Bounds in total variation distance for discrete-time processes on the sequence space

Flint, Ian; Privault, Nicolas; Torrisi, Giovanni Luca

2018

Flint, I., Privault, N., & Torrisi, G.L. (2018). Bounds in total variation distance for discrete-time processes on the sequence space. Potential Analysis, 52, 223–243. doi:10.1007/s11118-018-9744-0

https://hdl.handle.net/10356/137236

https://doi.org/10.1007/s11118-018-9744-0

Bounds in total variation distance for discrete-time processes on the sequence space

Ian Flint * Nicolas Privault[†] Giovanni Luca Torrisi[‡]

Abstract

Let $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ be the laws of two discrete-time stochastic processes defined on the sequence space $S^{\mathbb{N}}$, where S is a finite set of points. In this paper we derive a bound on the total variation distance $d_{TV}(\mathbb{P}, \widetilde{\mathbb{P}})$ in terms of the cylindrical projections of $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. We apply the result to Markov chains with finite state space and random walks on $\mathbb{Z}$ with not necessarily independent increments, and we consider several examples. Our approach relies on the general framework of stochastic analysis for discrete-time obtuse random walks and the proof of our main result makes use of the predictable representation of multidimensional normal martingales. Along the way, we obtain a sufficient condition for the absolute continuity of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$ which is of interest in its own right.

Key words: Total variation distance, Markov chains, random walks, normal martingales, obtuse random walks.

Mathematics Subject Classification (2010): 60J10, 60J05, 60G50, 60G42, 15A69.

1 Introduction

We consider the general problem of estimating the total variation distance between the laws of two discrete-time stochastic processes on the sequence space. Specifically, we provide an explicit upper bound on the total variation distance between two discrete-time stochastic processes in terms of the cylindrical projections of the laws of the processes. Our general estimate (see Theorem 4.1) yields: (i) an explicit upper bound on the total variation distance between the laws of two discrete-time Markov chains with finite state space in terms of the initial distribution and the coefficients of their transition matrices (see Corollary 5.1) (ii) an explicit upper bound on the total variation distance

*Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-04-02, 21 Nanyang Link Singapore 637371. e-mail: iflint@ntu.edu.sg

†Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-05-43, 21 Nanyang Link Singapore 637371. e-mail: nprivault@ntu.edu.sg

‡Istituto per le Applicazioni del Calcolo "Mauro Picone", CNR, Via dei Taurini 19, 00185 Roma, Italy. e-mail: torrisi@iac.rm.cnr.it

between the laws of two random walks on $\mathbb{Z}$ in terms of the initial distribution and the conditional distributions of their (not necessarily independent) increments (see Corollaries 5.2 and 5.3).

From the point of view of the applications, results of this kind can be used e.g. for parameter estimation in Markov chain models. Indeed, letting (X_n) denote a Markov chain based on some data set and (X_n^θ), $\theta \in \Theta \subset \mathbb{R}$, a parametrized family of Markov chains, one may minimize (with respect to θ) the upper bound on the total variation distance between (X_n) and (X_n^θ) to select the optimal estimation of (X_n) in the parametrized class of Markov chains (see the comment after Corollary 5.1). Our results can also be used to compute the distance between two inhomogeneous random walks. As an illustration, we present a specialization of our bounds to elephant random walks [12] and to a generalization of this model introduced in [9] (see Examples 2 and 3). We also mention that computable upper bounds on the distance between two discrete-time stochastic processes are potentially useful in the study of the performance measurement of learning algorithms (see e.g. Sections 3 and 4 in [13]).

We let S be a finite set of points, and for ease of notation we assume that $S = \{0, \ldots, d\}$ for some $d \geq 1$. The sequence space (see [5, p. 27]) is defined as $\Omega := \{0, \ldots, d\}^{\{0,\ldots,N\}}$, where $N \in \mathbb{N} := \{0, 1, \ldots\}$ or $N = \infty$, in which case we identify $\{0, \ldots, N\}$ with $\mathbb{N}$. The set S is regarded as the possible outcomes of an experiment, and Ω is the sample space corresponding to the repetition of the experiment. The space Ω is endowed with the filtration $(\mathcal{F}_n)_{n\in\{-1,0,\ldots,N\}}$ generated by the $\{0, \ldots, d\}$-valued coordinate maps $(\pi_n)_{n\in\{0,\ldots,N\}}$, $\pi_n((\omega_k)_{k\in\{0,\ldots,N\}}) := \omega_n$, i.e.

$$\mathcal{F}_n := \sigma(\pi_0, \ldots, \pi_n), \qquad n \in \{0, \ldots, N\}, \tag{1.1}$$

with $\mathcal{F}_{-1} := \{\emptyset, \Omega\}$, and we let $\mathcal{F} := \bigvee_{n=0}^N \mathcal{F}_n$. We fix two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$, we define the predictable processes (or cylindrical projections of $\mathbb{P}$ and $\widetilde{\mathbb{P}}$, respectively)

$$p_n^{(i)} := \mathbb{P}(\pi_n = i \mid \mathcal{F}_{n-1}) \quad \text{and} \quad \widetilde{p}_n^{(i)} := \widetilde{\mathbb{P}}(\pi_n = i \mid \mathcal{F}_{n-1}), \tag{1.2}$$

and we assume that

$$0 < p_n^{(i)}, \widetilde{p}_n^{(i)} < 1, \qquad i \in \{0, \ldots, d\}, \quad n \in \{0, \ldots, N\}.$$

We shall see that $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ can be the laws of two discrete-time Markov chains with finite state space and of two random walks on $\mathbb{Z}$, see Section 5. In this paper, we aim at providing a bound on the total variation distance $\mathrm{d}_{\mathrm{TV}}(\mathbb{P}, \widetilde{\mathbb{P}})$ between $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. Our approach relies on the construction of a suitable d-dimensional normal martingale (see [3]) correctly associated with the cylindrical projections (1.2).

The proof of the general bound on $\mathrm{d_{TV}}(\mathbb{P}, \widetilde{\mathbb{P}})$ (i.e. of Theorem 4.1) is based on the following decomposition of a square integrable random variable $F : \Omega \to \mathbb{R}$ obtained recently in [8]:

$$F(\omega) - \mathbb{E}[F] = \sum_{n=0}^{N} \langle \xi_n(\omega), Y_n(\omega) \rangle_{\mathbb{R}^d}, \tag{1.3}$$

where $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$, $\langle x, y \rangle_{\mathbb{R}^d}$ denotes the usual inner product of two elements $x = (x^1, \dots, x^d)$, $y = (y^1, \dots, y^d)$ of $\mathbb{R}^d$, $(\xi_n)_{n \in \{0,\dots,N\}}$ is a discrete-time predictable process given by the Clark-Ocone formula (see Proposition 2.2) and $(Y_n)_{n \in \{0,\dots,N\}}$ is the sequence of increments of a discrete-time d-dimensional normal martingale. The proof proceeds by taking the expectation under $\widetilde{\mathbb{P}}$ in (1.3), which yields

$$|\widetilde{\mathbb{E}}[F] - \mathbb{E}[F]| \le \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[|\langle \xi_n, Y_n \rangle_{\mathbb{R}^d}| \right], \tag{1.4}$$

and by bounding the right-hand side of this latter relation.

The inequality (1.4) is deduced under the crucial assumption $\widetilde{\mathbb{P}} \ll \mathbb{P}$, i.e. the absolute continuity of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$. Therefore, to apply our general estimate to most cases of interest, we need sufficient conditions for $\widetilde{\mathbb{P}} \ll \mathbb{P}$. We thus prove in Proposition 3.1 that if $\widetilde{\mathbb{E}}[\sum_{n=0}^{N} \widetilde{p}_n^{(i)}] < \infty$ and $\sum_{n=0}^{N} p_n^{(i)} < \infty$, $\widetilde{\mathbb{P}}$-a.s. for all i different from a fixed i_0, then $\widetilde{\mathbb{P}}$ is absolutely continuous with respect to $\mathbb{P}$. Additionally, the Radon-Nikodym derivative can be made explicit under slightly stronger conditions. We note that the result of Proposition 3.1 is quite general. As an illustration, if $V = (V_n)_{n \in \{0,\dots,N\}}$ and $W = (W_n)_{n \in \{0,\dots,N\}}$ are two sequences of $\{0, \dots, d\}$-valued random variables defined on Ω, Proposition 3.1 provides a sufficient condition for the absolute continuity of the law of V with respect to the law of W in terms of their respective cylindrical projections.

When $d = 1$ and $d = 2$ the general bound of Theorem 4.1 can be considerably simplified (see Corollaries 4.1 and 4.2). Such simplified estimates are applied to random walks in Section 5.2. In Corollary 4.3 we give a more explicit upper bound on $\mathrm{d_{TV}}(\mathbb{P}, \widetilde{\mathbb{P}})$ which is used in Section 5.1 to obtain estimates on the total variation distance between the laws of two discrete-time Markov chains with finite state space.

The paper is organized as follows. In Section 2, we give some preliminary results on obtuse random walks, and in Section 3 we provide sufficient conditions for the absolute continuity required in the statement of Theorem 4.1. In Section 4 we provide the general upper bound on $\mathrm{d_{TV}}(\mathbb{P}, \widetilde{\mathbb{P}})$ and we make it more explicit in some specific settings. The above-mentioned applications to discrete-time Markov chains with finite state space and random walks are given in Section 5.

2 Preliminaries

An obtuse system in $\mathbb{R}^d$, see [3], is a family of $d+1$ vectors $x_0, \ldots, x_d \in \mathbb{R}^d$ such that $\langle x_i, x_j \rangle_{\mathbb{R}^d} = -1$, $\forall i \neq j$. We let $\mathcal{OP}_p$ denote the collection of families $((v_n^{(i)})_{n\in\{0,\ldots,N\}})_{0\le i\le d}$ of $\mathbb{R}^d$-valued predictable processes such that $(v_n^{(i)}(\omega))_{0\le i\le d}$ is an obtuse system in $\mathbb{R}^d$ for all $\omega \in \Omega$ and $n \in \{0, \ldots, N\}$, and which is correctly associated to the predictable processes $p \equiv ((p_n^{(i)})_{n\in\{0,\ldots,N\}})_{0\le i\le d}$ defined in (1.2), in the sense that it verifies the structure equations

$$\sum_{i=0}^{d} v_n^{(i)}(\omega) p_n^{(i)}(\omega) = \mathbf{0} \quad \text{and} \quad \sum_{i=0}^{d} p_n^{(i)}(\omega)\, v_n^{(i)}(\omega) \otimes v_n^{(i)}(\omega) = \mathrm{I}_d, \tag{2.1}$$

$n \in \{0, \ldots, N\}$, $\omega \in \Omega$, where $\mathbf{0} \in \mathbb{R}^d$ is the null vector, I_d is the identity in the group $\mathcal{O}_d$ of $d \times d$ orthogonal real matrices, and

$$v_n^{(i)}(\omega) \otimes v_n^{(i)}(\omega) := ((v_n^{(i)}(\omega))^j (v_n^{(i)}(\omega))^k)_{1\le j,k\le d}.$$

For later purposes, we recall that

$$\|v_n^{(i)}\|_{\mathbb{R}^d}^2 = \frac{1-p_n^{(i)}}{p_n^{(i)}}, \quad n \in \{0, \ldots, N\}, \quad i = 0, \ldots, d. \tag{2.2}$$

The following proposition combines [3, Theorem 2(b)] and [2, Proposition 2.4].

Proposition 2.1. *The family $\mathcal{OP}_p$ is not empty and, for any $((v_n^{(i)})_{n\in\{0,\ldots,N\}})_{0\le i\le d} \in \mathcal{OP}_p$, we have*

$$\mathcal{OP}_p \equiv \{((Ov_n^{(i)})_{n\in\{0,\ldots,N\}})_{0\le i\le d} : O \in \mathcal{O}_d\}.$$

Hereon, we fix $((v_n^{(i)})_{n\in\{0,\ldots,N\}})_{0\le i\le d} \in \mathcal{OP}_p$, define $Y_n(\omega) := v_n^{(\pi_n(\omega))}(\omega) = v_n^{(\omega_n)}(\omega)$, $n \in \{0, \ldots, N\}$, $\omega \in \Omega$ and note that, since $v_n^{(i)} = v_n^{(j)}$ if and only if $i = j$ (which follows from the fact that $\langle v_n^{(i)}, v_n^{(j)} \rangle_{\mathbb{R}^d} = -1$ for all $i \neq j$), we have

$$\{\omega \in \Omega : \pi_n(\omega) = i\} = \{\omega \in \Omega : Y_n(\omega) = v_n^{(i)}(\omega)\}, \quad n \in \{0, \ldots, N\},\ i \in \{0, \ldots, d\},$$

which implies $\mathcal{F}_n = \sigma(Y_0, \ldots, Y_n)$, $n \in \{0, \ldots, N\}$, as well as

$$p_n^{(i)} = \mathbb{P}(\pi_n = i \mid \mathcal{F}_{n-1}) = \mathbb{P}\left(Y_n = v_n^{(i)} \mid \mathcal{F}_{n-1}\right), \quad n \in \{0, \ldots, N\},\ i \in \{0, \ldots, d\}.$$

By the structure equations (2.1), we then have that $(Y_0 + \cdots + Y_n)_{n\in\{0,\ldots,N\}}$ is a d-dimensional normal martingale, i.e.

$$\mathbb{E}[Y_n \mid \mathcal{F}_{n-1}] = 0 \text{ and } \mathrm{Var}\,(Y_n \mid \mathcal{F}_{n-1}) := \mathbb{E}[Y_n \otimes Y_n \mid \mathcal{F}_{n-1}] = \mathrm{I}_n, \quad n \in \{0, \ldots, N\}.$$

The process $(Y_0 + \cdots + Y_n)_{n\in\{0,\ldots,N\}}$ is called a d-dimensional obtuse random walk.

Finally we recall the following Clark-Ocone formula for square-integrable functions on $(\Omega, \mathcal{F}, \mathbb{P})$, see [8, Proposition 7.1 and Proposition 7.2].

Proposition 2.2. *For any real-valued $F \in \mathrm{L}^2\big(\Omega, \mathbb{P}\big)$ we have*

$$F = \mathbb{E}[F] + \sum_{n=0}^{N} \langle \mathbb{E}\left[D_n F \mid \mathcal{F}_{n-1}\right], Y_n \rangle_{\mathbb{R}^d}, \quad \mathbb{P} - a.s.,$$

where

$$D_n F(\omega) := \sum_{i=0}^{d} p_n^{(i)}(\omega) v_n^{(i)}(\omega) F(\omega_n^i), \quad n \in \{0, \ldots, N\},\ \omega \in \Omega, \tag{2.3}$$

and

$$\omega_n^i := (\omega_0, \ldots, \omega_{n-1}, i, \omega_{n+1}, \ldots, \omega_N), \quad n \in \{0, \ldots, N\},\ i \in \{0, \ldots, d\}.$$

Remark 1. *For later purposes, we note that when $d = 1$, setting*

$$p_n^{(0)}(\omega) := p_n(\omega), \quad p_n^{(1)}(\omega) := q_n(\omega) := 1 - p_n(\omega),$$

$\omega \in \Omega = \{0, 1\}^N$, $n \in \{0, \ldots, N\}$, by a simple computation we have that the corresponding set $\mathcal{OP}_p$ is given by the processes

$$(v_n^{(0)})_{n\in\{0,\ldots,N\}} := \left(-\sqrt{\frac{q_n}{p_n}}\right)_{n\in\{0,\ldots,N\}}, \quad (v_n^{(1)})_{n\in\{0,\ldots,N\}} := \left(\sqrt{\frac{p_n}{q_n}}\right)_{n\in\{0,\ldots,N\}},$$

and

$$((-1)v_n^{(0)})_{n\in\{0,\ldots,N\}} := \left(\sqrt{\frac{q_n}{p_n}}\right)_{n\in\{0,\ldots,N\}}, \quad ((-1)v_n^{(1)})_{n\in\{0,\ldots,N\}} := \left(-\sqrt{\frac{p_n}{q_n}}\right)_{n\in\{0,\ldots,N\}}.$$

The random variables

$$Y_n(\omega) = \begin{cases} v_n^{(0)}(\omega) = -\sqrt{\dfrac{q_n(\omega)}{p_n(\omega)}} & \text{if } \omega_n = 0, \\[2ex] v_n^{(1)}(\omega) = \sqrt{\dfrac{p_n(\omega)}{q_n(\omega)}} & \text{if } \omega_n = 1, \end{cases}$$

are the increments of a random walk on $\mathbb{R}$ (with possibly non-independent increments). In this case, the gradient operator defined in (2.3) *reduces to*

$$D_n F(\omega) = \sqrt{p_n(\omega)q_n(\omega)}\left(F(\omega_n^+) - F(\omega_n^-)\right), \quad \omega \in \Omega,$$

where $\omega_n^+ := (\omega_0, \ldots, \omega_{n-1}, 1, \omega_{n+1}, \ldots, \omega_N)$ *and* $\omega_n^- := (\omega_0, \ldots, \omega_{n-1}, 0, \omega_{n+1}, \ldots, \omega_N)$*, and this is precisely the setting of [11]. For background material on stochastic analysis in the discrete setting, see Chapter 1 of [11].*

3 Absolute continuity of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$

In this section we provide sufficient conditions for $\widetilde{\mathbb{P}} \ll \mathbb{P}$, for future use in combination with Theorem 4.1.

We begin with a preliminary lemma.

Lemma 3.1. *Let $i_0 \in \{0, \ldots, d\}$ be fixed, $N = +\infty$, define*

$$\Omega_0 := \bigcup_{n\geq 0} \bigcap_{k\geq n} \{\omega \in \Omega : \omega_k = i_0\} \tag{3.1}$$

and take $\overline{\omega} \in \Omega_0$. Then

$$\mathbb{P}(\{\overline{\omega}\}) = \prod_{n\geq 0} p_n^{(\overline{\omega}_n)}(\overline{\omega})$$

and $\mathbb{P}(\{\overline{\omega}\}) > 0$ if and only if

$$\sum_{n\geq 0} p_n^{(i)}(\overline{\omega}) < \infty, \quad \forall i \in \{0, \ldots, d\} \setminus \{i_0\}.$$

Proof. Define

$$A_k := \{\omega \in \Omega : \omega_0 = \overline{\omega}_0, \ldots, \omega_k = \overline{\omega}_k\}, \quad k \in \mathbb{N}.$$

By iterating (1.2) we have

$$\mathbb{P}\left(\{\omega \in \Omega : \omega_0 = \varepsilon_0, \ldots, \omega_n = \varepsilon_n\}\right) := \prod_{k=0}^{n} p_k^{(\varepsilon_k)}(\varepsilon),$$

$n \in \mathbb{N}$, $\varepsilon_0, \ldots, \varepsilon_n \in \{0, \ldots, d\}$, where $\varepsilon := (\varepsilon_0, \ldots, \varepsilon_n, 0, \ldots, 0) \in \Omega$, and thus

$$\mathbb{P}(A_k) = \prod_{n=0}^{k} p_n^{(\overline{\omega}_n)}(\overline{\omega}),$$

and so, since $(A_k)_{k\geq 0}$ is a non-increasing sequence of events, we have

$$\mathbb{P}\left(\{\overline{\omega}\}\right) = \mathbb{P}\left(\bigcap_{k\geq 0} A_k\right) = \lim_{k\to\infty} \mathbb{P}\left(A_k\right) = \prod_{n\geq 0} p_n^{(\overline{\omega}_n)}(\overline{\omega}).$$

Note that the infinite product is well-defined as a number in $[0,1)$. Letting $k(\overline{\omega})$ denote the integer such that $\overline{\omega}_n = i_0$ for any $n \geq k(\overline{\omega})$, we write

$$\begin{aligned}\mathbb{P}(\{\overline{\omega}\}) = \prod_{n\geq 0} p_n^{(\overline{\omega}_n)}(\overline{\omega}) &= \prod_{n=0}^{k(\overline{\omega})-1} p_n^{(\overline{\omega}_n)}(\overline{\omega}) \prod_{n\geq k(\overline{\omega})} p_n^{(i_0)}(\overline{\omega}) \\ &= \prod_{n=0}^{k(\overline{\omega})-1} p_n^{(\overline{\omega}_n)}(\overline{\omega}) \prod_{n\geq k(\overline{\omega})} \left(1 - \sum_{i\neq i_0} p_n^{(i)}(\overline{\omega})\right). \end{aligned} \tag{3.2}$$

The first product in (3.2) is positive since $p_n^{(i)}(\omega) > 0$ for any i, n and ω. The second product in (3.2) is positive if and only if

$$\sum_{n\geq 0}\sum_{i\neq i_0} p_n^{(i)}(\overline{\omega}) < \infty,$$

see e.g. [6, Theorem 1.9] p. 422. The proof is completed. □

The following proposition holds.

Proposition 3.1. *Assume that there exists* $i_0 \in \{0, \ldots, d\}$ *such that*

$$\widetilde{\mathbb{E}}\left[\sum_{n=0}^{N} \widetilde{p}_n^{(i)}\right] < \infty, \quad \forall i \in \{0, \ldots, d\} \setminus \{i_0\} \tag{3.3}$$

and

$$\sum_{n=0}^{N} p_n^{(i)} < \infty, \quad \widetilde{\mathbb{P}}\text{-a.s.}, \quad \forall i \in \{0, \ldots, d\} \setminus \{i_0\}. \tag{3.4}$$

Then $\widetilde{\mathbb{P}} \ll \mathbb{P}$. *If it is further assumed that*

$$\mathbb{E}\left[\sum_{n=0}^{N} p_n^{(i)}\right] < \infty, \quad \forall i \in \{0, \ldots, d\} \setminus \{i_0\}, \tag{3.5}$$

then the corresponding Radon-Nikodym derivative, say L, *is given by*

$$L(\omega) := \frac{\prod_{n=0}^{N} \widetilde{p}_n^{(\omega_n)}(\omega)}{\prod_{n=0}^{N} p_n^{(\omega_n)}(\omega)}$$

if $\prod_{n=0}^{N} p_n^{(\omega_n)}(\omega) \neq 0$, *and* $L(\omega) = 0$ *otherwise. Note that if* $N \in \mathbb{N}$ *then* (3.3), (3.4) *and* (3.5) *are clearly true and we have* $\prod_{n=0}^{N} p_n^{(\omega_n)}(\omega) \neq 0$.

Proof. We first consider the case $N \in \mathbb{N}$. For any $\omega \in \Omega$, we have

$$\mathbb{P}(\{\omega\}) = \prod_{n=0}^{N} p_n^{(\omega_n)}(\omega) \quad \text{and} \quad \widetilde{\mathbb{P}}(\{\omega\}) = \prod_{n=0}^{N} \widetilde{p}_n^{(\omega_n)}(\omega).$$

Since p and $\widetilde{p}$ are $(0,1)$-valued we deduce $\mathbb{P}(\{\omega\}) > 0$ and $\widetilde{\mathbb{P}}(\{\omega\}) > 0$ for any $\omega \in \Omega$, and so the probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are equivalent (the unique event with null probability is the empty set) and in particular we have the claimed expression for the Radon-Nikodym derivative L. Next, we consider the case where $N = \infty$. Let i_0 be chosen as in the main statement, and define

$$B_n = \{\omega \in \Omega : \omega_n = i_0\}, \quad n \geq 0.$$

By assumption (3.3) we have

$$\sum_{n\geq 0} \widetilde{\mathbb{P}}(\{\omega \in \Omega : \omega_n = i\}) < \infty, \quad \forall\, i \in \{0, \dots, d\} \setminus \{i_0\}.$$

Therefore

$$\sum_{n\geq 0} \widetilde{\mathbb{P}}(B_n^c) = \sum_{n\geq 0} \sum_{i \neq i_0} \widetilde{\mathbb{P}}(\{\omega \in \Omega : \omega_n = i\}) < \infty,$$

where B_n^c denotes the complement of the set B_n. So by Borel-Cantelli's lemma $\widetilde{\mathbb{P}}\big(\limsup_{n\to\infty} B_n^c\big) = 0$, i.e.

$$\widetilde{\mathbb{P}}(\Omega_0) = \widetilde{\mathbb{P}}\left(\liminf_{n\to\infty} B_n\right) = 1, \tag{3.6}$$

where Ω_0 is defined by (3.1). Let $A \in \mathcal{F}$ be such that $\mathbb{P}(A) = 0$. In particular, we also have $\mathbb{P}(A \cap \Omega_0) = 0$. Since Ω_0 is countable we have

$$\sum_{\eta \in \Omega_0 \cap A} \mathbb{P}(\{\eta\}) = 0.$$

It follows by Lemma 3.1 that for all $\eta \in \Omega_0 \cap A$, there exists $i \in \{0, \dots, d\} \setminus \{i_0\}$ (which depends on η) such that

$$\sum_{n\geq 0} p_n^{(i)}(\eta) = \infty.$$

Thus

$$\widetilde{\mathbb{P}}\,(A) = \widetilde{\mathbb{P}}\,(A \cap \Omega_0) \leq \widetilde{\mathbb{P}}\left(\left\{\eta \in \Omega_0 : \sum_{n\geq 0} \sum_{i \neq i_0} p_n^{(i)}(\eta) = \infty\right\}\right) = 0,$$

where we used (3.6) for the former equality and (3.4) for the latter. Let us now further assume that (3.5) holds. By arguments identical to those used in the beginning of the proof, we have $\mathbb{P}(\Omega_0) = 1$. Therefore, for any $A \in \mathcal{F}$, by Lemma 3.1 and (3.4), we deduce

$$\mathbb{E}\,[L\mathbf{1}_A] = \int_\Omega L(\omega)\mathbf{1}_A(\omega)\,\mathbb{P}(\mathrm{d}\omega) = \int_{\Omega_0} L(\eta)\mathbf{1}_A(\eta)\,\mathbb{P}(\mathrm{d}\eta)$$

$$
\begin{aligned}
&= \sum_{\eta \in \Omega_0 \cap A} L(\eta)\, \mathbb{P}(\{\eta\}) \\
&= \sum_{\eta \in \Omega_0 \cap A} L(\eta) \mathbf{1}_{\left\{\sum_{n\geq 0}\sum_{i\neq i_0} p_n^{(i)}(\eta)<\infty\right\}} \mathbb{P}(\{\eta\}) \\
&= \sum_{\eta \in \Omega_0 \cap A} \mathbf{1}_{\left\{\sum_{n\geq 0}\sum_{i\neq i_0} p_n^{(i)}(\eta)<\infty\right\}} \widetilde{\mathbb{P}}(\{\eta\}) \\
&= \sum_{\eta \in \Omega_0 \cap A} \widetilde{\mathbb{P}}(\{\eta\}) \\
&= \widetilde{\mathbb{P}}(A).
\end{aligned}
$$

The proof is completed. □

4 Upper bound on the total variation distance

Let $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ be two probability measures on $(\Omega, \mathcal{F})$ and consider the families

$$
p \equiv ((p_n^{(i)})_{n\in\{0,\dots,N\}})_{0\leq i\leq d} \quad \text{and} \quad \widetilde{p} \equiv ((\widetilde{p}_n^{(i)})_{n\in\{0,\dots,N\}})_{0\leq i\leq d}
$$

of $(0,1)$-valued predictable processes defined in (1.2). We recall that the total variation distance between the probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ is defined as

$$
\mathrm{d}_{\mathrm{TV}}(\mathbb{P},\widetilde{\mathbb{P}}) := \sup_{\substack{F\colon \Omega\to[0,1],\\ F \text{ measurable}}} \left|\mathbb{E}[F] - \widetilde{\mathbb{E}}[F]\right|.
$$

Theorem 4.1 below provides an upper bound on $\mathrm{d}_{\mathrm{TV}}(\mathbb{P},\widetilde{\mathbb{P}})$. In the following, for a vector $v \in \mathbb{R}^d$, we set

$$
|v| := (|v^1|,\dots,|v^d|),
$$

and, for a family of vectors $(v_i)_{i\in I} \subset \mathbb{R}^d$ indexed by a finite set $I \subset \mathbb{N}$, we define

$$
\max_{i\in I} v_i := \Big(\max_{i\in I} v_i^1, \dots, \max_{i\in I} v_i^d\Big).
$$

4.1 A general upper bound

The following theorem is our most general upper bound.

Theorem 4.1. *If $\widetilde{\mathbb{P}} \ll \mathbb{P}$ then for any orthogonal matrix $O \in \mathcal{O}_d$ such that*

$$
\sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left\langle \max_{i\in\{0,\dots,d\}} p_n^{(i)} |Ov_n^{(i)}|, \sum_{i=0}^{d} \widetilde{p}_n^{(i)} |Ov_n^{(i)}| \right\rangle_{\mathbb{R}^d}\right] < \infty, \tag{4.1}
$$

we have

$$\mathrm{d}_{\mathrm{TV}}(\mathbb{P},\widetilde{\mathbb{P}}) \leq d \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left\langle \max_{i\in\{0,\ldots,d\}} p_n^{(i)} |Ov_n^{(i)}|, \left|\sum_{i=0}^{d} \big(\widetilde{p}_n^{(i)} - p_n^{(i)}\big) Ov_n^{(i)}\right| \right\rangle_{\mathbb{R}^d}\right]. \tag{4.2}$$

Remark 2. *(i) Note that if p is equal to $\widetilde{p}$, i.e. $\mathbb{P} = \widetilde{\mathbb{P}}$, then the upper bound in (4.2) is zero.*

(ii) Theorem 4.1 provides an upper bound on the distance between the laws under $\mathbb{P}$, respectively $\tilde{\mathbb{P}}$, of a finite-dimensional random vector $(X_1,\ldots,X_m)$ defined on the sequence space. Indeed, letting $\mathcal{L}_{\mathbb{P}}((X_1,\ldots,X_m))$, respectively $\mathcal{L}_{\tilde{\mathbb{P}}}((X_1,\ldots,X_m))$, denote the law of $(X_1,\ldots,X_m)$ under $\mathbb{P}$, respectively $\tilde{\mathbb{P}}$, we have

$$\begin{aligned}
&\mathrm{d}_{\mathrm{TV}}\left(\mathcal{L}_{\mathbb{P}}((X_1,\ldots,X_m)), \mathcal{L}_{\widetilde{\mathbb{P}}}((X_1,\ldots,X_m))\right)\\
&\quad= \sup_{\substack{f:\mathbb{R}^m\to[0,1],\\ f \text{ measurable}}} \left|\mathbb{E}[f(X_1,\ldots,X_m)] - \widetilde{\mathbb{E}}[f(X_1,\ldots,X_m)]\right|\\
&\quad= \sup_{\substack{f:\mathbb{R}^m\to[0,1],\\ f \text{ measurable}}} \left|\int_\Omega f(X_1(\omega),\ldots,X_m(\omega))\,\mathbb{P}(\mathrm{d}\omega) - \int_\Omega f(X_1(\omega),\ldots,X_m(\omega))\,\widetilde{\mathbb{P}}(\mathrm{d}\omega)\right|\\
&\quad\leq \sup_{\substack{F:\Omega\to[0,1],\\ F \text{ measurable}}} \left|\int_\Omega F(\omega)\,\mathbb{P}(\mathrm{d}\omega) - \int_\Omega F(\omega)\,\widetilde{\mathbb{P}}(\mathrm{d}\omega)\right|\\
&\quad= \mathrm{d}_{\mathrm{TV}}(\mathbb{P},\widetilde{\mathbb{P}}).
\end{aligned}$$

In particular, if $(X_n)_{n\in\{0,\ldots,N\}}$ and $(\widetilde{X}_n)_{n\in\{0,\ldots,N\}}$ are two discrete-time processes defined on the sequence space and respective laws $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ such that $\widetilde{\mathbb{P}} \ll \mathbb{P}$, then the right-hand side of (4.2) yields an upper bound on the finite-dimensional distributions of the processes.

If additionally $X_1,\ldots,X_m$ are $\mathcal{F}_n$-measurable for some $n\in\{0,\ldots,N\}$, then the summation in (4.2) is truncated at n due to the equality $\mathrm{d}_{\mathrm{TV}}(\mathbb{P},\widetilde{\mathbb{P}}) = \mathrm{d}_{\mathrm{TV}}(\mathbb{P}|_{\mathcal{F}_n},\widetilde{\mathbb{P}}|_{\mathcal{F}_n})$ on the restricted laws $\mathbb{P}|_{\mathcal{F}_n}$, $\widetilde{\mathbb{P}}|_{\mathcal{F}_n}$.

The next lemma, whose proof is postponed to the end of the section, is used in the proof of Theorem 4.1.

Lemma 4.1. *Let $m \geq 2$ be a fixed integer, $(\alpha_k)_{k=1,\ldots,m} \subset \mathbb{R}$ be a sequence which is not of constant sign, and $(\beta_k)_{k=1,\ldots,m} \subset [0,1]$. Then,*

$$\left|\sum_{k=1}^{m} \alpha_k \beta_k\right| \leq (m-1) \max_{k=1,\ldots,m} |\alpha_k|.$$

Proof of Theorem 4.1. Let $O \in \mathcal{O}_d$ be an orthogonal matrix satisfying (4.1), and let $F : \Omega \to [0,1]$ be a measurable function. By Proposition 2.2 (with $Ov_n^{(i)}$ in place of $v_n^{(i)}$ in the definitions of D_nF and Y_n, see also Proposition 2.1) and the $(\mathcal{F}_n)_{n\in\{-1,\dots,N\}}$-predictability of the processes $(p_n^{(i)}Ov_n^{(i)})_{n\in\{0,\dots,N\}}$, $i \in \{0,\dots,d\}$, we have

$$\begin{aligned} F(\omega) &= \mathbb{E}[F] + \sum_{n=0}^{N} \langle \mathbb{E}\left[D_nF \mid \mathcal{F}_{n-1}\right](\omega), Y_n(\omega)\rangle_{\mathbb{R}^d} \\ &= \mathbb{E}[F] + \sum_{n=0}^{N}\sum_{i=0}^{d} p_n^{(i)}(\omega)\mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega)\langle Ov_n^{(i)}(\omega), Y_n(\omega)\rangle_{\mathbb{R}^d}, \end{aligned}$$

for $\mathbb{P}$-almost all $\omega \in \Omega$, where F_n^i is the random variable defined by $F_n^i(\omega) := F(\omega_n^i)$. Since $\widetilde{\mathbb{P}} \ll \mathbb{P}$, the previous equality holds $\widetilde{\mathbb{P}}$-almost surely. So, integrating with respect to $\widetilde{\mathbb{P}}$ yields

$$\widetilde{\mathbb{E}}[F] - \mathbb{E}[F] = \sum_{n=0}^{N}\int_\Omega \sum_{i=0}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega)p_n^{(i)}(\omega)\langle Ov_n^{(i)}(\omega), Y_n(\omega)\rangle_{\mathbb{R}^d}\,\widetilde{\mathbb{P}}(\mathrm{d}\omega).$$

Here the exchange between the sum and the integral is justified by Fubini's theorem. Indeed, as shall be checked later on, we have

$$(n,\omega) \mapsto \sum_{i=0}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega)p_n^{(i)}(\omega)\langle Ov_n^{(i)}(\omega), Y_n(\omega)\rangle_{\mathbb{R}^d} \in \mathrm{L}^1\big(\{0,\dots,N\}\times\Omega, \kappa\otimes\widetilde{\mathbb{P}}\big) \tag{4.3}$$

(where κ denotes the counting measure). For fixed $n \in \{0,\dots,N\}$, $\omega \in \Omega$ and $j \in \{1,\dots,d\}$, by the first relation in (2.1) $(p_n^{(i)}(\omega)(Ov_n^{(i)}(\omega))^j)_{0\le i\le d}$ is not of constant sign, and $\mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega) \in [0,1]$. Therefore by Lemma 4.1 we have

$$\begin{aligned} |\widetilde{\mathbb{E}}[F] - \mathbb{E}[F]| &= \left|\sum_{n=0}^{N}\int_\Omega \sum_{i=0}^{d}\sum_{j=1}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega)p_n^{(i)}(\omega)(Ov_n^{(i)}(\omega))^j Y_n(\omega)^j\,\widetilde{\mathbb{P}}(\mathrm{d}\omega)\right| \\ &= \left|\sum_{n=0}^{N}\int_\Omega \sum_{j=1}^{d}\sum_{i=0}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega)p_n^{(i)}(\omega)(Ov_n^{(i)}(\omega))^j Y_n(\omega)^j\,\widetilde{\mathbb{P}}(\mathrm{d}\omega)\right| \\ &\le \sum_{n=0}^{N}\sum_{j=1}^{d}\int_\Omega \left|\sum_{i=0}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega)p_n^{(i)}(\omega)(Ov_n^{(i)}(\omega))^j\right| |\widetilde{\mathbb{E}}\left[Y_n^j \mid \mathcal{F}_{n-1}\right](\omega)|\,\widetilde{\mathbb{P}}(\mathrm{d}\omega) \\ &\le d\sum_{n=0}^{N}\sum_{j=1}^{d}\int_\Omega \left(\max_{i\in\{0,\dots,d\}} p_n^{(i)}(\omega)|(Ov_n^{(i)}(\omega))^j|\right) |\widetilde{\mathbb{E}}\left[Y_n^j \mid \mathcal{F}_{n-1}\right](\omega)|\,\widetilde{\mathbb{P}}(\mathrm{d}\omega) \\ &= d\sum_{n=0}^{N}\widetilde{\mathbb{E}}\left[\left\langle \max_{i\in\{0,\dots,d\}} p_n^{(i)}|Ov_n^{(i)}|, |\widetilde{\mathbb{E}}\left[Y_n \mid \mathcal{F}_{n-1}\right]|\right\rangle_{\mathbb{R}^d}\right] \\ &= d\sum_{n=0}^{N}\widetilde{\mathbb{E}}\left[\left\langle \max_{i\in\{0,\dots,d\}} p_n^{(i)}|Ov_n^{(i)}|, \left|\sum_{i=0}^{d} \widetilde{p}_n^{(i)}Ov_n^{(i)}\right|\right\rangle_{\mathbb{R}^d}\right]. \end{aligned}$$

By (2.1) and Proposition 2.1 we have that $\sum_{i=0}^{d} p_n^{(i)} O v_n^{(i)} = \mathbf{0}$, and the claim (4.2) thus follows by taking the supremum over all the functions F. Recalling that here Y_n is defined with $Ov_n^{(i)}$ in place of $v_n^{(i)}$, the last equality above follows noticing that

$$\widetilde{\mathbb{E}}[Y_n \mid \mathcal{F}_{n-1}] = \widetilde{\mathbb{E}}\left[Ov_n^{(\pi_n)} \mid \mathcal{F}_{n-1}\right] = \sum_{i=0}^{d} \widetilde{p}_n^{(i)} O v_n^{(i)}.$$

It remains to check (4.3). By similar computations, by (4.1) we deduce

$$\begin{aligned}
&\sum_{n=0}^{N} \int_\Omega \left| \sum_{i=0}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega) p_n^{(i)}(\omega) \langle O v_n^{(i)}(\omega), Y_n(\omega)\rangle_{\mathbb{R}^d} \right| \widetilde{\mathbb{P}}(\mathrm{d}\omega) \\
&\quad \leq \sum_{n=0}^{N} \int_\Omega \sum_{j=1}^{d} \left| \sum_{i=0}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega) p_n^{(i)}(\omega) (O v_n^{(i)}(\omega))^j \right| |Y_n(\omega)^j| \,\widetilde{\mathbb{P}}(\mathrm{d}\omega) \\
&\quad = \sum_{n=0}^{N} \int_\Omega \sum_{j=1}^{d} \left| \sum_{i=0}^{d} \mathbb{E}\left[F_n^i \mid \mathcal{F}_{n-1}\right](\omega) p_n^{(i)}(\omega) (O v_n^{(i)}(\omega))^j \right| \widetilde{\mathbb{E}}[|Y_n^j| \mid \mathcal{F}_{n-1}](\omega) \,\widetilde{\mathbb{P}}(\mathrm{d}\omega) \\
&\quad \leq \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left\langle \max_{i\in\{0,\dots,d\}} p_n^{(i)} |O v_n^{(i)}|, \sum_{i=0}^{d} \widetilde{p}_n^{(i)} |O v_n^{(i)}| \right\rangle_{\mathbb{R}^d} \right] < \infty.
\end{aligned}$$

□

Proof of Lemma 4.1. Without loss of generality, assume that $\alpha_1, \dots, \alpha_n \geq 0$ and $\alpha_{n+1}, \dots, \alpha_m \leq 0$ for some $n \in \{1, \dots, m-1\}$. We have

$$\begin{aligned}
\left| \sum_{k=1}^{m} \alpha_k \beta_k \right| &\leq \max\left(\sum_{k=1}^{n} \alpha_k \beta_k, -\sum_{k=n+1}^{m} \alpha_k \beta_k \right) \\
&\leq \max\left(n \max_{k=1,\dots,n} \alpha_k, (m-n) \max_{k=n+1,\dots,m} (-\alpha_k) \right) \\
&\leq (m-1) \max_{k=1,\dots,m} |\alpha_k|.
\end{aligned}$$

□

4.2 Operative upper bounds

In this section, we provide corollaries of Theorem 4.1 which can be directly applied to the Markov chains and random walks considered in the next section. The first two corollaries are specializations of Theorem 4.1 to the cases $d = 1$ and $d = 2$ under a minor strengthening of the assumptions. More specifically, in the first corollary we consider $d = 1$, in which case there exists a unique orthogonal matrix and the bound given by (4.2) is computed explicitly. In the second corollary, we consider

$d = 2$ and for a particular orthogonal matrix we compute explicitly the upper bound in Theorem 4.1. Our last corollary provides a more tractable upper bound for the right-hand side of (4.2) which is valid for any d.

Corollary 4.1. *Let $d = 1$ and assume either $N < \infty$, or*

$$N = \infty \quad \text{and} \quad \sum_{n \geq 0} \widetilde{\mathbb{E}}[\widetilde{p}_n^{(i)} + p_n^{(i)}] < \infty, \quad \text{for some } i \in \{0, 1\}. \tag{4.4}$$

Then

$$\mathrm{d}_{\mathrm{TV}}(\mathbb{P}, \widetilde{\mathbb{P}}) \leq \sum_{n=0}^{N} \widetilde{\mathbb{E}}[|p_n^{(0)} - \widetilde{p}_n^{(0)}|]. \tag{4.5}$$

Proof. In order to simplify the notation, for $n \in \{0, \ldots, N\}$, define $p_n := p_n^{(0)}$, $q_n := p_n^{(1)} = 1 - p_n$, $\widetilde{p}_n := \widetilde{p}_n^{(0)}$ and $\widetilde{q}_n := \widetilde{p}_n^{(1)} = 1 - \widetilde{p}_n$. Let $((v_n^{(i)})_{n \in \{0,\ldots,N\}})_{i=0,1}$ be the stochastic processes defined in Remark 1. It is readily checked that

$$\begin{aligned}
\sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left(\max_{i \in \{0,1\}} p_n^{(i)} |v_n^{(i)}|\right) \sum_{i=0}^{1} \widetilde{p}_n^{(i)} |v_n^{(i)}|\right] &= \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left(\max_{i \in \{0,1\}} p_n^{(i)} |(-1) v_n^{(i)}|\right) \sum_{i=0}^{1} \widetilde{p}_n^{(i)} |(-1) v_n^{(i)}|\right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sqrt{p_n q_n} \left| \widetilde{p}_n \sqrt{\frac{q_n}{p_n}} + \widetilde{q}_n \sqrt{\frac{p_n}{q_n}} \right|\right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}}[\widetilde{p}_n q_n + \widetilde{q}_n p_n].
\end{aligned}$$

This latter quantity is clearly finite if either $N < \infty$ or condition (4.4) holds. Proposition 3.1 along with the assumptions ensure that $\widetilde{\mathbb{P}} \ll \mathbb{P}$. The claim thus follows by Theorem 4.1 noticing that

$$\begin{aligned}
\sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left(\max_{i \in \{0,1\}} p_n^{(i)} |v_n^{(i)}|\right) \left|\sum_{i=0}^{1} \widetilde{p}_n^{(i)} v_n^{(i)}\right|\right] &= \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left(\max_{i \in \{0,1\}} p_n^{(i)} |(-1) v_n^{(i)}|\right) \left|\sum_{i=0}^{1} \widetilde{p}_n^{(i)} (-1) v_n^{(i)}\right|\right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sqrt{p_n q_n} \left| \widetilde{p}_n \sqrt{\frac{q_n}{p_n}} - \widetilde{q}_n \sqrt{\frac{p_n}{q_n}} \right|\right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}}[|\widetilde{p}_n q_n - \widetilde{q}_n p_n|] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[|p_n - \widetilde{p}_n|\right].
\end{aligned}$$

□

Remark 3. *By a close inspection of the proof one may easily realize that condition* (4.4) *can be replaced by the slightly weaker condition:*

$$N = \infty, \quad \sum_{n\geq 0} \widetilde{\mathbb{E}}[\widetilde{p}_n^{(0)} p_n^{(1)} + \widetilde{p}_n^{(1)} p_n^{(0)}] < \infty,$$

as well as

$$\sum_{n\geq 0} \widetilde{\mathbb{E}}[\widetilde{p}_n^{(i)}] < \infty \quad \textit{and} \quad \sum_{n\geq 0} p_n^{(i)} < \infty, \quad \widetilde{\mathbb{P}} - a.s., \quad \textit{for some } i \in \{0,1\}.$$

The next corollary provides an upper bound on the total variation distance when $d = 2$. We mention that the upper bound corresponds to the choice of the orthogonal matrix $O = \mathrm{I}_2$ in (4.2). Clearly, different upper bounds may be obtained by considering other orthogonal matrices in the statement of Theorem 4.1.

Corollary 4.2. *Let $d = 2$ and assume either $N < \infty$ or*

$$N = \infty, \quad \sum_{n\geq 0} \widetilde{\mathbb{E}}[\widetilde{p}_n^{(1)} + \widetilde{p}_n^{(2)}] < \infty \quad \textit{and} \quad \sum_{n\geq 0} \widetilde{\mathbb{E}}[p_n^{(1)} + p_n^{(2)}] < \infty. \tag{4.6}$$

Then

$$\mathrm{d_{TV}}(\mathbb{P}, \widetilde{\mathbb{P}}) \leq 2 \sum_{n=0}^{N} \widetilde{\mathbb{E}} \left[\left| p_n^{(1)} + p_n^{(2)} - \widetilde{p}_n^{(1)} - \widetilde{p}_n^{(2)} + \frac{\widetilde{p}_n^{(1)} p_n^{(2)} - \widetilde{p}_n^{(2)} p_n^{(1)}}{p_n^{(1)} + p_n^{(2)}} \right| \right]. \tag{4.7}$$

Proof. The assumptions easily imply that $\widetilde{\mathbb{P}} \ll \mathbb{P}$ by Proposition 3.1 (take $i_0 = 0$ therein). In order to simplify the notation, for $n \in \{0, \ldots, N\}$, define $p_n := p_n^{(0)}$, $q_n := p_n^{(1)}$, $r_n := p_n^{(2)} = 1 - p_n - q_n$, $\widetilde{p}_n := \widetilde{p}_n^{(0)}$, $\widetilde{q}_n := \widetilde{p}_n^{(1)}$ and $\widetilde{r}_n := \widetilde{p}_n^{(2)} = 1 - \widetilde{p}_n - \widetilde{q}_n$. It is readily checked that, for fixed $n \in \{0, \ldots, N\}$,

$$v_n^{(0)} := \begin{pmatrix} \sqrt{\frac{1-p_n}{p_n}} \\ 0 \end{pmatrix}, \quad v_n^{(1)} := \begin{pmatrix} -\sqrt{\frac{p_n}{1-p_n}} \\ \sqrt{\frac{r_n}{q_n(1-p_n)}} \end{pmatrix}, \quad v_n^{(2)} := \begin{pmatrix} -\sqrt{\frac{p_n}{1-p_n}} \\ -\sqrt{\frac{q_n}{r_n(1-p_n)}} \end{pmatrix}$$

is an obtuse system of $\mathbb{R}^2$ verifying the structure equations (2.1). We first check that (4.1) holds with the identity matrix $O = \mathrm{I}_2$. Indeed, we have

$$\begin{aligned}
&\sum_{n=0}^{N} \widetilde{\mathbb{E}} \left[\left\langle \max_{i\in\{0,\ldots,d\}} p_n^{(i)} |O v_n^{(i)}|, \sum_{i=0}^{d} \widetilde{p}_n^{(i)} |O v_n^{(i)}| \right\rangle_{\mathbb{R}^d} \right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}} \left[\left\langle \begin{pmatrix} \sqrt{p_n(1-p_n)} \\ \sqrt{\frac{q_n r_n}{1-p_n}} \end{pmatrix}, \begin{pmatrix} \widetilde{p}_n \sqrt{\frac{1-p_n}{p_n}} + \widetilde{q}_n \sqrt{\frac{p_n}{1-p_n}} + \widetilde{r}_n \sqrt{\frac{p_n}{1-p_n}} \\ \widetilde{q}_n \sqrt{\frac{r_n}{q_n(1-p_n)}} + \widetilde{r}_n \sqrt{\frac{q_n}{r_n(1-p_n)}} \end{pmatrix} \right\rangle_{\mathbb{R}^d} \right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}} \left[\widetilde{p}_n(1-p_n) + \widetilde{q}_n p_n + \widetilde{r}_n p_n + \frac{\widetilde{q}_n r_n}{1-p_n} + \frac{\widetilde{r}_n q_n}{1-p_n} \right]
\end{aligned}$$

$$\leq \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[q_n + r_n + 2(\widetilde{q}_n + \widetilde{r}_n)\right] < \infty.$$

Hence by (4.2) in Theorem 4.1 and similar computations, we have

$$\begin{aligned} \mathrm{d}_{\mathrm{TV}}(\mathbb{P}, \widetilde{\mathbb{P}}) &\leq 2 \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left|\widetilde{p}_n(1-p_n) - \widetilde{q}_n p_n - \widetilde{r}_n p_n + \frac{\widetilde{q}_n r_n}{1-p_n} - \frac{\widetilde{r}_n q_n}{1-p_n}\right|\right] \\ &= 2 \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left|q_n - \widetilde{q}_n + r_n - \widetilde{r}_n + \frac{\widetilde{q}_n r_n - \widetilde{r}_n q_n}{1-p_n}\right|\right], \end{aligned}$$

which is precisely (4.7). □

Remark 4. *We limit ourselves to remark that the assumptions of Corollary 4.2 can be weakened. Indeed, the choice of $i_0 = 0$ in the proof is clearly arbitrary and, in view of Proposition 3.1, condition (4.6) is not the minimal one to guarantee $\widetilde{\mathbb{P}} \ll \mathbb{P}$.*

Corollary 4.3. *Assume that either $N < \infty$ or $N = \infty$ and there exists $i_0 \in \{0, \ldots, d\}$ such that*

$$\widetilde{\mathbb{E}}\left[\sum_{n \geq 0} \widetilde{p}_n^{(i)}\right] < \infty, \quad \forall i \in \{0, \ldots, d\} \setminus \{i_0\}, \tag{4.8}$$

$$\sum_{n \geq 0} p_n^{(i)} < \infty, \quad \widetilde{\mathbb{P}}\text{-a.s.}, \quad \forall i \in \{0, \ldots, d\} \setminus \{i_0\}, \tag{4.9}$$

and

$$\sum_{n \geq 0} \widetilde{\mathbb{E}}\left[\widetilde{p}_n^{(j)} \sqrt{\frac{p_n^{(i)}(1-p_n^{(i)})(1-p_n^{(j)})}{p_n^{(j)}}}\right] < \infty, \quad \forall i, j \in \{0, \ldots, d\}.$$

Then we have

$$\begin{aligned} \mathrm{d}_{\mathrm{TV}}(\mathbb{P}, \widetilde{\mathbb{P}}) &\leq d \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} \sqrt{p_n^{(i)}(1-p_n^{(i)})} \sum_{i=0}^{d} \sqrt{p_n^{(i)}(1-p_n^{(i)})} \left|1 - \frac{\widetilde{p}_n^{(i)}}{p_n^{(i)}}\right|\right] \\ &\leq \frac{d\sqrt{d+1}}{2} \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} \left|1 - \frac{\widetilde{p}_n^{(i)}}{p_n^{(i)}}\right|\right]. \end{aligned} \tag{4.10}$$

Proof. For ease of notation, in this proof we set $q_n^{(i)} := 1 - p_n^{(i)}$, $n \in \{0, \ldots, N\}$, $i \in \{0, \ldots, d\}$. For any orthogonal matrix $O \in \mathcal{O}_d$, we have

$$\sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left\langle \max_{i \in \{0, \ldots, d\}} p_n^{(i)} |O v_n^{(i)}|, \sum_{i=0}^{d} \widetilde{p}_n^{(i)} |O v_n^{(i)}| \right\rangle_{\mathbb{R}^d}\right]$$

$$
\begin{aligned}
&\le \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left\| \max_{i\in\{0,\dots,d\}} p_n^{(i)} |Ov_n^{(i)}| \right\|_{\mathbb{R}^d} \left\| \sum_{i=0}^{d} \widetilde{p}_n^{(i)} |Ov_n^{(i)}| \right\|_{\mathbb{R}^d}\right] \\
&\le \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} p_n^{(i)} \|Ov_n^{(i)}\|_{\mathbb{R}^d} \sum_{i=0}^{d} \widetilde{p}_n^{(i)} \|Ov_n^{(i)}\|_{\mathbb{R}^d}\right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} p_n^{(i)} \|v_n^{(i)}\|_{\mathbb{R}^d} \sum_{i=0}^{d} \widetilde{p}_n^{(i)} \|v_n^{(i)}\|_{\mathbb{R}^d}\right] \\
&= \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} \sqrt{p_n^{(i)} q_n^{(i)}} \sum_{i=0}^{d} \widetilde{p}_n^{(i)} \sqrt{\frac{q_n^{(i)}}{p_n^{(i)}}}\right] < \infty, \qquad (4.11)
\end{aligned}
$$

where in the last equality we have used (2.2). The quantity appearing in (4.11) is finite when $N = \infty$ by assumption, and it is finite when $N < \infty$ since by a conditioning argument each of the addends is finite. Next, by Theorem 4.1 and again (2.2) we have

$$
\begin{aligned}
\mathrm{d_{TV}}(\mathbb{P}, \widetilde{\mathbb{P}}) &\le d \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left\langle \max_{i\in\{0,\dots,d\}} p_n^{(i)} |Ov_n^{(i)}|, \left| \sum_{i=0}^{d} (\widetilde{p}_n^{(i)} - p_n^{(i)}) Ov_n^{(i)} \right| \right\rangle_{\mathbb{R}^d}\right] \\
&\le d \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\left\| \max_{i\in\{0,\dots,d\}} p_n^{(i)} |Ov_n^{(i)}| \right\|_{\mathbb{R}^d} \left\| \sum_{i=0}^{d} (\widetilde{p}_n^{(i)} - p_n^{(i)}) Ov_n^{(i)} \right\|_{\mathbb{R}^d}\right] \\
&\le d \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} p_n^{(i)} \|Ov_n^{(i)}\|_{\mathbb{R}^d} \sum_{i=0}^{d} |\widetilde{p}_n^{(i)} - p_n^{(i)}| \|Ov_n^{(i)}\|_{\mathbb{R}^d}\right] \\
&= d \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} p_n^{(i)} \|v_n^{(i)}\|_{\mathbb{R}^d} \sum_{i=0}^{d} |\widetilde{p}_n^{(i)} - p_n^{(i)}| \|v_n^{(i)}\|_{\mathbb{R}^d}\right] \\
&= d \sum_{n=0}^{N} \widetilde{\mathbb{E}}\left[\sum_{i=0}^{d} \sqrt{p_n^{(i)} q_n^{(i)}} \sum_{i=0}^{d} \sqrt{p_n^{(i)} q_n^{(i)}} \left| 1 - \frac{\widetilde{p}_n^{(i)}}{p_n^{(i)}} \right|\right].
\end{aligned}
$$

The last inequality in (4.10) follows since $\sqrt{p_n^{(i)} q_n^{(i)}} = \sqrt{p_n^{(i)}(1 - p_n^{(i)})} \le 1/2$ and

$$
\sum_{i=0}^{d} \sqrt{p_n^{(i)} q_n^{(i)}} \le (d+1) \sum_{i=0}^{d} \frac{1}{d+1} \sqrt{p_n^{(i)}} \le (d+1) \sqrt{\sum_{i=0}^{d} \frac{1}{d+1} p_n^{(i)}} = \sqrt{d+1}.
$$

□

5 Applications

In the next subsections we provide upper bounds on the total variation distance between the laws of two discrete-time Markov chains with finite state space on the one hand, and between those of two random walks on $\mathbb{Z}$ on the other hand. Without loss of generality, we shall assume that the above stochastic processes are defined on a same probability space $(A, \mathcal{A}, P)$, denoting by E the expectation under P.

5.1 Markov chains

In this section, we use Corollary 4.3 to bound the total variation distance between two Markov chains. We start by constructing a finite state space Markov chain in our framework.

Let $(X_n)_{n\in\{0,\dots,N\}}$ be a Markov chain on $(A, \mathcal{A}, P)$ with state space $\{0,\dots,d\}$, and define

$$p_0^{(i)} := P(X_0 = i) \in (0,1),$$

$i \in \{0,\dots,d\}$, and for each $n \in \{1,\dots,N\}$,

$$\alpha_n^{(k,i)} := P(X_n = i \mid X_{n-1} = k) \in (0,1), \quad k \in \{0,\dots,d\}.$$

For $\omega \in \Omega = \{0,\dots,d\}^{\{0,\dots,N\}}$, we define the predictable sequence $(p_n^{(i)})_{n\in\{0,\dots,N\}}$, $i \in \{0,\dots,d\}$, by

$$\begin{cases} p_n^{(i)}(\omega) := \alpha_n^{(\omega_{n-1},i)}, & 1 \le n \le N, \\ p_0^{(i)}(\omega) := p_0^{(i)}. \end{cases}$$

A straightforward computation shows that, under $\mathbb{P}$, $(\pi_n)_{n\in\{0,\dots,N\}}$ has the same distribution as $(X_n)_{n\in\{0,\dots,N\}}$.

In the following, we shall provide bounds on the total variation distance between the laws of $(X_n)_{n\in\{0,\dots,N\}}$ and of another Markov chain $(\widetilde{X}_n)_{n\in\{0,\dots,N\}}$ with state space $\{0,\dots,d\}$. For concreteness, we assume that $(\widetilde{X}_n)_{n\in\{0,\dots,N\}}$ is also defined on $(A, \mathcal{A}, P)$ and we put

$$\widetilde{p}_0^{(i)} := P(\widetilde{X}_0 = i) \in (0,1),$$

$i \in \{0,\dots,d\}$, and for each $n \in \{1,\dots,N\}$,

$$\widetilde{\alpha}_n^{(k,i)} := P(\widetilde{X}_n = i \mid \widetilde{X}_{n-1} = k) \in (0,1), \quad k \in \{0,\dots,d\},$$

and we construct analogously the predictable sequence $(\widetilde{p}_n^{(i)})_{n\in\{0,\dots,N\}}$, $i \in \{0,\dots,d\}$. Hereafter, we denote by $\mathcal{L}$ (respectively $\widetilde{\mathcal{L}}$) the law of $(X_n)_{n\in\{0,\dots,N\}}$ (respectively $(\widetilde{X}_n)_{n\in\{0,\dots,N\}}$). The inequality (5.4) below is an immediate consequence of Corollary 4.3, and (5.5) follows by a simple conditioning argument.

Corollary 5.1. *Assume either that $N < \infty$ or that $N = \infty$ and that there exists $i_0 \in \{0, \ldots, d\}$ such that*

$$\sum_{n\geq 1} E\Big[\widetilde{\alpha}_n^{(\widetilde{X}_{n-1},i)}\Big] < \infty, \quad \forall i \in \{0,\ldots,d\} \setminus \{i_0\}, \tag{5.1}$$

$$\sum_{n\geq 1} \alpha_n^{(\widetilde{X}_{n-1},i)} < \infty, \quad P\text{-a.s.}, \quad \forall i \in \{0,\ldots,d\} \setminus \{i_0\}, \tag{5.2}$$

and

$$\sum_{n\geq 1} E\left[\widetilde{\alpha}_n^{(\widetilde{X}_{n-1},j)} \sqrt{\frac{\alpha_n^{(\widetilde{X}_{n-1},i)}(1-\alpha_n^{(\widetilde{X}_{n-1},i)})(1-\alpha_n^{(\widetilde{X}_{n-1},j)})}{\alpha_n^{(\widetilde{X}_{n-1},j)}}}\right] < \infty, \quad \forall i, j \in \{0,\ldots,d\}. \tag{5.3}$$

Then we have

$$\begin{aligned} \mathrm{d_{TV}}(\mathcal{L}, \widetilde{\mathcal{L}}) &\leq d \sum_{i,j=0}^{d} \left(\sqrt{\frac{p_0^{(i)}(1-p_0^{(i)})(1-p_0^{(j)})}{p_0^{(j)}}} \left| p_0^{(j)} - \widetilde{p}_0^{(j)} \right| \right. \\ &\quad \left. + \sum_{n=1}^{N} E\left[\sqrt{\frac{\alpha_n^{(\widetilde{X}_{n-1},i)}(1-\alpha_n^{(\widetilde{X}_{n-1},i)})(1-\alpha_n^{(\widetilde{X}_{n-1},j)})}{\alpha_n^{(\widetilde{X}_{n-1},j)}}} \left| \alpha_n^{(\widetilde{X}_{n-1},j)} - \widetilde{\alpha}_n^{(\widetilde{X}_{n-1},j)} \right| \right] \right) &(5.4) \\ &\leq d \sum_{i,j=0}^{d} \left(\sqrt{\frac{p_0^{(i)}(1-p_0^{(i)})(1-p_0^{(j)})}{p_0^{(j)}}} \left| p_0^{(j)} - \widetilde{p}_0^{(j)} \right| \right. \\ &\quad \left. + \sum_{n=1}^{N} \sum_{k=0}^{d} \sqrt{\frac{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})(1-\alpha_n^{(k,j)})}{\alpha_n^{(k,j)}}} \left| \alpha_n^{(k,j)} - \widetilde{\alpha}_n^{(k,j)} \right| \right). &(5.5) \end{aligned}$$

For Markov chains with two and three states, we remark that the bound provided by Corollary 5.1 can be slightly improved by a direct application of Corollaries 4.1 and 4.2, respectively. Taking $\widetilde{\alpha}$ depending on a parameter $\theta \in \Theta$, one can minimize the right-hand side of (5.5) to obtain the optimal estimation of (X_n) in a parametrized class of Markov chains (X_n^θ).

In the following example we apply Corollary 5.1 to estimate the total variation distance between two particular finite state space Markov chains.

Example 1. We consider two finite state space Markov chains and compute an upper bound to their total variation distance by applying Corollary 5.1. Here, $(\widetilde{X}_n)$ is the Markov chain obtained via a multiplicative perturbation of the transition matrix and the initial law of (X_n). More specifically, for $\varepsilon \in (0,1)$ we set

$$\widetilde{p}_0^{(i)} := (1-\varepsilon)p_0^{(i)},$$
$$\widetilde{\alpha}_n^{(k,i)} := (1-\varepsilon)\alpha_n^{(k,i)},$$

and

$$\widetilde{\alpha}_n^{(k,0)} := 1 - (1-\varepsilon)\sum_{i=1}^d \alpha_n^{(k,i)}, \qquad \widetilde{p}_0^{(0)} := 1 - (1-\varepsilon)\sum_{i=1}^d p_0^{(i)}.$$

For $k \in \{0,\dots,d\}$ and $n \in \{1,\dots,N\}$ we have

$$\big|p_0^{(i)} - \widetilde{p}_0^{(i)}\big| = \varepsilon p_0^{(i)}, \ \big|\alpha_n^{(k,i)} - \widetilde{\alpha}_n^{(k,i)}\big| = \varepsilon \alpha_n^{(k,i)}, \quad i \in \{1,\dots,d\}, \tag{5.6}$$

and

$$\big|p_0^{(0)} - \widetilde{p}_0^{(0)}\big| = \varepsilon(1-p_0^{(0)}), \ \big|\alpha_n^{(k,0)} - \widetilde{\alpha}_n^{(k,0)}\big| = \varepsilon(1-\alpha_n^{(k,0)}). \tag{5.7}$$

We assume that

$$\sum_{n\ge1} \alpha_n^{(k,i)} < \infty, \quad \forall k \in \{0,\dots,d\},\ \forall i \in \{1,\dots,d\}, \tag{5.8}$$

and either $N < \infty$ or $N = \infty$ and

$$\sum_{n\ge1} \sqrt{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})\alpha_n^{(k,j)}(1-\alpha_n^{(k,j)})} < \infty, \quad \forall i,j,k \in \{0,\dots,d\}, \tag{5.9}$$

as well as

$$\sum_{n\ge1} \sqrt{\frac{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})(1-\alpha_n^{(k,0)})^3}{\alpha_n^{(k,0)}}} < \infty, \quad \forall i,k \in \{0,\dots,d\}. \tag{5.10}$$

In this case, for any $i,k \in \{0,\dots,d\}$ we have

$$\sum_{n\ge1} \widetilde{\alpha}_n^{(k,j)} \sqrt{\frac{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})(1-\alpha_n^{(k,j)})}{\alpha_n^{(k,j)}}} = (1-\varepsilon)\sum_{n\ge1}\sqrt{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})\alpha_n^{(k,j)}(1-\alpha_n^{(k,j)})} < \infty,$$

for any $j \in \{1,\dots,d\}$, and

$$\begin{aligned}
&\sum_{n\ge1} \widetilde{\alpha}_n^{(k,0)} \sqrt{\frac{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})(1-\alpha_n^{(k,0)})}{\alpha_n^{(k,0)}}} \\
&= \sum_{n\ge1}(\alpha_n^{(k,0)} + \varepsilon(1-\alpha_n^{(k,0)}))\sqrt{\frac{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})(1-\alpha_n^{(k,0)})}{\alpha_n^{(k,0)}}} \\
&= \sum_{n\ge1}\sqrt{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})\alpha_n^{(k,0)}(1-\alpha_n^{(k,0)})} + \varepsilon\sum_{n\ge1}\sqrt{\frac{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})(1-\alpha_n^{(k,0)})^3}{\alpha_n^{(k,0)}}} \\
&< \infty,
\end{aligned}$$

and so (5.3) holds. By Corollary 5.1, (5.6) and (5.7), we have

$$\mathrm{d_{TV}}(\mathcal{L},\widetilde{\mathcal{L}}) \le d\varepsilon \sum_{i=0}^d \left[\sqrt{\frac{p_0^{(i)}(1-p_0^{(i)})(1-p_0^{(0)})^3}{p_0^{(0)}}} + \sum_{n=1}^N\sum_{k=0}^d \sqrt{\frac{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})(1-\alpha_n^{(k,0)})^3}{\alpha_n^{(k,0)}}}\right.$$

$$+\sum_{j=1}^{d}\left(\sqrt{p_0^{(i)}(1-p_0^{(i)})p_0^{(j)}(1-p_0^{(j)})}+\sum_{n=1}^{N}\sum_{k=0}^{d}\sqrt{\alpha_n^{(k,i)}(1-\alpha_n^{(k,i)})\alpha_n^{(k,j)}(1-\alpha_n^{(k,j)})}\right)\Bigg]$$
$$\leq C\varepsilon,$$

where $C > 0$ is an explicit constant which does not depend on ε.

As an example of sequence satisfying conditions (5.8), (5.9) and (5.10), one can take $\alpha_n^{(k,i)} = a_n$, for any $i \in \{1, \dots, d\}$, $k \in \{0, \dots, d\}$ and $n \in \{1, 2, \dots\}$. Here, a_n is a deterministic sequence with values in $(0, 1/d)$ which is assumed to satisfy $\sum_{n\geq 1} a_n < \infty$.

5.2 Random walks

In this section, we use Corollaries 4.1 and 4.2 to bound the total variation distance between two different random walks on $\mathbb{Z}$. As will be made precise in Remark 5 below, the processes considered here are not in general Markov chains. Thus, the results of this section do not follow from those one in Section 5.1.

5.2.1 Nearest neighbor random walks

Let $(X_n)_{n\in\{0,\dots,N\}}$ be a sequence of (not necessarily i.i.d.) $\{-1, +1\}$-valued random variables defined on $(A, \mathcal{A}, P)$ and characterized by

$$p_0 := P(X_0 = +1) \in (0, 1),$$

and, for each $n \in \{0, \dots, N-1\}$,

$$\alpha_{n+1}(x_0, \dots, x_n) := P(X_{n+1} = +1 \mid X_0 = x_0, \dots, X_n = x_n) \in (0, 1), \quad \{x_0, \dots, x_n\} \in \{-1, +1\}^{n+1}. \tag{5.11}$$

We define the associated random walk by

$$S_n := \sum_{l=0}^{n} X_l, \quad n \in \{0, \dots, N\}.$$

Next, we note that the process $(X_n)_{n\in\{0,\dots,N\}}$ can be replicated in our framework by setting $d = 1$ and applying Corollary 4.1. For $\omega \in \Omega = \{0, 1\}^{\{0,1,\dots,N\}}$, we define the predictable sequence $(p_n^{(1)})_{n\in\{0,\dots,N\}}$ by

$$p_n^{(1)}(\omega) := \alpha_n\left((-1)^{1+\pi_0(\omega)}, \dots, (-1)^{1+\pi_{n-1}(\omega)}\right), \quad n \geq 1, \quad p_0^{(1)}(\omega) = p_0.$$

We accordingly define the predictable sequence $(p_n^{(0)})_{n\in\{0,\dots,N\}}$ by $p_n^{(0)} := 1 - p_n^{(1)}$. A straightforward computation shows that, under $\mathbb{P}$, $((-1)^{1+\pi_n})_{n\in\{0,\dots,N\}}$ has the same law as the increments

$(X_n)_{n\in\{0,\dots,N\}}$. Consequently, under $\mathbb{P}$, $(\sum_{l=0}^{n}(-1)^{1+\pi_l})_{n\in\{0,\dots,N\}}$ has the same law as $(S_n)_{n\in\{0,\dots,N\}}$. In the following, we shall provide bounds on the total variation distance between the laws of $(S_n)_{n\in\{0,\dots,N\}}$ and of another random walk $(\widetilde{S}_n)_{n\in\{0,\dots,N\}}$ defined according to another sequence $(\widetilde{X}_n)_{n\in\{0,\dots,N\}} \subset \{-1,+1\}$ on $(A,\mathcal{A},P)$. We put

$$\widetilde{p}_0 := P(\widetilde{X}_0 = +1) \in (0,1),$$

and for each $n \in \{1,\dots,N\}$,

$$\widetilde{\alpha}_{n+1}(x_0,\dots,x_n) := P(\widetilde{X}_{n+1} = +1 \mid \widetilde{X}_0 = x_0,\dots,\widetilde{X}_n = x_n) \in (0,1), \quad \{x_0,\dots,x_n\} \in \{-1,+1\}^{n+1},$$

and we construct analogously the predictable sequence $(\widetilde{p}_n^{(i)})_{n\in\{0,\dots,N\}}$, $i \in \{0,1\}$. Hereafter, we denote by $\mathcal{L}$ (respectively $\widetilde{\mathcal{L}}$) the law of $(S_n)_{n\in\{0,\dots,N\}}$ (respectively $(\widetilde{S}_n)_{n\in\{0,\dots,N\}}$). The following bound on the total variation distance between the (laws of the) two nearest neighbor random walks is an immediate consequence of the above construction and Corollary 4.1.

Corollary 5.2. *If either $N < \infty$, or $N = \infty$ and*

$$\max\left\{\sum_{n\ge1} E[\alpha_n(\widetilde{X}_0,\dots,\widetilde{X}_{n-1})], \sum_{n\ge1} E[\widetilde{\alpha}_n(\widetilde{X}_0,\dots,\widetilde{X}_{n-1})]\right\} < \infty,$$

then

$$\mathrm{d}_{\mathrm{TV}}(\mathcal{L},\widetilde{\mathcal{L}}) \le |p_0 - \widetilde{p}_0| + \sum_{n=1}^{N} E\big[|\alpha_n(\widetilde{X}_0,\dots,\widetilde{X}_{n-1}) - \widetilde{\alpha}_n(\widetilde{X}_0,\dots,\widetilde{X}_{n-1})|\big]. \tag{5.12}$$

Remark 5. *Note that $(S_n)_{n\in\{0,\dots,N\}}$ is in general not a Markov chain. Indeed, with obvious notation,*

$$\begin{aligned}
&P\big(S_n = s_n \mid S_{n-1} = s_{n-1},\dots,S_0 = s_0\big)\\
&\qquad = P\big(X_n = s_n - s_{n-1} \mid X_{n-1} = s_{n-1} - s_{n-2},\dots,X_1 = s_1 - s_0, X_0 = s_0\big)\\
&\qquad = f_n(s_{n-1} - s_{n-2},\dots,s_1 - s_0, s_0),
\end{aligned}$$

where $f_n = \alpha_n$ if $s_n = s_{n-1} + 1$ and $f_n = 1 - \alpha_n$ otherwise. The above is in general not equal to $P\big(S_n = s_n \mid S_{n-1} = s_{n-1}\big)$.

Example 2. We apply Corollary 5.2 to elephant random walks (see [12]; see also [4]). Given the parameters $p, r \in (0,1)$, the elephant random walks is defined as follows. At time zero, the probability that the random walk goes to the right is r, i.e. $p_0 = r$. At the time step $n > 0$, one draws an integer $k \in \{0,\dots,n-1\}$ with uniform probability. Then random walk makes an increment $X_n \in \{-1,+1\}$ equal to X_k with probability p, and equal to $-X_k$ with probability $1-p$ (note that if $p = 1/2$, then the elephant random walk corresponds to the simple and symmetric random walk on $\mathbb{Z}$). It is readily checked that the transition probabilities α_n defined in (5.11) are given by

$$\alpha_n(x_0,\dots,x_{n-1}) = \frac{1}{2n}\sum_{k=0}^{n-1}\bigl(1+(2p-1)x_k\bigr), \quad n>0,\ x_0,\dots,x_{n-1}\in\{-1,+1\}.$$

Let $\mathcal{L}$ denote the law of $(X_n)_{0\le n\le N}$ and $\widetilde{\mathcal{L}}$ the law of another elephant random walk $(\widetilde{X}_n)_{0\le n\le N}$ characterized by the parameters $\widetilde{r}$, $\widetilde{p}\in(0,1)$. If $N<\infty$, then by Corollary 5.2 we have

$$\mathrm{d}_{\mathrm{TV}}(\mathcal{L},\widetilde{\mathcal{L}}) \le |r-\widetilde{r}| + |p-\widetilde{p}|\sum_{n=1}^{N} E\Bigl[\Bigl|\frac{1}{n}\sum_{k=0}^{n-1}\widetilde{X}_k\Bigr|\Bigr]. \tag{5.13}$$

Due to recent results in [7], we remark that $E\bigl[\bigl|n^{-1}\sum_{k=0}^{n-1}\widetilde{X}_k\bigr|\bigr]$ is asymptotically equivalent to $(2\widetilde{q}-1)\Gamma(2\widetilde{p})^{-1}n^{2(\widetilde{p}-1)}$ as n goes to infinity, where Γ is the Euler gamma function.

5.2.2 Lazy nearest neighbor random walks

Let $(X_n)_{n\in\{0,\dots,N\}}$ be a sequence of (not necessarily i.i.d.) $\{-1,0,+1\}$-valued random variables defined on $(A,\mathcal{A},P)$ and characterized by

$$p_0 := P(X_0=1)\in(0,1), \quad q_0 := P(X_0=0)\in(0,1),$$

and, for each $n\in\{0,\dots,N-1\}$ and $x_0,\dots,x_n\in\{-1,0,+1\}$,

$$\alpha_{n+1}(x_0,\dots,x_n) := P(X_{n+1}=1 \mid X_0=x_0,\dots,X_n=x_n)\in(0,1), \tag{5.14}$$

$$\beta_{n+1}(x_0,\dots,x_n) := P(X_{n+1}=0 \mid X_0=x_0,\dots,X_n=x_n)\in(0,1), \tag{5.15}$$

We define the associated random walk by

$$S_n := \sum_{l=0}^{n} X_l, \quad n\in\{0,\dots,N\}.$$

Next, we note that the process $(X_n)_{n\in\{0,\dots,N\}}$ can be replicated in our framework by setting $d=2$ and applying Corollary 4.2. For $\omega\in\Omega=\{0,1,2\}^{\{0,1,\dots,N\}}$, we define the predictable sequences $(p_n^{(1)})_{n\in\{0,\dots,N\}}$, $(p_n^{(2)})_{n\in\{0,\dots,N\}}$, by

$$p_n^{(2)}(\omega) := \alpha_n\left((1-\pi_0(\omega))(-1)^{1+\pi_0(\omega)},\dots,(1-\pi_{n-1}(\omega))(-1)^{1+\pi_{n-1}(\omega)}\right), \quad p_0^{(1)}(\omega)=p_0,$$

$$p_n^{(1)}(\omega) := \beta_n\left((1-\pi_0(\omega))(-1)^{1+\pi_0(\omega)},\dots,(1-\pi_{n-1}(\omega))(-1)^{1+\pi_{n-1}(\omega)}\right), \quad p_0^{(2)}(\omega)=q_0,$$

where $1 \leq n \leq N$. We accordingly define the predictable sequences $(p_n^{(0)})_{n\in\{0,\ldots,N\}}$ by $p_n^{(0)} := 1 - p_n^{(1)} - p_n^{(2)}$. A straightforward computation shows that, under $\mathbb{P}$, $((1-\pi_n)(-1)^{1+\pi_n})_{n\in\{0,\ldots,N\}}$ has the same law as the increments $(X_n)_{n\in\{0,\ldots,N\}}$. As a consequence, under $\mathbb{P}$, $(\sum_{l=0}^n (1-\pi_l)(-1)^{1+\pi_l})_{n\in\{0,\ldots,N\}}$ has the same law as $(S_n)_{n\in\{0,\ldots,N\}}$.

In the following, we shall provide bounds on the total variation distance between the laws of $(S_n)_{n\in\{0,\ldots,N\}}$ and of another random walk $(\widetilde{S}_n)_{n\in\{0,\ldots,N\}}$ defined according to another sequence $(\widetilde{X}_n)_{n\in\{0,\ldots,N\}} \subset \{-1,0,+1\}$ on $(A, \mathcal{A}, P)$. We put

$$\widetilde{p}_0 := P(\widetilde{X}_0 = 1) \in (0,1), \quad \widetilde{q}_0 := P(\widetilde{X}_0 = 0) \in (0,1),$$

and for each $n \in \{1, \ldots, N\}$, we define

$$\widetilde{\alpha}_{n+1}(x_0, \ldots, x_n) := P(\widetilde{X}_{n+1} = +1 \mid \widetilde{X}_0 = x_0, \ldots, \widetilde{X}_n = x_n) \in (0,1),$$

and

$$\widetilde{\beta}_{n+1}(x_0, \ldots, x_n) := P(\widetilde{X}_{n+1} = 0 \mid \widetilde{X}_0 = x_0, \ldots, \widetilde{X}_n = x_n) \in (0,1),$$

and we construct analogously the predictable sequence $(\widetilde{p}_n^{(i)})_{n\in\{0,\ldots,N\}}$, $i \in \{0,1,2\}$. Hereafter, we denote by $\mathcal{L}$ (respectively $\widetilde{\mathcal{L}}$) the law of $(S_n)_{n\in\{0,\ldots,N\}}$ (respectively $(\widetilde{S}_n)_{n\in\{0,\ldots,N\}}$). The following bound on the total variation distance between the (laws of the) two lazy nearest neighbor random walks is an immediate consequence of the above construction and Corollary 4.2.

Corollary 5.3. *If either $N < \infty$, or $N = \infty$ and*

$$\sum_{n\geq 1} E[\alpha_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) + \beta_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) + \widetilde{\alpha}_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) + \widetilde{\beta}_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1})] < \infty,$$

then

$$\begin{aligned}
&\mathrm{d_{TV}}(\mathcal{L}, \widetilde{\mathcal{L}}) \\
&\leq 2\left| p_0 + q_0 - \widetilde{p}_0 - \widetilde{q}_0 + \frac{\widetilde{p}_0 q_0 - \widetilde{q}_0 p_0}{p_0 + q_0} \right| \\
&\quad + 2\sum_{n=1}^N E\Bigg[\Bigg| \alpha_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) + \beta_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) - \widetilde{\alpha}_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) - \widetilde{\beta}_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) \\
&\quad + \frac{\widetilde{\alpha}_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1})\beta_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) - \widetilde{\beta}_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1})\alpha_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1})}{\alpha_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1}) + \beta_n(\widetilde{X}_0, \ldots, \widetilde{X}_{n-1})} \Bigg| \Bigg]. \qquad (5.16)
\end{aligned}$$

Note that as in Remark 5, one can easily check that the lazy nearest neighbor random walks considered above are in general not Markov chains.

Example 3. A generalization of the elephant random walk considered in Example 2 has been proposed in [9]. Specifically, the random walk is parametrized by $p, q, r \in (0, 1)$ and, at time zero, $p_0 = r$ and $q_0 = 0$. At the time step $n > 0$, one draws an integer $k \in \{0, \ldots, n-1\}$ with uniform probability. Then the random walk makes an increment makes $X_n \in \{-1, 0, +1\}$ equal to X_k with probability p, equal to $-X_k$ with probability q, and stays put, i.e. $X_n = 0$, otherwise. The transition probabilities defined in (5.14) and (5.15) are given by

$$\alpha_n(x_0, \ldots, x_{n-1}) = \frac{1}{2n} \sum_{k=0}^{n-1} \big(x_k^2 (p+q) + x_k (p-q) \big), \quad n > 0, \ x_0, \ldots, x_{n-1} \in \{-1, 0, +1\},$$

and

$$\beta_n(x_0, \ldots, x_{n-1}) = 1 - \frac{p+q}{n} \sum_{k=0}^{n-1} x_k^2, \quad n > 0, \ x_0, \ldots, x_{n-1} \in \{-1, 0, +1\},$$

see [9]. Let $\mathcal{L}$ denote the law of $(X_n)_{0 \le n \le N}$ and $\widetilde{\mathcal{L}}$ the law of another random walk $(\widetilde{X}_n)_{0 \le n \le N}$ of this kind characterized by the parameters $\widetilde{p}, \widetilde{q}, \widetilde{r} \in (0, 1)$. Under the assumption $N < \infty$, by Corollary 5.3 we have

$$\begin{aligned}
&\mathrm{d}_{\mathrm{TV}}(\mathcal{L}, \widetilde{\mathcal{L}}) \\
&\le 2 \big| r - \widetilde{r} \big| \\
&+ 2 \sum_{n=1}^{N} E \Bigg[\Bigg| \frac{1}{2n} \sum_{k=0}^{n-1} \big(\widetilde{X}_k (p - q - (\widetilde{p} - \widetilde{q})) - \widetilde{X}_k^2 (p + q - (\widetilde{p} + \widetilde{q})) \big) \\
&+ \frac{(2n)^{-1} \sum_{k=0}^{n-1} \big(\widetilde{X}_k (\widetilde{p} - \widetilde{q} - (p - q)) + \widetilde{X}_k^2 (\widetilde{p} + \widetilde{q} - (p + q)) \big) + n^{-2} \big(p\widetilde{q} - \widetilde{p}q \big) \sum_{k=0}^{n-1} \widetilde{X}_k^2 \sum_{i=0}^{n-1} \widetilde{X}_i}{1 + (2n)^{-1} \sum_{k=0}^{n-1} \big(\widetilde{X}_k (p - q) - \widetilde{X}_k^2 (p + q) \big)} \Bigg| \Bigg].
\end{aligned}$$

References

[1] S. Attal and A. Dhahri. Repeated quantum interactions and unitary random walks. *Journal of Theoretical Probability*, 23: 345–361, 2010.

[2] S. Attal, J. Deschamps and C. Pellegrini. Complex obtuse random walks and their continuous-time limits. *Probability Theory and Related Fields*, 165: 65–116, 2016.

[3] S. Attal and M. Émery. Équations de structure pour des martingales vectorielles. In *Séminaire de Probabilités, XXVIII*, volume 1583 of *Lecture Notes in Mathematics*, pp. 256–278. Springer-Verlag, Berlin, 1994.

[4] E. Baur and J. Bertoin. Elephant random walks and their connection to Pólya-type urns. *Physical Review E*, 94: 051134, 2016.

[5] P. Billingsley. *Probability and Measure*, volume 245 of *Wiley Series in Probability and Statistics*, third edition. Wiley, 1995.

[6] P. Brémaud. *Markov Chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999.

[7] C.F. Coletti, R. Gava and G.M. Schütz Central limit theorem and related results for the elephant random walk. *Journal of Mathematical Physics*, 58: 053303, 2017.

[8] U. Franz and T. Hamdi. Stochastic analysis for obtuse random walks. *Journal of Theoretical Probability*, 28: 619–649, 2015.

[9] N. Kumar, U. Harbola and K. Lindenberg. Memory-induced anomalous dynamics: Emergence of diffusion, subdiffusion, and superdiffusion from a single random walk model. *Physical Review E*, 82: 021101, 2010.

[10] D.A. Levin, Y. Peres and E.L. Wilmer. *Markov chains and mixing times.* American Mathematical Society, Providence, RI, 2009.

[11] N. Privault. *Stochastic analysis in discrete and continuous settings with normal martingales*, volume 1982 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2009.

[12] G.M. Schütz and S. Trimper. Elephants can always remember: exact long-range memory effects in a non-Markovian random walk. *Physical Review E*, 70: 045101, 2004.

[13] J. Song, Y. Gao, H. Wang and B. An. Measuring the distance between finite Markov decision processes. In *Proceedings of the 15th International Conference on Autononous Agents and Multiagents Systems.* J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9/13, 2016, Singapore.