Convergence in total variation of an affine random recursion in $[0,p)^k$ to a uniform random vector

Claudio Asci
Dipartimento di Matematica e Geoscienze
Università degli Studi di Trieste
Via Valerio 12/1, 34127 Trieste, Italy
E-mail: casci@units.it

Abstract

We study the rate of convergence of the Markov chain $\mathbf{X}_{n+1}=A\mathbf{X}_n+\mathbf{B}_n$ (mod p), where A is an integer matrix with nonzero eigenvalues, p is real and positive, and $\{\mathbf{B}_n\}$ is a sequence of independent and identically distributed real random vectors. With some hypotheses on the law of $\mathbf{B}_n$, the sequence $\{\mathbf{X}_n\}$ converges to a random vector uniformly distributed in $[0,p)^k$. The rate of convergence is geometric and depends on A, p, k, and the distribution of $\mathbf{B}_n$. Moreover, if A has an eigenvalue that is a root of 1, then $n=O\left(p^2\right)$ steps are necessary to have $\mathbf{X}_n$ sampling from a nearly uniform law.

Key words and phrases. Continuous Markov chains; uniform ergodicity; generating random vectors; rate of convergence.

1 Introduction

In the mathematical literature, many authors studied the asymptotic behavior of the following affine random recursion on $\mathbf{Z}$:

$$X_{n+1}=aX_n+B_n \text{ (mod } p), \tag{1}$$

where $X_0 = x_0 \in \mathbf{Z}$, $a \in \mathbf{Z}^*$, $p \in \mathbf{N}^*$, and $\{B_n\}$ is a sequence of independent and identically distributed integer random variables. It is easy to prove that $\{X_n\}$ is a Markov chain that converges in law to the uniform distribution on $\mathbf{Z}_p$.

In the papers of Aldous and Diaconis (1986), Chung et al. (1987), and Hildebrand (1990 and 1993), the authors study the rate of convergence of the chain, by utilizing the Fourier analysis, a theory developed also in Diaconis (1988), Helleloid (2007), Rosenthal (1995), and Serre (1977).

When in the recursion (1) B_n is replaced by a fixed integer b, then the recursion is deterministic. The historical aim of this study has been the production of pseudorandom numbers on computers (see, for example, Knuth's book (1981)), for particular values of p.

In Asci (2001) and next in Hildebrand and McCollum (2008), there is the extension of the previous results to the higher-dimensional case, that is the study of the Markov chain on $\mathbf{Z}^k$ of the form

$$\mathbf{X}_{n+1} = A\mathbf{X}_n + \mathbf{B}_n \pmod{p},$$

where $A \in M_k(\mathbf{Z}) \cap GL_k(\mathbf{Q})$, $p \in \mathbf{N}^*$. However, in the last two works only some particular cases are considered. The general case is studied by Asci in two papers (Asci 2009a and Asci 2009b), where several results are obtained, depending on the size of the complex eigenvalues of A. In Asci 2009a, it is proved that, with some assumptions on p and on the distribution of $\mathbf{B}_n$, and without any assumptions on A, $n = O(p^2)$ steps are sufficient to have X_n sampling from a nearly uniform distribution on $\mathbf{Z}_p^k$. Moreover, if A has an eigenvalue of size 1, then $O\left(p^2\right)$ steps are also necessary. In Asci 2009b, it is shown that, if $|\lambda_i| \neq 1$ for all eigenvalues λ_i, then $n = O\left((\ln p)^2\right)$ steps are sufficient and $n = O(\ln p)$ steps are necessary to reach the uniform distribution.

In this paper the continuous $k-$ dimensional case is considered, that is the recursion on $\mathbf{R}^k$ defined by

$$\mathbf{X}_{n+1} = A\mathbf{X}_n + \mathbf{B}_n \pmod{p}, \tag{2}$$

where $\mathbf{X}_0 = \mathbf{x}_0 \in \mathbf{R}^k$, $A \in M_k(\mathbf{Z}) \cap GL_k(\mathbf{Q})$, $p \in \mathbf{R}^+$, and $\{\mathbf{B}_n\}$ is a sequence of independent and identically distributed real random vectors.

The aim of our work is to prove that, with some assumptions on the law of $\mathbf{B}_n$, the sequence $\{\mathbf{X}_n\}$ converges with geometric rate to a random vector uniformly distributed in

$[0,p)^k$ (Theorem 3.3). Moreover, we quantify the rate of convergence in terms of A, p, k, and the law of $\mathbf{B}_n$, and prove that, if A has an eigenvalue that is a root of 1, then $O\left(p^2\right)$ steps are necessary to achieve randomness (Theorem 3.4). We point out that our paper is mainly theoretical: we consider the recursion (2) in the most general context and provide some tools for further studies and applications.

In section 2, we recall some definitions and general results about homogeneous Markov chains, and we prove that the uniform distribution in $[0,p)^k$ is invariant for the chain $\{\mathbf{X}_n\}$. In section 3, we prove the irreducibility and the uniform ergodicity of the chain. In section 4, we expose some ideas for further study.

2 Preliminary results

In order to study the asymptotic behavior of the random sequence (2), we can suppose $\mathbf{X}_n \in [0,p)^k$. Indicate by $\mathcal{L}_{n,\mathbf{x}_0}$ and μ, $\forall\, n \in \mathbf{N}$, the laws of $\mathbf{X}_n$ and $\mathbf{B}_n$, respectively, and by $\mathbf{U}$ the random vector with uniform distribution on $[0,p)^k$; moreover, for any $n \in \mathbf{N}$, let $\mathbf{Y}_n$ be the random vector defined by the recursion $\mathbf{Y}_{n+1} = A\mathbf{Y}_n + \mathbf{B}_n$, $\mathbf{Y}_0 = \mathbf{x}_0$, and indicate by $\mu_{n,\mathbf{x}_0}$ its law.

Henceforth, $\forall\, \mathbf{x} \in \mathbf{R}^k$, we will indicate by $[\mathbf{x}]$ and $\{\mathbf{x}\}$ the vectors whose components are, respectively, the integer parts of the components of $\mathbf{x}$, and the fractional parts of the components of $\mathbf{x}$. Moreover, we will indicate by $\mathcal{B}^{(k)}$, by $\mathcal{B}([0,p)^k)$, and by $Leb^{(k)}$, respectively, the Borel σ-algebra on $\mathbf{R}^k$, the Borel σ-algebra on $[0,p)^k$, and the Lebesgue measure on $\mathbf{R}^k$. Finally, for any random vector $\mathbf{X}$, we will indicate by $\mathcal{L}_{\mathbf{X}}$ its law and by $f_{\mathbf{X}}$ its probability density function, if it exists.

Remark 2.1. $\forall\, n \in \mathbf{N}$, we have:

$$\mathbf{Y}_n = A^n\mathbf{x}_0 + \sum_{i=0}^{n-1} A^i \mathbf{B}_{n-1-i}.$$

Moreover, since $A \in M_k(\mathbf{Z})$, we have $\mathbf{X}_n = \mathbf{Y}_n \pmod p$; consequently, $\forall\, B \in \mathcal{B}([0,p)^k)$:

$$\mathcal{L}_{n,\mathbf{x}_0}(B) = \sum_{\mathbf{h}\in\mathbf{Z}^k} \mu_{n,\mathbf{x}_0}(B + p\mathbf{h}). \tag{3}$$

Define the variation distance between two probability measures φ and ψ on some measurable space $(E, \mathcal{E})$, in the following way:

$$||\varphi - \psi|| = \frac{1}{2} \sup_{f \in F} |E_\varphi[f] - E_\psi[f]| = \sup_{A \in \mathcal{E}} |\varphi(A) - \psi(A)|, \tag{4}$$

where $F \equiv \{f : E \longrightarrow \mathbf{C} : f$ is $\mathcal{E}$-measurable, $|f(x)| \leq 1 \; \forall x \in E\}$, and we say that a sequence $\{X_n\}$ of random variables converges in total variation to a random variable X if $\lim_{n \to +\infty} \|\mathcal{L}_{X_n} - \mathcal{L}_X\| = 0$.

Moreover, recall that, if $\{X_n\}$ is a homogeneous Markov chain on $(E, \mathcal{E})$, and if $P^n_{x_0}(A) = P(X_n \in A | X_0 = x_0)$, $\forall \, A \in \mathcal{E}$, then $\{X_n\}$ is called:

1. φ-irreducible, if there is a measure φ on $(E, \mathcal{E})$ such that, $\forall \, x_0 \in E$ and $\forall \, A \in \mathcal{E}$ such that $\varphi(A) > 0$, there is $n = n(x_0, A) \in \mathbf{N}^*$ such that $P^n_{x_0}(A) > 0$.

2. Uniformly ergodic, if

$$\lim_{n \to +\infty} \sup_{x_0 \in E} \|P^n_{x_0} - \pi\| = 0, \tag{5}$$

where π is a probability measure on $(E, \mathcal{E})$.

Our purpose is to prove (5) for $\pi = \mathcal{L}_\mathbf{U}$, in the case $(E, \mathcal{E}) = ([0, p)^k, \mathcal{B}([0, p)^k))$, $P^n_{x_0} = \mathcal{L}_{n, \mathbf{x}_0}$. We will use the following results, whose proofs can be found for example in Meyn and Tweedie (2005), Theorem 16.2.4, Proposition 10.1.1, and Theorem 10.0.1, respectively:

Theorem 2.2. If a homogeneous Markov chain $\{X_n\}$ on the state space $(E, \mathcal{E})$ verifies

$$P^m_{x_0}(A) \geq \rho_m(A), \; \forall \, x_0 \in E, \; \forall \, A \in \mathcal{E},$$

where $m \in \mathbf{N}$, ρ_m is a measure on $(E, \mathcal{E})$ (that is E is a small set: see the definition in Meyn and Tweedie (2005), page 109), then:

$$\|P^n_{x_0} - \pi\| \leq (1 - \rho_m(E))^{\left[\frac{n}{m}\right]}, \; \forall \, x_0 \in E, \; \forall \, n \in \mathbf{N}, \tag{6}$$

where π is a probability measure on $(E, \mathcal{E})$.

Proposition 2.3. If a homogeneous Markov chain $\{X_n\}$ on the state space $(E, \mathcal{E})$ is φ-irreducible and has an invariant probability measure, then $\{X_n\}$ is recurrent.

Theorem 2.4. If a homogeneous Markov chain $\{X_n\}$ on the state space $(E, \mathcal{E})$ is recurrent, then it has a unique invariant measure.

Let λ be a finite measure on $\big(\mathbf{R}^k, \mathcal{B}(\mathbf{R}^k)\big)$; define the Fourier transform $\widehat{\lambda} : \mathbf{R}^k \longrightarrow C$ by:

$$\widehat{\lambda}(\alpha) = \int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x}, \alpha\rangle\right) d\lambda(\mathbf{x}).$$

We have the following two results, whose proofs are similar to those of Lemmas 3.1 and 3.3 in Asci (2001).

Lemma 2.5. Suppose that $\mathbf{X}_0 = \mathbf{x}_0 \in \mathbf{R}^k$, $\alpha \in \mathbf{R}^k$. Then, $\forall\, n \in \mathbf{N}$:

$$\widehat{\mu_{n,\mathbf{x}_0}}(\alpha) = \exp\left(\frac{2\pi i}{p}\langle A^n \mathbf{x}_0, \alpha\rangle\right) \prod_{j=0}^{n-1} \widehat{\mu}\left({}^t A^j \alpha\right). \tag{7}$$

Moreover, if $\alpha \in \mathbf{Z}^k$:

$$\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha) = \widehat{\mu_{n,\mathbf{x}_0}}(\alpha); \tag{8}$$

$$|\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha)|^2 = \prod_{j=0}^{n-1}\left(\int\limits_{\mathbf{R}^k}\left(\int\limits_{\mathbf{R}^k} \cos\left(\frac{2\pi}{p}\left\langle \mathbf{x}-\mathbf{y}, {}^t A^j \alpha\right\rangle\right) d\mu(\mathbf{x})\right) d\mu(\mathbf{y})\right). \tag{9}$$

Proof. $\forall\, n \in \mathbf{N}$, we have:

$$\begin{aligned}
\widehat{\mu_{n,\mathbf{x}_0}}(\alpha) &= \int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x}, \alpha\rangle\right) d\mu_{n,\mathbf{x}_0}(\mathbf{x}) = E\left[\exp\left(\frac{2\pi i}{p}\langle \mathbf{Y}_n, \alpha\rangle\right)\right] \\
&= E\left[\exp\left(\frac{2\pi i}{p}\left\langle A^n \mathbf{x}_0 + \sum_{i=0}^{n-1} A^i \mathbf{B}_{n-1-i}, \alpha\right\rangle\right)\right] \\
&= \exp\left(\frac{2\pi i}{p}\langle A^n \mathbf{x}_0, \alpha\rangle\right) \prod_{j=0}^{n-1} \int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle A^j \mathbf{x}, \alpha\rangle\right) d\mu(\mathbf{x}) \\
&= \exp\left(\frac{2\pi i}{p}\langle A^n \mathbf{x}_0, \alpha\rangle\right) \prod_{j=0}^{n-1} \int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x},{}^t A^j \alpha\rangle\right) d\mu(\mathbf{x}) \\
&= \exp\left(\frac{2\pi i}{p}\langle A^n \mathbf{x}_0, \alpha\rangle\right) \prod_{j=0}^{n-1} \widehat{\mu}\left({}^t A^j \alpha\right).
\end{aligned}$$

Moreover, if $\alpha \in \mathbf{Z}^k$, from (3) we have:

$$\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha) = \int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x}, \alpha\rangle\right) d\mathcal{L}_{n,\mathbf{x}_0}(\mathbf{x})$$

$$= \sum_{\mathbf{h}\in\mathbf{Z}^k} \left(\int\limits_{[0,p)^k + p\mathbf{h}} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x} + p\mathbf{h}, \alpha\rangle\right) d\mu_{n,\mathbf{x}_0}(\mathbf{x}) \right)$$

$$= \int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x}, \alpha\rangle\right) d\mu_{n,\mathbf{x}_0}(\mathbf{x}) = \widehat{\mu_{n,\mathbf{x}_0}}(\alpha).$$

This implies

$$|\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha)|^2 = |\widehat{\mu_{n,\mathbf{x}_0}}(\alpha)|^2$$

$$= \prod_{j=0}^{n-1} \left(\widehat{\mu}\left({}^tA^j\alpha\right) \overline{\widehat{\mu}\left({}^tA^j\alpha\right)} \right)$$

$$= \prod_{j=0}^{n-1} \left(\int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x},^t A^j\alpha\rangle\right) d\mu(\mathbf{x}) \int\limits_{\mathbf{R^k}} \exp\left(-\frac{2\pi i}{p}\langle \mathbf{y},^t A^j\alpha\rangle\right) d\mu(\mathbf{y}) \right)$$

$$= \prod_{j=0}^{n-1} \left(\int\limits_{\mathbf{R}^k} \left(\int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x} - \mathbf{y},^t A^j\alpha\rangle\right) d\mu(\mathbf{x}) \right) d\mu(\mathbf{y}) \right)$$

$$= \prod_{j=0}^{n-1} \left(\int\limits_{\mathbf{R}^k} \left(\int\limits_{\mathbf{R}^k} \cos\left(\frac{2\pi}{p}\langle \mathbf{x} - \mathbf{y},^t A^j\alpha\rangle\right) d\mu(\mathbf{x}) \right) d\mu(\mathbf{y}) \right). \square$$

Lemma 2.6. Let $\alpha \in \mathbf{Z}^k - \{\mathbf{0}\}$. Then, $\forall\, n \in \mathbf{N}$:

$$\|\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U\| \geq \frac{1}{2}\left|\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha)\right|.$$

Proof. From (4), we have:

$$\|\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U\| = \frac{1}{2} \sup_{\|f\|_\infty \leq 1} |E_{\mathcal{L}_{n,\mathbf{x}_0}}[f] - E_U[f]|.$$

For all $\alpha \in \mathbf{Z}^k - \{\mathbf{0}\}$, define the following function $f : \mathbf{R}^k \longrightarrow \mathbf{C}$:

$$f(\mathbf{x}) = \exp\left(\frac{2\pi i}{p}\langle \mathbf{x}, \alpha\rangle\right).$$

Since $\|f\|_\infty = 1$, we obtain:

$$\|\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U\| = \frac{1}{2}|E_{\mathcal{L}_{n,\mathbf{x}_0}}[f] - E_U[f]|$$

$$= \frac{1}{2}\left|\int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x}, \alpha\rangle\right) d\mathcal{L}_{n,\mathbf{x}_0}(\mathbf{x}) - \frac{1}{p^k}\int\limits_{\mathbf{R}^k} \exp\left(\frac{2\pi i}{p}\langle \mathbf{x}, \alpha\rangle\right) 1_{[0,p)^k}(\mathbf{x})d\mathbf{x}\right|$$

$$= \frac{1}{2}\left|\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha) - \widehat{\mathcal{L}_U}(\alpha)\right|.$$

Moreover, since $\alpha \in \mathbf{Z}^k - \{\mathbf{0}\}$, there exists $j_0 \in \{1, ..., k\}$ such that $\alpha_{j_0} \in \mathbf{Z} - \{0\}$; then:

$$\widehat{\mathcal{L}_U}(\alpha) = \frac{1}{p^k}\prod_{j=1}^{k}\left(\int\limits_0^p \exp\left(\frac{2\pi i}{p}x_j\alpha_j\right)dx_j\right)$$

$$= \frac{1}{p^k}\prod_{j\in\{1,...,k\}-j_0}\left(\int\limits_0^p \exp\left(\frac{2\pi i}{p}x_j\alpha_j\right)dx_j\right)\int\limits_0^p \exp\left(\frac{2\pi i}{p}x_{j_0}\alpha_{j_0}\right)dx_{j_0}.$$

Observe that

$$\int\limits_0^p \exp\left(\frac{2\pi i}{p}x_{j_0}\alpha_{j_0}\right)dx_{j_0} = \frac{\exp(2\pi\alpha_{j_0}i) - 1}{\frac{2\pi i}{p}\alpha_{j_0}} = 0$$

$$\Rightarrow \widehat{\mathcal{L}_U}(\alpha) = 0,$$

from which

$$\|\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U\| \geq \frac{1}{2}\left|\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha)\right|. \ \square$$

Henceforth, by using the formula (6), we will quantify the rate of convergence of the Markov chain (2); moreover, by using Lemmas 2.5 and 2.6, we will find a lower bound for $\|\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U\|$. We start by proving the invariance of the uniform distribution on $[0,p)^k$ for the chain.

Proposition 2.7. Suppose that $\mathbf{X}$ and $\mathbf{Y}$ are two independent random vectors in $\mathbf{R}^k$, and $A \in M_k(\mathbf{Z}) \cap GL_k(\mathbf{Q})$; if $\mathbf{X}$ is uniformly distributed in the interval $[0,p)^k$, then the random vector $\mathbf{T} = A\mathbf{X} + \mathbf{Y} \pmod p$ is uniformly distributed in $[0,p)^k$.

Proof. $\forall\,\mathbf{z}\in\mathbf{R}^k$, we have:

$$f_{A\mathbf{X}}(\mathbf{z})=\frac{1}{|\det(A)|p^k}1_{A[0,p)^k}(\mathbf{z}),$$

from which, if $\mathbf{Z}=A\mathbf{X}$ (mod p):

$$\begin{aligned}f_{\mathbf{Z}}(\mathbf{z})&=\left(\sum_{\mathbf{h}\in\mathbf{Z}^k}f_{A\mathbf{X}}(\mathbf{z}+p\mathbf{h})\right)1_{[0,p)^k}(\mathbf{z})\\&=\frac{1}{|\det(A)|p^k}\cdot\left|\left\{\mathbf{h}\in\mathbf{Z}^k:\mathbf{z}+p\mathbf{h}\in A[0,p)^k\right\}\right|1_{[0,p)^k}(\mathbf{z}).\end{aligned}$$

Suppose that there exist $\mathbf{z}$, $\mathbf{x}\in[0,p)^k$ and $\mathbf{h}\in\mathbf{Z}^k$ such that $\mathbf{z}+p\mathbf{h}\in A[0,p)^k$, $\mathbf{x}+p\mathbf{h}\notin A[0,p)^k$, and define:

$$\bar{t}=\sup\left\{t\in[0,1]:t(\mathbf{x}+p\mathbf{h})+(1-t)(\mathbf{z}+p\mathbf{h})\in A[0,p)^k\right\},$$

$$\mathbf{y}_1=\bar{t}(\mathbf{x}+p\mathbf{h})+(1-\bar{t})(\mathbf{z}+p\mathbf{h}).$$

Observe that $\mathbf{x}_1=A^{-1}(\mathbf{y}_1)$ belongs to the affine space H_1 generated by 2^{k-1} vertices of the interval $[0,p)^k$. Consider the remaining 2^{k-1} vertices of $[0,p)^k$, indicate by H_2 the affine space that they generate, and define $\mathbf{x}_2$ as the projection of $\mathbf{x}_1$ into H_2; finally, define $\mathbf{y}_2=A\mathbf{x}_2$. Since H_1 and H_2 are parallel, so are AH_1 and AH_2, and consequently

$$\mathbf{z}-(\mathbf{y}_1-\mathbf{y}_2)+p\mathbf{h}\notin A[0,p)^k,\ \mathbf{x}-(\mathbf{y}_1-\mathbf{y}_2)+p\mathbf{h}\in A[0,p)^k.$$

Since $\mathbf{y}_1-\mathbf{y}_2=A(\mathbf{x}_1-\mathbf{x}_2)=p\overline{\mathbf{h}}$, for some $\overline{\mathbf{h}}\in\mathbf{Z}^k$, we have:

$$\left|\left\{\mathbf{h}\in\mathbf{Z}^k:\mathbf{z}+p\mathbf{h}\in A[0,p)^k\right\}\right|=\left|\left\{\mathbf{h}\in\mathbf{Z}^k:\mathbf{x}+p\mathbf{h}\in A[0,p)^k\right\}\right|,\ \forall\,\mathbf{z},\mathbf{x}\in[0,p)^k,$$

from which

$$f_{\mathbf{Z}}(\mathbf{z})=\frac{c}{|\det(A)|p^k}1_{[0,p)^k}(\mathbf{z}),\ \forall\,\mathbf{z}\in\mathbf{R}^k,\ \text{for some } c\in\mathbf{N},$$

and so

$$f_{\mathbf{Z}}(\mathbf{z})=\frac{1}{p^k}1_{[0,p)^k}(\mathbf{z}),\ \forall\,\mathbf{z}\in\mathbf{R}^k.$$

Then, the law of $\mathbf{Z}+\mathbf{Y}$ has a density g, with respect to the measure $Leb^{(k)}$, given by:

$$g(\mathbf{z})=\frac{1}{p^k}\int_{\mathbf{R}^k}1_{[0,p)^k}(\mathbf{z}-\mathbf{x})d\mathcal{L}_Y(\mathbf{x})=\frac{1}{p^k}\int_{\mathbf{R}^k}1_{\mathbf{z}-[0,p)^k}(\mathbf{x})d\mathcal{L}_Y(\mathbf{x}),\ \forall\,\mathbf{z}\in\mathbf{R}^k.$$

Finally, if $\mathbf{T} = A\mathbf{X} + \mathbf{Y} \pmod{p}$, since $\mathbf{T} = \mathbf{Z} + \mathbf{Y} \pmod{p}$, $\forall\, \mathbf{z} \in \mathbf{R}^k$, we have:

$$f_{\mathbf{T}}(\mathbf{z}) = \left(\sum_{\mathbf{h} \in \mathbf{Z}^k} g(\mathbf{z} + p\mathbf{h}) \right) 1_{[0,p)^k}(\mathbf{z}) = \frac{1}{p^k} \left(\sum_{\mathbf{h} \in \mathbf{Z}^k} \int\limits_{\mathbf{R}^k} 1_{\mathbf{z}+p\mathbf{h}-[0,p)^k}(\mathbf{x}) d\mathcal{L}_Y(\mathbf{x}) \right) 1_{[0,p)^k}(\mathbf{z})$$

$$= \frac{1}{p^k} \left(\int\limits_{\mathbf{R}^k} d\mathcal{L}_Y(\mathbf{x}) \right) 1_{[0,p)^k}(\mathbf{z}) = \frac{1}{p^k} 1_{[0,p)^k}(\mathbf{z}). \ \square$$

3 Convergence in total variation to the uniform distribution on $[0,p)^k$

Proposition 3.1. Suppose that $A \in M_k(\mathbf{Z}) \cap GL_k(\mathbf{Q})$, $\mu(B) \geq \alpha \int\limits_B 1_{[-a,a]^k}(\mathbf{x}) d\mathbf{x}$, $\forall\, B \in \mathcal{B}^{(k)}$, where $a, \alpha \in \mathbf{R}^+$; then, $\forall\, m \in \mathbf{N}$, $\forall\, \varepsilon \in \left[0, \frac{1}{2}\right]$, $\forall\, B \in \mathcal{B}^{(k)}$, and $\forall\, (t_0, t_1, ..., t_m) \in \mathbf{N}^{m+1}$ such that $t_i \neq t_j \ \forall\, i \neq j$, we have:

$$P\left(\sum_{i=0}^{m} A^i \mathbf{B}_{t_i} \in B \right) \geq \varphi_m(B), \tag{10}$$

where φ_m is the measure on $(\mathbf{R}^k, \mathcal{B}^{(k)})$ with density, with respect to the measure $Leb^{(k)}$, given by:

$$f_m(\mathbf{x}) = \left(\frac{2a\varepsilon}{k} \right)^{km} \frac{\alpha^{m+1}}{|\det(A)|^{\frac{m(m+1)}{2}}} 1_{\left[-\left(\frac{m(1-2\varepsilon)}{k}+1\right)a, \left(\frac{m(1-2\varepsilon)}{k}+1\right)a\right]^k}(\mathbf{x}), \ \forall\, \mathbf{x} \in \mathbf{R}^k. \tag{11}$$

In particular:

$$P\left(\sum_{i=0}^{m} A^i \mathbf{B}_{m-i} \in B \right) \geq \varphi_m(B).$$

Proof. In order to prove the formula (10), proceed by induction on m; if $m = 0$, the statement is true by assumption. Suppose that the statement is true for m; then, for

$m+1$, $\forall\, B \in \mathcal{B}^{(k)}$:

$$P\left(\sum_{i=0}^{m+1} A^i \mathbf{B}_{t_i} \in B\right) = \left(\mathcal{L}_{\sum\limits_{i=0}^{m} A^i \mathbf{B}_{t_i}} * \mathcal{L}_{A^{m+1}\mathbf{B}_{t_{m+1}}}\right)(B)$$

$$= \int\limits_{\mathbf{R}^k} 1_B(\mathbf{z}) d\left(\mathcal{L}_{\sum\limits_{i=0}^{m} A^i \mathbf{B}_{t_i}} * \mathcal{L}_{A^{m+1}\mathbf{B}_{t_{m+1}}}\right)(\mathbf{z}) = \int\limits_{\mathbf{R}^{2k}} 1_B(\mathbf{x}+\mathbf{y}) d\left(\mathcal{L}_{\sum\limits_{i=0}^{m} A^i \mathbf{B}_{t_i}} \otimes \mathcal{L}_{A^{m+1}\mathbf{B}_{t_{m+1}}}\right)(\mathbf{x},\mathbf{y})$$

$$= \int\limits_{\mathbf{R}^k} \left(\int\limits_{\mathbf{R}^k} 1_B(\mathbf{x}+\mathbf{y}) d\mathcal{L}_{A^{m+1}\mathbf{B}_{t_{m+1}}}(\mathbf{y})\right) d\mathcal{L}_{\sum\limits_{i=0}^{m} A^i \mathbf{B}_{t_i}}(\mathbf{x})$$

$$\geq \int\limits_{\mathbf{R}^k} \left(\int\limits_{\mathbf{R}^k} 1_B(\mathbf{x}+\mathbf{y}) d\mathcal{L}_{A^{m+1}\mathbf{B}_{t_{m+1}}}(\mathbf{y})\right) d\varphi_m(\mathbf{x}) \text{ (by the inductive assumption).} \qquad (12)$$

Observe that, $\forall\, C \in \mathcal{B}^{(k)}$, we have:

$$\mathcal{L}_{A^{m+1}\mathbf{B}_{t_{m+1}}}(C) = P\left(\mathbf{B}_{t_{m+1}} \in \left(A^{m+1}\right)^{-1} C\right) \geq \alpha \int\limits_{(A^{m+1})^{-1}C} 1_{[-a,a]^k}(\mathbf{x}) d\mathbf{x}$$

$$= \alpha \int\limits_{\{\mathbf{x}\in\mathbf{R}^k : A^{m+1}\mathbf{x}\in C\}} 1_{A^{m+1}[-a,a]^k}\left(A^{m+1}\mathbf{x}\right) d\mathbf{x} = \frac{\alpha}{|\det(A)|^{m+1}} \int\limits_{C} 1_{A^{m+1}[-a,a]^k}(\mathbf{y}) d\mathbf{y}.$$

Then, from (12) and (11):

$$P\left(\sum_{i=0}^{m+1} A^i \mathbf{B}_{t_i} \in B\right) \geq \frac{\alpha}{|\det(A)|^{m+1}} \int\limits_{\mathbf{R}^k} \left(\int\limits_{\mathbf{R}^k} 1_B(\mathbf{x}+\mathbf{y}) 1_{A^{m+1}[-a,a]^k}(\mathbf{y}) d\mathbf{y}\right) d\varphi_m(\mathbf{x})$$

$$= \frac{\alpha}{|\det(A)|^{m+1}} \int\limits_{\mathbf{R}^k} \left(\int\limits_{\mathbf{R}^k} 1_B(\mathbf{z}) 1_{A^{m+1}[-a,a]^k}(\mathbf{z}-\mathbf{x}) d\mathbf{z}\right) f_m(\mathbf{x}) d\mathbf{x}$$

$$= \left(\frac{2a\varepsilon}{k}\right)^{km} \frac{\alpha^{m+2}}{|\det(A)|^{\frac{(m+1)(m+2)}{2}}} \int\limits_{\mathbf{R}^k} \left(\int\limits_{B} 1_{A^{m+1}[-a,a]^k}(\mathbf{z}-\mathbf{x}) d\mathbf{z}\right) 1_{I_m}(\mathbf{x}) d\mathbf{x}$$

$$\text{(where } I_m = [-s_m, s_m]^k = \left[-\left(\frac{m(1-2\varepsilon)}{k}+1\right)a, \left(\frac{m(1-2\varepsilon)}{k}+1\right)a\right]^k \text{)}$$

$$= \left(\frac{2a\varepsilon}{k}\right)^{km} \frac{\alpha^{m+2}}{|\det(A)|^{\frac{(m+1)(m+2)}{2}}} \int\limits_{B} Leb^{(k)}\left(\left(\mathbf{z} - A^{m+1}[-a,a]^k\right) \cap I_m\right) d\mathbf{z}, \qquad (13)$$

where the last equality follows from Tonelli's theorem.

Set $D = \left\{\mathbf{x} \in \mathbf{R}^k : \sum\limits_{i=1}^{k} |x_i| \leq 1\right\}$. We have $A^{m+1}[-a,a]^k \supset A^{m+1}\left\{\mathbf{x}\in\mathbf{R}^k : \sum\limits_{i=1}^{k} |x_i| \leq a\right\} = aA^{m+1}D$; moreover, D is the convex hull of the set

$$E = \left\{\mathbf{x} \in \mathbf{R}^k : |x_i| = 1, \text{ for some } i \in \{1, ..., k\},\ x_j = 0\ \forall\, j \neq i\right\},$$

and so $A^{m+1}D$ is the convex hull of $A^{m+1}E$, by linearity. Since $\det(A) \neq 0$, $\forall\, \mathbf{x} \in E$ we have $A^{m+1}\mathbf{x} \neq \mathbf{0}$; this implies $A^{m+1}D \supset E$, and so $A^{m+1}D \supset D$, from which $A^{m+1}[-a,a]^k \supset aD = \left\{\mathbf{x} \in \mathbf{R}^k : \sum\limits_{i=1}^{k} |x_i| \leq a\right\} \supset \left[-\frac{a}{k}, \frac{a}{k}\right]^k$.

Consider the function $g_m : I_{m+1} \longrightarrow \mathbf{R}^+$ defined by

$$g_m(\mathbf{z}) = Leb^{(k)}\left(\left(\mathbf{z} - \left[-\frac{a}{k}, \frac{a}{k}\right]^k\right) \cap I_m\right), \forall\, \mathbf{z} \in I_{m+1}.$$

The previous arguments and formula (13) imply

$$P\left(\sum_{i=0}^{m+1} A^i \mathbf{B}_{t_i} \in B\right) \geq \left(\frac{2a\varepsilon}{k}\right)^{km} \frac{\alpha^{m+2}}{|\det(A)|^{\frac{(m+1)(m+2)}{2}}} \int\limits_B g_m(\mathbf{z}) 1_{I_{m+1}}(\mathbf{z}) d\mathbf{z}. \qquad (14)$$

Set $V_{m+1} = \big\{\mathbf{z} = (z_1, ..., z_k) \in \mathbf{R}^k : z_i \in \{s_{m+1}, -s_{m+1}\}, \forall\, i = 1, ..., k\big\} \subset \partial(I_{m+1})$. It is easy to prove that, $\forall\, \mathbf{z} \in I_{m+1}$ and $\forall\, \overline{\mathbf{z}} \in V_{m+1}$, we have

$$g_m(\mathbf{z}) \geq g_m(\overline{\mathbf{z}}) = \left(\frac{a}{k} - (s_{m+1} - s_m)\right)^k = \left(\frac{2a\varepsilon}{k}\right)^k.$$

Then, from formula (14):

$$P\left(\sum_{i=0}^{m+1} A^i \mathbf{B}_{t_i} \in B\right) \geq \left(\frac{2a\varepsilon}{k}\right)^{k(m+1)} \frac{\alpha^{m+2}}{|\det(A)|^{\frac{(m+1)(m+2)}{2}}} \int\limits_B 1_{I_{m+1}}(\mathbf{z}) d\mathbf{z},$$

that is the formula (10) for $m+1$. In particular, if $t_i = m - i$, $\forall\, i \in \{0, 1, ..., m\}$, we have:

$$P\left(\sum_{i=0}^{m} A^i \mathbf{B}_{m-i} \in B\right) \geq \varphi_m(B). \ \square$$

The following proposition prove that the state space $[0,p)^k$ is a small set.

Proposition 3.2. Suppose that $A \in M_k(\mathbf{Z}) \cap GL_k(\mathbf{Q})$, $p \in \mathbf{R}^+$, $\mu(B) \geq \alpha \int\limits_B 1_{[b,c]^k}(\mathbf{x}) d\mathbf{x}$, $\forall\, B \in \mathcal{B}^{(k)}$, for some $b, c \in \mathbf{R}$ such that $c - b = \delta \in \mathbf{R}^+$, and for some $\alpha \in \mathbf{R}^+$; then, there exist $m_0 = m_0(p, k, \delta) \in \mathbf{N}^*$ and a decreasing sequence $\{\sigma_m\}_{m \geq 1} \subset (0,1)$, where $\sigma_m = \sigma_m(k, \delta, \alpha, |\det(A)|)$, such that, $\forall\, m \geq m_0$, $\forall\, \mathbf{x}_0 \in [0,p)^k$, and $\forall\, C \in \mathcal{B}([0,p)^k)$, we have

$$\mathcal{L}_{m,\mathbf{x}_0}(C) \geq \frac{\sigma_m}{p^k} Leb^{(k)}(C).$$

Proof. Set $\mathbf{b} = (b, ..., b) \in \mathbf{R}^k$, $\mathbf{c} = (c, ..., c) \in \mathbf{R}^k$, $\mathbf{C}_n = \mathbf{B}_n - \frac{\mathbf{b}+\mathbf{c}}{2}$, $\forall\, n \in \mathbf{N}$; then, $\forall\, m \in \mathbf{N}^*$, $\forall\, \mathbf{x}_0 \in [0, p)^k$, and $\forall\, B \in \mathcal{B}^{(k)}$, we have:

$$\begin{aligned}\mathbf{Y}_m = A^m\mathbf{x}_0 + \sum_{i=0}^{m-1} A^i \mathbf{B}_{m-1-i} &= A^m\mathbf{x}_0 + \sum_{i=0}^{m-1} A^i \frac{\mathbf{b}+\mathbf{c}}{2} + \sum_{i=0}^{m-1} A^i \mathbf{C}_{m-1-i} \\ \Longrightarrow P(\mathbf{Y}_m \in B) &= P\left(\sum_{i=0}^{m-1} A^i \mathbf{C}_{m-1-i} \in B - A^m\mathbf{x}_0 - \sum_{i=0}^{m-1} A^i \frac{\mathbf{b}+\mathbf{c}}{2}\right).\end{aligned}$$

Observe that $P(\mathbf{C}_n \in B) \geq \alpha \int_B 1_{\left[-\frac{\delta}{2}, \frac{\delta}{2}\right]^k}(\mathbf{x}) d\mathbf{x}$, $\forall\, n \in \mathbf{N}$; then, from Proposition 3.1, $\forall\, \varepsilon \in \left[0, \frac{1}{2}\right]$, we have:

$$\begin{aligned}P(\mathbf{Y}_m \in B) &\geq \left(\frac{\varepsilon\delta}{k}\right)^{k(m-1)} \frac{\alpha^m}{|\det(A)|^{\frac{(m-1)m}{2}}} \int\limits_{B - A^m\mathbf{x}_0 - \sum\limits_{i=0}^{m-1} A^i \frac{\mathbf{b}+\mathbf{c}}{2}} 1_{[-\gamma_m, \gamma_m]^k}(\mathbf{x}) d\mathbf{x} \\ &= \left(\frac{\varepsilon\delta}{k}\right)^{k(m-1)} \frac{\alpha^m}{|\det(A)|^{\frac{(m-1)m}{2}}} \int_B 1_{A^m\mathbf{x}_0 + \sum\limits_{i=0}^{m-1} A^i \frac{\mathbf{b}+\mathbf{c}}{2} + [-\gamma_m, \gamma_m]^k}(\mathbf{t}) d\mathbf{t}, \qquad (15)\end{aligned}$$

where $\gamma_m = \left(\frac{(m-1)(1-2\varepsilon)}{k} + 1\right)\frac{\delta}{2}$.

Suppose $\varepsilon = \frac{1}{4}$ and set $m_0 = \max\left\{\left[\frac{2pk}{\delta}\right], 1\right\}$; $\forall\, m \geq m_0$, we have:

$$\gamma_m = \left(\frac{m-1}{2k} + 1\right)\frac{\delta}{2} \geq \frac{(m+1)\delta}{4k} \geq \frac{p}{2}\left[\frac{(m+1)\delta}{2kp}\right] \geq \frac{p}{2}.$$

Set $\mathbf{p} = (p, ..., p) \in \mathbf{R}^k$, $\mathbf{y} = (y_1, ..., y_k) = A^m\mathbf{x}_0 + \sum\limits_{i=0}^{m-1} A^i \frac{\mathbf{b}+\mathbf{c}}{2} - \frac{1}{2}\left[\frac{(m+1)\delta}{2kp}\right]\mathbf{p}$; then, from formula (15):

$$P(\mathbf{Y}_m \in B) \geq \left(\frac{\delta}{4k}\right)^{k(m-1)} \frac{\alpha^m}{|\det(A)|^{\frac{(m-1)m}{2}}} \int_B 1_{\mathbf{y} + \left[0, p\left[\frac{(m+1)\delta}{2kp}\right]\right]^k}(\mathbf{t}) d\mathbf{t}.$$

Moreover, $\forall\, C \in \mathcal{B}([0,p)^k)$, we have:

$$\begin{aligned}
P(\mathbf{X}_m \in C) \geq & \sum_{\mathbf{i}\in\left\{0,\dots,\left[\frac{(m+1)\delta}{2kp}\right]\right\}^k} P\left(\mathbf{Y}_m \in \left(C + p\left[\frac{\mathbf{y}}{p}\right] + p\mathbf{i}\right)\right) \\
\geq & \sum_{\mathbf{i}\in\left\{0,\dots,\left[\frac{(m+1)\delta}{2kp}\right]\right\}^k} \left(\frac{\delta}{4k}\right)^{k(m-1)} \frac{\alpha^m}{|\det(A)|^{\frac{(m-1)m}{2}}} \\
& \cdot Leb^{(k)}\left(\left(C + p\left[\frac{\mathbf{y}}{p}\right] + p\mathbf{i}\right) \cap \left(\mathbf{y} + \left[0, p\left[\frac{(m+1)\delta}{2kp}\right]\right]^k\right)\right) \\
= \left(\frac{\delta}{4k}\right)^{k(m-1)} & \frac{\alpha^m}{|\det(A)|^{\frac{(m-1)m}{2}}} Leb^{(k)}\left(D \cap \left(p\mathbf{z} + \left[0, p\left[\frac{(m+1)\delta}{2kp}\right]\right]^k\right)\right) \qquad (16)
\end{aligned}$$

(by the translation invariance of the Lebesgue measure),

where

$$D = \bigcup_{\mathbf{i}\in\left\{0,\dots,\left[\frac{(m+1)\delta}{2kp}\right]\right\}^k} C + p\mathbf{i},\ \mathbf{z} = (z_1, \dots, z_k) = \left\{\frac{\mathbf{y}}{p}\right\} \in [0,1)^k.$$

Observe that, $\forall\, i = 1, \dots, k$, since

$$[0, pz_i] + p\left[\frac{(m+1)\delta}{2kp}\right] = \left[p\left[\frac{(m+1)\delta}{2kp}\right], pz_i + p\left[\frac{(m+1)\delta}{2kp}\right]\right],$$

by definition of D we have

$$\begin{aligned}
& Leb^{(k)}\left(D \cap \left(\prod_{j=1}^{i-1}\left[pz_j, pz_j + p\left[\frac{(m+1)\delta}{2kp}\right]\right] \times [0, pz_i] \times \prod_{j=1+1}^{k}\left[pz_j, pz_j + p\left[\frac{(m+1)\delta}{2kp}\right]\right]\right)\right) \\
& = Leb^{(k)}\left(D \cap \left(\prod_{j=1}^{i-1}\left[pz_j, pz_j + p\left[\frac{(m+1)\delta}{2kp}\right]\right] \times \left[p\left[\frac{(m+1)\delta}{2kp}\right], pz_i + p\left[\frac{(m+1)\delta}{2kp}\right]\right]\right.\right. \\
& \qquad \left.\left. \times \prod_{j=1+1}^{k}\left[pz_j, pz_j + p\left[\frac{(m+1)\delta}{2kp}\right]\right]\right)\right).
\end{aligned}$$

This implies

$$\begin{aligned}
& Leb^{(k)}\left(D \cap \left(p\mathbf{z} + \left[0, p\left[\frac{(m+1)\delta}{2kp}\right]\right]^k\right)\right) \\
& = Leb^{(k)}\left(D \cap \left((0, pz_2, \dots, pz_k) + \left[0, p\left[\frac{(m+1)\delta}{2kp}\right]\right]^k\right)\right).
\end{aligned}$$

By iterating the argument for $i = 2, ..., k$, we have

$$Leb^{(k)}\left(D \cap \left(p\mathbf{z} + \left[0, p\left[\frac{(m+1)\delta}{2kp}\right]\right]^k\right)\right)$$
$$= Leb^{(k)}\left(D \cap \left[0, p\left[\frac{(m+1)\delta}{2kp}\right]\right]^k\right) = \left[\frac{(m+1)\delta}{2kp}\right]^k Leb^{(k)}(C).$$

Then, from formula (16):

$$P(\mathbf{X}_m \in C) \geq \left(\frac{\delta}{4k}\right)^{k(m-1)} \frac{\alpha^m}{|\det(A)|^{\frac{(m-1)m}{2}}} \left[\frac{(m+1)\delta}{2kp}\right]^k Leb^{(k)}(C)$$
$$\geq \left(\frac{\delta}{4k}\right)^{k(m-1)} \frac{\alpha^m}{|\det(A)|^{\frac{(m-1)m}{2}}} \left(\frac{(m+1)\delta}{2kp}\right)^k Leb^{(k)}(C) = \frac{\sigma_m}{p^k} Leb^{(k)}(C),$$

where $\sigma_m = \sigma_m(k, \delta, \alpha, |\det(A)|) = (m+1)^k \left(\left(\frac{\delta}{4k}\right)^k \alpha\right)^m \frac{1}{|\det(A)|^{\frac{(m-1)m}{2}}} \in (0, 1)$.

Finally, suppose $m \geq 1$ and observe that $\delta^k \alpha \leq 1$, since otherwise, by assumption, it would be the case that $\mu(\mathbf{R}^k) \geq \delta^k \alpha > 1$; we have:

$$\sigma_{m+1}(k, \delta, \alpha, |\det(A)|) = (m+2)^k \left(\left(\frac{\delta}{4k}\right)^k \alpha\right)^{m+1} \frac{1}{|\det(A)|^{\frac{m(m+1)}{2}}}$$
$$= \left(\frac{m+2}{m+1}\right)^k \cdot \frac{\delta^k \alpha}{(4k)^k} \frac{1}{|\det(A)|^m} \cdot (m+1)^k \left(\left(\frac{\delta}{4k}\right)^k \alpha\right)^m \frac{1}{|\det(A)|^{\frac{(m-1)m}{2}}}$$
$$\leq \left(\frac{m+2}{4(m+1)}\right)^k \sigma_m(k, \delta, \alpha, |\det(A)|) \leq \sigma_m(k, \delta, \alpha, |\det(A)|). \tag{17}$$

Then, $\{\sigma_m\}_{m \geq 1}$ is a decreasing sequence. □

Theorem 3.3. Suppose that $A \in M_k(\mathbf{Z}) \cap GL_k(\mathbf{Q})$, $p \in \mathbf{R}^+$, $\mu(B) \geq \alpha \int_B 1_{[b,c]^k}(\mathbf{x}) d\mathbf{x}$, $\forall B \in \mathcal{B}^{(k)}$, for some $b, c \in \mathbf{R}$ such that $c - b = \delta \in \mathbf{R}^+$, and for some $\alpha \in \mathbf{R}^+$; then, the Markov chain $\{\mathbf{X}_n\}$ in (2) is $Leb^{(k)}$-irreducible; moreover, there exist $\rho = \rho(p, k, \delta, \alpha, |\det(A)|) \in \mathbf{R}^+$ and $\tau = \tau(p, k, \delta, \alpha, |\det(A)|) \in (0, 1)$ such that, $\forall n \in \mathbf{N}$ and $\forall \mathbf{x}_0 \in [0, p)^k$, we have

$$\|\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U\| \leq \rho \tau^n. \tag{18}$$

Consequently, $\{X_n\}$ is uniformly ergodic.

Proof. From Proposition 3.2, there exist $m_0 = m_0(p, k, \delta) \in \mathbf{N}^*$ and a sequence $\{\sigma_m\}_{m \geq 1} \subset (0, 1)$, where $\sigma_m = \sigma_m(k, \delta, \alpha, |\det(A)|)$, such that, $\forall m \geq m_0$, $\forall \mathbf{x}_0 \in [0, p)^k$, and $\forall C \in \mathcal{B}([0,p)^k)$, we have $\mathcal{L}_{m,\mathbf{x}_0}(C) \geq \frac{\sigma_m}{p^k} Leb^{(k)}(C)$; then, if $Leb^{(k)}(C) > 0$, we have $\mathcal{L}_{m,\mathbf{x}_0}(C) > 0$, and so $\{\mathbf{X}_n\}$ is $Leb^{(k)}$-irreducible. Moreover, from Theorem 2.2, for $m = m_0$, there exists a probability distribution π_{m_0} on $([0,p)^k, \mathcal{B}([0,p)^k))$ such that $\|\mathcal{L}_{m,\mathbf{x}_0} - \pi_{m_0}\| \leq \rho\tau^n$, $\forall n \in \mathbf{N}$ and $\forall \mathbf{x}_0 \in [0,p)^k$, where $\tau = \tau(p,k,\delta,\alpha,|\det(A)|) = (1-\sigma_{m_0})^{\frac{1}{m_0}} \in (0,1)$, $\rho = \rho(p,k,\delta,\alpha,|\det(A)|) = \dfrac{1}{1-\sigma_{m_0}} \in \mathbf{R}^+$.

Finally, since $\{\mathbf{X}_n\}$ is $Leb^{(k)}$-irreducible and since π_{m_0} is invariant for $\{\mathbf{X}_n\}$ (see for example Meyn and Tweedie (2005), page 237), from Proposition 2.3 $\{\mathbf{X}_n\}$ is recurrent and so, from Theorem 2.4, it has a unique invariant measure. Since $\mathcal{L}_{\mathbf{U}}$ is invariant by Proposition 2.7, we have $\pi_{m_0} = \mathcal{L}_{\mathbf{U}}$, that is the formula (18). □

Theorem 3.4. Suppose that the matrix A has an eigenvalue $\lambda \in \mathbf{C}$ such that $\lambda^l = 1$, for some $l \in \mathbf{N}^*$ (hence, so does the matrix ${}^tA$), and that $\|\mathbf{B}_n\|_\infty \in L^2$ for all $n \in \mathbf{N}$. Then, there exist $\gamma \in \mathbf{R}^+$ such that, for all $p \in \mathbf{N}$, p sufficiently large, and for all $n \in \mathbf{N}$, we have:

$$||\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U|| \geq \frac{1}{2}\exp\left(-\frac{\gamma n}{p^2}\right).$$

Consequently, if also the assumptions of Theorem 3.3 hold, $O\left(p^2\right)$ steps are needed to reach the uniform distribution.

Proof. The assumption on λ implies ${}^tA^l\mathbf{x} = \mathbf{x}$ for some $\mathbf{x} \in \mathbf{C}^k - \{\mathbf{0}\}$, and so $({}^tA^l - I)\mathbf{x} = \mathbf{0}$, which implies $\mathbf{x} \in \mathbf{Q}^k - \{\mathbf{0}\}$; then there exists $\alpha \in \mathbf{Z}^k - \{\mathbf{0}\}$ such that ${}^tA^l\alpha = \alpha$. Then, for all $j \in \mathbf{N}$, there exists $i \in \{0,1,...,l-1\}$ such that ${}^tA^j\alpha = {}^tA^i\alpha$.

Moreover, from Lemmas 2.5 and 2.6, $\forall\, n \in \mathbf{N}$:

$$\|\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U\| \geq \frac{1}{2}\left|\widehat{\mathcal{L}_{n,\mathbf{x}_0}}(\alpha)\right|$$

$$= \frac{1}{2}\prod_{j=0}^{n-1}\left(\int\limits_{\mathbf{R}^k}\left(\int\limits_{\mathbf{R}^k}\cos\left(\frac{2\pi}{p}\langle\mathbf{x}-\mathbf{y},{}^t A^j\alpha\rangle\right)d\mu(\mathbf{x})\right)d\mu(\mathbf{y})\right)^{1/2}.$$

Since $\cos x \geq 1 - \dfrac{x^2}{2}\ \forall\, x \in \mathbf{R}$, we have:

$$\int\limits_{\mathbf{R}^k}\left(\int\limits_{\mathbf{R}^k}\cos\left(\frac{2\pi}{p}\langle\mathbf{x}-\mathbf{y},{}^t A^j\alpha\rangle\right)d\mu(\mathbf{x})\right)d\mu(\mathbf{y}) \geq 1-\frac{\gamma}{p^2},$$

where $\gamma = 2\pi^2 k^4 \left\|\alpha\right\|_\infty^2 \max\limits_{i\in\{0,1,\ldots,l-1\}} \left\|{}^tA\right\|_\infty^{2i} \int\limits_{\mathbf{R}^k} \left(\int\limits_{\mathbf{R}^k} \left\|\mathbf{x}-\mathbf{y}\right\|_\infty^2 d\mu(\mathbf{x}) \right) d\mu(\mathbf{y}) \in \mathbf{R}^+$. Moreover, there exists $d \in \mathbf{R}^+$ such that $1 - x \geq \exp(-2x) > 0$, $\forall\ x \in [0, d]$. For all p sufficiently large, we can suppose that $\dfrac{\gamma}{p^2} \in [0, d]$; hence:

$$\begin{aligned} ||\mathcal{L}_{n,\mathbf{x}_0} - \mathcal{L}_U|| &\geq \frac{1}{2}\left(1-\frac{\gamma}{p^2}\right)^{n/2} \\ &\geq \frac{1}{2}\exp\left(-\frac{\gamma n}{p^2}\right). \ \square \end{aligned}$$

4 Problems for further study

From Theorem 3.3, the rate of convergence of the recursion (2) to the uniform law is geometric and depends on A, p, k, and the law of $\mathbf{B}_n$; this is true for any integer matrix with nonzero eigenvalues. Moreover, if A has an eigenvalue that is a root of 1, a lower bound of the variation distance between $\mathcal{L}_{n,\mathbf{x}_0}$ and $\mathcal{L}_{\mathbf{U}}$ is obtained from Theorem 3.4, and we can establish that the number of steps necessary to reach the uniform distribution is $O\left(p^2\right)$. Our idea is that, without any assumptions on A, $n = O\left(p^2\right)$ steps are also sufficient. This result should agree with the discrete case studied in Asci 2009a. We hope to prove it and to develop some applications in a further paper.

References

[1] Aldous, D., and Diaconis, P., 1986. Shuffling cards and stopping times. *American Mathematical Monthly* **93**, 333-348.

[2] Asci, C., 2001. Generating uniform random vectors. *J. Theoret. Probab.* **14**(2), 333-356.

[3] Asci, C., 2009a. Asymptotic behavior of an affine random recursion in $\mathbf{Z}_p^k$ defined by a matrix with an eigenvalue of size 1. *Statistics and Probability Letters* **79**, 1421-1428.

[4] Asci, C., 2009b. Generating uniform random vectors in $\mathbf{Z}_p^k$: the general case. *J. Theoret. Probab.* **22**, 791-809.

[5] Chung, F.R.K., Diaconis, P., and Graham, R.L., 1987. Random walks arising in random number generation. *Ann. Probab.* **15**(3), 1148-1165.

[6] Diaconis, P., 1988. *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics, Hayward, California.

[7] Helleloid, G., 2007. *Automorphism Groups of Finite p-Groups: Structure and Applications.* Ph.D. thesis, Department of Mathematics, Stanford University.

[8] Hildebrand, M., 1990. *Rates of Convergence of Some Random Processes on Finite Groups.* Ph.D. thesis, Department of Mathematics, Harvard University.

[9] Hildebrand, M., 1993. Random processes of the form $X_{n+1} = a_n X_n + b_n \pmod{p}$. *Ann. Probab.* **21**(2), 710-720.

[10] Hildebrand, M., and McCollum, J., 2008. Generating random vectors in $(\mathbf{Z}/p\mathbf{Z})^d$ via an affine random process. *J. Theoret. Probab.* **21**, 802-811.

[11] Knuth, D.E., 1981. *The Art of Computer Programming* **2**, 2nd ed. Addison -Wesley, Reading, Massachusetts.

[12] Meyn, S.P., and Tweedie, R.L., 2005. *Markov chains and stochastic stability,* 2nd ed. Springer Verlag, London.

[13] Rosenthal, J.S., 1995. Convergence rates for Markov chains. *Siam Review* **37**(3), 387-405.

[14] Serre, J.P., 1977. *Linear Representations of Finite Groups.* Springer-Verlag, New York.