

# HITTING TIME, ACCESS TIME AND OPTIMAL TRANSPORT ON GRAPHS

MICHAEL C.H. CHOI

**ABSTRACT.** Given a discrete source distribution  $\mu$  and discrete target distribution  $\nu$  on a common finite state space  $\mathcal{X}$ , we are tasked with transporting  $\mu$  to  $\nu$  using a given discrete-time Markov chain  $X$  with the quickest possible time on average. We define the optimal transport time  $H(\mu, \nu)$  as stopping rule of  $X$  that gives the minimal expected transport time. This is also known as the access time from  $\mu$  to  $\nu$  of  $X$  in [L. Lovász and P. Winkler. Efficient Stopping Rules for Markov Chains. Proceedings of the Twenty-seventh Annual ACM Symposium on Theory of Computing (STOC '95) 76-82.]. We study bounds of  $H(\mu, \nu)$  in various special graphs, which are expressed in terms of the mean hitting times of  $X$  as well as parameters of  $\mu$  and  $\nu$  such as their moments. Among the Markov chains that we study, random walks on complete graphs is a good choice for transport as  $H(\mu, \nu)$  grows linearly in  $n$ , the size of the state space, while that of the winning streak Markov chain exhibits exponential dependence in  $n$ .

**AMS 2010 subject classifications:** 60J10

**Keywords:** hitting time; stopping rule; Markov chains; Skorohod embedding; optimal transport

## 1. INTRODUCTION AND MAIN RESULTS

Suppose that we are given a discrete-time ergodic Markov chain  $X = (X_n)_{n \in \mathbb{N}}$  on a finite state space  $\mathcal{X}$  with transition matrix  $P = (p_{i,j})_{i,j \in \mathcal{X}}$  that converges to the stationary distribution  $\pi = (\pi_i)_{i \in \mathcal{X}}$ . We are tasked with transporting from a source distribution  $X_0 \sim \mu = (\mu_i)_{i \in \mathcal{X}}$  to a given target distribution  $X_T \sim \nu = (\nu_i)_{i \in \mathcal{X}}$  using the shortest possible time on average, where  $T$  is a stopping time of  $X$ . Note that  $\mu$  and  $\nu$  need not have a common support as long as  $\sum_{i \in \mathcal{X}} \mu_i = \sum_{i \in \mathcal{X}} \nu_i = 1$ . Writing  $\Gamma$  to be the set of stopping time of  $X$  that transports  $\mu$  to  $\nu$ , we say that the stopping time  $T$  is an optimal transport time if it minimizes the mean transport time among  $\Gamma$ , that is,

**Definition 1.1** (optimal transport time).

$$H(\mu, \nu) := \inf_{T \in \Gamma} \mathbb{E}[T].$$

In the literature,  $H(\mu, \nu)$  is known as the access time from  $\mu$  to  $\nu$  of  $X$  [Lovász and Winkler \(1995\)](#). Quite a few well-studied parameters of Markov chains can be formulated in this way. For example, when  $\nu = \pi$ , this is known as the “mixing” time of  $X$  in [Lovász and Winkler \(1998\)](#) (This is not to be confused with the classical total-variation mixing time of  $X$ ). On the other hand, if  $\mu = \delta_i$  and  $\nu = \delta_j$ , the Dirac mass at state  $i$  and  $j$  respectively, then

$$H(i, j) := H(\delta_i, \delta_j) = \mathbb{E}_i[\tau_j]$$

is the mean hitting time of  $X$  from  $i$  to  $j$ , where we denote  $\tau_j := \inf\{n \geq 0; X_n = j\}$  and the usual convention of  $\inf \emptyset = 0$  applies. It is also obvious to see that, for  $j \in \mathcal{X}$ ,

$$H(\mu, j) := H(\mu, \delta_j) = \sum_{i \in \mathcal{X}} \mu_i H(i, j).$$

A closely related problem is the Skorohod embedding problem, where the underlying process is a Brownian motion instead of a finite Markov chain that we consider. We refer interested readers to [Obłój \(2004\)](#) for an excellent survey in this direction, and to [Mörters and Redl \(2016\)](#) in which they study the Skorohod embedding problem with non-randomized stopping time  $T$  in a Markov chain setting. A simple yet useful characterization of  $H(\mu, \nu)$  that this paper frequently exploits is the following representation:

**Theorem 1.1** (Theorem 5.2 of [Lovász and Winkler \(1995\)](#)).

$$H(\mu, \nu) = \max_{j \in \mathcal{X}} \sum_{i \in \mathcal{X}} (\mu_i - \nu_i) \mathbb{E}_i[\tau_j].$$

In view of Theorem 1.1,  $H(\mu, \nu)$  can be interpreted as a generalized notion of mean hitting time. Note that  $H(\mu, \nu)$  is not a metric on the space of probability measures, yet it defines a statistical divergence: it is clear from Definition 1.1 that  $H(\mu, \nu) \geq 0$  and is 0 if and only if  $\mu = \nu$ . It is also trivial from Theorem 1.1 that the triangle inequality  $H(\mu, \nu) \leq H(\mu, \rho) + H(\rho, \nu)$  holds where  $\rho$  is also a distribution on  $\mathcal{X}$ . However, it is in general asymmetric, that is,  $H(\mu, \nu) \neq H(\nu, \mu)$ .

The aim of this paper is to investigate the extend to which  $H(\mu, \nu)$  depends on the structure of the state space as well as to the initial and target distribution  $\mu$  and  $\nu$ . Writing  $n := |\mathcal{X}|$  to be the size of the state space, one important aspect of our investigation is how  $H(\mu, \nu)$  depends on  $n$ . This can shed lights on the time complexity of such transportation. In the following, we summarize our results in a table format:

TABLE 1. Main results

$X$	$H(\mu, \nu)$	Result
symmetric birth-death processes with parameter $p$	$O\left(\frac{n^2}{p}\right)$	Theorem 2.2
winning streak Markov chain	$O(2^n)$	Theorem 2.3
random walk on $d$ -dimensional hypercube	$O(n)$	Corollary 2.1
random walk on connected graphs	$O(n^3)$	Corollary 2.2
random walk on the $n$ -path	$O(n^2)$	Theorem 2.4
random walk on complete graphs	$O(n)$	Theorem 2.5
random walk on the $n$ -star	$O(n)$	Theorem 2.6

Among various Markov chains that we study in this paper, Table 1 seems to suggest that random walk on complete graphs and on the  $n$ -star are more suitable for such transport as the complexity of  $H(\mu, \nu)$  is at most  $O(n)$ . This is reasonable as walks on these graph structures can readily access most parts of the state space relatively quickly. This stands in contrast with walks on linear graphs such as  $n$ -path or symmetric birth-death processes, in which they can have worst-case optimal transport time to be of the order  $O(n^2)$  as the transitions are restricted to nearest-neighbour type. The worst example that we study is the winning streak Markov chains. It exhibits exponential dependence on  $n$  in the worst case.

The rest of the paper is organized as follows. In Section 1.1, we fix commonly used notations throughout the paper. In Section 2, we present our results on  $H(\mu, \nu)$  under different structures on the state space, namely for general finite Markov chains in Section 2.1, symmetric birth-death processes in Section 2.2, the winning streak Markov chain in Section 2.3, random walk on  $d$ -dimensional hypercube in Section 2.4 and random walks on connected graphs in Section 2.5. Afterwards, we proceed our investigation to various special graphs such as random walk on the  $n$ -path (Section 2.6), complete graph (Section 2.7) and  $n$ -star (Section 2.8).

**1.1. Notations.** Throughout the paper, we write  $1_A$  to be the indicator function of the set  $A$ , and for  $x \in \mathbb{R}$  we denote  $x_+ := \max\{x, 0\}$ . For  $a, b \in \mathbb{Z}$ , we write  $\llbracket a, b \rrbracket := \{a, a+1, \dots, b-1, b\}$ . For any two probability measures  $\mu, \nu$  on  $\mathcal{X}$ , we define the total variation distance between  $\mu$  and  $\nu$  to be

$$\|\mu - \nu\|_{TV} := \sup_{A \subset \mathcal{X}} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{j \in \mathcal{X}} |\mu(j) - \nu(j)|.$$

We also denote  $Z \sim \mu$  meaning the law of the random variable is  $\mu$ , and denote the expectation under  $\mu$  to be  $\mathbb{E}_{Z \sim \mu}$ .

## 2. OPTIMAL TRANSPORT ON ...

**2.1. ... finite Markov chains.** In this section, we consider a general finite Markov chain  $X$  and the task is to transport  $\mu$  to  $\nu$  using  $X$ . Our main result in this section below essentially states that we can bound the optimal transport time  $H(\mu, \nu)$  by the maximal mean hitting time. This result is particularly useful as for many Markov chains or random walks on graphs, information for maximal mean hitting time is readily available. We will see how we can apply this bound in subsequent sections which may give tight estimates on the order of  $H(\mu, \nu)$ .

**Theorem 2.1.** *Suppose that  $X$  is a finite Markov chain on  $\mathcal{X}$ . For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$ , we have*

$$(2.1) \quad H(\mu, \nu) \leq \max_{i, j \in \mathcal{X}} \mathbb{E}_i[\tau_j].$$

**Proof.** For a given state  $j \in \mathcal{X}$ , we see that

$$H(\mu, j) - H(\nu, j) \leq H(\mu, j) = \sum_{i \in \mathcal{X}} \mu_i H(i, j) \leq \max_{i, j \in \mathcal{X}} \mathbb{E}_i[\tau_j].$$

Maximizing over  $j$  together with Theorem 1.1 yields the desired result.  $\square$

Interestingly, this upper bound is independent of  $\mu$  or  $\nu$ , and depends on the choice of Markov chain through the maximal mean hitting time.

**2.2. ... symmetric birth-death processes.** In this section, we present results when the underlying Markov chain  $X$  is taken to be a symmetric birth-death process on  $\mathcal{X} = \llbracket 0, n \rrbracket$ . More precisely, the transition dynamics of  $X$  is assumed to be

$$(2.2) \quad p_{i,j} = \begin{cases} (\text{birth probability}) p, & \text{if } j = i+1, i \in \llbracket 0, n-1 \rrbracket, \\ (\text{death probability}) q := p, & \text{if } j = i-1, i \in \llbracket 1, n \rrbracket, \\ (\text{holding probability}) r := 1-2p, & \text{if } j = i, i \in \llbracket 0, n \rrbracket, \end{cases}$$

where  $p \in (0, 1/2]$  and  $p_{i,j} = 0$  otherwise. In other words,  $X$  has the same birth probability and death probability at each state, and this probability is parametrized by  $p$ .

**Theorem 2.2.** *Suppose that  $X$  is a birth-death process with the above transition dynamics (2.2). For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$  using  $X$ , we have*

$$(2.3) \quad H(\mu, \nu) = \frac{1}{2p} \left( \mathbb{E}_{Z \sim \nu}[Z] - \mathbb{E}_{Z \sim \mu}[Z] + \max_{j \in \llbracket 0, n \rrbracket} \left( \mathbb{E}_{Z \sim \mu} [\max\{Z, j\}^2 - \min\{Z, j\}^2] \right. \right. \\ \left. \left. - \mathbb{E}_{Z \sim \nu} [\max\{Z, j\}^2 - \min\{Z, j\}^2] \right) \right).$$

In particular,

$$(2.4) \quad \frac{1}{2p} (\mathbb{E}_{Z \sim \nu}[Z] - \mathbb{E}_{Z \sim \mu}[Z] + \mathbb{E}_{Z \sim \mu}[Z^2] - \mathbb{E}_{Z \sim \nu}[Z^2])_+ \leq H(\mu, \nu) \leq \frac{2n^2 + n}{2p} = O\left(\frac{n^2}{p}\right).$$

In Theorem 2.2 (2.3), we express the optimal transport time  $H(\mu, \nu)$  in terms of the moments of  $\mu$  and  $\nu$ . This formulation can perhaps lead to explicit results as we are tackling with moments of  $\mu$  and  $\nu$  instead of the mean hitting time of birth-death processes. In both (2.3) and (2.4), we observe that the optimal transport time  $H(\mu, \nu)$  is inversely proportional to the parameter  $p$ . This is reasonable as  $p$  gets smaller (resp. bigger), the birth-death chain moves slower (resp. faster) on the state space, making it harder (resp. easier) to transport  $\mu$  to  $\nu$ . Before we proceed to the proof of Theorem 2.2, we first state a lemma that yields a closed-form expression for the mean hitting time of birth-death processes. As a side remark, we note that closed-form formulae of mean hitting time of birth-death processes has been investigated in Palacios and Tetali (1996); Tetali (1991) using the electric network approach.

**Lemma 2.1.** *Suppose that  $X$  is a birth-death process with transition dynamics given by (2.2). Then the mean hitting times can be written as, for  $i, j \in \llbracket 0, n \rrbracket$ ,*

$$\mathbb{E}_i[\tau_j] = \begin{cases} \frac{(i+j+1)(j-i)}{2p}, & i < j, \\ \frac{(i+j-1)(i-j)}{2p}, & i > j. \end{cases}$$

**Proof.** We first consider the case when  $i < j$ . Using the strong Markov property and the birth-death property, we see that

$$\mathbb{E}_i[\tau_j] = \sum_{k=i}^{j-1} \mathbb{E}_k[\tau_{k+1}] = \sum_{k=i}^{j-1} \frac{k+1}{p} = \frac{(i+j+1)(j-i)}{2p},$$

where we use the equality  $\mathbb{E}_k[\tau_{k+1}] = (k+1)/p$  as in (Levin et al., 2009, Section 2.5). Next, we treat the case  $i > j$ . Using again the equality  $\mathbb{E}_{k+1}[\tau_k] = k/p$  as well as strong Markov property and birth-death property, we have

$$\mathbb{E}_i[\tau_j] = \sum_{k=j}^{i-1} \mathbb{E}_{k+1}[\tau_k] = \sum_{k=j}^{i-1} \frac{k}{p} = \frac{(i+j-1)(i-j)}{2p}.$$

□

With the above Lemma 2.1, we can prove the desired result of Theorem 2.2.

**Proof of Theorem 2.2.** Using Theorem 1.1, we have

$$\begin{aligned} H(\mu, \nu) &= \max_{j \in \llbracket 0, n \rrbracket} \sum_{i < j} (\mu_i - \nu_i) \frac{(i+j+1)(j-i)}{2p} + \sum_{i > j} (\mu_i - \nu_i) \frac{(i+j-1)(i-j)}{2p} \\ (2.5) \quad &= \frac{1}{2p} \max_{j \in \llbracket 0, n \rrbracket} \sum_{i < j} (\mu_i - \nu_i)(j^2 - i^2) + \sum_{i > j} (\mu_i - \nu_i)(i^2 - j^2) + j \sum_{i \neq j} (\mu_i - \nu_i) - \sum_{i \neq j} i(\mu_i - \nu_i). \end{aligned}$$

The third and the fourth term can be simplified into

$$(2.6) \quad j \sum_{i \neq j} (\mu_i - \nu_i) - \sum_{i \neq j} i(\mu_i - \nu_i) = j(\nu_j - \mu_j) + j\mu_j - \mathbb{E}_{Z \sim \mu}[Z] - j\nu_j + \mathbb{E}_{Z \sim \nu}[Z] = \mathbb{E}_{Z \sim \nu}[Z] - \mathbb{E}_{Z \sim \mu}[Z],$$

while the first and the second term reduces to

$$(2.7) \quad \begin{aligned} \sum_{i < j} (\mu_i - \nu_i)(j^2 - i^2) + \sum_{i > j} (\mu_i - \nu_i)(i^2 - j^2) &= j^2 (\mathbb{E}_{Z \sim \mu}[\mathbf{1}_{\{Z < j\}} - \mathbf{1}_{\{Z > j\}}] - \mathbb{E}_{Z \sim \nu}[\mathbf{1}_{\{Z < j\}} - \mathbf{1}_{\{Z > j\}}]) \\ &\quad - \mathbb{E}_{Z \sim \mu}[Z^2 \mathbf{1}_{\{Z < j\}} - Z^2 \mathbf{1}_{\{Z > j\}}] + \mathbb{E}_{Z \sim \nu}[Z^2 \mathbf{1}_{\{Z < j\}} - Z^2 \mathbf{1}_{\{Z > j\}}] \\ &= \mathbb{E}_{Z \sim \mu}[\max\{Z, j\}^2 - \min\{Z, j\}^2] \\ &\quad - \mathbb{E}_{Z \sim \nu}[\max\{Z, j\}^2 - \min\{Z, j\}^2] \end{aligned}$$

(2.3) follows by plugging (2.6) and (2.7) back into (2.5). For the upper bound of (2.4), using (2.3) and the fact that  $Z$  is supported on  $\llbracket 0, n \rrbracket$ , we have

$$H(\mu, \nu) \leq \frac{1}{2p} \left( \mathbb{E}_{Z \sim \nu}[Z] + \max_{j \in \llbracket 0, n \rrbracket} \left( \mathbb{E}_{Z \sim \mu}[\max\{Z, j\}^2] + \mathbb{E}_{Z \sim \nu}[\min\{Z, j\}^2] \right) \right) \leq \frac{2n^2 + n}{2p}.$$

As for the lower bound of (2.4), we observe that

$$\max_{j \in \llbracket 0, n \rrbracket} \left( \mathbb{E}_{Z \sim \mu}[\max\{Z, j\}^2 - \min\{Z, j\}^2] - \mathbb{E}_{Z \sim \nu}[\max\{Z, j\}^2 - \min\{Z, j\}^2] \right) \geq \mathbb{E}_{Z \sim \mu}[Z^2] - \mathbb{E}_{Z \sim \nu}[Z^2].$$

□

**2.3. ... winning streak Markov chain.** In this section, we take the so-called winning streak Markov chain  $X$  living on the state space  $\mathcal{X} = \llbracket 1, n \rrbracket$ . The transition dynamics  $P = (p_{i,j})_{i,j \in \mathcal{X}}$  is governed by

$$(2.8) \quad p_{i,j} = \begin{cases} 1/2, & \text{if } j = i + 1, i \in \llbracket 1, n - 1 \rrbracket, \\ 1/2, & \text{if } j = 1, i \in \llbracket 1, n \rrbracket, \\ 1/2, & \text{if } j = i = n, \end{cases}$$

and 0 otherwise. The expected hitting time of such chain has been calculated in (Chen and Zhang, 2008, Example 2), which is given by

$$\mathbb{E}_i[\tau_j] = \begin{cases} 2^j - 2^i, & i < j, \\ 2^j, & i > j. \end{cases}$$

The winning streak Markov chain is also studied in Levin et al. (2009). In our main result below, we formulate the optimal transport time  $H(\mu, \nu)$  in terms of the probability generating function (pgf) of  $\mu$  and  $\nu$ , which subsequently leads to a lower bound (2.10) expressed in the differences between the pgf of these two. However, this lower bound may not be tight as the value can be zero.

**Theorem 2.3.** *Suppose that  $X$  is a winning streak Markov chain with the above transition dynamics (2.8). For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$  using  $X$ , we have*

$$(2.9) \quad H(\mu, \nu) = \max_{j \in \llbracket 1, n \rrbracket} \mathbb{E}_{Z \sim \nu}[2^Z \mathbf{1}_{\{Z \leq j\}}] - \mathbb{E}_{Z \sim \mu}[2^Z \mathbf{1}_{\{Z \leq j\}}].$$

In particular,

$$(2.10) \quad (\mathbb{E}_{Z \sim \nu} [2^Z] - \mathbb{E}_{Z \sim \mu} [2^Z])_+ \leq H(\mu, \nu) \leq 2^n = O(2^n).$$

**Proof.** Using again Theorem 1.1, we have

$$\begin{aligned} H(\mu, \nu) &= \max_{j \in \llbracket 1, n \rrbracket} \sum_{i < j} (\mu_i - \nu_i)(2^j - 2^i) + \sum_{i > j} (\mu_i - \nu_i)2^j \\ &= \max_{j \in \llbracket 1, n \rrbracket} 2^j \sum_{i \neq j} (\mu_i - \nu_i) - \sum_{i < j} 2^i (\mu_i - \nu_i) \\ &= \max_{j \in \llbracket 1, n \rrbracket} 2^j (\nu_j - \mu_j) + \mathbb{E}_{Z \sim \nu} [2^Z \mathbf{1}_{\{Z < j\}}] - \mathbb{E}_{Z \sim \mu} [2^Z \mathbf{1}_{\{Z < j\}}] \\ &= \max_{j \in \llbracket 1, n \rrbracket} \mathbb{E}_{Z \sim \nu} [2^Z \mathbf{1}_{\{Z \leq j\}}] - \mathbb{E}_{Z \sim \mu} [2^Z \mathbf{1}_{\{Z \leq j\}}]. \end{aligned}$$

To see the upper bound in (2.10), we simply use Theorem 2.1 and  $\max_{i,j \in \llbracket 1, n \rrbracket} \mathbb{E}_i[\tau_j] = \mathbb{E}_1[\tau_n] = 2^n$ . As for the lower bound in (2.10), we note that

$$\max_{j \in \llbracket 1, n \rrbracket} \mathbb{E}_{Z \sim \nu} [2^Z \mathbf{1}_{\{Z \leq j\}}] - \mathbb{E}_{Z \sim \mu} [2^Z \mathbf{1}_{\{Z \leq j\}}] \geq \mathbb{E}_{Z \sim \nu} [2^Z] - \mathbb{E}_{Z \sim \mu} [2^Z].$$

□

To see that the upper bound of (2.10) is of the correct order in  $n$ , we now look at an example in which we take  $n > 1$ ,  $\mu = \delta_1$ , the Dirac mass at 1, and  $\nu_i = \frac{1}{2(n-1)}$  for  $i \in \llbracket 1, n-1 \rrbracket$ ,  $\nu_n = 1/2$ . This leads to

$$\max_{j \in \llbracket 1, n \rrbracket} \mathbb{E}_{Z \sim \nu} [2^Z \mathbf{1}_{\{Z \leq j\}}] = \mathbb{E}_{Z \sim \nu} [2^Z] = \frac{2^{n-1} - 1}{n-1} + 2^{n-1}, \quad \mathbb{E}_{Z \sim \delta_1} [2^Z \mathbf{1}_{\{Z \leq j\}}] = 2.$$

As a result, it follows from (2.9) that

$$H(\mu, \nu) = \max_{j \in \llbracket 1, n \rrbracket} \mathbb{E}_{Z \sim \nu} [2^Z \mathbf{1}_{\{Z \leq j\}}] - \mathbb{E}_{Z \sim \mu} [2^Z \mathbf{1}_{\{Z \leq j\}}] = \frac{2^{n-1} - 1}{n-1} + 2^{n-1} - 2 = O(2^n).$$

As another remark, if we are allowed to choose a Markov chain  $X$  to transport  $\mu$  to  $\nu$ , it seems that the winning streak Markov chain is a poor choice as the worst case optimal transport time can grow exponentially in  $n$ , the size of the state space.

**2.4. ... random walk on  $d$ -dimensional hypercube.** In this section, we study a multivariate example, namely a simple random walk on  $d$ -dimensional hypercube. The state space  $\mathcal{X}$  is the set of binary  $d$ -tuples  $\{0, 1\}^d$ , and the size of the state space is  $n := |\mathcal{X}| = 2^d$ . At each time step, one of the  $d$  coordinates, say  $x_j$ , is chosen uniformly at random and changed to  $1 - x_j$ . For optimal transport problem on hypercube, our result below shows that the optimal transport time is linear in the size  $n$  of the state space (or exponential in the dimension  $d$ ).

**Corollary 2.1.** *Suppose that  $X$  is a simple random walk on  $d$ -dimensional hypercube. For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$  using  $X$ , we have*

$$(2.11) \quad H(\mu, \nu) \leq (1 + o(1))n,$$

where  $n = 2^d$ .

**Proof.** According to (Palacios, 1994, Section 5.2), the maximal mean hitting time on  $d$ -dimensional hypercube is

$$\max_{i,j \in \mathcal{X}} \mathbb{E}_i[\tau_j] = \mathbb{E}_{(0,\dots,0)}[\tau_{(1,\dots,1)}] = (1 + o(1))n.$$

Desired result follows from Theorem 2.1.  $\square$

**2.5. ... random walks on connected graphs.** In this section and subsequent sections, we study random walk on connected graph  $G = (V, E)$ , where  $V$  is the vertex set and  $E$  is the edge set. We denote  $d_i$  to be the degree of  $i \in V$ . For random walk on such a graph, the transition dynamics on  $\mathcal{X} = V$  is

$$(2.12) \quad p_{i,j} = \begin{cases} \frac{1}{d_i}, & \text{if } (i,j) \in E, \\ 0, & \text{otherwise.} \end{cases}$$

In other words, at each vertex  $i$  the random walk pick a vertex  $j$  that are connected to  $i$  with an edge uniformly at random. For an overview of various results of random walk on graph, we refer readers to the excellent survey of Lovász (1996). In our main result of this section, we prove that  $H(\mu, \nu)$  is of  $O(n^3)$ . However, as we will see in subsequent sections, for a variety of special graphs this bound is quite loose.

**Corollary 2.2.** *Suppose that  $X$  is a random walk on connected graph  $G = (V, E)$ . For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$  using  $X$ , we have*

$$H(\mu, \nu) \leq n(n-1)^2 = O(n^3),$$

where  $n := |V|$ .

**Proof.** According to (Lawler, 1986, Theorem 2), the maximal mean hitting time on a connected graph is upper bounded by

$$\max_{i,j \in \mathcal{X}} \mathbb{E}_i[\tau_j] \leq n(n-1)^2.$$

Desired result follows from Theorem 2.1.  $\square$

The stationary distribution  $\pi = (\pi_i)$  of  $X$  is given by, for  $i \in V$ ,

$$\pi_i = \frac{d_i}{2|E|}.$$

In our next result of this section, we specialize into the case of random walk on graphs with symmetric hitting times. The cover time of such a random walk has been studied in Devroye and Sbihi (1990); Palacios (1992). Notable families of graphs that belong to this class are complete graphs, vertex-transitive graphs and distance-regular graphs. For random walk on graphs with symmetric hitting time, we can express the optimal transport time in terms of average hitting time as follows.

**Proposition 2.1** (Random walk with symmetric hitting times). *Suppose that  $X$  is a random walk on connected graph  $G = (V, E)$  with stationary distribution  $\pi$  and symmetric hitting time, i.e. for all  $i, j \in V$ ,  $\mathbb{E}_i[\tau_j] = \mathbb{E}_j[\tau_i]$ . For the optimal transport time  $H(\pi, \nu)$  and  $H(\mu, \pi)$ , we have*

$$(2.13) \quad H(\pi, \nu) = t_{av} - \min_j \sum_i \nu_i \mathbb{E}_i[\tau_j] \leq t_{av},$$



$$(2.14) \quad H(\mu, \pi) = \max_j \sum_i \mu_i \mathbb{E}_i[\tau_j] - t_{av} \leq \max_{i,j} \mathbb{E}_i[\tau_j] - t_{av},$$

where

$$n := |V|, \quad t_{av} := \sum_{i,j} \pi_i \pi_j \mathbb{E}_i[\tau_j] = \sum_{i=2}^n \frac{1}{1 - \lambda_i},$$

and  $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_n$  are the eigenvalues of the transition matrix  $P$  of the random walk.

**Proof.** We only prove (2.13), as (2.14) can be proved similarly. By Theorem 1.1, we see that

$$H(\pi, \nu) = \max_j \sum_i (\pi_i - \nu_i) \mathbb{E}_i[\tau_j] = \max_j \sum_i \pi_i \mathbb{E}_j[\tau_i] - \nu_i \mathbb{E}_i[\tau_j] = t_{av} - \min_j \sum_i \nu_i \mathbb{E}_i[\tau_j] \leq t_{av},$$

where we use the symmetric hitting time property in the second equality, and the random target lemma (see e.g. (Levin et al., 2009, Lemma 10.1)) in the third equality.  $\square$

Comparing (2.14) with the general result of Theorem 2.1, we see that  $t_{av}$  serves as a correction term. In addition, the results in Proposition 2.1 demonstrate a connection with the eigenvalues of the underlying random walk.

**2.6. ... random walk on the  $n$ -path.** In this section, we specialize into the case of random walk on the  $n$ -path. This is one among many classical examples of random walk on graphs that have been studied intensively, see for example Aldous and Fill (2002) or Lovász (1996). It can be regarded as a birth-death process on  $\llbracket 0, n \rrbracket$  with reflecting boundaries at 0 and  $n$ , where its transition dynamic is given by

$$(2.15) \quad p_{i,j} = \begin{cases} 1/2, & \text{if } j = i \pm 1, i \in \llbracket 1, n-1 \rrbracket, \\ 1, & \text{if } i = 0, j = 1 \text{ or } i = n, j = n-1, \\ 0, & \text{otherwise.} \end{cases}$$

**Theorem 2.4.** Suppose that  $X$  is a random walk on the  $n$ -path with transition given by (2.15). For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$  using  $X$ , we have

$$(2.16) \quad H(\mu, \nu) = \mathbb{E}_{Z \sim \nu}[Z^2] - \mathbb{E}_{Z \sim \mu}[Z^2] + 2n \max_{j \in \llbracket 0, n \rrbracket} (\mathbb{E}_{Z \sim \mu}(Z - j)_+ - \mathbb{E}_{Z \sim \nu}(Z - j)_+).$$

In particular,

$$(2.17) \quad (\mathbb{E}_{Z \sim \nu}[Z^2] - \mathbb{E}_{Z \sim \mu}[Z^2] + 2n (\mathbb{E}_{Z \sim \mu}[Z] - \mathbb{E}_{Z \sim \nu}[Z]))_+ \leq H(\mu, \nu) \leq n^2 = O(n^2).$$

As our first remark, we note that this upper bound (2.17) of the  $n$ -path is tighter than the general  $O(n^3)$  bound of Corollary 2.2. When  $\mu = \delta_0$  and  $\nu = \delta_n$ , the upper bound in (2.17) is trivially attained as  $H(\delta_0, \delta_n) = H(0, n) = \mathbb{E}_0[\tau_n] = n^2$ . To demonstrate a non-trivial instance in which both the upper and lower bounds in (2.17) are of the same order in  $n$ , i.e.  $H(\mu, \nu) = \Theta(n^2)$ , we consider the case when  $\mu$  is a discrete uniform on  $\llbracket 0, n \rrbracket$  while  $\nu$  is binomially distributed with parameters  $n$  and  $p \in (0, 1 - 1/\sqrt{3})$ . In this case, we have

$$\begin{aligned} \mathbb{E}_{Z \sim \mu}[Z] &= \frac{n}{2}, & \mathbb{E}_{Z \sim \mu}[Z^2] &= \frac{n(2n+1)}{6}, \\ \mathbb{E}_{Z \sim \nu}[Z] &= np, & \mathbb{E}_{Z \sim \nu}[Z^2] &= np(1-p) + n^2 p^2. \end{aligned}$$



As a result, the lower bound in (2.17) now reads

$$\mathbb{E}_{Z \sim \nu}[Z^2] - \mathbb{E}_{Z \sim \mu}[Z^2] + 2n(\mathbb{E}_{Z \sim \mu}[Z] - \mathbb{E}_{Z \sim \nu}[Z]) = n^2 \left( p^2 - 2p + \frac{2}{3} \right) + n \left( p(1-p) - \frac{1}{6} \right).$$

For  $p \in (0, 1 - 1/\sqrt{3})$ ,  $p^2 - 2p + 2/3 > 0$  and so both the lower and upper bound of (2.17) gives an order of  $n^2$ . Before we state the proof of Theorem 2.4, we first give a lemma that writes out the explicit expression for the mean hitting time of the random walk on the  $n$ -path.

**Lemma 2.2.** *Suppose that  $X$  is a random walk on the  $n$ -path with transition given by (2.15). Then the mean hitting times can be written as, for  $i, j \in \llbracket 0, n \rrbracket$ ,*

$$\mathbb{E}_i[\tau_j] = \begin{cases} j^2 - i^2, & i < j, \\ (i-j)2n - (i^2 - j^2), & i > j. \end{cases}$$

**Proof.** For  $i < j$ , the result can be found in (Aldous and Fill, 2002, Chapter 5 Example 8). For the case of  $i > j$ , we adapt a similar proof and proceed by induction on  $i$ . When  $i = j + 1$ , we observe that  $\mathbb{E}_{j+1}[\tau_j]$  is one less than the expected return time from  $j$  to  $j$  in a graph of  $n - j + 1$  nodes starting on the left, which is  $2(n - j) - 1$ . Suppose that  $\mathbb{E}_i[\tau_j] = (i - j)2n - (i^2 - j^2)$  holds for some  $i$ , then

$$\mathbb{E}_{i+1}[\tau_j] = \mathbb{E}_{i+1}[\tau_i] + \mathbb{E}_i[\tau_j] = 2(n - i) - 1 + (i - j)2n - (i^2 - j^2) = (i + 1 - j)2n - ((i + 1)^2 - j^2).$$

□

**Proof of Theorem 2.4.** Using Theorem 1.1 together with Lemma 2.2, we have

$$\begin{aligned} H(\mu, \nu) &= \max_{j \in \llbracket 0, n \rrbracket} \sum_{i < j} (\mu_i - \nu_i)(j^2 - i^2) + \sum_{i > j} ((\mu_i - \nu_i)(i - j)2n + (j^2 - i^2)) \\ &= \max_{j \in \llbracket 0, n \rrbracket} \sum_i (\mu_i - \nu_i)(j^2 - i^2) + 2n(\mathbb{E}_{Z \sim \mu}(Z - j)_+ - \mathbb{E}_{Z \sim \nu}(Z - j)_+) \\ &= \mathbb{E}_{Z \sim \nu}[Z^2] - \mathbb{E}_{Z \sim \mu}[Z^2] + 2n \max_{j \in \llbracket 0, n \rrbracket} (\mathbb{E}_{Z \sim \mu}(Z - j)_+ - \mathbb{E}_{Z \sim \nu}(Z - j)_+). \end{aligned}$$

For the upper bound of (2.17), we apply Theorem 2.1 and Lemma 2.2 to see that

$$\max_{i, j \in \llbracket 0, n \rrbracket} \mathbb{E}_i[\tau_j] = \mathbb{E}_0[\tau_n] = n^2.$$

As for the lower bound of (2.17), we use (2.16) and note that

$$2n \max_{j \in \llbracket 0, n \rrbracket} (\mathbb{E}_{Z \sim \mu}(Z - j)_+ - \mathbb{E}_{Z \sim \nu}(Z - j)_+) \geq 2n(\mathbb{E}_{Z \sim \mu}[Z] - \mathbb{E}_{Z \sim \nu}[Z]).$$

□

**2.7. ... random walk on complete graphs.** In this section, we look into the example of random walk on complete graphs on  $\mathcal{X} = \llbracket 0, n \rrbracket$ , with transition dynamics driven by, for  $i \neq j \in \mathcal{X}$ ,  $p_{i,i} = 0$  and  $p_{i,j} = 1/n$ , and the stationary distribution is well-known to be a discrete uniform  $\pi_i = 1/(n + 1)$ . Alternatively, we can take the degree of node  $i$  as  $d_i = n$  in (2.12). According to (Aldous and Fill, 2002, Chapter 5 Example 9), the mean hitting times of this random walk are

$$\mathbb{E}_i[\tau_j] = n.$$

As our main result below illustrates, the optimal transport time is of the order  $O(n)$ .

**Theorem 2.5.** *Suppose that  $X$  is a random walk on the complete graph. For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$  using  $X$ , we have*

$$(2.18) \quad H(\mu, \nu) = n \max_{j \in \llbracket 0, n \rrbracket} (\nu_j - \mu_j).$$

*In particular, for  $j \in \llbracket 0, n \rrbracket$ ,*

$$(2.19) \quad n(\nu_j - \mu_j)_+ \leq H(\mu, \nu) \leq n \|\mu - \nu\|_{TV} \leq n = O(n).$$

*The distribution within the class of Dirac mass  $\mathcal{P} := \{\delta_i; i \in \mathcal{X}\}$  that minimizes  $H(\mu, \nu)$  is  $\mu^* := \delta_{i^*}$  where  $i^* := \arg \max \nu_i$ , i.e.*

$$(2.20) \quad \min_{\mu \in \mathcal{P}} H(\mu, \nu) = H(\mu^*, \nu).$$

We first note that the  $O(n)$  bound of  $H(\mu, \nu)$  is much tighter than the  $O(n^3)$  bound in the general result Corollary 2.2. In addition, (2.20) has an operational meaning when we are given the freedom to choose within  $\mathcal{P}$  such that its optimal transport time  $H(\mu, \nu)$  is the fastest: we should pick the state  $i$  which maximizes  $\nu_i$ . This also makes sense intuitively as we initialize the chain at a high probability state of  $\nu$ , thus the transport time should be faster than other initializations.

**Proof of Theorem 2.5.** Thanks to Theorem 1.1, we have

$$H(\mu, \nu) = \max_{j \in \llbracket 0, n \rrbracket} \sum_{i \neq j} (\mu_i - \nu_i) \mathbb{E}_i[\tau_j] = n \max_{j \in \llbracket 0, n \rrbracket} (\nu_j - \mu_j).$$

For the upper bound in (2.19), we note that

$$H(\mu, \nu) \leq n \max_{j \in \llbracket 0, n \rrbracket} |\nu_j - \mu_j| \leq n \|\mu - \nu\|_{TV} \leq n.$$

As for the lower bound in (2.19), we simply use

$$n \max_{j \in \llbracket 0, n \rrbracket} (\nu_j - \mu_j) \geq n(\nu_j - \mu_j).$$

Finally, we prove (2.20). For  $\mu \in \mathcal{P}$  and  $\mu \neq \mu^*$ , we see that

$$H(\mu, \nu) = n \max_{j \in \llbracket 0, n \rrbracket} \nu_j \geq n \max_{j \in \llbracket 0, n \rrbracket, j \neq i^*} \nu_j = H(\mu^*, \nu).$$

□

**2.8. ... random walk on the  $n$ -star.** In this section, we consider the  $n$ -star graph on  $\mathcal{X} = \llbracket 0, n \rrbracket$ , where there is an edge connecting  $0 - 1, 0 - 2, \dots, 0 - n$  and 0 is the center of the star.

**Theorem 2.6.** *Suppose that  $X$  is a random walk on the  $n$ -star. For the optimal transport time  $H(\mu, \nu)$  from  $\mu$  to  $\nu$  using  $X$ , we have*

$$(2.21) \quad H(\mu, \nu) = \nu_0 - \mu_0 + 2n \left( \max_{j \in \llbracket 1, n \rrbracket} (\nu_j - \mu_j) \right)_+.$$

*In particular, for  $j \in \llbracket 1, n \rrbracket$ ,*

$$(2.22) \quad \nu_0 - \mu_0 + 2n(\nu_j - \mu_j)_+ \leq H(\mu, \nu) \leq (2n + 1) \|\mu - \nu\|_{TV} \leq 2n + 1 = O(n).$$

**Proof.** As in previous sections, the mean hitting times are vital for us in determining  $H(\mu, \nu)$ . According to (Aldous and Fill, 2002, Chapter 5 Example 10), the mean hitting times of the random walk on  $n$ -star are, for  $i \in \llbracket 1, n \rrbracket$ ,

$$\mathbb{E}_i[\tau_j] = \begin{cases} 1, & j = 0, \\ 2n, & j \neq i, 0. \end{cases}$$

As for  $i = 0$  and  $j \in \llbracket 1, n \rrbracket$ , we have

$$\mathbb{E}_0[\tau_j] = 2n - 1.$$

This leads to, for  $j \in \llbracket 1, n \rrbracket$ ,

$$\begin{aligned} \sum_i (\mu_i - \nu_i) \mathbb{E}_i[\tau_0] &= \nu_0 - \mu_0, \\ \sum_i (\mu_i - \nu_i) \mathbb{E}_i[\tau_j] &= (2n - 1)(\mu_0 - \nu_0) + 2n(\nu_0 + \nu_j - \mu_0 - \mu_j) = \nu_0 - \mu_0 + 2n(\nu_j - \mu_j). \end{aligned}$$

As a result, according to Theorem 1.1 we have

$$H(\mu, \nu) = \nu_0 - \mu_0 + 2n \left( \max_{j \in \llbracket 1, n \rrbracket} (\nu_j - \mu_j) \right)_+,$$

and (2.22) follows from

$$(\nu_j - \mu_j)_+ \leq \max_{j \in \llbracket 0, n \rrbracket} (\nu_j - \mu_j) \leq \|\mu - \nu\|_{TV}.$$

□

## REFERENCES

- D. Aldous and J. A. Fill. Reversible Markov Chains and Random Walks on Graphs, 2002. Unfinished monograph, recompiled 2014, available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- H. Chen and F. Zhang. The expected hitting times for finite Markov chains. *Linear Algebra Appl.*, 428(11-12):2730–2749, 2008.
- L. Devroye and A. Sbihi. Random walks on highly symmetric graphs. *J. Theoret. Probab.*, 3(4):497–514, 1990.
- G. F. Lawler. Expected hitting times for a random walk on a connected graph. *Discrete Math.*, 61(1):85–92, 1986.
- D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009.
- L. Lovász. Random walks on graphs: a survey. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 353–397. János Bolyai Math. Soc., Budapest, 1996.
- L. Lovász and P. Winkler. Efficient stopping rules for markov chains. In *Proceedings of the Twenty-seventh Annual ACM Symposium on Theory of Computing*, STOC '95, pages 76–82, New York, NY, USA, 1995. ACM.
- L. Lovász and P. Winkler. Reversal of Markov chains and the forget time. *Combin. Probab. Comput.*, 7(2):189–204, 1998.
- P. Mörters and I. Redl. Skorokhod embeddings for two-sided Markov chains. *Probab. Theory Related Fields*, 165(1-2):483–508, 2016.

- J. Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390, 2004.
- J. L. Palacios. Expected cover times of random walks on symmetric graphs. *J. Theoret. Probab.*, 5(3):597–600, 1992.
- J. L. Palacios. Expected hitting and cover times of random walks on some special graphs. In *Proceedings of the Fifth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science (Poznań, 1991)*, volume 5, pages 173–182, 1994.
- J. L. Palacios and P. Tetali. A note on expected hitting times for birth and death chains. *Statist. Probab. Lett.*, 30(2):119–125, 1996.
- P. Tetali. Random walks and the effective resistance of networks. *J. Theoret. Probab.*, 4(1): 101–109, 1991.

INSTITUTE FOR DATA AND DECISION ANALYTICS, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN,  
GUANGDONG, 518172, P.R. CHINA  
*Email address:* michaelchoi@cuhk.edu.cn