  $h_n^{(N)}(X_n, X_{n+1}) < \delta_K$ , then necessarily

$$\begin{aligned} & \mathbb{P} \left[ \Gamma \left( Z_{n-2} \frac{Z_{n-1} Y_n}{Y_{n-1} X_n} X_{n+1} \right) \neq 0 \middle| \begin{array}{ll} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{array} \right] \\ & \leq \mathbb{E} \left[ \Gamma \left( Z_{n-2} \frac{Z_{n-1} Y_n}{Y_{n-1} X_n} X_{n+1} \right)^2 \middle| \begin{array}{ll} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{array} \right] \\ & = h_n^{(N)}(X_n, X_{n+1})^2 < \delta_K^2, \end{aligned}$$

$$\text{whence } \mathbb{P} \left[ \Gamma \left( Z_{n-2} \frac{Z_{n-1} Y_n}{Y_{n-1} X_n} X_{n+1} \right) = 0 \middle| \begin{array}{ll} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{array} \right] > 1 - \delta_K^2.$$

At the same time, by the structure of the distribution of random hexagons,

$$\Omega_n := \left\{ \left( Z_{n-2} \frac{Z_{n-1} Y_n}{Y_{n-1} X_n} X_{n+1} \right) : \begin{array}{l} f_{n-1}(Z_{n-1}, Y_n) + f_n(Y_n, X_{n+1}) = a_{n+1}^{(N)}(X_{n+1}) \\ f_{n-2}(Z_{n-2}, Y_{n-1}) + f_{n-1}(Y_{n-1}, X_n) = a_n^{(N)}(X_n) \end{array} \right\}$$

$$\text{satisfies } \mathbb{P} \left[ \Omega_n \middle| \begin{array}{ll} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{array} \right] > \delta_K^2,$$

If the sum of the probabilities of two events is bigger than one, then they must intersect. It follows that there exist  $y_{n-1}, y_n$  such that

- $a_n^{(N)}(X_n) = f_{n-2}(z_{n-2}, y_{n-1}) + f_{n-1}(y_{n-1}, X_n);$
- $a_{n+1}^{(N)}(X_{n+1}) = f_{n-1}(z_{n-1}, y_n) + f_n(y_n, X_{n+1});$
- $\Gamma \begin{pmatrix} z_{n-2} & z_{n-1} & y_n & X_n & X_{n+1} \end{pmatrix} = 0.$

By the definition of  $g_n^{(N)}$ , this implies that  $g_n^{(N)}(X_n, X_{n+1}) = 0$ , which proves part (b).

Part (c) follows from part (b), Chebyshev's inequality, and the estimate  $\|g_n^{(N)}\|_\infty \leq 6K$  (as is true for the balance of every hexagon).  $\square$

Combining Lemmas 3.10 and 3.13 we obtain the following result

**Corollary 3.14 (Joint Reduction)** *There is a constant  $L = L(\varepsilon, K)$  such that under the conditions of the Reduction Lemma we can arrange, in addition to the other conclusions of Lemma 3.10, that  $\sum_{n=3}^{k_N} \|g_n^{(N)}\|_2^2 \leq LU_N$ .*

*Proof.* Apply Lemma 3.10 and then apply Lemma 3.13 to the resulting integer valued additive functional  $\frac{\xi g}{2\pi}$ . Notice that the reduction in this corollary depends on  $N$  even if  $f$  is an additive functional of a Markov chain.  $\square$

Corollary 3.14 says the following. Suppose we have an additive functional  $f$  such that both  $U_N$  is small and  $D_N(\xi)$  is small for some  $\xi$  (but  $D_N(\xi)$  can be much smaller than  $U_N$ ). Then we can adjust  $f$  such that at time  $N$ , the resulting functional will have a small norm as prescribed by  $U_N$  and small distance to  $\frac{2\pi}{\xi} \mathbb{Z}$  as prescribed by  $D_N$  at the same time.

### 3.3.2 The possible values of the co-range

We prove Theorem 3.1 in its version for Markov arrays: *The co-range of an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$  is equal to  $\mathbb{R}$  when  $f$  is center tight, and to  $\{0\}$  or  $t\mathbb{Z}$  ( $t > 0$ ) otherwise.*

Recall that the co-range is defined by

$$H := H(X, f) = \{\xi \in \mathbb{R} : \sup_N D_N(\xi) < \infty\}, \text{ where } D_N(\xi) = \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2.$$

STEP 1.  $H$  is a subgroup of  $\mathbb{R}$ .

*Proof.*  $H = -H$ , because  $d_n^{(N)}(-\xi) = d_n^{(N)}(\xi)$ .  $H \ni 0$ , because  $d_n^{(N)}(0) = 0$ .  $H$  is closed under addition, because if  $\xi, \eta \in H$ , then by Lemma 1.15,

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi + \eta)^2 \leq 8 \left[ \sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 + \sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\eta)^2 \right] < \infty.$$

STEP 2. If  $f$  is center-tight, then  $H = \mathbb{R}$ .

*Proof.* Suppose  $f$  is center-tight. By Corollary 2.8 and the center-tightness of  $f$ ,  $\sup_N \sum_{k=3}^{k_N} (u_k^{(N)})^2 < \infty$ . By Lemma 1.15(c),  $\sup_N \sum_{k=3}^{k_N} d_n^{(N)}(\xi)^2 < \infty$  for all  $\xi \in \mathbb{R}$ .

STEP 3. *If  $f$  is not center-tight, then  $\exists t_0$  s.t.*

$$H \cap (-t_0, t_0) = \{0\}. \quad (3.3.6)$$

*Proof.* Let  $K := \text{ess sup } |f|$ , then  $|\Gamma(P)| \leq 6K$  for a.e. hexagon  $P$ .

Fix  $\tau_0 > 0$  such that  $|e^{it} - 1|^2 \geq \frac{1}{2}t^2$  for all  $|t| < \tau_0$ , and let  $t_0 := \tau_0(6K)^{-1}$ . Then for all  $|\xi| < t_0$ ,  $|e^{i\xi\Gamma(P)} - 1|^2 \geq \frac{1}{2}\xi^2\Gamma(P)^2$  for all hexagons  $P$ .

Taking the expectation over  $P \in \text{Hex}(N, n)$ , we obtain that

$$d_n^{(N)}(\xi)^2 \geq \frac{1}{2}\xi^2(u_n^{(N)})^2 \text{ for all } |\xi| < t_0, 1 \leq n \leq k_N, N \geq 1. \quad (3.3.7)$$

Now assume by way of contradiction that there is  $0 \neq \xi \in H \cap (-t_0, t_0)$ , then  $\sup_N \sum_{n=3}^{k_N} (u_n^{(N)})^2 \leq \frac{2}{\xi^2} \sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 < \infty$ . By Corollary 2.8,  $f$  is center-tight, in contradiction to our assumption.

STEP 4. *If  $f$  is not center-tight, then  $H = \{0\}$ , or  $H = t\mathbb{Z}$  with  $t \geq \frac{\pi}{6\text{ess sup } |f|}$ .*

*Proof.* By steps 2 and 3,  $H$  is a proper closed subgroup of  $\mathbb{R}$ . So it must be equal to  $\{0\}$  or  $t\mathbb{Z}$  where  $t > 0$ . To see that  $t \geq \frac{\pi}{6\text{ess sup } |f|}$ , assume by contradiction that  $t = (\frac{\pi}{6\text{ess sup } |f|})\rho$  with  $0 < \rho < 1$ , and let  $\kappa := \min\{|e^{iu} - 1|^2/|u|^2 : |u| \leq \pi\rho\} > 0$ . Then  $|t\Gamma(P)| \leq 6t\text{ess sup } |f| = \pi\rho$  for every position  $n$  hexagon  $P$ , whence

$$d_n^2(t) = \mathbb{E}(|e^{it\Gamma} - 1|^2) \geq \kappa\mathbb{E}(\Gamma^2) = \kappa u_n^2.$$

This is impossible:  $t \in H$  so  $\sum d_n^2(t) < \infty$ , whereas  $f$  is not center-tight so  $\sum u_n^2 = \infty$ .  $\square$

### 3.3.3 Calculation of the essential range

We prove Theorem 3.2 in its version for Markov arrays: *For every a.s. uniformly bounded additive functional  $f$  on a uniformly elliptic Markov array  $X$ ,*

$$G_{ess}(X, f) = \begin{cases} \{0\} & H(X, f) = \mathbb{R} \\ \frac{2\pi}{\xi}\mathbb{Z} & H(X, f) = \xi\mathbb{Z} \\ \mathbb{R} & H(X, f) = \{0\}. \end{cases} \quad (3.3.8)$$

**Lemma 3.15** *Suppose  $f, g$  are two a.s. uniformly bounded additive functionals on the same uniformly elliptic Markov array. If  $f - g$  is center-tight, then  $f$  and  $g$  have the same co-range.*

*Proof.* By Corollary 2.8, if  $h = g - f$  is center-tight, then  $\sup_N \sum_{n=3}^{k_N} u_n^{(N)}(h)^2 < \infty$ . By Lemma 1.15 (b),(c),

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi, g)^2 \leq 8 \sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi, f)^2 + 8\xi^2 \sup_N \sum_{n=3}^{k_N} u_n^{(N)}(h)^2.$$

So the co-range of  $f$  is a subset of the co-range of  $g$ . By symmetry they are equal.  $\square$

**Proof of Theorem 3.2:** As we saw in the previous section, the possibilities for the co-range are  $\mathbb{R}$ ,  $t\mathbb{Z}$  with  $t \neq 0$ , and  $\{0\}$ .

CASE 1: *The co-range equals  $\mathbb{R}$ .* As we saw above, this can only happen if  $f$  is center-tight, in which case the essential range is  $\{0\}$  because we may subtract  $f$  from itself.

CASE 2.: *The co-range equals  $\xi\mathbb{Z}$  with  $\xi \neq 0$ .* We show that  $G_{ess}(X, f) = \frac{2\pi}{\xi}\mathbb{Z}$ .

By assumption,  $\xi$  is in the co-range:  $\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi, g)^2 < \infty$ . By the Reduction Lemma,  $f$  differs by a center-tight functional from a functional with algebraic range  $\subseteq \frac{2\pi}{\xi}\mathbb{Z}$ . So  $G_{ess}(X, f) \subseteq \frac{2\pi}{\xi}\mathbb{Z}$ .

Assume by way of contradiction that  $G_{ess}(X, f) \subsetneq \frac{2\pi}{\xi}\mathbb{Z}$ , then there exists a center-tight  $h$  such that the algebraic range of  $g := f - h$  is a subset of  $\frac{2\pi\ell}{\xi}\mathbb{Z}$  for some integer  $\ell > 1$ . The structure constants of  $g$  must satisfy  $d_n^{(N)}(\frac{\xi}{\ell}, g) \equiv 0$ , whence  $\frac{\xi}{\ell} \in$  co-range of  $g$ . By Lemma 3.15,  $\frac{\xi}{\ell} \in$  co-range of  $f$ , whence  $\frac{\xi}{\ell} \in \xi\mathbb{Z}$ . But this contradicts  $\ell > 1$ .

CASE 3.: *The co-range equals  $\{0\}$ .* We claim that the essential range is  $\mathbb{R}$ . Otherwise, there exists a center-tight  $h$  such that the algebraic range of  $g := f - h$  equals  $t\mathbb{Z}$  with  $t \neq 0$  or  $\{0\}$ . But this is impossible:

- (a) If the algebraic range of  $g$  is  $t\mathbb{Z}$ , then  $d_n^{(N)}(\frac{2\pi}{t}, g) = 0$  for all  $3 \leq n \leq k_N$ ,  $N \geq 1$ , so the co-range of  $g$  contains  $2\pi/t$ . By Lemma 3.15, the co-range of  $f$  contains  $2\pi/t$ , in contradiction to the assumption that it is  $\{0\}$ .
- (b) If the algebraic range of  $g$  is  $\{0\}$ , then  $f \equiv h$ , and  $f$  is center-tight. But by Theorem 3.1, the co-range of a center-tight functional is  $\mathbb{R}$ , whereas the co-range of our functional is  $\{0\}$ .  $\square$

### 3.3.4 Existence of irreducible reductions

We prove Theorem 3.3, in its version for Markov arrays: *For every a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$ , there exists an irreducible functional  $g$  such that  $f - g$  is center-tight and  $G_{alg}(X, g) = G_{ess}(X, g) = G_{ess}(X, f)$ .*

**Proof.** The essential range is a closed subgroup of  $\mathbb{R}$ , so  $G_{ess}(X, f) = \{0\}, t\mathbb{Z}$  or  $\mathbb{R}$ .

- (a) If  $G_{ess}(X, f) = \{0\}$ , then  $H(X, f) = \mathbb{R}$ , and  $f$  is center-tight. So take  $g \equiv 0$ .
- (b) If  $G_{ess}(X, f) = t\mathbb{Z}$  with  $t \neq 0$ , then by Theorem 3.2 the co-range of  $f$  is  $\xi\mathbb{Z}$  with  $\xi := 2\pi/t$ . So  $\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi, f)^2 < \infty$ . By the reduction lemma, there exists an additive functional  $g$  such that  $f - g$  is center-tight, and  $G_{alg}(X, g) \subseteq t\mathbb{Z}$ . By Lemma 3.15  $G_{ess}(X, f) = G_{ess}(X, g)$ , whence  $G_{ess}(X, f) = G_{ess}(X, g) \subseteq G_{alg}(X, g) \subseteq t\mathbb{Z} = G_{ess}(X, f)$ , and  $G_{ess}(X, g) = G_{alg}(X, g) = G_{ess}(X, f)$ .
- (c) If  $G_{ess}(X, f) = \mathbb{R}$ , take  $g := f$ .  $\square$

### 3.3.5 Proofs of results on hereditary arrays

**Proof of Theorem 3.6:** Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$ .

The first part of the theorem asks for the equivalence of the following conditions:

- (1)  $f$  is hereditary
- (2) for all  $\xi$ ,  $\liminf_{N \rightarrow \infty} \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty \Rightarrow \limsup_{N \rightarrow \infty} \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty$
- (3) for all  $\xi \notin H(X, f)$ ,  $D_N(\xi) \xrightarrow{N \rightarrow \infty} \infty$
- (4)  $H(X', f|_{X'}) = H(X, f)$  for every sub-array  $X'$  of  $X$ .

(1)  $\Rightarrow$  (2): Assume that  $f$  is hereditary and  $L_{\inf}(\xi) := \liminf D_N(\xi) < \infty$ . We'll show that  $L_{\sup}(\xi) := \limsup D_N(\xi) < \infty$ . This is obvious for  $\xi = 0$ , so suppose  $\xi \neq 0$ .

Choose  $N_\ell, M_\ell \uparrow \infty$  such that  $D_{N_\ell}(\xi) \xrightarrow{\ell \rightarrow \infty} L_{\inf}(\xi)$ ,  $D_{M_\ell}(\xi) \xrightarrow{\ell \rightarrow \infty} L_{\sup}(\xi)$ . Let

$$X' := \{X_k^{(N_\ell)}\} \text{ and } X'' := \{X_k^{(M_\ell)}\}.$$

Since  $L_{\inf}(\xi) < \infty$ ,  $H(X', f|_{X'})$  contains  $\xi$ , whence by (3.3.8),  $G_{ess}(X', f|_{X'}) \subseteq \frac{2\pi}{\xi} \mathbb{Z}$ . By the hereditary property,  $G_{ess}(X'', f|_{X''}) = G_{ess}(X, f) = G_{ess}(X', f|_{X'}) \subseteq \frac{2\pi}{\xi} \mathbb{Z}$ . This implies by (3.3.8) that  $H(X'', f|_{X''}) \ni \xi$ , whence  $L_{\sup}(\xi) < \infty$ .

(2)  $\Rightarrow$  (3): We assume that  $L_{\inf}(\xi) < \infty \Rightarrow L_{\sup}(\xi) < \infty$  and show that  $D_N(\xi) \rightarrow \infty$  for all  $\xi \notin H(X, f)$ . If  $\xi \notin H(X, f)$ , then  $\sup_N D_N(\xi) = \infty$ , so  $L_{\sup}(\xi) = \infty$ . By assumption, this forces  $L_{\inf}(\xi) = \infty$ , whence  $D_N(\xi) \rightarrow \infty$ .

(3)  $\Rightarrow$  (4): We assume that  $D_N(\xi) \rightarrow \infty$  for all  $\xi \notin H(X, f)$ , and show that  $H(X, f) = H(X', f|_{X'})$  for all sub-arrays  $X' = \{X_n^{(N_\ell)}\}$ . If  $\xi \in H(X, f)$ , then  $\sup_N D_N(\xi) < \infty$ , whence  $\sup_\ell D_{N_\ell}(\xi) < \infty$  and  $\xi \in H(X', f|_{X'})$ . If  $\xi \notin H(X, f)$ , then  $D_N(\xi) \rightarrow \infty$ , whence  $D_{N_\ell}(\xi) \rightarrow \infty$  and  $\xi \notin H(X', f|_{X'})$ .

(4)  $\Rightarrow$  (1): We assume that  $H(X', f|_{X'}) = H(X, f)$  for all sub-arrays  $X'$ , and show that  $G_{ess}(X', f|_{X'}) = G_{ess}(X, f)$  for all sub-arrays. The inclusion  $G_{ess}(X', f|_{X'}) \subseteq G_{ess}(X, f)$  is obvious, so we focus on  $G_{ess}(X', f|_{X'}) \supseteq G_{ess}(X, f)$ .

If  $G_{ess}(X', f|_{X'}) = \mathbb{R}$  then there is nothing to prove.

Suppose  $G_{ess}(X', f|_{X'}) \neq \mathbb{R}$ , then  $G_{ess}(X', f|_{X'}) = t\mathbb{Z}$  for some  $t \in \mathbb{R}$ . Let  $\xi := 2\pi/t$  when  $t \neq 0$  or any real number otherwise. By (3.3.8),

$$H(X', f|_{X'}) \ni \xi.$$

By assumption (4), this implies that  $H(X, f) \ni \xi$ , whence by (3.3.8),  $G_{ess}(X, f) \subseteq \frac{2\pi}{\xi} \mathbb{Z} = G_{ess}(X', f|_{X'})$ , and the proof of (1) is complete.

This finishes the proof that properties (1)–(4) are equivalent.

The second part of the theorem asks to show that  $f$  is *stably* hereditary iff  $D_N(\xi) \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(X, f)$ .

Suppose  $f$  is stably hereditary, then  $f$  is hereditary, whence  $D_N(\xi) \rightarrow \infty$  for all  $\xi \notin H(X, f)$ . To show that the convergence is uniform on compacts, we check that

$$\forall \xi \notin H(X, f), \forall M > 0, \exists N_\xi, \delta_\xi > 0 \left( \frac{N > N_\xi}{|\xi' - \xi| < \delta_\xi} \Rightarrow D_N(\xi') > M \right). \quad (3.3.9)$$

Suppose this were false for some  $\xi$  and  $M$ , then  $\exists \xi_N \rightarrow \xi$  such that  $D_N(\xi_N) \leq M$ . But this implies that  $\{(1 + \varepsilon_N) f_k^{(N)}\}$  is not hereditary for  $\varepsilon_N := \frac{\xi_N}{\xi} - 1$ , in contradiction to our assumptions.

Conversely, if  $D_N(\xi) \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(X, f)$ , and  $\varepsilon_N \rightarrow 0$ , then  $\{g_k^{(N)}\} = \{(1 + \varepsilon_N) f_k^{(N)}\}$  is hereditary, because for all  $\xi \notin H(X, f)$ ,  $D_N(\xi, g) \equiv D_N((1 + \varepsilon_N)\xi, f) \rightarrow \infty$ , and as we saw above (2)  $\Rightarrow$  (1).  $\square$

**Proof of Theorem 3.7:** The first part of the theorem assumes that  $G_{ess}(X, f) = t\mathbb{Z}$  or  $\{0\}$  and that  $f$  is hereditary, and asks to show that  $f$  is stably hereditary.

We begin with several reductions. It is sufficient to consider the case  $G_{ess}(X, f) = \mathbb{Z}$ : If  $G_{ess}(X, f) = t\mathbb{Z}$  with  $t \neq 0$  we work with  $t^{-1}f$ , and if  $G_{ess}(X, f) = \{0\}$  then  $H(X, f) = \mathbb{R}$  and  $D_N(\xi) \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(X, f)$  (vacuously), so  $f$  is stably hereditary by Theorem 3.6.

Next we claim that it is enough to treat the special case  $G_{alg}(X, f) = G_{ess}(X, f) = \mathbb{Z}$ . Otherwise we use Theorem 3.3 to write  $f = g - h$  where  $G_{alg}(X, g) = G_{ess}(X, g) = G_{ess}(X, f)$  and  $h$  is center-tight. By Lemma 3.15,  $H(X, g) = H(X, f)$ , and by Lemma 1.15 and Corollary 2.8,

$$D_N(\xi, f) \geq \frac{1}{8} D_N(\xi, g) - \frac{1}{8} \xi^2 \sup_n \sum_{k=3}^{k_n+1} u_k^{(n)}(h)^2 = \frac{1}{8} D_N(\xi, g) - O(1).$$

Thus, if  $D_N(\xi, g) \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(X, g)$ , then  $D_N(\xi, f) \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(X, f)$ .

By assumption,  $\text{ess sup } |f| \leq K$  for some integer  $K$ . Then for every hexagon  $P \in \text{Hex}(N, n)$ ,  $\Gamma(P) \in \mathbb{Z} \cap [-6K, 6K]$ .

Let  $m_n^{(N)}$  denote the probability measure on the space of hexagons  $\text{Hex}(N, n)$  and define for every  $\gamma \in \mathbb{Z} \cap [-6K, 6K]$ ,

$$\mu_N(\{\gamma\}) := \sum_{n=3}^{k_N} m_n^{(N)} \{P \in \text{Hex}(N, n) : \Gamma(P) = \gamma\}.$$

Using the identity  $|e^{i\xi\gamma} - 1|^2 = 4 \sin^2 \frac{\xi\gamma}{2}$ , we see that

$$d_N^2(\xi) = 4 \sum_{\gamma=-6K}^{6K} \mu_N(\gamma) \sin^2 \frac{\xi\gamma}{2}.$$

Since  $f$  is hereditary,  $D_N \rightarrow \infty$  on  $\mathbb{R} \setminus H(X, f)$ , and the expression for  $d_N^2(\xi)$  shows that if  $D_N \rightarrow \infty$  at  $\xi$ , then  $D_N \rightarrow \infty$  uniformly on an open neighborhood of  $\xi$ .

It follows that  $D_N \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{R} \setminus H(X, f)$ . By Theorem 3.6,  $f$  must be stably hereditary. This is the first part of the theorem.

The second part of the theorem says that if  $f$  is integer valued and not center-tight, and if  $\text{ess sup } |f| \leq K$ , then  $G_{ess}(X, f) = k\mathbb{Z}$  for some integer  $0 < k \leq 12K$ .

To see this recall that  $G_{ess}(X, f) \subset G_{alg}(X, f) \subset \mathbb{Z}$ , whence  $G_{ess}(X, f) = k\mathbb{Z}$  for some  $k \in \mathbb{Z}$ . Since  $f$  is not center-tight,  $k \neq 0$ . By (3.3.8),  $H(X, f) = \frac{2\pi}{k} \mathbb{Z}$ .

The inequality  $|f| \leq K$  implies that every hexagon  $P$  has balance  $|\Gamma(P)| \leq 6K$ . This implies that  $k \leq 12K$ : Otherwise  $|\frac{2\pi\Gamma(P)}{k}| < 0.95\pi$  and (3.3.2) gives

$$|e^{(2\pi i/k)\Gamma(P)} - 1|^2 \geq \text{const } \Gamma(P)^2.$$

But this implies that  $d_n^{(N)}(\frac{2\pi}{k}) \geq \text{const } u_n^{(N)}$ , whence

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)} \left(\frac{2\pi}{k}\right)^2 \geq \sup_N \sum_{n=3}^{k_N} (u_n^{(N)})^2 = \infty \text{ by non-center-tightness.}$$

This contradicts  $\frac{2\pi}{k} \in H(X, f)$ . Thus  $0 < k \leq 12K$ .

It follows from the first part of the theorem and from Theorem 3.6, that if  $f$  is integer valued and not center-tight, then the properties of being hereditary and of being stably hereditary are equivalent.  $\square$

### 3.4 Notes and references

In the stationary world, a center-tight cocycle is a coboundary (Schmidt [136]) and the problems discussed in this chapter reduce to the question how small can one make the range of a cocycle by subtracting from it a coboundary. The question appears naturally in the ergodic theory of group actions, because of its relation to the ergodic decomposition of skew-products [2, chapter 8], [136], [29], and to the structure of locally finite ergodic invariant measures for skew-products [6], [133], [124]. In the general setup of ergodic theory, minimal reductions such as in Theorem 3.3 are not always possible [93], although they do sometime exist [133], [124].

The relevance of (ir)reducibility to the local limit theorem appears in different form in the papers of Guivarc'h & Hardy [65], Aaronson & Denker [4], and Dolgopyat [49]. There “irreducibility” is expressed in terms of a condition which rules out non-trivial solutions for certain cohomological equations.

It is more difficult to uncover the irreducibility condition in the probabilistic literature on the LLT for sums of independent random variables. Rozanov's paper [128], for example, proves a LLT for independent  $\mathbb{Z}$ -valued random variables  $X_k$  assuming Lindeberg's condition (which is automatic for bounded random variables),  $\sum \text{Var}(X_k) = \infty$ , and subject to the assumption that

$$\prod_{k=1}^{\infty} \left( \max_{0 \leq m < t} \mathbb{P}(X_k = m \bmod t) \right) = 0 \text{ for all integers } t \geq 2. \quad (3.4.1)$$

Let  $X = \{X_k\}$  and  $f = \{f_k\}$  where  $f_k(x) = x$ . Clearly, (3.4.1) implies that  $G_{\text{alg}}(X, f) = \mathbb{Z}$ . We claim that (3.4.1) is equivalent to the irreducibility:  $G_{\text{ess}}(X, f) = \mathbb{Z}$ .

To see why, it is useful first to note that (3.4.1) is equivalent to

$$\sum_k \mathbb{P}[X_k \neq m_k \bmod t] = \infty \quad (3.4.2)$$

where  $m_k$  is the (smallest) most likely residue mod  $t$  for  $X_k$ .

**Irreducibility  $\Rightarrow$  Rozanov's condition:** Define for  $x \in \mathbb{Z}$  and  $2 \leq t \in \mathbb{Z}$ ,  $\{x\}_{t\mathbb{Z}} := t\{x/t\}$ ,  $[x]_{t\mathbb{Z}} := x - \{x\}_{t\mathbb{Z}}$ , and set

- $y_k(x) :=$  the (smallest) integer in  $m_k + t\mathbb{Z}$  closest to  $x$
- $z_k(x) := x - y_k(x)$
- $g_k(x) := (y_k(x) - m_k) + [x - y_k(x)]_{t\mathbb{Z}}$  ( $g_k$  takes values in  $t\mathbb{Z}$ )
- $h_k(x) := \{x - y_k(x)\}_{t\mathbb{Z}}$  ( $h_k$  takes values in  $\mathbb{Z}$ ). Then

$$X_k = g_k(X_k) + h_k(X_k) + m_k.$$

The algebraic range of  $g_k$  is inside  $t\mathbb{Z}$ , and by the Borel-Cantelli Lemma,

$$(3.4.2) \text{ fails} \Leftrightarrow X_k \neq m_k \pmod{t\mathbb{Z}} \text{ finitely often a.s.} \Leftrightarrow h_k(X_k) \neq 0 \text{ finitely often a.s.}$$

If (3.4.2) fails, then  $\sum_{k=0}^{\infty} h_k(X_k)$  converges a.s. (since a.s. there are only finitely non-zero terms).

Hence  $h$  is center-tight. Since  $G_{alg}(g) \subset t\mathbb{Z}$ , we have a contradiction to irreducibility.

**Rozanov's condition  $\Rightarrow$  irreducibility:** Fix  $\theta \in [0, t)$  and let  $m$  be the closest integer in  $[0, t) \cap \mathbb{Z}$  to  $\theta$ . Then  $|m' - \theta| \geq \frac{1}{2}$  for  $m' \neq m$ , whence

$$\mathbb{E}[\text{dist}^2(X_n, \theta + t\mathbb{Z})] \geq \frac{1}{4}\mathbb{P}(X_n \neq m \pmod{t}) \geq \frac{1}{4}[1 - \max_{0 \leq m < t} \mathbb{P}(X_n = m \pmod{t})].$$

Passing to the infimum over  $\theta$ , we obtain that

$$\mathfrak{D}^2(X_n, \frac{2\pi}{t}) \geq \frac{1}{4}[1 - \max_{0 \leq m < t} \mathbb{P}(X_n = m \pmod{t})].$$

(See §1.3.) We now obtain from Proposition 1.19 that

$$\begin{aligned} \sum_{n=3}^{\infty} d_n^2(\frac{2\pi}{t}) &\geq \text{const} \sum_{n=3}^{\infty} (\mathfrak{D}^2(X_{n-1}, \frac{2\pi}{t}) + \mathfrak{D}^2(X_n, \frac{2\pi}{t})) \\ &\geq \text{const} \sum_{n=2}^{\infty} \left(1 - \max_{0 \leq m < t} \mathbb{P}(X_k = m \pmod{t})\right) = \infty, \text{ by (3.4.2).} \end{aligned}$$

We find that the co-range does *not* contain  $2\pi/t$  for  $t = 2, 3, 4, \dots$ . We already know that the co-range does contain  $2\pi$  (because  $X_k$  are integer valued). The only closed sub-group of  $\mathbb{R}$  with these properties is  $2\pi\mathbb{Z}$ . So the co-range is  $2\pi\mathbb{Z}$ , and the essential range is  $\mathbb{Z}$  = the algebraic range

Other sufficient conditions for the LLT for sums of independent random variables such as those appearing in [104], [146] and [108] can be analyzed in a similar way. The reduction lemma was proved for sums of independent random variables in [49]. A version of Theorem 3.5 for sums of independent random variables appears in [108].

## Chapter 4

### The local limit theorem in the irreducible case

In this chapter we prove the local limit theorem for  $\mathbb{P}(S_N - z_N \in (a, b))$  when  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}}$  converges to a finite limit and  $f$  is irreducible. In this regime, the asymptotic behavior of  $\mathbb{P}(S_N - z_N \in (a, b))$  does not depend on the details of  $X$  and  $f$  (“universality”).

#### 4.1 Main results

##### 4.1.1 Local limit theorems for Markov chains

In the next two theorems, we assume that  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov **chain**  $X$ , and we let  $X = \{X_n\}$ ,  $f = \{f_n\}$ ,  $S_N = f_1(X_1, X_2) + \dots + f_N(X_N, X_{N+1})$ , and  $V_N := \text{Var}(S_N)$ . We make no assumptions on the initial distribution and allow  $\mathbb{P} = \mathbb{P}_x = \mathbb{P}(\cdot | X_1 = x)$ .

**Theorem 4.1.** Suppose  $f$  is irreducible, with algebraic range  $\mathbb{R}$ . Then  $V_N \rightarrow \infty$ , and for every interval  $(a, b)$  and  $z_N \in \mathbb{R}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}$  converges to a finite limit  $z$ ,

$$\mathbb{P}[S_N - z_N \in (a, b)] = [1 + o(1)] \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} (b - a), \text{ as } N \rightarrow \infty. \quad (4.1.1)$$

**Theorem 4.2.** Suppose  $t > 0$  and  $f$  is irreducible with algebraic range  $t\mathbb{Z}$ . Then  $V_N \rightarrow \infty$  and there are constants  $0 \leq \gamma_N < t$  such that for all  $k \in \mathbb{Z}$ , and for all  $z_N \in \gamma_N + t\mathbb{Z}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}$  converges to a finite limit  $z$ ,

$$\mathbb{P}[S_N - z_N = kt] = [1 + o(1)] \frac{e^{-z^2/2} t}{\sqrt{2\pi V_N}}, \quad \text{as } N \rightarrow \infty. \quad (4.1.2)$$

The constants  $\gamma_N$  are determined by the condition  $\mathbb{P}[S_N \in \gamma_N + t\mathbb{Z}] = 1$  for all  $N$ .

The conditions of the theorems can be checked from the data of  $X$  and  $f$  using the structure constants  $d_n(\xi)$  from §1.3:

**Lemma 4.1** Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . Then

- (1)  $f$  is non-lattice and irreducible iff  $\sum d_n^2(\xi) = \infty$  for all  $\xi \neq 0$ .
- (2)  $f$  is lattice and irreducible with algebraic range  $t\mathbb{Z}$ ,  $t > 0$ , iff  $\sum d_n^2(\xi) < \infty$  for  $\xi \in (2\pi/t)\mathbb{Z}$  and  $\sum d_n^2(\xi) = \infty$  for  $\xi \notin (2\pi/t)\mathbb{Z}$ .

(3)  $f$  is lattice and irreducible with algebraic range  $\{0\}$  iff  $f_n(X_n, X_{n+1})$  are a.s. constant for all  $n$ .

*Proof.*  $f$  is non-lattice and irreducible iff  $G_{ess}(X, f) = G_{alg}(X, f) = \mathbb{R}$ . By Theorem 3.1, this happens iff  $f$  has co-range  $\{0\}$ , which proves part (1). Part (2) is proved in a similar way, and part (3) is a triviality.  $\square$

#### 4.1.2 Local limit theorems for Markov arrays

In this section, we assume that  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$  with row lengths  $k_N + 1$ , and we let  $X = \{X_n^{(N)}\}$ ,  $f = \{f_n^{(N)}\}$ ,  $S_N = \sum_{i=1}^{k_N} f_i^{(N)}(X_i^{(N)}, X_{i+1}^{(N)})$ , and  $V_N := \text{Var}(S_N)$ . We make no assumptions on the initial distribution, and allow  $\mathbb{P} = \mathbb{P}_{x_1^{(N)}} = \mathbb{P}(\cdot | X_1^{(N)} = x_1^{(N)})$ .

The LLT for  $S_N$  may fail due to the possibility that  $f|_{X'}$  may have different essential range for different sub-arrays  $X'$ . To deal with this we need to assume hereditary behavior, see §3.2.3.

**Theorem 4.1'.** *Suppose  $f$  is stably hereditary, non-lattice and irreducible. Then  $V_N \rightarrow \infty$ , and for every interval  $(a, b)$  and  $z_N \in \mathbb{R}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N \rightarrow \infty]{} z \in \mathbb{R}$ ,*

$$\mathbb{P}[S_N - z_N \in (a, b)] = [1 + o(1)] \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} (b - a), \text{ as } N \rightarrow \infty. \quad (4.1.3)$$

**Theorem 4.2'.** *Suppose  $t > 0$  and  $f$  is hereditary, irreducible, and with algebraic range  $t\mathbb{Z}$ . Then  $V_N \rightarrow \infty$ , and there are  $0 \leq \gamma_N < t$  such that for all  $k \in \mathbb{Z}$  and  $z_N \in \gamma_N + t\mathbb{Z}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N \rightarrow \infty]{} z \in \mathbb{R}$ ,*

$$\mathbb{P}[S_N - z_N = kt] = [1 + o(1)] \frac{e^{-z^2/2} t}{\sqrt{2\pi V_N}}, \quad \text{as } N \rightarrow \infty. \quad (4.1.4)$$

*The constants  $\gamma_N$  are determined by the condition  $\mathbb{P}[S_N \in \gamma_N + t\mathbb{Z}] = 1$  for all  $N$ .*

Notice that whereas in the non-lattice case we had to assume that  $f$  is stably hereditary, in the lattice case it is sufficient to assume that  $f$  is hereditary. This is because in the lattice case the two assumptions are equivalent, see Theorem 3.7.

Again, it is possible to check the assumptions of the theorems from the data of  $X$  and  $f$  using the structure constants:

**Lemma 4.1'.** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array  $X$  with row lengths  $k_N + 1$ . Let  $d_n^{(N)}(\xi)$  be as §1.3, then*

(1)  $f$  is stably hereditary, irreducible, and with algebraic range  $\mathbb{R}$  iff

$$\sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 \xrightarrow[N \rightarrow \infty]{} \infty \text{ uniformly on compacts in } \mathbb{R} \setminus \{0\}.$$

(2) Suppose  $t \neq 0$ , then  $f$  is hereditary and irreducible with algebraic range  $t\mathbb{Z}$  if and only if  $\sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 \xrightarrow[N \rightarrow \infty]{} \infty$  for all  $\xi \notin \frac{2\pi}{t}\mathbb{Z}$ . In this case  $f$  is also stably hereditary.

*Proof.* As in the case of Markov chains,  $f$  is non-lattice and irreducible iff its co-range equals  $\{0\}$ . By Theorem 3.6,  $f$  is stably hereditary iff  $\sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 \xrightarrow[N \rightarrow \infty]{} \infty$  uniformly on compacts in  $\mathbb{R} \setminus \{0\}$ , which proves part (1).

Part (2) is proved in a similar way, with the additional observation that thanks to Theorem 3.7, in the irreducible lattice case, every hereditary additive functional is automatically stably hereditary.  $\square$

#### 4.1.3 Mixing local limit theorems

Let  $f$  be an additive functional on a Markov  $X$  with row lengths  $k_N + 1$ , and state spaces  $(\mathfrak{S}_n^{(N)}, \mathcal{B}(\mathfrak{S}_n^{(N)}))$ . Let  $S_N$  and  $V_N$  be as in the previous section.

**Theorem 4.3 (Mixing LLT).** *Suppose  $X$  is a uniformly elliptic Markov array, and  $f$  is an additive functional on  $X$  which is stably hereditary, a.s. uniformly bounded, and irreducible. Let  $\mathfrak{A}_N \subset \mathfrak{S}_{k_N+1}^{(N)}$  be measurable events such that  $\mathbb{P}[X_{k_N+1}^{(N)} \in \mathfrak{A}_N]$  is bounded away from zero, and let  $x_N \in \mathfrak{S}_1^{(N)}$ . Then for every  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  continuous with compact support,*

(1) **Non-lattice case:** *Suppose  $f$  has algebraic range  $\mathbb{R}$ . For every  $z_N \in \mathbb{R}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z \in \mathbb{R}$ ,*

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \mathbb{E}[\phi(S_N - z_N) | X_{k_N+1}^{(N)} \in \mathfrak{A}_N, X_1^{(N)} = x_N] = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(u) du.$$

(2) **Lattice case:** *Suppose  $f$  has algebraic range  $t\mathbb{Z}$  ( $t > 0$ ) and  $\mathbb{P}[S_N \in \gamma_N + t\mathbb{Z}] = 1$  for all  $N$ .*

*For every  $z_N \in \gamma_N + t\mathbb{Z}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z \in \mathbb{R}$ ,*

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \mathbb{E}[\phi(S_N - z_N) | X_{k_N+1}^{(N)} \in \mathfrak{A}_N, X_1^{(N)} = x_N] = \frac{e^{-z^2/2}|t|}{\sqrt{2\pi}} \sum_{u \in \mathbb{Z}} \phi(tu).$$

To understand what this means, think of  $\phi \approx 1_{(a,b)}$ .

In the next chapter, we will use mixing LLT for *irreducible* additive functionals to study the LLT for some *reducible* additive functionals, as follows. Suppose  $f = \tilde{f} + \nabla h$ , where  $\tilde{f}$  is irreducible and  $h$  is uniformly bounded. Then

$$S_N(f) = S_N(\tilde{f}) + h_1^{(N)}(X_1^{(N)}) - h_{k_N+1}^{(N)}(X_{k_N+1}^{(N)}).$$

To pass from the LLT for  $S_N(\tilde{f})$  (which we know since  $\tilde{f}$  is irreducible) to the LLT for  $S_N(f)$  (which we do not know because of the reducibility of  $f$ ), we need to understand the *joint* distribution of  $S_N(\tilde{f})$ ,  $h_1^{(N)}(X_1^{(N)})$  and  $h_{k_N+1}^{(N)}(X_{k_N+1}^{(N)})$ . This is the task achieved by the mixing LLT.

## 4.2 Proofs

We will provide the proofs in the general context of Markov arrays.

### Standing assumptions and notation for the remainder of the chapter:

$X = \{X_n^{(N)}\}$  is a Markov array with row lengths  $k_N + 1$ , state spaces  $\mathfrak{S}_n^{(N)}$ , and transition proba-

bilities  $\pi_{n,n+1}^{(N)}(x, dy)$ , and  $f = \{f_n^{(N)}\}$  is an additive functional on  $X$ . As always,  $d_n^{(N)}(\xi)$  are the structure constants of  $f$ .

We assume that  $\text{ess sup } |f| < K < \infty$ , and that  $X$  is uniformly elliptic with ellipticity constant  $\varepsilon_0$ . By the uniform ellipticity assumption,

$$\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y) \mu_{n+1}^{(N)}(dy)$$

with  $0 \leq p_n^{(N)}(x, y) < \varepsilon_0^{-1}$  such that  $\int p_n^{(N)}(x, y) p_{n+1}^{(N)}(y, z) \mu_{n+1}^{(N)}(dy) > \varepsilon_0$ . There is no loss of generality in assuming that  $\mu_k^{(N)}(E) = \mathbb{P}(X_k^{(N)} \in E)$ , see Proposition 1.12 and the discussion which follows it.

#### 4.2.1 Characteristic functions

The classical approach to limit theorems in probability theory, due to P. Lévy, is to apply the Fourier transform, and analyze the **characteristic functions** of the random variables in the problem. In our case the relevant characteristic functions are:

$$\begin{aligned} \Phi_N(x, \xi) &:= \mathbb{E}_x \left( e^{i\xi S_N} \right) \equiv \mathbb{E} \left( e^{i\xi S_N} \middle| X_1^{(N)} = x \right). \\ \Phi_N(x, \xi | \mathfrak{A}) &:= \mathbb{E}_x \left( e^{i\xi S_N} \middle| X_{k_N+1}^{(N)} \in \mathfrak{A} \right) \equiv \mathbb{E} \left( e^{i\xi S_N} \middle| X_{k_N+1}^{(N)} \in \mathfrak{A}, X_1^{(N)} = x \right). \end{aligned}$$

Here  $x \in \mathfrak{S}_1^{(N)}$ ,  $\mathfrak{A} \subset \mathfrak{S}_{k_N+1}^{(N)}$ ,  $\xi \in \mathbb{R}$ , and  $\mathbb{E}_x(\cdot) = \mathbb{E}(\cdot | X_1^{(N)} = x)$ .

We write these functions in terms of **perturbation operators** as in [109]. For every  $N \in \mathbb{N}$  and  $1 \leq n \leq k_N + 1$ , define  $\mathcal{L}_{n,\xi}^{(N)} : L^\infty(\mathfrak{S}_{n+1}^{(N)}) \rightarrow L^\infty(\mathfrak{S}_n^{(N)})$  by

$$\begin{aligned} \left( \mathcal{L}_{n,\xi}^{(N)} v \right) (x) &:= \int_{\mathfrak{S}_{n+1}^{(N)}} p_n^{(N)}(x, y) e^{i\xi f_n^{(N)}(x, y)} v(y) d\mu_{n+1}^{(N)}(y) \\ &\equiv \mathbb{E} \left( e^{i\xi f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})} v(X_{n+1}^{(N)}) \middle| X_n^{(N)} = x \right). \end{aligned}$$

**Lemma 4.2 (Nagaev)** *Let  $1(\cdot) \equiv 1$ , then the following identities hold:*

$$\mathbb{E} \left( e^{i\xi S_N} v(X_{k_N+1}^{(N)}) \middle| X_1^{(N)} = x \right) = \left( \mathcal{L}_{1,\xi}^{(N)} \mathcal{L}_{2,\xi}^{(N)} \dots \mathcal{L}_{k_N,\xi}^{(N)} v \right) (x), \quad (4.2.1)$$

$$\Phi_N(x, \xi) = \left( \mathcal{L}_{1,\xi}^{(N)} \mathcal{L}_{2,\xi}^{(N)} \dots \mathcal{L}_{k_N,\xi}^{(N)} 1 \right) (x), \quad (4.2.2)$$

$$\Phi_N(x, \xi | \mathfrak{A}) = \frac{\left( \mathcal{L}_{1,\xi}^{(N)} \mathcal{L}_{2,\xi}^{(N)} \dots \mathcal{L}_{N,\xi}^{(N)} 1_{\mathfrak{A}} \right) (x)}{\mathbb{P}_x[X_{k_N+1}^{(N)} \in \mathfrak{A}]} \quad (4.2.3)$$

*Proof.*  $\mathbb{E}(e^{i\xi S_N} v(X_{k_N+1}^{(N)}) | X_1^{(N)} = x) =$

$$\int p_1^{(N)}(x, y) e^{i\xi f_1^{(N)}(x, y)} \mathbb{E} \left( e^{i\xi \sum_{n=2}^N f_n^{(N)}} v | X_2^{(N)} = y \right) d\mu_2^{(N)}(y).$$

Proceeding by induction, we obtain (4.2.1), and (4.2.1) implies (4.2.2), (4.2.3).  $\square$

Let  $\|\cdot\|$  denote the operator norm on  $\text{Hom}(L^\infty, L^\infty)$ .

**Lemma 4.3**  $\mathcal{L}_{n,\xi}^{(N)}$  are bounded linear operators, and there is a positive constant  $\tilde{\varepsilon}$  which only depends on  $\varepsilon_0$  such that for all  $N \geq 1$  and  $5 \leq n \leq k_N$ ,  $\|\mathcal{L}_{n,\xi}^{(N)}\| \leq 1$ , and

$$\left\| \mathcal{L}_{n-4,\xi}^{(N)} \mathcal{L}_{n-3,\xi}^{(N)} \mathcal{L}_{n-2,\xi}^{(N)} \mathcal{L}_{n-1,\xi}^{(N)} \mathcal{L}_{n,\xi}^{(N)} \right\| \leq e^{-\tilde{\varepsilon} d_n^{(N)}(\xi)^2}.$$

*Proof.* Throughout this proof we fix  $N$  and drop the superscripts  $(N)$ , and we use the notation  $x_i, z_i$  etc. to denote points in  $\mathfrak{S}_i = \mathfrak{S}_i^{(N)}$ .

It is clear that  $\|\mathcal{L}_{n,\xi}^{(N)}\| \leq 1$ . To estimate the norm of

$$\mathcal{L} := \mathcal{L}_{n-4,\xi} \mathcal{L}_{n-3,\xi} \mathcal{L}_{n-2,\xi}^{(N)} \mathcal{L}_{n-1,\xi}^{(N)} \mathcal{L}_{n,\xi} : L^\infty(\mathfrak{S}_{n+1}) \rightarrow L^\infty(\mathfrak{S}_{n-4}),$$

we represent this operator as an integral operator, and analyze the kernel. Let

- $p(x_k, \dots, x_m) := \prod_{i=k}^{m-1} p_i(x_i, x_{i+1})$ ,
- $f(x_k, \dots, x_m) := \sum_{i=k}^{m-1} f_i(x_i, x_{i+1})$ ,
- $L(x_{n-4}, z_{n+1}) = \int_{\mathfrak{S}_{n-3} \times \dots \times \mathfrak{S}_n} p(x_{n-4}, z_{n-3}, \dots, z_{n+1}) e^{i\xi f(x_{n-4}, z_{n-3}, \dots, z_{n+1})} \mu_{n-3}(dz_{n-3}) \dots \mu_n(dz_n).$

Then  $(\mathcal{L}v)(x_{n-4}) = \int_{\mathfrak{S}_{n+1}} \left[ L(x_{n-4}, z_{n+1}) v(z_{n+1}) \right] \mu_{n+1}(dz_{n+1})$ , whence

$$\|\mathcal{L}v\|_\infty \leq \|v\|_\infty \sup_{x_{n-4} \in \mathfrak{S}_{n-4}} \int_{\mathfrak{S}_{n+1}} |L(x_{n-4}, z_{n+1})| \mu_{n+1}(dz_{n+1}).$$

To estimate this integral we change the order of integration:

$$\begin{aligned} \int_{\mathfrak{S}_{n+1}} |L(x_{n-4}, z_{n+1})| \mu_{n+1}(dz_{n+1}) &\leq \iint_{\mathfrak{S}_{n-2} \times \mathfrak{S}_{n+1}} \left[ |K_n(z_{n-2}, z_{n+1})| \right. \\ &\quad \left. \int_{\mathfrak{S}_{n-3}} p(x_{n-4}, z_{n-3}, z_{n-2}) \mu_{n-3}(dz_{n-3}) \right] \mu_{n-2}(dz_{n-2}) \mu_{n+1}(dz_{n+1}), \end{aligned} \quad (4.2.4)$$

where  $K_n(z_{n-2}, z_{n+1}) :=$

$$\iint_{\mathfrak{S}_{n-1} \times \mathfrak{S}_n} p(z_{n-2}, z_{n-1}, z_n, z_{n+1}) e^{i\xi f(z_{n-2}, z_{n-1}, z_n, z_{n+1})} \mu_{n-1}(dz_{n-1}) \mu_n(dz_n).$$

CLAIM: Let  $p(z_{n-2} \rightarrow z_{n+1}) := \mathbb{P}(X_{n+1} = z_{n+1} | X_{n-2} = z_{n-2})$ , then

$$\begin{aligned} |K_n(z_{n-2}, z_{n+1})| &\leq p(z_{n-2} \rightarrow z_{n+1}) - \\ &\quad - \frac{1}{4} p(z_{n-2} \rightarrow z_{n+1}) \mathbb{E} \left( |e^{i\xi \Gamma(P)} - 1|^2 \middle| \begin{array}{l} X_{n-2} = Y_{n-2} = z_{n-2} \\ X_{n+1} = Z_{n+1} = z_{n+1} \end{array} \right) \end{aligned} \quad (4.2.5)$$

*Proof of the claim.* Set  $\tilde{K}_n(z_{n-2}, z_{n+1}) := \frac{K_n(z_{n-2}, z_{n+1})}{p(z_{n-2} \rightarrow z_{n+1})}$ , then

$$\tilde{K}_n(z_{n-2}, z_{n+1}) = \mathbb{E} \left( e^{i\xi \sum_{k=n-2}^n f_k(X_k, X_{k+1})} \middle| \begin{array}{l} X_{n-2} = z_{n-2} \\ X_{n+1} = z_{n+1} \end{array} \right).$$

Writing  $|\tilde{K}_n(z_{n-2}, z_{n+1})|^2 = \tilde{K}_n(z_{n-2}, z_{n+1}) \overline{\tilde{K}_n(z_{n-2}, z_{n+1})}$ , we find that

$$|\tilde{K}_n(z_{n-2}, z_{n+1})|^2 = \mathbb{E} \left( e^{i\xi \Gamma(X_{n-2} \stackrel{X_{n-1}}{Y_{n-1}} X_{n+1})} \middle| \begin{array}{l} X_{n-2} = Y_{n-2} = z_{n-2} \\ X_{n+1} = Z_{n+1} = z_{n+1} \end{array} \right),$$

where  $\{Y_n\}$  is an independent copy of  $\{X_n\}$ , and  $\Gamma$  is as in (1.3.1).

The imaginary part is necessarily zero, so writing  $P = (X_{n-2} \stackrel{X_{n-1}}{Y_{n-1}} X_{n+1})$  we have by the identity  $1 - \cos \alpha = \frac{1}{2}|e^{i\alpha} - 1|^2$  that

$$\begin{aligned} |\tilde{K}_n(z_{n-2}, z_{n+1})|^2 &= 1 - \mathbb{E} \left( 1 - \cos(\xi \Gamma(P)) \middle| \stackrel{X_{n-2} = Y_{n-2} = z_{n-2}}{X_{n+1} = Z_{n+1} = z_{n+1}} \right) \\ &\equiv 1 - \frac{1}{2} \mathbb{E} \left( |e^{i\xi \Gamma(P)} - 1|^2 \middle| \stackrel{X_{n-2} = Y_{n-2} = z_{n-2}}{X_{n+1} = Z_{n+1} = z_{n+1}} \right). \end{aligned}$$

The claim follows, since  $\sqrt{1-t} \leq 1 - \frac{t}{2}$  for all  $0 \leq t \leq 1$ .

We now substitute (4.2.5) in (4.2.4). The result is a difference of two terms:

(a) The first term is obtained by replacing  $K_n(z_{n-2}, z_{n+1})$  in (4.2.4) by  $p(z_{n-2} \rightarrow z_{n+1})$ . It has the following *upper* bound:

$$\iint_{\mathfrak{S}_{n-2} \times \mathfrak{S}_{n+1}} \int_{\mathfrak{S}_{n-3}} p(x_{n-4}, z_{n-3}, z_{n-2}) p(z_{n-2} \rightarrow z_{n+1}) = 1.$$

(b) The second term is obtained by replacing  $K_n(z_{n-2}, z_{n+1})$  in (4.2.4) by

$$\frac{1}{4} p(z_{n-2} \rightarrow z_{n+1}) \mathbb{E} \left( |e^{i\xi \Gamma(P)} - 1|^2 \middle| \stackrel{X_{n-2} = Y_{n-2} = z_{n-2}}{X_{n+1} = Z_{n+1} = z_{n+1}} \right).$$

The inner-most integral satisfies  $\int_{\mathfrak{S}_{n-3}} p(x_{n-4}, z_{n-3}, z_{n-2}) \mu_{n-3}(dz_{n-3}) \geq \varepsilon_0$  because of uniform ellipticity. This leads to the following *lower* bound for the second term:

$$\frac{1}{4} \varepsilon_0^2 \mathbb{E} \left( |e^{i\xi \Gamma(P)} - 1|^2 \right) = \frac{1}{4} \varepsilon_0^2 d_n(\xi)^2.$$

In total we get:  $\int |L(x_{n-4}, z_{n+1})| \mu_{n+1}(dz_{n+1}) \leq 1 - \tilde{\varepsilon} d_n(\xi)^2$ , where  $\tilde{\varepsilon} := \frac{1}{4} \varepsilon_0^2$ . Since  $1 - t \leq e^{-t}$ , we are done.  $\square$

Recall that  $D_N(\xi) = \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2$ . Write  $D_N = \sum_{j=0}^4 D_{j,N}$  where

$$D_{j,N}(\xi) = \sum_{\substack{3 \leq n \leq k_N \\ n \equiv j \pmod{5}}} d_n^{(N)}(\xi)^2.$$

Applying Lemma 4.3 iteratively we conclude that there is a constant  $C$  independent of  $N$  s.t. for all  $N$ ,

$$|\Phi_N(x, \xi)| \leq Ce^{-\tilde{\varepsilon} \max(D_{0,N}, \dots, D_{4,N})} \leq Ce^{-\frac{1}{5}\tilde{\varepsilon} D_N(\xi)}. \quad (4.2.6)$$

If  $\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}) \geq \bar{\delta}$  then by (4.2.3),  $|\Phi_N(x, \xi | \mathfrak{A})| \leq \bar{\delta}^{-1} \|\mathcal{L}_{1,\xi}^{(N)} \mathcal{L}_{2,\xi}^{(N)} \dots \mathcal{L}_{k_N,\xi}^{(N)} 1_{\mathfrak{A}}\|$  whence

$$|\Phi_N(x, \xi | \mathfrak{A})| \leq Ce^{-\frac{1}{5}\tilde{\varepsilon} D_N(\xi)}. \quad (4.2.7)$$

The next result shows that if  $u_n^{(N)}$  is big, then  $d_n^{(N)}(\cdot)$  cannot be small at two nearby points. Recall the standing assumption  $\text{ess sup} \|f_n^{(N)}\|_\infty \leq K$ , and the definition of the structure constants  $u_n^{(N)}$  in (1.3.2).

**Lemma 4.4**  $\exists \tilde{\delta} = \tilde{\delta}(K) > 0$  s.t. if  $|\delta| \leq \tilde{\delta}$  then for all  $3 \leq n \leq k_N$ ,

$$d_n^{(N)}(\xi + \delta)^2 \geq \frac{2}{3} \delta^2 \left( u_n^{(N)} \right)^2 - 2|\delta| u_n^{(N)} d_n^{(N)}(\xi). \quad (4.2.8)$$

*Proof.* Fix a hexagon  $P = \begin{pmatrix} x_{n-1} & x_n \\ y_{n-1} & y_n \end{pmatrix} \in \text{Hex}(N, n)$ , and let

$$\mathfrak{u}_n := \Gamma(P), \quad \mathfrak{d}_n(\xi) := |e^{i\xi \mathfrak{u}_n} - 1|,$$

then the identity  $|e^{i\theta} - 1|^2 = 2(1 - \cos \theta)$  implies

$$\begin{aligned} \mathfrak{d}_n^2(\xi + \delta) &= |e^{i(\xi + \delta) \mathfrak{u}_n} - 1|^2 = 2[1 - \cos((\xi + \delta) \mathfrak{u}_n)] \\ &= 2[1 - \cos(\xi \mathfrak{u}_n) \cos(\delta \mathfrak{u}_n) + \sin(\xi \mathfrak{u}_n) \sin(\delta \mathfrak{u}_n)] \\ &= 2[(1 - \cos(\xi \mathfrak{u}_n)) \cos(\delta \mathfrak{u}_n) + (1 - \cos(\delta \mathfrak{u}_n)) \sin(\xi \mathfrak{u}_n) \sin(\delta \mathfrak{u}_n)] \\ &\geq 2[(1 - \cos(\delta \mathfrak{u}_n)) - |\sin(\xi \mathfrak{u}_n) \sin(\delta \mathfrak{u}_n)|] \text{ provided } |\tilde{\delta}| < \frac{\pi}{12K}, \end{aligned} \quad (4.2.9)$$

because in this case  $|\delta \mathfrak{u}_n| < \frac{\pi}{2}$ , so  $\cos(\delta \mathfrak{u}_n) \geq 0$ . Make  $\tilde{\delta}$  even smaller to guarantee  $0 \leq |t| \leq 6K\tilde{\delta} \Rightarrow \frac{1}{3}t^2 \leq 1 - \cos t \leq t^2$ , then

$$\begin{aligned} \mathfrak{d}_n^2(\xi + \delta) &\geq 2\left(\frac{1}{3}\delta^2 \mathfrak{u}_n^2 - |\delta \mathfrak{u}_n| \sqrt{1 - \cos^2(\xi \mathfrak{u}_n)}\right) \\ &= 2\left(\frac{1}{3}\delta^2 \mathfrak{u}_n^2 - |\delta \mathfrak{u}_n| \sqrt{(1 - \cos(\xi \mathfrak{u}_n))(1 + \cos(\xi \mathfrak{u}_n))}\right) \\ &\geq 2\left(\frac{1}{3}\delta^2 \mathfrak{u}_n^2 - |\delta \mathfrak{u}_n| \sqrt{2(1 - \cos(\xi \mathfrak{u}_n))}\right) = \frac{2}{3}\delta^2 \mathfrak{u}_n^2 - 2|\delta \mathfrak{u}_n| |e^{i\xi \mathfrak{u}_n} - 1| \\ &= \frac{2}{3}\delta^2 \mathfrak{u}_n^2 - 2|\delta \mathfrak{u}_n| \mathfrak{d}_n(\xi). \end{aligned}$$

Integrating on  $P \in \text{Hex}(N, n)$ , and using Cauchy-Schwarz to estimate the second term we obtain the lower bound for  $d_n(\xi + \delta)^2$ .  $\square$

Lemma 4.4 and the Cauchy-Schwarz inequality together give

$$D_N(\xi + \delta) \geq \frac{2}{3}\delta^2 U_N - 2|\delta| \sqrt{U_N D_N(\xi)} \quad (4.2.10)$$

where  $U_n := \sum_{k=3}^{k_N} (u_k^{(N)})^2$ . If  $V_N := \text{Var}(S_N) \rightarrow \infty$ , then as soon as  $V_N > 2C_2$  where  $C_2$  is the constant from Theorem 2.1, we have

$$\frac{U_N}{2C_1} \leq V_N \leq 2C_1 U_N. \quad (4.2.11)$$

So there are  $\widehat{\varepsilon}_1, \widehat{c}_1 > 0$  s.t.  $D_N(\xi + \delta) \geq \widehat{\varepsilon}_1 \delta^2 V_N - \widehat{c}_1 |\delta| \sqrt{V_N D_N(\xi)}$ . By (4.2.6), there are  $\widehat{\varepsilon}, \widehat{c} > 0$  s.t. for all  $N$  so large that  $V_N > 2C_2$ , for all  $\xi$  and  $|\delta| < \widetilde{\delta}$

$$|\Phi_N(x, \xi + \delta)| \leq C \exp \left( -\widehat{\varepsilon} V_N \delta^2 + \widehat{c} |\delta| \sqrt{V_N D_N(\xi)} \right). \quad (4.2.12)$$

We rephrase (4.2.12) as follows. Given a compact interval  $I \subset \mathbb{R}$ , let

$$A_N(I) := -\log \sup_{(x, \xi) \in \mathfrak{S}_1^{(N)} \times I} |\Phi_N(x, \xi)| \quad (4.2.13)$$

and choose some pair  $(\widetilde{x}_N, \widetilde{\xi}_N) \in \mathfrak{S}_1^{(N)} \times I$  such that

$$A_N(I) \leq -\log |\Phi_N(\widetilde{x}_N, \widetilde{\xi}_N)| \leq A_N(I) + \ln 2.$$

So  $|\Phi(\widetilde{x}_N, \widetilde{\xi}_N)| \geq \frac{1}{2} e^{-A_N(I)} = \frac{1}{2} \sup |\Phi_N(\cdot, \cdot)|$  on  $\mathfrak{S}_1^{(N)} \times I$ .

**Corollary 4.5** *For each  $\overline{\delta}$  there are  $\widetilde{C}, \widehat{\varepsilon}, \overline{c} > 0$  s.t. for every compact interval  $I$  s.t.  $|I| \leq \widetilde{\delta}$ , for all  $N$  for every  $(x, \xi) \in \mathfrak{S}_1^{(N)} \times I$ , for every  $\mathfrak{A} \subset \mathfrak{S}_{k_N+1}^{(N)}$  s.t.  $\mu_{k_N+1}^{(N)}(\mathfrak{A}) \geq \overline{\delta}$ ,*

$$\begin{aligned} |\Phi_N(x, \xi)| &\leq \widetilde{C} \exp \left( -\widehat{\varepsilon} V_N (\xi - \widetilde{\xi}_N)^2 + \overline{c} |\xi - \widetilde{\xi}_N| \sqrt{V_N A_N(I)} \right); \\ |\Phi_N(x, \xi | \mathfrak{A})| &\leq \widetilde{C} \exp \left( -\widehat{\varepsilon} V_N (\xi - \widetilde{\xi}_N)^2 + \overline{c} |\xi - \widetilde{\xi}_N| \sqrt{V_N A_N(I)} \right). \end{aligned}$$

*Proof.* We only give the proof in the case  $V_N$  is large, so that (4.2.12) holds. This is the case we need. We remark that the result also holds generally, because the estimate we seek is trivial when  $V_N$  is small.

Applying (4.2.12) with  $\widetilde{\xi}_N$  instead of  $\xi$  and  $\delta = \xi - \widetilde{\xi}_N$  gives

$$|\Phi_N(x, \xi)| \leq C \exp \left( -\widehat{\varepsilon} V_N (\xi - \widetilde{\xi}_N)^2 + \widehat{c} |\xi - \widetilde{\xi}_N| \sqrt{V_N D_N(\widetilde{\xi}_N)} \right).$$

By (4.2.6),  $e^{-A_N(\widetilde{\xi}_N)} \leq 2 |\Phi_N(\widetilde{x}_N, \widetilde{\xi}_N)| \leq 2C e^{-\frac{1}{5} \widetilde{\varepsilon} D_N(\widetilde{\xi}_N)}$ . We conclude that

$$D_N(\widetilde{\xi}) \leq C_1 A_N(I) + C_2$$

for some global constants  $C_1, C_2$ . The estimate of  $|\Phi_N(x, \xi)|$  follows. The second estimate is proved in the same way.  $\square$

#### 4.2.2 The LLT in the irreducible non-lattice case

We give the proof for arrays (Theorem 4.1'). Theorem 4.1 on chains follows, because every additive functional on a Markov chain is stably hereditary (Example 3.7).

We begin by proving that  $V_N \xrightarrow[N \rightarrow \infty]{} \infty$ . Otherwise  $\liminf V_N < \infty$ , and one can find  $N_\ell \uparrow \infty$  such that  $\text{Var}(S_{N_\ell}) = O(1)$ . Let  $\mathbf{X}'$  denote the sub-array with rows  $\mathbf{X}'^{(\ell)} = \mathbf{X}^{(N_\ell)}$ . By Theorem 2.2,

$f|_{X'}$  is center-tight, whence  $G_{ess}(X', f|_{X'}) = \{0\}$ . At the same time,  $G_{ess}(X, f) = G_{alg}(X, f) = \mathbb{R}$ , because  $f$  is irreducible and non-lattice. So  $G_{ess}(X', f|_{X'}) \neq G_{ess}(X, f)$ , in contradiction to the assumption that  $f$  is stably hereditary.

Next we fix  $z_N \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , and show that for every non-empty interval  $(a, b)$ , for every choice of  $x_1^{(N)} \in \mathfrak{S}_1^{(N)}$  ( $N \geq 1$ ),

$$\mathbb{P}_{x_1^{(N)}}[S_N - z_N \in (a, b)] \sim \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} (b - a), \text{ as } N \rightarrow \infty. \quad (4.2.14)$$

A well-known approximation argument [147], [17, chapter 10] reduces (4.2.14) to showing that for all  $\phi \in L^1(\mathbb{R})$  whose Fourier transform  $\widehat{\phi}(\xi) := \int_{\mathbb{R}} e^{-i\xi u} \phi(u) du$  has compact support,

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \mathbb{E}_{x_1^{(N)}}[\phi(S_N - z_N)] = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(u) du. \quad (4.2.15)$$

Fix  $\phi \in L^1$  such that  $\text{supp}(\widehat{\phi}) \subseteq [-L, L]$ . By the Fourier inversion formula,  $\mathbb{E}_{x_1^{(N)}}(\phi(S_N - z_N)) = \frac{1}{2\pi} \int_{-L}^L \widehat{\phi}(\xi) \Phi_N(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi$ . So (4.2.15) is equivalent to

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \cdot \frac{1}{2\pi} \int_{-L}^L \widehat{\phi}(\xi) \Phi_N(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \widehat{\phi}(0). \quad (4.2.16)$$

Below, we give a proof of (4.2.16).

We note for future reference that the proof of (4.2.16) below works under the milder assumption that  $\widehat{\phi}$  is bounded, continuous at zero and has compact support, e.g.  $\widehat{\phi} = \frac{1}{2\pi} 1_{[-\pi, \pi]}$  (which is the Fourier transform of  $\phi(u) = \frac{\sin(\pi u)}{\pi u} \notin L^1$ ).

Divide  $[-L, L]$  into segments  $I_j$  of length  $\leq \tilde{\delta}$  where  $\tilde{\delta}$  is given by Lemma 4.4, so that  $I_0$  is centered at 0. Let

$$J_{j,N} := \frac{1}{2\pi} \int_{I_j} \widehat{\phi}(\xi) \Phi_N(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi.$$

CLAIM 1 (CONTRIBUTION OF  $J_{0,N}$ ):

$$\sqrt{V_N} J_{0,N} \xrightarrow[N \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \widehat{\phi}(0). \quad (4.2.17)$$

*Proof of the claim.* Fix  $R > 0$ . Since  $J_{0,N} \geq 0$ ,  $A_N(J_{0,N}) = 0$ . By Corollary 4.5, given  $\varepsilon > 0$  there is  $R > 0$  such that

$$\left| \sqrt{V_N} \int_{\{\xi \in I_0 : |\xi| > R/\sqrt{V_N}\}} \widehat{\phi}(\xi) \Phi_N(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi \right| \leq \varepsilon.$$

Next, a change of variables  $\xi = s/\sqrt{V_N}$  gives

$$\sqrt{V_N} \int_{[|\xi| \leq R/\sqrt{V_N}]} \widehat{\phi}(\xi) \Phi_N(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi = \int_{[|s| \leq R]} \widehat{\phi}\left(\frac{s}{\sqrt{V_N}}\right) \mathbb{E}_{x_1^{(N)}}\left(e^{is \frac{S_N - z_N}{\sqrt{V_N}}}\right) ds.$$

By Dobrushin's CLT for inhomogeneous Markov arrays (Theorem 2.3)  $\frac{S_N - z_N}{\sqrt{V_N}}$  converges in distribution w.r.t.  $\mathbb{P}_{x_1^{(N)}}$  to the normal distribution with mean  $-z$  and variance 1. By Lévy's continuity theorem, this implies that

$$\mathbb{E}_{x_1^{(N)}}(e^{\frac{is\frac{S_N - z_N}{\sqrt{V_N}}}{\sqrt{V_N}}}) \xrightarrow[N \rightarrow \infty]{} e^{-isz - s^2/2}$$

uniformly on compacts, and so

$$\sqrt{V_N} \int_{|\xi| \leq R/\sqrt{V_N}} \widehat{\phi}(\xi) \Phi_N(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi = \widehat{\phi}(0) \int_{-R}^R e^{-isz} e^{-s^2/2} ds + o_{N \rightarrow \infty}(1).$$

Since this is true for all  $R$ , we can let  $R \rightarrow \infty$  sufficiently slow to obtain (4.2.17).

CLAIM 2 (CONTRIBUTION OF THE OTHER  $J_{j,N}$ ):  $\sqrt{V_N} J_{j,N} \xrightarrow[N \rightarrow \infty]{} 0$  for  $j \neq 0$ .

*Proof of the claim.* Since  $f$  is irreducible with algebraic range  $\mathbb{R}$ , the co-range of  $f$  is  $\{0\}$  (Theorems 3.1, 3.4). Since  $f$  is stably hereditary,

$$D_N(\xi) \xrightarrow[N \rightarrow \infty]{} \infty \text{ uniformly on compacts in } \mathbb{R} \setminus \{0\}.$$

By (4.2.6),  $\Phi_N(x_1^{(N)}, \xi) \rightarrow 0$  uniformly on compacts in  $\mathbb{R} \setminus \{0\}$ .

We will use this to show that for any interval  $I \subset \mathbb{R} \setminus \{0\}$

$$\sqrt{V_N} \int_I |\Phi(x_1^{(N)}, \xi)| d\xi \rightarrow 0. \quad (4.2.18)$$

By subdividing  $I$  into finitely many subintervals we see that it suffices to prove the claim for  $I = I_j$  for some  $j$ . Recall that  $A_N(I_j) = -\log \sup |\Phi_N(\cdot, \cdot)|$  on  $\mathfrak{S}_1^{(N)} \times I_j$ , and  $(\tilde{x}_{j,N}, \tilde{\xi}_{j,N})$  are points where this supremum is achieved up to factor 2. Set  $A_{j,N} := A_N(I_j)$ , then  $A_{j,N} \rightarrow \infty$  as  $N \rightarrow \infty$  for each  $j \neq 0$ .

Take large  $R$  and split  $I_j$  into two regions

$$I'_{j,N} := \left\{ \xi \in I_j : |\xi - \tilde{\xi}_{j,N}| \leq R \sqrt{\frac{A_{j,N}}{V_N}} \right\}, \quad I''_{j,N} := I_j \setminus I'_{j,N}.$$

Split the integral  $\int_{I_j} |\Phi(x_1^{(N)}, \xi)| d\xi$  into two integrals  $J'_{j,N}, J''_{j,N}$  accordingly.

o On  $I'_{j,N}$ ,  $|\Phi_N(x_1^{(N)}, \xi)| \leq e^{-A_{j,N}}$  and  $|I'_{j,N}| \leq 2R \sqrt{\frac{A_{j,N}}{V_N}}$ , so

$$\sqrt{V_N} |J'_{j,N}| \leq 2R \sqrt{A_{j,N}} e^{-A_{j,N}}.$$

o On  $I''_{j,N}$ , by Corollary 4.5,

$$\begin{aligned} |\Phi_N(x_1^{(N)}, \xi)| &\leq \tilde{C} \exp \left( -\widehat{\varepsilon} V_N |\xi - \tilde{\xi}_{j,N}| R \sqrt{\frac{A_{j,N}}{V_N}} + \bar{c} |\xi - \tilde{\xi}_{j,N}| \sqrt{V_N A_{j,N}} \right) \\ &\leq \tilde{C} \exp \left( -\frac{\widehat{\varepsilon}}{2} |\xi - \tilde{\xi}_{j,N}| \sqrt{A_{j,N} V_N} \right), \text{ provided } R \widehat{\varepsilon} > \bar{c} + \frac{\widehat{\varepsilon}}{2}. \end{aligned}$$

$$\text{Hence } \sqrt{V_N} J''_{j,N} \leq \sqrt{V_N} \tilde{C} \int_{-\infty}^{\infty} e^{-\frac{\tilde{\varepsilon}}{2}|s|} \sqrt{A_{j,N} V_N} ds = O(A_{j,N}^{-\frac{1}{2}}).$$

Combining these estimates, we obtain

$$\sqrt{V_N} \|\Phi_N(x_1^{(N)}, \cdot)\|_{L^1(I_j)} \leq 2R \sqrt{A_{j,N}} e^{-A_{j,N}} + \frac{C}{\sqrt{A_{j,N}}}. \quad (4.2.19)$$

Since  $A_{j,N} \rightarrow \infty$  as  $N \rightarrow \infty$  (4.2.18) follows.

$$\text{Since } |J_{j,N}| \leq \frac{\|\widehat{\phi}\|_\infty \|\Phi_N(x_1^{(N)}, \cdot)\|_{L^1(I_j)}}{2\pi}, \text{ claim 2 follows from (4.2.18).}$$

**Remark 4.6** Note that in the proof of (4.2.18) the irreducibility assumption is only used at the last sentence, namely, to conclude that  $A_{N,j} \rightarrow 0$  as  $N \rightarrow \infty$ . In particular, (4.2.19) holds for arbitrary arrays, irreducible or not.

Claims 1 and 2 imply (4.2.16), and (4.2.16) implies (4.2.14) by [17, chapter 10]. This proves the LLT theorem for initial distributions concentrated at single points (i.e.  $\mathbb{P} = \mathbb{P}_{x_1^{(N)}}$ ). To deduce the theorem for arbitrary initial distribution  $\mu_1^{(N)}(dx_1^{(N)})$ , it is sufficient to prove the following claim and then integrate:

**CLAIM 3:** (4.2.14) holds uniformly with respect to the choice of  $\{x_n^{(N)}\}$ .

*Proof of the claim.* Assume by contradiction that this is false, then there exists  $\varepsilon > 0$  and  $N_k \rightarrow \infty$  with  $y_1^{(N_k)}$  such that  $\mathbb{P}_{y_1^{(N_k)}}[S_{N_k} - z_{N_k} \in (a, b)] / \frac{e^{-z^2/2}(b-a)}{\sqrt{2\pi V_{N_k}}} \notin [e^{-\varepsilon}, e^\varepsilon]$ . But this contradicts (4.2.14) for any sequence  $\{x_1^{(N)}\}$  such that  $x_1^{(N_k)} = y_1^{(N_k)}$ .  $\square$

### 4.2.3 The LLT for the irreducible lattice case

We give the proof in the context of arrays (Theorem 4.2'):  $X$  is a uniformly elliptic array, and  $f$  is an additive functional on  $X$  which is a.s. uniformly bounded, hereditary, irreducible, and with algebraic range  $t\mathbb{Z}$  with  $t > 0$ . Without loss of generality,  $t = 1$ , otherwise work with  $t^{-1}f$ .

By Lemma 3.9 and the assumption that  $G_{alg}(\mathbb{R}) = \mathbb{Z}$ , there are constants  $c_n^{(N)}$  such that  $f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) - c_n^{(N)} \in \mathbb{Z}$  a.s. We may assume without loss of generality that  $c_n^{(N)} = 0$ , otherwise we work with  $f - c$ . So

$$S_N \in \mathbb{Z} \text{ a.s. for every } N \geq 1.$$

We will show that for every sequence of numbers  $z_N \in \mathbb{Z}$  such that  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , and for every  $x_1^{(N)} \in \mathfrak{S}_n^{(N)}$ ,

$$\mathbb{P}_{x_1^{(N)}}(S_N = z_N) = [1 + o(1)] \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}}, \text{ as } N \rightarrow \infty. \quad (4.2.20)$$

As in the irreducible case, once we prove (4.2.20) for all choices of  $\{x_1^{(N)}\}$ , it automatically follows that (4.2.20) holds uniformly in  $\{x_1^{(N)}\}$ . Integrating over  $(\mathfrak{S}_1^{(N)}, \mathcal{B}(\mathfrak{S}_1^{(N)}), \mu_1^{(N)})$  gives (4.1.4) with  $k = 0$ . For general  $k$ , take  $z'_N := z_N + k$ .

The assumptions on  $f$  imply that  $\text{Var}(S_N) \xrightarrow[N \rightarrow \infty]{} \infty$ . The proof is a routine modification of the argument we used in the non-lattice case, so we omit it.

Observe that  $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{im\xi} d\xi$  is equal to zero when  $m \in \mathbb{Z} \setminus \{0\}$ , and equal to one when  $m = 0$ . In particular, since  $S_N - z_N \in \mathbb{Z}$  almost surely, for every  $x_1^{(N)} \in \mathfrak{S}_1^{(N)}$

$$\mathbb{P}_{x_1^{(N)}}(S_N - z_N = 0) = \mathbb{E}_{x_1^{(N)}} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\xi(S_N - z_N)} d\xi \right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi.$$

Thus to prove (4.2.20) it is sufficient to show that

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_N(x_1^{(N)}, \xi) e^{-i\xi z_N} d\xi = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}. \quad (4.2.21)$$

Notice that (4.2.21) is (4.2.16) in the case  $\phi(u) = \frac{\sin(\pi u)}{\pi u}$ ,  $\widehat{\phi}(\xi) = \frac{1}{2\pi} 1_{[-\pi, \pi]}(\xi)$ , and can be proved in almost exactly the same way.

Here is a sketch of the proof. One divides  $[-\pi, \pi]$  into segments  $I_j$  of length less than the  $\tilde{\delta}$  of Lemma 4.4.

The contribution of the interval which contains zero is asymptotic to  $\frac{1}{\sqrt{2\pi V_N}} e^{-z^2/2}$ . This is shown as in claim 1 of the preceding proof.

The remaining intervals are bounded away from  $2\pi\mathbb{Z}$ . Their contribution is  $o(1/\sqrt{V_N})$ . This can be seen as in claim 2 of the preceding proof, using the facts that since  $f$  is irreducible with algebraic range  $\mathbb{Z}$ ,  $H(X, f) = 2\pi\mathbb{Z}$  (Theorems 3.1, 3.4), and since  $f$  is hereditary and  $G_{alg}(f) = \mathbb{Z}$ ,  $f$  is stably hereditary, whence  $D_N(\xi) \xrightarrow[N \rightarrow \infty]{} 0$  uniformly on compacts in  $\mathbb{R} \setminus 2\pi\mathbb{Z}$ .  $\square$

#### 4.2.4 The mixing LLT

The proof is very similar to the proof of the local limit theorem, except that we use  $\Phi(x, \xi | \mathfrak{A})$  instead of  $\Phi(x, \xi)$ .

We outline the proof in the non-lattice case, and leave the lattice case to the reader. Suppose  $X$  is a uniformly elliptic Markov array, and that  $f$  is a.s. uniformly bounded, stably hereditary, irreducible and with algebraic range  $\mathbb{R}$ .

Let  $\mathfrak{A}_N \in \mathfrak{S}_{k_N+1}^{(N)}$  be measurable sets s.t.  $\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N) > \delta > 0$ , and let  $x_N \in \mathfrak{S}_1^{(N)}$  be points. Suppose  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ . As before,  $V_N \rightarrow \infty$ , and a standard approximation argument ([17], chapter 10) says that it is enough to show that for every  $\phi \in L^1(\mathbb{R})$  s.t.  $\text{supp}(\widehat{\phi}) \subset [-L, L]$ ,

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \cdot \frac{1}{2\pi} \int_{-L}^L \widehat{\phi}(\xi) \Phi_N(x_N, \xi | \mathfrak{A}_N) e^{-i\xi z_N} d\xi = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \widehat{\phi}(0).$$

Divide  $[-L, L]$  as before into intervals  $I_j$  of length  $\leq \tilde{\delta}$  where  $\tilde{\delta}$  is given by Lemma 4.4 and  $I_0$  is centered at zero, and let

$$J_{j,N} := \frac{1}{2\pi} \int_{I_j} \widehat{\phi}(\xi) \Phi_N(x_N, \xi | \mathfrak{A}_N) e^{-i\xi z_N} d\xi.$$

CLAIM 1:  $\sqrt{V_N} J_{0,N} \xrightarrow[N \rightarrow \infty]{} (2\pi)^{-\frac{1}{2}} e^{-z^2/2} \widehat{\phi}(0)$ .

*Proof of the claim:* Fix  $R > 0$ . As before, applying Corollary 4.5 with  $A_N = 0$  we conclude that for each  $\varepsilon > 0$  there is  $R > 0$  such that

$$\left| \sqrt{V_N} \int_{\{\xi \in I_0: |\xi| > R/\sqrt{V_N}\}} \widehat{\phi}(\xi) \Phi_N(x_N, \xi | \mathfrak{A}_N) e^{-i\xi z_N} d\xi \right| \leq \varepsilon.$$

Next the change of variables  $\xi = s/\sqrt{V_N}$  gives

$$\begin{aligned} & \sqrt{V_N} \int_{\{\xi \in I_0: |\xi| \leq R/\sqrt{V_N}\}} \widehat{\phi}(\xi) \Phi(x_N, \xi | \mathfrak{A}_N) e^{-i\xi z_N} d\xi \\ &= \int_{-R}^R \widehat{\phi}\left(\frac{s}{\sqrt{V_N}}\right) \mathbb{E}_{x_N} \left( e^{is(\frac{S_N - z_N}{\sqrt{V_N}})} \middle| X_{k_N+1}^{(N)} \in \mathfrak{A}_N \right) d\xi \\ &= \frac{1}{\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N)} \int_{-R}^R \widehat{\phi}\left(\frac{s}{\sqrt{V_N}}\right) \mathbb{E}_{x_N} \left( e^{is(\frac{S_N - z_N}{\sqrt{V_N}})} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)}) \right) d\xi. \end{aligned} \quad (4.2.22)$$

We analyze the expectation in the integrand. Take  $1 \leq r_N \leq k_N$  such that  $r_N \rightarrow \infty$  and  $r_N/\sqrt{V_N} \rightarrow 0$ , and let

$$S_N^* := \sum_{j=1}^{k_N - r_N} f_j^{(N)}(X_j^{(N)}, X_{j+1}^{(N)}) \equiv S_N - \sum_{j=k_N - r_N + 1}^{k_N} f_j^{(N)}(X_j^{(N)}, X_{j+1}^{(N)}).$$

Since  $\text{ess sup } |\mathfrak{f}| < \infty$ ,  $|S_N - S_N^*| = o(\sqrt{V_N})$ , and so

$$\begin{aligned} & \mathbb{E}_{x_N} \left( e^{is(\frac{S_N - z_N}{\sqrt{V_N}})} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)}) \right) = \mathbb{E}_{x_N} \left( e^{is(\frac{S_N^* - z_N}{\sqrt{V_N}})} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)}) \right) + o(1) \\ &= \mathbb{E}_{x_N} \left( e^{is(\frac{S_N^* - z_N}{\sqrt{V_N}})} \mathbb{E} \left( 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)}) | X_1^{(N)}, \dots, X_{k_N - r_N}^{(N)} \right) \right) + o(1) \\ &= \mathbb{E}_{x_N} \left( e^{is(\frac{S_N^* - z_N}{\sqrt{V_N}})} \mathbb{E} \left( 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)}) | X_{k_N - r_N}^{(N)} \right) \right) + o(1) \text{ by the Markov property} \\ &\stackrel{!}{=} \mathbb{E}_{x_N} \left( e^{is(\frac{S_N^* - z_N}{\sqrt{V_N}})} \left[ \mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N) + O(\theta^{r_N}) \right] \right) + o(1), \text{ where } 0 < \theta < 1 \end{aligned}$$

and  $\stackrel{!}{=}$  uses the exponential mixing estimate (1.2.3). Since  $\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N)$  is bounded below, and  $\frac{S_N^* - z_N}{\sqrt{V_N}}$  converges in distribution to the standard normal distribution by Dobrushin's theorem, we may conclude that

$$\mathbb{E}_{x_N} \left( e^{is(\frac{S_N^* - z_N}{\sqrt{V_N}})} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)}) \right) = \frac{1 + o(1)}{\sqrt{2\pi}} e^{-z^2/2 - 2is} \mathbb{P} \left( X_{k_N+1}^{(N)} \in \mathfrak{A}_N \right).$$

Substituting this in (4.2.22) gives the claim.

CLAIM 2:  $\sqrt{V_N} J_{j,N} \xrightarrow[N \rightarrow \infty]{} 0$  for  $j \neq 0$ .

The claim is proved as in the previous proof, but with (4.2.7) replacing (4.2.6). Together, claims 1 and 2 imply the theorem.  $\square$

### 4.3 Notes and references

For a brief account of the history of the local limit theorem, see the end of the preface.

Many of the techniques we used in this chapter have a long history. The reduction of the LLT to the asymptotic analysis of the integrals (4.2.16) and (4.2.21) for  $\phi \in L^1$  with Fourier transforms with compact support was already used by Stone [147] for proving local limit theorems for sums of iid random variables. As mentioned at the end of the synopsis, the method of characteristic function operators is due to Nagaev [109], who used it to prove central and local limit theorems for homogeneous Markov chains, and this method was used extensively in dynamical systems. Hafouta & Kifer [69], Hafouta [66, 67], and Dragičević, Froyland, & González-Tokman [53], used this technique to prove the local limit theorem in a non-homogeneous setup.

The terminology “mixing LLT” is due to Rényi [125], who initiated the study of the stability of limit theorems under conditioning and changes of measure. The relevance of Mixing LLT to the study of reducible case is noted by Guivarc'h & Hardy [65]. Mixing LLT have numerous other applications including mixing of special flows [65, 50], homogenization [44] and skew products (see in particular, Theorem 5.2 in Chapter 5). Mixing LLT for additive functionals of (stationary) Gibbs-Markov processes were proved by Aaronson & Denker [4].

# Chapter 5

## The local limit theorem in the reducible case

In this chapter we prove the local limit theorem for  $\mathbb{P}(S_N - z_N \in (a, b))$  when  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}}$  converges to a finite limit and  $f$  is reducible. In the reducible case, the asymptotic behavior of  $\mathbb{P}(S_N - z_N \in (a, b))$  depends on the details of  $f_n(X_n, X_{n+1})$ . The dependence is strong for small intervals, and weak for large intervals.

### 5.1 Main results

#### 5.1.1 Heuristics and warm up examples

An additive functional is called **reducible** if

$$f = g + c$$

where  $c$  is center-tight, and the algebraic range of  $g$  is strictly smaller than the algebraic range of  $f$ . By the results of Chapter 3, if  $\text{Var}(S_N(f)) \rightarrow \infty$ ,  $X$  is uniformly elliptic, and  $f$  is a.s. bounded, then we can choose  $g$  to be irreducible. In this case

$$S_N(f) = S_N(g) + S_N(c).$$

where  $\text{Var}(S_N(g)) \sim \text{Var}(S_N(f)) \rightarrow \infty$ ,  $\text{Var}(S_N(c)) = O(1)$ , and  $S_N(g)$  satisfies the lattice local limit theorem. **The contribution of  $S_n(c)$  cannot be neglected.** In this chapter we give the corrections to the LLT needed to take  $S_n(c)$  into account.

Before stating our results in general, we discuss two simple examples which demonstrate some of the possible effects of  $S_N(c)$ .

#### Example 5.1 (Simple random walk with continuous first step and drift):

Suppose  $\{X_n\}_{n \geq 1}$  are independent real-valued random variables, where  $X_1$  is distributed like a random variable  $\mathfrak{F}$ , and  $X_i$  ( $i \geq 2$ ) are equal to 0, 1 with equal probabilities.

$\mathfrak{F}$  could be arbitrary, but we assume for simplicity that  $0 \leq \mathfrak{F} < 1$  a.s.,  $\mathbb{E}[\mathfrak{F}] = \frac{1}{2}$ , the distribution of  $\mathfrak{F}$  has a density, and  $\mathfrak{F}$  is *not* uniformly distributed on  $[0, 1]$ . Let  $\mu_{\mathfrak{F}}$  denote the probability measure associated with the distribution of  $\mathfrak{F}$ .

$S_N := X_1 + \dots + X_N$  is exactly  $S_N(f)$ , where  $f_n(x, y) := x$ . Since the distribution of  $\mathfrak{F}$  has a density,  $f$  has algebraic range  $\mathbb{R}$ .

The following decomposition shows that  $f$  is reducible, with essential range  $\mathbb{Z}$ : Let  $\delta_{ij}$  be Kronecker's delta, then  $f = g + c$  where

$$g_n(x, y) := (1 - \delta_{1,n})x, \quad c_n(x, y) := \delta_{1,n}x,$$

$g$  is irreducible with essential range  $\mathbb{Z}$ , and  $c$  is center tight.

We have  $S_N = \underbrace{(X_2 + \cdots + X_N)}_{S_N(g)} + \underbrace{X_1}_{S_N(c)}$ . Clearly,  $S_N(g), S_N(c)$  are independent;  $S_N(c) \sim \mathfrak{F}$ ; and  $S_N(g)$  has the binomial distribution  $B(\frac{1}{2}, N-1)$ . So  $S_N$  has distribution  $\mu_{\mathfrak{F}} * B(\frac{1}{2}, N-1)$ . This distribution has a density, which we denote by  $p_N(x)dx$ . The following holds as  $N \rightarrow \infty$ :

**(A) Non-uniform scaling limit for  $p_N(x)dx$ :**  $m_N := p_N(x)dx$  is a positive functional on  $C_c(\mathbb{R}) = \{\text{continuous functions with compact support}\}$ . Fix  $z_N := \mathbb{E}(S_N) = N/2$  and let  $V_N := \text{Var}(S_N) \sim N/4$ . Then for every  $\phi \in C_c(\mathbb{R})$  and  $N$  even,

$$\begin{aligned} \int \phi(x - z_N) p_N(x) dx &= \mathbb{E}[\phi(S_N - z_N)] = \mathbb{E}[\phi(S_N(g) + S_N(c) - z_N)] \\ &= \sum_{m \in \mathbb{Z}} \mathbb{E}[\phi(\mathfrak{F} + m - z_N)] \mathbb{P}[S_N(g) = m] = \sum_{m=0}^{N-1} \binom{N-1}{m} \frac{1}{2^{N-1}} \mathbb{E}[\phi(\mathfrak{F} + m - z_N)] \\ &= \frac{1}{2^{N-1}} \sum_{m=0}^{N-1} \binom{N-1}{m} \psi(m - \frac{N}{2}), \text{ where } \psi(m) := \mathbb{E}[\phi(\mathfrak{F} + m)] \\ &= \frac{1}{2^{N-1}} \sum_{m=-N/2}^{N/2-1} \binom{N-1}{m + N/2} \psi(m) \sim \frac{1}{\sqrt{2\pi V_N}} \sum_{m \in \mathbb{Z}} \psi(m) \text{ by Stirling's formula} \\ &\sim \frac{1}{\sqrt{2\pi V_N}} \sum_{m \in \mathbb{Z}} \mathbb{E}[\phi(\mathfrak{F} + m)], \text{ as } N \rightarrow \infty. \text{ This also holds for } N \text{ odd.} \end{aligned}$$

Thus the distribution of  $S_N - z_N$  tends to zero in the vague topology of Radon measure on  $\mathbb{R}$  “at a rate of  $1/\sqrt{2\pi N}$ ,” and if we inflate it by  $\sqrt{2\pi V_N}$  then it converges in the vague topology to  $\mu_{\mathfrak{F}}^*(\text{counting measure on } \mathbb{Z})$ .

By the assumptions on  $\mathfrak{F}$ , the scaling limit  $\mu_{\mathfrak{F}}^*(\text{counting measure on } \mathbb{Z})$  is not a Haar measure on a closed subgroup of  $\mathbb{R}$ . This is different from the irreducible case, when the scaling limit is the Haar measure on  $G_{\text{ess}}(\mathsf{X}, \mathsf{f})$ .

**(B) Non-standard limit for  $\sqrt{2\pi V_N} \mathbb{P}[S_N - z_N \in (a, b)]$ :** Fix  $a, b \in \mathbb{R} \setminus \mathbb{Z}$  s.t.  $|a - b| > 1$ . Repeating the previous calculation with  $\phi_i \in C_c(\mathbb{R})$  such that  $\phi_1 \leq 1_{(a,b)} \leq \phi_2$  and  $\sum_{m \in \mathbb{Z}} \mathbb{E}[\phi_i(m + \mathfrak{F})] \approx \sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a,b)}(m + \mathfrak{F})]$  gives for  $z_N = \mathbb{E}(S_N)$  that

$$\sqrt{2\pi V_N} \mathbb{P}[S_N - z_N \in (a, b)] \xrightarrow[N \rightarrow \infty]{} \sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a,b)}(m + \mathfrak{F})]. \quad (5.1.1)$$

This is different than the limit in the irreducible non-lattice LLT (Theorem 4.1):

$$\sqrt{2\pi V_N} \mathbb{P}[S_N - z_N \in (a, b)] \xrightarrow[N \rightarrow \infty]{} |a - b|; \quad (5.1.2)$$

or the limit in the irreducible lattice LLT with range  $\mathbb{Z}$  (Theorem 4.2):

$$\sqrt{2\pi V_N} \mathbb{P}[S_N - z_N \in (a, b)] \xrightarrow[N \rightarrow \infty]{} \sum_{m \in \mathbb{Z}} 1_{(a,b)}(m). \quad (5.1.3)$$

**(C) Robustness for large intervals:** Although different, the limits in (5.1.1), (5.1.3) and (5.1.2) are nearly the same as  $|a - b| \rightarrow \infty$ .

The ratio between the limits in (5.1.3), (5.1.2) tends to one as  $|a - b| \rightarrow \infty$ . The ratio between the limits in (5.1.1), (5.1.3) tends to one too, because  $\text{supp}(\mathfrak{F}) \subset [0, 1]$ , so  $|a - b| - 2 \leq \sum_{m \in \mathbb{Z}} 1_{(a,b)}(m + \mathfrak{F}) \leq |a - b| + 2$  a.s., whence

$$\left| \frac{\sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a,b)}(m + \mathfrak{F})]}{\sum_{m \in \mathbb{Z}} 1_{(a,b)}(m)} - 1 \right| \leq \frac{2}{|a - b|} \xrightarrow{|a-b| \rightarrow \infty} 0.$$

Example 5.1 is very special in that  $S_n(g), S_N(c)$  are independent. Nevertheless, we will see below that (A), (B), (C) are general phenomena, which also happen when  $S_N(g), S_N(h)$  are strongly correlated. The following simple example demonstrates another pathology that is quite general:

**Example 5.2 (Gradient perturbation of the lazy random walk) :**

Suppose  $X_n, Y_n$  are independent random variables such that  $X_n = -1, 0, +1$  with equal probabilities, and  $Y_n$  are uniformly distributed in  $[0, 1]$ . Let  $\mathbf{X} = \{(X_n, Y_n)\}_{n \geq 1}$ .

- The additive functional  $g_n((x_n, y_n); (x_{n+1}, y_{n+1})) = x_n$  generates the lazy random walk on  $\mathbb{Z}$ ,  $S_N(g) = X_1 + \dots + X_N$ . It is irreducible, and satisfies the lattice LLT with range  $\mathbb{Z}$ .
- The additive functional  $c_n((x_n, y_n), (x_{n+1}, y_{n+1})) = y_n - y_{n+1}$  is center-tight, and  $S_N(c) = Y_{N+1} - Y_1$ .
- The sum  $f = g + c$  is reducible, with algebraic range  $\mathbb{R}$  (because of  $c$ ) and essential range  $\mathbb{Z}$  (because of  $g$ ). It generates the process

$$S_N(f) = S_N(g) + Y_{N+1} - Y_1.$$

$S_N(f)$  lies in a *random coset*  $b_N + \mathbb{Z}$ , where  $b_N = Y_{N+1} - Y_1$ . Since the distribution of  $b_N$  is continuous,  $\mathbb{P}[S_N - z_N = k] = 0$  for all  $z_N, k \in \mathbb{Z}$ , and the standard lattice LLT fails. To deal with this, we must “shift”  $S_N - z_N$  back to  $\mathbb{Z}$ . This leads to the following (correct) statement: For all  $z_N \in \mathbb{Z}$  s.t.  $\frac{z_N}{\sqrt{V_N}} \rightarrow z$ , for all  $k \in \mathbb{Z}$ ,

$$\mathbb{P}[S_N - z_N - b_N = k] \sim \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}}$$

Notice the shift by a *random* bounded quantity  $b_N$ .

### 5.1.2 The LLT in the reducible case

**Theorem 5.1.** *Let  $\mathbf{X} = \{X_n\}$  be a uniformly elliptic Markov chain, and let  $f$  be a reducible a.s. uniformly bounded additive functional with essential range  $\delta(f)\mathbb{Z}$ , where  $\delta(f) \neq 0$ . Then there are random variables  $b_N = b_N(X_1, X_{N+1})$  and  $\mathfrak{F} = \mathfrak{F}(X_1, X_2, \dots)$  with the following properties:*

(1) *For every  $z_N \in \delta(f)\mathbb{Z}$  such that  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , for every  $\phi \in C_c(\mathbb{R})$  and  $x \in \mathfrak{S}_1$ ,*

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \mathbb{E}_x [\phi(S_N - z_N - b_N)] = \frac{\delta(f) e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} E_x [\phi(m\delta(f) + \mathfrak{F})].$$

(2) For every  $\mathfrak{A}_{N+1} \subset \mathfrak{S}_{N+1}$  measurable such that  $\mathbb{P}[X_{N+1} \in \mathfrak{A}_{N+1}]$  is bounded below, and for every  $x \in \mathfrak{S}_1$ ,

$$\lim_{N \rightarrow \infty} \sqrt{V_N} \mathbb{E}_x [\phi(S_N - z_N - b_N) | X_{N+1} \in \mathfrak{A}_{N+1}] = \frac{\delta(\mathfrak{f}) e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \mathbb{E}_x [\phi(m\delta(\mathfrak{f}) + \mathfrak{F})].$$

(3)  $\|b_N\|_\infty \leq 9\delta(\mathfrak{f})$ , and  $\mathfrak{F} \in [0, \delta(\mathfrak{f})]$ .

The statement may seem at first sight different from the previous LLT we discussed, so we'd like to spend some time on clarifying what it is saying.

o  $\mathbb{E}_x [\phi(S_N - z_N - b_N)]$ , when viewed as a positive functional on  $C_c(\mathbb{R})$ , represents the measure on  $\mathbb{R}$ ,  $m_{x,N}(E) = \mathbb{P}_x[S_N - z_N - b_N(X_1, X_{N+1}) \in E]$ . This is the distribution of  $S_N$ , conditioned on  $X_1 = x$ , after a shift by  $z_N + b_N(X_1, X_{N+1})$ . The deterministic shift by  $z_N$  cancels the drift of  $S_N$  (notice that  $z_N \approx \mathbb{E}(S_N) \approx \mathbb{E}_x(S_N)$ ). The random shift  $b_N$  is needed to force  $S_N$  to stay inside  $\delta(\mathfrak{f})\mathbb{Z}$ , see Example 5.2.

o The linear functional

$$\mathcal{A}_x(\phi) := \delta(\mathfrak{f}) \sum_{m \in \mathbb{Z}} \mathbb{E}_x [\phi(m\delta(\mathfrak{f}) + \mathfrak{F})] \quad (5.1.4)$$

defines the element of  $C_c(\mathbb{R})^*$  which represents the measure  $\mu_{x,\mathfrak{F}} * m_{\delta(\mathfrak{f})}$ , where  $\mu_{x,\mathfrak{F}}(E) = \mathbb{P}_x(\mathfrak{F} \in E)$  and  $m_{\delta(\mathfrak{f})} := \delta(\mathfrak{f}) \times$  counting measure on  $\delta(\mathfrak{f})\mathbb{Z}$ . So part (1) of Theorem 5.1 says that  $m_{x,N} \rightarrow 0$  in  $C_c(\mathbb{R})^*$  at rate  $1/\sqrt{V_N}$ , and gives the scaling limit  $\sqrt{2\pi V_N} m_N \xrightarrow[N \rightarrow \infty]{w^*} \mu_{x,\mathfrak{F}} * m_{\delta(\mathfrak{f})}$  when  $z = 0$ . See Example 5.1.

o As in Example 5.1, part (1) implies the following: For all  $a < b$  s.t.  $\mathfrak{F}$  has no atoms in  $\{a, b\} + \delta(\mathfrak{f})\mathbb{Z}$ , and for all  $z_N \in \delta(\mathfrak{f})\mathbb{Z}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ ,

$$\mathbb{P}_x[S_N - z_N - b_N \in (a, b)] = [1 + o(1)] \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} \cdot \mathcal{A}_x(1_{(a,b)}), \text{ and}$$

$$\mathcal{A}_x(1_{(a,b)}) \sim \begin{cases} |a - b| & \text{as } |a - b| \rightarrow \infty \\ \mathbb{P}_x[\mathfrak{F} \in (a, b)] & \text{for } (a, b) \subset [0, \delta(\mathfrak{f})]. \end{cases}$$

Viewed from this perspective,  $A_x(1_{(a,b)})$  is a “correction” to the term  $|a - b|$  in classical LLT (4.1.1), which is needed for intervals with length of order  $\delta(\mathfrak{f})$ .

These observations should be sufficient to understand the content of part (1). Part (2) is a “mixing” version of part (1), in the sense of §4.1.3. Such results are particularly useful in the reducible setup for the following reason. The random shift  $b_N(X_1, X_{N+1})$  is sometimes a nuisance, and it is tempting to turn it into a deterministic quantity by conditioning on  $X_1, X_{N+1}$ . We would have liked to say that part (1) survives such conditioning, but we cannot. The best we can say in general is that part (1) remains valid under conditioning of the form  $X_1 = x_1, X_{N+1} \in \mathfrak{A}_{N+1}$  provided  $\mathbb{P}(X_{N+1} \in \mathfrak{A}_{N+1})$  is bounded below. This is the content of part (2). For an example how to use such a statement, see §5.2.3.

In the following sections, we explore some of the consequences of Theorem 5.1.

### 5.1.3 Irreducibility as a necessary condition for the mixing LLT

Theorem 5.1 exposes the pathologies that could happen in the reducible case. But is irreducibility a necessary condition for the non-lattice LLT? No!

**Example 5.3** Take example 5.1 with fixed  $x$  and  $\mathfrak{F}$  uniformly distributed on  $[0, 1]$ , given  $X_1 = x$ . In this case,  $\delta(\mathfrak{f}) = 1$ ,  $\mu_{x, \mathfrak{F}} * m_{\delta(\mathfrak{f})} = \text{Lebesgue's measure}$ ,  $A_x(1_{(a, b)}) = |a - b|$ , and  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z \Rightarrow \mathbb{P}_x[S_N - z_N \in (a, b)] \sim \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} |a - b|$ , even though  $\mathfrak{f}$  is reducible, with essential range  $\mathbb{Z}$ . Of course, such behavior is immediately destroyed if we modify  $X_1$ .

In this section we show that irreducibility is a necessary condition for the mixing LLT, provided we impose the mixing LLT not just for  $(X, \mathfrak{f})$ , but also for all  $(X', \mathfrak{f}')$  obtained from  $(X, \mathfrak{f})$  by changing finitely many terms.

Let  $\mathfrak{f}$  be an additive functional on a Markov chain  $X$ . Denote the state spaces of  $X$  by  $\mathfrak{S}_n$ , and write  $X = \{X_n\}_{n \geq 1}$ ,  $\mathfrak{f} = \{f_n\}_{n \geq 1}$ . A sequence of events  $\mathfrak{A}_k \subset \mathfrak{S}_k$  is called **regular** if  $\mathfrak{A}_k$  are measurable, and  $\mathbb{P}(X_n \in \mathfrak{A}_n)$  is bounded away from zero.

- We say that  $(X, \mathfrak{f})$  satisfies the **mixing non-lattice local limit theorem** if  $V_N := \text{Var}(S_N) \rightarrow \infty$ , and for every regular sequence of events  $\mathfrak{A}_n \subset \mathfrak{S}_n$ ,  $x \in \mathfrak{S}_1$ , for all  $z_N \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , and for each non-empty interval  $(a, b)$ ,

$$\mathbb{P}_x\left(S_N - z_N \in (a, b) \mid X_{N+1} \in \mathfrak{A}_{N+1}\right) = [1 + o(1)] \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} |a - b| \text{ as } N \rightarrow \infty.$$

- Fix  $t > 0$ . We say that  $(X, \mathfrak{f})$  satisfies the **mixing uniform distribution mod  $t$  property**, if for every regular sequence of events  $\mathfrak{A}_n \subset \mathfrak{S}_n$ ,  $x \in \mathfrak{S}_1$ , and a non-empty interval  $(a, b)$  with length less than one,

$$\mathbb{P}_x(S_N \in (a, b) + t\mathbb{Z} \mid X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N \rightarrow \infty]{} \frac{|a - b|}{t}.$$

**Theorem 5.2.** Let  $\mathfrak{f}$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Given  $m$ , let  $(X_m, \mathfrak{f}_m) := (\{X_n\}_{n \geq m}, \{f_n\}_{n \geq m})$ . The following are equivalent:

- (1)  $\mathfrak{f}$  is irreducible with algebraic range  $\mathbb{R}$ ;
- (2)  $(X_m, \mathfrak{f}_m)$  satisfy the mixing non-lattice local limit theorem for all  $m$ ;
- (3)  $(X_m, \mathfrak{f}_m)$  satisfy the mixing uniform distribution mod  $t$  for all  $m$  and  $t > 0$ .

### 5.1.4 Universal bounds for $\mathbb{P}_x[S_N - z_N \in (a, b)]$

So far we have considered the problem of finding  $\mathbb{P}_x[S_N - z_N \in (a, b)]$  up to asymptotic equivalence. We now consider the problem of finding  $\mathbb{P}_x[S_N - z_N \in (a, b)]$  up to bounded multiplicative error, assuming only that  $V_N \rightarrow \infty$ .

We already saw that the predictions of the LLT for *large* intervals  $(a, b)$  are nearly the same both in the reducible and irreducible, lattice and non-lattice cases. Therefore we expect *universal* lower and upper bounds, for all sufficiently large intervals without further assumptions on irreducibility or on the arithmetic structure of the range. The question is how large is “sufficiently large.”

We certainly cannot expect universal lower and upper bounds for intervals smaller than the **graininess constant** of  $(X, f)$ :

$$\delta(f) := \begin{cases} t & G_{ess}(X, f) = t\mathbb{Z}, \ t > 0 \\ 0 & G_{ess}(X, f) = \mathbb{R} \\ \infty & G_{ess}(X, f) = \{0\}, \end{cases} \quad (5.1.5)$$

because intervals with length less than  $\delta(f)$  may fall in the gaps of the support of  $S_N - z_N$ . Theorem 5.1 can be used to see that universal bounds do apply as soon as  $|a - b| > \delta(f)$ :

**Theorem 5.3.** *Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . Then for every interval  $(a, b)$  of length  $L > \delta(f)$ , for all  $\varepsilon > 0$ ,  $x \in \mathfrak{S}_1$  and  $z_N \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , for all for all  $N$  large enough,*

$$\mathbb{P}_x(S_N - z_N \in (a, b)) \leq \frac{e^{-z^2/2}|a - b|}{\sqrt{2\pi V_N}} \left( 1 + \frac{21\delta(f)}{L} + \varepsilon \right), \quad (5.1.6)$$

$$\mathbb{P}_x(S_N - z_N \in (a, b)) \geq \frac{e^{-z^2/2}|a - b|}{\sqrt{2\pi V_N}} \left( 1 - \frac{\delta(f)}{L} - \varepsilon \right). \quad (5.1.7)$$

In addition, if  $0 < \delta(f) < \infty$  and  $k\delta(f) \leq L \leq (k+1)\delta(f)$ ,  $k \in \mathbb{N}$ , then

$$\left( \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} \right) k\delta(f) \lesssim \mathbb{P}_x(S_N - z_N \in (a, b)) \lesssim \left( \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} \right) (k+1)\delta(f).$$

Here  $A_N \lesssim B_N$  means that  $\limsup_{N \rightarrow \infty} (A_N/B_N) \leq 1$ .

We note that both upper and lower bound become asymptotic to the Gaussian density as  $L \rightarrow \infty$ . Notice also that the theorem makes no assumptions on the irreducibility of  $f$ .

Theorem 5.3 is an easy corollary of Theorem 5.1, see §5.2.4, but this is an overkill. At the end of the chapter we will supply a proof of universal bounds for intervals of length  $L > 2\delta(f)$ , which does not require the full force of Theorem 5.1, and which also applies to arbitrary initial distributions and to arrays.

## 5.2 Proofs

### 5.2.1 Characteristic functions

**Setup:** Throughout this section we assume that  $X = \{X_n\}$  is a uniformly elliptic Markov chain with state spaces  $\mathfrak{S}_n$ , marginals  $\mu_n(E) = \mathbb{P}(X_n \in E)$ , and transition probabilities  $\pi_{n,n+1}(x, dy) = p_n(x, y)\mu_{n+1}(dy)$  which satisfy the uniform ellipticity condition with ellipticity constant  $\varepsilon_0$ .

For every bounded measurable function  $\varphi : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \rightarrow \mathbb{R}$ , we let

$$\mathbb{E}(\varphi) := \mathbb{E}[\varphi(X_n, X_{n+1})], \ \sigma(\varphi) := \sqrt{\text{Var}(\varphi(X_n, X_{n+1}))}.$$

Next we assume that  $K > 0$ ,  $\varepsilon \in (0, 1)$  and  $f = \{f_n^{(N)} : 1 \leq n \leq N < \infty\}$  is an array of measurable functions  $f_n^{(N)} : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \rightarrow \mathbb{R}$  which satisfy the following assumptions for all  $N$ :

(I)  $\mathbb{E}(f_n^{(N)}) = 0$  and  $\text{ess sup } |f| < K$ .  
 (II) Let  $S_N := \sum_{n=1}^N f_n^{(N)}(X_n, X_{n+1})$  and  $V_N := \text{Var}(S_N)$ , then there exists  $\widehat{C} > 0$  s.t.

$$V_N \rightarrow \infty \quad \text{and} \quad \frac{1}{V_N} \sum_{n=1}^N \sigma^2(f_n^{(N)}) \leq \widehat{C}. \quad (5.2.1)$$

(III)  $f = \mathbb{F} + h + c$ , where

(a)  $\mathbb{F} = \{\mathbb{F}_n^{(N)}\}$  are measurable functions such that  $\text{ess sup } |\mathbb{F}| \leq K$ ,  $G_{alg}(\mathbb{X}, \mathbb{F}) \subset \mathbb{Z}$ .  
 (b)  $h = \{h_n^{(N)}\}$  are measurable functions such that

$$\mathbb{E}(h_n^{(N)}) = 0, \quad \text{ess sup } |h| < K, \quad \sum_{n=1}^N \sigma^2(h_n^{(N)}) \leq \varepsilon.$$

(c)  $c = \{c_n^{(N)}\}$  are constants. Necessarily  $|c_n^{(N)}| \leq 3K$  and  $c_n^{(N)} = -\mathbb{E}(\mathbb{F}_n^{(N)})$ . Let  $c^{(N)} := \sum_{n=1}^N c_n^{(N)}$ .

We are *not* assuming that  $\mathbb{E}(\mathbb{F}_n^{(N)}) = 0$ :  $\mathbb{F}_n^{(N)}$  are integer valued, and we do not wish to destroy this by subtracting the mean.

**Lemma 5.4** *Under the above assumptions, for every  $\bar{K} > 0$ ,  $m \in \mathbb{Z}$ , there are  $\bar{C}, \bar{N} > 0$  s.t. for every  $N > \bar{N}$ ,  $|s| \leq \bar{K}$ ,  $x \in \mathfrak{S}_1$ , and  $v_{N+1} : \mathfrak{S}_{N+1} \rightarrow \mathbb{R}$  with  $\|v_{N+1}\|_\infty \leq 1$ ,*

$$\mathbb{E}_x \left( e^{i(2\pi m + \frac{s}{\sqrt{V_N}})S_N} v_{N+1}(X_{N+1}) \right) = e^{2\pi imc^{(N)}} e^{-s^2/2} \mathbb{E}(v_{N+1}(X_{N+1})) + \eta_N(x)$$

where  $\mathbb{E}(|\eta|) \leq \bar{C} \left[ \sum_{n=1}^N \sigma^2(h_n^{(N)}) \right]^{1/2} \leq \bar{C}\sqrt{\varepsilon}$ .

*Proof.* In this proof we fix the value of  $N$ , and drop the superscripts  $N$  for the ease of notation (for example  $c^{(N)} = c$ ).

We develop a perturbation theory of transfer operators similar to [12]. Recall the operators  $\mathcal{L}_{n,\xi} : L^\infty(\mathfrak{S}_{n+1}) \rightarrow L^\infty(\mathfrak{S}_n)$  given by

$$(\mathcal{L}_{n,\xi} u)(x) = \int_{\mathfrak{S}_{n+1}} p_n(x, y) e^{i\xi f_n(x, y)} u(y) \mu_{n+1}(dy).$$

Let  $\xi = \xi(m, s) := 2\pi m + \frac{s}{\sqrt{V_N}}$ . Since  $\mathbb{F}_n$  is integer valued,

$$e^{i\xi f_n} = \exp[2\pi im\mathbb{F}_n + \frac{is}{\sqrt{V_N}}\mathbb{F}_n + i\xi c_n + i\xi h_n] = e^{2\pi imc_n} e^{i\left(\frac{s}{\sqrt{V_N}}(\mathbb{F}_n + c_n) + \xi h_n\right)}.$$

We now split  $e^{-2\pi imc_n} \mathcal{L}_{n,\xi} = \overline{\mathcal{L}}_{n,\xi} + \widehat{\mathcal{L}}_{n,\xi} + \widetilde{\mathcal{L}}_{n,\xi}$  where

$$\begin{aligned} (\overline{\mathcal{L}}_{n,\xi} u)(x) &= \int_{\mathfrak{S}_{n+1}} p_n(x, y) e^{\frac{is}{\sqrt{V_N}}(\mathbb{F}_n(x, y) + c_n)} u(y) \mu_{n+1}(dy), \\ (\widehat{\mathcal{L}}_{n,\xi} u)(x) &= i\xi \int_{\mathfrak{S}_{n+1}} p_n(x, y) h_n(x, y) u(y) \mu_{n+1}(dy), \text{ and} \end{aligned}$$

$$\left(\widetilde{\mathcal{L}}_{n,\xi} u\right)(x) = \int_{\mathfrak{S}_{n+1}} p_n(x,y) \left[ e^{i\xi h_n + \frac{is}{\sqrt{V_N}}(\mathbb{F}_n(x,y) + c_n)} - e^{-\frac{is}{\sqrt{V_N}}(\mathbb{F}_n(x,y) + c_n)} - i\xi h_n(x,y) \right] u(y) \mu_{n+1}(dy).$$

We claim that there exists  $C_1(\bar{K}, m) > 1$  such that for  $|s| \leq \bar{K}$ ,  $n \geq 1$

$$\|\mathcal{L}_{n,\xi}\| := \|\mathcal{L}_{n,\xi}\|_{L^\infty \rightarrow L^\infty} \leq 1, \quad (5.2.2)$$

$$\|\mathcal{L}_{n,\xi}\|_{L^1 \rightarrow L^\infty} \leq C_1(\bar{K}, m), \quad (5.2.3)$$

$$\|\widetilde{\mathcal{L}}_{n,\xi}\| := \|\widetilde{\mathcal{L}}_{n,\xi}\|_{L^\infty \rightarrow L^\infty} \leq 1, \quad (5.2.4)$$

$$\|\widehat{\mathcal{L}}_{n,\xi}\|_{L^\infty \rightarrow L^1} \leq C_1(\bar{K}, m) \sigma(h_n), \quad (5.2.5)$$

$$\|\widetilde{\mathcal{L}}_{n,\xi}\|_{L^\infty \rightarrow L^1} \leq C_1(\bar{K}, m) \left[ \sigma^2(h_n) + \frac{\sigma(h_n)\sigma(f_n)}{\sqrt{V_N}} \right]. \quad (5.2.6)$$

To see this, we represent these operators as integral operators, and estimate their kernels. For example,  $\widehat{\mathcal{L}}_{n,\xi}$  is an integral operator whose kernel has absolute value  $|i\xi p_n(x,y)h_n(x,y)| \leq \varepsilon_0^{-1}|\xi||h_n(x,y)|$ . So

$$\|\widehat{\mathcal{L}}_{n,\xi}\|_{L^\infty \rightarrow L^1} \leq \varepsilon_0^{-1} \sqrt{4\pi^2 m^2 + \bar{K}^2} \|h_n\|_{L^1} \leq \varepsilon_0^{-1} \sqrt{4\pi^2 m^2 + \bar{K}^2} \|h_n\|_{L^2},$$

and (5.2.5) follows from the identity  $\|h_n\|_{L^2} \equiv \sigma(h_n)$ . Similarly,  $\widetilde{\mathcal{L}}_{n,\xi}$  has kernel with absolute value

$$\begin{aligned} p_n(x,y) \left| e^{i\xi h_n + is\frac{\mathbb{F}_n(x,y) + c_n}{\sqrt{V_N}}} - e^{-is\frac{\mathbb{F}_n(x,y) + c_n}{\sqrt{V_N}}} - i\xi h_n(x,y) \right| &\leq \varepsilon_0^{-1} \left| e^{is\frac{\mathbb{F}_n(x,y) + c_n}{\sqrt{V_N}}} (e^{i\xi h_n} - 1) - i\xi h_n \right| \\ &= \varepsilon_0^{-1} \left| e^{is\frac{\mathbb{F}_n(x,y) + c_n}{\sqrt{V_N}}} (i\xi h_n + O(\xi^2 h_n^2)) - i\xi h_n \right| = \varepsilon_0^{-1} \left| e^{is\frac{\mathbb{F}_n(x,y) + c_n}{\sqrt{V_N}}} - 1 \right| |\xi h_n| + O(h_n^2) \\ &= O\left(\frac{1}{\sqrt{V_N}} |h_n(\mathbb{F}_n + c_n)|\right) + O(h_n^2) \end{aligned}$$

where the implicit constants in  $O(\cdot)$  are uniform on compact sets of  $\xi$ . It follows that uniformly on compact sets of  $\xi$ ,

$$\begin{aligned} \|\widetilde{\mathcal{L}}_{n,\xi}\|_{L^\infty \rightarrow L^1} &= O(V_N^{-1/2}) \mathbb{E}(|h_n(\mathbb{F}_n + c_n)|) + O(\|h_n\|_2^2) \\ &= O(V_N^{-1/2}) \|h_n\|_2 \|\mathbb{F}_n + c_n\|_2 + O(\|h_n\|_2^2) \\ &= O(V_N^{-1/2}) \|h_n\|_2 (\|f_n - h_n\|_2) + O(\|h_n\|_2^2) \\ &= O(V_N^{-1/2}) \|h_n\|_2 (\|f_n\|_2 + \|h_n\|_2) + O(\|h_n\|_2^2) \\ &= O\left(\frac{\|h_n\|_2 \|f_n\|_2}{\sqrt{V_N}} + \|h_n\|_2^2\right) = O\left(\frac{\sigma(h_n)\sigma(f_n)}{\sqrt{V_N}} + \sigma^2(h_n)\right), \end{aligned}$$

as claimed in (5.2.6).

Recall Nagaev's identity (4.2.1):  $\mathbb{E}_x[e^{i\xi S_N} v_{N+1}(X_{N+1})] = (\mathcal{L}_{1,\xi} \mathcal{L}_{2,\xi} \cdots \mathcal{L}_{N,\xi} v_{N+1})(x)$ . The decomposition  $e^{-2\pi imc_n} \mathcal{L}_{n,\xi} = \widehat{\mathcal{L}}_{n,\xi} + \widetilde{\mathcal{L}}_{n,\xi} + \widehat{\mathcal{L}}_{n,\xi}$  implies that

$$\mathbb{E}_x \left( e^{i\xi S_N} v_{N+1}(X_{N+1}) \right) = e^{2\pi imc} \left( \overline{\Phi}_N(x, \xi) + \widehat{\Phi}_N(x, \xi) + \widetilde{\Phi}_N(x, \xi) \right) \quad (5.2.7)$$

where  $c = c^{(N)} = c_1 + \dots + c_N$ , and

$$\begin{aligned}\overline{\Phi}_N(x, \xi) &:= (\overline{\mathcal{L}}_{1,\xi} \dots \overline{\mathcal{L}}_{N,\xi} v_{N+1})(x), \\ \widetilde{\Phi}_N(x, \xi) &:= \sum_{k=1}^{N-1} e^{-2\pi i m(c_1 + \dots + c_{k-1})} \left( \mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widetilde{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} \dots \overline{\mathcal{L}}_{N,\xi} v_{N+1} \right)(x), \\ \widehat{\Phi}_N(x, \xi) &:= \sum_{k=1}^{N-1} e^{-2\pi i m(c_1 + \dots + c_{k-1})} \left( \mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} \dots \overline{\mathcal{L}}_{N,\xi} v_{N+1} \right)(x).\end{aligned}$$

We will analyze each of these summands.

CLAIM 1: For every  $m \in \mathbb{Z}$ ,  $\left| \overline{\Phi}_N(x, \xi) - e^{-s^2/2} \mathbb{E}_x(v_{N+1}(X_{N+1})) \right| \xrightarrow[N \rightarrow \infty]{} 0$  uniformly in  $s$  on  $\{s \in \mathbb{R} : |s| \leq \bar{K}\}$ ,  $x \in \mathfrak{S}_1$ ,  $v_{N+1} \in \{v \in L^\infty(\mathfrak{S}_{N+1}) : \|v\| \leq 1\}$ .

PROOF:  $\overline{\Phi}_N(x, \xi) = \mathbb{E}_x \left( \exp \left( is \frac{\sum_{k=1}^N \mathbb{F}_k + c}{\sqrt{V_N}} \right) v_{N+1}(X_{N+1}) \right)$ , where  $\mathbb{E}(\sum_{k=1}^N \mathbb{F}_k) = -c$ . Fix  $1 \leq r \leq N$ . Using the decomposition  $f = \mathbb{F} + \mathbb{h} + \mathbb{c}$ , we find that

$$\frac{1}{\sqrt{V_N}} \left( \sum_{k=1}^N \mathbb{F}_k + c \right) = \frac{1}{\sqrt{V_N}} \sum_{k=1}^{N-r} f_k + \frac{1}{\sqrt{V_N}} \left( O(r) - \sum_{k=1}^N h_k \right).$$

By assumption III(b), the  $L^2$  norm of the second summand is  $O(1/\sqrt{V_N})$ . Therefore the second term converges to 0 in probability as  $N \rightarrow \infty$ , and

$$\overline{\Phi}_N(x, \xi) = \mathbb{E}_x \left( e^{\frac{is}{\sqrt{V_N}} S_{N-r}} v_{N+1}(X_{N+1}) \right) + o(1), \quad (5.2.8)$$

where we have abused notation and wrote  $S_{N-r} = f_1^{(N)} + \dots + f_{N-r}^{(N)}$ .

The rate of convergence to 0 depends on  $r$  and  $m$ , but is uniform when  $|s| \leq \bar{K}$  and  $\|v_{N+1}\|_\infty \leq 1$ . At the same time, by exponential mixing (see (1.2.3)), there is  $0 < \theta < 1$  such that

$$\begin{aligned}\mathbb{E}_x \left( e^{\frac{is}{\sqrt{V_N}} S_{N-r}} v_{N+1}(X_{N+1}) \right) &= \mathbb{E}_x \left[ e^{\frac{is}{\sqrt{V_N}} S_{N-r}} \mathbb{E}_x(v_{N+1}(X_{N+1}) | X_1, \dots, X_{N-r}) \right] \\ &= \mathbb{E}_x \left[ e^{\frac{is}{\sqrt{V_N}} S_{N-r}} \mathbb{E}_x(v_{N+1}(X_{N+1}) | X_{N-r}) \right] \text{ (Markov property)} \\ &= \mathbb{E}_x \left( e^{\frac{is}{\sqrt{V_N}} S_{N-r}} [\mathbb{E}_x(v_{N+1}(X_{N+1}) + O(\theta^r))] \right) \quad \text{(exponential mixing)} \\ &= \mathbb{E}_x(e^{isS_{N-r}/\sqrt{V_N}}) \mathbb{E}_x(v_{N+1}(X_{N+1})) + O(\theta^r) \quad (5.2.9)\end{aligned}$$

where the  $O(\theta^r)$  is uniform in  $\|v_{N+1}\|_\infty$ .

A similar mixing argument shows that

$$\mathbb{E}_x(S_{N-r}) = \mathbb{E}(S_{N-r} | X_1 = x) = \mathbb{E}(S_N) + O(1) = O(1)$$

uniformly in  $x \in \mathfrak{S}_1$ . By Dobrushin's CLT,

$$\mathbb{E}_x(e^{isS_{N-r}/\sqrt{V_N}}) = [1 + o(1)] \mathbb{E}_x(e^{\frac{isS_{N-r} - \mathbb{E}_x(S_N)}{\sqrt{V_N}}}) = [1 + o(1)] e^{-s^2/2} \text{ as } N \rightarrow \infty.$$

The claim follows from this, (5.2.8), and (5.2.9).

CLAIM 2. *There exists  $C_2(\bar{K}, m)$  s.t. for all  $|s| \leq \bar{K}$  and  $\|v_{N+1}\|_\infty \leq 1$ ,*

$$\|\tilde{\Phi}_N(x, \xi)\|_{L^1} \leq C_2(\bar{K}, m)\sqrt{\epsilon}.$$

PROOF:  $\|\tilde{\Phi}_N(x, \xi)\|_1 \leq \|\tilde{\mathcal{L}}_{1,\xi}\|_{L^\infty \rightarrow L^1} \|\tilde{\mathcal{L}}_{2,\xi}\| \cdots \|\tilde{\mathcal{L}}_{N,\xi}\|$

$$+ \sum_{k=2}^N \left( \|\mathcal{L}_{1,\xi}\| \cdots \|\mathcal{L}_{k-1,\xi} \tilde{\mathcal{L}}_{k,\xi}\| \|\tilde{\mathcal{L}}_{k+1,\xi}\| \cdots \|\tilde{\mathcal{L}}_{N,\xi}\| \right)$$

Suppose  $|s| \leq \bar{K}$ , then (5.2.3), (5.2.4) and (5.2.6) tell us that

$$\begin{aligned} \|\tilde{\mathcal{L}}_{1,\xi}\|_{L^\infty \rightarrow L^1} \|\tilde{\mathcal{L}}_{2,\xi}\| \cdots \|\tilde{\mathcal{L}}_{N,\xi}\| &\leq C_1(\bar{K}, m) \left[ \sigma(h_1)^2 + \frac{\sigma(h_1)\sigma(f_1)}{\sqrt{V_N}} \right], \\ \|\mathcal{L}_{k-1,\xi} \tilde{\mathcal{L}}_{k,\xi}\| &\leq \|\mathcal{L}_{k-1,\xi}\|_{L^1 \rightarrow L^\infty} \|\tilde{\mathcal{L}}_{k,\xi}\|_{L^\infty \rightarrow L^1} \leq C_1(\bar{K}, m)^2 \left[ \sigma(h_k)^2 + \frac{\sigma(h_k)\sigma(f_k)}{\sqrt{V_N}} \right]. \end{aligned}$$

Therefore  $\|\tilde{\Phi}_N(x, \xi)\|_1 \leq C_1(\bar{K}, m)^2 \sum_{k=1}^{N-1} \left[ \sigma(h_k)^2 + \frac{\sigma(h_k)\sigma(f_k)}{\sqrt{V_N}} \right]$ . By Cauchy-Schwarz,

$$\begin{aligned} \sum_{k=1}^{N-1} \sigma(h_k)^2 + \frac{\sigma(h_k)\sigma(f_k)}{\sqrt{V_N}} &\leq \sum_{k=1}^{N-1} \sigma(h_k)^2 + \sqrt{\sum_{k=1}^{N-1} \sigma^2(h_k) \cdot \frac{1}{V_N} \sum_{k=1}^{N-1} \sigma^2(f_k)} \\ &\leq \epsilon + \sqrt{\hat{C}\epsilon}, \text{ by assumptions II and III(b). The claim follows.} \end{aligned}$$

CLAIM 3. *There exists  $C_3(\bar{K}, m)$  s.t. for all  $|s| \leq \bar{K}$ , and  $\|v_{N+1}\|_\infty \leq 1$ ,*

$$\|\hat{\Phi}_N(x, \xi)\|_1 \leq C_3(\bar{K}, m)\sqrt{\epsilon}.$$

PROOF. Fix  $N$ ,  $v_{N+1} \in L^\infty(\mathfrak{S}_{N+1})$  such that  $\|v_{N+1}\|_\infty \leq 1$ , and define  $\zeta_k \in L^\infty(\mathfrak{S}_k)$ ,  $\eta_k \in \mathbb{R}$  s.t.

$$\phi_k(\cdot) := (\tilde{\mathcal{L}}_{k,\xi} \cdots \tilde{\mathcal{L}}_{N,\xi}) v_{N+1} = \zeta_k(\cdot) + \eta_k$$

where  $\eta_k := \mathbb{E}[(\tilde{\mathcal{L}}_{k,\xi} \cdots \tilde{\mathcal{L}}_{N,\xi}) v_{N+1}(X_k)]$ , and  $\mathbb{E}[\zeta_k(X_k)] = 0$ . Then

$$\|\hat{\Phi}_N(x, \xi)\|_1 \leq \sum_{k=1}^N \|\mathcal{L}_{1,\xi} \cdots \mathcal{L}_{k-1,\xi} \tilde{\mathcal{L}}_{k,\xi}(\zeta_{k+1} + \eta_{k+1} 1)\|_1. \quad (5.2.10)$$

By (5.2.4),  $|\eta_k| \leq 1$ . We will now work towards a control of  $\zeta_k$ :

SUB-CLAIM. *We can decompose  $\zeta_k = \zeta'_k + \zeta''_k$  so that for all  $|s| \leq \bar{K}$ , there exist  $\hat{C}_0, \hat{K}_0 > 0$  and  $0 < \hat{\theta}_0 < 1$  s.t. for all  $k = 1, \dots, N-2$*

$$\|\zeta'_k\|_\infty \leq \hat{\theta}_0^2 \|\zeta'_{k+2}\|_\infty + \hat{K}_0 \|\zeta''_{k+2}\|_1, \quad (5.2.11)$$

$$\|\zeta''_k\|_\infty \leq \hat{C}_0 \left( \frac{\sigma(f_k) + \sigma(f_{k+1}) + \sigma(h_k) + \sigma(h_{k+1})}{\sqrt{V_N}} \right). \quad (5.2.12)$$

*Proof.* In what follows,  $\mathcal{L}_k = \mathcal{L}_{k,0}$ . Write

$$\eta_k + \zeta_k = \phi_k = (\tilde{\mathcal{L}}_{k,\xi} \tilde{\mathcal{L}}_{k+1,\xi}) \phi_{k+2} = (\tilde{\mathcal{L}}_{k,\xi} \tilde{\mathcal{L}}_{k+1,\xi}) (\eta_{k+2} + \zeta_{k+2})$$

$$= (\mathcal{L}_k \mathcal{L}_{k+1}) \eta_{k+2} + (\mathcal{L}_k \mathcal{L}_{k+1}) \zeta_{k+2} + (\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1}) \phi_{k+2}.$$

Observe that  $\mathcal{L}_k 1 = 1$ , so  $(\mathcal{L}_k \mathcal{L}_{k+1}) \eta_{k+2} = \eta_{k+2}$ . This leads to the decomposition

$$\zeta_k = \underbrace{(\mathcal{L}_k \mathcal{L}_{k+1}) \zeta_{k+2}}_{\zeta'_k} + \underbrace{(\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1}) \phi_{k+2} + \eta_{k+2} - \eta_k}_{\zeta''_k}$$

We use this decomposition to define  $\zeta'_k, \zeta''_k$ . This gives the following recursion:

$$\begin{aligned} \zeta'_k &= (\mathcal{L}_k \mathcal{L}_{k+1}) \zeta'_{k+2} + (\mathcal{L}_k \mathcal{L}_{k+1}) \zeta''_{k+2}, \\ \zeta''_k &= (\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1}) \phi_{k+2} + \eta_{k+2} - \eta_k. \end{aligned} \quad (5.2.13)$$

Notice that  $\zeta'_k, \zeta''_k$  both have zero means. Indeed in our setup,  $\mu_j(E) = \mathbb{P}(X_j \in E)$  and  $(\mathcal{L}_k u)(x) = \mathbb{E}(u(X_{k+1})|X_k = x)$ , whence

$$\int \zeta'_k d\mu_k = \mathbb{E}(\zeta'_k(X_k)) = \mathbb{E}[\mathbb{E}(\mathbb{E}(\zeta_{k+2}(X_{k+2})|X_{k+1})|X_k)] = \mathbb{E}(\zeta_{k+2}(X_{k+2})) = 0,$$

and  $\mathbb{E}(\zeta''_k) = \mathbb{E}(\zeta_k) - \mathbb{E}(\zeta'_k) = 0 - 0 = 0$ .

To prove the estimates on  $\|\zeta'_k\|_\infty$ , we first make the following general observations. If  $\psi_{k+2} \in L^\infty(\mathfrak{S}_{k+2})$ , then  $(\mathcal{L}_k \mathcal{L}_{k+1} \psi_{k+2})(x) = \int \tilde{p}(x, z) \psi_{k+2}(z) \mu_{k+2}(dz)$ , where

$$\tilde{p}(x, z) = \int_{\mathfrak{S}_{k+1}} p_k(x, y) p_{k+1}(y, z) \mu_{k+1}(dy).$$

By uniform ellipticity,  $\tilde{p} \geq \varepsilon_0$  so we can decompose  $\tilde{p}_k = \varepsilon_0 + (1 - \varepsilon_0) \tilde{q}_k$  where  $\tilde{q}_k$  is a probability density. Hence if  $\psi_{k+2}$  has zero mean then

$$\begin{aligned} (\mathcal{L}_k \mathcal{L}_{k+1} \psi_{k+2})(x) &= \varepsilon_0 \int \psi_{k+2} d\mu_{k+2} + (1 - \varepsilon_0) \int \tilde{q}_k(x, y) \psi_{k+2}(y) \mu_{k+2}(dy) \\ &= (1 - \varepsilon_0) \int \tilde{q}_k(x, y) \psi_{k+2}(y) \mu_{k+2}(dy). \end{aligned}$$

Thus  $\|\mathcal{L}_k \mathcal{L}_{k+1} \psi_{k+2}\|_\infty \leq (1 - \varepsilon_0) \|\psi_{k+2}\|_\infty$ .

We apply this to  $\zeta'_{k+2} = (\mathcal{L}_k \mathcal{L}_{k+1}) \zeta'_{k+2} + (\mathcal{L}_k \mathcal{L}_{k+1}) \zeta''_{k+2}$ :

$$\begin{aligned} \|\zeta'_k\|_\infty &\leq (1 - \varepsilon_0) \|\zeta'_k\|_\infty + \|\mathcal{L}_k \mathcal{L}_{k+1} \zeta''_{k+2}\|_\infty \\ &\leq (1 - \varepsilon_0) \|\zeta'_k\|_\infty + \|\mathcal{L}_k\|_{L^1 \rightarrow L^\infty} \|\mathcal{L}_{k+1}\|_{L^1 \rightarrow L^1} \|\zeta''_{k+2}\|_1 \leq (1 - \varepsilon_0) \|\zeta'_k\|_\infty + \varepsilon_0^{-2} \|\zeta''_{k+2}\|_1. \end{aligned}$$

The last step is because  $0 \leq p_n(x, y) \leq \varepsilon_0^{-1}$ . This proves (5.2.11).

Next we analyze  $\|\zeta''_k\|_\infty$ . Since  $\zeta''_k$  has zero mean and  $\eta_{k+2} - \eta_k$  is constant, we can write  $\zeta''_k = \widehat{\zeta}''_k - \mathbb{E}(\widehat{\zeta}''_k)$  with  $\widehat{\zeta}''_k := (\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1}) \phi_{k+2}$ . Observe that the kernel of  $(\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1})$  is bounded by

$$\text{const} \frac{|s|}{\sqrt{V_N}} \int (\mathbb{F}_k(x, z) + \mathbb{F}_{k+1}(z, y) + c_k + c_{k+1}) \mu_{k+1}(dz).$$

By assumptions II and III, the  $L^1$ -norm of the kernel is bounded by

$$O\left(\frac{|s|}{\sqrt{V_N}}\right) \left( \|f_k - h_k\|_1 + \|f_{k+1} - h_{k+1}\|_1 \right) = O\left(\frac{|s|}{\sqrt{V_N}}\right) \left( \|f_k\|_1 + \|h_k\|_1 + \|f_{k+1}\|_1 + \|h_{k+1}\|_1 \right)$$

$$\leq O\left(\frac{|s|}{\sqrt{V_N}}\right) \left( \|f_k\|_2 + \|h_k\|_2 + \|f_{k+1}\|_2 + \|h_{k+1}\|_2 \right).$$

This implies that  $\|\widehat{\zeta}_k''\|_\infty = O\left(\frac{|s|}{\sqrt{V_N}}\right) (\sigma(f_k) + \sigma(f_{k+1}) + \sigma(h_k) + \sigma(h_{k+1}))$ , whence  $\|\zeta_k''\|_\infty \leq 2\|\widehat{\zeta}_k''\|_\infty = O\left(\frac{|s|}{\sqrt{V_N}}\right) (\sigma(f_k) + \sigma(f_{k+1}) + \sigma(h_k) + \sigma(h_{k+1}))$ . (5.2.12) and the sub-claim are proved.

We return to the proof of Claim 3. Iterating the estimate in the sub-claim, we conclude that for some constant  $\bar{C}$

$$\begin{aligned} \|\zeta_k'\|_\infty &\leq \bar{C} \left[ \widehat{\theta}_0^{2\lfloor \frac{N-k}{2} \rfloor} + \sum_{r=1}^{\lfloor \frac{N-k}{2} \rfloor - 1} \frac{\widehat{\theta}_0^{2r} (\sigma(f_{k+2r}) + \sigma(f_{k+2r+1}) + \sigma(h_{k+2r}) + \sigma(h_{k+2r+1}))}{\sqrt{V_N}} \right] \\ &\leq \bar{C} \widehat{\theta}_0^{-1} \left[ \widehat{\theta}_0^{N-k} + \sum_{r=1}^{N-k} \frac{\widehat{\theta}_0^r}{\sqrt{V_N}} (\sigma(f_{k+r}) + \sigma(h_{k+r})) \right]. \end{aligned}$$

Since  $\mathcal{L}_{j,\xi}$  are contractions and  $\|\widehat{\mathcal{L}}_{k,\xi}\|_{L^\infty \rightarrow L^1} \leq C_1(\bar{K}, m) \sigma(h_k)$ , this implies that

$$\begin{aligned} \sum_k \|\mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi}(\zeta_{k+1}')\|_{L^1} \\ \leq \bar{C} C_1(\bar{K}, m) \left[ \sum_r \widehat{\theta}_0^r \sum_k \sigma(h_k) \frac{\sigma(f_{k+r}) + \sigma(h_{k+r})}{\sqrt{V_N}} + \sum_k \sigma(h_k) \widehat{\theta}_0^{N-k} \right]. \end{aligned}$$

As in the proof of Claim 2, it follows from the Cauchy Schwartz inequality, (5.2.1), and assumption III(b) that the sum over  $k$  is  $O(\sqrt{\varepsilon})$ . Hence

$$\sum_k \|\mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi}(\zeta_{k+1}')\|_{L^1} = O(\sqrt{\varepsilon}). \quad (5.2.14)$$

Next we claim that

$$\sum_k \|\mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi}(\zeta_{k+1}'')\|_{L^1} = O(\sqrt{\varepsilon}). \quad (5.2.15)$$

The proof is similar to the proof of (5.2.14), except that now we use (5.2.13) to see that as in the proofs of (5.2.5), (5.2.6) and (5.2.12),

$$\|\mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi}(\zeta_{k+1}'')\|_{L^1} \leq C_4(\bar{K}, m) \sigma(h_k) \frac{\sigma(f_{k+1}) + \sigma(f_{k+2}) + \sigma(h_{k+1}) + \sigma(h_{k+2})}{\sqrt{V_N}}$$

for some constant  $C_4(\bar{K}, m)$ .

(5.2.14) and (5.2.15) give us an  $O(\sqrt{\varepsilon})$  bound for contribution of  $\zeta_{k+1}$  to (5.2.10). It remains to estimate the contribution of  $\eta_{k+1}$  to (5.2.10).

Split  $\mathcal{L}_{n,\xi} = e^{2\pi imc_n} \mathcal{L}_n + \mathcal{L}'_{n,\xi}$ . As before,

$$\begin{aligned} \mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi}(1) &= e^{2\pi im(c_1 + \dots + c_{k-1})} \mathcal{L}_1 \dots \mathcal{L}_{k-1} \widehat{\mathcal{L}}_{k,\xi}(1) \\ &+ \sum_j e^{2\pi im(c_{j+1} + \dots + c_{k-1})} \mathcal{L}_{1,\xi} \dots \mathcal{L}_{j-1,\xi} \mathcal{L}'_{j,\xi} \mathcal{L}_{j+1} \dots \mathcal{L}_{k-1} \widehat{\mathcal{L}}_{k,\xi}(1). \end{aligned} \quad (5.2.16)$$

Since  $\mathbb{E}(h_k) = 0$ ,  $\mathbb{E}[(\widehat{\mathcal{L}}_{k,\xi} 1)(X_k)] = 0$ . By exponential mixing (1.2.3), the first term on the RHS of (5.2.16) has  $L^\infty$  (whence  $L^1$ ) norm no larger than

$$C_{mix} \theta^{k-1} \|\widehat{\mathcal{L}}_{k,\xi} 1\|_\infty \leq \widehat{C}_3 \sigma(h_k) \theta^k$$

for some constant  $\widehat{C}_3 = \widehat{C}_3(\bar{K}, m)$  and  $0 < \theta < 1$ . Similarly each summand in the second term on the RHS of (5.2.16) has  $L^1$  norm less than

$$\begin{aligned} \|\mathcal{L}'_{j,\xi} \|\widehat{C}_3 \sigma(h_k) \theta^{k-j} &\leq \widehat{C}_4 \sigma(h_k) \theta^{k-j} \left\| \frac{s}{\sqrt{V_N}} (\mathbb{F}_j + c_j) + \xi h_j \right\|_2 \\ &\leq \widehat{C}_4 \sigma(h_k) \theta^{k-j} \left( \frac{\sigma(f_j)}{\sqrt{V_N}} + \sigma(h_j) \right), \end{aligned}$$

for  $\widehat{C}_4 = \widehat{C}_4(\bar{K}, m)$ . So the second term on the RHS of (5.2.16) has norm less than

$$\widehat{C}_5 \sigma(h_k) \sum_{j=1}^k \theta^{k-j} \left( \frac{\sigma(f_j)}{\sqrt{V_N}} + \sigma(h_j) \right) \quad (5.2.17)$$

for some constant  $\widehat{C}_5$ .

It follows that  $\sum_k \|\mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi}(1)\|_1$  is bounded by

$$\begin{aligned} &\sum_{k=1}^N \left( \widehat{C}_3 \sigma(h_k) \theta^k + \widehat{C}_5 \sigma(h_k) \sum_{j=1}^k \theta^{k-j} \left( \frac{\sigma(f_j)}{\sqrt{V_N}} + \sigma(h_j) \right) \right) \\ &\leq \widehat{C}_3 \sqrt{\sum_{k=1}^N \sigma^2(h_k)} \sqrt{\sum_{k=1}^N \theta^{2k}} + \widehat{C}_5 \sum_{r=0}^{N-1} \theta^r \sum_{j=1}^N \left( \frac{\sigma(f_j)}{\sqrt{V_N}} + \sigma(h_j) \right) \sigma(h_{j+r}) \\ &\leq \frac{\widehat{C}_3 \sqrt{\varepsilon}}{\sqrt{1-\theta^2}} + \widehat{C}_5 \sum_{r=0}^{N-1} \theta^r \left[ \left( \sqrt{\sum_{j=1}^N \frac{\sigma^2(f_j)}{V_N}} + \sqrt{\sum_{j=1}^N \sigma^2(h_j)} \right) \sqrt{\sum_{j=1}^N \sigma^2(h_{j+r})} \right]. \end{aligned}$$

By assumptions II and III, there is a constant  $\widehat{C}_6 = \widehat{C}_6(\bar{K}, m)$  such that

$$\sum_k \|\mathcal{L}_{1,\xi} \dots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi}(1)\|_1 \leq \widehat{C}_6 \sqrt{\varepsilon}. \quad (5.2.18)$$

Claim 3 now follows from (5.2.10), (5.2.14), (5.2.15), and (5.2.18).

Lemma 5.4 now follows from Claims 1–3 and (5.2.7).  $\square$

### 5.2.2 Proof of the LLT in the reducible case

**Setup and reductions.** Let  $\mathbf{f} = \{f_n\}$  be an a.s. uniformly bounded additive functional on a Markov chain  $\mathbf{X} = \{X_n\}$  with state spaces  $\mathfrak{S}_n$  and marginals  $\mu_n(E) = \mathbb{P}(X_n \in E)$ . We assume that  $\mathbf{f}$  is not center-tight, and that  $\mathbf{f}$  is reducible. In this case  $G_{ess}(\mathbf{X}, \mathbf{f}) = \delta(\mathbf{f})\mathbb{Z}$  with some  $\delta(\mathbf{f}) > 0$ . Without loss of generality,

$$\delta(f) = 1, G_{ess}(f) = \mathbb{Z}, \mathbb{E}(f_n) := \mathbb{E}[f_n(X_n, X_{n+1})] = 0 \text{ for all } n,$$

otherwise we center and rescale  $f$ .

By the reduction lemma (Lemma 3.10),  $f = \mathbb{F} + \nabla a + h + c$ , where

$$G_{alg}(\mathbb{X}, \mathbb{F}) = G_{ess}(\mathbb{X}, \mathbb{F}) = \mathbb{Z},$$

$h$  has summable variances and  $\mathbb{E}(h_n) := \mathbb{E}(h_n(X_n, X_{n+1})) = 0$ ,  $c = \{c_n\}$  are constants, and  $\mathbb{F}, a, h, c$  are a.s. uniformly bounded. There is no loss of generality in assuming that  $a \equiv 0$ , because Theorem 5.1 holds for  $f$  with  $b_N$  iff Theorem 5.1 holds for  $f - \nabla a$  with  $b'_N(X_1, X_{N+1}) := b_N(X_1, X_{N+1}) + a_{N+1}(X_{N+1}) - a_1(X_1)$ .

Henceforth we assume  $f = \mathbb{F} + h + c$ , and  $\mathbb{E}(f_n) = \mathbb{E}(h_n) = 0$ . So  $c_n = -\mathbb{E}(\mathbb{F}_n)$ . Let

$$c(N) := -\sum_{k=1}^N \mathbb{E}[\mathbb{F}_k(X_k, X_{k+1})]. \quad (5.2.19)$$

By Theorem 2.4, the following sum converges a.s.:

$$\mathfrak{H}(X_1, X_2, \dots) := \sum_{n=1}^{\infty} h_n(X_n, X_{n+1}).$$

**Lemma 5.5** *Under the previous assumptions, for every sequence of non-negative functions  $v_{N+1} \in L^\infty(\mathfrak{S}_{N+1})$  s.t.  $\|v_{N+1}\|_\infty \neq 0$  and for some  $\bar{\delta} > 0$*

$$\int_{\mathfrak{S}_{N+1}} v_{N+1} d\mu_{N+1} \geq \bar{\delta} \|v_{N+1}\|_\infty, \quad (5.2.20)$$

for all  $m \in \mathbb{Z}$ ,  $s \in \mathbb{R}$  and  $x \in \mathfrak{S}_1$ ,

$$\frac{\mathbb{E}_x \left( e^{i(2\pi m + \frac{s}{\sqrt{N}})S_N} v_{N+1}(X_{N+1}) \right)}{\mathbb{E}(v_{N+1}(X_{N+1}))} = e^{2\pi i m c(N) - s^2/2} \mathbb{E}_x \left( e^{2\pi m i \xi} \right) + o_{N \rightarrow \infty}(1). \quad (5.2.21)$$

where  $o(\cdot)$  term converges to 0 uniformly when  $|m + is|$  are bounded,  $v_{N+1}$  are bounded, and (5.2.20) holds.

*Proof.* Since the LHS of (5.2.21) remains unchanged upon multiplying  $v_{N+1}$  by a constant, we may assume that  $\|v_{N+1}\|_\infty = 1$ .

Fix  $\varepsilon > 0$  small and  $r$  so large that  $\sum_{k=r}^{\infty} \text{Var}(h_k) < \varepsilon$ . Fix  $N$ . Applying the Integer Reduction Lemma (Lemma 3.13) to  $\{\mathbb{F}_n\}_{n=r}^N$ , we obtain a decomposition

$$\mathbb{F}_n(x_n, x_{n+1}) = \mathfrak{a}_{n+1}^{(N)}(x_{n+1}) - \mathfrak{a}_n^{(N)}(x_n) + \mathfrak{c}_n^{(N)} + \tilde{f}_n^{(N)}(x_n, x_{n+1})$$

where  $\mathfrak{c}_n^{(N)}$  are bounded integers, and  $\mathfrak{a}_n^{(N)}(\cdot)$ ,  $\tilde{f}_n^{(N)}(\cdot, \cdot)$  are uniformly bounded measurable integer valued functions such that

$$\sum_{n=r}^N \|\tilde{f}_n^{(N)}\|_2^2 = O \left( \sum_{n=r}^N u_n^2(\mathbb{F}) \right).$$

There is no loss of generality in assuming that  $\mathfrak{a}_{N+1}^{(N)} = \mathfrak{a}_r^{(N)} = 0$ , otherwise replace  $\tilde{f}_r^{(N)}(x, y)$  by  $\tilde{f}_r^{(N)}(x, y) - \mathfrak{a}_r^{(N)}(x)$ , and  $\tilde{f}_N^{(N)}(x, y)$  by  $\tilde{f}_N^{(N)}(x, y) + \mathfrak{a}_{N+1}^{(N)}(y)$ . Then  $\sum_{n=r}^N \mathbb{F}_n = \sum_{n=r}^N (\mathfrak{c}_n^{(N)} + \tilde{f}_n^{(N)})$ , whence

$$S_N - S_{r-1} = \sum_{n=r}^N f_n = \sum_{n=r}^N \mathfrak{c}_n^{(N)} + \tilde{f}_n^{(N)} + h_n + c_n = \sum_{n=r}^N \tilde{f}_n^{(N)} + h_n - \mathbb{E}(\tilde{f}_n^{(N)} + h_n). \quad (5.2.22)$$

(The last equality is because  $\mathbb{E}(S_N - S_{r-1}) = 0$ .)

Let  $\mathbf{g}$  denote the array with rows  $g_n^{(N)} := \tilde{f}_n^{(N)} + h_n - \mathbb{E}(\tilde{f}_n^{(N)} + h_n)$  ( $n = r, \dots, N$ ),  $N > r$ . We claim that  $\mathbf{g}$  satisfies assumptions (I)–(III) of Lemma 5.4. (I) is clear, and (III) holds by choice of  $r$  and because  $\tilde{f}_n^{(N)}$  is integer valued. To see (II), note that

$$\begin{aligned} \sum_{n=1}^N \sigma^2(g_n^{(N)}) &= \sum_{n=1}^N \sigma^2(\tilde{f}_n^{(N)} + h_n) = \sum_{n=1}^N \sigma^2(\tilde{f}_n^{(N)}) + \sigma^2(h_n) + 2\text{Cov}(\tilde{f}_n^{(N)}, h_n) \\ &\leq \sum_{n=1}^N \sigma^2(\tilde{f}_n^{(N)}) + \sigma^2(h_n) + 2\sigma(\tilde{f}_n^{(N)})\sigma(h_n) \leq 2 \sum_{n=1}^N \sigma^2(\tilde{f}_n^{(N)}) + \sigma^2(h_n) \quad (\because 2ab \leq a^2 + b^2) \\ &= O\left(\sum_{n=r}^N u_n^2(\mathbb{F})\right) + O(1), \text{ by choice of } \tilde{\mathbf{f}} \text{ and } \mathbf{h}. \end{aligned}$$

Since  $\mathbf{f} = \mathbb{F} + \mathbf{h} + \mathbf{c}$ ,  $u_n^2(\mathbb{F}) = u_n^2(\mathbf{f} + \mathbf{h}) \leq 2[u_n^2(\mathbf{f}) + u_n^2(\mathbf{h})]$ , see Lemma 1.15(4). Thus by Theorem 2.7 and the assumption that  $\mathbf{h}$  has summable variances,

$$\sum_{n=r}^N u_n^2(\mathbb{F}) \leq 2 \sum_{n=r}^N u_n^2(\mathbf{f}) + u_n^2(\mathbf{h}) = O(V_N) + O(1) = O(V_N).$$

Assumption (II) is checked.

We now apply Lemma 5.4 to  $\mathbf{g}$ , and deduce that for every  $\bar{K} > 0$  and  $m \in \mathbb{Z}$  there are  $\bar{C}, \bar{N} > 0$  such that for all  $N > \bar{N} + r$ ,  $|s| \leq \bar{K}$ , and  $v_{N+1}$  in the unit ball of  $L^\infty$

$$\begin{aligned} &\mathbb{E}\left(e^{i(2\pi m + \frac{s}{\sqrt{V_N}})(S_N - S_{r-1})} v_{N+1}(X_{N+1}) \middle| X_r\right) \\ &= e^{2\pi imc^{(N)}} \cdot e^{-s^2/2} \mathbb{E}(v_{N+1}(X_{N+1})) + \eta_{N-r}(X_r), \end{aligned}$$

where  $c^{(N)} := -\sum_{n=r}^N \mathbb{E}(\tilde{f}_n^{(N)})$  and  $\|\eta_{N-r}\|_1 \leq \bar{C}\sqrt{\epsilon}$ . Since  $\|v_{N+1}\|_\infty = 1$ , we also have the trivial bound  $\|\eta_{N-r}\|_\infty \leq 2$ .

We are ready to prove the lemma. The left-hand-side of (5.2.21) equals

$$\begin{aligned} &\frac{\mathbb{E}_x\left(e^{i(2\pi m + \frac{s}{\sqrt{V_N}})S_N} v_{N+1}(X_{N+1})\right)}{\mathbb{E}(v_{N+1}(X_{N+1}))} = \\ &= \mathbb{E}_x\left(e^{i(2\pi m + \frac{s}{\sqrt{V_N}})S_{r-1}} \frac{\mathbb{E}\left(e^{i(2\pi m + \frac{s}{\sqrt{V_N}})(S_N - S_{r-1})} v_{N+1}(X_{N+1}) \middle| X_r\right)}{\mathbb{E}(v_{N+1}(X_{N+1}))}\right) \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}_x \left[ e^{i(2\pi m + \frac{s}{\sqrt{V_N}})S_{r-1}} \left( e^{2\pi i m c^{(N)} - s^2/2} + \frac{\eta_{N-r}(X_r)}{\mathbb{E}(\eta_{N+1}(X_{N+1}))} \right) \right] \\
&= \underbrace{e^{2\pi i m c^{(N)} - s^2/2} \mathbb{E}_x(e^{2\pi i m S_{r-1} + o(1)})}_{A} + O(\bar{\delta}^{-1}) \underbrace{\mathbb{E}_x(\eta_{N-r}(X_r))}_{B}, \text{ as } N \rightarrow \infty.
\end{aligned}$$

We examine  $A, B$ . Let  $\hat{c}^{(r-1)} := \sum_{k=1}^{r-1} c_k = -\mathbb{E}(\sum_{k=1}^{r-1} \mathbb{F}_k(X_k, X_{k+1}))$ . Since  $c(N) = -\sum_{k=1}^N \mathbb{E}(\mathbb{F}_k)$ ,

$$\begin{aligned}
c(N) &= -\sum_{k=1}^{r-1} \mathbb{E}(\mathbb{F}_k) - \sum_{k=r}^N (\mathbb{E}(\tilde{f}_n^{(N)}) + \mathbf{c}_n^{(N)}) \text{ because } \sum_{n=r}^N \mathbb{F}_n = \sum_{n=r}^N (\mathbf{c}_n^{(N)} + \tilde{f}_n^{(N)}) \\
&\equiv \hat{c}^{(r-1)} + c^{(N)} \pmod{\mathbb{Z}}, \text{ because } \mathbf{c}_n^{(N)} \in \mathbb{Z}.
\end{aligned}$$

By assumption,  $\mathbf{f} = \mathbb{F} + \mathbf{h} + \mathbf{c}$  with  $\mathbb{F}$  integer valued. Necessarily,

$$\exp(2\pi i m S_{r-1}) = \exp(2\pi i m \mathfrak{H}_r + 2\pi i m \hat{c}^{(r-1)}) \quad (5.2.23)$$

where  $\mathfrak{H}_r := \sum_{k=1}^{r-1} h_k(X_k, X_{k+1})$ . By choice of  $r$  and Lemma 2.5,

$$|\mathbb{E}_x(e^{i\xi \mathfrak{H}}) - \mathbb{E}_x(e^{i\xi \mathfrak{H}_r})| \leq |\xi| \mathbb{E}_x(|\mathfrak{H} - \mathfrak{H}_r|) \leq |\xi| \text{Var} \left( \sum_{k=r}^{\infty} h_k(X_k, X_{k+1}) \right)^{1/2} = O(\sqrt{\varepsilon})$$

uniformly when  $\xi$  varies in a compact domain. Substituting (5.2.23) in  $A$ , we obtain

$$A = [1 + o(1)] e^{2\pi i m c^{(N)} - \frac{s^2}{2}} \mathbb{E}_x \left( e^{2\pi i m \mathfrak{H}} \right) + O(\sqrt{\varepsilon}).$$

Next, the exponential mixing of  $\mathbf{X}$  implies that for all  $N$  large enough,

$$B := \mathbb{E}_x(\eta_{N-r}(X_r)) = \mathbb{E}(\eta_{N-r}(X_r)) + o(1) = O(\sqrt{\varepsilon}).$$

Thus the left-hand-side of (5.2.21) equals  $e^{2\pi i m c^{(N)} - s^2/2} \mathbb{E}_x(e^{2\pi i m \mathfrak{H} + o(1)}) + O(\sqrt{\varepsilon})$ . The lemma follows, because  $\varepsilon$  was arbitrary.  $\square$

**Proof of Theorem 5.1.** Suppose  $\mathbf{f}$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $\mathbf{X}$ , and assume  $G_{ess}(\mathbf{f}) = \delta(\mathbf{f})\mathbb{Z}$  with  $\delta(\mathbf{f}) \neq 0$ .

We begin with some reductions. By Theorem 3.3,  $\mathbf{f}$  has an optimal reduction, and we can write  $\mathbf{f} = \mathbb{F} + \mathbf{F}$  where  $\mathbb{F}$  has algebraic range  $\delta(\mathbf{f})\mathbb{Z}$  and  $\mathbf{F}$  is a.s. uniformly bounded and center-tight. There is no loss of generality in assuming that  $\text{ess sup } |\mathbf{F}| \leq \delta(\mathbf{f})$ , since this can always be arranged by replacing  $F_n$  by  $F_n \pmod{\delta(\mathbf{f})}$ . Next by the gradient lemma (Lemma 2.9), we decompose

$$\mathbf{F} = \nabla \mathbf{a} + \tilde{\mathbf{f}} + \tilde{\mathbf{c}}$$

where  $\text{ess sup } |\mathbf{a}| \leq 2 \text{ess sup } |\mathbf{F}|$ ,  $\tilde{\mathbf{f}}$  has summable variances, and  $\tilde{\mathbf{c}}$  are constants.

It is convenient to introduce  $f_n^* := \frac{1}{\delta(\mathbf{f})} [f_n - \nabla a_n - \mathbb{E}(f_n - \nabla a_n)]$ .  $G_{ess}(\mathbf{X}, \mathbf{f}^*) = \mathbb{Z}$ , and

$$\mathbf{f}^* = \frac{1}{\delta(\mathbf{f})} \mathbb{F} + \mathbf{h} + \mathbf{c}, \quad (5.2.24)$$

where  $h_n := \frac{1}{\delta(f)} [\tilde{f}_n - \mathbb{E}(\tilde{f}_n)]$  is a centered additive functional with summable variances, and  $c_n := \frac{1}{\delta(f)} [\tilde{c}_n + \mathbb{E}(\tilde{f}_n) - \mathbb{E}(f_n - \nabla a_n)]$ .

We first prove the theorem in the special case when

$$\delta(f) = 1, \mathbb{E}(f_n) = 0 \text{ for all } n, \text{ and } a \equiv 0. \quad (5.2.25)$$

In this case  $f = f^*$  and (5.2.24) places us in the setup of Lemma 5.5. Given this lemma, the proof is very similar to the proof of the local limit theorem in the irreducible non-lattice case, but we reproduce it for completeness. We focus on parts (2) and (3) of the theorem, because part (1) follows from them.

Define as in (5.2.19),  $c(N) := -\frac{1}{\delta(f)} \sum_{k=1}^N \mathbb{E}[\mathbb{F}_k(X_k, X_{k+1})]$ , and let

$$\mathfrak{H} := \sum_{n=1}^{\infty} h_n(X_n, X_{n+1}), \quad b_N := \{c(N)\}.$$

Fix  $\phi \in L^1(\mathbb{R})$  such that  $\text{supp}(\hat{\phi}) \subset [-L, L]$ , and let  $v_{N+1}$  denote the indicator function of  $\mathfrak{A}_{N+1}$ . By the Fourier inversion formula

$$\begin{aligned} & \mathbb{E}_x(\phi(S_N - b_N - z_N) | X_{N+1} \in \mathfrak{A}_{N+1}) \\ &= \frac{1}{2\pi} \int_{-L}^L \hat{\phi}(\xi) \frac{\mathbb{E}_x(e^{i\xi(S_N - b_N - z_N)} v_{N+1}(X_{N+1}))}{\mathbb{E}(v_{N+1}(X_{N+1}))} d\xi \end{aligned} \quad (5.2.26)$$

and the task is to find the asymptotic behavior of (5.2.26) in case  $z_N \in \mathbb{Z}$ ,  $\frac{z_N}{\sqrt{V_N}} \rightarrow z$ .

Let  $K := \text{ess sup } |f|$  and recall the constant  $\tilde{\delta} = \tilde{\delta}(K)$  from Lemma 4.4. Split  $[-L, L]$  into a finite collection of subintervals  $I_j$  of length less than  $\min\{\tilde{\delta}, \pi\}$ , in such a way that every  $I_j$  is either bounded away from  $2\pi\mathbb{Z}$ , or intersects it at a unique point  $2\pi m$  exactly at its center.

If  $I_j \cap 2\pi\mathbb{Z} = \emptyset$ , then  $\sum d_n^2(\xi) = \infty$  uniformly on  $I_j$  (Theorem 3.5). Thus by (4.2.7),  $\Phi_N(x, \xi) \rightarrow 0$  uniformly on  $I_j$ . In this case we can argue as in the proof of (4.2.18) and show that the contribution of  $I_j$  to the integral (5.2.26) is  $o(V_N^{-1/2})$ .

If  $I_j \cap 2\pi\mathbb{Z} \neq \emptyset$ , then the center of  $I_j$  equals  $2\pi m$  for some  $m \in \mathbb{Z}$ . Fix some large  $R$ . Let  $J'_{j,N}$  be the contribution to the integral from the set  $\{\xi \in I_j : |\xi - 2\pi m| \leq RV_N^{-1/2}\}$ , and let  $J''_{j,N}$  be the integral over  $\{\xi \in I_j : |\xi - 2\pi m| > RV_N^{-1/2}\}$ .

The main contribution comes from  $J'_{j,N}$ , because one can show as in Claim 2 in §4.2.2 that  $|J''_{j,N}| \leq C \int_{|u| > RV_N^{-1/2}} e^{-cV_N u^2} du \leq C \frac{e^{-cR^2}}{R\sqrt{V_N}}$ , which is negligible for  $R \gg 1$ .

To estimate  $J'_{j,N}$ , we make the change of variables  $\xi = 2\pi m + \frac{s}{\sqrt{V_N}}$ . Since  $z_N \in \mathbb{Z}$  and  $b_N = \{c(N)\}$ , we have  $\xi(S_N - b_N - z_N) = \xi S_N - 2\pi m c(N) - \frac{s}{\sqrt{V_N}} (z_N + \{c(N)\}) \pmod{2\pi}$ . So

$$J'_{j,N} = \frac{1}{2\pi\sqrt{V_N}} \left[ \int_{|s| < R} \hat{\phi}(2\pi m + \frac{s}{\sqrt{V_N}}) \frac{e^{-2\pi i m c(N)} \mathbb{E}_x(e^{i\xi S_N} v_{N+1}(X_{N+1}))}{\mathbb{E}(v_{N+1}(X_{N+1}))} e^{-is\frac{z_N + O(1)}{\sqrt{V_N}}} ds \right].$$

Fixing  $R$  and letting  $N \rightarrow \infty$ , we see by Lemma 5.5 that

$$\begin{aligned}\sqrt{V_N}J'_{j,N} &= \frac{\widehat{\phi}(2\pi m)}{2\pi} \mathbb{E}_x \left( e^{2\pi im\mathfrak{H}} \right) \int_{|s| < R} e^{-isz - s^2/2} ds + o_{N \rightarrow \infty}(1) \\ &= \frac{\widehat{\phi}(2\pi m)}{\sqrt{2\pi}} \mathbb{E}_x \left( e^{2\pi im\mathfrak{H}} \right) e^{-z^2/2} + o_{R \rightarrow \infty}(1) + o_{N \rightarrow \infty}(1).\end{aligned}$$

Combining the estimates for  $J_{j,N}$  we obtain that

$$\lim_{N \rightarrow \infty} \sqrt{V_N} J_{j,N} = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \mathbb{E}_x \left( e^{2\pi im\mathfrak{H}} \right) \widehat{\phi}(2\pi m),$$

if  $I_j$  intersects  $2\pi\mathbb{Z}$ , and this limit is zero otherwise. Hence

$$\begin{aligned}\lim_{n \rightarrow \infty} \sqrt{V_N} \mathbb{E}_x(\phi(S_N - b_N - z_N) | X_{N+1} \in \mathfrak{A}_{N+1}) \\ &= \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z} \cap [-L, L]} \mathbb{E}_x \left( e^{2\pi im\mathfrak{H}} \right) \widehat{\phi}(2\pi m) = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \mathbb{E}_x \left( e^{2\pi im\mathfrak{H}} \right) \widehat{\phi}(2\pi m) \\ &\equiv \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \mathbb{E}_x \left( e^{2\pi im\mathfrak{F}} \right) \widehat{\phi}(2\pi m), \text{ where } \mathfrak{F} \in [0, 1), \mathfrak{F} := \mathfrak{H} \pmod{\mathbb{Z}} \\ &= \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \widehat{(\mathcal{C}_x \phi)}(2\pi m), \text{ where } (\mathcal{C}_x \phi)(t) := \mathbb{E}_x[\phi(t + \mathfrak{F})] \\ &= \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} (\mathcal{C}_x \phi)(m) \equiv \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \mathbb{E}_x[\phi(m + \mathfrak{F})],\end{aligned}$$

by the Poisson summation formula.

This proves part (2) of the theorem in the special case (5.2.25), and in particular for the additive functional  $f^*$  defined above. Now consider the general case:

$$S_N(f) - \mathbb{E}[S_N(f)] \equiv \delta(f)S_N(f^*) + a_{N+1}(X_{N+1}) - a_1(X_1) + \mathbb{E}[a_1(X_1) - a_{N+1}(X_{N+1})].$$

Since part (2) of the theorem holds for  $f^*$  with  $\mathfrak{F} = \{\sum h_n\} \in [0, 1)$  and  $b_N = \{c(N)\}$ , it must hold for  $f$  with  $\delta(f)\mathfrak{F}$  and

$$b_N(X_1, X_{N+1}) := \delta(f)\{c(N)\} + a_{N+1}(X_{N+1}) - a_1(X_1) + \mathbb{E}[a_1(X_1) - a_{N+1}(X_{N+1})].$$

Clearly  $|b_N| \leq \delta(f) + 4\text{ess sup } |a|$ . Recalling that  $\text{ess sup } |a| \leq 2\text{ess sup } |F| \leq 2\delta(f)$ , we find that  $\text{ess sup } |b_N| \leq 9\delta(f)$ , proving part (3) as well.  $\square$

### 5.2.3 Necessity of the irreducibility assumption

Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . Recall that  $f_r = \{f_n\}_{n \geq r}$  and  $X_r = \{X_n\}_{n \geq r}$ . In this section we prove Theorem 5.2, which asserts the equivalence of the following three conditions:

- (a)  $f$  is irreducible with algebraic range  $\mathbb{R}$ .
- (b)  $(X_r, f_r)$  satisfies the mixing non-lattice local limit theorem, for all  $r$ .

(c)  $(X_r, f_r)$  satisfies the mixing uniform distribution mod  $t$  for all  $r$  and  $t > 0$ .

**(a)  $\Rightarrow$  (b):** To see this recall that additive functionals on uniformly elliptic Markov chains are special cases of stably hereditary additive functionals on uniformly elliptic Markov arrays, and apply Theorem 4.3(1) to  $\phi$  continuous with compact support which approximate indicators of intervals in  $L^1(\mathbb{R})$ .

**(b)  $\Rightarrow$  (a):** Assume  $f$  satisfies the “mixing non-lattice LLT” property. By definition,  $V_N \rightarrow \infty$ , and therefore  $f$  is not center-tight.

Also,  $G_{alg}(X, f) = \mathbb{R}$ , otherwise  $\mathbb{P}_x(S_N - z_N \in (a, b) | X_{N+1} \in \mathfrak{A}_{N+1}) = 0$  for  $z_N$  and  $(a, b)$  such that  $z_N + (a, b) \subset \mathbb{R} \setminus G_{alg}(X, f)$ .

If  $G_{ess}(X, f) = \mathbb{R}$  then  $f$  is irreducible and we are done. Assume by way of contradiction that  $G_{ess}(X, f) \neq \mathbb{R}$ , then  $G_{ess}(X, f) = t\mathbb{Z}$  for some  $t > 0$  ( $t = 0$  is impossible because  $f$  is not center-tight). There is no loss of generality in assuming that

$$G_{ess}(X, f) = \mathbb{Z} \quad \text{and} \quad \mathbb{E}(f_n(X_n, X_{n+1})) = 0 \text{ for all } n.$$

Let  $S_N^{(r)} := f_r(X_r, X_{r+1}) + \dots + f_N(X_N, X_{N+1})$  and  $V_N^{(r)} := \text{Var}(S_N^{(r)})$ . By the exponential mixing of  $X$  (Proposition 1.11),

$$|V_N - V_N^{(r)}| = |V_{r-1} + 2\text{Cov}(S_N^{(r)}, S_{r-1})| \leq V_r + 2 \sum_{j=1}^{r-1} \sum_{k=r}^{\infty} \text{Cov}(f_j, f_k) = O(1).$$

Therefore, for fixed  $r$ ,  $V_N/V_N^{(r)} \xrightarrow[N \rightarrow \infty]{} 1$ .

Since  $G_{alg}(X, \mathbb{R}) = \mathbb{R}$  and  $G_{ess}(X, f) = \mathbb{Z}$ ,  $f$  is reducible, and we can write

$$f = \mathbb{F} + \nabla a + h + c,$$

where  $\mathbb{F}$  is irreducible with algebraic range  $\mathbb{Z}$ ,  $a_n(x)$  are uniformly bounded (say by  $K$ ),  $h$  has summable variances,  $\mathbb{E}(h_n) = 0$ , and  $c$  are constants.

Let

$$b_N^{(r)}(X_r, X_{N+1}) := a_{N+1}(X_{N+1}) - a_r(X_r) + \left\{ - \sum_{k=r}^N \mathbb{E}(\mathbb{F}_k(X_k, X_{k+1})) \right\},$$

$$\mathfrak{F} := \sum_{n=1}^{\infty} h_n(X_n, X_{n+1}), \quad \mathfrak{F}_r := \sum_{n=r}^{\infty} h_n(X_n, X_{n+1}).$$

By Theorem 2.4, these sums converge almost surely and in  $L^2$ .

As we saw in the proof of Theorem 5.1, if  $\frac{z_N - \mathbb{E}(S_N^{(r)})}{\sqrt{V_N^{(r)}}} \rightarrow 0$  and  $\mathbb{P}(X_n \in \mathfrak{A}_n)$  is bounded below, then for all  $\phi \in C_c(\mathbb{R})$  and  $x_r \in \mathfrak{S}_r$ ,

$$\lim_{N \rightarrow \infty} \sqrt{2\pi V_N} \mathbb{E}_{x_r} [\phi(S_N^{(r)} - b_N^{(r)} - z_N)] = \sum_{m \in \mathbb{Z}} \mathbb{E}[\phi(m + \mathfrak{F}_r)]. \quad (5.2.27)$$

We are going to choose  $r, x_r, z_N, \mathfrak{A}_N$  and  $\phi$  in such a way that (5.2.27) is inconsistent with (b). Here are the choices:

- o *Choice of  $r$ :* Since  $\mathfrak{F}_r$  is the tail of a convergent series,  $\mathfrak{F}_r \xrightarrow[r \rightarrow \infty]{} 0$  a.s., whence in probability. Choose  $r$  s.t.  $\mathbb{P}(|\mathfrak{F}_r| \geq 0.2) < 10^{-3}$ .

- *Choice of  $x_r$ :*  $\mathbb{P}(|\mathfrak{F}_r| \geq 0.2) = \int \mathbb{P}_x(|\mathfrak{F}_r| \geq 0.2) \mu_r(dx)$ . So there exist  $x_r \in \mathfrak{S}_r$  s.t.

$$\mathbb{P}_{x_r}(|\mathfrak{F}_r| > 0.2) < 10^{-3}.$$

- *Choice of  $\mathfrak{A}_N$ :* By construction,  $\text{ess sup } |b_N^{(r)}| \leq 2K + 1$ . Divide  $[-2K - 1, 2K + 1]$  into equal intervals of length less than  $10^{-2}$ . At least one such interval, call it  $J_N$ , satisfies  $\mathbb{P}(b_N^{(r)} \in J_N) \geq 10^{-2}(4K + 2)^{-1}$  and  $|J_N| \leq 10^{-2}$ . Let

$$\mathfrak{A}_{N+1}^{(r)} := [b_N^{(r)} \in J_N].$$

- $z_N := -\text{center of } J_N$ , then  $z_N = O(1)$  and  $\frac{z_N - \mathbb{E}(S_N^{(r)})}{\sqrt{V_N^{(r)}}} \rightarrow 0$ .
- Choose a sequence  $N_k \rightarrow \infty$  such that  $z_{N_k} \rightarrow a$ . Let  $I := -a + [0.4, 0.6]$ .
- Choose  $\phi \in C_c(\mathbb{R})$  s.t.  $0 \leq \phi \leq 1$ ,  $\phi|_{[0.3, 0.7]} \equiv 1$  and  $\phi|_{\mathbb{R} \setminus [0.2, 0.8]} \equiv 0$ .

With these choices,

$$\begin{aligned} & \liminf_{N \rightarrow \infty} \sqrt{2\pi V_N^{(r)}} \mathbb{P}_{x_r}(S_N^{(r)} - z_N \in I \mid X_{N+1} \in \mathfrak{A}_{N+1}^{(r)}) \\ & \leq \lim_{k \rightarrow \infty} \sqrt{2\pi V_{N_k}} \mathbb{P}_{x_r}(S_{N_k}^{(r)} - z_{N_k} \in I \mid b_{N_k}^{(r)}(X_r, X_{N_k+1}) \in J_{N_k}) \quad \left( \text{because } \frac{V_N^{(r)}}{V_N} \rightarrow 1 \right) \\ & \leq \lim_{k \rightarrow \infty} \sqrt{2\pi V_{N_k}} \mathbb{P}_{x_r}(S_{N_k}^{(r)} - b_{N_k}^{(r)} - z_{N_k} \in [0.3, 0.7] \mid b_{N_k}^{(r)} \in J_{N_k}), \quad (\text{because for } k \gg 1 \\ & \quad I - b_{N_k} \subset I - J_{N_k} \subset I + \left( z_{N_k} - \frac{|J_{N_k}|}{2}, z_{N_k} + \frac{|J_{N_k}|}{2} \right) \subset I + (a - 0.1, a + 0.1) \subset [0.3, 0.7]) \\ & \leq \lim_{k \rightarrow \infty} \sqrt{2\pi V_{N_k}} \mathbb{E}_{x_r}(\phi(S_{N_k}^{(r)} - b_{N_k}^{(r)} - z_{N_k}) \mid b_{N_k}^{(r)} \in J_{N_k+1}) \\ & = \sum_{m \in \mathbb{Z}} \mathbb{E}_{x_r}[\phi(m + \mathfrak{F}_r)], \text{ by (5.2.27)} \\ & \leq \sum_{m \in \mathbb{Z}} \mathbb{P}_{x_r}(m + \mathfrak{F}_r \in [0.2, 0.8]) \leq \mathbb{P}_{x_r}(|\mathfrak{F}_r| \geq 0.2) < 10^{-3} < |I|. \end{aligned}$$

But this contradicts (b).

**(a)  $\Rightarrow$  (c):** Suppose  $(X, f)$  is non-lattice and irreducible, then  $(X_r, f_r)$  is non-lattice and irreducible for all  $r$ . Fix  $t > 0$ ,  $x_1 \in \mathfrak{S}_1$ , and some sequence of measurable events  $\mathfrak{A}_n \subset \mathfrak{S}_n$  such that

$\mathbb{P}(X_n \in \mathfrak{A}_n)$  is bounded below. Let  $S_N^{(r)} := \sum_{k=r}^N f_k(X_k, X_{k+1})$ .

We show that for every continuous and periodic  $\phi(x)$  with period  $t$ ,

$$\mathbb{E}_x(\phi(S_N^{(r)}) \mid X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N \rightarrow \infty]{} \frac{1}{t} \int_0^t \phi(x) dx. \quad (5.2.28)$$

It is enough to show (5.2.28) for trigonometric polynomials  $\phi(u) = \sum_{|n| < L} c_n e^{2\pi i n u / t}$ , as these are dense in  $C[0, t]$ . For such functions,

$$\begin{aligned}
\mathbb{E}_x(\phi(S_N^{(r)})|X_{N+1} \in \mathfrak{A}_{N+1}) &= \sum_{|n| < L} c_n \mathbb{E}_x(e^{2\pi i n S_N^{(r)}/t} | X_{N+1} \in \mathfrak{A}_{N+1}) \\
&= c_0 + \sum_{0 < |n| < L} \Phi_N(x, \frac{2\pi n}{t} | \mathfrak{A}_{N+1}), \text{ where } \Phi_N \text{ are the characteristic functions of } (\mathbf{X}_r, \mathbf{f}_r) \\
&= c_0 + o(1), \text{ by irreducibility and (4.2.7).}
\end{aligned}$$

Since  $c_0 = \frac{1}{t} \int_0^t \phi(u) du$ , (5.2.28) follows. Standard approximation arguments show that (5.2.28) implies that

$$\mathbb{P}_x(S_N^{(r)} \in (a, b) | X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N \rightarrow \infty]{} \frac{|a-b|}{t} \text{ for all intervals } (a, b).$$

(c)  $\Rightarrow$  (a): We need the following lemma.

**Lemma 5.6** *Fix a regular sequence of sets  $\mathfrak{A}_N$ ,  $x$ , and  $t > 0$ , and suppose that*

$$\mathbb{P}_x(S_N^{(r)} \in (a, b) + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N \rightarrow \infty]{} \frac{|a-b|}{t}$$

for all intervals  $(a, b)$  s.t.  $0 < |a-b| < t$ . Then the convergence is uniform in  $(a, b)$ .

*Proof.* Without loss of generality,  $(a, b) \subset [0, t]$ . We are asked to find for each  $\varepsilon > 0$  an  $N_0$  such that

$$|\mathbb{P}_x(S_N^{(r)} \in (a, b) + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) - \frac{|a-b|}{t}| < \varepsilon \text{ for all } N > N_0 \text{ and } a < b.$$

Choose  $0 < \delta < \min\{\frac{\varepsilon}{5}, 1\}$ , and divide  $[0, t]$  into finitely many equal disjoint intervals  $\{I_j\}$  with length  $|I_j| < \delta$ . Choose  $N_0$  so that for all  $N > N_0$ , for all  $I_j$ ,

$$|\mathbb{P}_x(S_N^{(r)} \in I_j + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) - \frac{|I_j|}{t}| < \frac{\delta|I_j|}{t}. \quad (5.2.29)$$

$I := (a, b)$  can be approximated from within and from outside by finite (perhaps empty) unions of intervals  $I_j$  whose total length differs from  $|a-b|$  by no more than  $2\delta$ . Summing (5.2.29) over these unions we see that for all  $N > N_0$ ,

$$\begin{aligned}
\mathbb{P}_x(S_N^{(r)} \in I + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) &\leq \frac{|a-b| + 2\delta}{t} + \frac{\delta(|a-b| + 2\delta)}{t} \\
\mathbb{P}_x(S_N^{(r)} \in I + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) &\geq \frac{|a-b| - 2\delta}{t} - \frac{\delta|a-b|}{t}.
\end{aligned}$$

By choice of  $\delta$ ,  $|\mathbb{P}_x(S_N^{(r)} \in I + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) - \frac{|a-b|}{t}| < \varepsilon$ .  $\square$

We can now prove that (c)  $\Rightarrow$  (a). Suppose  $(\mathbf{X}_r, \mathbf{f}_r)$  has the “mixing uniform distribution mod  $t$ ” property for all  $r$  and  $t$ . This property is invariant under centering, because of Lemma 5.6. So we may assume without loss of generality that  $\mathbb{E}[f_n(X_n, X_{n+1})] = 0$  for all  $n$ .

First we claim that  $(\mathbf{X}, \mathbf{f})$  is not center-tight. Otherwise there are constants  $c_N$  and  $M$  such that  $\mathbb{P}(|S_N - c_N| > M) < 0.1$  for all  $N$ . Take  $t := 5M$  and  $N_k \rightarrow \infty$  such that  $c_{N_k} \xrightarrow[k \rightarrow \infty]{} c \pmod{t\mathbb{Z}}$ , then by the bounded convergence theorem and (c),

$$\begin{aligned}
0.9 &\leq \lim_{k \rightarrow \infty} \mathbb{P}(S_{N_k} \in [c - 2M, c + 2M]) \leq \lim_{N \rightarrow \infty} \mathbb{P}(S_N \in [c - 2M, c + 2M] + t\mathbb{Z}) \\
&= \int_{\mathfrak{S}_1} \lim_{N \rightarrow \infty} \mathbb{P}_x(S_N \in [c - 2M, c + 2M] + t\mathbb{Z} | X_{N+1} \in \mathfrak{S}_{N+1}) \mu_1(dx) = \frac{4M}{t} = 0.8,
\end{aligned}$$

a contradiction. Thus  $(X, f)$  is not center-tight and  $V_N \rightarrow \infty$ .

Assume by way of contradiction that  $G_{ess}(X, f) \neq \mathbb{R}$ , then  $G_{ess}(X, f) = t\mathbb{Z}$  for some  $t$ , and  $t \neq 0$  because  $V_N \rightarrow \infty$ . Without loss of generality  $t = 1$ , otherwise we can rescale  $f$ . By the integer reduction lemma, we can write

$$f_n(x, y) + a_n(x) - a_{n+1}(y) = \mathbb{F}_n(x, y) + h_n(x, y) + c_n$$

where  $a_k, \mathbb{F}_k, h_k, c_k$  are uniformly bounded,  $\mathbb{F}_n$  are integer valued,  $h_n$  have summable variances, and  $\mathbb{E}(h_n) = 0$ . Then  $\mathfrak{F} := \sum_{n \geq 1} h_n(X_n, X_{n+1})$  converges a.s., and  $\mathfrak{F}_r := \sum_{n \geq r} h_n(X_n, X_{n+1}) \xrightarrow[r \rightarrow \infty]{} 0$  almost surely.

Working as in the proof of  $(b) \Rightarrow (a)$ , we construct  $x \in \mathfrak{S}_1$  and  $r > 1$  such that

$$|\mathbb{E}_x(e^{2\pi i \mathfrak{F}_r})| > 0.999.$$

Next we construct a regular sequence of measurable sets  $\mathfrak{A}_{N+1}$ , and intervals  $J_N$  with lengths  $< 0.0001$  and centers  $z_N = O(1)$  such that  $a_{N+1}(X_{N+1}) - a_1(X_r) \in J_N$ , whenever  $X_{N+1} \in \mathfrak{A}_{N+1}, X_r = x$ .

By Lemma 5.5 with  $s = 0$ ,  $m = 1$ , and  $v_{N+1} \equiv 1$ , there are  $c(r, N) \in \mathbb{R}$  s.t.

$$\mathbb{E}_x(e^{2\pi i (S_N^{(r)} + a(X_1) - a(X_{N+1}) - z_N)}) |X_{N+1} \in \mathfrak{A}_{N+1}| = e^{2\pi i (c(r, N) - z_N)} \mathbb{E}_x(e^{2\pi i \mathfrak{F}_r}) + o(1),$$

as  $N \rightarrow \infty$ . Since

$$\left\| \left( e^{2\pi i (S_N^{(r)} + a(X_r) - a(X_{N+1}) - z_N)} - e^{2\pi i (S_N^{(r)})} \right) 1_{[X_{N+1} \in \mathfrak{A}_{N+1}, X_r = x]} \right\|_{\infty} < 0.1,$$

we find that for all  $N$  large enough,  $|\mathbb{E}_x(e^{2\pi i (S_N^{(r)})})| |X_{N+1} \in \mathfrak{A}_{N+1}| > \frac{1}{2}$ .

But this is a contradiction, since (c) implies that

$$\mathbb{E}_x(e^{2\pi i S_N^{(r)}}) |X_{N+1} \in \mathfrak{A}_{N+1}| \xrightarrow[N \rightarrow \infty]{} \frac{1}{2\pi} \int_0^{2\pi} e^{iu} du = 0.$$

So  $G_{ess}(X, f) = \mathbb{R}$  and (a) is proved.  $\square$

#### 5.2.4 Universal bounds for Markov chains

**Lemma 5.7** Suppose  $\mathfrak{F}$  is a real random variable such that  $0 \leq \mathfrak{F} < \delta$  almost surely. Then for every interval  $(a, b)$  of length  $L > \delta$ ,

$$\left(1 - \frac{\delta}{L}\right) |a - b| < \delta \sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a, b)}(m\delta + \mathfrak{F})] < \left(1 + \frac{\delta}{L}\right) |a - b|.$$

*Proof.* Fix  $k$  large, and divide  $[0, \delta)$  into  $k$  intervals  $I_j := \frac{j\delta}{k} + [0, \frac{\delta}{k})$ . For each  $j$ ,

$$\begin{aligned} \delta \sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a, b)}(m\delta + \mathfrak{F}) | \mathfrak{F} \in I_j] &\leq \delta \sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a + \frac{(j-1)\delta}{k}, b + \frac{(j+1)\delta}{k})}(m\delta) | \mathfrak{F} \in I_j] \\ &= \delta \sum_{m \in \mathbb{Z}} 1_{(a + \frac{(j-1)\delta}{k}, b + \frac{(j+1)\delta}{k})}(m\delta) \leq |a - b| + 1 + \frac{2\delta}{k} \xrightarrow[k \rightarrow \infty]{} |a - b| + 1. \end{aligned}$$

Multiplying by  $\mathbb{P}[\mathfrak{F} \in I_j]$  and summing over  $j = 0, \dots, k-1$  gives the bound  $\delta \sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a,b)}(m\delta + \mathfrak{F})] \leq |a-b| + \delta$ . Similarly,  $\delta \sum_{m \in \mathbb{Z}} \mathbb{E}[1_{(a,b)}(m\delta + \mathfrak{F})] \geq |a-b| - \delta$ . The lemma follows.  $\square$

**Proof of Theorem 5.3:** If  $\delta(f) = \infty$  then there is nothing to prove, and if  $\delta(f) = 0$  then  $(X, f)$  is non-lattice and irreducible, and the universal bounds follow from Theorem 4.1. So assume  $\delta(f)$  is finite and positive.

Suppose  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ . Let  $\mathfrak{F}$  and  $b_N(X_1, X_N)$  be as in Theorem 5.1.

**Upper bound (5.1.6):** Fix  $x \in \mathfrak{S}_1$ , let  $\delta := \delta(f)$  and suppose  $(a, b)$  is an interval of length  $L > \delta$ . We may assume without loss of generality that  $a - 10\delta, b + 10\delta$  are not atoms of the distribution of  $\mathfrak{F}$  given  $X_1 = x$  (otherwise change  $a, b$  a little).

Suppose  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , and write  $z_N = \bar{z}_N + \zeta_N$ ,  $\bar{z}_N \in \delta\mathbb{Z}$ ,  $|\zeta_N| \leq \delta$ . Recall that by Theorem 5.1,  $|b_N| \leq 9\delta$ . Therefore

$$S_N - z_N \in (a, b) \Rightarrow S_N - \bar{z}_N - b_N \in (a - 10\delta, b + 10\delta)$$

So

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \sqrt{2\pi V_N} \mathbb{P}_x[S_N - z_N \in (a, b)] \\ & \leq \limsup_{N \rightarrow \infty} \sqrt{2\pi V_N} \mathbb{P}_x[S_N - \bar{z}_N - b_N \in (a - 10\delta, b + 10\delta)] \\ & = e^{-z^2/2} \delta \sum_{m \in \mathbb{Z}} \mathbb{E}_x[1_{(a-10\delta, b+10\delta)}(m\delta + \mathfrak{F})] \text{ by Theorem 5.1} \\ & \leq \left(1 + \frac{\delta}{|a-b| + 20\delta}\right) e^{-z^2/2} (|a-b| + 20\delta) \text{ by Lemma 5.7} \quad (5.2.30) \\ & \leq (|a-b| + 21\delta) e^{-z^2/2} \leq \left(1 + \frac{21\delta}{L}\right) e^{-z^2/2} |a-b|. \end{aligned}$$

**Lower bound (5.1.7):** Fix  $x \in \mathfrak{S}_1$  and an interval  $(a, b)$  with length bigger than some  $L > \delta(f)$ . Recall that  $|b_N|$  are uniformly bounded. Choose some  $\bar{K}$  so that  $\mathbb{P}[|b_N| \leq \bar{K}] = 1$  and fix  $x \in \mathfrak{S}_1$  s.t.  $\mathbb{P}_x[\sup |b_N| \leq \bar{K}] = 1$ .

Next, divide  $[-\bar{K}, \bar{K}]$  into  $k$  disjoint intervals  $I_{j,N}$  of equal length  $\frac{2\bar{K}}{k}$ , with  $k$  large. For each  $N$ ,

$$\sum_{\mathbb{P}_x[b_N \in I_{j,N}] \geq k^{-2}} \mathbb{P}_x[b_N \in I_{j,N}] \geq 1 - \frac{1}{k},$$

because to complete the left-hand-side to one we need to add the probabilities of  $[b_N \in I_{j,N}]$  for the  $j$  s.t.  $\mathbb{P}_x[b_N \in I_{j,N}] < k^{-2}$ , and there are at most  $k$  such events.

Therefore, we can divide  $\{I_{j,N}\}$  into two groups of size at most  $k$ : The first contains the  $I_{j,N}$  with  $\mathbb{P}_x[b_N \in I_{j,N}] \geq k^{-2}$ , and the second corresponds to events with total probability less than  $\frac{1}{k}$  (conditioned on  $X_1 = x$ ).

Re-index the intervals in the first group (perhaps with repetitions) in such a way that it takes the form  $I_{j,N}$  ( $j = 1, \dots, k$ ) for all  $N$ . Then for each  $j$ ,  $\mathfrak{A}_{j,N} := [b_N \in I_{j,N}, X_1 = x]$  is a regular sequence of events.

Let  $\beta_{j,N} :=$  center of  $I_{j,N}$  and set  $z_{j,N} := z_N - \beta_{j,N}$ . Every sequence has a subsequence s.t.  $z_{j,N}$  converges mod  $\delta(f)$ . We will henceforth assume that  $z_{j,N} = \bar{z}_{j,N} + \zeta_0 + \zeta_{j,N}$  where  $\bar{z}_{j,N} \in \delta(f)\mathbb{Z}$  and  $|\zeta_{j,N}| < \frac{K}{k}$ , and  $|\zeta_0| < \delta(f)$  is fixed.

Recall that  $|I_{j,N}| = \frac{2\bar{K}}{k}$ . Conditioned on  $\mathfrak{A}_{j,N}$ ,  $b_N = \beta_{j,N} \pm \frac{2\bar{K}}{k}$ , therefore  $\bar{z}_{j,N} + \zeta_0 + b_N = z_N \pm \frac{3K}{k}$ , whence

$$S_N - \bar{z}_{j,N} - b_N \in \left( a - \zeta_0 + \frac{3\bar{K}}{k}, b - \zeta_0 - \frac{3\bar{K}}{k} \right) \Rightarrow S_N - z_N \in (a, b).$$

There is no loss of generality in assuming that the endpoints of this interval are not atoms of the distribution of  $\mathfrak{F}$  given  $X_1 = x$ , otherwise perturb  $K$  a little. Since  $\mathfrak{A}_{j,N}$  is a regular sequence, we have by Theorem 5.1 part (2) and the lemma that

$$\begin{aligned} & \liminf_{N \rightarrow \infty} \sqrt{2\pi V_N} \mathbb{P}_x(S_N - z_N \in (a, b) | \mathfrak{A}_{j,N}) \\ & \geq \liminf_{N \rightarrow \infty} \sqrt{2\pi V_N} \mathbb{P}_x(S_N - \bar{z}_{j,N} - b_N \in (a - \zeta_0 + \frac{3\bar{K}}{k}, b - \zeta_0 - \frac{3\bar{K}}{k}) | \mathfrak{A}_{j,N}) \\ & = \delta(f) e^{-z^2/2} \sum_{m \in \mathbb{Z}} \mathbb{E}_x[1_{(a - \zeta_0 + \frac{3\bar{K}}{k}, b - \zeta_0 - \frac{3\bar{K}}{k})}(m\delta(f) + \mathfrak{F})] \\ & \geq \left(1 - \frac{\delta}{L}\right) \left(|a - b| - \frac{6\bar{K}}{k}\right) e^{-z^2/2}. \end{aligned} \quad (5.2.31)$$

We now multiply these bounds by  $\mathbb{P}_x[\mathfrak{A}_{j,N}]$  and sum over  $j$ . This gives

$$\begin{aligned} & \liminf_{N \rightarrow \infty} \sqrt{2\pi V_N} \mathbb{P}_x \left( [S_N - z_N \in (a, b)] \bigcap \bigcup_{j=1}^k \mathfrak{A}_{j,N} \right) \\ & \geq \left(1 - \frac{\delta}{L}\right) \left(|a - b| - \frac{6\bar{K}}{k}\right) e^{-z^2/2} \left(1 - \frac{1}{k}\right). \end{aligned}$$

Passing to the limit  $k \rightarrow \infty$ , we obtain

$$\liminf_{N \rightarrow \infty} \sqrt{2\pi V_N} \mathbb{P}_x([S_N - z_N \in (a, b)]) \geq \left(1 - \frac{\delta}{L}\right) e^{-z^2/2} |a - b|,$$

and the lower bound is proved.

To prove the last statement of the theorem let  $\mathcal{A}_x$  be the positive functional on  $C_c(\mathbb{R})$  defined by (5.1.4), and let  $\mu_{\mathcal{A}_x}$  be the Radon measure on  $\mathbb{R}$  s.t.  $\mu_{\mathcal{A}_x}(\phi) = \mathcal{A}_x[\phi]$  for  $\phi \in C_c(\mathbb{R})$ .

The inequalities (5.2.30), (5.2.31) can be used to see that

$$(1 - \delta L^{-1})(|a - b| - O(1)) \leq \mu_{\mathcal{A}_x}(a, b) \leq (1 + 21\delta L^{-1})(|a - b| + O(1)),$$

whence  $\lim_{L \rightarrow \infty} \frac{\mu_{\mathcal{A}_x}[0, L]}{L} = 1$ . Since  $\mu_{\mathcal{A}_x}$  is clearly invariant under translation by  $\delta(f)$ , it must be the case that for each  $a$ ,  $\mu_{\mathcal{A}_x}[a, a + \delta(f)] = \delta(f)$ , whence

$$\forall k \in \mathbb{N} \quad \mu_{\mathcal{A}_x}([a, a + \delta(kf)]) = k\delta(f). \quad (5.2.32)$$

Given an interval  $(a, b)$  of length  $L$  with  $k\delta(f) < L < (k+1)\delta(f)$  take two intervals  $I^-, I^+$  such that

$$I^- \subset (a, b) \subset I^+, \quad \mu_{\mathcal{A}_x}(\partial I^-) = \mu_{\mathcal{A}_x}(\partial I^+) = 0, \quad |I^-| = k\delta(f), \quad |I^+| = (k+1)\delta(f).$$

Next let  $\phi^-, \phi^+$  be continuous functions with compact support such that

$$1_{I^-} < \phi^- < 1_{[a,b]} < \phi^+ < 1_{I^+}.$$

Then for large  $N$ ,  $\sqrt{V_N} \mathbb{P}(S_N - z_N \in (a, b))$  is sandwiched between  $\mathcal{A}(\phi^-)$  and  $\mathcal{A}(\phi^+)$  which in turn is sandwiched between

$$\mu_{\mathcal{A}}(I^-) = k\delta(f) \quad \text{and} \quad \mu_{\mathcal{A}}(I^+) = (k+1)\delta(f)$$

where the equalities rely on (5.2.32). The proof of the theorem is complete.  $\square$

### 5.2.5 Universal bounds for Markov arrays

Next, we give a different proof of universal lower and upper bounds, which does not rely on Theorem 5.1, and which also applies to arrays and to arbitrary initial distributions.

**Theorem 5.4.** *Let  $X$  be a uniformly elliptic Markov array, and  $f$  an a.s. uniformly bounded additive functional which is stably hereditary and not center tight. For every  $\varepsilon > 0$  there is  $N_\varepsilon > 0$  as follows. Suppose  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N \rightarrow \infty]{} z \in \mathbb{R}$ , and  $|a - b| > 2\delta(f) + \varepsilon$ , then for all  $N > N_\varepsilon$ ,*

$$\frac{1}{3} \left( \frac{e^{-z^2/2} |a - b|}{\sqrt{2\pi V_N}} \right) \leq \mathbb{P}(S_N - z_N \in (a, b)) \leq 3 \left( \frac{e^{-z^2/2} |a - b|}{\sqrt{2\pi V_N}} \right).$$

Recall that by our conventions, the **Fourier transform** of an  $L^1$  function  $\gamma: \mathbb{R} \rightarrow \mathbb{R}$  is  $\widehat{\gamma}(x) = \int_{-\infty}^{\infty} e^{-itx} \gamma(t) dt$ . Fix some  $b > 0$ , and define the Fourier pair

$$\psi_b(t) := \frac{\pi}{4b} 1_{[-b, b]}(t), \quad \widehat{\psi}_b(x) = \frac{\pi}{2b} \left( \frac{\sin(bx)}{x} \right).$$

**Lemma 5.8**  $1 \leq \widehat{\psi}_b(x) \leq \frac{\pi}{2b}$  for  $|x| \leq \frac{\pi}{2b}$ ; and  $|\widehat{\psi}_b(x)| < 1$  for  $|x| > \frac{\pi}{2b}$ .

*Proof.* The function  $\widehat{\psi}_b(x)$  is even, with zeroes at  $z_n = \pi n/b$ ,  $n \in \mathbb{Z} \setminus \{0\}$ . The critical points are  $c_0 = 0$  and  $\pm c_n$  where  $n \geq 1$  and

$$c_n := \text{the unique solution of } \tan(bc_n) = bc_n \text{ in } \left( z_n, z_n + \frac{\pi}{2b} \right).$$

It is easy to see that  $c_n = z_n + \frac{\pi}{2b} - o(1)$  as  $n \rightarrow \infty$ , and that

$$\text{sgn}[\widehat{\psi}_b(c_n)] = (-1)^n, \quad |\widehat{\psi}_b(c_n)| \leq \frac{1}{2n}, \quad \widehat{\psi}_b(c_n) \sim \frac{(-1)^n}{2n} \text{ as } n \rightarrow \infty.$$

So  $\widehat{\psi}_b$  attains global maximum  $\widehat{\psi}_b(0) = \frac{\pi}{2}$  at  $c_0$ , and  $|\widehat{\psi}_b(t)| \leq \frac{1}{2n}$  everywhere on  $[\pi n/b, \pi(n+1)/b]$ .

In particular,  $|\widehat{\psi}_b(t)| < 1/2$  for  $|t| \geq \pi/b$ . On  $(0, \pi/b)$  the function is decreasing from its global maximum  $\widehat{\psi}_b(0) = \frac{\pi}{2}$  to  $\widehat{\psi}_b(\frac{\pi}{b}) = 0$ , passing through  $\widehat{\psi}_b(\frac{\pi}{2b}) = 1$ . It follows that  $1 \leq \widehat{\psi}_b(t) \leq \frac{\pi}{2}$  on  $(0, \frac{\pi}{2b})$  and  $|\widehat{\psi}_b(t)| < 1$  for  $t > \frac{\pi}{2b}$ . The lemma follows, because  $\widehat{\psi}_b(-t) = \widehat{\psi}_b(t)$ .  $\square$

**Lemma 5.9** *There exist two continuous functions  $\gamma_1(x), \gamma_2(x)$  s.t.  $\text{supp}(\gamma_i) \subset [-2, 2]$ ;  $\gamma_1(0) > \frac{1}{3}$ ;  $\gamma_2(0) < 3$ ; and  $\widehat{\gamma}_1(x) \leq 1_{[-\pi, \pi]}(x) \leq \widehat{\gamma}_2(x)$  ( $x \in \mathbb{R}$ ).*

*Proof.* Throughout this proof,  $\psi^{*n} := \psi * \dots * \psi$  ( $n$  times), where  $*$  denotes the convolution. Let  $\gamma_1(t) := \frac{1}{4}[\psi_{\frac{1}{2}}^{*4}(t) - \psi_{\frac{1}{2}}^{*2}(t)]$ . Then  $\widehat{\gamma}_1(x) = \frac{1}{4}[\widehat{\psi}_{\frac{1}{2}}(x)^4 - \widehat{\psi}_{\frac{1}{2}}(x)^2]$ . By Lemma 5.8,  $1 \leq \widehat{\psi}_{\frac{1}{2}} \leq \frac{\pi}{2}$  on  $[-\pi, \pi]$  and  $|\widehat{\psi}_{\frac{1}{2}}| < 1$  outside  $[-\pi, \pi]$ . So

$$\max_{|x| \leq \pi} \widehat{\gamma}_1(x) \leq \max_{1 \leq y \leq \frac{\pi}{2}} \frac{1}{4} (y^4 - y^2) = \frac{1}{4} \left[ \left(\frac{\pi}{2}\right)^4 - \left(\frac{\pi}{2}\right)^2 \right] < 1,$$

$$\max_{|x| \geq \pi} \widehat{\gamma}_1(x) \leq \max_{|y| \leq 1} \frac{1}{4} (y^4 - y^2) = 0.$$

So  $\widehat{\gamma}_1(x) \leq 1_{[-\pi, \pi]}(x)$  for all  $x \in \mathbb{R}$ .

It is obvious from the definition of the convolution that

$$\text{supp}(\gamma_1) = \{x + y + z + w : x, y, z, w \in [-\frac{1}{2}, \frac{1}{2}]\} = [-2, 2].$$

Here is the calculation showing that  $\gamma_1(0) > \frac{1}{3}$ :

$$\begin{aligned} (\psi_b^{*2})(t) &= \frac{\pi^2}{16b^2} (1_{[-b, b]} * 1_{[-b, b]})(t) = \frac{\pi^2}{16b^2} 1_{[-2b, 2b]}(t)(2b - |t|) \\ (\psi_b^{*4})(0) &= (\psi_b^{*2} * \psi_b^{*2})(0) \\ &= \frac{\pi^4}{256b^4} \int_{-\infty}^{\infty} 1_{[-2b, 2b]}(t)(2b - |t|) 1_{[-2b, 2b]}(-t)(2b - |-t|) dt \\ &= \frac{\pi^4}{256b^4} \int_{-2b}^{2b} (2b - |t|)^2 dt = \frac{\pi^4}{128b^4} \int_0^{2b} (2b - t)^2 dt = \frac{\pi^4}{128b^4} \cdot \frac{(2b)^3}{3} = \frac{\pi^4}{48b}. \end{aligned}$$

So  $\psi_{\frac{1}{2}}^{*4}(0) = \frac{\pi^4}{24}$ ,  $\psi_{\frac{1}{2}}^{*2}(0) = \frac{\pi^2}{4}$ , and  $\gamma_1(0) = \frac{1}{4}(\frac{\pi^4}{24} - \frac{\pi^2}{4}) > \frac{1}{3}$ .

Next we set  $\gamma_2(t) := (\psi_{\frac{1}{2}} * \psi_{\frac{1}{2}})(t) \equiv \frac{\pi^2}{4} 1_{[-1, 1]}(t)(1 - |t|)$ . Then  $\text{supp}(\gamma_2) = [-1, 1]$  and  $\gamma_2(0) = \frac{\pi^2}{4} < 3$ . Finally,  $\widehat{\gamma}_2 \geq 1_{[-\pi, \pi]}(x)$ , because by Lemma 5.8,

- $\widehat{\gamma}_2(t) = (\widehat{\psi}_{\frac{1}{2}})^2(x) \geq 1$  for all  $|x| \leq \frac{\pi}{2 \cdot \frac{1}{2}} = \pi$ , and
- $\widehat{\gamma}_2(t) = (\widehat{\psi}_{\frac{1}{2}})^2(x) \geq 0$  for all  $|x| \geq \pi$ .

□

**Proof of Theorem 5.4.** If  $G_{ess}(X, f) = \mathbb{R}$  then the theorem follows from the LLT in the irreducible case. Otherwise (since  $f$  is not center-tight),  $G_{ess}(X, f) = t\mathbb{Z}$  for some  $t > 0$ , and there is no loss of generality in assuming that  $G_{ess}(X, f) = \mathbb{Z}$ .

Henceforth we assume that  $G_{ess}(X, f) = \mathbb{Z}$ . In this case our interval  $I := [a, b]$  has length bigger than 2. Notice that we can always center  $I$  by modifying  $z_N$  by a constant. So we may take our interval to be of the form

$$I = [-a, a], \text{ with } a > 1.$$

Let  $\gamma_i(t)$  be the functions constructed in Lemma 5.9, then

$$\widehat{\gamma}_1\left(\frac{\pi t}{a}\right) \leq 1_I(t) \leq \widehat{\gamma}_2\left(\frac{\pi t}{a}\right).$$

Therefore, for every choice of  $x_1^{(N)} \in \mathfrak{S}_1^{(N)}$  ( $N \geq 1$ ),

$$\begin{aligned} \mathbb{P}_{x_1^{(N)}}(S_N - z_N \in I) &= \mathbb{E}_{x_1^{(N)}}[1_I(S_N - z_N)] \geq \mathbb{E}_{x_1^{(N)}}\left[\widehat{\gamma}_1\left(\frac{\pi(S_N - z_N)}{a}\right)\right] \\ &= \mathbb{E}_{x_1^{(N)}}\left[\int_{-\infty}^{\infty} e^{-i\frac{\pi t}{a}(S_N - z_N)} \gamma_1(t) dt\right] = \int_{-\infty}^{\infty} \mathbb{E}_{x_1^{(N)}}(e^{-i\frac{\pi t}{a}(S_N - z_N)}) \gamma_1(t) dt. \end{aligned}$$

Recalling that  $\text{supp}(\gamma_1) \subset [-2, 2]$ , and substituting  $t = a\xi/\pi$ , we obtain

$$\mathbb{P}_{x_1^{(N)}}(S_N - z_N \in I) \geq \frac{|I|}{2\pi} \int_{-2\pi/a}^{2\pi/a} \mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)}) \gamma_1\left(\frac{a\xi}{\pi}\right) d\xi. \quad (5.2.33)$$

Similarly, we have

$$\mathbb{P}_{x_1^{(N)}}(S_N - z_N \in I) \leq \frac{|I|}{2\pi} \int_{-2\pi/a}^{2\pi/a} \mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)}) \gamma_2\left(\frac{a\xi}{\pi}\right) d\xi. \quad (5.2.34)$$

Next we claim that under the assumptions of Theorem 5.4:

**Lemma 5.10** *If  $G_{ess}(X, f) = \mathbb{Z}$  and  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N \rightarrow \infty]{} z \in \mathbb{R}$ , then for every  $a > 1$*

$$\sqrt{V_N} \int_{-2\pi/a}^{2\pi/a} \mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)}) \gamma_i\left(\frac{a\xi}{\pi}\right) d\xi \xrightarrow[N \rightarrow \infty]{} \sqrt{2\pi} e^{-\frac{1}{2}z^2} \gamma_i(0)$$

and the convergence is uniform in  $a$  on compact subsets of  $\mathbb{R} \setminus [-1, 1]$ .

*Proof.* In what follows we fix  $i \in \{1, 2\}$  and let  $\gamma(\xi) := \gamma_i\left(\frac{a\xi}{\pi}\right)$ . Divide  $[-\frac{2\pi}{a}, \frac{2\pi}{a}]$  into segments  $I_j$  of length at most  $\tilde{\delta}$ , where  $\tilde{\delta}$  is given by Lemma 5.5, making sure that  $I_0$  is centered at zero. Let

$$J_{j,N} := \int_{I_j} \mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)}) \gamma(\xi) d\xi.$$

CLAIM 1.  $\sqrt{V_N} J_{0,N} \xrightarrow[N \rightarrow \infty]{} \sqrt{2\pi} e^{-z^2/2} \gamma(0)$ .

*Proof.* The proof is similar to the proof of (4.2.17).

Applying Corollary 4.5 to the interval  $I_0$ , and noting that  $A_N(I_0) = 0$  and  $\tilde{\xi}_N = 0$  we find that

$$|\mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)})| \leq \tilde{C} \exp(-\tilde{\varepsilon} \xi^2 V_N).$$

So for every  $R > 1$ ,

$$\sqrt{V_N} \int_{\xi \in I_0: |\xi| > \frac{R}{\sqrt{V_N}}} \mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)}) \gamma(\xi) d\xi = O(e^{-\tilde{\varepsilon} R^2}).$$

Similarly, for all  $N$  large enough

$$\begin{aligned} \sqrt{V_N} \int_{\xi \in I_0: |\xi| \leq \frac{R}{\sqrt{V_N}}} \mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)}) \gamma(\xi) d\xi &= \int_{-R}^R \mathbb{E}_{x_1^{(N)}}(e^{-i\eta(\frac{S_N - z_N}{\sqrt{V_N}})}) \gamma\left(\frac{\eta}{\sqrt{V_N}}\right) d\eta \\ &= \int_{-R}^R \mathbb{E}_{x_1^{(N)}}(e^{-i\eta(\frac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}})}) e^{i\eta(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}})} \gamma_i\left(\frac{a\eta}{\pi\sqrt{V_N}}\right) d\eta \\ &\stackrel{!}{=} \int_{-R}^R e^{-\frac{1}{2}\eta^2 + i\eta z} \gamma(0) d\eta + o_{N \rightarrow \infty}(1) \text{ uniformly on compact sets of } a \\ &= \sqrt{2\pi} e^{-\frac{1}{2}z^2} \gamma(0) + o_{R \rightarrow \infty}(1) + o_{N \rightarrow \infty}(1), \end{aligned}$$

where  $\stackrel{!}{=}$  is a consequence of Dobrushin's CLT and the bounded convergence theorem. (When applying Dobrushin's Theorem it is useful to recall that by the exponential mixing of uniformly elliptic arrays,  $|\mathbb{E}(S_N) - \mathbb{E}_{x_1^{(N)}}(S_N)| = O(1)$ , therefore the condition  $\frac{z - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$  is equivalent to the condition  $(z - \mathbb{E}_{x_1^{(N)}}(S_N)) / \sqrt{V_N} \rightarrow z$ .) In summary,

$$\sqrt{V_N} J_{0,N} = \sqrt{2\pi} e^{-\frac{1}{2}z^2} \gamma(0) + o_{R \rightarrow \infty}(1) + o_{N \rightarrow \infty}(1).$$

Fixing  $R$ , we see that  $\limsup \sqrt{V_N} J_{0,N}$  and  $\liminf \sqrt{V_N} J_{0,N}$  are both equal to

$$\sqrt{2\pi} e^{-\frac{1}{2}z^2} \gamma(0) + o_{R \rightarrow \infty}(1).$$

Passing to the limit  $R \rightarrow \infty$  gives us that the limit exists and is equal to  $\sqrt{2\pi} e^{-\frac{1}{2}z^2} \gamma(0)$ .

It is not difficult to see that the convergence is uniform on compact subsets of  $a$ .

CLAIM 2.  $\sqrt{V_N} J_{j,N} \xrightarrow[N \rightarrow \infty]{} 0$  for every  $j \neq 0$ .

*Proof.* Since  $G_{ess}(f) = \mathbb{Z}$ , the co-range is  $H(f) = 2\pi\mathbb{Z}$ . So

$$I_j \subset [-\frac{2\pi}{a}, \frac{2\pi}{a}] \setminus \text{int}(I_0) \subset \text{a compact subset of } \mathbb{R} \setminus H(f).$$

This implies by the stable hereditary property of  $f$  that

$$D_N(\xi) \xrightarrow[N \rightarrow \infty]{} \infty \text{ uniformly on } I_j,$$

whence by (4.2.6),  $|\mathbb{E}_{x_1^{(N)}}(e^{-i\eta(S_N - z_N)})| \xrightarrow[N \rightarrow \infty]{} 0$  uniformly on  $I_j$ .

Let  $A_{j,N} := -\log\{\sup|\mathbb{E}_{x_1^{(N)}}(e^{-i\xi(S_N - z_N)})| : (x, \xi) \in \mathfrak{S}_1^{(N)} \times I_j\}$ , then  $A_{j,N} \xrightarrow[N \rightarrow \infty]{} \infty$ , and this divergence is uniform for  $a$  ranging over compact subsets of  $\mathbb{R} \setminus [-1, 1]$ .

From this point onward, the proof of the claim is identical to the proof of (4.2.18). We omit the details.

The Lemma follows by summing over all subintervals  $I_j$  in  $[-\frac{2\pi}{a}, \frac{2\pi}{a}]$ , and noting that the number of these intervals is uniformly bounded (by  $1 + \frac{4\pi}{\delta}$ ).  $\square$

We now return to the proof of theorem. Lemma 5.10, (5.2.33), (5.2.34), and the inequalities  $\gamma_1(0) > \frac{1}{3}$  and  $\gamma_2(0) < 3$  imply that for every choice of  $\{x_1^{(N)}\}_{N \geq 1}$ , for all  $N$  sufficiently large

$$\frac{1}{3} \cdot \frac{|I|}{\sqrt{2\pi V_N}} e^{-z^2/2} \leq \mathbb{P}_{x_1^{(N)}}(S_N - z_N \in I) \leq 3 \cdot \frac{|I|}{\sqrt{2\pi V_N}} e^{-z^2/2}. \quad (5.2.35)$$

This estimate is uniform in  $\{x_1^{(N)}\}_{N \geq 1}$ : There is an  $N_0$  such that (5.2.36) holds for all  $N \geq N_0$  and for all choices of  $\{x_1^{(N)}\}_{N \geq 1}$ . Otherwise, there exist  $N_k \rightarrow \infty$  and  $x_1^{(N_k)} \in \mathfrak{S}_1^{(N_k)}$  which violate (5.2.36). But then (5.2.36) fails for any choice of  $x_1^{(N)}$  which contains  $x_1^{(N_k)}$  as a subsequence, whereas (5.2.36) holds for all possible choices.

Since (5.2.36) holds uniformly in  $\{x_1^{(N)}\}_{N \geq 1}$ , we can integrate and deduce that for all  $N$  sufficiently large

$$\frac{1}{3} \cdot \frac{|I|}{\sqrt{2\pi V_N}} e^{-z^2/2} \leq \mathbb{P}(S_N - z_N \in I) \leq 3 \cdot \frac{|I|}{\sqrt{2\pi V_N}} e^{-z^2/2} \quad (5.2.36)$$

for any initial distributions  $\mu_1^{(N)}(dx_1^{(N)})$  on  $\mathfrak{S}_1^{(N)}$ .  $\square$

We end this section by recording a useful consequence of the previous proof: The upper bound in Theorem 5.4 does not require any information about the arithmetic properties of  $f$ .

**Lemma 5.11** *For each  $K, \varepsilon_0$  and  $\ell$  there is a constant  $C^* = C^*(K, \varepsilon_0, \ell)$  s.t. if  $f$  is an additive functional of a uniformly elliptic Markov chain with ellipticity constant  $\varepsilon_0$ , and if  $|f| \leq K$ , then for every  $x \in \mathfrak{S}_1$ ,  $N \geq 1$ , and for each interval  $J$  of length  $\ell$ ,*

$$\mathbb{P}_x(S_N \in J) \leq \frac{C^*}{\sqrt{V_N}}.$$

*Proof.* It suffices to prove the result for  $\ell = 4$  since longer intervals could be covered by a finite number of intervals of length 4. Thus  $J = z_N + I$  with  $I = [-2, 2]$ . Applying (5.2.34) with  $a = 2$  we get

$$\mathbb{P}_x(S_N \in J) \leq \hat{C} \int_{-\pi}^{\pi} |\Phi_N(x, -\xi)| d\xi.$$

where  $\hat{C} = \frac{2}{\pi} \|\gamma_2\|_\infty$ . Dividing  $[-\pi, \pi]$  into finitely many subintervals of length  $\tilde{\delta}/2$  where  $\tilde{\delta}$  comes from Lemma 4.4, and applying (4.2.19) on each subinterval we obtain the result.  $\square$

### 5.3 Notes and references

Dolgopyat proved a version of Theorem 5.1 for sums of independent random variables. The connection between the LLT and uniform distribution modulo  $t$  was considered for sums of independent random variables by Prohorov [122], Rozanov [128], and Gamkrelidze [58].

The question of estimating  $\mathbb{P}[S_N - z_N \in (a, b)]$  is related to the study of the rate of convergence in the CLT. In particular, a Berry-Esseen type result on the rate of convergence in the CLT would certainly imply that  $\exists M$  s.t. for all  $|a - b| > M$ , if  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow z$ , then for all  $N$  large enough,

$\mathbb{P}[S_N - z_N \in (a, b)]$  equals  $\frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}$  up to bounded multiplicative error. Such results were shown to us by Y. Hafouta. The Berry-Esseen approach has the advantage of giving information on the time  $N$  when the universal estimates kick in, but has the disadvantage that it only applies to very large intervals (how large depends on the growth of the third moment of  $S_N$ ). By contrast, the results of this chapter apply to intervals of length  $> \delta(f)$ , which is optimal, but do not say on how large  $N$  should be for the estimates to work.

Lemma 5.11 for the sums of independent random variables appears in [117, Section III.1]. The proof in the Markov case is essentially the same.



## Chapter 6

# Local limit theorems for large and moderate deviations

In this chapter we prove the local limit theorem in the regimes of moderate and large deviations. In these cases the asymptotic behavior of  $\mathbb{P}(S_N - z_N \in (a, b))$  is determined by the “rate functions,” the Legendre transforms of the log-moment generating functions of  $S_N$ .

### 6.1 The moderate deviations and large deviations regimes

Suppose  $f$  is an irreducible, a.s. uniformly bounded, additive functional on a uniformly elliptic Markov chain  $X$ , with algebraic range  $\mathbb{R}$  or  $t\mathbb{Z}$  with  $t > 0$ . Let

$$S_N = f_1(X_1, X_2) + \cdots + f_N(X_N, X_{N+1}), \quad V_N := \text{Var}(S_N).$$

In the previous chapters, we analyzed  $\mathbb{P}(S_N - z_N \in (a, b))$  as  $N \rightarrow \infty$ , in the **regime of local deviations**,  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}} \rightarrow \text{const}$ . In this chapter we ask what happens when  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}} \rightarrow \infty$ .

Usually in the literature the large deviations regime is defined by the condition  $|z_N - \mathbb{E}(S_N)| \geq \bar{\varepsilon} \text{Var}(S_N)$  for some fixed  $\bar{\varepsilon} > 0$ . However, to get meaningful results we need to assume some upper bounds  $|z_N - \mathbb{E}(S_N)|$  as well. We will study the following regimes:

(1) **Moderate deviations:**  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}} \rightarrow \infty$  and  $z_N - \mathbb{E}(S_N) = o(\text{Var}(S_N))$ ,

(2) **Large deviations:**  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\text{Var}(S_N)}} \rightarrow \infty$  and  $|z_N - \mathbb{E}(S_N)| \leq \varepsilon \text{Var}(S_N)$  for some  $\varepsilon > 0$  “small enough.”

In some cases we can take  $\varepsilon = \infty$ , see e.g. §7.3, but in others  $\varepsilon$  must really be finite, see Example 6.21. To see why it is forced on us, let us consider a few examples of what might go wrong when  $|z_N - \mathbb{E}(S_N)|/\text{Var}(S_N)$  is “too big.”

If  $\frac{z_N - \mathbb{E}(S_N)}{V_N}$  grows too fast, e.g. if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} > \frac{2\text{ess sup}|S_N|}{V_N}$ , then the probabilities  $\mathbb{P}[S_N - z_N \in (0, \infty)]$  are all equal to zero, and our problem is vacuous. A more subtle but related issue arises when  $\frac{z_N - \mathbb{E}(S_N)}{V_N}$  falls at the boundary of the domain of the Legendre transforms of  $t \mapsto \frac{1}{V_N} \log \mathbb{E}(e^{t(S_N - \mathbb{E}(S_N))})$ . Why this matters will be clear once we explain the strategy of our proofs (see the end of §6.3.1 and §6.4). At this point we can only present an example:

**Example 6.1** If  $\frac{z_N - \mathbb{E}(S_N)}{V_N}$  falls near the boundary of the domain of the Legendre transforms of  $t \mapsto \frac{1}{V_N} \log \mathbb{E}(e^{t(S_N - \mathbb{E}(S_N))})$ , then the behavior of  $\mathbb{P}[S_N - z_N \in (a, b)]$  may depend not just on  $\lim_{N \rightarrow \infty} \frac{z_N - \mathbb{E}(S_N)}{V_N}$  but also on  $z_N$  itself.

*Proof.* Let  $S_N := X_1 + \dots + X_N$  where  $X_i$  are identically distributed independent random variables equal to  $-1, 0, 1$  with equal probabilities. Here  $\mathbb{E}(S_N) = 0$ ,  $V_N = 2N/3$ , the Legendre transforms of the log-moment generating functions have domains  $(-\frac{3}{2}, \frac{3}{2})$ , and the classical theory of large deviations says that if  $z \in (-\frac{3}{2}, \frac{3}{2})$ , then  $\lim_{z_N/V_N \rightarrow z} \frac{1}{V_N} \log \mathbb{P}[S_N - z_N > 0]$  exists and is finite. But no such conclusion holds when  $z = \frac{3}{2}$ :

- If  $z_N = N$ , then  $[S_N - z_N > 0] = \emptyset$  and  $\frac{1}{V_N} \log \mathbb{P}[S_N - z_N > 0] = -\infty$ ;
- If  $z_N = N - 1$ , then  $[S_N - z_N > 0] = [S_N = N]$ , and  $\frac{1}{V_N} \log \mathbb{P}[S_N - z_N > 0] = -\frac{3}{2} \log 3$ .

So the limit depends on how  $z_N/V_N$  approaches  $\frac{3}{2}$ , and it could be infinite.  $\square$

For general additive functionals on Markov chains (homogeneous or not), we do not know how to determine the asymptotic behavior of  $\mathbb{P}[S_N - z_N \in (a, b)]$  when  $\frac{z_N}{V_N}$  is close to  $\partial C_N$ , where

$$C_N := \text{domain of the Legendre transform of } \frac{1}{V_N} \log \mathbb{E}(e^{t(S_N - \mathbb{E}(S_N))}).$$

We can only analyze the case where  $\frac{z_N - \mathbb{E}(S_N)}{\text{Var}(S_N)}$  is well inside the interior of  $C_N$  for all  $N$ . This is why we must assume that  $|z_N - \mathbb{E}(S_N)| \leq \varepsilon \text{Var}(S_N)$  for  $\varepsilon$  small enough.

It is instructive to compare the regime of large deviations to the regime of the LLT from the point of view of universality.

The asymptotic behavior of  $\mathbb{P}[S_N - z_N \in (a, b)]$  in the regime of local deviations does not depend on the details of the distributions of  $f_n(X_n, X_{n+1})$ . It depends only on rough features such as  $\text{Var}(S_N)$ , the algebraic range, and (in case the algebraic range is  $t\mathbb{Z}$ ) on the constants  $c_N$  s.t.  $S_N \in c_N + t\mathbb{Z}$  almost surely.

By contrast, in the regime of large deviations the asymptotic behavior of  $\mathbb{P}[S_N - z_N \in (a, b)]$  depends on the entire distribution of  $S_N$ . The dependence is through the Legendre transform of  $\log \mathbb{E}(e^{tS_N})$ , a function which encodes the entire distribution of  $S_N$ , not just its rough features.

We will consider two partial remedies to the lack of universality:

- (a) *Conditioning*: The conditional distributions of  $S_N - z_n$  given that  $S_N - z_N > a$  has a universal scaling limit, see Corollary 6.4.
- (b) *Moderate deviations*: If  $|z_N - \mathbb{E}(S_N)| = o(\text{Var}(S_N))$ , then  $\mathbb{P}[S_N - z_N \in (a, b)]$  have universal lower and upper bounds (Theorems 6.3, 6.4).

## 6.2 Local limit theorems for large deviations

### 6.2.1 The log moment generating functions

Suppose  $|f| < K$  almost surely. For every  $N$  such that  $V_N \neq 0$ , we define the **normalized log moment generating function** of  $S_N$  to be

$$\mathcal{F}_N(\xi) := \frac{1}{V_N} \log \mathbb{E}(e^{\xi S_N}) \quad (\xi \in \mathbb{R}).$$

The a.s. uniform boundedness of  $f$  guarantees the finiteness of the expectation, and the real analyticity of  $\mathcal{F}_N(\xi)$  on  $\mathbb{R}$ .

**Example 6.2 (Sums of iid's)**

Suppose that  $S_N = \sum_{n=1}^N X_n$  where  $X_n$  where  $X_N$  are i.i.d. bounded random variables with non-zero variance. Let  $X$  denote the common law of  $X_n$ . Then

$$\mathcal{F}_N(\xi) = \mathcal{F}_X(\xi) := \frac{1}{\text{Var}(X)} \log \mathbb{E}(e^{\xi X})$$

is independent of  $n$ . In addition,

- (i)  $\mathcal{F}_X(\xi)$  is strictly convex, by Hölder's inequality and because  $X \neq \text{const}$  a.s. Since  $\mathcal{F}_X(\xi)$  is smooth, its second derivative must be bounded away from zero on compacts. So  $\mathcal{F}_N(\xi)$  are uniformly strictly convex on compacts.
- (ii)  $\lim_{\xi \rightarrow -\infty} \mathcal{F}'_N(\xi) = \text{ess inf}(X)/\text{Var}(X)$ ,  $\lim_{\xi \rightarrow +\infty} \mathcal{F}'_N(\xi) = \text{ess sup}(X)/\text{Var}(X)$ . To see this, use convexity to see that  $\lim \mathcal{F}'_N(\xi)$  are the slopes of the asymptotes of  $\mathcal{F}_X(\xi)$ , or equivalently  $\lim \frac{1}{\xi} \mathcal{F}_N(\xi)$ . The last limits can be easily found to be equal to  $\text{ess sup}(X)/\text{Var}(X)$  as  $\xi \rightarrow \infty$ , and  $\text{ess inf}(X)/\text{Var}(X)$  as  $\xi \rightarrow -\infty$ .

Properties (i) and (ii) play a key role in the study of large deviations for sums of i.i.d. random variables. A significant part of the effort in this chapter is to understand to which extent similar results holds in the setting of bounded additive functionals of uniformly elliptic Markov chains. We start with the following facts.

**Theorem 6.1.** *Let  $f$  be an a.s. uniformly bounded additive functional of a uniformly elliptic Markov chain  $X$ , and assume  $V_N \neq 0$  for all  $N \geq N_0$ , then*

- (1) *For all  $N \geq N_0$ ,  $\mathcal{F}_N(0) = 0$ ,  $\mathcal{F}'_N(0) = \frac{\mathbb{E}(S_N)}{V_N}$ ,  $\mathcal{F}''_N(0) = 1$ .*
- (2) *For every  $N \geq N_0$ ,  $\mathcal{F}_N(\xi)$  is strictly convex on  $\mathbb{R}$ .*
- (3) *The convexity is uniform on compacts: For every  $R > 0$  there is  $C = C(R)$  positive s.t. for all  $N \geq N_0$ ,  $C^{-1} \leq \mathcal{F}''_N(\xi) \leq C$  on  $[-R, R]$ .*
- (4) *Suppose  $V_N \rightarrow \infty$ . For every  $\varepsilon > 0$  there are  $\delta, N_\varepsilon > 0$  s.t. for all  $|\xi| \leq \delta$ ,  $N > N_\varepsilon$ , we have  $e^{-\varepsilon} \leq \mathcal{F}''_N(\xi) \leq e^\varepsilon$ , and*

$$e^{-\varepsilon} \frac{1}{2} \left( \xi - \frac{\mathbb{E}(S_N)}{V_N} \right)^2 \leq \mathcal{F}_N(\xi) - \frac{\mathbb{E}(S_N)}{V_N} \xi \leq e^\varepsilon \frac{1}{2} \left( \xi - \frac{\mathbb{E}(S_N)}{V_N} \right)^2.$$

This is very similar to what happens for iid's, but there is one important difference: In our setting  $V_N$  may be much smaller than  $N$ .

For the proof of this theorem see §6.3.5. Here is an immediate corollary:

**Corollary 6.3** *Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . If  $V_N := \text{Var}(S_N) \rightarrow \infty$ , then for all  $0 < \alpha < \frac{1}{2}$  and  $\kappa > 0$ , if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \sim \kappa V_N^{-\alpha}$  as  $N \rightarrow \infty$ , then*

$$\lim_{N \rightarrow \infty} \frac{1}{V_N^{2\alpha-1}} \log \mathbb{P}[S_N - z_N \geq 0] = -\frac{1}{2} \kappa^2.$$

*Proof.* There is no loss of generality in assuming that  $\mathbb{E}(S_N) = 0$  for all  $N$ . Let  $a_n := V_n^{1-2\alpha}$ ,  $b_n := V_n^\alpha$ ,  $W_n := S_n/b_n$ . Then  $a_n \rightarrow \infty$ , whence by Theorem 6.1(4),

$$\mathcal{F}(\xi) := \lim_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{E}(e^{\xi W_n}) = \lim_{n \rightarrow \infty} V_n^{2\alpha} \mathcal{F}_N\left(\frac{\xi}{V_n^\alpha}\right) = \frac{1}{2} \xi^2.$$

We may now use the Gärtner-Ellis Theorem see e.g. [55, Thm II.6.1]) and  $\frac{z_n}{a_n b_n} \rightarrow \kappa$  to deduce that  $\lim_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{P}[S_n - z_n \geq 0] = \lim_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{P}\left[\frac{W_n}{a_n} \geq \frac{z_n}{a_n b_n}\right] = -\frac{1}{2} \kappa^2$ .  $\square$

### 6.2.2 The rate functions

Suppose  $V_N \neq 0$ . The **rate functions**  $\mathcal{I}_N(\eta)$  are the **Legendre transforms** of  $\mathcal{F}_N(\xi)$ . Specifically, let  $a_N := \inf \mathcal{F}'_N$  and  $b_N := \sup \mathcal{F}'_N$ ; then  $\mathcal{I}_N : (a_N, b_N) \rightarrow \mathbb{R}$  is

$$\mathcal{I}_N(\eta) := \xi \eta - \mathcal{F}_N(\xi) \text{ for the unique } \xi \text{ s.t. } \mathcal{F}'_N(\xi) = \eta.$$

The existence and uniqueness of  $\xi$  is because of the smoothness and strict convexity of  $\mathcal{F}_N$  on  $\mathbb{R}$ . We call  $(a_N, b_N)$  the **domain** of  $\mathcal{I}_N$ , and denote it by

$$\text{dom}(\mathcal{I}_N) := (a_N, b_N).$$

Equivalently,  $\text{dom}(\mathcal{I}_N) = (\mathcal{F}'(-\infty), \mathcal{F}'(+\infty))$ , where  $\mathcal{F}'(\pm\infty) := \lim_{t \rightarrow \pm\infty} \mathcal{F}'(t)$ . Later we will also need the sets  $(a_N^R, b_N^R) \subset \text{dom}(\mathcal{I}_N)$ , where  $R > 0$  and

$$a_N^R := \mathcal{F}'_N(-R), \quad b_N^R := \mathcal{F}'_N(R). \quad (6.2.1)$$

The functions  $\mathcal{I}_N$  and their domains depend on  $N$ . The following theorem identifies certain uniformity and universality in their behavior.

**Theorem 6.2.** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ , and assume  $V_N \neq 0$  for all  $N$  large enough, then*

- (1)  $\exists c, N_1, R > 0$  s.t. for all  $N > N_1$ ,  $\text{dom}(\mathcal{I}_N) \supset [a_N^R, b_N^R] \supseteq \left[ \frac{\mathbb{E}(S_N)}{V_N} - c, \frac{\mathbb{E}(S_N)}{V_N} + c \right]$ .
- (2) For each  $R$  there exists  $\rho = \rho(R)$  s.t.  $\rho^{-1} \leq \mathcal{I}_N'' \leq \rho$  on  $[a_N^R, b_N^R]$  for all  $N > N_1$ .
- (3) Suppose  $V_N \rightarrow \infty$ . For every  $\varepsilon > 0$  there exists  $\delta > 0$  and  $N_\varepsilon$  such that for all  $\eta \in [\frac{\mathbb{E}(S_N)}{V_N} - \delta, \frac{\mathbb{E}(S_N)}{V_N} + \delta]$  and  $N > N_\delta$ ,

$$e^{-\varepsilon} \frac{1}{2} \left( \eta - \frac{\mathbb{E}(S_N)}{V_N} \right)^2 \leq \mathcal{I}_N(\eta) \leq e^\varepsilon \frac{1}{2} \left( \eta - \frac{\mathbb{E}(S_N)}{V_N} \right)^2.$$

- (4) Suppose  $V_N \rightarrow \infty$  and  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then

$$V_N \mathcal{I}_N \left( \frac{z_N}{V_N} \right) = \frac{1 + o(1)}{2} \left( \frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \right)^2 \quad \text{as } N \rightarrow \infty.$$

The proof of the theorem will be given in §6.3.6.

The significance of part (4) will become apparent in §6.2.3.

### 6.2.3 The LLT for moderate deviations.

Recall that the state spaces of  $X$  are denoted by  $\mathfrak{S}_i$  ( $i \geq 1$ ), and that  $\mathbb{P}_x$  denotes the conditional probability given  $X_1 = x$ .

**Theorem 6.3.** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . Suppose  $f$  is irreducible with algebraic range  $\mathbb{R}$ . If  $z_N \in \mathbb{R}$  satisfy  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow 0$ , then for every non-empty  $(a, b)$  and  $x \in \mathfrak{S}_1$ ,*

$$\begin{aligned}\mathbb{P}_x[S_N - z_N \in (a, b)] &= [1 + o(1)] \frac{|a - b|}{\sqrt{2\pi V_N}} \exp\left(-V_N \mathcal{J}_N\left(\frac{z_N}{V_N}\right)\right) \text{ as } N \rightarrow \infty, \\ \mathbb{P}_x[S_N - z_N \in (a, b)] &= [1 + o(1)] \frac{|a - b|}{\sqrt{2\pi V_N}} \exp\left[-\frac{1 + o(1)}{2} \left(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2\right] \text{ as } N \rightarrow \infty.\end{aligned}$$

**Theorem 6.4.** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . Assume  $f$  is irreducible with algebraic range  $\mathbb{Z}$ , and  $S_N \in c_N + \mathbb{Z}$  almost surely. If  $z_N \in c_N + \mathbb{Z}$  satisfy  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \rightarrow 0$ , then for every  $x \in \mathfrak{S}_1$ ,*

$$\begin{aligned}\mathbb{P}_x[S_N = z_N] &= \frac{[1 + o(1)]}{\sqrt{2\pi V_N}} \exp\left(-V_N \mathcal{J}_N\left(\frac{z_N}{V_N}\right)\right) \text{ as } N \rightarrow \infty, \\ \mathbb{P}_x[S_N = z_N] &= \frac{[1 + o(1)]}{\sqrt{2\pi V_N}} \exp\left[-\frac{1 + o(1)}{2} \left(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2\right] \text{ as } N \rightarrow \infty.\end{aligned}$$

We will obtain these results as special cases of a more complicated and general asymptotic relation which we will state in the next section.

The two asymptotic relations in Theorems 6.3 and 6.4 complement each other. The first is a precise asymptotic, but it is not universal, because it is expressed in terms of the rate functions, which depend on the fine details of the distributions of  $S_N$ . The second is universal, but it is not an asymptotic equivalence because the right-hand-side is only determined up to a multiplicative error of size  $\exp[o(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}})^2]$ .

### 6.2.4 The LLT for large deviations.

Recall the definition of the subsets  $(a_N^R, b_N^R) := (\mathcal{F}'_N(-R), \mathcal{F}'_N(R)) \subset \text{dom}(\mathcal{J}_N)$  from (6.2.1). It is convenient to define

$$[\widehat{a}_N^R, \widehat{b}_N^R] := \left[a_N^R - \frac{\mathbb{E}(S_N)}{\sqrt{V_N}}, b_N^R - \frac{\mathbb{E}(S_N)}{\sqrt{V_N}}\right].$$

**Theorem 6.5.** *Let  $f$  be an a.s. uniformly bounded, irreducible, additive functional on a uniformly elliptic Markov chain  $X$ . For every  $R$  large enough there are functions  $\rho_N : \mathfrak{S}_1 \times [\widehat{a}_N^R, \widehat{b}_N^R] \rightarrow \mathbb{R}^+$ ,  $\xi_N : [\widehat{a}_N^R, \widehat{b}_N^R] \rightarrow \mathbb{R}$  as follows:*

(1)  $\exists c > 0$  such that  $[\widehat{a}_N^R, \widehat{b}_N^R] \supset [-c, c]$  for all  $N$  large enough.

(2) **Non Lattice case:** Suppose  $G_{\text{alg}}(X, f) = \mathbb{R}$ , then for every sequence of  $z_N \in \mathbb{R}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ , for all finite non-empty intervals  $(a, b)$ , and for every  $x \in \mathfrak{S}_1$ , we have the following

asymptotic as  $N \rightarrow \infty$ :

$$\mathbb{P}_x[S_N - z_N \in (a, b)] = [1 + o(1)] \cdot \frac{e^{-V_N \mathcal{J}_N(\frac{z_N}{V_N})}}{\sqrt{2\pi V_N}} \rho_N \left( x, \frac{z_N - \mathbb{E}(S_N)}{V_N} \right) \int_a^b e^{-t \xi_N \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right)} dt.$$

(3) **Lattice case:** Suppose  $G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{Z}$  and  $S_N \in c_N + \mathbb{Z}$  a.s., then for every sequence of  $z_N \in c_N + \mathbb{Z}$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$ , for all finite non-empty intervals  $(a, b)$  and  $x \in \mathfrak{S}_1$ , the following asymptotic holds when  $N \rightarrow \infty$ :

$$\mathbb{P}_x[S_N - z_N \in (a, b)] = [1 + o(1)] \cdot \frac{e^{-V_N \mathcal{J}_N(\frac{z_N}{V_N})}}{\sqrt{2\pi V_N}} \rho_N \left( x, \frac{z_N - \mathbb{E}(S_N)}{V_N} \right) \cdot \sum_{t \in (a, b) \cap \mathbb{Z}} e^{-t \xi_N \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right)}.$$

(4) **Properties of the error terms:**

- (a)  $\rho_N(x, \eta)$  are bounded away from 0,  $\infty$  on  $\mathfrak{S}_1 \times [\hat{a}_N^R, \hat{b}_N^R]$  uniformly in  $N$ , and  $\rho_N(x, \eta) \xrightarrow[\eta \rightarrow 0]{} 1$  uniformly in  $N$  and  $x$ .
- (b) For each  $R > 0$  there exists  $C = C_R > 0$  such that for all  $\eta \in [\hat{a}_N^R, \hat{b}_N^R]$  and  $N$ ,  $C^{-1}|\eta| \leq |\xi_N(\eta)| \leq C|\eta|$  and  $\text{sgn}(\xi(\eta)) = \text{sgn}(\eta)$ .

The proof of this result will occupy us in §§6.3.1–6.3.7.

Theorem 6.5 above assumes irreducibility. Without this assumption we have a following weaker bound.

**Theorem 6.6.** Let  $K := \text{ess sup} |\mathsf{f}|$ , and suppose  $V_N \rightarrow \infty$ . For each  $\varepsilon, R$  there is  $D(\varepsilon, R, K)$  and  $N_0$  such that for all  $z_N \in [\mathcal{F}'_N(\varepsilon), b_N^R]$  and  $N > N_0$ ,

$$D^{-1} \leq \frac{\sqrt{V_N} \mathbb{P}(S_N \geq z_N)}{e^{-V_N \mathcal{J}_N(\frac{z_N}{V_N})}} \leq D.$$

To assist the reader in digesting the statement of Theorem 6.5, we now explain how to use it to obtain Theorems 6.3, 6.4 on moderate deviations, as well as other consequences.

**Proof of Theorems 6.3 and 6.4:** By Theorem 6.5(1),  $\exists R > 0$  s.t. if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  for all  $N$  large enough, and

$$\rho_N(x, \frac{z_N - \mathbb{E}(S_N)}{V_N}) \xrightarrow[N \rightarrow \infty]{} 1, \quad \xi_N(\frac{z_N - \mathbb{E}(S_N)}{V_N}) \rightarrow 0, \quad \frac{1}{b-a} \int_a^b e^{-t \xi_N(\frac{z_N - \mathbb{E}(S_N)}{V_N})} dt \rightarrow 1.$$

Suppose  $G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$ , then theorem 6.5(2) implies that

$$\mathbb{P}[S_N - z_N \in (a, b)] \sim \frac{|a - b|}{\sqrt{2\pi V_N}} \exp(-V_N \mathcal{J}_N(z_N/V_N)).$$

Next, by Theorem 6.2(2), if  $\frac{z_n - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then

$$V_N \mathcal{J}_N \left( \frac{z_N}{V_N} \right) \sim \frac{1}{2} \left( \frac{z_n - \mathbb{E}(S_N)}{\sqrt{V_N}} \right)^2,$$

whence  $\mathbb{P}[S_N - z_N \in (a, b)] \sim \frac{|a - b|}{\sqrt{2\pi V_N}} \exp(-\frac{1+o(1)}{2} (\frac{z_n - \mathbb{E}(S_N)}{\sqrt{V_N}})^2)$ . This proves Theorem 6.3. The proof of Theorem 6.4 is similar, and we leave it to the reader.  $\square$

Here are some other consequences of Theorem 6.5.

**Corollary 6.4** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Suppose  $f$  is irreducible, with algebraic range  $\mathbb{R}$ .*

- (1) *If  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$  then for any finite non empty interval  $(a, b)$  the distribution of  $S_N - z_N$  conditioned on  $S_N - z_N \in (a, b)$  is asymptotically uniform on  $(a, b)$ .*
- (2) *If  $\liminf \frac{z_N - \mathbb{E}(S_N)}{V_N} > 0$  and there exists  $R$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  for all sufficiently large  $N$ , then the distribution of*

$$\xi_N \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right) \cdot (S_N - z_N) \text{ conditioned on } S_N \geq z_N$$

*is asymptotically exponential with parameter 1.*

*Remark.* The condition in (2) is satisfied whenever  $\liminf \frac{z_N - \mathbb{E}(S_N)}{V_N} > 0$ , and  $\limsup \frac{z_N - \mathbb{E}(S_N)}{V_N} > 0$  is small enough, see Theorem 6.5(1).

*Proof.* To see part (1), note first that if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then  $\xi_N = \xi_N \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right) \rightarrow 0$ , whence  $\frac{1}{\beta - \alpha} \int_\alpha^\beta e^{-t\xi_N} dt \xrightarrow[N \rightarrow \infty]{} 1$  for every non-empty interval  $(\alpha, \beta)$ . Thus by Theorem 6.5, for every interval  $[c, d] \subset [a, b]$ ,

$$\lim_{N \rightarrow \infty} \frac{\mathbb{P}_x[S_N - z_N \in (c, d)]}{\mathbb{P}_x[S_N - z_N \in (a, b)]} = \frac{|c - d|}{|a - b|}.$$

(the prefactors  $\rho_N$  are identical, and they cancel out).

To see part (2), note first that our assumptions on  $z_N$  guarantee that  $\xi_N = \xi_N \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right)$  is bounded from away from zero and infinity, and that all its limit points are strictly positive.

Suppose  $\xi_{N_k} \rightarrow \xi$ . Then arguing as in part (1) it is not difficult to see that for all  $(a, b) \subset (0, \infty)$  and  $r > 0$ ,

$$\lim_{k \rightarrow \infty} \frac{\mathbb{P}_x[\xi_{N_k}(S_{N_k} - z_{N_k}) \in (a + r, b + r) | S_{N_k} > z_{N_k}]}{\mathbb{P}_x[\xi_{N_k}(S_{N_k} - z_{N_k}) \in (a, b) | S_{N_k} > z_{N_k}]} = e^{-r}.$$

Since this is true for all convergent  $\{\xi_{N_k}\}$ , and since any subsequence of  $\{\xi_N\}$  has a convergent subsequence,

$$\begin{aligned} \liminf_{N \rightarrow \infty} \frac{\mathbb{P}_x[\xi_N(S_N - z_N) \in (a + r, b + r) | S_N > z_N]}{\mathbb{P}_x[\xi_N(S_N - z_N) \in (a, b) | S_N > z_N]} &= e^{-r}, \\ \limsup_{N \rightarrow \infty} \frac{\mathbb{P}_x[\xi_N(S_N - z_N) \in (a + r, b + r) | S_N > z_N]}{\mathbb{P}_x[\xi_N(S_N - z_N) \in (a, b) | S_N > z_N]} &= e^{-r}, \end{aligned}$$

and so  $\lim_{N \rightarrow \infty} \frac{\mathbb{P}_x[\xi_N(S_N - z_N) \in (a + r, b + r) | S_N > z_N]}{\mathbb{P}_x[\xi_N(S_N - z_N) \in (a, b) | S_N > z_N]} = e^{-r}$ . So conditioned on  $S_N > z_N$ ,  $\xi_N(S_N - z_N)$  is asymptotically exponential with parameter 1.  $\square$

**Corollary 6.5** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Suppose  $f$  is irreducible, with algebraic range  $\mathbb{Z}$ . Let  $z_N$  be a sequence of integers.*

- (1) *If  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$  then for any  $a < b$  in  $\mathbb{Z}$  the distribution of  $S_N - z_N$  conditioned on  $S_N - z_N \in [a, b]$  is asymptotically uniform on  $[a, b]$ .*

(2) If  $\liminf \frac{z_N - \mathbb{E}(S_N)}{V_N} > 0$  and there exists  $R$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  for all sufficiently large  $N$ ,  $\xi_N \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right) \rightarrow \xi$ , then

$(S_N - z_N)$  conditioned on  $S_N \geq z_N$

is asymptotically geometric with parameter  $e^{-\xi}$ .

The proof is similar to the proof in the non-lattice case, so we omit it.

It is worthwhile to note the following consequence of this result. In the following statement, “local distribution” means a functional on  $C_c(\mathbb{R})$  and “vague convergence” means convergence on all continuous functions with compact support.

**Corollary 6.6** *Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Let  $z_N$  be a sequence s.t. for some  $R$ ,  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  for large  $N$ . Let  $\zeta_N$  be the local distribution of  $S_N$  around  $z_N$ , that is  $\zeta_N(\phi) = \mathbb{E}_x(\phi(S_N - z_N))$ . Let  $\zeta$  be a vague limit of  $\{q_N \zeta_N\}$  for some sequence  $q_N > 0$ . If  $f$  is irreducible then  $\zeta$  has density  $c_1 e^{c_2 t}$  with respect to the Haar measure on the algebraic range of  $f$  for some  $c_1 \in \mathbb{R}_+, c_2 \in \mathbb{R}$ .*

If the restriction  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  is dropped, then it is likely that  $\zeta$  is either as above, or an atomic measure with one atom, but our methods are insufficient for proving this.

### 6.3 Proofs

We prove Theorems 6.1, 6.2, 6.5 and 6.6. (Theorems 6.3 and 6.4 are direct consequences, and were proved in §6.2.4.)

We assume throughout that  $\{X_n\}$  is a uniformly elliptic Markov chain with state spaces  $\mathfrak{S}_n$ , transition probabilities  $\pi_{n,n+1}(x, dy)$ , and stationary distributions  $\mu_k(E) := \mathbb{P}(X_k \in E)$ . Let  $f = \{f_n\}$  be an a.s. uniformly bounded additive functional on  $X$ . Let  $\varepsilon_0$  denote the ellipticity constant of  $X$ , and  $K = \text{ess sup } |f|$ .

#### 6.3.1 Strategy of proof

The proof can be briefly described as an implementation of “change of measure” technique (aka “Cramér’s transform”).

We explain the idea. Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ , and let  $z_N$  be as in Theorem 6.5. We will modify the transition probabilities of  $X = \{X_n\}$  to generate a Markov array  $\tilde{X} = \{\tilde{X}_n^{(N)}\}$  whose row sums  $\tilde{S}_N = f_1(\tilde{X}_1^{(N)}, \tilde{X}_2^{(N)}) + \dots + f_N(\tilde{X}_N^{(N)}, \tilde{X}_{N+1}^{(N)})$  satisfy

$$z_N - \mathbb{E}(\tilde{S}_N) = o\left(\sqrt{\text{Var}(\tilde{S}_N)}\right). \quad (6.3.1)$$

(6.3.1) places us in the regime of local deviations which we have analyzed in Chapter 4. The results of that chapter provide asymptotics for  $\mathbb{P}(\tilde{S}_N - z_N \in (a, b))$ , and these can be translated into asymptotics for  $\mathbb{P}(S_N - z_N \in (a, b))$ .

The array  $\tilde{X}$  is constructed from  $(X, f)$  as follows: Let  $\mathfrak{S}_n$  and  $\pi_{n,n+1}(x, dy)$  denote the state spaces and transition probabilities of the original Markov chain  $X$ , then we take  $f_n^{(N)} = f_n$ ,  $\mathfrak{S}_n^{(N)} = \mathfrak{S}_n$ , and we let  $\tilde{X}$  be the Markov array with state spaces  $\mathfrak{S}_n^{(N)}$  and transition probabilities

$$\tilde{\pi}_{n,n+1}^{(N)}(x, dy) := e^{\xi_N f_n(x, y)} \frac{h_{n+1}(y, \xi_N)}{e^{p_n(\xi_N)} h_n(x, \xi_N)} \cdot \pi_{n,n+1}(x, dy).$$

Here  $\xi_N$  is a parameter that is calibrated to get (6.3.1), and  $p_n, h_n, h_{n+1}$  are chosen to guarantee that  $\tilde{\pi}_{n,n+1}^{(N)}(x, dy)$  has total mass equal to one. This technique is called a “change of measure.”

The value of  $\xi_N$  depends on  $\frac{z_N - \mathbb{E}(S_N)}{V_N}$ . To construct  $\xi_N$  and to control it, we must know that  $\frac{z_N}{V_N}$  belong to a sets where  $\mathcal{F}_N$  are strictly convex, uniformly in  $N$ . This is the reason why we need to assume that  $\exists R$  s.t.  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  for all  $N$ , a condition we can check as soon as  $|\frac{z_N - \mathbb{E}(S_N)}{V_N}| < c$  with  $c$  small enough.<sup>1</sup>

We remark that the dependence of  $\xi_N$  on  $N$  means that  $\{\tilde{X}_n^{(N)}\}$  is an array, not a chain. The fact that the change of measure produces arrays from chains is the reason we insisted on working with arrays in the first part of this work.

### 6.3.2 A parameterized family of changes of measure

In this section we construct, for an arbitrary given sequence of constants  $\xi_N \in \mathbb{R}$ , transition probabilities of the form

$$\tilde{\pi}_{n,n+1}^{(N)}(x, dy) := e^{\xi_N f_n(x, y)} \frac{h_{n+1}(y, \xi_N)}{e^{p_n(\xi_N)} h_n(x, \xi_N)} \cdot \pi_{n,n+1}(x, dy), \quad (6.3.2)$$

where  $p_n(\xi_N)$  are real numbers and  $h_k^{\xi_N}(\cdot) = h_k(\cdot, \xi_N)$  are positive functions on  $\mathfrak{S}_k$  which are chosen to guarantee that  $\tilde{\pi}_{n,n+1}^{(N)}(x, dy)$  has total mass equal to one.

We treat the sequence of parameters  $\xi_N$  as arbitrary. In the next section we will explain how to choose a particular  $\{\xi_N\}$  to guarantee (6.3.1).

**Lemma 6.7** *Given  $\xi \in \mathbb{R}$  and a sequence of real numbers  $\{a_n\}_{n \in \mathbb{N}}$ , there are unique numbers  $p_n(\xi) \in \mathbb{R}$ , and unique non-negative  $h_n(\cdot, \xi) \in L^\infty(\mathfrak{S}_n, \mathcal{B}(\mathfrak{S}_n), \mu_n)$  s.t.  $\int_{\mathfrak{S}_n} h_n(x, \xi) \mu_n(dx) = \exp(a_n \xi)$  for all  $n$ , and for a.e.  $x$*

$$\int_{\mathfrak{S}_{n+1}} e^{\xi f_n(x, y)} \frac{h_{n+1}(y, \xi)}{e^{p_n(\xi)} h_n(x, \xi)} \pi_{n,n+1}(x, dy) = 1. \quad (6.3.3)$$

*The unique solution is positive almost everywhere.*

*Remark.:* Notice that if  $\{\bar{h}_n(\cdot, \xi)\}$ ,  $\{\bar{p}_n(\xi)\}$  satisfy the Lemma with  $a_n = 0$ , then the unique solution with general  $\{a_n\}$  is given by

$$h_n(\cdot, \xi) := e^{a_n \xi} \bar{h}_n(\cdot, \xi), \quad p_n(\xi) := \bar{p}_n(\xi) - a_n \xi + a_{n+1} \xi. \quad (6.3.4)$$

<sup>1</sup> Other situations where the condition  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  can be checked are discussed in §6.4.

Evidently,  $h_n, p_n$  give rise to the same probability kernel (6.3.2) as do  $\bar{h}_n, \bar{p}_n$ . We call  $\{\bar{h}_n\}$  and  $\{\bar{p}_n\}$  the *fundamental solution*.

*Proof.* It is enough to prove the existence and uniqueness of the fundamental solution, so henceforth we assume  $a_n = 0$ . We may also assume without loss of generality that  $|\xi| \leq 1$ , else scale  $f$ .

Set  $V_n := L^\infty(\mathfrak{S}_n, \mathcal{B}(\mathfrak{S}_n), \mu_n)$ , and define operators  $L_n^\xi : V_{n+1} \rightarrow V_n$  by

$$(L_n^\xi h)(x) = \int_{\mathfrak{S}_{n+1}} e^{\xi f_n(x,y)} h(y) \pi_{n,n+1}(x, dy). \quad (6.3.5)$$

The operators  $L_n^\xi$  are linear, bounded, and positive.

For (6.3.3) to hold, it is necessary and sufficient that  $h_n^\xi(\cdot) := h_n(\cdot, \xi)$  be positive a.e., and  $L_n^\xi h_{n+1}^\xi = e^{p_n(\xi)} h_n^\xi$  for some  $p_n(\xi) \in \mathbb{R}$ .

Positivity everywhere may be replaced by the weaker property that  $h_n^\xi \in L^\infty \setminus \{0\}$  are all non-negative a.e., because for such functions, since  $|f| \leq K$  a.s. and  $X$  is uniformly elliptic with ellipticity constant  $\varepsilon_0$ ,

$$h_n^\xi(x) = e^{-p_n(\xi) - p_{n+1}(\xi)} (L_n^\xi L_{n+1}^\xi h_{n+2}^\xi)(x) \geq e^{-p_n(\xi) - p_{n+1}(\xi) - 2K} \varepsilon_0 \|h_{n+2}^\xi\|_1.$$

Thus to prove the lemma it is enough to find a sequence numbers  $p_n(\xi) \in \mathbb{R}$  and non-negative  $h_n^\xi \in L^\infty \setminus \{0\}$  such that  $L_n^\xi h_{n+1}^\xi = e^{p_n(\xi)} h_n^\xi$  for some  $p_n(\xi) \in \mathbb{R}$ .

The existence and uniqueness of such “generalized eigenvectors” can be proved as in [56], [15], [81] using **Hilbert’s projective metrics**. We recall what these are. Let  $C_n := \{h \in V_n : h \geq 0 \text{ a.e.}\}$ . These are closed cones and  $L_n^\xi(C_{n+1}) \subset C_n$ . Define

$$d_n(h, g) := \log \left( \frac{M(h|g)}{m(h|g)} \right) \in [0, \infty], \quad (h, g \in C_n),$$

where  $M = M(f|g), m = m(f|g)$  are the best constants in the estimate  $mh \leq f \leq Mh$ . This is a pseudo-metric on the interior of  $C_n$ , and  $d(h, g) = 0 \Leftrightarrow h, g$  are proportional. Also, for all  $h, g \in C_n \setminus \{0\}$ ,

$$\left\| \frac{h}{\int h} - \frac{g}{\int g} \right\|_1 \leq e^{d_n(h,g)} - 1. \quad (6.3.6)$$

Birkhoff’s theorem [13] says that any linear map  $T : C_{n+2} \rightarrow C_n$  such that the  $d_n$ -diameter of  $T(C_{n+2})$  in  $C_n$  is less than some  $\Delta > 0$ , contracts the Hilbert’s projective metric at least by a factor  $\theta := \tanh(\Delta/4) \in (0, 1)$ .

We will apply Birkhoff’s theorem to the linear transformations

$$T_n^\xi := L_n^\xi L_{n+1}^\xi : C_{n+2} \rightarrow C_n.$$

One checks using the standing assumptions and  $|\xi| \leq 1$  that

$$e^{-2K} \varepsilon_0 \|h\|_1 \leq (T_n^\xi h)(x) \leq e^{2K} \varepsilon_0^{-2} \|h\|_1 \quad (h \in C_{n+2}), \quad (6.3.7)$$

whence  $d_n(T_n^\xi h, 1) \leq 4K + 3 \log(1/\varepsilon_0)$ . So the diameter of  $T_n^\xi(C_{n+2})$  in  $C_n$  is less than  $\Delta := 8K + 6 \log(1/\varepsilon_0)$ . Hence by Birkhoff’s Theorem mentioned above,

$$d_n(T_{n+1}^\xi h, T_{n+1}^\xi g) \leq \theta d_{n+2}(h, g) \quad (h, g \in C_{n+2}). \quad (6.3.8)$$

where  $\theta := \tanh(2K + \frac{3}{2} \log(1/\varepsilon_0)) \in (0, 1)$ .

It follows that for every  $n$ ,  $\{L_n^\xi L_{n+1}^\xi \cdots L_{n+k-1}^\xi 1_{\mathfrak{S}_{n+k}}\}_{k \geq 1} \subset C_n$  is a Cauchy sequence with respect to  $d_n$ . By (6.3.6),

$$\frac{L_n^\xi L_{n+1}^\xi \cdots L_{n+k-1}^\xi 1_{\mathfrak{S}_{n+k}}}{\|L_n^\xi L_{n+1}^\xi \cdots L_{n+k-1}^\xi 1_{\mathfrak{S}_{n+k}}\|_1}$$

is a Cauchy sequence in  $L^1$ .

The limiting function  $h_n^\xi$  has integral one, and is positive and bounded, because of (6.3.7). Clearly,  $L_n^\xi h_{n+1}^\xi = e^{p_n} h_n^\xi$  for some  $p_n \in \mathbb{R}$ . So  $\{h_n^\xi\}, \{p_n\}$  exist.

Moreover, the proof shows that  $\text{diam} \left( \bigcap_{k \geq 1} L_n^\xi \cdots L_{n+k-1}^\xi (C_{n+k}) \right) = 0$ . It follows that  $h_n^\xi$  is unique up to multiplicative constant, whence by the normalization condition, unique. The lemma is proved.  $\square$

The proof has the following consequence, which we mention for future reference: For every  $R > 0$ , there exists  $C_0 > 0$  and  $\theta \in (0, 1)$  (depending on  $R$ ) such that for every  $|\xi| \leq R$

$$d_1 \left( L_1^\xi \cdots L_N^\xi h_{N+1}^\xi, L_1^\xi \cdots L_N^\xi 1 \right) \leq C_0 \theta^{N/2} d_{N+1} \left( h_{N+1}^\xi, 1 \right). \quad (6.3.9)$$

The case when  $N$  is even follows directly from (6.3.8) and does not require the constant  $C_0$ . The case of odd  $N$  is obtained from the even case by using the exponential contraction of  $L_2^\xi \cdots L_N^\xi$  and the fact that one additional application of  $L_1^\xi$  (or any other positive linear operator) does not increase the Hilbert norm. This implies (6.3.9) with  $C_0 := \theta^{-1/2}$ .

**Lemma 6.8** *Let  $h_n^\xi(\cdot) = h(\cdot, \xi)$  be as in Lemma 6.7. If  $a_n$  is bounded, then for every  $R > 0$  there is  $C = C(R)$  s.t. for all  $n \geq 1$ , a.e.  $x \in \mathfrak{S}_n$  and  $|\xi| \leq R$ ,*

$$C^{-1} \leq h_n(x, \xi) \leq C \quad \text{and} \quad C^{-1} < e^{p_n(\xi)} < C.$$

*Proof.* It is enough to consider the fundamental solution ( $a_n = 0$ ,  $\int h_n = 1$ ); the general case follows from (6.3.4). It is also sufficient to consider the case  $|\xi| \leq 1$ ; the general case follows by scaling  $f$ .

Let  $\{h_n^\xi\}$  be the fundamental solution, then in the notation of the previous proof,  $T_n^\xi h_{n+2}^\xi = e^{p_n(\xi) + p_{n+1}(\xi)} h_n^\xi$ , whence by (6.3.7),

$$e^{-2K} \varepsilon_0 \leq e^{p_n(\xi) + p_{n+1}(\xi)} h_{n+2}^\xi \leq e^{2K} \varepsilon_0^{-2}.$$

Integrating, and recalling that  $\int h_{n+2}^\xi d\mu_{n+2} = \exp(a_{n+2}\xi) = 1$ , we obtain

$$e^{-2K} \varepsilon_0 \leq e^{p_n(\xi) + p_{n+1}(\xi)} \leq e^{2K} \varepsilon_0^{-2}.$$

So  $e^{-4K} \varepsilon_0^2 \leq h_n^\xi(\cdot) \leq e^{4K} \varepsilon_0^{-4}$ .

Observe that  $e^{p_n} = \int L_n^\xi h_{n+1}^\xi d\mu_{n+1} = e^{\pm K} \int h_{n+1}^\xi d\mu_{n+1}$ . So  $e^{p_n}$  is also uniformly bounded away from zero and infinity.  $\square$

In the next section we will choose  $\xi_N$  to guarantee (6.3.1), and as it turns out, the choice involves a condition on  $\frac{\partial p_n}{\partial \xi}$ . Later, we will also require information on  $\frac{\partial^2 p_n}{\partial \xi^2}$ . In preparation for this, we will now study the differentiability of

$$\xi \mapsto h_n^\xi \text{ and } \xi \mapsto p_n(\xi).$$

The map  $\xi \mapsto h_n^\xi$  takes values in the Banach space  $L^\infty$ . To analyze it, we will use the theory of **real-analytic maps into Banach spaces** [43].

Let us briefly review this theory. Suppose  $\mathfrak{X}, \mathfrak{Y}$  are Banach spaces. Let  $a_n : \mathfrak{X}^n \rightarrow \mathfrak{Y}$  be a multilinear map. The **norm** of  $a_n$  is

$$\|a_n\| := \sup\{\|a_n(x_1, \dots, x_n)\| : x_i \in \mathfrak{X}, \|x_i\| \leq 1 \text{ for all } i\}.$$

A multilinear map is called **symmetric** if it is invariant under the permutation of its coordinates. Given  $x \in \mathfrak{X}$ , we denote

$$a_n x^n := a_n(x, \dots, x).$$

A **power series** is a formal expression  $\sum_{n \geq 1} a_n x^n$  where  $a_n : \mathfrak{X}^n \rightarrow \mathfrak{Y}$  are multilinear and symmetric.

A function  $\phi : \mathfrak{X} \rightarrow \mathfrak{Y}$  is called **real analytic** at  $x_0$  if there is some  $r > 0$  and a power series  $\sum a_n x^n$  (called the **Taylor series at  $x_0$** ) such that  $\sum \|a_n\| r^n < \infty$  and

$$\phi(x) = \phi(x_0) + \sum_{n \geq 1} a_n(x - x_0)^n$$

whenever  $\|x - x_0\| < r$ . One can check that if this happens, then

$$a_n(x_1, \dots, x_n) = \frac{1}{n!} \left. \frac{d}{dt_1} \right|_{t_1=0} \dots \left. \frac{d}{dt_n} \right|_{t_n=0} \phi(x_0 + \sum_{i=1}^n t_i x_i). \quad (6.3.10)$$

Conversely, if  $\sum a_n(x - x_0)^n$  has positive radius of convergence with  $a_n$  as in (6.3.10), then  $\phi$  is real-analytic, and equal to its Taylor series  $\phi(x_0) + \sum a_n(x - x_0)^n$  on a neighborhood of  $x_0$ .

**Example 6.9** Let  $\phi : \mathfrak{X} \times \mathfrak{X} \times \mathbb{R} \rightarrow \mathfrak{X}$  be the map  $\phi(x, y, z) := x - y/z$ . Then  $\phi$  is real-analytic at every  $(x_0, y_0, z_0)$  such that  $z_0 \neq 0$ , with Taylor series

$$\phi(x, y, z) = \phi(x_0, y_0, z_0) + \sum_{n=1}^{\infty} a_n(x - x_0, y - y_0, z - z_0)^n,$$

where  $\|a_n\| = O(\|y_0\|/|z_0|^{n+1}) + O(n/|z_0|^{n+1})$ .

*Proof.* If  $|z - z_0| < |z_0|$ , then  $x - y/z = x - \frac{y}{z_0} \sum_{k \geq 0} (-1)^k \frac{1}{z_0^k} (z - z_0)^k$ . For each  $n \geq 1$ ,  $\underline{x}_0 := (x_0, y_0, z_0)$ ,  $\underline{x}_i := (x_i, y_i, z_i)$  ( $1 \leq i \leq n$ ), and  $(t_1, \dots, t_n) \in \mathbb{R}^n$ ,

$$\phi(\underline{x}_0 + \sum_{i=1}^n t_i \underline{x}_i) = x_0 + \sum_{i=1}^n t_i x_i + \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{z_0^{k+1}} \left( y_0 + \sum_{i=1}^n t_i y_i \right) \left( \sum_{i=1}^n t_i z_i \right)^k$$

converges in norm whenever  $(t_1, \dots, t_n) \in A_n := [|\sum_{i=1}^n t_i z_i| < |z_0|]$ . In particular, on  $A_n$ , this series is real-analytic in each  $t_i$ , and can be differentiated term-by-term infinitely many times.

To find  $a_n(\underline{x}_1, \dots, \underline{x}_n)$  we observe that the differential (6.3.10) is equal to the coefficient of  $t_1 \cdots t_n$  in the previous series. So for  $n > 2$ ,

$$a_n(\underline{x}_1, \dots, \underline{x}_n) = \frac{(-1)^{n+1} y_0}{z_0^{n+1}} \cdot z_1 \cdots z_n + \frac{(-1)^n}{z_0^n} \sum_{i=1}^n y_i z_1 \cdots \widehat{z_i} \cdots z_n$$

where the hat above  $z_i$  indicates that the  $i$ -th term should be omitted. It follows that  $\|a_n\| = O(\|y_0\|/|z_0|^{n+1}) + O(n/|z_0|^n)$ .  $\square$

**Lemma 6.10** *The functions  $\xi \mapsto h_n^\xi, p_n(\xi)$  are real-analytic. If  $a_n$  is bounded, then for every  $R > 0$  there is  $C(R) > 0$  s.t. for every  $|\xi| \leq R$  and  $n \geq 1$ ,*

$$\left\| \frac{\partial}{\partial \xi} h_n(\cdot, \xi) \right\|_\infty \leq C(R), \quad \left\| \frac{\partial^2}{\partial \xi^2} h_n(\cdot, \xi) \right\|_\infty \leq C(R).$$

*Proof.* The proof is based on §3.3 in [54], although it is somewhat simpler because our setup is more elementary.

It is enough to consider the special case  $R = 1$  and  $a_n = 0$ . In particular,  $\int h_n^\xi = 1$ .

Fix  $|\xi| \leq 1$  and let  $T_n := T_n^\xi$ ,  $h_n(\cdot) = h_n(\cdot, \xi)$  be as in the proof of Lemma 6.7. Define two Banach spaces:

$$X := \left\{ (S_n)_{n \in \mathbb{N}} : \begin{array}{l} S_n : L^\infty(\mathfrak{S}_{n+2}) \rightarrow L^\infty(\mathfrak{S}_n) \text{ are bounded linear} \\ \text{operators, and } \|S\| := \sup_n \|S_n\| < \infty \end{array} \right\}$$

$$Y := \{(\varphi_n)_{n \in \mathbb{N}} : \varphi_n \in L^\infty(\mathfrak{S}_{n+2}), \|\varphi\| := \sup \|\varphi_n\|_\infty < \infty\}$$

Using (6.3.7), it is not difficult to see that  $T := (T_n)$  belongs to  $X$ . By Lemma 6.8,  $h := (h_n)_{n \in \mathbb{N}}$  belongs to  $Y$ .

STEP 1. *There exists  $0 < \delta < 1$  s.t. for every  $(S, \varphi) \in X \times Y$ , for all  $|\xi| \leq 1$ , if  $\|S - T\| < \delta$  and  $\|\varphi - h\| < \delta$ , then  $\inf |\int (S_n \varphi_{n+2})| > \delta$ .*

*Proof.* By (6.3.7),  $\|T_n\| \leq M$  where  $M := e^{2K} \varepsilon_0^{-2}$ , and by Lemma 6.8, there is a constant  $\varepsilon_1 > 0$  so that for all  $n$  and  $|\xi| \leq 1$

$$\varepsilon_1 \leq (T_n h_{n+2})(x) \leq \varepsilon_1^{-1}.$$

So if  $\|S - T\| < \delta$  and  $\|\varphi - h\| < \delta$ , then for a.e.  $x$ ,

$$\begin{aligned} S_n \varphi_{n+2}(x) &= (T_n h_{n+2})(x) - (T_n - S_n) h_{n+2}(x) - S_n (h_{n+2} - \varphi_{n+2})(x) \\ &\geq \varepsilon_1 - \|T - S\| \|h\| - (\|S - T\| + \|T\|) \|\varphi - h\| \\ &\geq \varepsilon_1 - \delta \|h\| - (\delta + M) \delta. \end{aligned}$$

Let  $C$  be a uniform upper bound for  $\|h\|$  which holds for all  $|\xi| \leq 1$ . If  $0 < \delta < (\frac{\varepsilon_1}{C+M+2}) \wedge 1$ , then  $S_n \varphi_{n+2} > \delta$  a.e., and the step follows.

Henceforth we fix  $\delta$  as in step 1. Let  $B_\delta(T) := \{S \in X : \|S - T\| < \delta\}$  and  $B_\delta(h) := \{\varphi \in Y : \|\varphi - h\| < \delta\}$ , and define

$$\Upsilon : B_\delta(T) \times B_\delta(h) \rightarrow Y, \quad \Upsilon(S, \varphi) := \left( \varphi_n - \frac{S_n \varphi_{n+2}}{\int (S_n \varphi_{n+2}) d\mu_{n+2}} \right)_{n \in \mathbb{N}}.$$

This is well-defined by the choice of  $\delta$ , and  $\Upsilon(T, h) = 0$ .

STEP 2.  *$\Upsilon$  is real-analytic on  $B_\delta(T) \times B_\delta(h)$ .*

*Proof.* First we write  $\Upsilon = \Phi(\Upsilon^{(1)}, \Upsilon^{(2)}, \Upsilon^{(3)})$  with

$$\circ \quad \Upsilon^{(1)} : X \times Y \rightarrow Y, \quad \Upsilon^{(1)}(S, \varphi) = \varphi$$

- o  $\Upsilon^{(2)} : X \times Y \rightarrow Y$ ,  $\Upsilon^{(2)}(S, \varphi) = (S_n \varphi_{n+2})_{n \in \mathbb{N}}$ .
- o  $\Upsilon^{(3)} : X \times Y \rightarrow \ell^\infty$ ,  $\Upsilon^{(3)}(S, \varphi) = (\int (S_n \varphi_{n+2}) d\mu_{n+2})_{n \in \mathbb{N}}$ .
- o  $\Phi : \{(\varphi, \psi, \xi) \in Y \times Y \times \ell^\infty : \inf |\xi_i| > 0\} \rightarrow Y$ ,

$$\Phi((\varphi, \psi, \xi)_{i \geq 1}) = (\varphi_i - \xi_i^{-1} \psi_i)_{i \geq 1}.$$

By step 1,  $\vec{\Upsilon} := (\Upsilon^{(1)}, \Upsilon^{(2)}, \Upsilon^{(3)})$  maps  $B_\delta(T) \times B_\delta(h)$  into

$$U := \{(\varphi, \psi, \xi) \in Y \times Y \times \ell^\infty : \|\varphi\| < C + \delta, \|\psi\| < M + \delta, \inf |\xi_i| > \delta/2\},$$

whence into the domain of  $\Phi$ .

We claim that for each of the functions  $\Upsilon^{(i)}$ , some high enough derivative of  $\Upsilon^{(i)}$  is identically zero. Let  $D$  be the derivative, and let  $D_i$  be the partial derivative with respect to the  $i$ -th variable, then

(1)  $\Upsilon^{(1)}$  is linear, so  $(D\Upsilon^{(1)})(S, \varphi)$  is constant, and  $D^2\Upsilon^{(1)} = 0$ .

(2)  $\Upsilon^{(2)} : X \times Y \rightarrow Y$ ,  $\Upsilon^{(2)}(S, \varphi) = (S_n \varphi_{n+2})_{n \in \mathbb{N}}$ . Here

$$\begin{aligned} (D_1\Upsilon^{(2)})(S, \varphi)(S') &= (S'_n \varphi_{n+2})_{n \in \mathbb{Z}}, & (D_1^2\Upsilon^{(2)})(S, \varphi) &= 0 \\ (D_2\Upsilon^{(2)})(S, \varphi)(\varphi') &= (S_n \varphi'_{n+2})_{n \in \mathbb{Z}}, & (D_2^2\Upsilon^{(2)})(S, \varphi) &= 0 \\ (D_1 D_2 \Upsilon^{(2)})(S, \varphi)(S', \varphi') &= (S'_n \varphi'_{n+2})_{n \in \mathbb{Z}} \end{aligned}$$

We see that  $D^2\Upsilon^{(2)}$  does not depend on  $(S, \varphi)$ , so  $D^3\Upsilon = 0$

(3)  $\Upsilon^{(3)} : X \times Y \rightarrow \ell^\infty$ ,  $\Upsilon^{(3)}(S, \varphi) = (\int (S_n \varphi_{n+2}) d\mu_{n+2})_{n \in \mathbb{N}}$ . As before, the third derivative is zero.

Consequently,  $\Upsilon^{(i)}$  are real-analytic on its domain (with finite Taylor series at every point). Next we show that  $\Phi$  is real-analytic on  $U$ . To do this we recall that by Example 6.9,  $x - \frac{y}{z} = \sum_{n=0}^{\infty} a_n(x_0, y_0, z_0)(x - x_0, y - y_0, z - z_0)^n$  where  $a_n(x_0, y_0, z_0) : (\mathbb{R}^3)^n \rightarrow \mathbb{R}$  are symmetric multilinear functions depending on  $(x_0, y_0, z_0)$ , s.t.  $\|a_n(x_0, y_0, z_0)\| = O(|y_0|/|z_0|^{n+1}) + O(n/|z_0|^n)$ . So

$$\Phi(\varphi, \psi, \xi) = \Phi(\varphi^{(0)}, \psi^{(0)}, \xi^{(0)}) + \sum_{n=1}^{\infty} A_n(\varphi - \varphi^{(0)}, \psi - \psi^{(0)}, \xi - \xi^{(0)})^n, \quad (6.3.11)$$

where  $A_n : (Y \times Y \times \ell^\infty)^n \rightarrow Y$ , has entries

$$\begin{aligned} A_n((\varphi^{(1)}, \psi^{(1)}, \xi^{(1)}), \dots, (\varphi^{(n)}, \psi^{(n)}, \xi^{(n)}))_i(x) &:= \\ a_n(\varphi_i^{(0)}(x), \psi_i^{(0)}(x), \xi_i^{(0)})(\varphi_i^{(1)}(x), \psi_i^{(1)}(x), \xi_i^{(1)}), \dots, (\varphi_i^{(n)}(x), \psi_i^{(n)}(x), \xi_i^{(n)}) \end{aligned}$$

$A_n$  inherits multilinearity and symmetry from  $a_n$ , and by construction,

$$\|A_n(\varphi^{(0)}, \psi^{(0)}, \xi^{(0)}))\| \leq \sup \left\{ \|a_n(x_0, y_0, z_0)\| : |x_0|, |y_0| \leq C + M + \delta, |z_0| > \frac{\delta}{2} \right\} = O(2^n n / \delta^n).$$

So the right-hand-side of (6.3.11) has positive radius of convergence, proving the analyticity of  $\Phi : U \rightarrow Y$ .

The step follows from the well-known result that the composition of real-analytic functions is real-analytic, see [43].

CLAIM 4.  $(D_2\Upsilon)(T, h) : Y \rightarrow Y$ , the partial derivative of  $\Upsilon$  at  $(T, h)$  with respect to the second variable, has bounded inverse.

*Proof.* A direct calculation shows that  $(D_2\Upsilon)(T, h)(\varphi) = \varphi - \Lambda\varphi$ , where

$$(\Lambda\varphi)_n = \frac{T_n\varphi_{n+2}}{\int(T_nh_{n+2})d\mu_n} - \left( \frac{\int(T_n\varphi_{n+2})d\mu_n}{\int(T_nh_{n+2})d\mu_n} \right) h_n.$$

To prove the claim, we show that  $\Lambda$  has spectral radius  $< 1$ .

Let  $T_n^{(k)} := T_n T_{n+2} \cdots T_{n+2(k-1)}$ , then we claim that

$$(\Lambda^k\varphi)_n = \frac{T_n^{(k)}\varphi_{n+2k}}{\int(T_n^{(k)}h_{n+2k})d\mu_n} - \left( \frac{\int(T_n^{(k)}\varphi_{n+2k})d\mu_n}{\int(T_n^{(k)}h_{n+2k})d\mu_n} \right) h_n. \quad (6.3.12)$$

To see this we first note, using  $T_m h_{m+2} \propto h_m$  and  $\int h_m d\mu_m = 1$ , that

$$\int(T_n^{(k+1)}h_{n+2(k+1)})d\mu_n = \int(T_nh_{n+2})d\mu_n \int(T_{n+2}^{(k)}h_{n+2(k+1)})d\mu_{n+2}.$$

With this identity in mind, the formula for  $\Lambda^k$  follows by induction.

We now explain why (6.3.12) implies that the spectral radius of  $\Lambda$  is less than one. Fix  $\varphi \in Y$ . Recall that  $C^{-1} \leq h_n \leq C$  for all  $n$ , and let

$$\psi := \varphi + 2C\|\varphi\|h.$$

Then  $\psi \in Y$ ,  $\Lambda^k\psi = \Lambda^k\varphi$  for all  $k$  (because  $\Lambda h = 0$ ), and for all  $n$

$$C\|\varphi\|h_n \leq \psi_n \leq 3C\|\varphi\|h_n \quad (6.3.13)$$

In particular, if  $C_n$  is the cone from the proof of Lemma 6.7, and  $d_n$  is its projective Hilbert metric, then  $\psi_n \in C_n$  and  $d_n(\psi_n, h_n) \leq \log 3$ . Since  $T_n$  contracts the Hilbert projective norm by a factor  $\theta \in (0, 1)$ ,

$$d_n(T_n^{(k)}\psi_{n+2k}, T_n^{(k)}h_{n+2k}) \leq \theta^k \log 3.$$

This implies by the definition of  $d_n$  that for a.e.  $x \in \mathfrak{S}_n$ ,

$$\left| \frac{(T_n^{(k)}\psi_{n+2k})(x) / \int(T_n^{(k)}\psi_{n+2k})}{(T_n^{(k)}h_{n+2k})(x) / \int(T_n^{(k)}h_{n+2k})} - 1 \right| \leq \max\{3^{\theta^k} - 1, 1 - 3^{-\theta^k}\} = 3^{\theta^k} - 1 =: \varepsilon_k.$$

The denominator simplifies to  $h_n$ . So

$$\left\| \frac{(T_n^{(k)}\psi_{n+2k})}{\int(T_n^{(k)}\psi_{n+2k})} - h_n \right\|_{\infty} \leq \varepsilon_k \|h\|. \quad (6.3.14)$$

Next we use the positivity of  $T_n^{(k)}$  and (6.3.13) to note that

$$C\|\varphi\|T_n^{(k)}h_{n+2k} \leq T_n^{(k)}\psi_{n+2k} \leq 3C\|\varphi\|T_n^{(k)}h_{n+2k}.$$

We deduce that

$$C\|\varphi\| \leq \frac{\int(T_n^{(k)}\psi_{n+2k})}{\int(T_n^{(k)}h_{n+2k})} \leq 3C\|\varphi\|. \quad (6.3.15)$$

By (6.3.12), (6.3.14) and (6.3.15),

$$\begin{aligned} \|\Lambda^k \varphi\|_\infty &\equiv \|\Lambda^k \psi\| = \sup_n \left\| \frac{T_n^{(k)} \psi_{n+2k}}{\int T_n^{(k)} h_{n+2k}} - \frac{\int T_n^{(k)} \psi_{n+2k}}{\int T_n^{(k)} h_{n+2k}} \cdot h_n \right\|_\infty \\ &\leq \sup_n \left\| \frac{T_n^{(k)} \psi_{n+2k}}{\int T_n^{(k)} \psi_{n+2k}} - h_n \right\|_\infty \cdot \sup_n \left\| \frac{\int T_n^{(k)} \psi_{n+2k}}{\int T_n^{(k)} h_{n+2k}} \right\|_\infty \leq 3C\epsilon_k \|h\| \cdot \|\varphi\|, \end{aligned}$$

whence  $\rho(\Lambda) \leq \lim \sqrt[k]{\epsilon_k} = \theta < 1$ .

COMPLETION OF THE PROOF OF THE LEMMA. We constructed a real-analytic function  $\Upsilon : X \times Y \rightarrow Y$  such that  $\Upsilon(T, h) = 0$  and  $(D_2 \Upsilon)(T, h) : Y \rightarrow Y$  has a bounded inverse. By the implicit function theorem for real-analytic functions on Banach spaces [151],  $T$  has a neighborhood  $W \subset X$  where one can define a real-analytic function  $h : W \rightarrow Y$  so that  $\Upsilon(S, h(S)) = 0$ .

Recall that  $T = T^\xi := \{T_n^\xi\}_{n \in \mathbb{N}}$  and  $h = \{h_n(\cdot, \xi)\}_{n \geq 1}$ . By the uniqueness part of Lemma 6.7,  $h(T) = h(\cdot, \xi)$ . It is easy to see using  $\text{ess sup } |f| < \infty$  that  $\xi \mapsto T^\xi$  is real-analytic (even holomorphic). So  $\xi \mapsto h(T^\xi)$  is real-analytic, whence continuously differentiable infinitely many times. Thus  $\xi \mapsto h_n(\cdot, \xi)$  is real-analytic for all  $n$ , and  $\left\{ \frac{\partial^k}{\partial \xi^k} h_n(\cdot, \xi) \right\}_{n \geq 1} = \frac{\partial}{\partial \xi^k} h(T^\xi) \in Y$  for all  $k$ . By the definition of  $Y$ ,  $\sup_{|\xi| \leq 1} \sup_{n \geq 1} \left\| \frac{\partial}{\partial \xi} h_n(\cdot, \xi) \right\|_\infty = \left\| \frac{\partial}{\partial \xi} h(T) \right\| < \infty$  and  $\sup_{|\xi| \leq 1} \sup_{n \geq 1} \left\| \frac{\partial^2}{\partial \xi^2} h_n(\cdot, \xi) \right\|_\infty = \left\| \frac{\partial^2}{\partial \xi^2} h(T) \right\| < \infty$ .  $\square$

### 6.3.3 Choosing the parameters

Given  $\xi \in \mathbb{R}$  and  $\{a_n\} \subset \mathbb{R}$  bounded, let  $\{\tilde{X}_n^\xi\}_{n \geq 1}$  denote the Markov chain with the initial distribution and state spaces of  $X$ , but with transition probabilities

$$\tilde{\pi}_{n,n+1}^\xi(x, dy) = e^{\xi f_n(x, y)} \frac{h_{n+1}(y, \xi)}{e^{p_n(\xi)} h_n(x, \xi)} \cdot \pi_{n,n+1}(x, dy),$$

where  $p_n(\xi)$  and  $h_k^\xi(\cdot) = h_k(\cdot, \xi)$  are as in Lemma 6.7. (This chain does not depend on the choice of  $\{a_n\}$ , see the remark after the statement of Lemma 6.7.) Denote the expectation and variance operators of this chain by  $\tilde{\mathbb{E}}^\xi, \tilde{V}^\xi$ .

In this section we show that if  $V_N := \text{Var}(S_N) \rightarrow \infty$  and  $\frac{z_N - \mathbb{E}(S_N)}{V_N}$  is sufficiently small, then it is possible to choose  $\xi_N$  and  $a_n$  bounded s.t.

$$\frac{z_N - \tilde{\mathbb{E}}^{\xi_N}(S_N)}{\sqrt{\tilde{V}^{\xi_N}(S_N)}} \xrightarrow{N \rightarrow \infty} 0 \quad \text{and} \quad \sum_{n=1}^N p'_n(0) = \mathbb{E}(S_N).$$

Indeed, we will find  $\xi_N$  so that  $\tilde{\mathbb{E}}^{\xi_N}(S_N) = z_N + O(1)$ . The construction will show that if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then  $\xi_N \rightarrow 0$ .

Let  $\bar{h}_n^\xi := \bar{h}_n(\cdot, \xi) : \mathfrak{S}_n \rightarrow (0, \infty)$  and  $\bar{p}_n(\xi) \in \mathbb{R}$  be the fundamental solution:  $L_n^\xi \bar{h}_{n+1}^\xi = e^{\bar{p}_n(\xi)} \bar{h}_n^\xi$  and  $\int \bar{h}_n(x, \xi) \mu_n(dx) = 1$ . Then  $\bar{h}_n^\xi = e^{-a_n \xi} h_n(\cdot, \xi)$  and  $\bar{p}_n(\xi) = p_n(\xi) + a_n \xi - a_{n+1} \xi$

so

$$\tilde{\pi}_{n,n+1}^{\xi}(x, dy) = e^{\xi f_n(x, y)} \frac{\bar{h}_{n+1}(x, \xi)}{e^{\bar{p}_n(\xi)} \bar{h}_n(y, \xi)} \pi_{n,n+1}(x, dy).$$

Let  $\bar{P}_N(\xi) := \bar{p}_1(\xi) + \dots + \bar{p}_N(\xi)$ .

**Lemma 6.11**  $\xi \mapsto \bar{P}_N(\xi)$  is real analytic, and for every  $R > 0$  there is a constant  $C(R)$  such that for all  $|\xi| \leq R$  and  $N \in \mathbb{N}$ ,

- (1)  $|\bar{P}'_N(\xi) - \tilde{\mathbb{E}}^{\xi}(S_N)| \leq C(R)$ ;
- (2) Suppose  $V_N \rightarrow \infty$ . Then  $C(R)^{-1} \leq \tilde{V}^{\xi}(S_N)/V_N \leq C(R)$  for all  $N$  and  $|\xi| \leq R$ , and

$$\bar{P}''_N(\xi)/\tilde{V}^{\xi}(S_N) \xrightarrow[N \rightarrow \infty]{} 1 \text{ uniformly in } |\xi| \leq R.$$

*Proof.* We have the identity  $e^{\bar{P}_N(\xi)} = \int (L_1^{\xi} \dots L_N^{\xi} \bar{h}_{N+1}^{\xi})(x) \mu_1(dx)$ . Since  $\xi \mapsto \bar{h}^{\xi}$  and  $\xi \mapsto L_n^{\xi}$  are real-analytic,  $\xi \mapsto \bar{P}_N(\xi)$  is real-analytic.

Given  $x \in \mathfrak{S}_1$  (the state space of  $X_1$ ), define two measures on  $\prod_{i=2}^{N+1} \mathfrak{S}_i$  so that for every  $E_i \in \mathcal{B}(\mathfrak{S}_i)$  ( $1 \leq i \leq N+1$ ),

$$\begin{aligned} \pi_x(E_2 \times \dots \times E_{N+1}) &:= \mathbb{P}(X_2 \in E_2, \dots, X_{N+1} \in E_{N+1} | X_1 = x_1), \\ \tilde{\pi}_x^{\xi}(E_2 \times \dots \times E_{N+1}) &:= \tilde{\mathbb{P}}^{\xi}(\tilde{X}_2^{\xi} \in E_2, \dots, \tilde{X}_{N+1}^{\xi} \in E_{N+1} | \tilde{X}_1^{\xi} = x_1). \end{aligned}$$

Let  $S_N(x, \underline{y}) := f(x, y_1) + \sum_{i=1}^N f_i(y_i, y_{i+1})$ , then

$$\frac{d\tilde{\pi}_x^{\xi}}{d\pi_x}(y_2, \dots, y_{N+1}) = e^{\xi S_N(x, \underline{y})} e^{-\bar{P}_N(\xi)} \left( \frac{\bar{h}_{N+1}(y_{N+1}, \xi)}{\bar{h}_1(x, \xi)} \right).$$

By Lemma 6.10,  $\xi \mapsto \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x}(y_2, \dots, y_{N+1})$  is real-analytic. Differentiating, gives

$$\frac{d}{d\xi} \left[ \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x} \right] = \left[ S_N(x, \underline{y}) - \bar{P}'_N(\xi) \frac{\bar{h}_1(x, \xi)}{\bar{h}_{N+1}(y_{N+1}, \xi)} \frac{d}{d\xi} \left( \frac{\bar{h}_{N+1}(y_{N+1}, \xi)}{\bar{h}_1(x, \xi)} \right) \right] \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x}. \text{ We write this as}$$

$$\frac{d}{d\xi} \left[ \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x} \right] = \left[ S_N(x, \underline{y}) - \bar{P}'_N(\xi) + \varepsilon_N(x, y_{N+1}, \xi) \right] \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x}, \quad (6.3.16)$$

where  $\varepsilon_N(x, y_{N+1}, \xi) := \frac{\bar{h}_1(x, \xi)}{\bar{h}_{N+1}(y_{N+1}, \xi)} \frac{d}{d\xi} \left( \frac{\bar{h}_{N+1}(y_{N+1}, \xi)}{\bar{h}_1(x, \xi)} \right)$ . By Lemmas 6.8 and 6.10,  $\varepsilon_N(x, y_{N+1}, \xi)$  is uniformly bounded in  $N, x, \underline{y}$ , and  $|\xi| \leq R$ .

By the intermediate value theorem and the uniform boundedness of  $\xi \mapsto \frac{d}{d\xi} \left[ \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x} \right]$  on compact subsets of  $\xi \in \mathbb{R}$ ,  $\frac{1}{\delta} \left[ \frac{d\tilde{\pi}_x^{\xi+\delta}}{d\pi_x} - \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x} \right]$  is uniformly bounded for  $0 < \delta < 1$ . So by the bounded convergence theorem

$$\int \lim_{\delta \rightarrow 0} \frac{1}{\delta} \left[ \frac{d\tilde{\pi}_x^{\xi+\delta}}{d\pi_x} - \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x} \right] d\pi_x = \lim_{\delta \rightarrow 0} \int \frac{1}{\delta} \left[ \frac{d\tilde{\pi}_x^{\xi+\delta}}{d\pi_x} - \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x} \right] d\pi_x = 0.$$

So  $\int \frac{d}{d\xi} \left[ \frac{d\tilde{\pi}_x^{\xi}}{d\pi_x} \right] d\pi_x = 0$ , whence by (6.3.16),  $0 = \tilde{\mathbb{E}}_x^{\xi}(S_N) - \bar{P}'_N(\xi) + O(1)$ , where  $\tilde{\mathbb{E}}_x^{\xi} = \tilde{\mathbb{E}}^{\xi}(\cdot | \tilde{X}_1^{\xi} = x)$ . Integrating with respect to  $x$  we obtain that

$$\bar{P}'_N(\xi) = \tilde{\mathbb{E}}^\xi(S_N) + O(1)$$

uniformly in  $|\xi| \leq R, N \rightarrow \infty$ .

Differentiating (6.3.16) again we obtain

$$\begin{aligned} \frac{d^2}{d\xi^2} \left[ \frac{d\tilde{\pi}_x^\xi}{d\pi_x} \right] &= \frac{d}{d\xi} \left[ \frac{d\tilde{\pi}_x^\xi}{d\pi_x} \left( S_N(x, \underline{y}) - \bar{P}'_N(\xi) + \varepsilon_N(x, y_{N+1}, \xi) \right) \right] \\ &= \frac{d\tilde{\pi}_x}{d\pi_x} \left[ \left( S_N(x, \underline{y}) - \bar{P}'_N(\xi) + \varepsilon_N(x, y_{N+1}, \xi) \right)^2 - \bar{P}''_N(\xi) + \frac{d\varepsilon_N}{d\xi} \right]. \end{aligned}$$

By Lemmas 6.8 and 6.10,  $\frac{d\varepsilon_N}{d\xi}$  is uniformly bounded in  $x, y_{N+1}, N$  and  $|\xi| \leq R$ . As before,  $\int \frac{d^2}{d\xi^2} \frac{d\tilde{\pi}_x^\xi}{d\pi_x} d\pi_x = \frac{d^2}{d\xi^2} \int \frac{d\tilde{\pi}_x^\xi}{d\pi_x} d\pi_x = 0$ , whence

$$\begin{aligned} 0 &= \tilde{\mathbb{E}}^\xi \left[ \left( S_N - \bar{P}'_N(\xi) + O(1) \right)^2 \right] - \bar{P}''_N(\xi) + O(1) \\ &= \tilde{\mathbb{E}}^\xi \left[ \left( S_N - \tilde{\mathbb{E}}^\xi(S_N) + O(1) \right)^2 \right] - \bar{P}''_N(\xi) + O(1), \\ &= \tilde{V}^\xi(S_N) - \bar{P}''_N(\xi) + O\left(\sqrt{\tilde{V}^\xi(S_N)}\right) \end{aligned} \tag{6.3.17}$$

where the  $O(1)$  terms are uniformly bounded in  $N$  when  $|\xi| \leq R$ .

If  $|\xi| \leq R$ , then  $\tilde{\pi}_{n,n+1}^\xi(x, dy)$  are uniformly elliptic with  $\varepsilon_0$  replaced by  $\varepsilon_0/(C^2 e^{KR})$  for the  $C$  in Lemma 6.8. Therefore by Theorem 2.7,  $\tilde{V}^\xi(S_N) \asymp \sum_{n=3}^N u_n^2(\xi)$  where  $u_n(\xi)$  are the structure constants of  $\{\tilde{X}_n^\xi\}$ . Clearly,  $u_n(\xi) \asymp u_n$  where  $u_n = u_n(0)$  are the structure constants of  $\{X_n\}$ . So  $\tilde{V}^\xi(S_N) \asymp V_N \rightarrow \infty$  where the multiplicative error bounds is uniform in  $N$  and  $|\xi| \leq R$ . By (6.3.17),  $\bar{P}''_N(\xi)/\tilde{V}^\xi(S_N) \xrightarrow[N \rightarrow \infty]{} 1$ .  $\square$

**The choice of  $a_N$ :** Lemma 6.11(1) with  $\xi = 0$  says that  $\bar{P}'_N(0) = \mathbb{E}(S_N) + O(1)$ . The error term is a nuisance, and we will choose  $a_n$  to get rid of it. Given  $N$ , let

$$a_n := \mathbb{E}(S_{n-1}) - \bar{P}'_{n-1}(0), \quad a_1 := 0 \tag{6.3.18}$$

This is a bounded sequence, because of Lemma 6.11(1). The choice of  $\{a_n\}$  leads to the following objects:

$$\begin{aligned} h_n^\xi(x) &= h_n(x, \xi) := \exp(a_n \xi) \bar{h}_n(x, \xi), \\ p_n(\xi) &:= \bar{p}_n(\xi) + (a_{n+1} - a_n) \xi. \end{aligned} \tag{6.3.19}$$

The transition kernel  $\tilde{\pi}_{n,n+1}^\xi$  is left unchanged, because the differences between  $\bar{h}_n$  and  $h_n$  and between  $\bar{p}_n$  and  $p_n$  cancel out. But now,

$$P_N(\xi) := p_1(\xi) + \cdots + p_N(\xi) \equiv \bar{P}_N(\xi) + (\mathbb{E}(S_N) - \bar{P}'_N(0))\xi, \tag{6.3.20}$$

satisfies  $P'_N(0) = \mathbb{E}(S_N)$ .

**Properties of  $P_N(\xi)$ :** These functions turn out to be closely related to the distributional properties of  $X$  and its change of measure  $X^\xi$ .

Recall that  $\mathcal{F}_N(\xi) := \frac{1}{V_N} \log \mathbb{E}(e^{\xi S_N})$ , and that  $\tilde{V}^\xi$  is the variance of  $S_N$  with respect to the change of measure  $\tilde{X}^\xi$ . Then:

**Lemma 6.12** Suppose  $V_N \rightarrow \infty$  then  $\xi \mapsto P_N(\xi)$  is real analytic, and

- (1)  $P'_N(0) = \mathbb{E}(S_N)$
- (2) For every  $R > 0$ , there exists  $C(R) > 0$  s.t.

$$|P'_N(\xi) - \tilde{\mathbb{E}}^\xi(S_N)| \leq C(R) \text{ for all } |\xi| \leq R, N \in \mathbb{N}.$$

- (3) For every  $R > 0$ , there exists  $C(R) > 0$  s.t.

$$C(R)^{-1} \leq \tilde{V}^\xi(S_N)/V_N \leq C(R) \text{ for all } |\xi| \leq R, N \in \mathbb{N}.$$

- (4)  $P''_N(\xi)/\tilde{V}^\xi(S_N) \xrightarrow[N \rightarrow \infty]{} 1$  uniformly on compact subsets of  $\xi$ .

(5)  $P_N(\xi)/V_N = \mathcal{F}_N(\xi) + o(V_N^{-1})$  uniformly on compact subsets of  $\xi$ , as  $N \rightarrow \infty$ . Specifically, let  $\Delta_N(R) := \sup_{|\xi| \leq R} V_N \left| \mathcal{F}_N(\xi) - \frac{P_N(\xi)}{V_N} \right|$ . Then  $\sup_N \Delta_N(R) < \infty$  for all  $R > 0$ , and  $\sup_N \Delta_N(R) \xrightarrow[R \rightarrow 0^+]{} 0$ .

(6)  $P'_N(\xi)/V_N = \mathcal{F}'_N(\xi) + O(V_N^{-1})$  uniformly on compact subsets of  $\xi$ , as  $N \rightarrow \infty$ . Specifically, let  $\bar{\Delta}_N(R) := \sup_{|\xi| \leq R} V_N \left| \mathcal{F}'_N(\xi) - \frac{P'_N(\xi)}{V_N} \right|$ . Then  $\sup_{N \geq N_0} \bar{\Delta}_N(R) < \infty$ .

*Proof.* The real analyticity of  $P_N(\xi)$  and parts (1)–(4) follow directly from Lemma 6.11, the identity  $P_N(\xi) = \bar{P}_N(\xi) + (a_{N+1} - a_1)\xi$ , and the boundedness of  $a_n$ .

The proof of part (5) uses the operators  $L_n^\xi : L^\infty(\mathfrak{S}_{n+1}) \rightarrow L^\infty(\mathfrak{S}_n)$  from (6.3.5),  $(L_n^\xi h)(x) := \int_{\mathfrak{S}_{n+1}} e^{\xi f_n(x,y)} h(y) \pi_{n,n+1}(x,dy) \equiv \mathbb{E}_x[e^{\xi f_n(x, X_{n+1})} h(X_{n+1})]$ .

Let  $h_n^\xi := h_n(\cdot, \xi) \in L^\infty(\mathfrak{S}_n)$  be the unique positive functions constructed so that  $L_n^\xi h_{n+1}^\xi = e^{p_n(\xi)} h_n^\xi$ , where  $p_1(\xi) + \dots + p_N(\xi) = P_N(\xi)$ . (To construct  $h_n^\xi$ , apply Lemma 6.7 with  $a_n$  as in (6.3.19).) In particular,  $h_n^0 \equiv 1$  and

$$\mathbb{E}_x \left( e^{\xi S_N} h_{N+1}(X_{N+1}) \right) = e^{P_N(\xi)} h_1^\xi(x). \quad (6.3.21)$$

By Lemma 6.8, there exists  $C_1 = C_1(R) > 1$  such that  $C_1^{-1} \leq h_{N+1}^\xi \leq C_1$  for all  $|\xi| \leq R$  and  $N \geq 1$ . Thus by (6.3.21),

$$C_1(R)^{-2} e^{P_N(\xi)} \leq \mathbb{E} \left( e^{\xi S_N} \right) \leq C_1(R)^2 e^{P_N(\xi)}.$$

Taking logarithms, we deduce that  $|\mathcal{F}_N(\xi) - P_N(\xi)/V_N| \leq 2C_1(R)/V_N$  for all  $N \geq 1$  and  $|\xi| \leq R$ . Equivalently,  $\sup_N \Delta_N(R) \leq 2C_1$ .

Next, by Lemma 6.10 and the identity  $h_n^0 \equiv 1$ ,  $\|h_N^\xi - 1\|_\infty \xrightarrow[N \rightarrow \infty]{} 0$  uniformly on compact subsets of  $\xi$ . Returning to the definition of  $C_1(R)$  we find that we may choose  $C_1(R) \xrightarrow[R \rightarrow 0^+]{} 1$ . As before, this implies that  $\sup_N \Delta_N(R) \xrightarrow[R \rightarrow 0]{} 0$ .

Here is the proof of part (6). Fix  $R > 0$  and let  $\tilde{\mathbb{E}}^\xi$  denote the expectation operator with respect to the change of measure  $\tilde{X}^\xi$ , then

$$V_N \mathcal{F}'_N(\xi) = \frac{\mathbb{E}(S_N e^{\xi S_N})}{\mathbb{E}(e^{\xi S_N})} = \frac{\tilde{\mathbb{E}}^\xi(S_N(h_1^\xi/h_{N+1}^\xi))}{\tilde{\mathbb{E}}^\xi(h_1^\xi/h_{N+1}^\xi)}. \quad (6.3.22)$$

We have already remarked that  $X^\xi$  are uniformly elliptic, and that their uniform ellipticity constants are bounded away from zero for  $\xi$  ranging on a compact set. This gives us the mixing bounds in Proposition 1.11 with the same  $C_{mix} > 0$ ,  $0 < \theta < 1$  for all  $|\xi| \leq R$ . So

$$\tilde{\mathbb{E}}^\xi \left( \frac{h_1^\xi S_N}{h_{N+1}^\xi} \right) = \tilde{\mathbb{E}}^\xi(h_1^\xi) \tilde{\mathbb{E}}^\xi \left( 1/h_{N+1}^\xi \right) \tilde{\mathbb{E}}^\xi(S_N) + O(1) \text{ as } N \rightarrow \infty,$$

$$\tilde{\mathbb{E}}^\xi \left( \frac{h_1^\xi}{h_{N+1}^\xi} \right) = \tilde{\mathbb{E}}^\xi(h_1^\xi) \tilde{\mathbb{E}}^\xi \left( 1/h_{N+1}^\xi \right) + O(\theta^N), \text{ as } N \rightarrow \infty$$

where the big oh's are uniform for  $|\xi| \leq R$ . Plugging this into (6.3.22) gives

$$V_N \mathcal{F}'_N(\xi) = \tilde{\mathbb{E}}^\xi(S_N) + O(1) \text{ as } N \rightarrow \infty, \text{ uniformly for } |\xi| \leq R.$$

Part (6) follows from this from part (2) of the lemma.  $\square$

**The choice of  $\xi_N$ :** We choose  $\xi_N$  so that  $P'_N(\xi_N) = z_N$ ,  $\tilde{\mathbb{E}}^{\xi_N}(S_N) = z_N + O(1)$ . The following lemma gives sufficient conditions for the existence of such  $\xi_N$ .

**Lemma 6.13** *Suppose  $V_N \rightarrow \infty$ ,  $R > 0$ , and*

$$[\hat{a}_N^R, \hat{b}_N^R] := \left[ \mathcal{F}'_N(-R) - \frac{\mathbb{E}(S_N)}{V_N}, \mathcal{F}'_N(R) - \frac{\mathbb{E}(S_N)}{V_N} \right].$$

(1) *For each  $R$  there is  $C(R)$ ,  $N(R)$  s.t. if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$ , and  $N > N(R)$  then*

- (a)  $\exists! \xi_N \in [-(R+1), (R+1)]$  s.t.  $P'_N(\xi_N) = z_N$ ;
- (b)  $C(R)^{-1} \left| \frac{z_N - \mathbb{E}(S_N)}{V_N} \right| \leq |\xi_N| \leq C(R) \left| \frac{z_N - \mathbb{E}(S_N)}{V_N} \right|$ ;
- (c)  $\text{sgn}(\xi_N) = \text{sgn}(\frac{z_N - \mathbb{E}(S_N)}{V_N})$ ;
- (d)  $\left| \tilde{\mathbb{E}}^{\xi_N}(S_N) - z_N \right| \leq C(R)$ .

(2) *For every  $R > 1$  there exists  $c(R) > 0$  such that for all  $N$  large enough,*

$$\text{if } \left| \frac{z_N - \mathbb{E}(S_N)}{V_N} \right| \leq c(R), \text{ then } \frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]. \quad (6.3.23)$$

Consequently, if  $\left| \frac{z_N - \mathbb{E}(S_N)}{V_N} \right| < c(R)$ , then there exists a unique  $\xi_N$  with (a)–(d) above.

*Proof.* Let  $[\tilde{a}_N^R, \tilde{b}_N^R] := \left[ \frac{P'_N(-R) - \mathbb{E}(S_N)}{V_N}, \frac{P'_N(R) - \mathbb{E}(S_N)}{V_N} \right]$ .

CLAIM: *For all  $R > 0$ , for all  $N$  large enough,*

$$[\hat{a}_N^R, \hat{b}_N^R] \subset [\tilde{a}_N^{R+1}, \tilde{b}_N^{R+1}] \subset [\hat{a}_N^{R+2}, \hat{b}_N^{R+2}].$$

*Proof of the claim:* By parts (3) and (4) of Lemma 6.12, there exists  $\delta > 0$  such that  $P''_N(\xi)/V_N \geq \delta$  on  $[-(R+2), (R+2)]$ . Thus by the mean value theorem,

$$\tilde{b}_N^{R+2} \geq \tilde{b}_N^{R+1} + \delta, \quad \tilde{b}_N^{R+1} \geq \tilde{b}_N^R + \delta, \quad \tilde{a}_N^{R+2} \leq \tilde{a}_N^{R+1} - \delta, \quad \tilde{a}_N^{R+1} \leq \tilde{a}_N^R - \delta.$$

Next by part (6) of Lemma 6.12,  $|\hat{b}_N^{R'} - \tilde{b}_N^{R'}| = O(V_N^{-1})$  and  $|\hat{a}_N^{R'} - \tilde{a}_N^{R'}| = O(V_N^{-1})$  for all  $R' \leq R+2$ . For all  $N$  large enough  $|O(V_N^{-1})| < \delta$ , and

$$\hat{a}_N^{R+2} < \tilde{a}_N^{R+1} < \hat{a}_N^R < \hat{b}_N^R < \tilde{b}_N^{R+1} < \hat{b}_N^{R+2},$$

which proves the claim.

We can now prove part (1) of the lemma. Let  $\varphi_N(\xi) := \frac{P_N(\xi) - \xi P'_N(0)}{V_N}$ . By Lemma 6.12,  $\varphi_N(\xi)$  is strictly convex, smooth, and

$$P'_N(\xi_N) = z_N \quad \text{iff} \quad \varphi'_N(\xi_N) = \frac{z_N - P'_N(0)}{V_N}.$$

Fix  $R > 0$ . By the claim, for all  $N$  large enough, if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\tilde{a}_N^R, \hat{b}_N^R]$ , then  $\frac{z_N - P'_N(0)}{V_N} \equiv \frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\tilde{a}_N^{R+1}, \hat{b}_N^{R+1}] \equiv \varphi'_N[-(R+1), (R+1)]$ . Since  $\varphi'_N$  is continuous and strictly increasing, there  $\exists! \xi_N \in [-(R+1), (R+1)]$  such that  $\varphi'_N(\xi_N) = \frac{z_N - P'_N(0)}{V_N}$ . Equivalently, there exists a unique  $|\xi_N| \leq R+1$  such that  $P'_N(\xi_N) = z_N$ .

This argument shows that for every  $N$  sufficiently large, for every  $\eta \in [\tilde{a}_N^R, \hat{b}_N^R]$  there exists a unique  $\xi = \xi(\eta) \in [-(R+1), (R+1)]$  such that

$$\varphi'_N(\xi(\eta)) = \eta.$$

By Lemma 6.12,  $\exists \delta(R) > 0$  so that  $\delta(R) \leq \varphi''_N \leq \delta(R)^{-1}$  on  $[-(R+1), (R+1)]$ . So  $\eta \mapsto \xi(\eta)$  is  $\frac{1}{\delta(R)}$ -bi-Lipschitz on  $[\tilde{a}_N^R, \hat{b}_N^R]$ . By construction,  $\varphi'_N(0) = 0$ . So  $\xi(0) = 0$ , whence by the bi-Lipschitz property

$$\delta(R)|\eta| \leq |\xi(\eta)| \leq \delta(R)^{-1}|\eta| \text{ on } [\tilde{a}_N^R, \hat{b}_N^R].$$

Since  $\varphi_N$  is real-analytic and strictly convex,  $\varphi'_N$  is smooth and strictly increasing. By the inverse mapping theorem,  $\eta \mapsto \xi(\eta)$  is smooth and strictly increasing. So

$$\text{sgn}(\xi(\eta)) = \text{sgn}(\eta) \text{ on } [\tilde{a}_N^R, \hat{b}_N^R].$$

Specializing to the case  $\eta = \frac{z_N - \mathbb{E}(S_N)}{V_N}$ , gives properties (a)–(c) of  $\xi_N$ .

Property (d) is because of by Lemma 6.12, which says that

$$z_N = P'_N(\xi_N) = \tilde{E}^{\xi_N}(S_N) + O(1).$$

Notice that the big oh is uniform because  $|\xi_N| \leq R+1$ . This completes the proof of part (1).

Here is the proof of part (2): For every  $R > 1$ , for all  $N$  large enough

$$\begin{aligned} [\tilde{a}_N^R, \hat{b}_N^R] &\supset [\tilde{a}_N^{R-1}, \hat{b}_N^{R-1}] \equiv \varphi'_N[-(R-1), (R-1)] \quad (\because \text{claim, } \varphi'_N \text{ is increasing}) \\ &\supset [-\delta(R-1)(R-1), \delta(R-1)(R-1)] \quad (\because \varphi'_N(0) = 0, \varphi''_N \geq \delta(R+1)). \end{aligned}$$

So  $[\tilde{a}_N^R, \hat{b}_N^R] \supset [-c, c]$  for  $R \geq 2$  where  $c := \delta(1)$ . □

**Corollary 6.14** Suppose  $V_N \rightarrow \infty$  and  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then for all  $N$  large enough, there exists a unique  $\xi_N$  such that  $P'_N(\xi_N) = z_N$ . Furthermore,  $\xi_N \rightarrow 0$ .

### 6.3.4 The asymptotic behavior of $\tilde{V}_N^{\xi_N}(S_N)$

Let  $\tilde{V}_N^{\xi}$  denote the variance of  $S_N$  with respect to the change of measure  $X^{\xi}$ . We compare  $\tilde{V}_N^{\xi}$  to  $V_N$ .

**Lemma 6.15** Suppose  $V_N \xrightarrow[N \rightarrow \infty]{} \infty$ , and define  $\xi_N$  as in Lemma 6.13.

- (1) Suppose  $R > 0$  and  $\frac{\xi_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N^R, \hat{b}_N^R]$  for all  $N$ , then  $\tilde{V}_N^{\xi_N} \asymp V_N$  as  $N \rightarrow \infty$ .
- (2) If  $\frac{\xi_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then  $\tilde{V}_N^{\xi_N} \sim V_N$  as  $N \rightarrow \infty$ .
- (3)  $\tilde{V}_N^{\xi} \sim V_N$  as  $N \rightarrow \infty$  uniformly on compact subsets of  $\xi$ : For every  $\varepsilon > 0$  there are  $\xi^* > 0$  and  $N_0 > 1$ , so that  $\tilde{V}_N^{\xi}/V_N \in [e^{-\varepsilon}, e^{\varepsilon}]$  for all  $|\xi| < \xi^*, N > N_0$ .

*Proof.* Part (1) is because of Lemma 6.12(3) and the bound  $|\xi_N| \leq R + 1$  from Lemma 6.12. Part (2) follows from part (3) and Corollary 6.14. It remains to prove part (3).

To do this we decompose  $S_N$  into weakly correlated large blocks of roughly the same  $X$ -variance, and check that the  $X^{\xi}$ -variance of the  $i$ -th block converges *uniformly* in  $i$  to its  $X$ -variance.

Let  $\{X_n\}$  and  $\{\tilde{X}_n^{\xi}\}$  denote the Markov chains with transition kernels  $\{\pi_{n,n+1}(x, dy)\}$ ,  $\{\tilde{\pi}_{n,n+1}^{\xi}(x, dy)\}$  and initial distribution  $\mu_1(dx)$ . Given natural numbers  $n > m$ , let

$$\begin{aligned} S_{n,m} &:= X_n + \cdots + X_{m-1} \\ \tilde{S}_{n,m}^{\xi} &:= \tilde{X}_n^{\xi} + \cdots + \tilde{X}_{m-1}^{\xi} \\ p_{n,m}(\xi) &:= p_n(\xi) + \cdots + p_{m-1}(\xi). \end{aligned}$$

Notice that for all  $R > 0$ ,  $n < m$ , and  $|\xi| \leq R$ ,

$$p_{n,m}(0) = 0, \quad p'_{n,m}(0) = \mathbb{E}(S_{n,m}), \quad |p'_{n,m}(\xi) - \tilde{\mathbb{E}}^{\xi}(\tilde{S}_{n,m}^{\xi})| \leq C(R). \quad (6.3.24)$$

The first identity is because  $\bar{h}_n(\cdot, 0) \equiv 1$ ,  $\bar{p}_n(0) = 1$  by the uniqueness of the fundamental solution. The second identity is because

$$p'_{n,m}(0) = P'_{m-1}(0) - P'_{n-1}(0) = \mathbb{E}(S_{m-1}) - \mathbb{E}(S_{n-1})$$

by choice of  $\{a_n\}$ . The inequality can be proved by applying Lemma 6.12 to the shifted Markov chain  $\{X_k\}_{k \geq n}$ .

Let  $V(S_{n,m}) := \text{Var}(S_{n,m})$ . The application of Lemma 6.12 to the shifted Markov chain  $\{X_k\}_{k \geq n}$  also gives a constant  $M_0$  s.t. for all  $|\xi| \leq R$ ,

$$V(S_{n,m}) \geq M_0 \Rightarrow \begin{cases} C(R)^{-1} \leq \tilde{V}^{\xi}(\tilde{S}_{n,m}^{\xi})/V(S_{n,m}) \leq C(R) \\ 2^{-1} \leq p''_{n,m}(\xi)/\tilde{V}^{\xi}(\tilde{S}_{n,m}^{\xi}) \leq 2. \end{cases} \quad (6.3.25)$$

$M_0$  is independent of  $n$ : It is a function of  $R, K, \varepsilon_0$ , and the uniform bounds on  $h_n(\cdot, \xi)$  and its derivatives.

**STEP 1 (UNIFORM EXPONENTIAL MIXING).** There are  $C_{mix}^* = C_{mix}(R) > 0$ ,  $\eta = \eta(R) \in (0, 1)$  such that for every  $|\xi| \leq R$ , for all  $n < m$ ,

$$|\text{Cov}(f_m(\tilde{X}_m^{\xi}, \tilde{X}_{m+1}^{\xi}), f_n(\tilde{X}_n^{\xi}, \tilde{X}_{n+1}^{\xi}))| \leq C_{mix}^* \eta^{m-n}.$$

Proof: If  $|\xi| \leq R$ , then the Markov chain  $\tilde{X}_n^\xi$  is uniformly elliptic with ellipticity constant  $\varepsilon_0(R) > 0$ . The step follows from Proposition 1.11.

STEP 2 (BLOCK DECOMPOSITION). *For every  $\varepsilon > 0$  small enough, for every  $R > 1$ , there exists  $M > 1$  and integers  $n_i \uparrow \infty$  such that:*

- (1)  $M \leq V(S_{n_i, n_{i+1}}) \leq 2M$ ;
- (2)  $|\text{Cov}(\tilde{S}_{n_i, n_{i+1}}^\xi, \tilde{S}_{n_j, n_{j+1}}^\xi)| \leq C_{\text{mix}}^{\#} \eta^{n_j - n_{i+1}}$  for all  $|\xi| \leq R$  and  $i < j$ , where the constant  $C_{\text{mix}}^{\#}$  is independent of  $M, i, j$ ;
- (3) For all  $|\xi| \leq R$ , for all  $i > 3$ , for all  $n \in [n_i, n_{i+1}]$ ,

$$e^{-\varepsilon} \leq \frac{\tilde{V}^\xi(\tilde{S}_{1,n}^\xi)}{\sum_{k=1}^{i-1} \tilde{V}^\xi(\tilde{S}_{n_k, n_{k+1}}^\xi) + \tilde{V}_{n_i, n}^\xi} \leq e^\varepsilon. \quad (6.3.26)$$

$$(4) M^* := \sup_i \sup_{n \in [n_i, n_{i+1}]} \sup_{|\xi| \leq R} |p''_{n_i, n}(\xi)| < \infty$$

*Proof.* We write  $V_{n,m} := V(S_{n,m})$  and  $\tilde{V}_{n,m}^\xi := \tilde{V}^\xi(\tilde{S}_{n,m}^\xi)$ , and fix

$$M > \max \left\{ 2 \left( K^2 + \frac{C_{\text{mix}}^*}{1 - \eta} \right), \frac{4C_{\text{mix}}^* C(R)}{\varepsilon^{-1} (1 - \eta)^3} \right\}.$$

Construct  $n_i = n_i(M) \in \mathbb{N}$  by induction as follows:  $n_1 := 1$ , and

$$n_{i+1} := \min\{n > n_i : V_{n_i, n_{i+1}} > M\}.$$

There does indeed exist  $n > n_i$  with  $V_{n_i, n_{i+1}} > M$ , because  $V_{n_i, n} \xrightarrow{n \rightarrow \infty} \infty$ , as can be seen from the following calculation:

$$\begin{aligned} \infty &\xleftarrow[\infty \leftarrow n]{} V_{1,n} = V_{1,n_i} + V_{n_i,n} + 2\text{Cov}(S_{1,n_i}, S_{n_i,n}) \\ &= V_{n_i,n} + V_{1,n_i} + O\left(\sum_{m=1}^{n_i-1} \sum_{k=0}^{\infty} |\text{Cov}(X_m, X_{n_i+k})|\right) \\ &= V_{n_i,n} + O(1), \text{ by step 1 with } \xi = 0. \end{aligned}$$

By construction,  $V_{n_i, n_{i+1}} > M$ , and

$$\begin{aligned} V_{n_i, n_{i+1}} &\leq V_{n_i, n_{i+1}-1} + |V_{n_i, n_{i+1}} - V_{n_i, n_{i+1}-1}| \\ &\leq M + |V_{n_i, n_{i+1}} - V_{n_i, n_{i+1}-1}| \quad \text{by the minimality of } n_{i+1} \\ &\leq M + V(f_{n_{i+1}-2}(X_{n_{i+1}-2}, X_{n_{i+1}-1})) \\ &\quad + 2|\text{Cov}(f_{n_{i+1}-2}(X_{n_{i+1}-2}, X_{n_{i+1}-1}), S_{n_{i+1}-1})| \\ &\leq M + 2 \left( K^2 + \frac{C_{\text{mix}}^*}{1 - \eta} \right) \leq 2M \quad \text{by the choice of } M. \end{aligned}$$

So  $M < V_{n_i, n_{i+1}} \leq 2M$ , and  $\{n_i\}$  satisfies part (1).

If  $i < j$ , then  $|\text{Cov}(\tilde{S}_{n_i, n_{i+1}}^\xi, \tilde{S}_{n_j, n_{j+1}}^\xi)| \leq \sum_{k=n_i}^{n_{i+1}-1} \sum_{\ell=n_j}^{n_{j+1}-1} C_{\text{mix}}^* \eta^{\ell-k}$

$$\leq C_{mix}^* \sum_{k=n_i}^{n_{i+1}-1} \frac{\eta^{n_j-k}}{1-\eta} = \frac{C_{mix}^* \eta^{n_j-n_{i+1}}}{1-\eta} \sum_{k=n_i}^{n_{i+1}-1} \eta^{n_{i+1}-k} = \frac{C_{mix}^* \eta^{n_j-n_{i+1}}}{(1-\eta)^2}.$$

Part (2) follows with  $C_{mix}^\# := C_{mix}^* / (1-\eta)^2$ .

Part (3) follows from parts (1),(2). Namely, fix  $n \in [n_i, n_{i+1}]$ , then

$$\begin{aligned} & \left| \tilde{V}^\xi(\tilde{S}_{1,n}^\xi) - \sum_{k=1}^{i-1} \tilde{V}_{n_k, n_{k+1}}^\xi - \tilde{V}_{n_i, n}^\xi \right| \\ & \leq 2 \sum_{1 \leq k < \ell \leq i-1} |\text{Cov}(\tilde{S}_{n_k, n_{k+1}}^\xi, \tilde{S}_{n_\ell, n_{\ell+1}}^\xi)| + 2 \sum_{1 \leq k \leq i-1} |\text{Cov}(\tilde{S}_{n_k, n_{k+1}}^\xi, \tilde{S}_{n_i, n}^\xi)| \\ & \leq 2 \sum_{1 \leq k < \ell \leq i-1} C_{mix}^\# \eta^{n_\ell - n_{k+1}} + 2 \sum_{1 \leq k \leq i-1} C_{mix}^\# \eta^{n_i - n_{k+1}} \leq 2 \sum_{1 \leq k < \ell \leq i} C_{mix}^\# \eta^{\ell - k - 1} \\ & = 2C_{mix}^\# \sum_{k=1}^{i-1} \sum_{\ell=k+1}^{i-1} \eta^{\ell - k - 1} \leq \frac{2C_{mix}^\# i}{1-\eta} = \frac{2C_{mix}^* i}{(1-\eta)^3}. \end{aligned}$$

By (6.3.25),  $\sum_{k=1}^{i-1} \tilde{V}_{n_k, n_{k+1}}^\xi \geq \frac{M(i-1)}{C(R)}$ . So

$$\left| \frac{\tilde{V}_{n_i, n_{i-1}}^\xi}{\sum_{k=1}^{i-1} \tilde{V}_{n_k, n_{k+1}}^\xi + \tilde{V}_{n_i, n}^\xi} - 1 \right| \leq \frac{\left( \frac{2C_{mix}^*}{(1-\eta)^3} \right) i}{C(R)^{-1} M(i-1)} \leq \frac{1}{M} \cdot \frac{2C_{mix}^* C(R)}{(1-\eta)^3} \cdot \frac{i}{i-1} \leq \frac{\varepsilon}{2} \cdot \frac{i}{i-1},$$

where the last inequality is by the choice of  $M$ . If  $i > 3$ , the last bound is less than  $\frac{3}{4}\varepsilon$ , and (6.3.26) follows for all  $\varepsilon$  sufficiently small.

Part (4) is a uniform bound on  $|p''_{n_i, n}(\xi)|$  for  $i \in \mathbb{N}$ ,  $n \in [n_i, n]$ ,  $|\xi| \leq R$ . By construction,  $V_{n_i, n} \leq 2M$ . By Theorem 2.7, this implies a uniform upper bound on  $\sum_{k=n_i}^{n-1} u_k^2$ . The structure constants of  $\{X_n\}$  and  $\{\tilde{X}_n^\xi\}$  are equal up to a bounded multiplicative error. So the same theorem, applied to the Markov chain  $\{\tilde{X}_k^\xi\}_{k \geq n_i}$ , gives a uniform upper bound for  $\tilde{V}_{n_i, n}^\xi$ , whence  $\sup_i \sup_{n \in [n_i, n_{i+1}]} \sup_{|\xi| \leq R} \tilde{V}_{n_i, n}^\xi < \infty$ .

A routine modification of the argument we used to show (6.3.17) shows that

$$\left| p''_{n, m}(\xi) - \tilde{\mathbb{E}}^\xi \left[ \left( \tilde{S}_{n_i, n}^\xi - \tilde{\mathbb{E}}^\xi(\tilde{S}_{n_i, n}^\xi) + O(1) \right)^2 \right] \right| \leq \text{const.}$$

The expectation term is uniformly bounded because of the bound on  $\tilde{V}_{n_i, n}^\xi$  and the Minkowski inequality, so part (4) follows.

STEP 3 (BLOCK EXPECTATION). *For every  $\varepsilon > 0$  there exists  $\xi^* > 0$  such that for all  $|\xi| \leq \xi^*$ ,*

$$\left| \tilde{\mathbb{E}}^\xi(\tilde{S}_{n_i, n}^\xi) - \mathbb{E}(S_{n_i, n}) \right| \leq \varepsilon \text{ for all } i \in \mathbb{N} \text{ and } n_i \leq n \leq n_{i+1}.$$

*Proof.* By Lemma 6.8  $h_k(\cdot, \xi)$  is uniformly bounded away from zero and infinity when  $|\xi| \leq R$ . By Lemma 6.10,  $\xi \mapsto h_k(\cdot, \xi)$  is uniformly Lipschitz on  $[-R, R]$ . It follows that

$$\frac{h_{n+1}(y, \xi)}{h_{n_i}(x, \xi)} \xrightarrow[\xi \rightarrow 0]{} 1 \text{ uniformly for } i \in \mathbb{N}, n \in [n_i, n_{i+1}], (x, y) \in \mathfrak{S}_{n_i} \times \mathfrak{S}_{n+1}.$$

In particular, there is a  $\xi_1^*$  s.t. for all  $|\xi| \leq \xi_1^*$

$$2^{-1} \leq \frac{h_{n+1}(y, \xi)}{h_{n_i}(x, \xi)} \leq 2 \text{ for all } i \in \mathbb{N}, n \in [n_i, n_{i+1}], \text{ and } (x, y) \in \mathfrak{S}_{n_i} \times \mathfrak{S}_n.$$

This has a useful consequence. Since

$$\begin{aligned} \mathbb{E} \left( e^{\xi S_{n_i, n}} \frac{h_{n+1}(X_{n+1}, \xi)}{e^{p_{n_i, n}(\xi)} h_{n_i}(X_{n_i}, \xi)} \right) &= \mathbb{E} \left( \mathbb{E}_{X_{n_i}} \left( e^{\xi S_{n_i, n}} \frac{h_{n+1}(X_{n+1}, \xi)}{e^{p_{n_i, n}(\xi)} h_{n_i}(X_{n_i}, \xi)} \right) \right) = \mathbb{E}(1) = 1, \\ 2^{-1} \mathbb{E} \left( e^{\xi S_{n_i, n}} \right) &\leq e^{p_{n_i, n}(\xi)} \leq 2 \mathbb{E} \left( e^{\xi S_{n_i, n}} \right) \text{ whenever } |\xi| \leq \xi_1^*. \end{aligned} \quad (6.3.27)$$

Fix  $L > 0$  and let  $A_L := [|S_{n_i, n} - \mathbb{E}(S_{n_i, n})| \leq L]$ , then:

$$\begin{aligned} \widetilde{\mathbb{E}}_{X_{n_i}}^{\xi} (\widetilde{S}_{n_i, n}^{\xi}) - \mathbb{E}(S_{n_i, n}) &= \mathbb{E} \left( (S_{n_i, n} - \mathbb{E}(S_{n_i, n})) e^{\xi S_{n_i, n}} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{e^{p_{n_i, n}(\xi)} h_{n_i}(X_{n_i}, \xi)} \right) \\ &= \mathbb{E}_{X_{n_i}} \left( (S_{n_i, n} - \mathbb{E}(S_{n_i, n})) e^{\xi S_{n_i, n} - p_{n_i, n}(\xi)} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{h_{n_i}(X_{n_i}, \xi)} \cdot 1_{A_L} \right) \\ &\quad + \mathbb{E}_{X_{n_i}} \left( (S_{n_i, n} - \mathbb{E}(S_{n_i, n})) e^{\xi S_{n_i, n} - p_{n_i, n}(\xi)} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{h_{n_i}(X_{n_i}, \xi)} \cdot 1_{A_L^c} \right). \end{aligned}$$

*Expectation of the first summand:*  $M^* := \sup_i \sup_{n \in [n_i, n_{i+1}]} \sup_{|\xi| \leq R} |p_{n_i, n}''(\xi)| < \infty$ . Therefore by (6.3.24), for all  $|\xi| \leq R$ ,  $n \in [n_i, n_{i+1}]$ ,

$$p_{n_i, n}(\xi) = p_{n_i, n}(0) + \xi p'_{n_i, n}(0) + O(\xi^2) = \xi \mathbb{E}(S_{n_i, n}) + O(\xi^2), \quad (6.3.28)$$

where  $|O(\xi^2)| \leq M^* \xi^2$ .

So on  $A_L$ ,  $|\xi S_{n_i, n} - p_{n_i, n}(\xi)| \leq |\xi| \cdot |S_{n_i, n} - \mathbb{E}(S_{n_i, n})| + M^* \xi^2 \leq L|\xi| + M^* \xi^2$ , uniformly in  $i, n \in [n_i, n]$ . In particular,

$$e^{\xi S_{n_i, n} - p_{n_i, n}(\xi)} \xrightarrow[\xi \rightarrow 0]{} 1 \text{ on } A_L, \text{ uniformly in } i, n \in [n_i, n].$$

Together with the uniform convergence  $\frac{h_{n+1}(X_{n+1}, \xi)}{h_{n_i}(X_{n_i}, \xi)} \xrightarrow[\xi \rightarrow 0]{} 1$ , this implies that the first summand converges to  $\mathbb{E}_{X_{n_i}}[(S_{n_i, n} - \mathbb{E}(S_{n_i, n})) 1_{A_L}]$  uniformly in  $i, X_{n_i}$ , and  $n \in [n_i, n_{i+1}]$ .

The expectation of the limit satisfies

$$\begin{aligned} |\mathbb{E}[(S_{n_i, n} - \mathbb{E}(S_{n_i, n})) 1_{A_L}]| &= |\mathbb{E}[(S_{n_i, n} - \mathbb{E}(S_{n_i, n})) 1_{A_L^c}]| \\ &\leq \mathbb{E}[L^{-1} (S_{n_i, n} - \mathbb{E}(S_{n_i, n}))^2 1_{A_L^c}] \leq \frac{V(S_{n_i, n})}{L} \leq \frac{2M}{L}. \end{aligned}$$

Thus, for every  $\varepsilon > 0$ , for every  $L$  large enough, for all  $|\xi|$  sufficiently small, for all  $i, n \in [n_i, n_{i+1}]$ , the first summand has expectation  $< \varepsilon/2$ .

*Expectation of the second summand:* Fix  $0 < \delta \ll \xi_1^*$ . Assume  $L$  is so large s.t.  $|t| < \delta e^{\delta|t|}$  for all  $|t| > L$ .

Decompose  $A_L^c := A_+^c \uplus A_-^c$ , where  $A_+^c := [S_{n_i, n} - \mathbb{E}(S_{n_i, n}) > L]$  and  $A_-^c := [S_{n_i, n} - \mathbb{E}(S_{n_i, n}) < -L]$ . Then

$$\begin{aligned}
& \mathbb{E}_{X_{n_i}} \left( |S_{n_i, n} - \mathbb{E}(S_{n_i, n})| e^{\xi S_{n_i, n} - p_{n_i, n}(\xi)} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{h_{n_i}(X_{n_i}, \xi)} \cdot 1_{A_+^c} \right) \\
& \leq 2 \mathbb{E}_{X_{n_i}} \left( |S_{n_i, n} - \mathbb{E}(S_{n_i, n})| e^{\xi S_{n_i, n} - p_{n_i, n}(\xi)} \cdot 1_{A_+^c} \right), \text{ provided } |\xi| \leq \xi_1^* \\
& \leq 4 \mathbb{E}_{X_{n_i}} \left( (S_{n_i, n} - \mathbb{E}(S_{n_i, n})) e^{\xi S_{n_i, n}} \cdot 1_{A_+^c} \right) / \mathbb{E}(e^{\xi S_{n_i, n}}), \text{ by (6.3.27)} \\
& = 4 \mathbb{E}_{X_{n_i}} \left( (S_{n_i, n} - \mathbb{E}(S_{n_i, n})) e^{\xi (S_{n_i, n} - \mathbb{E}(S_{n_i, n}))} \cdot 1_{A_+^c} \right) / \mathbb{E}(e^{\xi (S_{n_i, n} - \mathbb{E}(S_{n_i, n}))}) \\
& \leq 4\delta \mathbb{E}_{X_{n_i}} (e^{(\xi + \delta)(S_{n_i, n} - \mathbb{E}(S_{n_i, n}))}) / \mathbb{E}(e^{\xi (S_{n_i, n} - \mathbb{E}(S_{n_i, n}))}) \\
& \leq 16\delta \exp(p_{n_i, n}(\xi + \delta) - p_{n_i, n}(\xi) - \delta \mathbb{E}(S_{n_i, n})), \text{ provided } |\xi + \delta| < \xi_1^*.
\end{aligned}$$

(see (6.3.27)). Expanding  $p_{n_i, n}(\xi + \delta)$  into Taylor series around  $\xi$ , and recalling  $|p''_{n_i, n}(\xi)| \leq M^*$  for  $|\xi| \leq R$ , we find that the term in the exponent is bounded above by

$$\begin{aligned}
& \delta |p'_{n_i, n}(\xi) - \mathbb{E}(S_{n_i, n})| + M^* \delta^2 = \delta |p'_{n_i, n}(\xi) - p'_{n_i, n}(0)| + M^* \delta^2 \\
& \leq M^* (\delta |\xi| + \delta^2) \leq M^* (R\delta + \delta^2),
\end{aligned}$$

which can be made as small as we wish by choosing  $\delta$  properly.

The conclusion is that for all  $L$  large enough, for all  $|\xi|$  sufficiently small, for all  $i, n \in [n_i, n_{i+1}]$ ,

$$\mathbb{E} \left( |S_{n_i, n} - \mathbb{E}(S_{n_i, n})| e^{\xi S_{n_i, n} - p_{n_i, n}(\xi)} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{h_{n_i}(X_{n_i}, \xi)} \cdot 1_{A_+^c} \right) \leq \frac{\varepsilon}{4}.$$

Similarly, one can show that for all  $L$  large enough, for all  $|\xi|$  sufficiently small, for all  $i, n \in [n_i, n_{i+1}]$ ,

$$\mathbb{E} \left( |S_{n_i, n} - \mathbb{E}(S_{n_i, n})| e^{\xi S_{n_i, n} - p_{n_i, n}(\xi)} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{h_{n_i}(X_{n_i}, \xi)} \cdot 1_{A_-^c} \right) < \frac{\varepsilon}{4}.$$

Thus, for every  $\varepsilon > 0$ , for all  $L$  sufficiently large, for all  $|\xi|$  sufficiently small, for all  $i, n \in [n_i, n_{i+1}]$ , the expectation of the second summand is less than  $\varepsilon/2$  in absolute value.

**STEP 4 (BLOCK VARIANCE).** For every  $\varepsilon > 0$  there exists  $\xi^* > 0$  such that for all  $|\xi| \leq \xi^*$ ,  $|\widetilde{V}_{n_i, n}^{\xi} - V_{n_i, n}| \leq \varepsilon$  for all  $i \in \mathbb{N}$  and  $n_i \leq n \leq n_{i+1}$ .

*Proof.* The proof is similar to the proof of step 3. Fix  $L$  to be determined later and let  $A_L := [|S_{n_i, n} - \mathbb{E}(S_{n_i, n})| \leq L]$ , then

$$\begin{aligned}
\widetilde{V}_{n_i, n}^{\xi} (\widetilde{S}_{n_i, n}^{\xi}) &= \mathbb{E} \left( (S_{n_i, n} - \widetilde{\mathbb{E}}(\widetilde{S}_{n_i, n}^{\xi}))^2 e^{\xi S_{n_i, n}} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{e^{p_{n_i, n}(\xi)} h_{n_i}(X_{n_i}, \xi)} 1_{A_L} \right) \\
&+ \mathbb{E} \left( (S_{n_i, n} - \widetilde{\mathbb{E}}(\widetilde{S}_{n_i, n}^{\xi}))^2 e^{\xi S_{n_i, n}} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{e^{p_{n_i, n}(\xi)} h_{n_i}(X_{n_i}, \xi)} 1_{A_L^c} \right).
\end{aligned}$$

The second summand can be analyzed as in step 3, this time with the inequality  $t^2 < \delta e^{\delta|t|}$  for all  $|t|$  large enough. The conclusion is that for every  $\varepsilon > 0$ , for all  $L$  sufficiently large, for all  $|\xi|$  sufficiently small, for all  $i, n \in [n_i, n_{i+1}]$ ,

$$\mathbb{E} \left( (S_{n_i, n} - \widetilde{\mathbb{E}}(\widetilde{S}_{n_i, n}^{\xi}))^2 e^{\xi S_{n_i, n}} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{e^{p_{n_i, n}(\xi)} h_{n_i}(X_{n_i}, \xi)} 1_{A_L^c} \right) < \frac{\varepsilon}{2}. \quad (6.3.29)$$

The first summand converges to  $\mathbb{E}((S_{n_i,n} - \mathbb{E}(S_{n_i,n}))^2 1_{A_L})$  as  $\xi \rightarrow 0$  uniformly in  $i, n \in [n_i, n_{i+1}]$  because

- $e^{\xi S_{n_i,n}} \cdot \frac{h_{n+1}(X_{n+1}, \xi)}{e^{p_{n_i,n}(\xi)} h_{n_i}(X_{n_i}, \xi)} 1_{A_L} \xrightarrow[\xi \rightarrow 0]{} 1_{A_L}$  uniformly in  $i \in \mathbb{N}, n \in [n_i, n_{i+1}]$ , see the proof of step 3; and
- $(S_{n_i,n} - \widetilde{\mathbb{E}}(S_{n_i,n}^\xi))^2 1_{A_L} \xrightarrow[\xi \rightarrow 0]{} (S_{n_i,n} - \mathbb{E}(S_{n_i,n}))^2 1_{A_L}$  uniformly in  $i, n \in [n_i, n_{i+1}]$ , because for some  $t$  between  $\widetilde{\mathbb{E}}(S_{n_i,n}^\xi)$  and  $\mathbb{E}(S_{n_i,n})$ ,

$$\begin{aligned} & |(S_{n_i,n} - \widetilde{\mathbb{E}}(S_{n_i,n}^\xi))^2 - (S_{n_i,n} - \mathbb{E}(S_{n_i,n}))^2| \\ &= 2|S_{n_i,n} - t| |\widetilde{\mathbb{E}}(S_{n_i,n}^\xi) - \mathbb{E}(S_{n_i,n})| \\ &\leq 2(L + |\widetilde{\mathbb{E}}(S_{n_i,n}^\xi) - \mathbb{E}(S_{n_i,n})|) |\widetilde{\mathbb{E}}(S_{n_i,n}^\xi) - \mathbb{E}(S_{n_i,n})| \text{ on } A_L \\ &\xrightarrow[\xi \rightarrow 0]{} 0 \text{ uniformly on } A_L \text{ in } i \in \mathbb{N}, n \in [n_i, n_{i+1}], \text{ by step 3.} \end{aligned}$$

The limit of the first summand  $\mathbb{E}((S_{n_i,n} - \mathbb{E}(S_{n_i,n}))^2 1_{A_L}) \xrightarrow[L \rightarrow \infty]{} V_{n_i,n}$  uniformly in  $i, n \in [n_i, n_{i+1}]$ . Indeed, applying (6.3.29) with  $\xi = 0$

$$|V_{n_i,n} - \mathbb{E}((S_{n_i,n} - \mathbb{E}(S_{n_i,n}))^2 1_{A_L})| = \mathbb{E}((S_{n_i,n} - \mathbb{E}(S_{n_i,n}))^2 1_{A_L^c}) < \frac{\varepsilon}{2}$$

for all  $L$  large enough, for all  $i \in \mathbb{N}, n \in [n_i, n_{i+1}]$ . Step 4 follows.

PROOF OF PART (3) OF THE LEMMA. Fix  $\varepsilon > 0$ , and construct the block decomposition as in step 2.

By step 4 there exists  $\xi^* > 0$  s.t. for all  $|\xi| < \xi^*$ , for all  $k \in \mathbb{N}, n \in [n_k, n_{k+1}]$ ,  $e^{-\varepsilon} V_{n_k,n} \leq \widetilde{V}_{n_k,n}^\xi \leq e^\varepsilon V_{n_k,n}$ . Therefore

$$e^{-\varepsilon} \leq \frac{\sum_{k=1}^{i-1} \widetilde{V}_{n_k,n_{k+1}}^\xi + \widetilde{V}_{n_i,n}^\xi}{\sum_{k=1}^{i-1} V_{n_k,n_{k+1}} + V_{n_i,n}} \leq e^\varepsilon.$$

By part (3) of step 2, for all  $n > n_3$ , for all  $|\xi| < \xi^*$ ,  $e^{-3\varepsilon} \leq \widetilde{V}_n^\xi / V_n \leq e^{3\varepsilon}$ .  $\square$

### 6.3.5 Asymptotics of the log moment generating functions

We need an elementary observation from probability theory. Let  $X, Y$  be two random variables on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Suppose  $X$  has finite non-zero variance, and  $Y$  is positive and bounded. Let  $\text{Var}^Y(X)$  be the variance of  $X$  with respect to the change of measure  $\frac{Y}{\mathbb{E}(Y)} d\mathbb{P}$ , i.e.

$$\text{Var}^Y(X) := \frac{\mathbb{E}(X^2 Y)}{\mathbb{E}(Y)} - \left( \frac{\mathbb{E}(XY)}{\mathbb{E}(Y)} \right)^2.$$

**Lemma 6.16** Suppose  $0 < \text{Var}(X) < \infty$  and  $C^{-1} \leq Y \leq C$  with  $C$  a positive constant, then  $C^{-4} \leq \frac{\text{Var}^Y(X)}{\text{Var}(X)} \leq C^4$ .

*Proof.* For every random variable  $W$ , if  $W_1, W_2$  are two independent copies of  $W$  then  $\text{Var}(W) = \frac{1}{2}\mathbb{E}[(W_1 - W_2)^2]$ . In particular, if  $(X_1, Y_1), (X_2, Y_2)$  are two independent copies of the random vector  $(X, Y)$ , then

$$\text{Var}^Y(X) = \frac{1}{2} \frac{\mathbb{E}[(X_1 - X_2)^2 Y_1 Y_2]}{\mathbb{E}(Y_1 Y_2)} = C^{\pm 4} \frac{1}{2} \mathbb{E}[(X_1 - X_2)^2] = C^{\pm 4} \text{Var}(X). \quad \square$$

**Proof of Theorem 6.1 on the asymptotic behavior of  $\mathcal{F}_N(\xi) := \frac{1}{V_N} \log \mathbb{E}(e^{\xi S_N})$ :** Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ , s.t.  $V_N := \text{Var}(S_N) \neq 0$  for  $N \geq N_0$ .

Since  $\|S_N\|_\infty < \infty$ , we may differentiate under the expectation and obtain that for all  $k$ ,  $\frac{d^k}{d\xi^k} \mathbb{E}(e^{\xi S_N}) = \mathbb{E}(S_N^k e^{\xi S_N})$ . A direct calculation now shows that

$$\begin{aligned} \mathcal{F}'_N(\xi) &= \frac{1}{V_N} \frac{\mathbb{E}(S_N e^{\xi S_N})}{\mathbb{E}(e^{\xi S_N})} = \frac{1}{V_N} \mathbb{E}^{Y_N^\xi}(S_N), \\ \mathcal{F}''_N(\xi) &= \frac{1}{V_N} \left[ \frac{\mathbb{E}(S_N^2 e^{\xi S_N})}{\mathbb{E}(e^{\xi S_N})} - \left( \frac{\mathbb{E}(S_N e^{\xi S_N})}{\mathbb{E}(e^{\xi S_N})} \right)^2 \right] = \frac{\text{Var}^{Y_N^\xi}(S_N)}{\text{Var}(S_N)}, \text{ where } Y_N^\xi := e^{\xi S_N}. \end{aligned}$$

**Part 1:** Substituting  $\xi = 0$  gives  $\mathcal{F}_N(0) = 0$ ,  $\mathcal{F}'_N(0) = \frac{\mathbb{E}(S_N)}{V_N}$ ,  $\mathcal{F}''_N(0) = 1$ .

**Part 2:**  $\mathcal{F}''_N(\xi) = 0 \Leftrightarrow \text{Var}^{Y_N^\xi}(S_N) = 0 \Leftrightarrow S_N = \text{const} \frac{Y_N^\xi}{\mathbb{E}(Y_N^\xi)} d\mathbb{P}\text{-a.s.} \Leftrightarrow S_N = \text{const } \mathbb{P}\text{-a.s.} \Leftrightarrow \text{Var}(S_N) = 0$ . So  $\mathcal{F}_N$  is strictly convex on  $\mathbb{R}$  for all  $N > N_0$ .

**Part 3:**  $\frac{\tilde{V}_N^\xi(S_N)}{\text{Var}(S_N)} \equiv \frac{\text{Var}^{Z_N}(S_N)}{\text{Var}(S_N)}$ , where  $Z_N^\xi := e^{\xi S_N} \frac{h_{N+1}^\xi}{h_1^\xi}$  (the normalization constant does not matter). Next,  $Z_N^\xi \equiv Y_N^\xi W_N^\xi$ , where  $W_N^\xi := h_{N+1}^\xi / h_1^\xi$ . Lemma 6.8 says that for every  $R > 0$  there is a constant  $C = C(R)$  s.t.  $C^{-1} \leq W_N^\xi \leq C$  for all  $N$  and  $|\xi| \leq R$ . Lemma 6.10 and the obvious identity  $h_n^0 \equiv 1$  imply that  $W_N^\xi \xrightarrow[\xi \rightarrow 0]{} 1$  uniformly in  $N$ . So there is no loss of generality in assuming that  $C(R) \xrightarrow[R \rightarrow 0]{} 1$ .

By Lemma 6.16 with the probability measure  $\frac{e^{\xi S_N}}{\mathbb{E}(e^{\xi S_N})} d\mathbb{P}$  and  $Y = W_N^\xi$ ,

$$\frac{\tilde{V}_N^\xi(S_N)}{V_N \mathcal{F}''_N(\xi)} = \frac{\text{Var}^{Y_N^\xi W_N^\xi}(S_N)}{\text{Var}^{Y_N^\xi}(S_N)} \in [C(R)^{-4}, C(R)^4], \quad \forall |\xi| \leq R, \quad N \geq 1. \quad (6.3.30)$$

By Lemma 6.12(3),  $\tilde{V}_N^\xi(S_N) \asymp V_N$  uniformly on compact sets of  $\xi$ , and by Lemma 6.15 for every  $\varepsilon$  there exists  $\delta, N_\varepsilon > 0$  s.t.  $e^{-\varepsilon} < \tilde{V}_N^\xi(S_N)/V_N < e^\varepsilon$  for all  $N > N_\varepsilon$  and  $|\xi| \leq \delta$ . It follows that for every  $R$  there exists  $C_2(R) > 1$  such that  $C_2(R) \xrightarrow[R \rightarrow 0]{} 1$  and  $C_2(R)^{-1} \leq \mathcal{F}''_N(\xi) \leq C_2(R)$  for all  $|\xi| \leq R$ .

**Part 4:** Suppose  $\varepsilon > 0$ . We saw in part 3 that there exist  $\delta, N_\varepsilon$  s.t.  $e^{-\varepsilon} \leq \mathcal{F}''_N(\xi) \leq e^\varepsilon$  for all  $|\xi| \leq \delta, N \geq N_\varepsilon$ .

Recall that  $\mathcal{F}_N(0) = 0$  and  $\mathcal{F}'_N(0) = \mathbb{E}(S_N)/V_N$ . So for all  $|\xi| \leq \delta$ ,

$$\mathcal{F}_N(\xi) = \mathcal{F}_N(0) + \int_0^\xi \left( \mathcal{F}'_N(0) + \int_{\mathcal{F}'_N(0)}^\eta \mathcal{F}''_N(\alpha) d\alpha \right) d\eta.$$

Since  $\mathcal{F}_N'' = e^{\pm \varepsilon}$  on  $[-\delta, \delta]$  and  $|\eta| \leq |\xi| \leq \delta$ ,

$$\mathcal{F}_N(\xi) = \frac{\mathbb{E}(S_N)}{V_N} \xi + \frac{1}{2} e^{\pm \varepsilon} \left( \xi - \frac{\mathbb{E}(S_N)}{V_N} \right)^2. \quad \square$$

### 6.3.6 Asymptotics of the rate functions.

The *rate functions*  $\mathcal{J}_N(\eta)$  are the Legendre transforms of  $\mathcal{F}_N(\xi) = \frac{1}{V_N} \log \mathbb{E}(e^{\xi S_N})$ . Recall that the Legendre transform of a strictly convex function  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  is the function  $\varphi^* : (\inf \varphi', \sup \varphi') \rightarrow \mathbb{R}$ ,

$$\varphi^*(\eta) = \xi \eta - \varphi(\xi) \text{ for the unique } \xi \text{ s.t. } \varphi'(\xi) = \eta.$$

On its domain,  $\varphi^*(\eta) = \max\{\xi \eta - \varphi(\xi)\}$ .

**Lemma 6.17** Suppose  $\varphi(\xi)$  is strictly convex and twice differentiable on  $\mathbb{R}$ , and let  $\varphi'(\pm\infty) := \lim_{\xi \rightarrow \pm\infty} \varphi'(\xi)$ . Then the Legendre transform  $\varphi^*$  is strictly convex and twice differentiable on  $(\varphi'(-\infty), \varphi'(+\infty))$ , and for every  $\xi \in \mathbb{R}$ ,

$$\varphi^*(\varphi'(t)) = t\varphi'(t) - \varphi(t), \quad (\varphi^*)'(\varphi'(t)) = t, \quad (\varphi^*)''(\varphi'(t)) = \frac{1}{\varphi''(t)} \quad (6.3.31)$$

*Proof.* Under the assumptions of the lemma,  $\varphi'$  is strictly increasing and differentiable. So  $(\varphi')^{-1} : (\varphi'(-\infty), \varphi'(\infty)) \rightarrow \mathbb{R}$  is well-defined, strictly increasing and differentiable, and

$$\varphi^*(\eta) = \eta(\varphi')^{-1}(\eta) - \varphi[(\varphi')^{-1}(\eta)]$$

The lemma follows by differentiation of right-hand-side.  $\square$

#### Proof of Theorem 6.2 on the asymptotics of the rate functions $\mathcal{J}_N := \mathcal{F}_N^*$ :

**Part 1:** Since  $\mathcal{F}_N$  is strictly convex and smooth,  $\mathcal{F}_N'$  is strictly increasing and continuous. So  $\mathcal{F}_N'[-1, 1] = [\mathcal{F}_N'(-1), \mathcal{F}_N'(1)] \equiv [a_N^1, b_N^1]$ , and for every  $\eta \in [a_N^1, b_N^1]$ , there exists a unique  $\xi \in [-1, 1]$  such that  $\mathcal{F}_N'(\xi) = \eta$ . So  $\text{dom}(\mathcal{J}_N) \supset [a_N^1, b_N^1]$ .

By Theorem 6.1 there is  $C > 0$  such that  $C^{-1} \leq \mathcal{F}_N'' \leq C$  on  $[-1, 1]$  for all  $N \geq N_0$ . Since  $\mathcal{F}_N'(0) = \frac{\mathbb{E}(S_N)}{V_N}$  and  $\mathcal{F}_N'(\rho) = \mathcal{F}_N'(0) + \int_0^\rho \mathcal{F}_N''(\xi) d\xi$ , we have

$$b_N^1 \equiv \mathcal{F}_N'(1) \geq \frac{\mathbb{E}(S_N)}{V_N} + C^{-1}, \quad a_N^1 \equiv \mathcal{F}_N'(-1) \leq \frac{\mathbb{E}(S_N)}{V_N} - C^{-1}.$$

So  $\text{dom}(\mathcal{J}_N) \supset [a_N^1, b_N^1] \supset \left[ \frac{\mathbb{E}(S_N)}{V_N} - C^{-1}, \frac{\mathbb{E}(S_N)}{V_N} + C^{-1} \right]$  for all  $N \geq N_0$ .

**Part 2** follows from Lemma 6.17 and the strict convexity of  $\mathcal{F}_N$  on  $[-R, R]$ .

**Part 3:** Let  $J_N := \left[ \frac{\mathbb{E}(S_N)}{V_N} - C^{-1}, \frac{\mathbb{E}(S_N)}{V_N} + C^{-1} \right]$ . In part 1 we constructed functions  $\xi_N : J_N \rightarrow [-1, 1]$  such that  $\mathcal{F}_N'(\xi_N(\eta)) = \eta$ .

Clearly  $\xi_N \left( \frac{\mathbb{E}(S_N)}{V_N} \right) = 0$ . Recalling that  $C^{-1} \leq \mathcal{F}_N'' \leq C$  on  $[-1, 1]$ , we see that  $\xi_N'(\eta) = \frac{1}{\mathcal{F}_N''(\xi_N(\eta))} \in [C^{-1}, C]$  on  $J_N$ . Hence

$$|\xi_N(\eta)| \leq C|\eta - \frac{\mathbb{E}(S_N)}{V_N}| \text{ for all } \eta \in J_N, N \geq N_0.$$

Fix  $0 < \varepsilon < 1$ . By Theorem 6.1(4) there are  $\delta, N_\varepsilon > 0$  s.t.  $e^{-\varepsilon} \leq \mathcal{F}_N'' \leq e^\varepsilon$  on  $[-\delta, \delta]$  for all  $N > N_\varepsilon$ . If  $|\eta - \frac{\mathbb{E}(S_N)}{V_N}| < \delta/C$ , then  $|\xi_N(\eta)| < \delta$ , and  $\mathcal{F}_N''(\xi_N(\eta)) \in [e^{-\varepsilon}, e^\varepsilon]$ .

Since  $\mathcal{F}_N(0) = 0$  and  $\mathcal{F}_N'(0) = \frac{\mathbb{E}(S_N)}{V_N}$ , we have by (6.3.31) that  $\mathcal{J}_N(\frac{\mathbb{E}(S_N)}{V_N}) = \mathcal{J}_N'(\frac{\mathbb{E}(S_N)}{V_N}) = 0$  and  $\mathcal{J}_N''(\eta) = 1/\mathcal{F}_N''(\xi_N(\eta)) \in [e^{-\varepsilon}, e^\varepsilon]$ . Writing

$$\mathcal{J}_N(\eta) = \mathcal{J}_N(\frac{\mathbb{E}(S_N)}{V_N}) + \int_{\frac{\mathbb{E}(S_N)}{V_N}}^{\eta} \left( \mathcal{J}_N'(\frac{\mathbb{E}(S_N)}{V_N}) + \int_{\frac{\mathbb{E}(S_N)}{V_N}}^{\alpha} \mathcal{J}_N''(\beta) d\beta \right) d\alpha,$$

we find that  $\mathcal{J}_N(\eta) = e^{\pm\varepsilon} \frac{1}{2} (\eta - \frac{\mathbb{E}(S_N)}{V_N})^2$  for all  $\eta$  s.t.  $|\eta - \frac{\mathbb{E}(S_N)}{V_N}| \leq \delta/C$ .

**Part 4:** If  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then  $\frac{z_N}{V_N} \in \left[ \frac{\mathbb{E}(S_N)}{V_N} - \delta_N, \frac{\mathbb{E}(S_N)}{V_N} + \delta_N \right]$  with  $\delta_N \rightarrow 0$ . By part 3,  $\mathcal{J}_N(\frac{z_N}{V_N}) \sim \frac{1}{2} \left( \frac{z_N - \mathbb{E}(S_N)}{V_N} \right)^2$ , whence  $V_N \mathcal{J}_N(\frac{z_N}{V_N}) \sim \frac{1}{2} \left( \frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \right)^2$ .  $\square$

Let  $H_N(\eta)$  denote the Legendre transform of  $P_N(\xi)/V_N$ . We will compare  $H_N(\eta)$  to  $\mathcal{J}_N(\eta)$ . This is needed to link the change of measure we performed in section §6.3.3 to the functions  $\mathcal{J}_N$  which appear in the statement of the local limit theorem for large deviations.

**Lemma 6.18** Suppose  $R > 0$  and  $V_N \neq 0$  for all  $N$  large enough. Then

- (1)  $H_N$  is well-defined and real-analytic on  $\left[ \frac{P'_N(-R)}{V_N}, \frac{P'_N(R)}{V_N} \right]$  for all  $N$  large enough.
- (2) There exists  $c > 0$  such that  $H_N(\cdot)$  is well-defined and real-analytic on  $\left( \frac{\mathbb{E}(S_N)}{V_N} - c, \frac{\mathbb{E}(S_N)}{V_N} + c \right)$  for all  $N$  large enough.

*Proof.* Lemma 6.13 and its proof provide real analytic maps

$$\xi_N : \left[ \frac{P'_N(-R)}{V_N}, \frac{P'_N(R)}{V_N} \right] \rightarrow [-R, R] \quad \text{s.t.} \quad \frac{P'_N(\xi_N(\eta))}{V_N} = \eta.$$

Hence  $H_N(\eta) = \frac{1}{V_N} [\xi_N(\eta) P'_N(\xi_N(\eta)) - P_N(\xi_N(\eta))]$  is well-defined and real-analytic on the interval  $\left[ \frac{P'_N(-R)}{V_N}, \frac{P'_N(R)}{V_N} \right]$ . This proves part (1). Part (2) follows from Lemma 6.13(2).  $\square$

**Lemma 6.19** Suppose  $V_N \neq 0$  for all  $N \geq N_0$ , then  $\exists c > 0$  such that

- (1)  $\text{dom}(\mathcal{J}_N) \cap \text{dom}(H_N) \supset \left[ \frac{\mathbb{E}(S_N)}{V_N} - c, \frac{\mathbb{E}(S_N)}{V_N} + c \right]$  for all  $N \geq N_0$ .
- (2) Recall that  $[a_N^R, b_N^R] = [\mathcal{F}_N'(-R), \mathcal{F}_N'(R)]$ . For every  $R > 0$  there exists  $C(R) > 0$  s.t. if  $z/V_N \in [a_N^R, b_N^R]$  and  $N \geq N_0$ , then

$$\left| V_N \mathcal{J}_N\left(\frac{z}{V_N}\right) - V_N H_N\left(\frac{z}{V_N}\right) \right| \leq C(R).$$

- (3) For every  $\varepsilon > 0$ ,  $\exists \delta, N_\varepsilon > 0$  s.t. if  $N \geq N_\varepsilon$  and  $\left| \frac{z - \mathbb{E}(S_N)}{V_N} \right| < \delta$ , then

$$\left| V_N \mathcal{J}_N\left(\frac{z}{V_N}\right) - V_N H_N\left(\frac{z}{V_N}\right) \right| \leq \varepsilon.$$

*Proof.* Part (1) is a direct consequence of Lemma 6.18 and Theorem 6.2(1).

To prove the other parts of the lemma, we use the following consequence of Lemma 6.12(6): For every  $R > 0$ , for all  $N$  large enough, for every  $\eta \in [a_N^R, b_N^R]$ , there exist  $\xi_N^{(1)}, \xi_N^{(2)} \in [-(R+1), (R+1)]$  such that

$$\frac{P'_N(\xi_N^{(1)})}{V_N} = \eta, \quad \mathcal{F}'_N(\xi_N^{(2)}) = \eta.$$

Arguing as in the proof of part 3 of Theorem 6.2, we can also find a constant  $C(R)$  such that  $|\xi_N^{(i)}| \leq C(R)|\eta - \frac{\mathbb{E}(S_N)}{V_N}|$ .

It is a general fact that the Legendre transform of a convex function  $\varphi$  is equal on its domain to  $\varphi^*(\eta) = \sup_{\xi} \{\xi \eta - \varphi(\xi)\}$ . Thus for every  $z \in [a_N^R V_N, b_N^R V_N]$ ,

$$\begin{aligned} V_N \mathcal{J}_N \left( \frac{z}{V_N} \right) &= V_N \sup_{\xi} \left\{ \xi \frac{z}{V_N} - \mathcal{F}_N(\xi) \right\} = V_N \left( \xi_N^{(2)} \frac{z}{V_N} - \mathcal{F}_N(\xi_N^{(2)}) \right) \\ &\leq V_N \left( \xi_N^{(2)} \frac{z}{V_N} - \frac{P_N(\xi_N^{(2)})}{V_N} \right) + \Delta_N(R+1), \text{ see Lemma 6.12(5)} \\ &\leq V_N \sup_{\xi} \left\{ \xi \frac{z}{V_N} - \frac{P_N(\xi)}{V_N} \right\} + \Delta_N(R+1) \equiv V_N H_N \left( \frac{z}{V_N} \right) + \Delta_N(R+1). \end{aligned}$$

So  $V_N \mathcal{J}_N \left( \frac{z}{V_N} \right) - V_N H_N \left( \frac{z}{V_N} \right) \leq \Delta_N(R+1)$ .

Similarly, one can show that  $V_N H_N \left( \frac{z}{V_N} \right) - V_N \mathcal{J}_N \left( \frac{z}{V_N} \right) \leq \Delta_N(R+1)$ , whence

$$\sup_{N \geq N_0} \sup_{z \in [a_N^R V_N, b_N^R V_N]} \left| V_N \mathcal{J}_N \left( \frac{z}{V_N} \right) - V_N H_N \left( \frac{z}{V_N} \right) \right| \leq \sup_{N \geq N_0} \Delta_N(R+1).$$

Part (2) now follows from Lemma 6.12(5).

If instead of taking  $z/V_N \in [a_N^R, b_N^R]$  we take  $z/V_N \in \left( \frac{\mathbb{E}(S_N)}{V_N} - \delta, \frac{\mathbb{E}(S_N)}{V_N} + \delta \right)$ , then  $|\xi_N^{(i)}| < C\delta$ , and the same argument will show that

$$\sup_{N \geq N_0} \sup_{\left| \frac{z - \mathbb{E}(S_N)}{V_N} \right| \leq \delta} \left| V_N \mathcal{J}_N \left( \frac{z}{V_N} \right) - V_N H_N \left( \frac{z}{V_N} \right) \right| \leq \sup_{N \geq N_0} \Delta_N(C\delta).$$

Part (3) follows from Lemma 6.12(5).  $\square$

### 6.3.7 The local limit theorem for large deviations.

**Proof of Theorem 6.5.** We give the proof in the non-lattice case; the modifications needed for the lattice case are routine.

Suppose  $f$  is an a.s. uniformly bounded additive functional of a uniformly elliptic Markov chain  $X$ . We assume that  $f$  is irreducible, and that  $f$  has algebraic range  $\mathbb{R}$ . In this case  $f$  is not center-tight, and  $V_N := \text{Var}(S_N) \rightarrow \infty$  (see §2.1). There is no loss of generality in assuming that  $V_N \neq 0$  for all  $N$ .

Recall that  $[\hat{a}_N, \hat{b}_N] = [\mathcal{F}'_N(-R) - \frac{\mathbb{E}(S_N)}{V_N}, \mathcal{F}'_N(R) - \frac{\mathbb{E}(S_N)}{V_N}]$ , and suppose

$$\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\hat{a}_N, \hat{b}_N].$$

Let  $h_n^\xi(\cdot) := h_n(\cdot, \xi)$ ,  $p_n(\xi)$ , and  $P_N(\xi)$  be as in §§6.3.2, 6.3.3. The assumption on  $z_N$  allows us to construct  $\xi_N \in [-(R+1), (R+1)]$  as in Lemma 6.13:

$$P'_N(\xi_N) = z_N \text{ and } \xi_N = O\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right).$$

Define a Markov array  $\tilde{X} := \{\tilde{X}_n^{(N)} : 1 \leq n \leq N+1\}$  with state spaces  $(\mathfrak{S}_n, \mathcal{B}(\mathfrak{S}_n), \mu_n)$  (the state spaces of  $X$ ), and transition probabilities

$$\tilde{\pi}_{n,n+1}^{(N)}(x, dy) := e^{\xi_N f_n(x, y)} \frac{h_{n+1}(y, \xi_N)}{e^{p_n(\xi_N)} h_n(x, \xi_N)} \cdot \pi_{n,n+1}(x, dy).$$

Let  $\tilde{f} = \{f_n^{(N)} : 1 \leq n \leq N+1, N \in \mathbb{N}\}$  where  $f_n^{(N)} := f_n$ , and set

$$\tilde{S}_N := f_1(\tilde{X}_1^{(N)}, \tilde{X}_2^{(N)}) + \cdots + f_N(\tilde{X}_N^{(N)}, \tilde{X}_{N+1}^{(N)}).$$

Recall that  $e^{\xi_N f_n}$ ,  $h_n$ , and  $e^{p_n(\xi_N)}$  are uniformly bounded away from zero and infinity, by the assumption on  $f$ , and Lemma 6.8. So  $\tilde{\pi}_{n,n+1}^{(N)}(x, dy)$  differ from  $\pi_{n,n+1}(x, dy)$  by densities which are bounded away from zero and infinity uniformly in  $N$ . It follows that  $\tilde{X}$  is uniformly elliptic,  $\tilde{f}$  is a.s. uniformly bounded, and the structure constants of  $(\tilde{X}, \tilde{f})$  are equal to the structure constants of  $(X, f)$  up to a uniformly bounded multiplicative error. Thus

- (1)  $(\tilde{X}, \tilde{f})$  and  $(X, f)$  have the same algebraic ranges, co-ranges, and essential ranges. In particular,  $(\tilde{X}, \tilde{f})$  is irreducible and non-lattice.
- (2)  $(\tilde{X}, \tilde{f})$  is stably hereditary (see Examples 3.5 and 3.6 in §3.2.3).
- (3)  $\tilde{V}_N := \text{Var}(\tilde{S}_N) \xrightarrow[N \rightarrow \infty]{} \infty$  (because  $\tilde{V}_N \asymp \sum_{n=3}^N u_n^2 \asymp V_N \rightarrow \infty$ ).

Furthermore, by the choice of  $\xi_N$ ,  $\mathbb{E}(\tilde{S}_N) \equiv \tilde{\mathbb{E}}^{\xi_N}(S_N) = z_N + O(1)$ , so

$$\frac{z_N - \mathbb{E}(\tilde{S}_N)}{\sqrt{V_N}} = O\left(\frac{1}{\sqrt{V_N}}\right) \xrightarrow[N \rightarrow \infty]{} 0.$$

Therefore  $\tilde{S}_N$  satisfies the local limit theorem (Theorem 4.1):

$$\mathbb{P}_x(\tilde{S}_N - z_N \in (a, b)) \sim |a - b| \Big/ \sqrt{2\pi \tilde{V}_N^{\xi_N}}$$

for every  $x \in \mathfrak{S}_1$  and  $(a, b) \neq \emptyset$ .

We will translate this into an asymptotic for  $\mathbb{P}(S_N - z_N \in (a, b))$ . For all  $N$  large enough, for every  $x \in \mathfrak{S}_1$ ,

$$\begin{aligned} \mathbb{P}_x[S_N - z_N \in (a, b)] &= e^{P_N(\xi_N) - \xi_N z_N} \times \\ &\times \mathbb{E}_x \left( e^{\xi_N S_N} \frac{h_{N+1}^{\xi_N}(X_{N+1}^{(N)})}{e^{P_N(\xi_N)} h_1^{\xi_N}(x)} \cdot \frac{h_1^{\xi_N}(x)}{h_{N+1}^{\xi_N}(X_{N+1}^{(N)})} \cdot e^{\xi_N(z_N - S_N)} 1_{(a,b)}(S_N - z_N) \right) \\ &= e^{P_N(\xi_N) - \xi_N z_N} h_1^{\xi_N}(x) \tilde{\mathbb{E}}_x \left( h_{N+1}^{\xi_N}(\tilde{X}_{N+1}^{(N)})^{-1} \phi_{a,b}(\tilde{S}_N - z_N) \right) \end{aligned} \tag{6.3.32}$$

where  $\phi_{a,b}(t) := 1_{(a,b)}(t) e^{-\xi_N t}$ .

The pre-factor simplifies as follows. By construction  $\frac{P'_N(\xi_N)}{V_N} = \frac{z_N}{V_N}$ . Thus

$$\xi_N z_N - P_N(\xi_N) = V_N \left( \xi_N \frac{z_N}{V_N} - \frac{P_N(\xi_N)}{V_N} \right) = V_N \left( \xi_N \frac{P'_N(\xi_N)}{V_N} - \frac{P_N(\xi_N)}{V_N} \right).$$

So

$$e^{P_N(\xi_N) - \xi_N z_N} = e^{-V_N H_N\left(\frac{z_N}{V_N}\right)}, \quad (6.3.33)$$

where  $H_N(\eta)$  is the Legendre transform of  $P_N(\xi)/V_N$ .

Using the mixing LLT for Markov arrays Theorem 4.3, one can see that

$$\mathbb{E}_x \left( h_{N+1}^{\xi_N} (\tilde{X}_{N+1}^{(N)})^{-1} \phi_{a,b} (\tilde{S}_N - z_N) \right) \sim \frac{\mu_{N+1} \left( 1/h_{N+1}^{\xi_N} \right)}{\sqrt{2\pi \tilde{V}_N^{\xi_N}}} \int_a^b e^{-\xi_N t} dt, \quad (6.3.34)$$

as  $N \rightarrow \infty$ . To do this approximate  $\phi_{a,b}$  in  $L^1(\mathbb{R})$  from below and above continuous functions with compact support, and approximate  $h_{N+1}^{\xi_N}$  in  $L^1(\mathfrak{S}_{N+1}^{(N)}, \mathcal{B}(\mathfrak{S}_{N+1}^{(N)}), \mu_{N+1}^{(N)})$  from above and below by finite linear combinations of indicators of sets with uniformly bounded measure (here  $\mu_{N+1}^{(N)}$  is the distribution of  $X_{N+1}^{(N)}$ ).

Since  $\xi_N$  is bounded, Lemma 6.12(4) tells us that  $\tilde{V}_N^{\xi_N} \sim P_N''(\xi_N)$  as  $N \rightarrow \infty$ . Since  $H_N(\eta)$  is the Legendre transform of  $P_N(\xi)/V_N$ , and  $P'_N(\xi_N)/V_N = z_N/V_N$ ,

$$\tilde{V}_N^{\xi_N} \sim V_N \cdot \left( \frac{P_N''(\xi_N)}{V_N} \right) = \frac{V_N}{H_N''\left(\frac{z_N}{V_N}\right)} \quad \text{as } N \rightarrow \infty. \quad (6.3.35)$$

Substituting (6.3.33), (6.3.34), and (6.3.35) in (6.3.32), we obtain the following:

$$\begin{aligned} \mathbb{P}_x[S_N - z_N \in (a, b)] &\sim \left[ \frac{e^{-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)}}{\sqrt{2\pi V_N}} \int_a^b e^{-\xi_N t} dt \right] \times \\ &\times \underbrace{\left[ e^{V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right) - V_N H_N\left(\frac{z_N}{V_N}\right)} \sqrt{H_N''\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right)} \right]}_{\hat{\rho}_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right)} \times \underbrace{\left[ h_1^{\xi_N}(x) \mu_{N+1} \left( \frac{1}{h_{N+1}^{\xi_N}} \right) \right]}_{\bar{\rho}_N\left(x, \frac{z_N - \mathbb{E}(S_N)}{V_N}\right)} \end{aligned}$$

Let  $\eta_N := \frac{z_N - \mathbb{E}(S_N)}{V_N}$ , then  $\xi_N = \xi_N(\eta_N)$  where  $\xi_N : [\hat{a}_N, \hat{b}_N] \rightarrow [-(R+1), (R+1)]$  is defined implicitly by  $P'_N(\xi_N(\eta)) = \eta V_N + \mathbb{E}(S_N)$ . Lemma 6.13 shows that  $\xi_N(\cdot)$  is well-defined.

Notice that there exists a constant  $L = L(R)$  such that  $|\eta_N| \leq L(R)$ . Indeed,  $\eta_N \in [\hat{a}_N^R, \hat{b}_N^R]$  and  $|\hat{a}_N^R|, |\hat{b}_N^R| \leq |\mathcal{F}'(\pm R) - \mathcal{F}'(0)| \leq R \sup_{[-R, R]} \mathcal{F}_N''$ , which is uniformly bounded by Theorem 6.1(3).

The functions  $\hat{\rho}_N : [-L, L] \rightarrow \mathbb{R}$  are defined by

$$\hat{\rho}_N(\eta) := e^{V_N \mathcal{I}_N\left(\eta + \frac{\mathbb{E}(S_N)}{V_N}\right) - V_N H_N\left(\eta + \frac{\mathbb{E}(S_N)}{V_N}\right)} \sqrt{H_N''(\eta)}.$$

Lemma 6.19 and Theorem 6.2 say that there exists  $C$  such that

$$C^{-1} \leq \hat{\rho}_N(\eta) \leq C \text{ for all } N \text{ and } |\eta| \leq L.$$

They also say that for every  $\varepsilon > 0$  there are  $\delta, N_\varepsilon > 0$  s.t.

$$e^{-\varepsilon} \leq \widehat{\rho}_N(\eta) \leq e^{\varepsilon} \text{ for all } N > N_{\varepsilon} \text{ and } |\eta| \leq \delta.$$

In particular, if  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \rightarrow 0$ , then  $\widehat{\rho}_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right) \xrightarrow[N \rightarrow \infty]{} 1$ .

The functions  $\overline{\rho}_N : \mathfrak{S}_1 \times (-c, c) \rightarrow \mathbb{R}$  are defined by

$$\overline{\rho}_N(x, \eta) := h_1(x, \xi(\eta)) \mu_{N+1} \left( \frac{1}{h_{N+1}(x, \xi(\eta))} \right).$$

By Lemma 6.8, there exists a constant  $C$  such that

$$C^{-1} \leq \overline{\rho}_N(x, \eta) \leq C \text{ for all } N \text{ and } |\eta| \leq L.$$

By Lemma 6.10 and the obvious identity  $h_n(\cdot, 0) \equiv 1$ ,  $\|h_n^{\xi} - 1\|_{\infty} \xrightarrow[\xi \rightarrow 0]{} 0$  uniformly in  $n$ . Since  $|\xi(\eta)| \leq C|\eta|$ , for every  $\varepsilon > 0$  there are  $\delta, N_{\varepsilon} > 0$  such that

$$e^{-\varepsilon} \leq \overline{\rho}_N(x, \eta) \leq e^{\varepsilon} \text{ for all } x \in \mathfrak{S}_1, N > N_{\varepsilon}, \text{ and } |\eta| \leq \delta.$$

Setting  $\rho_N := \widehat{\rho}_N \cdot \overline{\rho}_N$  we complete the proof of theorem in the non-lattice case. The modifications needed for the lattice case are routine, and are left to the reader.  $\square$

### 6.3.8 Rough bounds in the reducible case.

**Proof of Theorem 6.6:** We proceed as in the proof of Theorem 6.5 in §6.3.7, but using the rough bounds of §5.2.5 instead of the precise LLT to estimate the probabilities for the change of measure.

Let  $\hbar = 100K + 1$  where  $K = \text{ess sup}(\mathbf{f})$ . Then using Theorem 5.4 and the assumption that  $z_N \in [\mathcal{F}'_N(\varepsilon), b_N^R]$  we get that there exist a constant  $\bar{c} = \bar{c}(R)$  and  $\xi_N := \xi_N\left(\frac{z_N}{V_N}\right) \in [\varepsilon, R + 1]$  such that for all  $N$  large enough,

$$\bar{c} e^{-\xi_N \hbar} \hbar \leq \frac{\sqrt{V_N} \mathbb{P}(S_N - z_N \in [0, \hbar])}{e^{-V_N \mathcal{J}_N\left(\frac{z_N}{V_N}\right)}}. \quad (6.3.36)$$

Note that Theorem 5.4 is applicable since  $\hbar > 2\delta(\mathbf{f})$  due to Corollary 3.3).

Since  $\mathbb{P}(S_N \geq z_N) \geq \mathbb{P}(S_N - z_N \in [0, \hbar])$  the lower bound follows.

Likewise applying Lemma 5.11 we conclude that there is a constant  $C^* = C^*(R)$  s.t for all  $N$  large enough we have, uniformly in  $j \in \mathbb{N} \cup \{0\}$ ,

$$\frac{\sqrt{V_N} \mathbb{P}(S_N - z_N \in [\hbar j, \hbar(j+1)])}{e^{-V_N \mathcal{J}_N\left(\frac{z_N}{V_N}\right)}} \leq C^* e^{-\xi_N \hbar j}.$$

Summing over  $j$  we obtain the lower bound.  $\square$

## 6.4 Large deviations threshold

The results of this chapter are all stated for  $z_N$  s.t. for some  $R > 0$  and all sufficiently large  $N$ ,  $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ . In this section we will discuss how restrictive is this assumption.

We say that a sequence  $\{z_N\}$  is  **$R$ -admissible** if there is a constant  $N_0$  s.t. for  $N \geq N_0 \exists \xi_N \in [-R, R]$  such that  $P'_N(\xi_N) = z_N$ . A sequence  $\{z_N\}$  is **admissible** if it is  $R$ -admissible for some  $R$ .

A number  $z$  is called **reachable** (respectively  **$R$ -reachable**) if the sequence  $\{zV_N\}$  is admissible (respectively  $R$ -admissible).

We denote the set of  $R$ -reachable points by  $\mathcal{C}_R$  and the set of reachable points by  $\mathcal{C}$ . Since  $P'_N$  is monotone increasing,

$$\text{int}(\mathcal{C}) = (\mathfrak{c}_-, \mathfrak{c}_+)$$

for some  $\mathfrak{c}_\pm = \mathfrak{c}_\pm(X)$ .

### Example 6.20 (Sums of iid's)

Let  $S_N = \sum_{n=1}^N X_n$  where  $X_n$  are iid random variables having law  $X$  with expectation zero and variance one. Recall from Example 6.2 that in this case  $\mathcal{F}_N$  does not depend on  $N$ <sup>2</sup> so by property (ii) of Example 6.2 we obtain

$$\mathfrak{c}_- = \text{ess inf}(X), \quad \mathfrak{c}_+ = \text{ess sup}(X). \quad (6.4.1)$$

Then  $S_N/N \in [\mathfrak{c}_-, \mathfrak{c}_+]$  almost surely for all  $N$ , and therefore  $\mathbb{P}[S_N - zN \in (a, b)]$  is zero when  $z \notin [\mathfrak{c}_-, \mathfrak{c}_+]$ . Henceforth we refer to such  $z$  as “irrelevant.”

Not all relevant  $z$  are reachable:  $z$  is reachable only when  $z \in (\mathfrak{c}_-, \mathfrak{c}_+)$ . Our results do not apply for  $z = \mathfrak{c}_\pm$ . Indeed different asymptotic behavior may hold for  $z_N$  s.t.  $\frac{z_N}{V_N} \rightarrow \mathfrak{c}_\pm$ , see Example 6.1. Still, the large deviation LLT for  $\mathbb{P}[S_N - zN \in (a, b)]$  holds for most “relevant” values of  $z$ . Our next example shows that this is not always the case:

### Example 6.21

Let  $X_n = (Y_n, Z_n)$  where  $\{Y_n\}$ ,  $\{Z_n\}$  are two independent sequences of iid random variables having uniform distribution on  $[0, 1]$ . Fix a sequence  $\{p_n\}$  and let

$$f_n(Y_n, Z_n) = \begin{cases} Z_n & \text{if } Y_n > p_n \\ 2 & \text{if } Y_n \leq p_n. \end{cases}$$

We now discuss two possible choices of  $\{p_n\}$ .

(a) Let  $f'$  be defined as above with  $p_n \equiv \frac{1}{2}$ . Then  $f'_n$  are iid so by discussion of the Example 6.20 the results of the present chapter apply to  $\mathbb{P}(S'_N \in zN + (a, b))$  provided that  $z \in (0, 2)$  while the possible range of  $\frac{S_N(f')}{N}$  is  $[0, 2]$ .

(b) Let  $f''$  be defined as above with  $p_n$  tending to 0 as  $n \rightarrow \infty$ . Since  $\text{Var}(Z_n) = \frac{1}{12}$  it follows that  $V_N = (1 + o(1)) \frac{N}{12}$ . We shall show below that in case (b)

$$\mathfrak{c}_- = 0, \quad \mathfrak{c}_+ = 12. \quad (6.4.2)$$

In other words the results of the present chapter apply to  $\mathbb{P}(S_N(f'') \in zN + (a, b))$  provided that  $z \in (0, 1)$ . On the other hand, the possible range of  $\frac{S_N(f'')}{N}$  is  $[0, 2]$  since for each fixed  $N$  the distributions of  $S_N(f')$  and  $S_N(f'')$  are absolutely continuous with respect to each other. We will see that the reason our results do not apply for  $z > 1$  is that in that case  $\mathbb{P}(S_N(f'') \geq zN)$  decays super exponentially.

<sup>2</sup> Note that in this case we also have  $P_N(\xi)/N = \mathcal{F}_N(\xi)$ , since  $\mathcal{L}_{N,\xi}(e^{\xi x_{n+1}}) = \mathbb{E}(e^{\xi X}) \cdot e^{\xi x_n}$ , whence  $p_n(\xi) = \ln \mathbb{E}(e^{\xi X})$ .

In this section we discuss methods for computing  $c_{\pm}$  (in particular, proving (6.4.2)) and provide sufficient conditions for good behavior, when  $(c_-, c_+)$  covers “most” relevant  $z$ .

**Lemma 6.22**  $\forall R > 0 \exists \varepsilon = \varepsilon(R) > 0$  s.t. if  $\{z_N\}$  is  $R$ -admissible, and  $|\bar{z}_N - z_N| \leq \varepsilon V_N$ , then  $\{\bar{z}_N\}$  is  $(R+1)$ -admissible.

*Proof.* By the uniform strict convexity of  $\frac{P_N}{V_N}$  on  $[-(R+1), (R+1)]$ , there exists  $\varepsilon > 0$  such that  $P'_N(R+1) \geq z_N + \varepsilon V_N$  and  $P'_N(-(R+1)) \leq z_N - \varepsilon V_N$ .  $\square$

**Corollary 6.23** (a)  $\mathcal{C}$  is open, and (b) if  $\mathbb{E}(S_n) \equiv 0$ , then  $\mathcal{C}$  is a non-empty neighborhood of zero.

*Proof.* Part (a) follows from Lemma 6.22. Part (b) follows from (6.3.23).  $\square$

Without the assumption  $\mathbb{E}(S_N) = 0$ ,  $\mathcal{C}$  may be empty. Even though Theorem 6.2 provides many admissible sequences, the associated  $\frac{z_N}{V_N}$  need not converge:

**Example 6.24** An example with  $\mathcal{C} = \emptyset$  and with admissible sequences  $\{z_N\}$  such that  $z_N/V_N$  does not converge.

Let  $N_k = 10^k$ . Consider  $X_n = a_n + U_n$  where  $U_n$  are iid having uniform distribution on  $[0, 1]$  and

$$a_n = \begin{cases} 10 & \text{if } N_{2k} \leq n < N_{2k+1}, \\ -10 & \text{if } N_{2k+1} \leq n < N_{2k+2}. \end{cases}$$

With probability one  $S_{N_{2k+1}} > N_{2k+1}$ ,  $S_{N_{2k}} < -N_{2k}$ . The first inequality gives  $\mathcal{C} \cap (-\infty, 0] = \emptyset$ , the second one gives  $\mathcal{C} \cap [0, +\infty) = \emptyset$ . Hence  $\mathcal{C} = \emptyset$ .

In this example, if  $\{z_N\}$  is an admissible sequence then  $\frac{z_{N_{2k+1}}}{N_{2k+1}} \geq 1$  and  $\frac{z_{N_{2k}}}{N_{2k}} \leq -1$ . Since  $V_N = \frac{N}{12}$  the ratio  $\frac{z_N}{V_N}$  does not converge.

**Theorem 6.7.** Let  $f$  be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ , with essential range  $\mathbb{Z}$  or  $\mathbb{R}$ . The following are equivalent:

- (a)  $\{z_N\}$  is admissible.
- (b)  $\exists \varepsilon > 0, \eta > 0$  s.t.  $\forall \{\bar{z}_N\}$  with  $|\bar{z}_N - z_N| \leq \varepsilon V_N$  and  $\forall a_N, b_N$  s.t.  $|a_N|, |b_N| \leq 10$  and  $b_N - a_N > 1$  we have  $\mathbb{P}(S_N \in \bar{z}_N + (a_N, b_N)) \geq \eta^{V_N}$ .
- (c)  $\exists \varepsilon > 0, \eta > 0$  s.t.  $\mathbb{P}(S_N \geq z_N + \varepsilon V_N) \geq \eta^{V_N}$  and  $\mathbb{P}(S_N \leq z_N - \varepsilon V_N) \geq \eta^{V_N}$ .

**Example 6.25** The case  $\varepsilon = 0$ .

Let  $S_N = \sum_{n=1}^N X_n$  where  $X_n$  are iid supported on  $[\alpha, \beta]$  and such that  $X$  has an atom on the right edge:  $\mathbb{P}(X = \beta) = \gamma > 0$ . Then  $\mathbb{P}[S_N \geq \beta N] = \mathbb{P}[S_N = \beta N] = \gamma^N$  while  $\mathbb{P}[S_N \geq \beta N + 1] = 0$ . Thus  $\{\beta N\}$  is not admissible. This example shows that taking  $\varepsilon = 0$  in part (c) of Theorem 6.7 gives a condition which is **not** equivalent to the conditions (a)–(c) of the theorem.

*Proof.* (a)  $\Rightarrow$  (b) : If  $\{z_N\}$  is admissible then by Lemma 6.22  $\exists \varepsilon > 0$  such that if  $|\bar{z}_N - z_N| \leq \varepsilon V_N$  then  $\{\bar{z}_N\}$  is admissible. Now (b) follows from formula (6.3.36) in the proof of Theorem 6.6.

(b)  $\Rightarrow$  (c) : The bound  $\mathbb{P}[S_N \geq z_N + \varepsilon V_N] \geq \eta^{V_N}$  follows from part (b) with  $\bar{z}_N = z_N + \varepsilon V_N$ ,  $a_N = 0$ ,  $b_N = 1.1$ . The lower bound is similar.

(c)  $\Rightarrow$  (a) : Our assumptions on the essential range imply that  $(X, f)$  is not center-tight, and therefore  $V_N \rightarrow \infty$ . By Lemma 6.12(5)  $P_N(R) - V_N \mathcal{F}_N(R)$  is eventually bounded, and therefore for some  $c(R) > 0$  and all  $N > N(R)$ ,

$$e^{P_N(R)} \geq c(R) \mathbb{E} \left( e^{RS_N} \right) \geq c(R) \mathbb{E} \left( e^{RS_N} 1_{[S_N \geq z_N + \varepsilon V_N]} \right) \geq c(R) \eta^{V_N} e^{R(z_N + \varepsilon V_N)}.$$

This implies that for all  $N$  large enough  $P_N(R) \geq R(z_N + (\varepsilon/2))V_N$ .

Since  $P_N(0) = 0$  the Mean Value Theorem tells us that  $\exists \xi_N^+ \in [0, R]$  such that  $P'_N(\xi_N^+) \geq z_N + \frac{\varepsilon V_N}{2}$ . Likewise we can find  $\xi_N^- \in [-R, 0]$  such that  $P'_N(\xi_N^-) \leq z_N - \frac{\varepsilon V_N}{2}$ . By the Intermediate Value Theorem  $\exists \xi_N \in [\xi_N^-, \xi_N^+]$  s.t.  $P'_N(\xi_N) = z_N$ .  $\square$

**Corollary 6.26** *Under the assumptions of the previous theorem, if  $\mathbb{E}(S_N) \equiv 0$  then  $\mathfrak{c}_+ = \sup\{z : \mathfrak{J}(z) < \infty\}$ , where*

$$\mathfrak{J}(z) = \limsup_{N \rightarrow \infty} \frac{|\log \mathbb{P}(S_N \in zV_N + [-1, 1])|}{\log V_N}.$$

*Proof.* By Theorem 6.7(b), if  $z \in (\mathfrak{c}_-, \mathfrak{c}_+)$ , then  $\mathfrak{J}(z) < \infty$ . So  $\mathfrak{c}_+ \leq \sup\{z : \mathfrak{J}(z) < \infty\}$ .

To see the other inequality, note that  $\mathfrak{c}_+ > 0$  (by Corollary 6.23), and  $\mathfrak{J}(0) < \infty$  (by (6.3.36)). We will show that

$$\frac{1}{2} \sup\{z : \mathfrak{J}(z) < \infty\} < \bar{z} < \sup\{z : \mathfrak{J}(z) < \infty\} \Rightarrow \bar{z} \text{ is admissible,} \quad (6.4.3)$$

and deduce that  $\mathfrak{c}_+ \geq \sup\{z : \mathfrak{J}(z) < \infty\}$ .

Fix  $\bar{z}$  as in (6.4.3), then  $\exists \varepsilon > 0$  s.t.  $\mathfrak{J}(\bar{z} + 2\varepsilon) < \infty$  and  $\bar{z} - \varepsilon > 0$ . Necessarily  $\exists \eta > 0$  s.t. for all  $N$  large enough

$$\begin{aligned} \mathbb{P}[S_N \geq (\bar{z} + \varepsilon)V_N] &\geq \mathbb{P}[S_N \in (\bar{z} + 2\varepsilon)V_N + [-1, 1]] \geq \eta^N \\ \mathbb{P}[S_N \leq (\bar{z} - \varepsilon)V_N] &\geq \mathbb{P}[S_N \leq 0] = \frac{1}{2} + o(1) \geq \eta^N \end{aligned}$$

By Theorem 6.7(c),  $\bar{z}$  is admissible.  $\square$

We say that  $(X, f)$  and  $(\tilde{X}, \tilde{f})$  are related by the change of measure if  $f_n \equiv \tilde{f}_n$  and  $\pi_n(x, dy)$  is equivalent to  $\tilde{\pi}_n(x, dy)$  with

$$\bar{\varepsilon} \leq \frac{\tilde{\pi}_n(x, dy)}{\pi_n(x, dy)} \leq \bar{\varepsilon}^{-1}.$$

**Lemma 6.27** *Suppose  $f$  is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain  $X$ . If  $(X, f)$  and  $(\tilde{X}, \tilde{f})$  are related by the change of measure and  $V_N \geq cN$  for some  $c > 0$ , then  $\{z_N\}$  is  $(X, f)$ -admissible iff  $\{z_N\}$  is  $(\tilde{X}, \tilde{f})$ -admissible.*

*Proof.* Since  $X$  is uniformly elliptic,  $\tilde{X}$  is uniformly elliptic. The exponential mixing bounds for uniformly elliptic chains imply that  $\tilde{V}_N := \text{Var}[S_N(\tilde{X}, \tilde{f})]$  and  $V_N := \text{Var}[S_N(X, f)]$  are both  $O(N)$ . Without loss of generality,  $cN \leq V_N \leq c^{-1}C_N$ .

Under the assumptions of the Lemma, the structure constants of  $(X, f)$  are equal to the structure constants of  $(\tilde{X}, \tilde{f})$  up to bounded multiplicative error. By Theorem 2.7,  $\tilde{V}_N := \text{Var}[S_N(\tilde{X}, \tilde{f})] \asymp V_N$  as  $N \rightarrow \infty$ . So  $\exists \tilde{c} > 0$  s.t.  $\tilde{c}N \leq \tilde{V}_N \leq \tilde{c}^{-1}N$ .

Let  $\{z_N\}$  be  $(X, f)$ -admissible. Then there are  $\varepsilon > 0, \eta > 0$  such that

$$\mathbb{P}[S_N \geq z_N + \varepsilon V_N] \geq \eta^N, \quad \mathbb{P}[S_N \leq z_N - \varepsilon V_N] \geq \eta^N.$$

It follows that  $\tilde{\mathbb{P}}[S_N(\tilde{X}, \tilde{f}) \geq z_N + \tilde{\varepsilon} \tilde{V}_N] \geq \tilde{\eta}^N$ ,  $\tilde{\mathbb{P}}[S_N(\tilde{X}, \tilde{f}) \leq z_N - \tilde{\varepsilon} \tilde{V}_N] \geq \tilde{\eta}^N$  where  $\tilde{\eta} = \eta \bar{\varepsilon}$  and  $\tilde{\varepsilon} := \tilde{c}c\varepsilon$ . Hence  $\{z_N\}$  is  $\tilde{X}$ -admissible.  $\square$

**Lemma 6.28** *Let  $f$  and  $\tilde{f}$  be two a.s. uniformly bounded additive functionals on the same uniformly elliptic Markov chain. Suppose  $V_N := \text{Var}[S_N(f)] \rightarrow \infty$  and*

$$\lim_{N \rightarrow \infty} \frac{\|S_N(\tilde{f}) - S_N(f)\|_\infty}{V_N} = 0. \quad (6.4.4)$$

*Then  $\{z_N\}$  is  $f$ -admissible iff  $\{z_N\}$  is  $\tilde{f}$ -admissible.*

*Proof.* We write  $S_N = S_N(f)$ , and  $\tilde{S}_N = S_N(\tilde{f})$ . By the assumptions of the lemma,  $\tilde{V}_N := \text{Var}(\tilde{S}_N) \sim V_N$  as  $N \rightarrow \infty$ .

Let  $\{z_N\}$  be  $f$ -admissible. By Theorem 6.7(b), there are  $\varepsilon > 0, \eta > 0$  such that

$$\mathbb{P}[S_N \geq z_N + \varepsilon V_N] \geq \eta^N, \quad \mathbb{P}[S_N \leq z_N - \varepsilon V_N] \geq \eta^N.$$

It now follows from (6.4.4) that for large  $N$

$$\tilde{\mathbb{P}}\left[\tilde{S}_N \geq z_N + \frac{\varepsilon}{2}\tilde{V}_N\right] \geq \eta^N, \quad \tilde{\mathbb{P}}\left[\tilde{S}_N \leq z_N - \frac{\varepsilon}{2}\tilde{V}_N\right] \geq \eta^N.$$

Hence  $\{z_N\}$  is  $\tilde{f}$ -admissible.  $\square$

We end this section by proving (6.4.2).

*Proof of (6.4.2).* To show that  $c_+ \leq 12$  assume by contradiction that  $\text{int}(\mathcal{C})$  contained some  $z > 12$ . Then Theorem 6.6 would imply that

$$\mathbb{P}[S_N \geq zV_N] \geq \eta^{V_N} \text{ for some } \eta > 0. \quad (6.4.5)$$

Note that

$$\begin{aligned} \log \mathbb{E}(e^{\xi f_n(Y_n, Z_n)}) &= \log \left( p_n e^{2\xi} + (1 - p_n) \mathbb{E}(e^{\xi U[0,1]}) \right) \\ &= \log \left( p_n e^{2\xi} + (1 - p_n) \frac{e^\xi - 1}{\xi} \right) \xrightarrow{n \rightarrow \infty} \log \frac{e^\xi - 1}{\xi} \end{aligned}$$

because  $p_n \rightarrow 0$ . So

$$\mathcal{F}_N(\xi) = \frac{1}{V_N} \log \prod_{n=1}^N \mathbb{E}(e^{\xi f_n(Y_n, Z_n)}) \sim \frac{12}{N} \sum_{n=1}^N \log \mathbb{E}(e^{\xi f_n(Y_n, Z_n)}) \xrightarrow{N \rightarrow \infty} 12 \log \left( \frac{e^\xi - 1}{\xi} \right).$$

The last expression is strictly smaller than  $12\xi$  if  $\xi > 0$ . Therefore for any  $\xi > 0$  we have for sufficiently large  $N$ ,  $\mathbb{E}(e^{\xi S_N}) \leq e^{12V_N\xi}$ . By Markov's inequality,

$$\mathbb{P}[S_N \geq zV_N] \leq e^{(12-z)V_N\xi} \text{ for all } \xi > 0 \text{ and } N \text{ sufficiently large.}$$

But this is incompatible with (6.4.5), since  $z > 12$ . Therefore  $c_+ \leq 12$ .

Next we show that  $(0, 12) \in \text{Int}(\mathcal{C})$ . By Theorem 6.7 it suffices to show that for every  $\mathfrak{z} := \frac{z}{12} \in (0, 1)$ ,  $\exists \varepsilon, \eta > 0$  such that  $\mathbb{P}[A_\varepsilon^\pm(N)] \geq \eta^N$  where

$$A_\varepsilon^+(N) = \mathbb{P}[S_N \geq (\mathfrak{z} + \varepsilon)N], \quad A_\varepsilon^-(N) = \mathbb{P}[S_N \leq (\mathfrak{z} - \varepsilon)N].$$

Take  $\varepsilon > 0$  so small that  $\bar{\mathfrak{z}} := z + \varepsilon < 1$ . Since  $S_N \geq Z_N$  we have

$$\mathbb{P}[S_N \geq \bar{\mathfrak{z}}N] \geq \mathbb{P}\left[\sum_{n=1}^N Z_n \geq \bar{\mathfrak{z}}N\right].$$

The RHS is greater than some  $\bar{\eta}^N$  in view of Theorem 6.6 and equation (6.4.1) from Example 6.20. It follows that  $\mathfrak{c}^+ = 12$ .

The proof of the fact that  $\mathfrak{c}^- = 0$  is similar but easier.  $\square$

## 6.5 Notes and references

The reader should note the difference between the LLT for large deviations and the large deviations principle (LDP): LLT for large deviations give the asymptotics of  $\mathbb{P}[S_N - z_N \in (a, b)]$  or  $\mathbb{P}[S_N > z_N]$ ; The LDP gives the asymptotics of the *logarithm* of  $\mathbb{P}[S_N > z_N]$ , see Dembo & Zeitouni [35] and Varadhan [149].

The interest in precise asymptotics for  $\mathbb{P}[S_N > z_N]$  in the regime of large deviations goes back to the first paper on large deviations, by Cramér [31]. That paper gave an asymptotic series expansion for  $\mathbb{P}[S_N - \mathbb{E}(S_N) > x]$  for  $S_N$  =sums of iid's. The first sharp asymptotics for  $\mathbb{P}[S_N - z_N \in (a, b)]$  appear to be the work of Richter [126], [74, chapter 7] and Blackwell & Hodges [14].

These results were refined by many authors, with important contributions by Petrov [116], Linnik [96], Moskvin [106], Bahadur & Ranga Rao [11], Statulavicius [144] and Saulis [134]. We refer the reader to the books of Ibragimov & Linnik [74], Petrov [117], and of Saulis & Statulavicius [135] for accounts of these and other results, and also to the survey of Nagaev [110] for a discussion of the case of sums of independent random variables which are not necessarily identically distributed.

Plachky and Steinebach [118] and Chaganty & Sethuraman [22, 23] proved LLT for large deviations for arbitrary sequences of random variables  $T_n$  (e.g. sums of dependent random variables), subject only to assumptions on the asymptotic behavior of the normalized log-moment generating functions of  $T_n$  and their Legendre-Fenchel transforms (their rate functions). Our LLT for large deviations are in the spirit of these results.

Corollary 6.4 is an example of a limit theorem conditioned on a large deviation. For other examples of such results, in the context of statistical physics, see [39].

We comment on some of the technical devices in the proofs. The “change of measure” trick discussed in section 6.3.1 goes back to Cramér [31] and is a standard idea in large deviations. In the classical homogeneous setup, a single parameter  $\xi_N = \xi$  works for all times  $N$ , but in our inhomogeneous setup, we need to allow the parameter  $\xi_N$  to depend  $N$ . For other instances of changes of measure which involve a time dependent parameter, see Dembo & Zeitouni [34] and references therein.

Birkhoff’s Theorem on the contraction of Hilbert’s projective metric is proved in [13]. Results similar to Lemma 6.7 on the existence of the generalized eigenfunction  $h_n^\xi$  were proved by many authors in many different contexts, see for example [81], [56], [15], [131], [54], [69], [67]. The analytic dependence of the generalized eigenvalue and eigenvector on the parameter  $\xi$  was considered in a different context (the top Lyapunov exponent) by Ruelle [129] and Peres [115]. Our proof of Lemma 6.10 follows closely a proof in [54]. For an account of the theory of real-analyticity for vector valued functions, see [43] and [151].



## Chapter 7

# Miscellaneous examples and special cases

In this chapter we consider several special cases where our general results take stronger form. These include homogeneous Markov chains, asymptotically homogeneous additive functionals. We also explain how continuity assumptions can be used to strengthen the results of the previous chapters.

### 7.1 Homogenous Markov chains

A Markov chain  $X = \{X_n\}$  is called **homogeneous** if its state spaces and transition probabilities do not depend on  $n$

$$\mathfrak{S}_n = \mathfrak{S}, \quad \mu_n = \mu, \quad \pi_n(x, dy) = \pi(x, dy) \quad \text{for all } n,$$

and  $X_n$  is stationary.

An additive functional on a homogeneous Markov chain is called **homogeneous** if  $f = \{f_n\}$  and

$$f_n(x, y) = f(x, y) \quad \text{for all } n.$$

The LLT for homogeneous countable state Markov chains is due to Nagaev. The following version, which allows continuous spaces, follows from results in [73].

**Theorem 7.1.** *Let  $f$  denote an a.s. uniformly bounded homogeneous additive functional on a uniformly elliptic homogeneous Markov chain  $X$ .*

(1) **Asymptotic Variance:** *The limit  $\sigma^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \text{Var}(S_N)$  exists, and  $\sigma^2 = 0$  iff we can represent  $f(X_1, X_2) = a(X_2) - a(X_1) + \kappa$  a.s. where  $a : \mathfrak{S} \rightarrow \mathbb{R}$  is a bounded measurable function and  $\kappa$  is a constant, equal to  $\mathbb{E}(f(X_1, X_2))$ .*

(2) **CLT:** *If  $\sigma^2 > 0$ , then  $\frac{S_N - \mathbb{E}(S_N)}{\sqrt{N}}$  converges in probability as  $N \rightarrow \infty$  to the Gaussian distribution with mean zero and variance  $\sigma^2$ .*

(3) **LLT:** *If  $\sigma^2 > 0$  then exactly one of the following options holds:*

(a) **Non-Lattice LLT:** *If  $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{N}} \rightarrow z$ , then for every interval  $[a, b]$ ,*

$$\mathbb{P}[S_N - z_N \in [a, b]] = [1 + o(1)] \frac{e^{-z^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2 N}} (b - a), \text{ as } N \rightarrow \infty;$$

(b) Periodicity: There exist  $\kappa \in \mathbb{R}, t > 0$  and a bounded measurable function  $a : \mathfrak{S} \rightarrow \mathbb{R}$  such that  $f(X_1, X_2) + a(X_1) - a(X_2) + \kappa \in t\mathbb{Z}$  a.s.

*Proof.* Let  $V_N := \text{Var}(S_N)$  and  $f_k := f(X_k, X_{k+1})$ , and assume without loss of generality that  $\mathbb{E}[f(X_1, X_2)] = 0$ .

**Proof of part (1):** By stationarity,  $\mathbb{E}(f_n) = 0$  for all  $n$ , and so

$$V_N = \mathbb{E}(f_n^2) = \sum_{n=1}^N \mathbb{E}(f_n^2) + 2 \sum_{1 \leq m < n \leq N} \mathbb{E}(f_n f_m).$$

By stationarity,  $\mathbb{E}(f_n f_m) = \mathbb{E}(f_0 f_{n-m})$  and

$$\frac{1}{N} V_N = \mathbb{E}(f_0^2) + 2 \sum_{k=1}^{N-1} \mathbb{E}(f_0 f_k) \left(1 - \frac{k}{N}\right).$$

$|\mathbb{E}(f_0 f_m)|$  decays exponentially (Prop. 1.11), so  $\sum |\mathbb{E}(f_0 f_k)| < \infty$ , whence

$$\sigma^2 := \lim_{N \rightarrow \infty} \frac{1}{N} \text{Var}(S_N) = \mathbb{E}(f_0^2) + 2 \sum_{k=1}^{\infty} \mathbb{E}(f_0 f_k). \quad (7.1.1)$$

(This identity for  $\sigma^2$  is called the **Green-Kubo formula**.)

Let  $u_n$  denote the structure constants of  $(X, f)$ . The homogeneity assumptions implies that  $u_n$  is independent of  $n$ , say  $u_n = u$  for all  $n$ . It follows that  $U_N \equiv u_3^2 + \dots + u_N^2 = (N-2)u^2$ . Now we have two cases:

- (I)  $u > 0$ : In this case by Theorem 2.7,  $V_N \asymp U_N \asymp N$ , whence  $\sigma^2 > 0$ .
- (II)  $u = 0$ : In this case,  $\text{Var}(S_N) = O(1)$  by Theorem 2.7, whence  $\sigma^2 = 0$  and  $f$  is center-tight.

By the Gradient Lemma, (Lemma 2.9),  $f(X_1, X_2) = a_2(X_2) - a_1(X_1) + \kappa$  for some  $a_1, a_2 : \mathfrak{S} \rightarrow \mathbb{R}$  bounded and measurable and  $\kappa \in \mathbb{R}$ . In the homogeneous case, we may take  $a_1 \equiv a_2$ , see (2.2.4) in the proof of the Gradient Lemma. So  $f(X_1, X_2) = a(X_2) - a(X_1) + \kappa$  a.s.

**Proof of part (2):** This follows from part (1) and Dobrushin's CLT.

**Proof of part (3):** By homogeneity, the structure constants  $d_n(\xi)$  are independent of  $n$ , and they are all equal to  $d(\xi) := \mathbb{E}(|e^{i\xi \Gamma} - 1|^2)^{1/2}$ , where  $\Gamma$  is the balance of a random hexagon at position 3. So  $D_N(\xi) = \sum_{k=3}^N d_k^2(\xi) = (N-3)d^2(\xi)$ .

If  $d(\xi) \neq 0$  for all  $\xi \neq 0$ , then  $D_N(\xi) \rightarrow \infty$  for all  $\xi \neq 0$ ,  $f$  is irreducible by Theorem 3.2. and the LLT follows from Theorem 4.1.

If  $d(\xi) = 0$  for some  $\xi \neq 0$ , then  $D_N(\xi) = 0$  for all  $N$ ,  $\xi$  is in the co-range of  $(X, f)$ , and our reduction lemma says that there exist  $c_n \in \mathbb{R}$  and uniformly bounded measurable  $a_n : \mathfrak{S} \rightarrow \mathbb{R}$  and  $h_n(X_n, X_{n+1})$  such that  $\sum h_n(X_n, X_{n+1})$  converges a.s., and  $f(X_n, X_{n+1}) + a_n(X_n) - a_{n+1}(X_{n+1}) + h_n(X_n, X_{n+1}) + \kappa_n \in \frac{2\pi}{\xi} \mathbb{Z}$  a.s.

Let  $A_n(X_n, X_{n+1}, \dots) := a_n(X_n) + \sum_{k \geq n} h_k(X_k, X_{k+1})$ , then for all  $n$

$$f_n(X_n, X_{n+1}) + A_n(X_n, X_{n+1}, \dots) - A_{n+1}(X_{n+1}, X_{n+2}, \dots) + \kappa_n \in \frac{2\pi}{\xi} \mathbb{Z} \text{ a.s.} \quad (7.1.2)$$

We need to replace  $A_i(X_i, X_{i+1}, \dots)$  by  $a(X_i)$ . This is the purpose of the following proposition, whose proof will complete the proof of the theorem:

**Proposition 7.1** *Let  $X$  be a uniformly elliptic homogeneous Markov chain with state space  $(\mathfrak{S}, \mathcal{B}, \mu)$ , and let  $f : \mathfrak{S} \times \mathfrak{S} \rightarrow \mathbb{R}$  be a measurable function such that  $\text{ess sup } |f(X_1, X_2)| < \infty$ . If there exist measurable functions  $A_n : \mathfrak{S}^{\mathbb{N}} \rightarrow \mathbb{R}$  and  $\kappa_n \in \mathbb{R}$  satisfying (7.1.2), then there exist  $\kappa \in \mathbb{R}$  and a measurable  $a : \mathfrak{S} \rightarrow \mathbb{R}$  such that*

$$f(X_n, X_{n+1}) + a(X_n) - a(X_{n+1}) + \kappa \in \mathbb{Z} \text{ a.s. for all } n.$$

*Proof.* Throughout this proof, let  $\Omega := \mathfrak{S}^{\mathbb{N}}$ , equipped with the  $\sigma$ -algebra  $\mathcal{F}$  generated by the cylinder sets

$$[A_1, \dots, A_n] := \{x \in \mathfrak{S}^{\mathbb{N}} : x_i \in A_i \ (i = 1, \dots, n)\} \ (A_i \in \mathcal{B})$$

and the unique probability measure  $m$  on  $(\Omega, \mathcal{F})$  s.t.

$$m[A_1, \dots, A_n] = \mathbb{P}[X_1 \in A_1, \dots, X_n \in A_n]$$

Let  $\sigma : \Omega \rightarrow \Omega$  denote the left-shift map,  $\sigma[(x_n)_{n \geq 1}] = (x_{n+1})_{n \geq 1}$ . The stationarity of  $X$  translates to the shift invariance of  $m$ :  $m \circ \sigma^{-1} = m$ .

**STEP 1 (Zero-One Law):** Let  $\sigma^{-n}\mathcal{F} := \{\sigma^{-n}(A) : A \in \mathcal{F}\}$ , then for every  $A \in \bigcap_{n \geq 1} \sigma^{-n}\mathcal{F}$ , either  $m(A) = 0$  or  $m(A) = 1$ .

*Proof.* Fix a cylinder  $A := [A_1, \dots, A_\ell]$ .

By uniform ellipticity, for every cylinder  $B = [B_1, \dots, B_n]$ ,

$$m(A \cap \sigma^{-(\ell+1)}B) = m([A_1, \dots, A_\ell, *, B_1, \dots, B_n]) \geq \varepsilon_0 m(A)m(B).$$

Applying this to cylinders  $\underbrace{[\Omega, \dots, \Omega]}_{k \text{ times}}, B_1, \dots, B_n$  we find that

$$m(A \cap \sigma^{-(\ell+k)}[B_1, \dots, B_n]) \geq \varepsilon_0 m(A)m(B) \text{ for all } k \geq 1.$$

By the monotone class theorem,

$$m(A \cap \sigma^{-(\ell+k)}E) \geq \varepsilon_0 m(A)m(E) \text{ for every } \mathcal{F}\text{-measurable } E \text{ and } k \geq 1. \quad (7.1.3)$$

Suppose  $E \in \bigcap_{k \geq 1} \sigma^{-n}\mathcal{F}$ , and let  $A$  be an arbitrary cylinder of length  $\ell$ . By the assumption on  $E$ ,  $E = \sigma^{-n}E_n$  with  $E_n \in \mathcal{F}$  and  $n > \ell$ . So

$$m(A \cap E) = m(E \cap \sigma^{-n}E_n) \geq \varepsilon_0 m(A)m(E_n) = \varepsilon_0 m(A)m(E).$$

We see that  $\frac{m(A \cap E)}{m(A)} \geq \varepsilon_0 m(E)$  for all cylinders  $A$ , whence

$$\mathbb{E}(1_E | X_1, \dots, X_\ell) \geq \varepsilon_0 m(E) \text{ for all } \ell.$$

By the martingale convergence theorem,  $1_E \geq \varepsilon_0 m(E)$  a.e., whence  $m(E) = 0$  or 1.

**STEP 2:** Identify  $f$  with a function  $f : \Omega \rightarrow \mathbb{R}$  s.t.  $f[(x_i)_{i \geq 1}] = f(x_1, x_2)$ . Then there exist  $A : \Omega \rightarrow \mathbb{R}$  measurable and  $\kappa \in \mathbb{R}$  s.t.  $f + A - A \circ \sigma + \kappa \in \mathbb{Z}$  almost surely.

*Proof.* The assumptions of the proposition say that there exist  $A_n : \Omega \rightarrow \mathbb{R}$  measurable and  $\kappa_n \in \mathbb{R}$  s.t.

$$f \circ \sigma^n + A_n \circ \sigma^n - A_{n+1} \circ \sigma^{n+1} + \kappa_n \in \mathbb{Z} \text{ } m\text{-a.e. for every } n.$$

Let  $w_n := e^{2\pi i A_n}$ , then  $e^{2\pi i f \circ \sigma^n} \frac{w_n \circ \sigma^n}{w_{n+1} \circ \sigma^{n+1}} = 1$   $m$ -a.s. Since  $m \circ \sigma^{-1} = m$  we have  $e^{2\pi i f} \frac{w_n}{w_{n+1} \circ \sigma} = 1$  a.s. for all  $n$ . This gives the chain of identities

$$w_n = e^{-2\pi i f} w_{n+1} \circ \sigma = e^{-2\pi i (f + f \circ \sigma)} w_{n+2} \circ \sigma^2 = \dots = e^{-2\pi i \sum_{j=0}^{k-1} f \circ \sigma^k} w_{n+k} \circ \sigma^k.$$

It follows that  $w_n/w_{n+1} = (w_{n+k}/w_{n+k+1}) \circ \sigma^k$  for all  $k$ . Hence  $w_n/w_{n+1}$  is  $\sigma^{-k}\mathcal{F}$ -measurable for all  $k$ . By the zero-one law,  $w_n/w_{n+1}$  is constant almost surely. In particular, there exists a constant  $c$  such that  $A_2 - A_1 \in c + \mathbb{Z}$   $m$ -a.e., and the step follows with  $A := A_1$  and  $\kappa := \kappa_1 + c$ .

**STEP 3:** *There exists  $a : \Omega \rightarrow \mathbb{R}$  constant on cylinders of length one such that  $f + a - a \circ \sigma + \kappa \in \mathbb{Z}$   $m$ -a.e.*

*Proof.* Let  $L : L^1(\Omega) \rightarrow L^1(\Omega)$  denote the **transfer operator** of  $\sigma : \Omega \rightarrow \Omega$ , which describes the action of  $\sigma$  on mass densities on  $\Omega$ :  $\sigma_*[\varphi dm] = L\varphi dm$ . Formally,  $L\varphi := \frac{dm_\varphi \circ \sigma^{-1}}{dm}$ , where  $m_\varphi := \varphi dm$ . We will need the following (standard) facts:

(a) If  $\varphi$  depends only on the first  $m$ -coordinates, then  $L\varphi$  depends only on the first  $(m-1) \vee 1$ -coordinates. Specifically,  $(L\varphi)[(y_i)_{i \geq 1}] = \Phi(y_1, \dots, y_{m-1})$  where

$$\Phi(y_1, \dots, y_{m-1}) := \mathbb{E}[\varphi(X_1, \dots, X_m) | X_i = y_i \ (1 \leq i \leq m-1)];$$

(b)  $L\varphi$  is characterized by the condition  $\int \psi L\varphi dm = \int \psi \circ \sigma \varphi dm \ \forall \psi \in L^\infty(\mathcal{S})$ ;

(c)  $L(\psi \circ \sigma) = \psi L\varphi \ \forall \varphi \in L^1, \psi \in L^\infty$ ;

(d)  $L1 = 1$ ;

(e)  $\forall \varphi \in L^\infty, L^n \varphi \xrightarrow[n \rightarrow \infty]{} \int \varphi dm$  in  $L^1$ .

Part (b) is standard. Parts (c) and (d) follow from (b) and the  $\sigma$ -invariance of  $m$ . Part (a) follows from (b), and the identity

$$\begin{aligned} \int \psi L\varphi dm &= \int \psi dm_\varphi \circ \sigma^{-1} = \int \psi \circ \sigma \varphi dm = \mathbb{E}[\psi(X_2, X_3, \dots) \varphi(X_1, \dots, X_m)] \\ &= \mathbb{E}(\psi(X_2, X_3, \dots) \mathbb{E}[\varphi(X_1, \dots, X_m) | X_2, X_3, \dots]) \stackrel{!}{=} \mathbb{E}(\psi(X_2, X_3, \dots) \mathbb{E}[\varphi | X_2, \dots, X_m]) \\ &= \int \psi \Phi dm \end{aligned}$$

where  $\stackrel{!}{=}$  is because of the Markov property. To see part (e) note that it is enough to consider  $\varphi \in L^\infty$  such that  $\int \varphi dm = 0$  (otherwise work with  $\varphi - \int \varphi dm$ ). For such functions,

$$\begin{aligned} \|L^n \varphi\|_1 &= \int \text{sgn}(L^n \varphi) L^n \varphi dm = \int \text{sgn}(L^n \varphi) \circ \sigma^n \cdot \varphi dm \\ &= \int \text{sgn}(L^n \varphi) \circ \sigma^n \mathbb{E}(\varphi | \sigma^{-n} \mathcal{F}) dm \leq \int |\mathbb{E}(\varphi | \sigma^{-n} \mathcal{F})| dm \end{aligned}$$

The integrand is uniformly bounded (by  $\|\varphi\|_\infty$ ), and it converges pointwise to  $\mathbb{E}(\varphi | \cap \sigma^{-n} \mathcal{F}) = \mathbb{E}(\varphi | \{\emptyset, \Omega\}) = \mathbb{E}(\varphi) = 0$ .

Let  $w := e^{2\pi i A}$  where  $A : \Omega \rightarrow \mathbb{R}$  is as in step 2, and assume w.l.o.g. that  $\kappa = 0$  (else absorb it into  $f$ ). Set  $S_n = f + f \circ \sigma + \dots + f \circ \sigma^{n-1}$ , then  $e^{-2\pi i f} = w/w \circ \sigma$ , whence  $e^{-2\pi i S_n} = w/w \circ \sigma^n$ . By (c), for all  $\varphi \in L^1(\Omega)$ ,

$$w L^n (e^{-2\pi i S_n} \varphi) = L^n (e^{-2\pi i S_n} w \circ \sigma^n \varphi) = L^n (w \varphi) \xrightarrow[n \rightarrow \infty]{L^1} \int w \varphi dm.$$

Since  $|w| = 1$  a.e.,  $\exists m \geq 2$  and  $\exists \varphi = \varphi(x_1, \dots, x_m)$  bounded measurable so that  $\int w \varphi dm \neq 0$ . For this  $\varphi$ , we have

$$w^{-1} = L^1 \text{-} \lim_{n \rightarrow \infty} \frac{L^n(e^{-2\pi i S_n} \varphi)}{\int w \varphi dm}.$$

We claim that the right-hand-side depends only on the first coordinate. This is because  $e^{-2\pi i f} \varphi$  is function of the first  $m$  coordinates, whence by (a),  $L(e^{-2\pi i f} \varphi)$  is a function of the first  $(m-1) \vee 1$  coordinates. Applying this argument again we find that  $L^2(e^{-2\pi i S_2} \varphi) = L[e^{-2\pi i f} L(e^{-2\pi i f} \varphi)]$  is a function of the first  $(m-2) \vee 1$  coordinates. Continuing by induction, we find that  $L^n(e^{-2\pi i S_n} \varphi)$  is a function of  $(m-n) \vee 1$ -coordinates, and eventually of the first coordinate only.

So  $w^{-1}$  is an  $L^1$ -limit of a functions of the first coordinate. Therefore we can write  $w[(x_i)_{i \geq 1}] = \exp[2\pi i a(x_1)]$  a.e., where  $a : \mathfrak{S} \rightarrow \mathbb{R}$  is measurable. By construction  $e^{2\pi i f} w / w \circ \sigma = 1$ , so  $f(X_1, X_2) + a(X_1) - a(X_2) \in \mathbb{Z}$  almost surely. By stationarity,  $f(X_n, X_{n+1}) + a(X_n) - a(X_{n+1}) \in \mathbb{Z}$  almost surely for all  $n$ .  $\square$

We now determine the domain of the rate functions for large deviations. We note that the results of Chapter 6 concern  $\mathbb{P}[S_N \geq zV_N] = \mathbb{P}[S_N \geq z\sigma^2(1+o(1))N]$ , while in large deviation literature it is common to use the normalization  $\mathbb{P}[S_N \geq zN]$ . To simplify the comparison with other results we will assume till the end of this section that  $\sigma^2 = 1$  which can always be achieved by scaling  $f$ .

Let  $\mathcal{S}_N = \text{ess sup } S_N$ . Using the stationarity of  $\{X_n\}$  and the homogeneity of  $f$  it is not difficult to see that  $\mathcal{S}_{n+m} \leq \mathcal{S}_n + \mathcal{S}_m$ , and therefore the limit

$$\mathfrak{s}_+ = \lim_{N \rightarrow \infty} \frac{\text{ess sup } S_N}{N} = \lim_{N \rightarrow \infty} \frac{\mathcal{S}_N}{N}$$

exists. Repeating the same argument for  $(-f)$  gives that

$$\mathfrak{s}_- = \lim_{N \rightarrow \infty} \frac{\text{ess inf } S_N}{N}$$

exists as well.

Recall the notation for large deviation thresholds  $\mathfrak{c}_-, \mathfrak{c}_+$  introduced in §6.4.

**Theorem 7.2.** *Let  $f$  be an a.s. uniformly bounded homogenous additive functional on a uniformly elliptic homogeneous Markov chain, and assume  $f$  has zero mean and asymptotic variance  $\sigma^2 = 1$ . Then  $\mathfrak{c}_+ = \mathfrak{s}_+$  and  $\mathfrak{c}_- = \mathfrak{s}_-$ .*

*Proof.* We prove the first identity, the second one is similar.

First, for any  $\varepsilon > 0$ ,  $\mathbb{P}[S_N \geq (\mathfrak{s}_+ + \varepsilon)N] = 0$  for sufficiently large  $N$ , whence by Theorem 6.7,  $\mathfrak{c}_+ \leq \mathfrak{s}_+$ .

Let  $K := \text{ess sup } |f|$ . For every  $\varepsilon > 0$ , for all sufficiently large  $M$ ,

$$\delta_M := \mathbb{P}[S_M \geq (\mathfrak{s}_+ - \varepsilon)M] > 0.$$

Let  $\sigma(X_i, \dots, X_j)$  denote the  $\sigma$ -field generated by  $X_i, \dots, X_j$ . By uniform ellipticity, if  $E \in \sigma(X_1, \dots, X_{M+1})$  and  $F \in \sigma(X_{M+3}, \dots, X_{2M+3})$ , then  $\mathbb{P}[E \cap F] \geq \varepsilon_0 \mathbb{P}(E) \mathbb{P}(F)$  (see (7.1.3)). Consequently,

$$\begin{aligned}
& \mathbb{P}[S_{2(M+2)} \geq 2(\mathfrak{s}_+ - \varepsilon)M - 2K] \\
& \geq \mathbb{P} \left[ \sum_{k=1}^M f_k(X_k, X_{k+1}) \geq M(\mathfrak{s}_+ - \varepsilon), \sum_{k=M+3}^{2M+2} f_k(X_k, X_{k+1}) \geq M(\mathfrak{s}_+ - \varepsilon) \right] \\
& \geq \varepsilon_0 \mathbb{P} \left[ \sum_{k=1}^M f_k(X_k, X_{k+1}) \geq M(\mathfrak{s}_+ - \varepsilon) \right] \mathbb{P} \left[ \sum_{k=M+3}^{2M+2} f_k(X_k, X_{k+1}) \geq M(\mathfrak{s}_+ - \varepsilon) \right].
\end{aligned}$$

Thus by stationarity,  $\mathbb{P}[S_{2(M+2)} \geq 2(\mathfrak{s}_+ - \varepsilon)M - 2K] \geq \varepsilon_0 \delta_M^2$ . Applying this argument repeatedly, we find that for each  $\ell$ ,

$$\mathbb{P}[S_{(M+2)\ell} \geq ((\mathfrak{s}_+ - \varepsilon)M - 2K)\ell] \geq (\varepsilon_0 \delta_M^2)^\ell.$$

Now Corollary 6.26 tells us that for all sufficiently large  $M$ ,  $\mathfrak{c}_+ \geq \frac{(\mathfrak{s}_+ - \varepsilon)M - 2K}{M+2}$ . Letting  $M \rightarrow \infty$  we obtain  $\mathfrak{c}_+ \geq \mathfrak{s}_+ - \varepsilon$ . Since  $\varepsilon$  is arbitrary,  $\mathfrak{c}_+ \geq \mathfrak{s}_+$ .  $\square$

## 7.2 Perturbations of homogeneous chains

Let  $(X, f)$  be a bounded homogenous additive functional on a uniformly elliptic Markov chain with stationary measure  $\mu$  and transition probability  $\pi(x, dy) = p(x, y)\mu(dy)$ . We consider *non-homogeneous* perturbations  $(\tilde{X}, \tilde{f})$  of the form

$$\tilde{f}_n(x, y) = f(x, y) + g_n(x, y), \tilde{\pi}_n(x, dy) = \tilde{p}_n(x, y)\mu(dy).$$

We assume that the strength of the perturbation decays at infinity. Namely for each  $\varepsilon > 0$  there is  $n_0$  such that for  $n \geq n_0$

$$\|g_n\|_\infty \leq \varepsilon \quad \text{and} \quad 1 - \varepsilon \leq \frac{\tilde{p}_n(x, y)}{p(x, y)} \leq 1 + \varepsilon.$$

**Theorem 7.3.** *If the additive functional  $g$  is center tight on  $X$ , then  $G_{ess}(\tilde{X}, \tilde{f}) = G_{ess}(X, f)$ . If  $g$  is not center tight then  $G_{ess}(\tilde{X}, \tilde{f}) = \mathbb{R}$ .*

*Proof.* We note that it suffices to prove the result in the case  $\tilde{p}_n \equiv p$ . Indeed by our assumptions,

$$\frac{1}{2} \leq \frac{\tilde{p}_n(x, y)}{p(x, y)} \leq 2 \tag{7.2.1}$$

if  $n$  is sufficiently large. Since discarding a finite number of terms does not change the essential range (since any functional vanishing for large  $n$  is center tight) we may assume that (7.2.1) holds for all  $n$ . Now Example 3.8 shows that the essential range of the functionals defined via  $p$  and via  $\tilde{p}_n$  are the same. Thus we assume henceforth that  $\tilde{p}_n \equiv p$  for all  $n$ .

If  $g$  is center tight then the essential ranges of  $f$  and  $\tilde{f}$  are the same, so we shall assume that  $g$  is not center tight, and prove that  $D_N(\xi, \tilde{f}) \rightarrow \infty$  for every  $\xi \neq 0$ . Let  $\mathfrak{d} := d_n(\xi, f)$  (the RHS does not depend on  $n$  by stationarity).

Suppose first that  $\mathfrak{d} \neq 0$ . By Lemma 1.15(2) we have

$$\mathfrak{d}^2 = d_n^2(\xi, f) \leq 8 \left[ d_n(\xi, \tilde{f})^2 + d_n(\xi, g)^2 \right].$$

Next, the assumption  $\|g_n\|_\infty \xrightarrow{n \rightarrow \infty} 0$  implies that  $d_n^2(\xi, g) \xrightarrow{n \rightarrow \infty} 0$ . Accordingly  $d_n(\tilde{f}, \xi)^2 \geq \frac{\mathfrak{d}^2}{10}$  for all  $n$  large enough, so that  $D_N(\xi, \tilde{f}) \rightarrow \infty$  as needed.

Next assume  $\mathfrak{d} = 0$ . In this case for any hexagon  $P_n$  we have  $e^{i\xi\Gamma(f, P_n)} = 1$ , where  $\Gamma(f, \cdot)$  denotes the balance for the additive functional  $f$ . Hence  $e^{i\xi\Gamma(\tilde{f}, \cdot)} = e^{i\xi\Gamma(g, \cdot)}$ , and so

$$d_n(\xi, \tilde{f}) = d_n(\xi, g).$$

Let  $\gamma_N := \max_{n \geq N} \text{ess sup} |g_n|$ , and fix  $\tau_0 > 0$  such that  $|e^{it} - 1|^2 \geq \frac{1}{2}t^2$  for all  $|t| < \tau_0$ . If  $n \geq N$  and  $0 < |\eta| \leq \tau_0(6\gamma_N)^{-1}$ , then (3.3.7) tells us that

$$d_n^2(\eta, g) \geq \frac{\eta^2}{2} u_n^2(g) \text{ for all } n > N + 3.$$

By assumption,  $g$  is not center-tight, so  $\sum u_n^2(g) = \infty$ . It follows that  $D_N(\eta, g) \rightarrow \infty$  for all  $0 < |\eta| \leq \tau_0(6\gamma_N)^{-1}$ .

By assumption,  $\gamma_N \rightarrow 0$ , so  $D_N(\eta, g) \rightarrow \infty$  for all  $\eta \neq 0$ . It follows that the co-range of  $g$  equals  $\{0\}$ , and the essential range of  $g$  equals  $\mathbb{R}$ .  $\square$

Next we discuss the large deviation thresholds for  $\tilde{f}$ .

**Theorem 7.4.** (a) *If  $f$  is not a coboundary then  $\mathfrak{c}_\pm(\tilde{f}) = \mathfrak{c}_\pm(f) = \mathfrak{s}_\pm(f)$ .*

(b) *If  $f$  is a homogeneous gradient,  $\mathbb{E}(g_n) = 0$  for all  $n$ , and  $g$  is not center tight, then  $\mathfrak{c}_+(\tilde{f}) = +\infty$ ,  $\mathfrak{c}_-(\tilde{f}) = -\infty$ .*

*Proof.* The proof of part (a) is very similar to the proof of Lemma 6.28 so we omit it.

In the proof of part (b) we may assume that  $f = 0$  since adding a homogeneous gradient does not change the large deviation threshold. In particular in the rest of the proof we will abbreviate  $S_N = S_N(g)$ ,  $S_{n_1, n_2} = S_{n_1, n_2}(g) = \sum_{k=n_1}^{n_2-1} g_k(X_k, X_{k+1})$ ,  $V_N = \text{Var}(S_N(g))$ . Since  $g$  is not center tight,  $V_N \rightarrow \infty$ .

Assume without loss of generality that  $\text{ess sup} |g| \leq \frac{1}{2}$ , then  $\text{Var}[g_k(X_k, X_{k+1})] \leq 1$  for all  $k$ . Divide the interval  $[0, N]$  into blocks

$$\left[ n_1, n_2 \right] \cup \{n_2 + 1\} \cup \left[ n_3, n_4 \right] \cup \cdots \cup \left[ n_k, n_{k+1} \right] \cup \{n_{k+1} + 1\} \cup \left[ n_{k+2}, N \right]$$

where  $n_i$  is increasing,  $1 \leq \text{Var}(S_{n_j, n_{j+1}}) \leq 2$  for  $j \leq k + 1$ , and  $\text{Var}(S_{n_{k+2}, N}) \leq 1$ .

Since  $\|g_n\|_\infty \rightarrow 0$ ,  $\min\{n_{j+1} - n_j : \ell \leq j \leq k\} \xrightarrow{\ell \rightarrow \infty} \infty$ . Also, the analysis of §6.3.4 shows that

$$\frac{1}{V_N} \sum_j \text{Var}(S_{n_j, n_{j+1}}) \rightarrow 1.$$

In particular, the number of blocks  $\beta_N$ , is between  $V_N/2$  and  $3V_N/2$ .

Let  $M_j = \max_{n_j \leq l \leq n_{j+1}} \|g_l\|_\infty$ . Note that  $M_j \rightarrow 0$ . Therefore applying Dobrushin CLT to the array  $\{g_l/M_j\}_{n_j \leq l \leq n_{j+1}}$  we conclude that  $S_{n_j, n_{j+1}} / \sqrt{\text{Var}(S_{n_j, n_{j+1}})}$  is asymptotically normal. In particular, for each  $z > 0$  there exists  $\eta = \eta(z) > 0$  such that for  $j$  large enough and all  $x_j \in \mathfrak{S}_{n_j}$

$$\mathbb{P}_{x_j}(S_{n_j, n_{j+1}} \geq 3z) \geq \eta. \quad (7.2.2)$$

A uniform ellipticity argument similar to the one we used in the proof of Theorem 7.2 gives

$$\mathbb{P}_{x_j}(S_N \geq \beta_N z) \geq c\eta^{\beta_N},$$

where  $c$  incorporates the contribution of blocks (with small  $j$ ) where (7.2.2) fails.

Now Corollary 6.26 implies that  $\mathfrak{c}^+ \geq z$ . Since  $z$  is arbitrary,  $\mathfrak{c}^+ = +\infty$ . A similar argument shows that  $\mathfrak{c}_-(g) = -\infty$ .  $\square$

### 7.3 Small additive functionals.

The perturbations of  $f \equiv 0$  were analyzed in the previous section, however, since this case is of independent interest it makes sense to summarize the results obtained for this particular case.

**Theorem 7.5.** *Let  $g$  be a uniformly bounded additive functional of uniformly elliptic Markov chain. Suppose that  $\mathbb{E}(g_n) = 0$  and that  $\lim_{n \rightarrow \infty} \|g_n\|_\infty = 0$ . Then*

*either  $g$  is center tight in which case  $\sum_{n=1}^{\infty} g_n$  converges almost surely*

*or  $g$  is not center tight in which case  $S_N(g)$  satisfies non lattice LLT (4.1.1) and  $\mathfrak{c}_\pm(g) = \pm\infty$ .*

*Proof.* The non-center tight case was analyzed in §7.2. In the center tight case the results of Chapter 2 tell us that  $g$  can be decomposed as

$$g_n(x, y) = c_n + a_{n+1}(y) - a_n(x) + h_n(x, y) \quad \text{where} \quad \sum_n \text{Var}(h_n) < \infty.$$

Changing  $a_n$  if necessary we may assume that  $\mathbb{E}(a_n) = 0$  in which case

$$\mathbb{E}(g_n) = 0 = \mathbb{E}(h_n + c_n).$$

Therefore the additive functional  $\tilde{h} = h + c$  has zero mean and finite variance. Hence by Theorem 2.4  $\sum_{n=1}^{\infty} (h_n + c_n)$  converges almost surely. In summary  $S_N(g) - a_N + a_1$  converges almost surely, and hence  $S_N(g) - a_N$  converges almost surely. On the other hand equation (2.2.4) shows that  $\lim_{N \rightarrow \infty} a_N = 0$  completing the proof.  $\square$

The following result which a direct consequence of Theorem 7.5 shows that for small additive functionals a vague limit of the local distribution of  $S_N$  always exists.

**Corollary 7.2** *Let  $g$  satisfy the assumptions of Theorem 7.5. Then either and  $S_N$  converges a.s. to some random variable  $\mathcal{S}$  in which case for each continuous compactly supported function  $\phi$*

$$\lim_{N \rightarrow \infty} \mathbb{E}(\phi(S_N)) = \mathbb{E}(\phi(\mathcal{S}))$$

*or  $S_N$  satisfies a non-lattice LLT. That is, for each continuous compactly supported function  $\phi$  for each sequence  $z_N$  such that the limit  $z = \lim_{N \rightarrow \infty} \frac{z_N}{\sqrt{V_N}}$  exists we have*

$$\lim_{N \rightarrow \infty} V_N \mathbb{E}(\phi(S_N)) = \mathbb{E}(\phi(\mathcal{S})) = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(s) ds.$$

## 7.4 Equicontinuous additive functionals

In this section we examine the consequences of *topological* assumptions on  $f$  and  $X$ . Specifically we will say that  $(X, f)$  is **equicontinuous** if

- (T)  $(\mathfrak{S}_n, \mathcal{B}_n, \mu_n)$  are complete separable metric spaces,  $\mathcal{B}_n$  are the Borel  $\sigma$ -algebras, and  $\mu_n$  are Borel probability measures;
- (S) for every  $\varepsilon > 0$  there exists  $\delta > 0$  such that for all  $x_n \in \mathfrak{S}_n$  and  $n \geq 1$ ,  $\mu_n[B(x_n, \varepsilon)] > \delta$ . Here  $B(x, \varepsilon) := \{y \in \mathfrak{S}_n : \text{dist}(x, y) < \varepsilon\}$ .
- (U) for every  $\varepsilon > 0$  there exists  $\delta > 0$  such that for all  $n \geq 1$  and  $x_n, y_n \in \mathfrak{S}_n$ ,  $\text{dist}(x_n, y_n) < \delta \Rightarrow |f_n(x_n) - f_n(y_n)| < \varepsilon$ .

### 7.4.1 Range.

**Theorem 7.6.** Suppose  $(X, f)$  is equicontinuous and a.s. uniformly bounded. Assume in addition the following:

- (a) **One-step ellipticity condition:**  $\exists \varepsilon_0$  s.t. for every  $n$ ,  $\pi_n(x, dy) = p_n(x, y) \mu_{n+1}(dy)$  where  $\varepsilon_0 \leq p_n(x, y) \leq \varepsilon_0^{-1}$ .
- (b)  $\mathfrak{S}_n$  are all connected.

Then  $f$  is either irreducible with algebraic range  $\mathbb{R}$ , or it is center tight.

*Proof.* Choose  $c_1 > 0$  such that  $|e^{i\theta} - 1|^2 = 4 \sin^2\left(\frac{\theta}{2}\right) \geq c_1 \theta^2$  for all  $|\theta| \leq 0.1$ . We fix  $\xi \neq 0$ , and consider the following two cases:

- (I)  $\exists N_0$  such that  $|\xi \Gamma(P)| < 0.1$  for every position  $n$  hexagon  $P$ , for each  $n \geq N_0$ .
- (II)  $\exists n_k \uparrow \infty$  and  $\exists$  position  $n_k$  hexagons  $P_{n_k}$  such that  $|\xi \Gamma(P_{n_k})| \geq 0.1$ .

In case (I), for all  $n \geq N_0$ ,  $d_n^2(\xi) = \mathbb{E}(|e^{i\xi \Gamma} - 1|^2) \geq c_1 \mathbb{E}(\Gamma^2) \equiv c_1 u_n^2$ . So either  $\sum u_n^2 = \infty$  and then  $\sum d_n^2(\xi) = \infty$  for all  $\xi \neq 0$ , and  $f$  is irreducible with essential range  $\mathbb{R}$ ; or  $\sum u_n^2 < \infty$  and then  $f$  is center-tight by Corollary 2.8.

In case (II), for every  $k$  there is a position  $n_k$  hexagon  $P_{n_k}$  with  $|\xi \Gamma(P_{n_k})| \geq 0.1$ . There is also a position  $n_k$  hexagon  $P'_{n_k}$  with balance zero (such hexagons always exist because we can take  $y_{n_k-1} = x_{n_k-1}$ ,  $y_{n_k} = x_{n_k}$ ). We would like to apply the intermediate value to deduce the existence of a position  $n_k$  hexagon  $\bar{P}_{n_k}$  such that  $0.05 < \xi \Gamma(\bar{P}_{n_k}) < 0.1$ . To do this we note that:

- Because of the one-step ellipticity condition, the space of position  $n_k$  hexagons is homeomorphic to  $\mathfrak{S}_{n_k-2} \times \mathfrak{S}_{n_k-1}^2 \times \mathfrak{S}_{n_k}^2 \times \mathfrak{S}_{n_k}$ .
- The product of connected topological spaces is connected.
- Real-valued continuous functions on connected topological spaces satisfy the intermediate value theorem.
- The balance of hexagon depends continuously on the hexagon.

So  $\bar{P}_{n_k}$  exists. Necessarily,  $|e^{i\xi \Gamma(\bar{P}_{n_k})} - 1| \geq c_1 \xi^2 \Gamma^2(\bar{P}_{n_k}) =: c_2$ .

Write  $\bar{P}_{n_k}$  in coordinates:  $\bar{P}_{n_k} := \left( \bar{x}_{n_k-2}; \frac{\bar{x}_{n_k-1}}{\bar{y}_{n_k-1}}; \frac{\bar{x}_{n_k}}{\bar{y}_{n_k}}; \bar{y}_{n_k+1} \right)$ . By the equicontinuity of  $f$ ,  $\exists \varepsilon > 0$  such that  $|e^{i\xi \Gamma(P)} - 1| > \frac{1}{2} c_2$  for every hexagon  $P$  whose coordinates are in the  $\varepsilon$ -neighborhood of the coordinates of  $\bar{P}_{n_k}$ . By the equicontinuity of  $\mu_n$  and the one-step ellipticity condition, this

collection of hexagons  $P$  have hexagon measure  $\geq \delta$  for some  $\delta > 0$  independent of  $k$ . So  $d_{n_k}^2(\xi) \geq \frac{1}{2}c_2\delta$ .

Summing over all  $k$ , we find that  $\sum d_{n_k}^2(\xi) = \infty$ . Since  $\xi \neq 0$  was arbitrary,  $(X, f)$  has essential range  $\mathbb{R}$ .  $\square$

#### 7.4.2 Large deviation threshold.

**Lemma 7.3** Suppose that  $\mathfrak{S}_n$  are metric spaces,  $f_n$  are equicontinuous, and for each  $\varepsilon > 0$  there exists  $\delta > 0$  such that if  $p_n(x, y) > 0$  then

$$\pi_n(x, B(y, \varepsilon)) > \delta. \quad (7.4.1)$$

Suppose that  $V_N > cN$  and that

$$\limsup_{N \rightarrow \infty} \frac{\inf_{x_1, \dots, x_{N+1}} \sum_{j=1}^N f_j(x_j, x_{j+1})}{V_N} < z < \liminf_{N \rightarrow \infty} \frac{\sup_{x_1, \dots, x_{N+1}} \sum_{j=1}^N f_j(x_j, x_{j+1})}{V_N}.$$

Then  $z \in \mathcal{C}$ .

Note that assumption (7.4.1) is satisfied whenever  $X$  satisfies (S) and the one step ellipticity condition. We also remark that Example 6.21 shows that equicontinuity assumption on  $f$  is essential.

*Proof.* Fix  $N$ . Consider first the case where  $z > \frac{\mathbb{E}(S_N)}{V_N}$ . By assumption there is an  $\varepsilon$  such that for all sufficiently large  $n$  there is a sequence  $\bar{x}_1, \dots, \bar{x}_{N+1}$  such that

$$\sum_{j=1}^N f_j(\bar{x}_j, \bar{x}_{j+1}) \geq (z + \varepsilon)V_N.$$

By ellipticity, for each  $x \in \mathfrak{S}_1$  there a sequence  $\tilde{x}_1, \tilde{x}_2 \dots \tilde{x}_{N+1}$  such that  $\tilde{x}_1 = x$  and

$$\sum_{j=1}^N f_j(\tilde{x}_j, \tilde{x}_{j+1}) \geq (z + \varepsilon)V_N - 4K, \text{ where } K := \text{ess sup } |f|.$$

(In fact one can take  $\tilde{x}_j = \bar{x}_j$  for  $j \geq 3$ .) By uniform continuity of  $f_j$  and the fact that  $V_N$  grows linearly, there is  $r$  such that if  $X_j \in B(\tilde{x}_j, r)$  for  $j \leq N+1$  then

$$\sum_{j=1}^N f_j(X_j, X_{j+1}) \geq (z + \varepsilon/2)V_N - 4K.$$

By (7.4.1) there is  $\delta > 0$  such that  $\mathbb{P}_x(X_j \in B(\tilde{x}_j, r)) \geq \delta^N$ . Hence

$$\mathbb{P}(S_N \geq (z + \varepsilon/3)V_N) \geq \delta^N.$$

Next, by the CLT and the assumption that  $z > \frac{\mathbb{E}(S_N)}{V_N}$ , if  $\varepsilon$  is small enough and  $N$  is large enough, then<sup>1</sup>  $\mathbb{P}(S_N \leq (z - \varepsilon)V_N) \geq \delta^N$ . Now Theorem 6.7 shows that  $z \in \mathcal{C}$ .

The case  $z < \frac{\mathbb{E}(S_N)}{V_N}$  is analyzed similarly now using the estimate

$$\limsup_{N \rightarrow \infty} \frac{\inf_{x_1, \dots, x_{N+1}} \sum_{j=1}^N f_j(x_j, x_{j+1})}{V_N} \leq z - \varepsilon.$$

□

**Corollary 7.4** *Under the assumptions of Lemma 7.3 if*

$$\mathfrak{z}^- = \lim_{N \rightarrow \infty} \frac{\inf_{x_1, \dots, x_{N+1}} \sum_{j=1}^N f_j(x_j, x_{j+1})}{V_N} \quad \text{exists then} \quad \mathfrak{z}^- = \mathfrak{c}^-,$$

$$\mathfrak{z}^+ = \lim_{N \rightarrow \infty} \frac{\sup_{x_1, \dots, x_{N+1}} \sum_{j=1}^N f_j(x_j, x_{j+1})}{V_N} \quad \text{exists then} \quad \mathfrak{z}^+ = \mathfrak{c}^+.$$

*Proof.* We will prove the second statement, the first one is similar.  $\mathfrak{z}^+ \leq \mathfrak{c}^+$  by Lemma 7.3. On the other hand if  $z > \mathfrak{z}^+$  then for large  $N$ ,  $\mathbb{P}(S_N > V_N z) = 0$ . Hence  $\mathfrak{c}^+ \leq \mathfrak{z}^+$ . □

We now restate the result of the last corollary in a slightly different way under an extra assumption. Namely, we suppose that

$$\mathfrak{S}_n \text{ are compact \& } \forall x_n, x_{n+1} : p_n(x_n, x_{n+1}) > 0 \quad (7.4.2)$$

**Definition 7.5** *Let  $\mathcal{M}_N$  denote the space of sequences  $\mathbf{x} \in \prod_n \mathfrak{S}_n$  such that if  $\mathbf{y}_n = \mathbf{x}_n$  for  $n \geq N+1$  then*

$$\sum_{n=1}^N f(\mathbf{y}_n, \mathbf{y}_{n+1}) \geq \sum_{n=1}^N f(\mathbf{x}_n, \mathbf{x}_{n+1}).$$

*Denote  $\mathcal{M} = \bigcap_{N=1}^{\infty} \mathcal{M}_N$ . The elements of  $\mathcal{M}$  will be called **minimizers**.*

The properties of  $\mathcal{M}_N$  are summarized below.

**Lemma 7.6** *Suppose that (7.4.2) holds and that  $f_n : \mathfrak{S}_n \rightarrow [-K, K]$  are continuous. Then*

- (a)  $\mathcal{M}_N$  are closed sets.
- (b) If  $N \geq M$  then  $\mathcal{M}_N \subset \mathcal{M}_M$ .
- (c)  $\mathcal{M}$  is non empty.

(d) If  $\sum_{n=1}^N f(\mathbf{x}_n, \mathbf{x}_{n+1}) = \inf_{\mathbf{y}} \sum_{n=1}^N f(\mathbf{y}_n, \mathbf{y}_{n+1})$  then  $\mathbf{x} \in \mathcal{M}_N$ .

(e) If  $\mathbf{x} \in \mathcal{M}_N$  then  $\sum_{n=1}^N f(\mathbf{x}_n, \mathbf{x}_{n+1}) \leq \inf_{\mathbf{y}} \sum_{n=1}^N f(\mathbf{y}_n, \mathbf{y}_{n+1}) + 2K$ .

<sup>1</sup> Alternatively, combining Theorem 6.2(1)) and Theorem 6.6 (applied to  $-S_N$ ) we get that  $\mathbb{P}\left(S_N \leq -\frac{\mathbb{E}(S_N)}{V_N} - \varepsilon\right) \geq \delta^N$  provided that  $\varepsilon$  is small enough and  $\delta$  is close to 1.

*Proof.* (a) If  $\mathcal{M}_N \ni \mathbf{x}^j \xrightarrow{j \rightarrow \infty} \mathbf{x} \notin \mathcal{M}_N$ , then there would exist  $\bar{\mathbf{x}}$  such that  $\bar{\mathbf{x}}_n = \mathbf{x}_n$  for  $n \geq N+1$ , and

$$\sum_{n=1}^N f(\bar{\mathbf{x}}_n, \bar{\mathbf{x}}_{n+1}) < \sum_{n=1}^N f(\mathbf{x}_n, \mathbf{x}_{n+1}).$$

Let  $\mathbf{y}^j$  be the sequence such that  $\mathbf{y}_n^j = \mathbf{x}_n^j$  for  $n \geq N+1$ ,  $\mathbf{y}_n^j = \bar{\mathbf{x}}_n$  for  $n \leq N$ . By the continuity of  $f$ ,

$$\sum_{n=1}^N f(\mathbf{y}_n^j, \mathbf{y}_{n+1}^j) < \sum_{n=1}^N f(\mathbf{x}_n^j, \mathbf{x}_{n+1}^j)$$

for large  $j$  contradicting,  $\mathbf{x}^j \in \mathcal{M}_N$ .

Next let  $\mathbf{x} \in \mathcal{M}_N$  and  $\mathbf{x}_n = \mathbf{y}_n$  for  $n \geq M$  with  $N > M$ . Then

$$\sum_{n=1}^M [f_n(\mathbf{y}_n, \mathbf{y}_{n+1}) - f_n(\mathbf{x}_n, \mathbf{x}_{n+1})] = \sum_{n=1}^N [f_n(\mathbf{y}_n, \mathbf{y}_{n+1}) - f_n(\mathbf{x}_n, \mathbf{x}_{n+1})] \geq 0.$$

This proves (b).

Combining (a) and (b) we see that  $\mathcal{M}_n$  are nested compact sets, hence their intersection is non-empty.

(d) is clear.

Next, let  $\mathbf{x}$  be the argmin of  $\sum_{n=1}^N (\mathbf{z}_n, \mathbf{z}_{n+1})$  and  $\mathbf{y} \in \mathcal{M}_N$ . Let  $\mathbf{z}$  be such that  $\mathbf{z}_n = \mathbf{x}_n$ , for  $1 \leq n \leq N-1$  and  $\mathbf{z}_n = \mathbf{y}_n$  for  $n \geq N$ . Then

$$\sum_{n=1}^N f_n(\mathbf{y}_n, \mathbf{y}_{n+1}) \leq \sum_{n=1}^N f_n(\mathbf{z}_n, \mathbf{z}_{n+1}) \leq \sum_{n=1}^N f_n(\mathbf{x}_n, \mathbf{x}_{n+1}) + 2K$$

proving (e).  $\square$

If, in addition to the assumptions of Lemma 7.6 we also suppose that  $f$  satisfies (U), then part (e) of the lemma implies that for each  $\mathbf{x} \in \mathcal{M}$  (which is non-empty by part (c))

$$\mathfrak{c}_- = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\mathbf{x}_n, \mathbf{x}_{n+1}).$$

## 7.5 Notes and references

Theorem 7.1 is well-known, see [109, 73, 127, 65, 112]. We note that in the homogeneous setting the assumptions on  $f$  can be significantly weakened. In particular, the assumption that  $f$  is bounded can be replaced by the assumption that the distribution of  $f$  is in the domain of attraction of the Gaussian distribution [109], one can allow  $f$  to depend on infinitely many  $X_n$  assuming that the dependence of  $f(x_1, x_2, \dots)$  on  $(x_n, x_{n+1}, \dots)$  decays exponentially in  $n$  [65], and the ellipticity assumption can be replaced by the assumption that the generator has a spectral gap [109, 73]. In particular, the LLT holds under the Doeblin condition saying that  $\exists \varepsilon_0 > 0$  and a measure  $\zeta$  on  $\mathfrak{S}$  such that

$$\pi(x, dy) = \varepsilon_0 \zeta + (1 - \varepsilon_0) \tilde{\pi}(x, dy)$$

where  $\tilde{\pi}$  is an arbitrary transition probability (cf. equation (1.2.2) in the proof of Lemma 1.10). There are also versions of this theorem for  $f$  in the domain of attraction of a stable law, see [4].

The aforementioned weaker conditions however are not sufficient to get LLT in the large deviation regime, in fact large deviation probabilities could be polynomially small for unbounded functions, see [150].

There is a vast literature on the sufficient conditions for the Central Limit Theorem for homogenous chains, see [40, 62, 63, 73, 82, 85, 102] and references wherein, however, the local limit theorem is much less understood, see notes to Chapter 4.

The characterization of coboundaries in terms of vanishing of the asymptotic variance  $\sigma^2$  is due to Leonov [94]. A large number of papers discuss the regularity of the gradients in case an additive functional is a gradient, see [20, 32, 75, 97, 98, 113, 112, 152] and the references wherein. Our approach is closest to [48, 75, 112]. We note that the condition  $u(f) = 0$  which is sufficient for  $f$  being a coboundary, is simpler than the equivalent condition  $\sigma^2 = 0$ . For example for finite chains, to compute  $\sigma^2$  one needs to compute infinitely many correlations  $\mathbb{E}(f_0 f_n)$  while checking that  $u = 0$  involves checking balance of finitely many hexagons.

Inhomogeneous Markov processes arising from perturbations of homogeneous Markov chains as in section 7.2 arise naturally in some stochastic optimization algorithms such as the Metropolis algorithm. For large deviations and other limit theorems for such examples, see [42, 41] and references therein.

Minimizers play important role in statistical mechanics where they are called *ground states*. See e.g. [57, 123]. In the case the phase spaces  $\mathfrak{S}_n$  are non-compact and/or the observable  $f(x, y)$  is unbounded, the minimizers have an interesting geometry, see e.g. [26]. For finite states we have the following remarkable result [18]: for each  $d$  there is a constant  $p(d)$  such that for any homogeneous Markov chain with  $d$  states for any additive functional we have

$$\mathfrak{s}_+ = \max_{q \leq p} \frac{1}{q} \max_{x_1, \dots, x_q} [f(x_1, x_2) + \dots + f(x_{q-1}, x_q) + f(x_q, x_1)].$$

This result is false for more general homogenous chains, consider for example the case  $\mathfrak{S} = \mathbb{N}$  and  $f(x, y) = 1$  if  $y = x + 1$  and  $f(x, y) = 0$  otherwise.

Corollary 7.2 was proven in [49] for inhomogeneous sums of independent random variables (in the independent case one does not need the assumption that  $\lim_{n \rightarrow \infty} \|g_n\|_\infty = 0$  since the gradient obstruction does not appear in the independent case).



# Chapter 8

## LLT for Markov chains in random environment

We prove quenched local limits theorems for Markov chains in random environment with stationary ergodic noise processes.

### 8.1 Markov chains in random environment

Informally, Markov chains in random environment (MCRE) are Markov chains whose transition probabilities depend on a noisy parameter  $\omega$  which varies in time.<sup>1</sup> It is customary to model the time evolution of  $\omega$  by orbits of a dynamical system called the “noise process.” Here are the formal definitions:

**Noise process:** This is an ergodic measure preserving invertible Borel transformation  $T$  on a standard measure space  $(\Omega, \mathcal{F}, m)$ . “**Invertible**” means that there exists  $\Omega_1 \subset \Omega$  of full measure such that  $T : \Omega_1 \rightarrow \Omega_1$  is injective and surjective, and  $T^{-1}, T : \Omega_1 \rightarrow \Omega_1$  are measurable.<sup>2</sup> “**Measure preserving**” means that for every  $E \in \mathcal{F}$ ,  $m(T^{-1}E) = m(E)$ . “**Ergodic**” means that for every  $E \in \mathcal{F}$ ,  $T^{-1}E = E \Rightarrow m(E) = 0$  or  $m(E^c) = 0$ .

If  $m(\Omega) < \infty$  then we will speak of a **finite noise process**, and we will always normalize  $m$  so that  $m(\Omega) = 1$ . If  $m(\Omega) = \infty$ , then we will speak of an **infinite noise process**. The infinite noise processes we consider here will all be defined on  $\sigma$ -finite non-atomic measure spaces. Such processes arise naturally in the study of noise driven by a null recurrent Markov chain, see Example 8.5 below.

**Markov chains in Random Environment (MCRE):** A MCRE with noise process  $(\Omega, \mathcal{F}, m, T)$  is given by the following data:

- **State space:** A separable complete metric space  $\mathfrak{S}$ , with its Borel  $\sigma$ -algebra  $\mathcal{B}$ .
- **Random transition kernel:** A measurable family of Borel probability measures  $\pi(\omega, x, dy)$  on  $(\mathfrak{S}, \mathcal{B})$ , indexed by  $(\omega, x) \in \Omega \times \mathfrak{S}$ . Measurability means that  $(\omega, x) \mapsto \int \varphi(y) \pi(x, \omega, dy)$  is measurable for every bounded Borel  $\varphi : \mathfrak{S} \rightarrow \mathbb{R}$ .
- **Initial probability distribution:** A measurable family of Borel probability measures  $\mu_\omega$  on  $(\mathfrak{S}, \mathcal{B})$  indexed by  $\omega \in \Omega$ , Measurability means that for all bounded Borel  $\varphi : \mathfrak{S} \rightarrow \mathbb{R}$ ,  $\omega \mapsto \int \varphi(x) \mu_\omega(dx)$  is measurable.

This data gives for each  $\omega$  an inhomogeneous Markov chain  $X^\omega = \{X_n^\omega\}$  with state space  $\mathfrak{S}$ , initial distribution  $\mu_\omega$ , and transition kernels  $\pi_n^\omega(x, dy) = \pi(T^n \omega, x, dy)$ .

<sup>1</sup> MCRE should not be confused with “random walks in random environment,” see §8.4.

<sup>2</sup> Invertibility is convenient, but not necessary. Non-invertible ergodic noise processes can always be replaced by their ergodic and invertible natural extensions. See [30, Ch. 10]

Here are some examples. Suppose  $(\mathfrak{S}, \mathcal{B}, \mu_0)$  is a standard probability space,  $S$  is a finite or countable set, and  $\{\pi_i(x, dy)\}_{i \in S}$  are transition probabilities on  $\mathfrak{S}$ .

### Example 8.1 (Bernoulli noise)

Consider the noise process  $(\Omega, \mathcal{F}, m, T)$  where

- o  $\Omega = S^{\mathbb{Z}} = \{(\dots, \omega_{-1}, \omega_0, \omega_1, \dots) : \omega_i \in S\}$ ;
- o  $\mathcal{F}$  is generated by the **cylinders**  $k[a_k, \dots, a_n] := \{\omega \in \Omega : \omega_i = a_i, k \leq i \leq n\}$ ;
- o  $\{p_i\}_{i \in S}$  are non-negative numbers s.t.  $\sum p_i = 1$ , and  $m$  is the unique measure s.t.  $m(k[a_k, \dots, a_n]) = p_{a_k} \cdots p_{a_n}$  for all cylinders.
- o  $T : \Omega \rightarrow \Omega$  is the **left shift map**,  $T[(\omega_i)_{i \in \mathbb{Z}}] = (\omega_{i+1})_{i \in \mathbb{Z}}$

It's well-known that  $(\Omega, \mathcal{F}, \mu, T)$  is an ergodic probability preserving map.

Define  $\pi(\omega, x, dy) := \pi_{\omega_0}(x, dy)$ . Notice that  $\pi(T^n \omega, x, dy) = \pi_{\omega_n}(x, dy)$ , and  $\omega_n$  are iid's taking the values  $i \in S$  with probabilities  $p_i$ . Since  $\omega_n$  are iid,  $\{X^{\omega} : \omega \in \Omega\}$  represent a random Markov chain whose transition probabilities vary randomly and independently in time.

### Example 8.2 (Positive recurrent Markov noise)

Suppose  $(Y_n)_{n \in \mathbb{Z}}$  is a stationary ergodic Markov chain with state space  $S$  and a stationary probability vector  $(p_s)_{s \in S}$ . In particular,  $(Y_n)_{n \in \mathbb{Z}}$  is positive recurrent. Let:

- o  $\Omega := \{(\omega_i) \in S^{\mathbb{Z}} : \mathbb{P}[Y_1 = \omega_i, Y_2 = \omega_{i+1}] \neq 0 \text{ for all } i \in \mathbb{Z}\}$ ;
- o  $\mathcal{F}$  is the  $\sigma$ -algebra generated by the cylinders (see above);
- o  $m$  is the unique (probability) measure such that  $m(k[a_k, \dots, a_n]) = \mathbb{P}[Y_k = a_k, \dots, Y_n = a_n]$  for all cylinders;
- o  $T$  is the left shift map (see above).

Define as before,  $\pi(\omega, x, dy) := \pi_{\omega_0}(x, dy)$ . The resulting MCRE represents a Markov chain whose transition probabilities at time  $n = 1, 2, 3, \dots$  are  $\pi_{Y_{n-1}}(x, dy)$ .

### Example 8.3 (General stationary ergodic noise processes)

The previous construction works verbatim with any stationary ergodic stochastic process  $\{Y_n\}$  taking values in  $S$ . The assumption that  $S$  is countable can be replaced by requiring only that  $S$  be complete, separable, metric space, see e.g. [52].

### Example 8.4 (Quasi-periodic noise)

Let  $(\Omega, \mathcal{F}, m, T)$  be the circle rotation:  $\Omega = \mathbb{T}^1 := \{\omega \in \mathbb{C} : |\omega| = 1\}$ ;  $\mathcal{F}$  is the Borel  $\sigma$ -algebra;  $m$  is the normalized Lebesgue measure; and  $T : \Omega \rightarrow \Omega$  is the rotation by an angle  $\alpha$ :  $T(\omega) = e^{i\alpha}\omega$ .  $T$  is probability preserving, and it is well-known that  $T$  is ergodic iff  $\alpha/2\pi$  is irrational.

Choose a partition of the unit circle  $\Omega$  into disjoint arcs  $\{I_i\}_{i \in S}$  and define  $\varphi : \Omega \rightarrow S$  by  $\varphi(\omega) = i$  for  $\omega \in I_i$ . For example, if  $S = \{1, 2\}$  we can take  $I_1, I_2$  to be two equal halves of the circle. Next define

$$\pi(\omega, x, dy) = \pi_{\varphi(\omega)}(x, dy)$$

Now  $X^{\omega}$  are inhomogeneous Markov chains whose transition probabilities vary quasi-periodically: They are given by  $\pi_{\varphi(e^{in\alpha}\omega)}(x, dy)$ .

More generally, one can take a  $d$  parameter measurable family of transition probabilities  $\pi_{\omega}(x, y)$ , where  $\omega = (\omega_1, \omega_2, \dots, \omega_d) \in \mathbb{R}^d / \mathbb{Z}^d$ , fix some “initial phase”  $(\bar{\omega}_1, \dots, \bar{\omega}_d)$ , and consider the chain with transition probabilities

$$\pi_n(x, y) = \pi_{(\bar{\omega}_1 + n\alpha_1, \dots, \bar{\omega}_d + n\alpha_d) \bmod \mathbb{Z}^d}(x, y).$$

### Example 8.5 (Null recurrent Markov noise)

This is an example with infinite noise process. Suppose  $(Y_n)_{n \in \mathbb{Z}}$  is an ergodic null recurrent Markov chain with countable state space  $S$ , and stationary positive vector  $(p_i)_{i \in S}$ . Here  $p_i > 0$  and (by null recurrence)  $\sum p_i = \infty$ . For example,  $(Y_n)_{n \in \mathbb{Z}}$  could be the simple random walk on  $\mathbb{Z}^d$  for  $d = 1, 2$ , with the stationary measure which assigns the same mass to each site of  $\mathbb{Z}^d$ . Let

- $\Omega = \{(\omega_i)_{i \in \mathbb{Z}} \in S^{\mathbb{Z}} : \mathbb{P}[Y_1 = \omega_i, Y_2 = \omega_{i+1}] \neq 0 \text{ for all } i \in \mathbb{Z}\}$ ;
- $\mathcal{F}$  is the  $\sigma$ -algebra generated by the cylinders;
- $m$  is the unique (infinite) Borel measure which satisfies for each cylinder

$$m_k[a_k, \dots, a_n] = p_{a_k} \mathbb{P}[Y_i = a_i \text{ for } k \leq i \leq n | Y_k = a_k]$$

- $T : \Omega \rightarrow \Omega$  is the left shift map  $T[(\omega_i)_{i \in \mathbb{Z}}] = \omega_{i+1}$ .

Then it is well-known that  $(\Omega, \mathcal{F}, m, T)$  is an *infinite* ergodic measure preserving invertible map, see [2].

Just as in Example 8.2, one can easily construct many MCRE with transition probabilities  $\pi_{Y_n}(x, dy)$  which vary randomly in time according to  $(Y_n)_{n \in \mathbb{Z}}$ . For each particular realization of  $\omega = (\omega_i)_{i \in \mathbb{Z}}$ ,  $X^\omega$  is an ordinary inhomogeneous Markov chain (on a probability space). But as we shall see below, some features of  $X^\omega$  such as the growth of variance, are different than in the finite noise process case.

### Example 8.6 (Transient Markov noise: a non-example)

The previous construction fails for transient Markov chains such as the random walk on  $\mathbb{Z}^d$  for  $d \geq 3$ , because in the transient case,  $(\Omega, \mathcal{F}, m, T)$  is not ergodic, [2].

We could try to work with the ergodic components of  $m$ , but this does not yield a new mathematical object, because of the following general fact [2]: Every ergodic component of an *invertible* totally dissipative infinite measure preserving map is concentrated on a single orbit  $\{T^n(\omega)\}_{n \in \mathbb{Z}}$ . MCRE with such noise processes have just one possible realization of noise up to time shift. Their theory is the same as the theory of general inhomogeneous Markov chains, and does not merit separate treatment.

Suppose  $X^\Omega$  is a MCRE with noise space  $(\Omega, \mathcal{F}, m, T)$ . A **Random additive functional** is a measurable function  $f : \Omega \times \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$ . This induces the additive functional  $f^\omega$  on  $X^\omega$

$$f_n^\omega(x, y) = f(T^n \omega, x, y).$$

For each  $\omega \in \Omega$  we define

$$\begin{aligned} S_N^\omega &:= \sum_{n=1}^N f_n^\omega(X_n^\omega, X_{n+1}^\omega) \equiv \sum_{n=1}^N f(T^n \omega, X_n^\omega, X_{n+1}^\omega), \\ V_N^\omega &:= \text{Var}(S_N^\omega) \text{ w.r.t. the distribution of } X^\omega. \end{aligned}$$

Throughout this chapter, we make the following **standing assumptions**:

- (B) **Uniform boundedness:**  $|f| \leq K$  where  $K < \infty$  is a constant;
- (E) **Uniform ellipticity:** There is a constant  $0 < \varepsilon_0 < 1$  and a Borel function  $p : \Omega \times \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$  such that
  - (a)  $\pi(\omega, x, dy) = p(\omega, x, y) \mu_\omega(dy)$ ;
  - (b)  $0 \leq p \leq 1/\varepsilon_0$ ;

(c)  $\int_{\mathfrak{S}} p(\omega, x, y) p(T\omega, y, z) \mu_{T\omega}(dy) > \varepsilon_0$  for all  $\omega, x, z$ .

(S) **Stationarity:** For every  $\varphi : \mathfrak{S} \rightarrow \mathbb{R}$  bounded and Borel, for every  $\omega \in \Omega$ ,

$$\int \varphi(y) \mu_{T\omega}(dy) = \int_{\mathfrak{S}} \left( \int_{\mathfrak{S}} \varphi(y) \pi(\omega, x, dy) \right) \mu_{\omega}(dx).$$

(B) and (E) imply that  $f^{\omega}$  is a uniformly bounded additive functional and that  $X^{\omega}$  is uniformly elliptic for every  $\omega$ . (S) is equivalent to saying that if  $X_0$  is distributed according to  $\mu_{\omega}$  then  $X_n$  is distributed according to  $\mu_{T^n\omega}$  for all  $n > 0$ . Subject to (E), (S) can always be assumed without loss of generality, because of Proposition 1.12 and the discussion which follows it.

Some of our results will require the following **continuity hypothesis**:

(C) The Borel structure of  $\Omega$  and  $\mathfrak{S}$  is generated by a topologies so that  $\Omega$  and  $\mathfrak{S}$  are complete and separable metric spaces, and

- (C1)  $T : \Omega \rightarrow \Omega$  is a homeomorphism and  $\text{supp}(m) = \Omega$ .
- (C2)  $(\omega, x, y) \mapsto p(\omega, x, y)$  is continuous, and  $\omega \mapsto \int_{\mathfrak{S}} \varphi d\mu_{\omega}$  is continuous for every bounded continuous  $\varphi : \mathfrak{S} \rightarrow \mathbb{R}$ .
- (C3)  $(\omega, x, y) \mapsto f(\omega, x, y)$  are continuous.

(C) is not part of our standing assumptions, and we will state it explicitly whenever it is used.

## 8.2 Main results

Let  $\mathbb{P}$  denote the measure on  $\Omega \times \mathfrak{S} \times \mathfrak{S}$  which represents the joint distribution of  $(\omega, X_1^{\omega}, X_2^{\omega})$ :

$$\mathbb{P}(d\omega, dx, dy) := \int_{\mathfrak{S}} \int_{\mathfrak{S}} \int_{\Omega} m(d\omega) \mu_{\omega}(dx) \pi(\omega, x, dy). \quad (8.2.1)$$

(1)  $f(\omega, x, y)$  is called **relatively cohomologous to a constant** if there are bounded measurable functions  $a : \Omega \times \mathfrak{S} \rightarrow \mathbb{R}$  and  $c : \Omega \rightarrow \mathbb{R}$  such that

$$f(\omega, x, y) = a(\omega, x) - a(T\omega, y) + c(\omega) \quad \mathbb{P}\text{-a.e.}$$

(2) Fix  $t \neq 0$ , then  $f(\omega, x, y)$  is **relatively cohomologous to a coset of  $t\mathbb{Z}$**  if there are measurable functions  $a : \Omega \times \mathfrak{S} \rightarrow S^1$  and  $\lambda : \Omega \rightarrow S^1$  s.t.

$$e^{(2\pi i/t)f(\omega, x, y)} = \lambda(\omega) \frac{a(\omega, x)}{a(T\omega, y)} \quad \mathbb{P}\text{-a.e.}$$

**Theorem 8.1.** Assume  $f$  is an additive functional on a MCRE with finite noise process. Under the standing assumptions (B), (E), (S):

- (1) If  $f$  is relatively cohomologous to a constant, then  $|V_N^{\omega}| \leq C$  for all  $N$ , for a.e.  $\omega$ , where  $C = C(\varepsilon_0, K)$  is a constant.
- (2) If  $f$  is not relatively cohomologous to a constant, then there is a constant  $\sigma^2 > 0$  such that for a.e.  $\omega$ ,  $V_N^{\omega} \sim N\sigma^2$  as  $N \rightarrow \infty$ .

**Theorem 8.2.** Let  $f$  be an additive functional on a MCRE with finite noise process. Assume the standing assumptions (B), (E), (S) and that

- (a) Either  $|\mathfrak{S}| \leq \aleph_0$ , or  $|\mathfrak{S}| > \aleph_0$  and the continuity hypothesis (C) holds.

(b)  $f$  is not relatively cohomologous to a coset of  $t\mathbb{Z}$  for any  $t \neq 0$ .

Then there exists  $\sigma^2 > 0$  such that for a.e.  $\omega$ , for every open interval  $(a, b)$ , and for every  $z_N, z \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{N}} \rightarrow z$ ,

$$\mathbb{P}[S_N^\omega - z_N \in (a, b)] \sim \frac{1}{\sqrt{N}} \left( \frac{e^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \right) |a - b| \text{ as } N \rightarrow \infty.$$

**Theorem 8.3.** Let  $f$  be an additive functional on a MCRE with finite noise process. Assume the standing assumptions (B), (E), (S), and that all the values of  $f$  are integers. If  $f$  is not relatively cohomologous to a coset of  $t\mathbb{Z}$  with  $t \neq 1$ , then there exists  $\sigma^2 > 0$  such that for a.e.  $\omega$ , and for every  $z_N, z \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{N}} \rightarrow z$ ,

$$\mathbb{P}[S_N^\omega = z_N] \sim \frac{1}{\sqrt{N}} \left( \frac{e^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \right) \text{ as } N \rightarrow \infty.$$

**Theorem 8.4.** Let  $f$  be an additive functional on a MCRE with finite noise process  $(\Omega, \mathcal{F}, m, T)$ . Assume (B), (E), (S). If  $f$  is not relatively cohomologous to a constant, then

- (1) There exists a continuously differentiable and strictly convex function  $\mathcal{F} : \mathbb{R} \rightarrow \mathbb{R}$  such that for a.e.  $\omega \in \Omega$ ,  $\mathcal{F}(\xi) = \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}(e^{\xi S_N^\omega})$  for all  $\xi \in \mathbb{R}$ .
- (2)  $\frac{1}{N} \mathbb{E}(S_N^\omega) \xrightarrow[N \rightarrow \infty]{} \mathcal{F}'(0)$  for a.e.  $\omega$ .
- (3) Let  $\mathcal{F}'(\pm\infty) := \lim_{\xi \rightarrow \pm\infty} \mathcal{F}'(\xi)$ , and let  $\mathcal{I}_N(\eta, \omega)$ ,  $\mathcal{I}(\eta)$  denote the Legendre transforms of  $\mathcal{F}_N(\xi) := \frac{1}{N} \log \mathbb{E}(e^{\xi S_N^\omega})$ ,  $\mathcal{F}(\xi)$ . Then for a.e.  $\omega$ , for every  $\eta \in (\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$ ,  $\mathcal{I}_N(\eta, \omega) \xrightarrow[N \rightarrow \infty]{} \mathcal{I}(\eta)$ .
- (4)  $\mathcal{I}(\eta)$  is strictly convex, has compact level sets, is equal to zero at  $\eta = \mathcal{F}'(0)$ , and is strictly positive elsewhere.
- (5) With probability one

$$\mathfrak{c}_- = \mathcal{F}(-\infty) = \lim_{N \rightarrow \infty} \frac{\text{ess inf } S_N^\omega}{N}, \quad \mathfrak{c}_+ = \mathcal{F}(\infty) = \lim_{N \rightarrow \infty} \frac{\text{ess sup } S_N^\omega}{N}.$$

**Corollary 8.7** Under the conditions of the previous theorem, for a.e.  $\omega$ ,  $S_N^\omega/N$  satisfies the large deviations principle with the rate function  $\mathcal{I}(\eta)$ :

- (1)  $\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[S_N^\omega/N \in K] \leq -\inf_{z \in K} \mathcal{I}(z)$  for all closed sets  $K \subset \mathbb{R}$ .
- (2)  $\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[S_N^\omega/N \in G] \geq -\inf_{z \in G} \mathcal{I}(z)$  for all open sets  $G \subset \mathbb{R}$ .

*Proof.* This is a consequence of the Gärtner-Ellis Theorem. □

So far we have only considered MCRE with finite noise spaces. We will now discuss the case of infinite noise spaces  $(\Omega, \mathcal{F}, m, T)$ . The main new phenomena in this case are:

**Example 8.8** For MCRE with an infinite noise process:

- (a) It is possible that  $V_N^\omega \rightarrow \infty$  m-a.e., but that  $V_N^\omega = o(N)$  a.e.
- (b) It is possible that  $\exists a_N$  s.t.  $V_N^\omega \sim a_N$  for m-a.e.  $\omega$ .

*Proof.* Let  $X_n$  be iid bounded real random variables with variance one and distribution  $\mu$ . Let  $f_n(x) = x$ . Let  $(\Omega, \mathcal{F}, m, T)$  be an infinite noise process, and fix  $E \in \mathcal{F}$  of finite positive measure. Let

$$\pi(\omega, x, dy) := \mu(dy), \quad f(\omega, x, y) := 1_E(\omega)x$$

Then  $S_N^\omega = \sum_{n=1}^N 1_E(T^n \omega) X_n$ , and  $V_N^\omega = \sum_{n=1}^N 1_E(T^n \omega)$ .

We now appeal to the following general results from infinite ergodic theory. Let  $(\Omega, \mathcal{F}, m, T)$  be an ergodic, invertible, measure preserving map on a non-atomic  $\sigma$ -finite measure space, and let  $L_+^1 := \{A \in L^1(\Omega, \mathcal{F}, m) : A \geq 0, \int A dm > 0\}$ . If  $m(\Omega) = \infty$ , then

- (1)  $\sum_{n=1}^N A \circ T^n = \infty$  almost everywhere for all  $A \in L_+^1$ ;
- (2)  $\frac{1}{N} \sum_{n=1}^N A \circ T^n \xrightarrow[N \rightarrow \infty]{} 0$  almost everywhere for all  $A \in L^1$ ;
- (3) Let  $a_N$  be a sequence of positive real numbers, then at least one of the following possibilities happens:

- (a)  $\liminf_{N \rightarrow \infty} \frac{1}{a_N} \sum_{n=1}^N A \circ T^n = 0$  a.e. for all  $A \in L_+^1$ ;
- (b)  $\limsup_{N \rightarrow \infty} \frac{1}{a_N} \sum_{n=1}^N A \circ T^n = \infty$  a.e. for all  $A \in L_+^1$ .

So  $\exists a_N \uparrow \infty$  s.t.  $\sum_{n=1}^N A(T^n \omega) \sim a_N$  for a.e.  $\omega$ , even for a single  $A \in L_+^1$ .

These results can all be found in [2]: (1) is a consequence of the Halmos Recurrence Theorem; (2) follows from the Ratio Ergodic Theorem; and (3) is a theorem of J. Aaronson. Specializing to the case  $A = 1_E$  we find that  $V_N^\omega \rightarrow \infty$  a.e.;  $V_N^\omega = o(N)$  a.e. as  $N \rightarrow \infty$ ; and  $\exists a_N$  so that  $V_N^\omega \sim a_N$  for a.e.  $\omega \in \Omega$ .

Here are our general results on MCRE with infinite noise spaces.

**Theorem 8.5.** Suppose  $f^\omega$  is a random additive functional on a MCRE with infinite noise space on a non-atomic  $\sigma$ -finite measure space. Under the standing assumptions (B), (E), (S):

- (1) If  $f$  is relatively cohomologous to a constant, then  $|V_N^\omega| \leq C$  for all  $N$ , for a.e.  $\omega$ , where  $C = C(\epsilon_0, K)$  is a constant.
- (2) If  $f$  is not relatively cohomologous to a constant then  $V_N^\omega \rightarrow \infty$  a.s.

**Theorem 8.6.** Suppose  $f^\omega$  is a random additive functional on a MCRE with infinite noise space on a non-atomic  $\sigma$ -finite measure space. Assume the standing assumptions (B), (E), (S) and that

- (a) Either  $|\mathfrak{S}| \leq \aleph_0$ , or  $|\mathfrak{S}| > \aleph_0$  and the continuity hypothesis (C) holds.
- (b)  $f$  is not relatively cohomologous to a coset of  $t\mathbb{Z}$  for any  $t \neq 0$ .

Then for a.e.  $\omega$ , for every open interval  $(a, b)$ , and for every  $z_N, z \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{V_N^\omega} \rightarrow z$ ,  $\mathbb{P}[S_N^\omega - z_N \in (a, b)] \sim \frac{e^{-z^2/2}}{\sqrt{2\pi V_N^\omega}} |a - b|$  as  $N \rightarrow \infty$ .

**Theorem 8.7.** Suppose  $f^\omega$  is a random additive functional on a MCRE with infinite noise space on a non-atomic  $\sigma$ -finite measure space. Assume the standing assumptions (B), (E), (S), and that all the values of  $f$  are integers. If  $f$  is not relatively cohomologous to a coset of  $t\mathbb{Z}$  with  $t \neq 1$ , then for every  $z_N, z \in \mathbb{R}$  such that  $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{V_N^\omega}} \rightarrow z$ , for a.e.  $\omega$ ,  $\mathbb{P}[S_N^\omega = z_N] \sim \frac{e^{-z^2/2}}{\sqrt{2\pi V_N^\omega}}$  as  $N \rightarrow \infty$ .

### 8.3 Proofs

Throughout this section  $X^\omega$  is a Markov chain in random environment with stationary ergodic, possibly infinite, noise process  $(\Omega, \mathcal{F}, m, T)$ , and  $f^\omega$  is a random additive functional on  $X^\omega$ . We assume throughout (B),(E),(S).

#### 8.3.1 The essential range is a.s. constant

The purpose of this section is to prove the following result:

**Proposition 8.9** *There exist closed subgroups  $H, G_{ess} \leq \mathbb{R}$  s.t. for  $m$ -a.e.  $\omega$ , the co-range of  $(X^\omega, f^\omega)$  equals  $H$ , the essential range of  $(X^\omega, f^\omega)$  equals  $G_{ess}$ , and*

$$G_{ess} = \begin{cases} \mathbb{R} & H = \{0\}, \\ \frac{2\pi}{t} \mathbb{Z} & H = t\mathbb{Z}, t \neq 0, \\ \{0\} & H = \mathbb{R}. \end{cases}$$

We call  $H$  and  $G_{ess}$  the **a.s. co-range** and **a.s. essential range**.

We begin with a calculation of the structure constants of  $(X^\omega, f^\omega)$ . Fix an element  $\omega$  in the noise space, and let  $\text{Hex}(\omega)$  denote the probability space of position 3 hexagons for  $X^\omega$ . Let  $m_\omega$  denote the hexagon measure, as defined in §1.3.1. Recall the definition of the balance  $\Gamma(P)$  of a hexagon  $P$ , and define

$$\begin{aligned} u(\omega) &:= \mathbb{E}(|\Gamma(P)|^2)^{1/2} \\ d(\xi, \omega) &:= \mathbb{E}(|e^{i\xi\Gamma(P)} - 1|^2)^{1/2} \text{ (expectation on } P \in \text{Hex}(\omega) \text{ w.r.t. } m_\omega\text{).} \end{aligned}$$

Since the space of position  $n + 3$  hexagons for  $X^\omega$  is  $\text{Hex}(T^n \omega)$ , together with the hexagon measure  $m_{T^n \omega}$ , it follows that the structure constants of  $(X^\omega, f^\omega)$  are

$$d_{n+3}(\xi, f^\omega) = d(T^n \omega, \xi) \quad \text{and} \quad u_{n+3}(f^\omega) = u(T^n \omega) \quad (n \geq 0). \quad (8.3.1)$$

**Lemma 8.10**  *$u(\cdot), d(\cdot, \cdot)$  are Borel measurable, and for every  $\omega$ ,  $d(\cdot, \omega)$  is continuous. Under the continuity hypothesis (C),  $u(\cdot), d(\cdot, \cdot)$  are continuous.*

*Proof.* To check this, express the hexagon measure explicitly as a measure on  $\mathfrak{S}^6$  in terms of the transition kernel  $\pi(\omega, x, y)$ , using the formulas for the bridge distributions of §1.2.3, and write  $\Gamma(P)$  explicitly a function on  $\mathfrak{S}^6$  in terms of  $f(\omega, x, y)$ . We omit the details, which are routine.  $\square$

**Proof of Proposition 8.9.** Let  $H_\omega := H(X^\omega, f^\omega)$  be the essential range of  $(X^\omega, f^\omega)$ . By Theorem 3.1,  $H_\omega$  is either  $\mathbb{R}$  or  $t\mathbb{Z}$  for some  $t \geq 0$ . By (8.3.1)

$$D_N(\xi, \omega) := \sum_{n=3}^N d_n(\xi, f^\omega)^2 \equiv \sum_{n=0}^{N-3} d(T^n \omega, \xi)^2.$$

STEP 1:  $U(a, b) := \{\omega \in \Omega : D_N(\cdot, \omega) \xrightarrow[N \rightarrow \infty]{} \infty \text{ uniformly on } (a, b)\}$  is measurable and  $T$ -invariant for all  $a < b$ .

*Proof.*  $T$ -invariance is because  $d^2 \leq 4$  whence  $|D_N(\xi, T\omega) - D_N(\xi, \omega)| \leq 8$ . Measurability is because of the identity

$$U(a, b) = \left\{ \omega \in \Omega : \begin{array}{l} \forall M \in \mathbb{Q} \exists N \in \mathbb{N} \text{ s.t.} \\ \text{for all } \xi \in (a, b) \cap \mathbb{Q}, D_N(\omega, \xi) > M \end{array} \right\}.$$

The inclusion  $\subseteq$  is obvious. The inclusion  $\supseteq$  is because if  $\omega \notin U(a, b)$  then for some  $M \in \mathbb{Q}$ , for all  $N \in \mathbb{N}$  there exists some  $\eta \in (a, b)$  s.t.  $D_N(\omega, \eta) < M$ , whence by the continuity of  $\eta \mapsto D_N(\omega, \eta)$  there is some  $\xi \in (a, b) \cap \mathbb{Q}$  such that  $D_N(\omega, \xi) < M$ . So  $\omega \notin U(a, b) \Rightarrow \omega \notin \text{RHS}$ .

STEP 2: The sets  $\Omega_1 := \{\omega \in \Omega : H_\omega = \{0\}\}$ ,  $\Omega_2 := \{\omega \in \Omega : H_\omega = \mathbb{R}\}$ , and  $\Omega_3 := \{\omega \in \Omega : H_\omega = t\mathbb{Z} \text{ for some } t \neq 0\}$  are measurable and  $T$ -invariant. Therefore by ergodicity, for each  $i$ , either  $m(\Omega_i) = 0$  or  $m(\Omega_i^c) = 0$ .

*Proof.* Recall that for Markov chains,  $D_N \rightarrow \infty$  uniformly on compact subsets of the complement of the co-range (Theorem 3.5). So

$$\Omega_1 = \bigcap_{n=1}^{\infty} U\left(\frac{1}{n}, n\right), \quad \Omega_2 = \bigcap_{0 < a < b \text{ rational}} U(a, b)^c, \quad \Omega_3 = \Omega_1^c \cap \Omega_2^c.$$

By step 1,  $\Omega_i$  are  $T$ -invariant and measurable. Since  $T$  is ergodic, these sets are either of measure zero or of full measure.

By Theorem 3.2, if  $\Omega_1$  has full measure, then the essential range is a.s.  $\mathbb{R}$ . Similarly, if  $\Omega_2$  has full measure, then the essential range is  $\{0\}$  almost surely. It remains to consider the case when  $\Omega_3$  has full measure.

STEP 3: If  $\Omega_3$  has full measure, then there exist  $t \neq 0$  such that  $\Omega_3(t) := \{\omega \in \Omega : H_\omega = t\mathbb{Z}\}$  has full measure, and then the essential range is  $(2\pi/t)\mathbb{Z}$  almost surely.

*Proof.* For every  $\omega \in \Omega_3$  there exists  $t(\omega) > 0$  such that  $H_\omega = t(\omega)\mathbb{Z}$ . We can characterize  $t(\omega)$  as follows:

$$t(\omega) = \sup \left\{ t \in \mathbb{Q} \cap (0, \infty) : \begin{array}{l} D_N(\omega, \cdot) \rightarrow \infty \text{ uniformly} \\ \text{on compact subsets of } (0, t) \end{array} \right\}.$$

It is clear from this expression that  $t(T\omega) = t(\omega)$ , and that for every  $A > 0$ ,

$$[t(\omega) \geq A] = \bigcap_{0 < a < b < A \text{ rational}} U(a, b).$$

So  $t(\cdot)$  is a measurable  $T$ -invariant function, whence by ergodicity constant. Let  $t$  denote this constant, then  $H_\omega = t\mathbb{Z}$  for a.e.  $\omega$ . By Theorem 3.2,  $G_{ess}(X^\omega, f^\omega) = (2\pi/t)\mathbb{Z}$  almost surely.  $\square$

### 8.3.2 Variance growth

In this section we prove Theorems 8.1 and 8.5 on the behavior of  $V_N^\omega$  as  $N \rightarrow \infty$ .

**Lemma 8.11** Suppose  $(\Omega, \mathcal{F}, m, T)$  is an invertible, ergodic, measure preserving map of a probability space or of a non-atomic infinite measure space. Let  $A : \Omega \rightarrow \mathbb{R}$  be a non-negative measurable function. Either  $A = 0$  a.e., or  $\sum_{n \geq 0} A \circ T^n = \infty$  a.e.

*Proof.* If  $m(\Omega) < \infty$ , then the Lemma follows from the Birkhoff ergodic theorem. In the more general case  $m(\Omega) \leq \infty$ , the lemma follows from the well-known fact that invertible ergodic measure preserving maps on non-atomic measure spaces are conservative. We supply the details, for completeness.

If  $A$  is not equal to 0 a.e., then there is  $\varepsilon > 0$  s.t.  $E := \{\omega \in \Omega : A(\omega) \geq \varepsilon\}$  has positive measure. We claim that

$$\sum_{n \geq 0} 1_E(T^n \omega) = \infty \quad (8.3.2)$$

almost everywhere on  $E$ . Since  $A \geq \varepsilon 1_E$  (8.3.2) implies that  $\sum_{n \geq 0} A(T^n \omega) = \infty$  almost everywhere on  $E$ , and, by ergodicity, almost everywhere on  $\Omega$ , proving the lemma.

It remains to prove (8.3.2). Suppose by way of contradiction that it is not true that  $\sum_{n \geq 0} 1_E(T^n \omega) = \infty$  almost everywhere on  $E$ . Then there exists  $N$  s.t.

$$W := \{\omega \in E : \sum_{n=0}^{\infty} 1_E(T^n \omega) = N\}$$

has positive measure. The invertibility and measurability of  $T$  implies that  $T^n(W)$  is measurable for all  $n \in \mathbb{Z}$ , and that  $\{T^n(W)\}_{n \in \mathbb{Z}}$  are pairwise disjoint.

Since  $(\Omega, \mathcal{F}, m)$  is non-atomic, we can break  $W = W_1 \cup W_2$  where  $W_i$  are measurable, disjoint, and with positive measure. By invertibility,  $\widehat{W}_i := \bigcup_{n \in \mathbb{Z}} T^n W_i$  are disjoint  $T$ -invariant sets with positive measure. But this contradicts ergodicity.  $\square$

**Part 1:  $V_N^\omega$  is bounded, or tends to infinite almost surely.** Recall that  $K$  is a bound for  $\text{ess sup } |f|$ , and  $\varepsilon_0$  is a uniform ellipticity constant for  $X^\omega$ . By Theorem 2.7 and (8.3.1) there are positive constants  $C_i = C_i(\varepsilon_0, K)$  ( $i = 1, 2$ ) such that for all  $N$ ,

$$C_1^{-1} \sum_{n=3}^N u(T^n \omega)^2 - C_2 \leq V_N^\omega \leq C_1 \sum_{n=3}^N u(T^n \omega)^2 + C_2.$$

If  $u(\omega) = 0$   $m$ -a.e., then for a.e.  $\omega$ ,  $V_N^\omega \leq C_2$  for all  $N$ . Otherwise, by Lemma 8.11,  $\sum_{n=3}^N u(T^n \omega)^2 \xrightarrow[N \rightarrow \infty]{} \infty$ , whence  $V_N^\omega \rightarrow \infty$  almost everywhere.

**Part 2: Linear growth of variance when  $V_N^\omega \rightarrow \infty$  a.e. and  $m(\Omega) = 1$ .** Suppose  $m(\Omega) = 1$  and  $V_N^\omega \rightarrow \infty$  almost surely. We claim that

$$\exists \sigma^2 > 0 \text{ s.t. } V_N^\omega \sim N\sigma^2 \text{ a.s.} \quad (8.3.3)$$

Let  $\sigma_0^2 := \int_{\Omega} u^2 dm$ . This is a finite number, because  $\|u\|_{\infty} \leq 6K$  and  $m(\Omega) = 1$ . This is a positive number, because as we saw in part 1, if  $u = 0$  a.e., then  $V_N^\omega = O(1)$  a.e. contrary to our assumptions. By the pointwise ergodic theorem,  $\sum_{n=3}^N u(T^n \omega)^2 = [1 + o(1)]\sigma_0^2 N$ . Hence  $V_N^\omega \geq [1 + o(1)]C_1(\varepsilon_0, K)^{-1}N\sigma_0^2 \rightarrow \infty$ .

Let  $F_n := f(T^{n-1} \omega, X_n^\omega, X_{n+1}^\omega)$  and let  $\mathbb{E}^\omega$ ,  $\text{Var}^\omega$ ,  $\text{Cov}^\omega$  denote the expectation, variance and covariance with respect to  $X^\omega$ , then

$$\begin{aligned} V_N^\omega &= \sum_{n=1}^N \text{Var}^\omega(F_n) + 2 \sum_{n=1}^N \sum_{m=n+1}^N \text{Cov}^\omega(F_n, F_m) \\ &= \sum_{n=1}^N \text{Var}^\omega(F_n) + 2 \sum_{n=1}^N \sum_{k=1}^{N-n} \text{Cov}^\omega(F_n, F_{n+k}). \end{aligned}$$

By assumption (S)  $\{\mu_\omega\}$  are stationary, so  $\{X_i^{T^n\omega}\}_{i \geq 1}$  has the same distribution as  $\{X_1^\omega\}_{i \geq n}$ . Therefore  $\text{Var}^\omega(F_n) = \text{Var}^{T^{n-1}\omega}(F_1)$  and  $\text{Cov}^\omega(F_n, F_{n+k}) = \text{Cov}^{T^{n-1}\omega}(F_1, F_{1+k})$ . Thus

$$V_N^\omega = \sum_{n=0}^{N-1} \psi_0(T^n\omega) + 2 \sum_{n=0}^{N-1} \sum_{k=1}^{N-n} \psi_k(T^n\omega),$$

where  $\psi_0(\omega) := \text{Var}^\omega[f(\omega, X_1^\omega, X_2^\omega)]$  and

$$\psi_k(\omega) = \text{Cov}^\omega[f(\omega, X_1^\omega, X_2^\omega), f(T^k\omega, X_{k+1}^\omega, X_{k+2}^\omega)].$$

By the ergodic theorem  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \psi_0(T^n\omega) = \int \psi_0 dm$ . To find the limit of the normalized double sum we first recall that by the uniform mixing of  $\{X_n^\omega\}$  (a consequence of the ellipticity assumption),  $\|\psi_k\|_\infty \leq C_{mix} \|f\|_\infty^2 \theta^k$  with  $C_{mix}, 0 < \theta < 1$  which only depend on  $\varepsilon_0$  (Proposition 1.11). Therefore for every  $M$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=1}^{N-n} \psi_k(T^n\omega) = \left[ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=1}^{M-1} \psi_k(T^n\omega) \right] + O(\theta^M),$$

whence by the ergodic theorem  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=1}^{N-n} \psi_k(T^n\omega) = \sum_{k=1}^{\infty} \int \psi_k dm$ , with the last sum converging exponentially fast. In summary,

$$\frac{1}{N} V_N^\omega \xrightarrow[N \rightarrow \infty]{} \sigma^2 := \int \left[ \psi_0 + 2 \sum_{k=1}^{\infty} \psi_k \right] dm.$$

Since as we saw above  $\liminf \frac{1}{N} V_N^\omega \geq C_1 \sigma_0^2$  and  $\sigma_0^2 > 0$ , it must be the case that  $\sigma^2 > 0$ , and (8.3.3) is proved.

We now relate the following two properties:

- (a)  $f$  is relatively cohomologous to a constant;
- (b)  $V_N^\omega$  is bounded  $m$ -a.e.

**Part 3: (a)  $\Rightarrow$  (b):** Suppose  $f$  is relatively cohomologous to a constant. By Fubini's theorem, for  $m$ -a.e.  $\omega$ , for every  $n$ ,

$$f_n^\omega(X_n^\omega, X_{n+1}^\omega) = a(T^n\omega, X_n^\omega) - a(T^{n+1}\omega, X_{n+1}^\omega) + c(T^n\omega) \text{ a.s.}$$

with respect to the distribution of  $\{X_n^\omega\}$ .<sup>3</sup>

Summing over  $n$ , we obtain that for a.e.  $\omega$ , for every  $N$ ,

$$|S_N^\omega - \sum_{n=1}^N c(T^n\omega)| = |a(\omega, X_1^\omega) - a(T^N\omega, X_{N+1}^\omega)| \leq 2 \text{ess sup } a(\cdot, \cdot).$$

In particular, for every  $\omega$ ,  $V_N^\omega$  is bounded. By the first part of the proof, for a.e.  $\omega$ , for all  $N$ ,  $|V_N^\omega| \leq C_2(\varepsilon_0, K)$ .

**Part 4: (b)  $\Rightarrow$  (a):** Next suppose that  $f$  is not relatively cohomologous to a constant. Recall that  $\sigma_0^2 = \int u^2 dm$ .

<sup>3</sup> Here we use the assumption that  $(\Omega, \mathcal{F}, m)$  is  $\sigma$ -finite. Fubini's theorem may be false otherwise.

We claim that  $\sigma_0^2 > 0$ , and deduce from the first part of the proof that  $V_N^\omega \rightarrow +\infty$  a.e.

Assume by way of contradiction that  $\sigma_0^2 = 0$ , then  $u(\omega) = 0$  a.e., whence for a.e.  $\omega$ , for every  $n$ , almost every position  $n$  hexagon of  $X^\omega$  has balance zero. Applying the gradient lemma to  $X^\omega$ , we find bounded functions  $g_n^\omega$  and constants  $c_n^\omega$  such that

$$f_n^\omega(X_n^\omega, X_{n+1}^\omega) = g_n^\omega(X_n^\omega) - g_{n+1}^\omega(X_{n+1}^\omega) + c_n^\omega \text{ a.s.}$$

The issue is to show that  $g_n^\omega, c_n^\omega$  can be given the form  $g_n^\omega(x) = a(T^n \omega, x)$  and  $c_n^\omega = c(T^n \omega)$  where  $a(\cdot, \cdot), c(\cdot)$  are measurable.

This is indeed the case, because the proof of the gradient lemma shows that we can take

$$\begin{aligned} c_n^\omega &= \mathbb{E}^\omega[f_{n-2}^\omega(X_{n-2}^\omega, X_{n-1}^\omega)] \\ g_n^\omega(z) &= \mathbb{E}\left(f_{n-2}^\omega(X_{n-2}^\omega, X_{n-1}^\omega) + f_{n-1}^\omega(X_{n-1}^\omega, X_n^\omega) \middle| X_n^\omega = z\right). \end{aligned}$$

So  $c_n^\omega = c(T^n \omega)$  and  $g_n^\omega(z) = a(T^n \omega, z)$  for

$$\begin{aligned} c(\omega) &:= \int_{\mathfrak{S}^2} \mu_{T^{-2}\omega}(x) \mu_{T^{-1}\omega}(dy) p(T^{-2}\omega, x, y) f(T^{-2}\omega, x, y). \\ a(\omega, z) &:= \int_{\mathfrak{S}^3} \mu_{T^{-2}\omega}(dx) \mu_{T^{-1}\omega}(dy) p(T^{-2}\omega, x, y) p(T^{-1}\omega, y, z) \times \\ &\quad \frac{f(T^{-2}\omega, x, y) + f(T^{-1}\omega, y, z)}{\int_{\mathfrak{S}^2} \mu_{T^{-2}\omega}(dx) \mu_{T^{-1}\omega}(dy) [p(T^{-2}\omega, x, y) p(T^{-1}\omega, y, z)]}. \end{aligned}$$

These are measurable functions, and our standing assumptions imply that they are bounded.

We see that  $f$  is relatively cohomologous to a constant in contradiction to our assumption. So  $\sigma_0^2 > 0$ , whence by the first part of the proof  $V_N^\omega$  tends to infinity.  $\square$

### 8.3.3 The local limit theorem

In this section, we prove Theorems 8.2, 8.3, 8.6 and 8.7 on the local limit theorem for Markov chains in random environment. We need the following lemmas:

**Lemma 8.12** *Suppose  $\Omega$  is a Borel space,  $\mathfrak{S}$  is a separable metric space, and  $\psi : \Omega \times \mathfrak{S} \rightarrow \mathbb{R}$  is a Borel function such that for every  $\omega \in \Omega$ ,  $\psi(\omega, \cdot)$  is continuous on  $\mathfrak{S}$  and positive somewhere. Then there exists a Borel measurable  $x : \Omega \rightarrow \mathfrak{S}$  such that  $\psi(\omega, x(\omega)) > 0$ .*

*Proof.* Fix a countable dense set  $\{x_i\} \subset \mathfrak{S}$ . Our assumptions on  $\psi$  imply that for every  $\omega$  there exists an  $i$  such that  $\psi(\omega, x_i) > 0$ . So

$$i(\omega) := \min\{i \in \mathbb{N} : \psi(\omega, x_i) > 0\}$$

is well-defined and Borel measurable. Take  $x(\omega) := x_{i(\omega)}$ .  $\square$

**Lemma 8.13** *If  $W_1, W_2$  are two independent random variables such that for some  $a, t \in \mathbb{R}$ ,  $W_1 + W_2 \in a + t\mathbb{Z}$  with full probability, then  $a = a_1 + a_2$  where  $W_1 \in a_1 + t\mathbb{Z}$ ,  $W_2 \in a_2 + t\mathbb{Z}$  with full probability.*

*Proof.* Without loss of generality  $a = 0$ ,  $t = 2\pi$ . Then

$$|\mathbb{E}(e^{iW_1})| \cdot |\mathbb{E}(e^{iW_2})| = |\mathbb{E}(e^{i(W_1+W_2)})| = 1,$$

whence  $|\mathbb{E}(e^{iW_k})| = 1$  ( $k = 1, 2$ ). Choose  $a_k$  such that  $\mathbb{E}(e^{i(W_k - a_k)}) = 1$ , then  $\mathbb{E}(\cos(W_k - a_k)) = 1$ , whence  $W_k - a_k \in 2\pi\mathbb{Z}$  almost surely. Necessarily  $a_1 + a_2 \in 2\pi\mathbb{Z}$ , and there is no problem in adjusting  $a_1$  to get that the sum zero.  $\square$

**Proof of Theorems 8.2 and 8.6 on the non-lattice case.** Theorems 8.2 and 8.6 provide the LLT for Markov chains in random environment with finite and infinite noise process, under the assumption that  $f$  is not relatively cohomologous to a coset of  $t\mathbb{Z}$  with  $t \neq 0$ .

In this case,  $f$  is also not relatively cohomologous to a constant, and by Theorems 8.1 and 8.5,  $V_N^\omega \rightarrow \infty$  as  $N \rightarrow \infty$ . Moreover, if the noise process  $(\Omega, \mathcal{F}, m, T)$  satisfies  $m(\Omega) = 1$ , then  $\exists \sigma^2 > 0$  s.t.  $V_N^\omega \sim N\sigma^2$ .

To prove the theorems it is sufficient to show that for a.e.  $\omega$ ,  $G_{ess}(X^\omega, f^\omega) = \mathbb{R}$ , as this will imply the LLT by the general results of Chapter 4.

Assume by way of contradiction that  $G_{ess}(X^\omega, f^\omega) \neq \mathbb{R}$  on a set of  $\omega$ 's of positive measure. By Proposition 8.9,  $G_{ess}(X^\omega, f^\omega) = G_{ess}$  a.e. where  $G_{ess} = \{0\}$  or  $\frac{2\pi}{t}\mathbb{Z}$  with  $t \neq 0$ . The first possibility cannot happen, because it implies that  $f^\omega$  is center-tight, whence  $V_N^\omega = O(1)$ , whereas  $V_N^\omega \rightarrow \infty$ . So there exists  $t \neq 0$  such that  $G_{ess}(X^\omega, f^\omega) = (2\pi/t)\mathbb{Z}$  a.s., and  $H_\omega := H(X^\omega, f^\omega) = t\mathbb{Z}$  a.e.

By the reduction lemma, for every  $\omega$  s.t.  $H_\omega = t\mathbb{Z}$  there are measurable functions  $g_n^\omega(x)$ ,  $h_n^\omega(x, y)$  with  $\sum \text{Var}[h_n^\omega] < \infty$ , and constants  $c_n^\omega$  such that

$$\exp[i(t(f_n^\omega(x, y) - g_n^\omega(x) + g_{n+1}^\omega(y) + h_n^\omega(x, y) - c_n^\omega))] = 1$$

a.s. with respect to the distribution of  $(X_n^\omega, X_{n+1}^\omega)$ . So  $e^{it(f(\omega, x, y) + h_n^\omega(x, y))} = \lambda_n^\omega \frac{a_n^\omega(x)}{a_{n+1}^\omega(y)}$ , where  $\lambda_n^\omega = e^{itc_n^\omega}$ ,  $a_n^\omega(x) = e^{itg_n^\omega(x)}$ .

But now we run into a problem: Our proof of the reduction lemma does not provide  $g_n^\omega$  and  $c_n^\omega$  of the form  $c_n^\omega = c(T^n \omega)$ ,  $a_n^\omega = a(T^n \omega, x)$  with  $c(\cdot), a(\cdot, \cdot)$  measurable, and we need to show that  $h_n^\omega = 0$ .

To this end we use the following additional structure: For a.e.  $\omega$ ,  $H_\omega = t\mathbb{Z}$  so  $\sum d(T^n \omega, t)^2 < \infty$   $\mu$ -almost everywhere. By the ergodic theorem, this can only happen if  $d(\omega, t) = 0$  almost everywhere. Hence

$$\Gamma\left(Z_1^\omega, \frac{Z_2^\omega}{Y_2^\omega}, \frac{Y_3^\omega}{X_3^\omega}, X_4^\omega\right) \in \frac{2\pi}{t}\mathbb{Z} \text{ a.e. in } \text{Hex}(\omega) \text{ for } m\text{-a.e. } \omega. \quad (8.3.4)$$

Recall the ladder process  $\underline{L}_n^\omega = (Z_{n-2}^\omega, Y_{n-2}^\omega, X_n^\omega)$  associated with  $\{X_n^\omega\}$ . Let  $\mathbb{P}^\omega$  denote its distribution, and define as in the proof of the reduction lemma,

$$\begin{aligned} H^\omega(\underline{L}_n^\omega, \underline{L}_{n+1}^\omega) &:= \Gamma\left(Z_{n-2}^\omega, \frac{Z_{n-1}^\omega}{Y_{n-1}^\omega}, \frac{Y_n^\omega}{X_n^\omega}, X_{n+1}^\omega\right) \\ &\Gamma\left(Z_1^\omega, \frac{Z_2^\omega}{Y_2^\omega}, \frac{Z_3^\omega}{X_3^\omega}, \frac{Y_4^\omega}{X_4^\omega}, X_5^\omega\right) \stackrel{!}{=} H^\omega(\underline{L}_3^\omega, \underline{L}_4^\omega) + H^\omega(\underline{L}_4^\omega, \underline{L}_5^\omega) \end{aligned}$$

The last definition requires justification because the RHS seems at first sight to depend on  $Y_3^\omega$ . In fact it does not. To see this observe that the last expression is the balance of the octagon obtained by stacking  $\left(Z_2^\omega, \frac{Z_3^\omega}{Y_3^\omega}, \frac{Y_4^\omega}{X_4^\omega}, X_5^\omega\right)$  on top of  $\left(Z_1^\omega, \frac{Z_2^\omega}{Y_2^\omega}, \frac{Y_3^\omega}{X_3^\omega}, X_4^\omega\right)$  and removing the common edge  $(Z_2^\omega, Y_3^\omega, X_4^\omega)$  which “cancels out.”

CLAIM 1. Let  $\mathbb{P}^\omega$  denote the distribution of  $\{\underline{L}_n^\omega\}$ , then there exist measurable functions  $\zeta_1(\omega), \zeta_2(\omega) \in \mathfrak{S}$  such that for a.e.  $\omega$

$$\Gamma \left( \zeta_1(\omega), \frac{\zeta_2(\omega)}{Y_2^\omega}, \frac{\zeta_1(\omega)}{X_3^\omega}, \frac{Y_4^\omega}{X_4^\omega}, X_5^\omega \right) \in \frac{2\pi}{t} \mathbb{Z} \quad ; \quad \mathbb{P}^\omega \left( \cdot \middle| \begin{array}{l} Z_3^\omega = \zeta_1(\omega) \\ Z_2^\omega = \zeta_2(\omega) \\ Z_1^\omega = \zeta_1(\omega) \end{array} \right) \text{-a.e.}$$

*Proof.* By (8.3.4),  $\Gamma \in \frac{2\pi}{t} \mathbb{Z}$  with full  $\mathbb{P}^\omega$ -probability, for a.e.  $\omega$ . The point is to obtain this a.s. with respect to the conditional measures.

Suppose first that  $\mathfrak{S}$  is countable, then for fixed  $\omega$ , the  $\mathbb{P}^\omega$ -distribution of  $(\underline{L}_3^\omega, \underline{L}_2^\omega, \underline{L}_3^\omega)$  is purely atomic, and  $\Gamma \in \frac{2\pi}{t} \mathbb{Z}$  for every octagon with positive  $\mathbb{P}^\omega$ -probability. So the claim holds for any pair  $(\zeta_1(\omega), \zeta_2(\omega)) \in \mathfrak{S}$  such that

$$\mathbb{P}^\omega[(Z_1^\omega, Z_2^\omega, Z_3^\omega) = (\zeta_1(\omega), \zeta_2(\omega), \zeta_1(\omega))] > 0.$$

Such pairs exist by the ellipticity assumption. Since  $\mathfrak{S}$  is countable there is no problem to choose such  $(\zeta_1, \zeta_2)$  measurably.

Now suppose  $\mathfrak{S}$  is uncountable but with the continuity property (C). By Fubini's theorem and (8.3.4), for a.e.  $\omega \in \Omega$ , for a.e.  $(\zeta_1, \zeta_2, \zeta_3)$  with respect to the distribution  $(\zeta_1, \zeta_2, \zeta_3) \sim (Z_1^\omega, Z_2^\omega, Z_3^\omega)$ ,

$$\mathbb{E}_{\mathbb{P}^\omega} \left( \left| e^{(2\pi i/t)\Gamma(Z_1^\omega, \frac{Z_2^\omega}{Y_2^\omega}, \frac{Z_3^\omega}{X_3^\omega}, \frac{Y_4^\omega}{X_4^\omega}, X_5^\omega)} - 1 \right|^2 \middle| \begin{array}{l} Z_1^\omega = \zeta_1 \\ Z_2^\omega = \zeta_2 \\ Z_3^\omega = \zeta_3 \end{array} \right) = 0. \quad (8.3.5)$$

By the Markov property, this conditional expectation has canonical interpretation for *every*  $(\omega, \zeta_1, \zeta_2, \zeta_3)$  in the set

$$A = \{(\omega, a, b, c) : p(\omega, a, b) p(T\omega, b, c) > 0\}.$$

By assumption (C2),  $A$  is open. By assumption (C1), every open subset of  $A$  has positive measure with respect to the measure  $\int \mathbb{P}^\omega m(d\omega)$ . By assumption (C2), the left-hand-side of (8.3.5) depends continuously on  $(\omega, \zeta_1, \zeta_2, \zeta_3)$ . Therefore (8.3.5) is true for *all*  $(\zeta_1, \zeta_2, \zeta_3) \in A$ .

Thus to prove the claim it remains to construct measurable functions  $\zeta_1(\omega), \zeta_2(\omega)$  such that  $(\omega, \zeta_1(\omega), \zeta_2(\omega), \zeta_1(\omega)) \in A$  for all  $\omega$ .

By the ellipticity condition  $\int_{\mathfrak{S}} p(\omega, a, \zeta) p(T\omega, \zeta, a) \mu_{T\omega}(d\zeta) > \varepsilon_0$ , so for every  $\omega$  there are  $(\zeta_1, \zeta_2)$  s.t.

$$\psi(\omega, (\zeta_1, \zeta_2)) := p(\omega, \zeta_1, \zeta_2) p(T\omega, \zeta_2, \zeta_1) > 0.$$

By Lemma 8.12 it is possible to choose measurable  $\zeta_1(\omega), \zeta_2(\omega)$  like this. Claim 1 is proved.

Given  $\omega \in \Omega$  and  $a, b \in \mathfrak{S}$ , construct the bridge distribution  $\mathbb{P}_{ab}^\omega(E) = \mathbb{P}^\omega(Y_2^\omega \in E | Z_1^\omega = a, X_3^\omega = b)$  as in §1.2.3.

CLAIM 2. For a.e.  $\omega$ , for a.e.  $(\xi_3, \xi_4, \xi_5)$  sampled from the joint distribution of  $(X_3^\omega, X_4^\omega, X_5^\omega)$ , the random variables

$$\begin{aligned} W_3^\omega &:= f(\omega, \zeta_1(\omega), Y_2) + f(T\omega, Y_2, \xi_3), \quad Y_2 \sim \mathbb{P}_{\zeta_1, \xi_3}^\omega \\ W_5^{T^2\omega} &:= f(T^2\omega, \zeta_1(\omega), Y_4) + f(T^3\omega, Y_4, \xi_5), \quad Y_4 \sim \mathbb{P}_{\zeta_1, \xi_5}^{T^2\omega} \end{aligned}$$

are purely atomic, and belong to some coset of  $\frac{2\pi}{t} \mathbb{Z}$  with full probability. (These cosets could be different.)

*Proof.* By choice of  $\zeta_i(\omega)$  and Fubini's theorem, for a.e.  $(\xi_3, \xi_4, \xi_5) \sim (X_3^\omega, X_4^\omega, X_5^\omega)$ ,

$$\Gamma\left(\zeta_1(\omega), \frac{\zeta_2(\omega)}{Y_2^\omega}, \frac{\zeta_1(\omega)}{\xi_3}, \frac{Y_4^\omega}{\xi_4}, \xi_5\right) \in \frac{2\pi}{t}\mathbb{Z} \quad \mathbb{P}^\omega\left(\cdot \middle| \begin{array}{l} Z_0^\omega = \zeta_1 \\ Z_2^\omega = \zeta_2 \\ Z_1^\omega = \zeta_1 \\ X_3^\omega = \xi_3 \\ X_4^\omega = \xi_4 \\ X_5^\omega = \xi_5 \end{array}\right) \text{-a.e.}$$

Notice that  $\Gamma\left(\zeta_1(\omega), \frac{\zeta_2(\omega)}{Y_2^\omega}, \frac{\zeta_1(\omega)}{\xi_3}, \frac{Y_4^\omega}{\xi_4}, \xi_5\right)$  is equal to the *independent* difference of  $W_3^\omega$  and  $W_5^{T^2\omega}$ , plus a constant which only depends on  $\omega$ . The claim now follows from Lemma 8.13.

CLAIM 3. *Given  $\omega$  and  $(\xi_3, \xi_4, \xi_5)$  as in claim 2, let*

$$g(\omega, \xi_3) := \begin{cases} \text{the smallest positive atom of } W_3^\omega \text{ if } \exists \text{ positive atoms,} \\ \text{otherwise, the largest non-positive atom of } W_3^\omega \end{cases} \quad c(\omega) := -f(\omega, \zeta_1(\omega), \zeta_2(\omega)) - f(T\omega, \zeta_2(\omega), \zeta_1(\omega)).$$

These functions are well-defined, measurable, and

$$[f(T^2\omega, \xi_3, \xi_4) + f(T^3\omega, \xi_4, \xi_5)] + g(\omega, \xi_3) - g(T^2\omega, \xi_5) + c(\omega) \in \frac{2\pi}{t}\mathbb{Z} \quad (8.3.6)$$

for  $\mu$ -a.e.  $\omega$ , for a.e.  $(\xi_3, \xi_4, \xi_5) \sim (X_3^\omega, X_4^\omega, X_5^\omega)$ .

*Proof.* The function  $g(\omega, \xi_3)$  is well-defined for a.e.  $\omega$  because of claim 2. To see that it is measurable, we note that  $(\omega, \xi_3) \mapsto \mathbb{P}(W_3^\omega \in (a, b))$  are measurable, and

$$\begin{aligned} [g(\omega, \xi_3) > a] &= \{(\omega, \xi_3) : \mathbb{P}(0 < W_3^\omega \leq a) = 0, \mathbb{P}(W_3^\omega > a) \neq 0\} & (a > 0) \\ [g(\omega, \xi_3) > a] &= \{(\omega, \xi_3) : \mathbb{P}(W_3^\omega > a) \neq 0\} & (a \leq 0) \end{aligned}$$

are measurable. The measurability of  $c(\omega)$  is clear.

Equation (8.3.6) holds because the left-hand-side of (8.3.6) is, up to a sign, an atom of the random variable

$$\Gamma\left(\zeta_1(\omega), \frac{\zeta_2(\omega)}{Y_2^\omega}, \frac{\zeta_1(\omega)}{\xi_3}, \frac{Y_4^\omega}{\xi_4}, \xi_5\right), \quad (\underline{L}_3^\omega, \underline{L}_4^\omega) \sim \mathbb{P}^\omega\left(\cdot \middle| \begin{array}{l} Z_0^\omega = \zeta_1(\omega) \\ Z_2^\omega = \zeta_2(\omega) \\ Z_1^\omega = \zeta_1(\omega) \\ X_3^\omega = \xi_3 \\ X_4^\omega = \xi_4 \\ X_5^\omega = \xi_5 \end{array}\right)$$

and we chose  $(\zeta_1(\omega), \zeta_2(\omega))$  so that this random variable takes values in  $\frac{2\pi}{t}\mathbb{Z}$  a.s.

Claim 3 gives us measurable functions  $a(\omega, x) := \exp(-itg(T^{-2}\omega, x))$  and  $\lambda(\omega) := \exp(-itc(T^{-2}\omega))$  such that

$$e^{it[f(\omega, X_1^\omega, X_2^\omega) + f(T\omega, X_2^\omega, X_3^\omega)]} = \lambda(\omega) \frac{a(\omega, X_1^\omega)}{a(T^2\omega, X_3^\omega)}.$$

Multiplying both sides of the equation by  $e^{it[f^\omega - f^{\omega \circ T}]}$  gives

$$e^{2itf(\omega, X_1^\omega, X_2^\omega)} = \lambda(\omega) \frac{b(\omega, X_1^\omega, X_2^\omega)}{b(T\omega, X_2^\omega, X_3^\omega)},$$

where  $b(\omega, x, y) := a(\omega, x)a(T\omega, y)e^{itf(\omega, x, y)}$ .

This is not quite a relative cohomology to a coset of  $(\pi/t)\mathbb{Z}$ , because  $b(\omega, x, y)$  seems to depend not just on  $x$  but also on  $y$ . In fact there is a bounded measurable function  $\beta(\omega, x)$  such that

$$b(\omega, X_1^\omega, X_2^\omega) = \beta(\omega, X_1^\omega) \quad \mathbb{P} - \text{almost everywhere,}$$

where  $\mathbb{P}$  is given by (8.2.1). This can be seen as follows. Rearrange terms to see that

$$b(T\omega, X_2^\omega, X_3^\omega) = \lambda(\omega) e^{-2itf(\omega, X_1^\omega, X_2^\omega)} b(\omega, X_1^\omega, X_2^\omega).$$

By the Markov property of  $X^\omega$ , for fixed  $\omega$ , the left-hand-side and the right-hand-side of this equation are conditionally independent given  $X_2^\omega$ . Two independent random variables which are equal, must be constant. So for  $m$ -a.e.  $\omega$ ,

$$b(T\omega, X_2^\omega, X_3^\omega) = \mathbb{E}^\omega(b(T\omega, X_2^\omega, X_3^\omega) | X_2^\omega).$$

Setting

$$\beta(\omega, X_1^\omega) := \mathbb{E}^\omega(b(\omega, X_1^\omega, X_2^\omega) | X_1^\omega)$$

and using stationarity to shift indices where needed, we find that

$$b(\omega, X_1^\omega, X_2^\omega) = \beta(\omega, X_1^\omega) \mathbb{P}-a.e., \quad b(T\omega, X_2^\omega, X_3^\omega) = \beta(T\omega, X_2^\omega) \mathbb{P}-a.e.$$

Hence

$$e^{2itf(\omega, X_1^\omega, X_2^\omega)} = \lambda(\omega) \frac{\beta(\omega, X_1^\omega)}{\beta(T\omega, X_2^\omega)} \mathbb{P}\text{-a.e.}$$

So  $f$  is relatively cohomologous to a coset of  $\frac{\pi}{t}\mathbb{Z}$ .

We obtained a contradiction to our assumptions. This contradiction shows that  $G_{ess}(X^\omega, f^\omega) = \mathbb{R}$  for a.e.  $\omega$ . The local limit theorem now follows from Theorem 4.1, applied to  $(X^\omega, f^\omega)$ , since Theorem 8.1 gives the a.s. asymptotic  $V_N^\omega \sim N\sigma^2$  for some  $\sigma^2 > 0$  independent of  $\omega$ .  $\square$

**Proofs of Theorem 8.3 and 8.7 on the lattice case.** Theorems 8.3 and 8.7 provide the LLT for Markov chains in random environment with finite and infinite noise processes for integer valued additive functionals, under the assumption that  $f$  is not relatively cohomologous to a coset of  $t\mathbb{Z}$  with  $t \neq 1$ .

The proof is similar to the proof in the non-lattice case, except that now to check irreducibility we need to show that  $H_\omega = \mathbb{Z}$  almost surely. Since  $f$  is integer valued,  $1 \in H_\omega$ , so if this is not the case then necessarily  $H_\omega = t\mathbb{Z}$  for  $t = \frac{1}{n}$  and  $n \in \mathbb{N}$ . Now repeat the proof of Theorems 8.2 and 8.6 verbatim.  $\square$

### 8.3.4 Log-moment generating functions and rate functions

We prove Theorem 8.4 on the a.s. convergence of the log-moment generating functions of  $(X^\omega, f^\omega)$  and their Legendre transforms. Suppose  $f$  is an essentially bounded additive functional on a MCRE with a finite noise space  $(\Omega, \mathcal{B}, m, T)$ . Without loss of generality,  $m(\Omega) = 1$ .

**Part (1): Convergence of log-moment generating functions:** We are asked to show that for a.e.  $\omega$ ,  $\mathcal{G}_N^\omega(\xi) := \frac{1}{N} \log \mathbb{E}(e^{\xi S_N^\omega})$  converge pointwise on  $\mathbb{R}$ . To do this we recall three facts from chapter 6:

FACT 1: Given  $\xi \in \mathbb{R}$ , for every  $\omega \in \Omega$  there are unique numbers  $\bar{p}_n(\xi, \omega) \in \mathbb{R}$  and unique non-negative functions  $\bar{h}_n(\cdot, \xi, \omega) \in L^\infty(\mathfrak{S}, \mathcal{B}(\mathfrak{S}), \mu_{T^{n-1}\omega})$  such that  $\int_{\mathfrak{S}} \bar{h}_n(x, \xi, \omega) \mu_{T^{n-1}\omega}(dx) = 1$  for all  $n \geq 1$ , and

$$\int_{\mathfrak{S}} e^{\xi f(T^n \omega, x, y)} \frac{\bar{h}_{n+1}(y, \xi, \omega)}{e^{\bar{p}_n(\xi, \omega)} \bar{h}_n(x, \xi, \omega)} \pi(T^n \omega, x, dy) = 1. \quad (8.3.7)$$

Furthermore,  $\bar{p}_n(\xi, \omega) = \bar{p}(\xi, T^n \omega)$  for all  $n$ , where and  $\bar{p}(\xi, \omega)$  is measurable.

*Proof.* The existence and uniqueness of  $\bar{h}_n, \bar{p}_n$  follows from Lemma 6.7, applied to  $(X^\omega, f^\omega)$  with  $a_n = 0$ . Writing (8.3.7) first for  $(n, \omega)$  and then for  $(n-1, T\omega)$ , and then invoking uniqueness, we find that  $\bar{p}_n(\xi, \omega) = \bar{p}_{n-1}(\xi, T\omega)$ . So

$$\bar{p}_n(\xi, \omega) = \bar{p}_{n-1}(\xi, T\omega) = \dots = \bar{p}_1(\xi, T^{n-1}\omega) = \bar{p}(\xi, T^n\omega),$$

where  $\bar{p}(\xi, \omega) := \bar{p}_1(\xi, T^{-1}\omega)$ . The proof of Lemma 6.7 represents  $h_n(x, \xi, \omega)$  as a limit of expressions which are measurable in  $(x, \xi, \omega)$ , so  $(x, \xi, \omega) \mapsto \bar{h}_n(x, \xi, \omega)$  is measurable. By (8.3.7),  $(\omega, \xi) \mapsto \bar{p}(\xi, \omega)$  is measurable.

FACT 2: Let  $K := \text{ess sup } |f|$  and let  $\varepsilon_0$  denote a uniform ellipticity bound for  $X^\omega$ . For every  $R > 0$  there exists a constant  $C(\varepsilon_0, K, R)$  such that  $|\bar{p}(\xi, \omega)| \leq C(\varepsilon_0, K, R)$  for all  $\omega \in \Omega$  and  $|\xi| \leq R$ .

*Proof.* See the proof of Lemma 6.8.

FACT 3: Let  $\bar{P}_N(\xi, \omega) := \sum_{k=1}^N \bar{p}(\xi, T^k \omega)$ , then for a.e.  $\omega \in \Omega$ ,

$$\mathcal{G}_N^\omega(\xi) = \left( \frac{V_n^\omega}{N} \right) \left[ \frac{\bar{P}_N(\xi, \omega)}{V_N^\omega} + O\left( \frac{1}{V_N^\omega} \right) \right] \text{ uniformly on compact subsets of } \xi \in \mathbb{R}.$$

*Proof.* It is convenient to work with  $\mathcal{F}_N^\omega(\xi) := \frac{1}{V_N^\omega} \log \mathbb{E}(e^{\xi S_N^\omega}) \equiv (N/V_N^\omega) \mathcal{G}_N^\omega(\xi)$ . Let  $P_N(\xi, \omega) := \bar{P}_N(\xi, \omega) + (\mathbb{E}(S_N^\omega) - \frac{d}{d\xi} \big|_{\xi=0} \bar{P}'_N(0, \omega)) \xi$ . For each  $\omega \in \Omega$  such that  $V_N^\omega \rightarrow \infty$ ,

- (1)  $\frac{d}{d\xi} \big|_{\xi=0} \bar{P}'_N(0, \omega)$  exists, by Lemma 6.10.
- (2)  $|P_N(\xi, \omega) - \bar{P}_N(\xi, \omega)| = O(1)$  uniformly on compact subsets of  $\xi \in \mathbb{R}$ , by Lemmas 6.11 and 6.12.
- (3)  $|\mathcal{F}_N^\omega(\xi) - P_N(\xi)/V_N^\omega| = O(1/V_N^\omega)$  uniformly on compact subsets of  $\xi \in \mathbb{R}$ , by Lemma 6.12.

Fact 3 follows.

We can now prove the a.s. convergence of  $\mathcal{F}_N^\omega(\xi)$ . By the assumptions of the theorem,  $f$  is not relatively cohomologous to a constant. Therefore, by Theorem 8.1, there exists  $\sigma^2 > 0$  such that  $V_N^\omega \sim \sigma^2 N$  as  $N \rightarrow \infty$  for a.e.  $\omega$ .

Fix a countable dense set  $\{\xi_1, \xi_2, \dots\} \subset \mathbb{R}$ . For each  $i$ ,  $\omega \mapsto \bar{p}(\xi_i, \omega)$  is bounded and measurable. So for a.e.  $\omega$ ,

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathcal{G}_N^\omega(\xi_i) &= \sigma^2 \lim_{N \rightarrow \infty} \frac{1}{V_N^\omega} \sum_{k=1}^N \bar{p}(\xi_i, T^k \omega) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \bar{p}(\xi_i, T^k \omega) \\ &= \int_{\Omega} \bar{p}(\xi_i, \omega) m(d\omega), \text{ by the Birkhoff ergodic theorem.} \end{aligned}$$

This shows that for all  $i$  there exists  $\mathcal{G}(\xi_i) \in \mathbb{R}$  such that  $\lim_{N \rightarrow \infty} \mathcal{G}_N^\omega(\xi_i) = \mathcal{G}(\xi_i)$  for a.e.  $\omega$ . Let  $\Omega'$  denote the set of full measure of  $\omega$  where this holds for all  $i \in \mathbb{N}$ .

Fix  $\omega \in \Omega'$ , then the functions  $\xi \mapsto \mathcal{F}_N^\omega(\xi)$  are equicontinuous on compacts, because if  $K := \text{ess sup } |f|$ , then

$$|(\mathcal{F}_N^\omega)'(\xi)| \leq \left| \frac{|\xi| \mathbb{E}(|S_N^\omega| e^{\xi S_N^\omega})}{V_N^\omega \mathbb{E}(e^{\xi S_N^\omega})} \right| = \frac{|\xi| K N}{V_N^\omega} = O(|\xi|).$$

Therefore for a.e.  $\omega$ , the functions  $\xi \mapsto \mathcal{G}_N^\omega(\xi)$  are equicontinuous on compacts.

Recall that if a sequence of functions  $\varphi_n(\xi)$  which is equicontinuous on compacts converges on a dense subset of  $\mathbb{R}$ , then  $\varphi_n(\xi)$  converges for all  $\xi \in \mathbb{R}$ . Moreover, the limit is continuous. So there is a continuous function  $\mathcal{F}^\omega(\xi)$  such that

$$\lim_{N \rightarrow \infty} \mathcal{G}_N^\omega(\xi) = \mathcal{F}^\omega(\xi) \text{ for all } \xi \in \mathbb{R}, \omega \in \Omega'.$$

In fact  $\mathcal{F}^\omega(\xi)$  does not depend on  $\omega$ , because by virtue of continuity,

$$\mathcal{F}^\omega(\xi) = \lim_{k \rightarrow \infty} \mathcal{F}^\omega(\xi_{i_k}) = \lim_{k \rightarrow \infty} \mathcal{G}(\xi_{i_k}), \text{ whenever } \xi_{i_k} \xrightarrow{k \rightarrow \infty} \xi.$$

We are therefore free to write  $\mathcal{F}^\omega(\xi) = \mathcal{F}(\xi)$ .

It remains to show that  $\mathcal{F}(\xi)$  is differentiable and strictly convex on  $\mathbb{R}$ . Fix  $\omega \in \Omega'$ . Applying Theorem 6.1 to  $(X^\omega, f^\omega)$  we find that for every  $R > 0$  there is a  $C = C(R)$  such that  $C^{-1} \leq (\mathcal{F}_N^\omega)'' \leq C$  on  $[-R, R]$ . This implies that  $\mathcal{F}$  is differentiable and strictly convex on  $(-R, R)$  because of the following general lemma:

**Lemma 8.14** Suppose  $\varphi_n : \mathbb{R} \rightarrow \mathbb{R}$  are twice differentiable convex functions such that  $C^{-1} \leq \varphi_n'' \leq C$  with  $C > 0$ , on  $(-R, R)$ . If  $\varphi_n \xrightarrow{N \rightarrow \infty} \varphi$  pointwise on  $(-R, R)$ , then  $\varphi$  is continuously differentiable and strictly convex on  $(-R, R)$ .

*Proof.* A pointwise limit of convex functions is convex, and convex functions have one sided derivatives. Let  $\varphi'_\pm(\xi)$  denote the one-sided derivatives of at  $\xi$ .

DIFFERENTIABILITY: For all  $|\xi| < R$ ,

$$\begin{aligned} |\varphi'_+(\xi) - \varphi'_-(\xi)| &= \lim_{h \rightarrow 0^+} \left| \frac{\varphi(\xi + h) - \varphi(\xi)}{h} - \frac{\varphi(\xi - h) - \varphi(\xi)}{h} \right| \\ &= \lim_{h \rightarrow 0^+} \lim_{n \rightarrow \infty} \left| \frac{\varphi_n(\xi + h) - \varphi_n(\xi)}{h} - \frac{\varphi_n(\xi - h) - \varphi_n(\xi)}{h} \right| \\ &= \lim_{h \rightarrow 0^+} \lim_{n \rightarrow \infty} |\varphi'_n(\xi_n) - \varphi'_n(\eta_n)| \text{ for some } \xi_n, \eta_n \in (\xi - h, \xi + h) \\ &\leq \lim_{h \rightarrow 0^+} \lim_{n \rightarrow \infty} 2Ch = 0, \text{ because } |\varphi_n''| \leq C \text{ on a neighborhood of } \xi. \end{aligned}$$

We find that  $\varphi'_+(\xi) = \varphi'_-(\xi)$ , whence  $\varphi$  is differentiable at  $\xi$ .

STRICT CONVEXITY: Suppose  $-R < \xi < \eta < R$ , then

$$\begin{aligned} \varphi'(\eta) - \varphi'(\xi) &= \varphi'_+(\eta) - \varphi'_-(\xi) = \lim_{h \rightarrow 0^+} \frac{\varphi(\eta + h) - \varphi(\eta)}{h} - \frac{\varphi(\xi - h) - \varphi(\xi)}{h} \\ &= \lim_{h \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{\varphi_n(\eta + h) - \varphi_n(\eta)}{h} - \frac{\varphi_n(\xi - h) - \varphi_n(\xi)}{h} \\ &= \lim_{h \rightarrow 0^+} \varphi'_n(\eta_n) - \varphi'_n(\xi_n) \text{ for some } \xi_n \in [\xi - h, \xi], \eta_n \in [\eta, \eta + h] \\ &\geq \liminf_{n \rightarrow \infty} C^{-1} |\eta_n - \xi_n| \geq C^{-1} (\eta - \xi), \text{ because } \varphi_n'' > C^{-1} \text{ on } (-R, R). \end{aligned}$$

It follows that  $\varphi'$  is strictly increasing on  $(-R, R)$ .

THE DERIVATIVE IS CONTINUOUS: The same calculation as before shows that if  $-R < \xi < \eta < R$ , then  $|\varphi'(\eta) - \varphi'(\xi)| \leq C|\xi - \eta|$ , whence  $\varphi'$  is (Lipschitz) continuous on  $(-R, R)$ .  $\square$

**Part (2): Convergence of  $\mathbb{E}(S_N^\omega)/N$ :** We need the following standard fact.

**Lemma 8.15** Suppose  $\varphi_n(\xi), \varphi(\xi)$  are finite, convex, and differentiable on  $(-R, R)$ . If  $\varphi_n(\xi) \xrightarrow[n \rightarrow \infty]{} \varphi(\xi)$  on  $(-R, R)$ , then  $\varphi'_n(\xi) \xrightarrow[n \rightarrow \infty]{} \varphi'(\xi)$  on  $(-R, R)$ .

*Proof.* Fix  $\xi \in (-R, R)$ . By convexity, for every  $h > 0$  sufficiently small,

$$\frac{\varphi_n(\xi) - \varphi_n(\xi - h)}{h} \leq \varphi'_n(\xi) \leq \frac{\varphi_n(\xi + h) - \varphi_n(\xi)}{h}. \quad (8.3.8)$$

To see this note that the LHS is at most  $(\varphi_n)'_-(\xi)$ , the RHS is at least  $(\varphi_n)'_+(\xi)$ , and both one-sided derivatives equal  $\varphi'_n(\xi)$ .

Passing to the limit  $n \rightarrow \infty$  in (8.3.8), we find that

$$\limsup \varphi'_n(\xi), \liminf \varphi'_n(\xi) \in \left[ \frac{\varphi(\xi) - \varphi(\xi - h)}{h}, \frac{\varphi(\xi + h) - \varphi(\xi)}{h} \right].$$

We now invoke the differentiability of  $\varphi$ , pass to the limit  $h \rightarrow 0^+$ , and discover that  $\limsup \varphi'_n(\xi)$  and  $\liminf \varphi'_n(\xi)$  are both equal to  $\varphi'(\xi)$ .  $\square$

For a.e.  $\omega$ ,  $(V_N^\omega/N)\mathcal{F}_N^\omega(\xi) \equiv \mathcal{G}_N(\xi, \omega) \xrightarrow[N \rightarrow \infty]{} \mathcal{F}(\xi)$ . So by the lemma

$$(V_N^\omega/N) \frac{d}{d\xi} \Big|_{\xi=0} \mathcal{F}_N^\omega(\xi, \omega) \xrightarrow[N \rightarrow \infty]{} \mathcal{F}'(0).$$

A calculation shows that the derivative equals  $\mathbb{E}(S_N^\omega)/V_N^\omega$ . So  $\mathbb{E}(S_N^\omega)/N \rightarrow \mathcal{F}'(0)$ .

**Part (3): Convergence of Legendre transforms.** Again, the proof is based on a general property of convex functions.

**Lemma 8.16** Suppose  $\varphi_n(\xi), \varphi(\xi)$  are finite, strictly convex, continuously differentiable functions on  $\mathbb{R}$ , s.t.  $\varphi_n(\xi) \rightarrow \varphi(\xi)$  for all  $\xi \in \mathbb{R}$ . Let  $\varphi'(\pm\infty) := \lim_{\xi \rightarrow \pm\infty} \varphi'(\xi)$ . Let  $\varphi_n^*, \varphi^*$  denote the Legendre transforms of  $\varphi_n, \varphi$ . For all  $\eta \in (\varphi'(-\infty), \varphi'(+\infty))$ ,  $\varphi_n^*(\eta)$  is well-defined for all  $n$  sufficiently large, and  $\varphi_n^*(\eta) \rightarrow \varphi^*(\eta)$ .

*Proof.* Fix  $\eta \in (\varphi'(-\infty), \varphi'(+\infty))$ . By assumption,  $\varphi'$  is continuous and strictly increasing. Therefore, there exists  $\xi$  such that  $\varphi'(\eta) = \xi$ .

Fix  $\varepsilon > 0$  and  $\xi_1 < \xi < \xi_2$  such that  $|\xi_1 - \xi_2| < \varepsilon$ . Then  $\varphi'(\xi_1) < \eta < \varphi'(\xi_2)$ . By Lemma 8.15,  $\varphi'_n(\xi_i) \rightarrow \varphi'(\xi_i)$ , and therefore there exists  $N$  such that for all  $n > N$ ,

$$\varphi'(\xi_1) - 1 < \varphi'_n(\xi_1) < \eta < \varphi'_n(\xi_2) < \varphi'(\xi_2) + 1.$$

Since  $\eta \in (\varphi'_n(\xi_1), \varphi'_n(\xi_2))$  and  $\varphi'_n$  is continuous and strictly increasing, there exists a unique  $\xi_n \in (\xi_1, \xi_2)$  so that  $\varphi'_n(\xi_n) = \eta$ . So  $\varphi_n^*(\eta)$  is well-defined, and

$$\varphi_n^*(\eta) = \xi_n \eta - \varphi_n(\xi_n).$$

Similarly,  $\varphi^*(\eta) = \xi \eta - \varphi(\xi)$ .

We now estimate the distance between  $\varphi_n^*(\eta)$  and  $\varphi^*(\eta)$ . Recall first that for all  $n > N$ ,  $\varphi'(\xi_1) - 1 < \varphi'_n(\xi_1) < \varphi'_n(\xi_2) < \varphi'(\xi_2) + 1$ . Let

$$M := \max\{|\varphi'(\xi_1) - 1|, |\varphi'(\xi_2) + 1|\},$$

then  $|\varphi'_n| \leq M$  on  $(\xi_1, \xi_2)$  for all  $n > N$ . Consequently,

$$\begin{aligned} |\varphi_n^*(\eta) - \varphi^*(\eta)| &\leq |\xi_n - \xi| \cdot |\eta| + |\varphi_n(\xi_n) - \varphi(\xi)| \\ &\leq |\xi_1 - \xi_2| \cdot |\eta| + |\varphi_n(\xi_n) - \varphi_n(\xi)| + |\varphi_n(\xi) - \varphi(\xi)| \\ &\leq \varepsilon |\eta| + M |\xi_n - \xi| + |\varphi_n(\xi) - \varphi(\xi)| \leq \varepsilon (M + |\eta|) + o(1), \text{ as } n \rightarrow \infty, \end{aligned}$$

because  $\varphi_n(\xi) \rightarrow \varphi(\xi)$ ,  $\xi, \xi_n \in (\xi_1, \xi_2)$ , and  $|\xi_1 - \xi_2| \leq \varepsilon$ . Since  $\varepsilon$  is arbitrary, we have that  $\varphi_n^*(\eta) \rightarrow \varphi^*(\eta)$ .  $\square$

**Part (4): Properties of  $\mathcal{I}(\eta)$ .** Fix  $\omega$  such that  $\varphi_N(\xi) := \frac{1}{N} \log \mathbb{E}(e^{\xi S_N^\omega})$  converges pointwise to  $\mathcal{F}$ . By Lemma 8.16,  $\varphi_N^*$  converges pointwise to  $\mathcal{I}$ . Since  $\varphi_N''$  is uniformly bounded away from zero and infinity on compacts (see the first part of the proof),  $(\varphi_N^*)''$  is uniformly bounded away from zero and infinity on compacts. Hence by Lemma 8.14

$\mathcal{I} = \lim \varphi_N^*$  is strictly convex and continuously differentiable.

By Lemma 8.15,  $(\varphi_N^*)'(\eta) \xrightarrow[N \rightarrow \infty]{} \mathcal{I}'(\eta)$  for all  $\eta$  in the interior of the range of  $\varphi'$ , and  $\varphi_N(\xi) \xrightarrow[N \rightarrow \infty]{} \mathcal{F}'(\xi)$  for all  $\xi \in \mathbb{R}$ . The convergence is uniform on compacts, because  $(\varphi_N^*)'', \varphi_N''$  are bounded on compacts.

It is easy to verify that  $\varphi_N$  is twice differentiable. Therefore by Lemma 6.17,  $\varphi_N^*$  is twice differentiable and  $(\varphi_N^*)'(\varphi_N'(\xi)) = \xi$  for all  $\xi$ . Passing to the limit as  $N \rightarrow \infty$  we obtain the important identity  $\mathcal{I}'(\mathcal{F}'(\xi)) = \xi$  for all  $\xi \in \mathbb{R}$ .

One consequence of the identity  $\mathcal{I}'(\mathcal{F}'(\xi)) = \xi$  is that  $\mathcal{I}'(\mathcal{F}'(0)) = 0$ , so  $\eta = \mathcal{F}'(0)$  is a critical point of  $\mathcal{I}(\cdot)$ . By strict convexity,  $\mathcal{I}$  attains its global minimum at  $\mathcal{F}'(0)$ . The value there is zero:

$$\mathcal{I}(\mathcal{F}'(0)) = 0 \cdot \mathcal{F}'(0) - \mathcal{F}(0) = 0.$$

We conclude that  $\mathcal{I}(\eta) = 0$  when  $\eta = \mathcal{F}'(0)$ , and  $\mathcal{I}(\eta) > 0$  for  $\eta \neq \mathcal{F}'(0)$ .

Another consequence of the identity  $\mathcal{I}'(\mathcal{F}'(\xi)) = \xi$  (and the fact that  $\mathcal{F}'$  is increasing) is that  $\mathcal{I}'(\xi) \xrightarrow[\xi \rightarrow \pm\infty]{} \pm\infty$ , and therefore  $\mathcal{I}$  has compact level sets.  $\square$

**Part (5): Large deviation threshold.** We prove the identity for  $\mathcal{F}'(\infty)$ , the identity for  $\mathcal{F}'(-\infty)$  follows by replacing  $f \rightarrow -f$ .

*Step I.*  $c_+ \geq \mathcal{F}'(+\infty)$ .

*Proof.* Given  $\eta \in (\mathcal{F}'(-\infty), \mathcal{F}(+\infty))$  choose

$$\mathcal{F}'(-\infty) < \eta^- < \eta < \eta^+ < \mathcal{F}(+\infty).$$

Take  $\xi^\pm$  be s.t.  $\mathcal{F}'(\xi^\pm) = \eta^\pm$ . By Lemma 8.15,  $\lim_{N \rightarrow \infty} \mathcal{F}'_N(\xi^\pm) = \eta^\pm$ . Hence for large  $N$   $\mathcal{F}'_N(\xi^-) \leq \eta \leq \mathcal{F}_N(\xi^+)$  and so  $\eta$  is reachable.

*Step II.*  $c_+ \leq \mathcal{F}'(+\infty)$ .

*Proof.* Take  $\eta > c_+$ . If  $\eta \in \mathcal{C}_R$  for some  $R$  we would have that for some  $R$ , for all  $N$  large enough  $\mathcal{F}'_N(R) \geq \eta$  (see Lemma 6.12(5)). However by Lemma 8.15,  $\lim_{N \rightarrow \infty} \mathcal{F}_N(R) = \mathcal{F}'(R) < \mathcal{F}'(+\infty) < \eta$  contradicting our assumption that  $\eta$  is reachable.

*Step III.* Denote  $\mathcal{S}_N(\omega) = \text{ess sup } S_N^\omega$ . Then the limit  $s^+ := \lim_{N \rightarrow \infty} \frac{\mathcal{S}_N(\omega)}{N}$  exists and is independent of  $\omega$  with probability one.

*Proof.* By our ellipticity assumption

$$\mathcal{S}_{N+M}(\omega) \leq \mathcal{S}_N(\omega) + S_M(T^N \omega) - 4K.$$

Thus the sequence  $\mathcal{T}_N(\omega) = \mathcal{S}_N(\omega) - 4K$  is subadditive. Since  $\mathcal{S}_N(\omega) \geq -KN$  the Subadditive Ergodic Theorem implies that the limit  $\lim_{N \rightarrow \infty} \frac{\mathcal{S}_N(\omega)}{N} = \lim_{N \rightarrow \infty} \frac{\mathcal{T}_N(\omega)}{N}$  exists and is independent of  $\omega$  with probability one.

*Step IV.*  $\mathfrak{c}_+ \leq \mathfrak{s}^+$  because for each  $\varepsilon > 0$  we have that with probability one for large  $N$ ,  $\mathbb{P}^\omega(S_N \geq (\mathfrak{s}^+ + \varepsilon)N) = 0$ .

*Step V.*  $\mathfrak{c}_+ \geq \mathfrak{s}^+$

*Proof.* Fix  $\varepsilon > 0$ . By Step III for each sufficiently large  $N_0$  there exists  $\gamma_{\varepsilon, N_0} > 0$  and a set  $\Omega_{\varepsilon, N_0}$  s.t.  $m(\Omega_{\varepsilon, N_0}) \geq 1 - \varepsilon$  and for all  $\omega \in \Omega_{\varepsilon, N_0}$ , for  $\mu_\omega$ -a.e.  $x \in \mathfrak{S}$ ,

$$\mathbb{P}^\omega(S_{N_0} \geq (\mathfrak{s}^+ - \varepsilon)N_0 | X_1 = x) \geq \gamma_{\varepsilon, N_0}. \quad (8.3.9)$$

Given  $M$  let  $j_1(\omega) < j_2(\omega) < \dots < j_{n_M(\omega)}(\omega)$  be all the times  $1 \leq j < M$  when  $T^{jN_0}(\omega) \in \Omega_{\varepsilon, N_0}$ , then

$$\mathbb{P}^\omega(\mathcal{S}_{N_0 M} \geq n_M(\mathfrak{s}^+ - \varepsilon)N_0 - (M - n_M)N_0 K) \geq \gamma_{\varepsilon, N_0}^{N_0 M}.$$

(To see this, estimate conditional probabilities of this event given  $X_{j_1}^\omega, \dots, X_{j_{n_M}}^\omega$  using (8.3.9), and take expectation over  $X_{j_1}^\omega, \dots, X_{j_{n_M}}^\omega$ .)

By the Ergodic Theorem, for a.e.  $\omega$  there is a limit

$$\beta(\omega) = \lim_{M \rightarrow \infty} \frac{n_M(\omega)}{M} \quad \text{and} \quad \int \beta(\omega) dm = m(\Omega_{\varepsilon, N_0}).$$

So for large  $M$ , and on a set  $\overline{\Omega}_\varepsilon$  of positive measure,  $n_M/M > 1 - 2\varepsilon$  whence

$$n_M(\mathfrak{s}^+ - \varepsilon)N_0 - (M - n_M)N_0 K \geq [(1 - 2\varepsilon)(\mathfrak{s}^+ - \varepsilon) - 2\varepsilon K]N_0 M.$$

Now Theorem 6.7(c) shows that on  $\overline{\Omega}_\varepsilon$ ,  $\mathfrak{c}^+(\omega) \geq (1 - 2\varepsilon)(\mathfrak{s}^+ - \varepsilon) - 2\varepsilon K$ .

By steps I and II above,  $\mathfrak{c}^+$  actually does not depend on  $\omega$  (in fact, using Theorem 6.7, it is easy to verify directly that  $\mathfrak{c}^+$  is  $T$ -invariant, and therefore by ergodicity, constant). we get that

$$\mathfrak{c}^+(\omega) \geq (1 - 2\varepsilon)(\mathfrak{s}^+ - \varepsilon) - 2\varepsilon K$$

almost surely. Since  $\varepsilon$  is arbitrary the result follows.  $\square$

## 8.4 Notes and references

Markov chains in random environment (MCRE) should not be confused with “random walks in random environment” (RWRE). In the RWRE model, the transition kernel at time  $n$  depends on the *position* of random walk at time  $n$ , i.e.  $\pi_n(x, dy) = \pi(S_n, x, dy)$ . In a MCRE, the transition kernel at time  $n$  depends on the *noise* at time  $n$ , i.e.  $\pi_n(x, dy) = \pi(T^n \omega, x, dy)$ . For a recent treatment of the LLT for RWRE, see [21] and references therein.

Markov chains in random environment were introduced by Cogburn [24]. The setup is a particular case of a “random dynamical system.” For a fixed realization of noise, a Markov chain in random environment reduces to an inhomogeneous Markov chain, and a random dynamical system reduces to a “sequential” (aka “time-dependent” or “non-autonomous”) dynamical system.

Various authors considered probabilistic limit theorems in these contexts. Limit theorems for Markov chains in random environment are given in Cogburn [25], Seppäläinen [138], Kifer [79], [80] and Hafouta & Kifer [69, chapters 6,7],[68]. Results for random dynamical systems can be found in Kifer [80], Conze, Le Borgne & Roger [27], Denker & Gordin [37], Aimino, Nicol & Vaienti [8], Nicol, Török & Vaienti [111], and Dragičević, Froyland & González-Tokman [53]. For limit theorems for sequential dynamical systems, see Bakhtin [12], Conze & Raugi [28], Haydn, Nicol & Török [72], Korepanov, Kosloff & Melbourne [86], and Hafouta [66, 67].

If we set the noise process to be the identity on the one point space, then the LLT in this chapter reduce to LLT for homogeneous stationary Markov chains, as in Theorem 7.1. For more general LLT for homogeneous Markov chains, see Nagaev [109], Guivarc'h & Hardy [65].

The results of this chapter are all essentially known in the case  $T$  preserves a finite measure. Theorem 8.1 was proved in the more general setup of random dynamical systems by Kifer [80],[78]. Theorems 8.2 and 8.3 are close to the (earlier) results of Dragičević, Froyland & González-Tokman [53], and Hafouta & Kifer [69, chapter 7, Theorem 7.1.5]. The main difference is in the irreducibility assumptions. Our condition of non-relative cohomology to a coset is replaced in [69] by what these authors call the “lattice” and “non-lattice” cases (this is not the same as our terminology). In the paper [53], the non-cohomology condition is replaced by a condition on the decay of the norms of certain perturbed characteristic function operators, and a connection to a non-cohomology condition is made under additional assumptions.

The results for infinite noise processes seem to be new. The reason we can also treat this case, is that the LLT we provide in this work do not require any assumptions on the rate of growth of  $V_N$ , and they also work when it grows sub-linearly.

## References

1. J. Aaronson and M. Denker. Distributional limits for hyperbolic infinite volume geodesic flows. *Tr. Mat. Inst. Steklova*, 216(Din. Sist. i Smezhnye Vopr.):181–192, 1997.
2. Jon Aaronson. *An introduction to infinite ergodic theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.
3. Jon Aaronson and Manfred Denker. The Poincaré series of  $\mathbf{C} \setminus \mathbf{Z}$ . *Ergodic Theory Dynam. Systems*, 19(1):1–20, 1999.
4. Jon Aaronson and Manfred Denker. Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stoch. Dyn.*, 1(2):193–237, 2001.
5. Jon Aaronson and Manfred Denker. Group extensions of Gibbs-Markov maps. *Probab. Theory Related Fields*, 123(1):28–40, 2002.
6. Jon Aaronson, Hitoshi Nakada, Omri Sarig, and Rita Solomyak. Invariant measures and asymptotics for some skew products. *Israel J. Math.*, 128:93–134, 2002.
7. Jon Aaronson and Benjamin Weiss. Remarks on the tightness of cocycles. *Colloq. Math.*, 84/85(part 2):363–376, 2000. Dedicated to the memory of Anzelm Iwanik.
8. Romain Aimino, Matthew Nicol, and Sandro Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. *Probab. Theory Related Fields*, 162(1-2):233–274, 2015.
9. A. Avila, D. Dolgopyat, E. Duryev, and O. Sarig. The visits to zero of a random walk driven by an irrational rotation. *Israel J. Math.*, 207(2):653–717, 2015.
10. Martine Babillot and François Ledrappier. Lalley’s theorem on periodic orbits of hyperbolic flows. *Ergodic Theory Dynam. Systems*, 18(1):17–39, 1998.
11. R. R. Bahadur and R. Ranga Rao. On deviations of the sample mean. *Ann. Math. Statist.*, 31:1015–1027, 1960.
12. V. I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings. I. *Izv. Ross. Akad. Nauk Ser. Mat.*, 58(2):40–72, 1994.
13. Garrett Birkhoff. Extensions of Jentzsch’s theorem. *Trans. Amer. Math. Soc.*, 85:219–227, 1957.
14. David Blackwell and J. L. Hodges, Jr. The probability in the extreme tail of a convolution. *Ann. Math. Statist.*, 30:1113–1120, 1959.
15. Thomas Bogenschütz and Volker Mathias Gundlach. Ruelle’s transfer operator for random subshifts of finite type. *Ergodic Theory Dynam. Systems*, 15(3):413–447, 1995.
16. Richard C. Bradley. *Introduction to strong mixing conditions. Vol. 1,2,3*. Kendrick Press, Heber City, UT, 2007.
17. Leo Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.
18. Julien Brémont. Gibbs measures at temperature zero. *Nonlinearity*, 16(2):419–426, 2003.
19. Anne Broise. Transformations dilatantes de l’intervalle et théorèmes limites. *Astérisque*, (238):1–109, 1996. Études spectrales d’opérateurs de transfert et applications.
20. Henk Bruin, Mark Holland, and Matthew Nicol. Livšic regularity for Markov systems. *Ergodic Theory Dynam. Systems*, 25(6):1739–1765, 2005.
21. Fabienne Castell, Nadine Guillotin-Plantard, Françoise Pène, and Bruno Schapira. A local limit theorem for random walks in random scenery and on randomly oriented lattices. *Ann. Probab.*, 39(6):2079–2118, 2011.
22. Narasinga R. Chaganty and J. Sethuraman. Large deviation local limit theorems for arbitrary sequences of random variables. *Ann. Probab.*, 13(1):97–114, 1985.
23. Narasinga Rao Chaganty and Jayaram Sethuraman. Strong large deviation and local limit theorems. *Ann. Probab.*, 21(3):1671–1690, 1993.
24. Robert Cogburn. Markov chains in random environments: the case of Markovian environments. *Ann. Probab.*, 8(5):908–916, 1980.
25. Robert Cogburn. On the central limit theorem for Markov chains in random environments. *Ann. Probab.*, 19(2):587–604, 1991.
26. Gonzalo Contreras and Renato Iturriaga. *Global minimizers of autonomous Lagrangians*. 22º Colóquio Brasileiro de Matemática. [22nd Brazilian Mathematics Colloquium]. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999.
27. Jean-Pierre Conze, Stéphane Le Borgne, and Mikaël Roger. Central limit theorem for stationary products of toral automorphisms. *Discrete Contin. Dyn. Syst.*, 32(5):1597–1626, 2012.
28. Jean-Pierre Conze and Albert Raugi. Limit theorems for sequential expanding dynamical systems on  $[0, 1]$ . In *Ergodic theory and related fields*, volume 430 of *Contemp. Math.*, pages 89–121. Amer. Math. Soc., Providence, RI, 2007.
29. Jean-Pierre Conze and Albert Raugi. On the ergodic decomposition for a cocycle. *Colloq. Math.*, 117(1):121–156, 2009.
30. I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai. *Ergodic theory*, volume 245 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskiĭ.
31. H. Cramér. Sur une nouveau théorème-limite de la théorie des probabilités, 1938. *Reprinted in H. Cramér, Collected works*, A. Martin-Löf (Ed.), vol II, Springer, Berlin, 1994, pages 895–913.
32. R. de la Llave, J. M. Marco, and R. Moriyón. Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation. *Ann. of Math.* (2), 123(3):537–611, 1986.
33. A. de Moivre. *The doctrine of chances or, a method of calculating the probabilities of events in play*. New impression of the second edition, with additional material. Cass Library of Science Classics, No. 1. Frank Cass & Co., Ltd., London, 1967.
34. Amir Dembo and Ofer Zeitouni. Large deviations via parameter dependent change of measure, and an application to the lower tail of Gaussian processes. In *Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993)*, volume 36 of *Progr. Probab.*, pages 111–121. Birkhäuser, Basel, 1995.
35. Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

36. Mark Demers, Françoise Pene, and Ofer Zeitouni. Local limit theorem for randomly deforming billiards. *Preprint* (2019), *arXiv:1902.06850*.
37. Manfred Denker and Mikhail Gordin. The central limit theorem for random perturbations of rotations. *Probab. Theory Related Fields*, 111(1):1–16, 1998.
38. Manfred Denker and Xiaofei Zheng. On the local times of stationary processes with conditional local limit theorems. *Stochastic Process. Appl.*, 128(7):2448–2462, 2018.
39. Bernard Derrida and Tridib Sadhu. Large deviations conditioned on large deviations I: Markov chain and Langevin equation. *J. Stat. Phys.*, 176(4):773–805, 2019.
40. Yves Derriennic and Michael Lin. The central limit theorem for Markov chains with normal transition operators, started at a point. *Probab. Theory Related Fields*, 119(4):508–528, 2001.
41. Zach Dietz and Sunder Sethuraman. Large deviations for a class of nonhomogeneous Markov chains. *Ann. Appl. Probab.*, 15(1A):421–486, 2005.
42. Zach Dietz and Sunder Sethuraman. Occupation laws for some time-nonhomogeneous Markov chains. *Electron. J. Probab.*, 12:no. 23, 661–683, 2007.
43. J. Dieudonné. *Foundations of modern analysis*. Academic Press, New York-London, 1969. Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.
44. Péter Nándori Dmitry Dolgopyat. Infinite measure mixing for some mechanical systems. *Preprint* (2018), *arXiv:1812.01174*.
45. R. Dobrushin. Central limit theorem for non-stationary Markov chains. i, ii. *Theory of Probab. & Appl.*, 1:65–80, 329–383, 1956.
46. W. Doeblin. Le cas discontinu des probabilités en chaîne. *Publ. Fac. Sci. Univ. Masaryk (Brno)*, (236), 1937.
47. W. Doeblin. Sur les propriétés asymptotiques de mouvement régis par certains types de chaînes simples. *Bull. Math. Soc. Roum. Sci.*, 39(1):57–115, 1937.
48. Dmitry Dolgopyat. Prevalence of rapid mixing in hyperbolic flows. *Ergodic Theory Dynam. Systems*, 18(5):1097–1114, 1998.
49. Dmitry Dolgopyat. A local limit theorem for sums of independent random vectors. *Electron. J. Probab.*, 21:Paper No. 39, 15, 2016.
50. Dmitry Dolgopyat and Peter Nandori. On mixing and central limit theorem for hyperbolic flows. *Ergodic Theory Dynam. Systems*, 20(1):142–174, 2020.
51. Dmitry Dolgopyat, Domokos Szász, and Tamás Varjú. Recurrence properties of planar Lorentz process. *Duke Math. J.*, 142(2):241–281, 2008.
52. J. L. Doob. *Stochastic processes*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1990. Reprint of the 1953 original, A Wiley-Interscience Publication.
53. D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti. A spectral approach for quenched limit theorems for random expanding dynamical systems. *Comm. Math. Phys.*, 360(3):1121–1187, 2018.
54. Loc Dubois. Real cone contractions and analyticity properties of the characteristic exponents. *Nonlinearity*, 21(11):2519–2536, 2008.
55. Richard S. Ellis. *Entropy, large deviations, and statistical mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1985 original.
56. P. Ferrero and B. Schmitt. Produits aléatoires d’opérateurs matrices de transfert. *Probab. Theory Related Fields*, 79(2):227–248, 1988.
57. Sinai Ya. G. *Theory of phase transitions: rigorous results*, volume 108 of *International Series in Natural Philosophy*. Pergamon Press, Oxford-Elmsford, N.Y., 1982. Translated from the Russian by J. Fritz, A. Krámli, P. Major and D. Szász.
58. N. G. Gamkrelidze. On a local limit theorem for lattice random variables. *Teor. Verojatnost. i Primenen*, 9:733–736, 1964.
59. B. V. Gnedenko. On a local limit theorem of the theory of probability. *Uspehi Matem. Nauk (N. S.)*, 3(3(25)):187–194, 1948.
60. B. V. Gnedenko. On a local theorem for the region of normal attraction of stable laws. *Doklady Akad. Nauk SSSR (N.S.)*, 66:325–326, 1949.
61. B. V. Gnedenko and A. N. Kolmogorov. *Limit distributions for sums of independent random variables*. Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1954. Translated and annotated by K. L. Chung. With an Appendix by J. L. Doob.
62. M. I. Gordin. The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR*, 188:739–741, 1969.
63. M. I. Gordin and B. A. Lifšic. Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR*, 239(4):766–767, 1978.
64. Walter Helbig Gottschalk and Gustav Arnold Hedlund. *Topological dynamics*. American Mathematical Society Colloquium Publications, Vol. 36. American Mathematical Society, Providence, R. I., 1955.
65. Y. Guivarc’h and J. Hardy. Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d’Anosov. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(1):73–98, 1988.
66. Yeor Hafouta. Limit theorems for some skew-products with mixing base maps. *Preprint* (2018), *arXiv:1808.00735v5*.
67. Yeor Hafouta. A sequential rpf theorem and its applications to limit theorems for time dependent dynamical systems and inhomogeneous markov chains. *Preprint* (2019), *arXiv:1903.04018v4*.
68. Yeor Hafouta and Yuri Kifer. A nonconventional local limit theorem. *J. Theoret. Probab.*, 29(4):1524–1553, 2016.
69. Yeor Hafouta and Yuri Kifer. *Nonconventional limit theorems and random dynamics*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2018.
70. J. Hajnal. Weak ergodicity in non-homogeneous Markov chains. *Proc. Cambridge Philos. Soc.*, 54:233–246, 1958.
71. P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980. Probability and Mathematical Statistics.
72. Nicolai Haydn, Matthew Nicol, Andrew Török, and Sandro Vaienti. Almost sure invariance principle for sequential and non-stationary dynamical systems. *Trans. Amer. Math. Soc.*, 369(8):5293–5316, 2017.
73. Hubert Hennion and Loïc Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, volume 1766 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.
74. I. A. Ibragimov and Yu. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing, Groningen, 1971. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman.

75. A. Katok and A. Kononenko. Cocycles' stability for partially hyperbolic systems. *Math. Res. Lett.*, 3(2):191–210, 1996.
76. Atsushi Katsuda and Toshikazu Sunada. Closed orbits in homology classes. *Inst. Hautes Études Sci. Publ. Math.*, (71):5–32, 1990.
77. A. Y. Khinchin. *Mathematical foundations of quantum statistics*. Dover Publications, Inc., Mineola, NY, english edition, 1998. Translated from the Russian by E. J. Kelly, Jr., M. D. Friedman, W. H. Furry and A. H. Halperin, Translation edited and with a preface by Irwin Shapiro.
78. Yuri Kifer. *Ergodic theory of random transformations*, volume 10 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1986.
79. Yuri Kifer. Perron-Frobenius theorem, large deviations, and random perturbations in random environments. *Math. Z.*, 222(4):677–698, 1996.
80. Yuri Kifer. Limit theorems for random transformations and processes in random environments. *Trans. Amer. Math. Soc.*, 350(4):1481–1518, 1998.
81. Yuri Kifer. Thermodynamic formalism for random transformations revisited. *Stoch. Dyn.*, 8(1):77–102, 2008.
82. C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.
83. A. N. Kolmogorov. A local limit theorem for classical Markov chains. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 13:281–300, 1949.
84. A. N. Kolmogorov and A. Ya. Khinchin. Über konvergenz von reihen, deren glieder durch den zufall bestimmt weden. *Mat. Sb.*, 32:668–667, 1925.
85. Tomasz Komorowski, Claudio Landim, and Stefano Olla. *Fluctuations in Markov processes*, volume 345 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2012. Time symmetry and martingale approximation.
86. A. Korepanov, Z. Kosloff, and I. Melbourne. Martingale-coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 35(4):859–885, 2018.
87. Steven P. Lalley. Closed geodesics in homology classes on surfaces of variable negative curvature. *Duke Math. J.*, 58(3):795–821, 1989.
88. Steven P. Lalley. Renewal theorems in symbolic dynamics, with applications to geodesic flows, non-Euclidean tessellations and their fractal limits. *Acta Math.*, 163(1-2):1–55, 1989.
89. Pierre-Simon Laplace. *Théorie analytique des probabilités. Vol. I.* Éditions Jacques Gabay, Paris, 1995. Introduction: *Essai philosophique sur les probabilités*. [Introduction: Philosophical essay on probabilities], Livre I: *Du calcul des fonctions génératrices*. [Book I: On the calculus of generating functions], Reprint of the 1819 fourth edition (Introduction) and the 1820 third edition (Book I).
90. Pierre-Simon Laplace. *Théorie analytique des probabilités. Vol. II.* Éditions Jacques Gabay, Paris, 1995. Livre II: *Théorie générale des probabilités*. [Book II: General probability theory], Suppléments. [Supplements], Reprint of the 1820 third edition (Book II) and of the 1816, 1818, 1820 and 1825 originals (Supplements).
91. François Ledrappier and Omri Sarig. Unique ergodicity for non-uniquely ergodic horocycle flows. *Discrete Contin. Dyn. Syst.*, 16(2):411–433, 2006.
92. François Ledrappier and Omri Sarig. Fluctuations of ergodic sums for horocycle flows on  $\mathbb{Z}^d$ -covers of finite volume surfaces. *Discrete Contin. Dyn. Syst.*, 22(1-2):247–325, 2008.
93. M. Lemańczyk. Analytic nonregular cocycles over irrational rotations. *Comment. Math. Univ. Carolin.*, 36(4):727–735, 1995.
94. V. P. Leonov. On the dispersion of time means of a stationary stochastic process. *Teor. Verojatnost. i Primenen.*, 6:93–101, 1961.
95. J. W. Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.*, 15(1):211–225, 1922.
96. Ju. V. Linnik. Limit theorems for the sums of independent variables taking into account the large deviations. I. *Teor. Verojatnost. i Primenen.*, 6:145–163, 1961.
97. A. N. Livšic. Certain properties of the homology of  $Y$ -systems. *Mat. Zametki*, 10:555–564, 1971.
98. A. N. Livšic. Cohomology of dynamical systems. *Izv. Akad. Nauk SSSR Ser. Mat.*, 36:1296–1320, 1972.
99. A.M. Lyapunov. Sur une proposition de la théorie des probabilités. *Bull. de l'Academie Imperiale des Sci. de St. Petersbourg*, 13(4):359–386, 1900.
100. A.A. Markov. Extension of the law of large numbers to dependent events (russian). *Bull. Soc. Phys. Math. Kazan*, 15(2):135–156, 1906.
101. Przemysław Matuła. A note on the almost sure convergence of sums of negatively dependent random variables. *Statist. Probab. Lett.*, 15(3):209–213, 1992.
102. Michael Maxwell and Michael Woodroffe. Central limit theorems for additive functionals of Markov chains. *Ann. Probab.*, 28(2):713–724, 2000.
103. D. L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Probability*, 2:620–628, 1974.
104. J. Mineka and S. Silverman. A local limit theorem and recurrence conditions for sums of independent non-lattice random variables. *Ann. Math. Statist.*, 41:592–600, 1970.
105. Calvin C. Moore and Klaus Schmidt. Coboundaries and homomorphisms for nonsingular actions and a problem of H. Helson. *Proc. London Math. Soc. (3)*, 40(3):443–475, 1980.
106. D. A. Moskvin. A local limit theorem for large deviations in the case of differently distributed lattice summands. *Teor. Verojatnost. i Primenen.*, 17:716–722, 1972.
107. D. A. Moskvin, G. A. Freiman, and A. A. Judin. Structural theory of set summation, and local limit theorems for independent lattice random variables. *Teor. Verojatnost. i Primenen.*, 19:52–62, 1974.
108. A. B. Mukhin. Local limit theorems for lattice random variables. *Teor. Verojatnost. i Primenen.*, 36(4):660–674, 1991.
109. S. V. Nagaev. Some limit theorems for stationary Markov chains. *Teor. Verojatnost. i Primenen.*, 2:389–416, 1957.
110. S. V. Nagaev. Large deviations of sums of independent random variables. *Ann. Probab.*, 7(5):745–789, 1979.
111. Matthew Nicol, Andrew Török, and Sandro Vaienti. Central limit theorems for sequential and random intermittent dynamical systems. *Ergodic Theory Dynam. Systems*, 38(3):1127–1153, 2018.

112. William Parry and Mark Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, (187-188):268, 1990.
113. William Parry and Mark Pollicott. Skew products and Livsic theory. In *Representation theory, dynamical systems, and asymptotic combinatorics*, volume 217 of *Amer. Math. Soc. Transl. Ser. 2*, pages 139–165. Amer. Math. Soc., Providence, RI, 2006.
114. M. Peligrad. On the local limit theorems for lower psi-mixing markov chains. *Preprint*, page 12 pages, 2018.
115. Yuval Peres. Domains of analytic continuation for the top Lyapunov exponent. *Ann. Inst. H. Poincaré Probab. Statist.*, 28(1):131–148, 1992.
116. V. V. Petrov. Generalization of Cramér's limit theorem. *Uspehi Matem. Nauk (N.S.)*, 9(4(62)):195–202, 1954.
117. V. V. Petrov. *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Band 82.
118. D. Plachky and J. Steinebach. A theorem about probabilities of large deviations with an application to queuing theory. *Period. Math. Hungar.*, 6(4):343–345, 1975.
119. Mark Pollicott and Richard Sharp. Asymptotic expansions for closed orbits in homology classes. *Geom. Dedicata*, 87(1-3):123–160, 2001.
120. Mark Pollicott and Richard Sharp. Chebotarev-type theorems in homology classes. *Proc. Amer. Math. Soc.*, 135(12):3887–3894, 2007.
121. Georg Pólya. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Math. Ann.*, 84(1-2):149–160, 1921.
122. Yu. V. Prokhorov. On a local limit theorem for lattice distributions. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:535–538, 1954.
123. Firas Rassoul-Agha and Timo Seppäläinen. *A course on large deviations with an introduction to Gibbs measures*, volume 162 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.
124. Albert Raugi. Mesures invariantes ergodiques pour des produits gauches. *Bull. Soc. Math. France*, 135(2):247–258, 2007.
125. Alfréd Rényi. Contributions to the theory of independent random variables. *Acta Math. Acad. Sci. Hungar.*, 1:99–108, 1950.
126. Vol'fgang Rihter. Lokale Grenzwertsätze für grosse Abweichungen. *Teor. Veroyatnost i Primenen.*, 2:214–229, 1957.
127. J. Rousseau-Egele. Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux. *Ann. Probab.*, 11(3):772–788, 1983.
128. Yu. A. Rozanov. On a local limit theorem for lattice distributions. *Teor. Veroyatnost. i Primenen.*, 2:275–281, 1957.
129. D. Ruelle. Analyticity properties of the characteristic exponents of random matrix products. *Adv. in Math.*, 32(1):68–80, 1979.
130. David Ruelle. *Thermodynamic formalism*, volume 5 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1978. The mathematical structures of classical equilibrium statistical mechanics, With a foreword by Giovanni Gallavotti and Gian-Carlo Rota.
131. Hans Henrik Rugh. Cones and gauges in complex spaces: spectral gaps and complex Perron-Frobenius theory. *Ann. of Math.* (2), 171(3):1707–1752, 2010.
132. E. L. Rvacheva. On domains of attraction of multidimensional distributions. *L'vov Gos. Univ. Uč. Zap. Ser. Meh.-Mat.*, 29(6):5–44, 1954.
133. Omri Sarig. Invariant Radon measures for horocycle flows on abelian covers. *Invent. Math.*, 157(3):519–551, 2004.
134. L. Saulis. Asymptotic expansion for probabilities with large deviations. *Litovsk. Mat. Sb.*, 9:605–625, 1969.
135. L. Saulis and V. A. Statulevičius. *Limit theorems for large deviations*, volume 73 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1991. Translated and revised from the 1989 Russian original.
136. Klaus Schmidt. *Cocycles on ergodic transformation groups*. Macmillan Company of India, Ltd., Delhi, 1977. Macmillan Lectures in Mathematics, Vol. 1.
137. E. Seneta. On the historical development of the theory of finite inhomogeneous Markov chains. *Proc. Cambridge Philos. Soc.*, 74:507–513, 1973.
138. Timo Seppäläinen. Large deviations for Markov chains with random transitions. *Ann. Probab.*, 22(2):713–748, 1994.
139. S. Sethuraman and S. R. S. Varadhan. A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains. *Electron. J. Probab.*, 10:no. 36, 1221–1235, 2005.
140. Richard Sharp. Local limit theorems for free groups. *Math. Ann.*, 321(4):889–904, 2001.
141. L. A. Shepp. A local limit theorem. *Ann. Math. Statist.*, 35:419–423, 1964.
142. S. H. Sirazdinov. *Limit theorems for stationary Markov chains*. Izdat. Akad. Nauk Uzbeksko SSR, Taškent, 1955.
143. S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.
144. V. A. Statulevičius. On large deviations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 6:133–144, 1966.
145. V. Statuljavičius. Local limit theorems and asymptotic expansions for non-stationary Markov chains. *Litovsk. Mat. Sb.*, 1(1-2):231–314, 1961.
146. V. A. Statuljavičius. Limit theorems for densities and the asymptotic expansions for distributions of sums of independent random variables. *Teor. Veroyatnost. i Primene*, 10:645–659, 1965.
147. Charles Stone. A local limit theorem for nonlattice multi-dimensional distribution functions. *Ann. Math. Statist.*, 36:546–551, 1965.
148. P. Tchebycheff. Sur deux théorèmes relatifs aux probabilités. *Acta Math.*, 14(1):305–315, 1890.
149. S. R. S. Varadhan. *Large deviations*, volume 27 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2016.
150. A. D. Wentzell. *Limit theorems on large deviations for Markov stochastic processes*, volume 38 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990. Translated from the Russian.
151. E. F. Whittlesey. Analytic functions in Banach spaces. *Proc. Amer. Math. Soc.*, 16:1077–1083, 1965.
152. Amie Wilkinson. The cohomological equation for partially hyperbolic diffeomorphisms. *Astérisque*, (358):75–165, 2013.



# Index

*R*-admissible, 165  
additive functional  
  a.s. uniformly bounded, 24  
  center tight, 41  
  equicontinuous, 179  
  gradient, 41  
  homogeneous, 171  
  of a Markov array, 23  
  of a Markov chain, 23  
  reducible and irreducible, 65  
  uniformly bounded, 24  
  with summable variance, 42  
admissible, 165  
algebraic range  
  arrays, 67  
  Markov chains, 65  
balance, 31  
Bernoulli noise process, 186  
bridge distribution, 30  
center tightness  
  characterization, 43  
  definition, 41  
  obstruction to the LLT, 41  
center-tightness  
  and the co-range, 66  
Central Limit Theorem, 50  
change of measure, 23, 69  
  and large deviations, 138  
characteristic functions, 90  
circular mean, 72  
circular variance, 72  
CLT, 43  
co-range, 66, 67  
  for Markov chains in random environment, 191  
coboundary, 42  
cohomologous, 188  
contraction coefficient, 25  
cylinders, 186  
decay of correlations, 29  
Dirac measure, 21  
Dobrushin's Theorem, 43  
ellipticity condition, 24  
ellipticity constant, 24  
equicontinuous additive functional, 179  
ergodic, 185  
essential range, 67  
for Markov chains in random environment, 191  
Fourier transform, 126  
gradient, 41  
graininess constant, 106  
Green-Kubo formula, 172  
hereditary, 69  
  stably hereditary, 69  
hexagon, 30  
  admissible, 31  
  balance, 31  
  distribution, 31  
  measure, 31  
Hilbert's projective metric  
  contraction properties, 140  
  definition, 140  
homogeneous additive functional, 171  
homogeneous Markov chain, 20, 171  
independent  $\sigma$ -algebras, 37  
inhomogeneous Markov chain, 20  
Initial distribution  
  of Markov chains, 19  
initial distribution  
  of Markov array, 22  
irreducible, 65, 67  
joint distribution  
  of a Markov chain, 19  
Kolmogorov two-series theorem, 44  
ladder process, 34  
large deviations, 131  
lattice case, 65  
Legendre transform, 134  
  definition, 159  
  domain, 134  
Local limit theorem  
  mixing LLT, 89  
  mixing non-lattice, 105  
log-moment generating function, 132  
Markov array, 22  
  *N*-th row, 22  
  row lengths, 22  
  sub-arrays, 69  
  why study them, 23  
Markov chain, 19  
  with finite state spaces, 20  
  homogeneous, 20, 171

- inhomogeneous, 20
- initial distribution, 19
- joint distribution, 19
- random environment, 185
- state spaces, 19
- transition probabilities, 19
- with finite memory, 21
- Markov chains
  - Doeblin, 24
- Markov operator
  - contraction coefficient, 25
- Markov operators, 25
- Markov property, 19
- Markovian noise process, 186
- Martingale CLT, 50
- martingale difference array, 49
- measure preserving transformatios, 185
- mixing, 29
- mixing conditions
  - $\alpha$ -mixing, 37
  - $\phi$ -mixing, 37
  - $\psi$ -mixing, 37
  - $\rho$ -mixing, 37
- Mixing LLT, 89
- mixing uniform distribution mod  $t$ , 105
- Nagaev's identities, 90, 109
- Noise process, 185
  - Bernoulli, 186
  - Markovian, 186
  - null recurrent Markov, 187
  - Quasi-periodic, 186
- non-lattice case, 65
- normalized log-moment generating function, 132
- obstructions to the LLT
  - center tightness, 41
  - gradients, 41, 42
  - lattice case, 65
  - reducibility, 66
- one-step ellipticity condition, 179
- oscillation, 25
- perturbation operators, 90
- potential, 42
- power series on Banach spaces, 142
- Quasi-periodic noise process, 186
- range
  - algebraic, 65, 67
  - reduced, 65
- rate functions, 134
- reachable, 165
- real-analyticity for functions on Banach spaces, 142
- reduced range, 65
- reducible, 65
  - local limit theorems, 101
- reduction, 65
- Reduction lemma, 70
- regime
  - of large deviations, 131
  - local, 23
  - local deviations, 131
  - moderate, 132
  - of large deviations, 23, 131
  - universality, 132
- relatively cohomologous, 188
- restriction, 69
- row lengths, 22
- Rozanov's condition, 85
- simple random walk, 65
- stably hereditary, 69
- State spaces, 19
- state spaces
  - of Markov array, 22
- structure constants
  - and center tightness, 43
  - and growth of variance, 43
  - definition, 31
- sub-array, 69
- summable variance, 42
  - a.s. convergence of  $\{S_N\}_{N \geq 1}$ , 44
- symmetric multilinear function, 142
- Taylor series on Banach spaces, 142
- total variation distance, 25
- transfer function, 42
- transfer operator, 174
- transition kernel, 25
  - contraction coefficient, 25
- Transition probabilities
  - of Markov chains, 19
- transition probabilities
  - of a Markov array, 22
- uniform ellipticity
  - $\gamma$ -step, 36
  - $\gamma$ -step ellipticity condition, 24
  - and decay of correlations, 29
  - for Markov arrays, 24
  - for Markov chains, 24
  - one-step, 179
- uniform integrability, 50
- universality, 132
- variance
  - circular, 72
- Variance estimate, 43
- Zero-one law, 173