SIAM J. MATRIX ANAL. APPL.
Vol. 37, No. 3, pp. 933–954

THE HITTING TIME OF MULTIPLE RANDOM WALKS*

RUSHABH PATEL†, ANDREA CARRON‡, AND FRANCESCO BULLO§

Abstract. This work provides generalized notions and analysis methods for the *hitting time* of random walks on graphs. The hitting time, also known as the Kemeny constant or the mean first passage time, of a random walk is widely studied; however, only limited work is available for the multiple random walker scenario. In this work we provide a novel method for calculating the hitting time for a single random walker as well as the first analytic expression for calculating the hitting time for multiple random walkers, which we denote as the *group hitting time*. We also provide a closed form solution for calculating the hitting time between specified nodes for both the single and multiple random walker cases. Our results allow for the multiple random walks to be different and, moreover, for the random walks to operate on different subgraphs. Finally, using sequential quadratic programming, we show that the combination of transition matrices that generate the minimal group hitting time for various graph topologies is often different.

Key words. hitting time, Markov chain, surveillance, robotics, Kemeny constant

AMS subject classifications. 60J10, 60J20, 60J22

DOI. 10.1137/15M1010737

1. Introduction.

Random walks appear in many active research areas. They are used to describe electrical networks [10, 39], how information propagates in social networks [15], vehicle surveillance strategies [37], and network searches [36], among others. Our initial motivation comes from one application in particular: surveillance strategies. Vehicle surveillance and routing strategies are used in various fields in both single-agent and multiagent cases. In the context of robotics, single- and multiagent surveillance/routing strategies appear in environmental monitoring [12, 31], minimizing emergency vehicle response times [4], and traffic routing and border patrol [30, 38].

In this paper, we focus on multiagent stochastic surveillance strategies. More specifically, we look at the *hitting time* of a random walk for multiple random walkers; each random walk is governed by its own Markov chain. This problem is not only of interest in the context of robotic surveillance, but is also of general interest in various other fields. Some direct applications include, but are not limited to, determining how quickly epidemics spread [41], how information propagates in a social network [3], and how quickly information packets get transferred in a wireless node network [34].

The *hitting time* of a random walk governed by a Markov chain is the expected time taken by a random walker to travel between *any* two nodes in a network. For a single finite discrete-time Markov chain, this quantity is also well known as the *Kemeny constant* of the Markov chain, the *mean first passage time* of a Markov chain, the *first hitting time* of a Markov chain or, for the case of reversible Markov chains, the *eigentime* of a Markov chain. We refer to this quantity as the *first hitting time* or simply *hitting time* of a Markov chain, both due to the descriptive nature of

*Received by the editors March 2, 2015; accepted for publication (in revised form) by B. Meini May 6, 2016; published electronically July 21, 2016. This work was supported in part by AFOSR FA9550-15-1-0138.

http://www.siam.org/journals/simax/37-3/M101073.html

†Mechanical Engineering, University of California, Santa Barbara, CA 93106-5070 (r_patel@engineering.ucsb.edu).

‡Information Engineering, University of Padova, Padova 35131, Italy (carronan@dei.unipd.it).

§Center for Control, Dynamical Systems and Computation, University of California, Santa Barbara, CA 93106-5070 (bullo@engineering.ucsb.edu).

this coinage as well as its prevalence in the literature. The hitting time of a finite irreducible Markov chain first appeared in [18], but it was rediscovered for finite *reversible* Markov chains in [5]. Since its original discovery, the hitting time has been further developed by several groups [17, 19, 32]. The authors of [17, 23] give bounds on the hitting time for various graph topologies, and in [7] an alternate formulation is explored. Recently, the authors of [32] extended the notion of the hitting time to networks with travel distances and provided a scheme for its optimization, and the authors of [1] provided the first formulation of the hitting time for continuous time reversible Markov chains.

The hitting time is closely related to several other well-studied Markov chain properties. We focus on three quantities that are of particular interest; however, several others exist. The first and most closely related quantity is the *pairwise* hitting time between two nodes, which is the expected time to travel between a *specified* pair of nodes. Clearly, the hitting time is simply the expectation of all possible pairwise hitting times of a Markov chain, where the underlying probability measure is determined by the Markov chain's stationary distribution. Using the relation between reversible Markov chains and electrical networks [10], the authors of [39] gave analytic expressions for hitting times in terms of effective resistance. Using the electrical framework, closed form expressions for pairwise hitting times have also been given for special cases of certain Markov chains [29, 28]. Second, the *cover time* of a graph is the expected time it takes to reach every node in the graph at least once. This quantity is sometimes interpreted as a function of a single node or, more generally, in the context of an arbitrary set of nodes. There are several works relating the cover time to the hitting time of a transition matrix [27]; many of these works bound the cover time in terms of pairwise hitting time (most often the worst-case pairwise hitting time). Finally, the *mixing rate* of an irreducible Markov chain is the rate at which an arbitrary distribution converges to the chain's stationary distribution. A good relationship between Markov chain mixing and hitting times can be found in [24].

In this work, we look at the hitting time of multiple random walkers. More specifically, we analyze the expected first time to reach any single node in a network given that there are an arbitrary number of random walkers in that network. Recently, the authors of [11] looked at bounds on hitting times and cover times for multiple random walkers. In [2] alternate bounds on cover time were formulated for reversible Markov chains, and tight bounds were explored in [13, 14]. The authors of [9, 8] found solutions for cover times in the limit as the number of nodes in the graph becomes infinite. However, as far as we can discern, a key assumption made by all work presented thus far in the literature is that results are based on k copies of a *simple* random walk over a single graph. We make no such assumptions in our work; every random walker can move according to a different and arbitrary random walk and need not share the same underlying graph topology. Also, as opposed to prior work, our results are not bounds but exact analytic expressions.

To achieve our results, we utilize the notion of Kronecker graphs. Preliminary results for undirected Kronecker graphs were introduced in [40], showing conditions under which the Kronecker product of two graphs generates a connected graph. In [21, 22] the authors consider and analyze special Kronecker graphs that are created by products of the same $n \times n$ edge matrix in order to model large networks and also introduce the concept of "stochastic" Kronecker graphs to generate large networks. This method of generating networks has been further refined for the special case of $n = 2$ in [25, 35]. In our work we also utilize the notion of Kronecker products between stochastic matrices, but this should not be confused with the notion of "stochastic"

Kronecker graphs previously mentioned. In this work we are *not* attempting to generate network models, but instead are utilizing novel aspects of Kronecker products and stochastic matrices that have, to the best of our knowledge, not been deeply explored; the ideas presented here would most closely be linked to that of [40] of the previously mentioned works.

Given the above, there are several key contributions of this paper. First, we provide a novel method for the computation of the mean first passage time (hitting time) of a Markov chain. In the process, we also provide an alternate closed form method for calculating pairwise hitting times between any two nodes for an arbitrary irreducible Markov chain. Second, we define and provide the first closed form solution for computing the hitting time given multiple (different) Markov chains on the same graph. We denote this extension the "group" hitting time. Third, our results also allow for the extension and calculation of the pairwise hitting time to the hitting time between any set of nodes for multiple random walkers; for any combination of specified starting nodes, we can calculate the first hitting time to a specified desired node. Fourth, we further extend the notion of group hitting time and hitting times between sets of nodes from multiple random walks on the same graph to random walks on multiple subgraphs. Finally, we provide a detailed numerical analysis to build intuition on the transition matrices that generate a minimal group hitting time. Before stating the paper organization, it is worthwhile to note that we achieve our analytic results by introducing a method of proof that utilizes the Kronecker product; thus our work not only provides results in Markov chain behavior, but also gives general insight into Kronecker graphs and stochastic matrices.

The paper is organized as follows. In section 1.1 we introduce notation that will be used throughout the paper and review useful concepts of Kronecker products and Markov chains. In section 2 we provide background on the hitting time of a Markov chain and provide our alternate formulation. In section 3 we introduce the group hitting time and hitting time between sets of nodes for a Markov chain and provide a detailed characterization. In section 4 we provide insight into optimal group hitting times through numerical optimization, and finally in section 5 we present our conclusions and future research directions.

1.1. Notation. In this section we define various useful concepts and notation. First, we provide an overview of some facts and results on Markov chains. Then we introduce the notation that will be used throughout the paper to deal with tensors, and we conclude with a brief summary of the Kronecker product (also known as a tensor product) and some of its properties.

1.1.1. Markov chains. A *Markov chain* is a sequence of random variables taking value in the *finite* set $\{1, \dots, n\}$ with the Markov property, namely, that the future state depends only on the present state. Let $X_k \in \{1, \dots, n\}$ denote the location of a random walker at time $k \in \{0, 1, 2, \dots\}$; then a Markov chain is *time-homogeneous* if $\mathbb{P}[X_{n+1} = j | X_n = i] = \mathbb{P}[X_n = j | X_{n-1} = i] = p_{i,j}$, where $P \in \mathbb{R}^{n \times n}$ is the *transition matrix* of the Markov chain. By definition, each transition matrix P is row-stochastic, i.e., $\sum_{i=j}^{n} p_{i,j} = 1$ for all $i \in \{1, \dots, n\}$. The vector $\boldsymbol{\pi} \in \mathbb{R}^{n \times 1}$ is a *stationary distribution* of P if $\sum_{i=1}^{n} \boldsymbol{\pi}_i = 1$, $0 \leq \boldsymbol{\pi}_i \leq 1$ for all $i \in \{1, \dots, n\}$, and $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$.

We will use the following well-known results on Markov chains. A Markov chain is *irreducible* if, for each $i, j \in \{1, \dots, n\}$, there exists $t \in N$ such that $(P^t)_{i,j} > 0$. If the Markov chain is *irreducible*, then there is a unique stationary distribution $\boldsymbol{\pi}$, and the corresponding eigenvalues of the transition matrix, λ_i for $i \in \{1, \dots, n\}$, are such

that $\lambda_1 = 1$, $|\lambda_i| \leq 1$, and $\lambda_i \neq 1$ for $i \in \{2, \dots, n\}$. We let $\rho[P]$ denote the spectral radius of the matrix P, i.e., the largest magnitude of the eigenvalues of P.

Unless otherwise mentioned, in this paper we consider finite irreducible time-homogeneous Markov chains. For more details on Markov chains or irreducible matrices, see [18] or [26, Chapter 8], respectively.

1.1.2. Tensor notation. We use the notation $A = [a_{i_1 \dots i_k, j}]$ to denote the matrix generated by elements $a_{i_1 \dots i_k, j}$, where the j index denotes the jth column of A and the rows of A are determined by cycling through index i_k, then i_{k-1}, and so forth. For example, consider $i_1, i_2, j \in \{1, \dots, n\}$; then

$$A = [a_{i_1 i_2, j}] = \begin{bmatrix} a_{11,1} & a_{11,2} & \dots & a_{11,n} \\ a_{12,1} & a_{12,2} & \dots & a_{12,n} \\ \vdots & \vdots & \dots & \vdots \\ a_{1n,1} & a_{1n,2} & \dots & \vdots \\ a_{21,1} & a_{21,2} & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots \\ a_{nn,1} & \dots & \dots & a_{nn,n} \end{bmatrix}.$$

For the case where $A = [a_{i,j}]$ this corresponds to the classic interpretation with element $a_{i,j}$ in the ith row and jth column of A. To avoid ambiguity, especially in cases when $A = [a_{i_1 \dots i_k, j}]$, at times we will refer to the (i, j) element of A by $A(i, j)$. Unless otherwise mentioned, vectors will be denoted by boldfaced letters (i.e., $\boldsymbol{a}$). We use the notation $\text{diag}[\boldsymbol{a}]$ to indicate the diagonal matrix generated by vector $\boldsymbol{a}$ and $\text{vec}(A)$ to indicate the vectorization of a matrix $A \in \mathbb{R}^{n \times m}$, where

$$\text{vec}(A) = [A(1,1), \dots, A(n,1), A(1,2), \dots, A(n,2), \dots, A(m,1), \dots, A(n,m)]^T.$$

In other words, even if we define A as $A = [a_{i_1 \dots i_k, j}]$, the vector $\text{vec}(A) = \text{vec}([a_{i_1 \dots i_k, j}])$ is simply a stacking of the columns of A. We also define the special matrix $[\mathcal{I}^{h_1 \dots h_n, k}_{i_1 \dots i_n, j}]$ as the matrix whose entries are all zero except for a single entry at $(h_1 \dots h_n, k)$ which has a value of 1, where $h_1, \dots, h_n, k$ can only take values within the range of values that $i_1, \dots, i_n, j$ take. With this matrix definition, it is easy to verify for $A = [a_{i_1 \dots i_n, j}]$ that $a_{h_1 \dots h_n, k} = \text{vec}([\mathcal{I}^{h_1 \dots h_n, k}_{i_1 \dots i_n, j}])^T \text{vec}(A)$. This enables us to go back and forth between the vectorized notation to the individual matrix elements. We denote $I_n \in \mathbb{R}^{n \times n}$ as the identity matrix of size n, $\mathbb{1}_n$ as the vector of ones of size n, and $\mathbb{0}_{n \times n}$ as the matrix zeros of size $n \times n$. We define a generalized Kronecker delta function $\delta_{i_1 i_2 \dots i_n, j}$, by

$$\delta_{i_1 \dots i_n, j} = \begin{cases} 1 & \text{if there exists } k \in \{1, \dots, n\} \text{ such that } i_k = j, \\ 0 & \text{otherwise.} \end{cases}$$

Then, with a slight abuse of notation, $A_d = [\delta_{i_1 i_2 \cdots_n, j} a_{i_1 \dots i_k, j}]$ represents the "diagonal" matrix generated by the elements of A. In reality, only when $A = [a_{i,j}]$ is the matrix truly diagonal. Finally, since the subscript operator is already in use, we use the superscript operation in parenthesis to delineate between two variables with the same name. For example, we write $A^{(1)}$ and $A^{(2)}$ to distinguish that the matrices are

different. A superscript without parentheses denotes a matrix raised to that power (i.e., $A^2 = AA$ and $(A^{(2)})^2 = A^{(2)}A^{(2)}$).

We are now ready to review some useful facts about Kronecker products. The Kronecker product, represented by the symbol $\otimes$, of two matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{q\times r}$ is an $nq \times mr$ matrix given by

$$A \otimes B = \begin{bmatrix} a_{1,1}B & \dots & a_{1,m}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \ddots & a_{n,m}B \end{bmatrix}.$$

To build some intuition, notice for $A \in \mathbb{R}^{n\times n}$ that $I_n \otimes A$ is the block diagonal matrix with n copies of A on the diagonal:

$$I_n \otimes A = \begin{bmatrix} A & \mathbb{0}_{n\times n} & \dots & \mathbb{0}_{n\times n} \\ \mathbb{0}_{n\times n} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{n\times n} \\ \mathbb{0}_{n\times n} & \dots & \mathbb{0}_{n\times n} & A \end{bmatrix}. \tag{1.1}$$

This implies for $A = I_n$ that $I_n \otimes A = I_n \otimes I_n = I_{n^2}$. The Kronecker product is bilinear and has many useful properties, two of which are summarized in the following lemma; see [16] for more information.

LEMMA 1.1. *Given the matrices A, B, C, and D, the following relations hold for the Kronecker product:*

(i) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$,

(ii) $(B^T \otimes A)\,\mathrm{vec}(C) = \mathrm{vec}(ACB)$,

where it is assumed that the matrices are of appropriate dimension when matrix multiplication or addition occurs. In addition, for matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$ with respective eigenvalues λ_i^A, $i \in \{1, \dots, n\}$, and λ_j^B, $j \in \{1, \dots, m\}$,

(iii) *the eigenvalues of $A \otimes B$ are $\lambda_i^A \lambda_j^B$ for $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$.*

2. The hitting time of a Markov chain. We begin by first reviewing properties of the single random walker hitting time and then provide a new formulation for determining this quantity. We show that this alternative formulation gives rise to new results for the pairwise hitting time between two specified nodes, and, in the subsequent section, we utilize it to extend the notion of the hitting time and pairwise hitting time to the multiple random walker case.

2.1. The hitting time of a Markov chain. Consider a strongly connected directed weighted graph $\mathcal{G} = (V, \mathcal{E}, P)$ with node set $V := \{1, \dots, n\}$, edge set $E \subset V \times V$, and *irreducible* row-stochastic matrix $P = [p_{i,j}]$ with the property that $p_{i,j} \geq 0$ if $(i, j) \in \mathcal{E}$ and $p_{i,j} = 0$ otherwise, and $\sum_{j=1}^{n} p_{i,j} = 1$ for all $i \in \{1, \dots, n\}$. We interpret the weight, $p_{i,j}$, of edge (i, j) as the probability of moving along that edge and, therefore, P as the transition matrix of a Markov chain.

Let $X_t \in \{1, \dots, n\}$ denote the location of a random walker at time $t \in \{0, 1, 2, \dots\}$. For any two nodes $i, j \in \{1, \dots, n\}$, the *first passage time from i to j*, denoted by $T_{i,j}$, is the first time that the random walker starting at node i at time 0 reaches node j, that is,

$$T_{i,j} = \min\{t \geq 1 \mid X_t = j \text{ given that } X_0 = i\}.$$

The mean first passage time from i to j is then given by $m_{i,j} = \mathbb{E}[T_{i,j}]$. It is convenient to define the *mean first passage time matrix* M as the matrix whose (i,j)th entries are given by $m_{i,j}$. The *mean first passage time from start node* i, denoted by $\boldsymbol{h}_i$, is the expected first passage time from node i to any other node in the graph. For a random walk described by transition matrix P with stationary distribution $\boldsymbol{\pi}$, the mean first passage time from node i is given by

$$\boldsymbol{h}_i = \sum_{j=1}^{n} m_{i,j}\boldsymbol{\pi}_j.$$

Remarkably, the mean first passage time is independent of the start node, that is, $\boldsymbol{h}_i = \boldsymbol{h}_j$ for all $i,j \in \{1,\dots,n\}$ [17]. For convenience, we let $H(P) = \boldsymbol{h}_i$, for all $i \in \{1,\dots,n\}$, denote the *hitting time*. Given the definition of the hitting time, one quickly sees that it can also be determined using the matrix M as follows:

$$\boldsymbol{\pi}^T M \boldsymbol{\pi} = (\boldsymbol{\pi}\otimes\boldsymbol{\pi})^T \operatorname{vec}(M) = \sum_{i=1}^{n} \boldsymbol{\pi}_i \sum_{j=1}^{n} \boldsymbol{\pi}_j m_{i,j} = H(P), \tag{2.1}$$

where the relation $\boldsymbol{\pi}^T M \boldsymbol{\pi} = (\boldsymbol{\pi}\otimes\boldsymbol{\pi})^T \operatorname{vec}(M)$ follows from identity (ii) in Lemma 1.1. We now provide an overview of the formulas for these quantities. The first passage time from i to j satisfies the recursive formula

$$T_{i,j} = \begin{cases} 1 & \text{with probability } p_{i,j}, \\ T_{k,j} + 1 & \text{with probability } p_{i,k}, k \neq j. \end{cases}$$

Taking the expectation, we get the following formula for $m_{i,j}$:

$$m_{i,j} = p_{i,j} + \sum_{k=1,k\neq j}^{n} p_{i,k}(m_{k,j}+1) = 1 + \sum_{k=1,k\neq j}^{n} p_{i,k} m_{k,j}.$$

The above formula can be expressed in various vectorized forms. The classical form as seen in the literature utilizes the matrix representation of M and is given by

$$(I-P)M = \mathbb{1}_n \mathbb{1}_n^T - PM_{\mathrm{d}}, \tag{2.2}$$

where P is the transition matrix of the Markov chain and $M_{\mathrm{d}} = \operatorname{diag}[\{1/\boldsymbol{\pi}_1,\dots,1/\boldsymbol{\pi}_n\}]$; the value for M_d is determined by premultiplying (2.2) by $\boldsymbol{\pi}^T$. For reasons which will become clear later, we vectorize the matrix M to generate a different representation of (2.2). Applying property (ii) of Lemma 1.1 to (2.2) gives

$$(I_n \otimes (I_n - P)) \operatorname{vec}(M) = \mathbb{1}_{n^2} - (I_n \otimes P) \operatorname{vec}(M_{\mathrm{d}}). \tag{2.3}$$

In a similar way as before, M_d can be determined from (2.3) with appropriate vector premultiplication.

The following classic result, derived from the matrix representation of M, shows that the hitting time can be written as a function of the eigenvalues of P [17].

THEOREM 2.1 (hitting time of an irreducible Markov chain). *Consider a Markov chain with an irreducible transition matrix P with eigenvalues $\lambda_1 = 1$ and λ_i, $i \in \{2,\dots,n\}$. The hitting time of the Markov chain is given by*

$$H(P) = 1 + \sum_{i=2}^{n} \frac{1}{1-\lambda_i}.$$

The proof for Theorem 2.1 relies on the fact that M_d is known or determined a priori; however, as will be seen next, there exists an equivalent expression for $H(P)$ which requires no such knowledge. Before presenting the alternative formulation for $H(P)$, we introduce the following useful result.

LEMMA 2.2 (eigenvalue shifting for stochastic matrices). *Let $P \in \mathbb{R}^{n\times n}$ be an irreducible row-stochastic matrix, and let E be any diagonal matrix with diagonal elements $E_{ii} \in \{0,1\}$, with at least one diagonal element which is zero. Then the eigenvalues λ_i of PE satisfy $|\lambda_i| < 1$ for all $i \in \{1,\dots,n\}$.*

Proof. The stochastic matrix P is nonnegative by definition, and therefore so is PE. Since P is irreducible, then $0 < PE < P$ and $\rho[PE] < \rho[P] = 1$ [33, Chapter 1, Exercise 1.16]. ∎

We are almost ready to present our alternative representation of the hitting time, but first we must introduce the following equality: notice that $\operatorname{vec}(M_d) = E\operatorname{vec}(M)$, where E is defined by $E = \operatorname{diag}[\boldsymbol{\delta}]$, where $\boldsymbol{\delta} = \operatorname{vec}([\delta_{i,j}])$. Using this interpretation of $\operatorname{vec}(M_d)$, we are ready to state our result.

THEOREM 2.3 (hitting times of an irreducible Markov chain). *Consider a Markov chain with an irreducible transition matrix $P \in \mathbb{R}^{n\times n}$. Then the following properties hold:*

(i) *the hitting time of the Markov chain is given by*

$$H(P) = (\boldsymbol{\pi}\otimes\boldsymbol{\pi})^T \operatorname{vec}(M), \quad \textit{where}$$
$$\operatorname{vec}(M) = \big(I_{n^2} - (I_n \otimes P)(I_{n^2} - E)\big)^{-1}\mathbb{1}_{n^2}; \quad \textit{and}$$

(ii) *the pairwise hitting time between nodes h and k, denoted $m_{h,k}$, of the Markov chain is given by*

$$m_{h,k} = \operatorname{vec}([\mathcal{I}_{i,j}^{h,k}])^T \operatorname{vec}(M).$$

Proof. First, notice that rearranging (2.3) and substituting $E\operatorname{vec}(M)$ for $\operatorname{vec}(M_d)$ gives that

$$\big(I_{n^2} - (I_n \otimes P)(I_{n^2} - E)\big)\operatorname{vec}(M) = \mathbb{1}_{n^2}. \tag{2.4}$$

From (2.1) we know that $H(P) = (\boldsymbol{\pi}\otimes\boldsymbol{\pi})\operatorname{vec}(M)$, and therefore it only remains to show that $I_{n^2} - (I_n \otimes P)(I_{n^2} - E)$ is in fact invertible. First, recall from (1.1) that $I_n \otimes P$ results in the block diagonal matrix, whose diagonal blocks consist of copies of P. Second, notice that $(I_{n^2} - E)$ is simply the identity matrix with some diagonal entries set to zero. It can be easily verified that $(I_n \otimes P)(I_{n^2} - E)$ results in the block diagonal matrix where each block contains the matrix P with one column set to zero. For example, for $P \in \mathbb{R}^{3\times 3}$ we have that

$$(I_n \otimes P)(I_{n^2} - E) = \begin{bmatrix} \begin{bmatrix} 0 & p_{12} & p_{13} \\ 0 & p_{22} & p_{23} \\ 0 & p_{32} & p_{33} \end{bmatrix} & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{bmatrix} p_{11} & 0 & p_{13} \\ p_{21} & 0 & p_{23} \\ p_{31} & 0 & p_{33} \end{bmatrix} & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{bmatrix} p_{11} & p_{12} & 0 \\ p_{21} & p_{22} & 0 \\ p_{31} & p_{32} & 0 \end{bmatrix} \end{bmatrix}.$$

Notice that each diagonal block will have at least one column set to zero. Hence, using Lemma 2.2, we have that the maximum eigenvalue of each block is strictly less than one in magnitude, and thus $\rho[(I_n \otimes P)(I_{n^2} - E)] < 1$. Let λ_i denote the eigenvalues of $(I_n \otimes P)(I_{n^2} - E)$. Then, since the eigenvalues of $I_{n^2} - (I_n \otimes P)(I_{n^2} - E)$ are simply $1-\lambda_i$ and $|\lambda_i| < 1$ for all $i \in \{1, \dots, n^2\}$, this implies the matrix $I_{n^2} - (I_n \otimes P)(I_{n^2} - E)$ is invertible and $\mathrm{vec}(M)$ is the unique solution to (2.4). ∎

Remark 2.4. Notice that when the transition matrix P is reducible, the hitting time $H(P)$ as defined in Theorem 2.3 is ill-posed (i.e., due to $I_{n^2} - (I_n \otimes P)(I_{n^2} - E)$ being singular). However, for reducible Markov chains with a single or multiple essential classes, i.e., for Markov chains whose condensation has a single or multiple sinks, there is a natural notion of hitting time for each essential class.

It should be noted that determining the closed form solution for pairwise-hitting time, and hence hitting time, can also be found using a technique which leverages the absorbing states of the chain, the details of which can be found in [33, Chapter 4]. We leverage our alternate construct in order to generalize the hitting time to multiple random walkers, as will be seen in the following sections.

3. Group hitting time of multiple Markov chains. In the following sections, we will expand the single-agent hitting time to the N-agent group hitting time. To build intuition, we initially assume that every agent can reach all nodes in the graph, and then move to the case where each agent only needs to have access to a subset of nodes in the graph.

3.1. Random walkers covering the full graph. Consider $h \in \{1, \dots, N\}$ strongly connected directed weighted graphs $\mathcal{G}^{(h)} = (V, E^{(h)}, P^{(h)})$ with same node sets $V := \{1, \dots, n\}$ and different edge sets $E^{(h)} \subset V \times V$ with corresponding irreducible row-stochastic matrices $P^{(h)} = [p^{(h)}_{i,j}]$ for $h \in \{1, \dots, N\}$ satisfying the property $p^{(h)}_{i,j} \geq 0$ if $(i,j) \in E^{(h)}$ and $p^{(h)}_{i,j} = 0$ otherwise, and $\sum_{j=1}^{n} p^{(h)}_{i,j} = 1$ for all $i \in \{1, \dots, n\}$. As before, each matrix $P^{(h)}$ is a transition matrix and describes a Markov chain on the graph.

Let $X^{(1)}_t, X^{(2)}_t, \dots, X^{(N)}_t \in \{1, \dots, n\}$ denote the location of N random walkers at time $t \in \{0, 1, 2, \dots\}$. For any $N+1$ nodes $i_1, \dots, i_N, j \in \{1, \dots, n\}$, the *first passage time from any i_h, $h \in \{1, \dots, N\}$ to j*, denoted by $T_{i_1 \dots i_N, j}$, is the first time that any random walker reaches node j when starting from nodes i_h, $h \in \{1, \dots, N\}$. More formally,

$$T_{i_1 \dots i_N, j} = \min \big\{ t \geq 1 \big| X^{(1)}_t = j \text{ or } \dots \text{ or } X^{(N)}_t = j \text{ given that } X^{(h)}_0 = i_h \text{ for all } h \in \{1, \dots, N\} \big\}. \tag{3.1}$$

With this definition, we are ready to state our first result for the N-random-walker system.

LEMMA 3.1 (recursive formulation of first passage time for multiple random walkers). *Let $m_{i_1 \dots i_N, j} = \mathbb{E}[T_{i_1 \dots i_N, j}]$ denote the first passage time from any i_h, $h \in \{1, \dots, N\}$ to j. Also, let $P^{(h)}$ be the irreducible transition matrix associated with $\mathcal{G}^{(h)}$. Then*

$$m_{i_1 \dots i_N, j} = 1 + \sum_{k_1, \dots, k_N \neq j} m_{k_1 \dots k_N, j} p^{(1)}_{i_1, k_1} \dots p^{(N)}_{i_N, k_N},$$

or, in matrix notation,

$$M = \mathbb{1}_{n^N}\mathbb{1}_n^T + (P^{(1)} \otimes \cdots \otimes P^{(N)})M - (P^{(1)} \otimes \cdots \otimes P^{(N)})M_d, \tag{3.2}$$

where $M = [m_{i_1 \ldots i_N, j}]$ *and* $M_d = [\delta_{i_1 \ldots i_N, j} m_{i_1 \ldots i_N, j}]$.

Proof. For clarity, we first study the 2-agent case and then generalize. By definition, the 2-agent first passage time satisfies the recursive formula

$$T_{i_1 i_2, j} = \begin{cases} 1 & \text{with probability } p^{(1)}_{i_1,j} + p^{(2)}_{i_2,j} - p^{(1)}_{i_1,j} p^{(2)}_{i_2,j}, \\ T_{k_1 k_2, j} + 1 & \text{with probability } p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2} \text{ such that } k_1, k_2 \neq j. \end{cases}$$

First, notice that the probability of moving to node j directly from nodes i_1 or i_2 is given by the non-mutually-exclusive probability $p^{(1)}_{i_1,j} + p^{(2)}_{i_2,j} - p^{(1)}_{i_1,j} p^{(2)}_{i_2,j}$ and that this probability can also be described by the equivalent expression $1 - (1 - p^{(1)}_{i_1,j})(1 - p^{(2)}_{i_2,j})$. Substituting in the alternate expression for the probability and taking the expectation of $T_{i_1 i_2, j}$, we have that

$$\begin{aligned} \mathbb{E}[T_{i_1 i_2, j}] &= 1 - (1 - p^{(1)}_{i_1,j})(1 - p^{(2)}_{i_2,j}) + \sum_{k_1, k_2 \neq j} (\mathbb{E}[T_{k_1 k_2, j}] + 1) p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2} \\ &= 1 - (1 - p^{(1)}_{i_1,j})(1 - p^{(2)}_{i_2,j}) + \sum_{k_1, k_2 \neq j} m_{k_1 k_2, j} p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2} + \sum_{k_1, k_2 \neq j} p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2}. \end{aligned} \tag{3.3}$$

Utilizing the row-stochastic property of $P^{(1)}$ and $P^{(2)}$, expand $(1 - p^{(1)}_{i_1,j})(1 - p^{(2)}_{i_2,j})$ above to get

$$\begin{aligned} (1 - p^{(1)}_{i_1,j})(1 - p^{(2)}_{i_2,j}) &= \left(\sum_{k_1 \neq j} p^{(1)}_{i_1,k_1} \right) \left(\sum_{k_2 \neq j} p^{(2)}_{i_2,k_2} \right) \\ &= \sum_{k_1, k_2 \neq j} p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2}; \end{aligned}$$

substituting this back into (3.3) gives the result

$$m_{i_1 i_2, j} = 1 + \sum_{k_1, k_2 \neq j} m_{k_1 k_2, j} p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2}$$

or, in matrix notation,

$$M = \mathbb{1}_{n^2}\mathbb{1}_n^T + (P^{(1)} \otimes P^{(2)})M - (P^{(1)} \otimes P^{(2)})M_d,$$

where $M = [m_{i_1 i_2, j}]$ and $M_d = [\delta_{i_1 i_2, j} m_{i_1 i_2, j}]$.

Similarly to the 2-agent case, the N-agent first passage time satisfies the recursive formula

$$T_{i_1 \ldots i_N, j} = \begin{cases} 1 & \text{with probability } 1 - (1 - p^{(1)}_{i_1,j})(1 - p^{(2)}_{i_2,j}) \ldots (1 - p^{(N)}_{i_N,j}), \\ T_{k_1 \ldots k_N, j} + 1 & \text{with probability } p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2} \ldots p^{(N)}_{i_N,k_N} \\ & \text{such that } k_1, \ldots, k_N \neq j. \end{cases}$$

As in the 2-agent case, we take expectations and utilize the row-stochastic properties of each $P^{(h)}$ (i.e., $(1 - p^{(h)}_{i_h,j}) = \sum_{k_h \neq j} p^{(1)}_{i_h,k_h}$) to reach the result. ∎

Similarly to how the random variable $T_{i,j}$ arises from the Markov chain described in the transition matrix P, we can think of the random variable $T_{i_1 \dots i_N, j}$ as being generated from the Markov chain described by the stochastic matrix $P^{(1)} \otimes \cdots \otimes P^{(N)}$.

Given the formulation for the N-agent first passage time matrix, we can define a quantity similar to the mean first passage time given by (2.1). In order to do this, we first need to define the frequency of being at any given node in the graph. This quantity should take into account the probability of being at one node instead of another in the limit of the random walks. Since the random walks are evolving in parallel, the relative frequency of being at a specific node is simply the average of the N random walkers visit frequency at that node. More explicitly, $\boldsymbol{\pi}_{\text{ave}} = \sum_{h=1}^{N} (\boldsymbol{\pi}^{(h)})/N$. Then, similarly to the single-agent case, the *N-agent mean first passage time from start nodes i_h, $h \in \{1, \dots, N\}$*, denoted $\boldsymbol{h}_{i_1 \dots i_N}$, is given by

$$\boldsymbol{h}_{i_1 \dots i_N} = \sum_{j=1}^{n} m_{i_1 \dots i_N, j} \boldsymbol{\pi}_{\text{ave},j}.$$

Therefore, the average time to go from any set of N nodes to a single node in a graph is given by

$$\begin{aligned} H_N &= \sum_{i_1=1}^{n} \boldsymbol{\pi}_{i_1}^{(1)} \cdots \sum_{i_N=1}^{n} \boldsymbol{\pi}_{i_N}^{(N)} \sum_{j=1}^{n} m_{i_1 \dots i_N, j} \boldsymbol{\pi}_{\text{ave},j} \\ &= (\boldsymbol{\pi}^{(1)} \otimes \cdots \otimes \boldsymbol{\pi}^{(N)})^T M \boldsymbol{\pi}_{\text{ave}} \\ &= (\boldsymbol{\pi}_{\text{ave}} \otimes \boldsymbol{\pi}^{(1)} \otimes \cdots \otimes \boldsymbol{\pi}^{(N)})^T \operatorname{vec}(M), \end{aligned}$$

where we denote H_N as being the N-agent *group* hitting time. It is clear the group hitting time can be written as the function $P^{(1)} \times \cdots \times P^{(N)} \mapsto H_N(P^{(1)}, \dots, P^{(N)})$, but to ease notation we simply write H_N.

Given the definition of H_N we are almost ready to state our next result. First, we must introduce the following equality: notice that $\operatorname{vec}(M_d) = E \operatorname{vec}(M)$, where E is defined by $E = \operatorname{diag}[\boldsymbol{\delta}]$, where $\boldsymbol{\delta} = \operatorname{vec}([\delta_{i_1 \dots i_N, j}])$. Using this interpretation of $\operatorname{vec}(M_d)$ we are ready to state our result.

THEOREM 3.2 (group hitting time for irreducible Markov chains). *Consider N multiple Markov chains, each with an irreducible transition matrix $P^{(h)} \in \mathbb{R}^{n \times n}$. Let $\boldsymbol{\pi}_{\text{ave}} = (\sum_{h=1}^{N} \boldsymbol{\pi}^{(h)})/N$, and let $E \in \mathbb{R}^{n^{N+1} \times n^{N+1}}$ be the diagonal matrix which satisfies the equality $E \operatorname{vec}(M) = \operatorname{vec}(M_d)$. Then the following properties hold:*

(i) *the group hitting time of the Markov chain is given by*

(3.4)
$$H_N = (\boldsymbol{\pi}_{\text{ave}} \otimes \boldsymbol{\pi}^{(1)} \otimes \cdots \otimes \boldsymbol{\pi}^{(N)})^T \operatorname{vec}(M), \quad \textit{where}$$
$$\operatorname{vec}(M) = \left(I_{n^{N+1}} - (I_n \otimes P^{(1)} \otimes \cdots \otimes P^{(N)})(I_{n^{N+1}} - E)\right)^{-1} \mathbb{1}_{n^{N+1}}; \ \textit{and}$$

(ii) *the hitting time between nodes $h_1, \dots, h_N$ and k, denoted $m_{h_1 \dots h_N, k}$, of the Markov chain is given by*

$$m_{h_1 \dots h_N, k} = \operatorname{vec}([\mathcal{I}_{i_1 \dots i_N, j}^{h_1 \dots h_N, k}])^T \operatorname{vec}(M).$$

Proof. Letting $P = P^{(1)} \otimes \cdots \otimes P^{(N)}$, notice that (3.2) can be written in the vectorized form

$$\operatorname{vec}(M) = \mathbb{1}_{n^{N+1}} + (I_n \otimes P) \operatorname{vec}(M) - (I_n \otimes P) \operatorname{vec}(M_d).$$

Rearranging terms and substituting $E\,\mathrm{vec}(M)$ for $\mathrm{vec}(M_d)$ gives

$$\big(I_{n^{N+1}} - (I_n \otimes P)(I_{n^{N+1}} - E)\big)\,\mathrm{vec}(M) = \mathbb{1}_{n^{N+1}}. \tag{3.5}$$

To ease the complexity of our notation, we move forward by looking at the 2-agent case and then generalizing from there. For the 2-agent case with $P = P^{(1)} \otimes P^{(2)}$ the system (3.5) becomes

$$\big(I_{n^3} - (I_n \otimes P)(I_{n^3} - E)\big)\,\mathrm{vec}(M) = \mathbb{1}_{n^3},$$

indicating a unique solution exists if $I_{n^3} - (I_n \otimes P)(I_{n^3} - E)$ is invertible. Let the eigenvalues $\lambda_i^{(1)}$ and $\lambda_i^{(2)}$ be associated with transition matrices $P^{(1)}$ and $P^{(2)}$, respectively. Then, from property (iii) of Lemma 1.1, the eigenvalues of $P^{(1)} \otimes P^{(2)}$ are $\lambda_j^{(1)}\lambda_k^{(2)}$ for $j, k \in \{1, \ldots, n\}$. This means that when just two $P^{(i)}$ are periodic, the Kronecker product can result in a Markov chain which has multiple eigenvalues at 1, making this chain reducible. Therefore, we must show that irreducible blocks can be constructed that allow the application of Lemma 2.2 as before. We begin by noting that for a reducible $P = P^{(1)} \otimes P^{(2)}$ we can apply a series of permutation matrices $S^{(i)}$ for $i \in \{1, \ldots, m\}$ to P [26, Chapter 8.3] such that

$$\bar{P} = \begin{bmatrix} A^{(1)} & * & \ldots & * \\ \mathbb{0} & A^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ \mathbb{0} & \ldots & \mathbb{0} & A^{(k)} \end{bmatrix}, \tag{3.6}$$

where $\bar{P} = SPS^T$ with permutations $S = (S^{(m)}) \ldots (S^{(1)})$, and each $A^{(i)}$ is irreducible. Since P is row-stochastic, each irreducible component is either substochastic ($\rho[A^{(i)}] < 1$) or stochastic ($\rho[A^{(i)}] = 1$). Since we only need to consider the case in which $A^{(i)}$ is stochastic, let $B \in \mathbb{R}^{r \times r}$ denote an irreducible stochastic matrix $A^{(i)} \in \mathbb{R}^{r \times r}$ in (3.6). For B irreducible, $\bar{P}$ has the form

$$\bar{P} = \begin{bmatrix} A^{(1)} & * & \ldots & \ldots & \ldots & * \\ \mathbb{0} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & B & \mathbb{0} & \ldots & \mathbb{0} \\ \vdots & \ddots & \ddots & A^{(l)} & \ldots & * \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbb{0} & \ldots & \ldots & \ldots & \mathbb{0} & A^{(k)} \end{bmatrix}.$$

Notice from the definition of $\mathrm{vec}(M)$ and the block diagonal structure of $(I_n \otimes P)$ that the jth block in $(I_n \otimes P)$ corresponds to mean first passage times, $m_{i_1 i_2, j}$, to node j. Since a permutation matrix simply acts as a relabeling of elements, assume without loss of generality that the elements associated with B vary from $i_1, i_2 \in \{1, \ldots, r\}$.

Therefore, the equations associated with B have the form

$$\begin{aligned} m_{i_1 i_2, j} &= 1 + \sum_{k_1,k_2} m_{k_1 k_2, j} p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2} - \sum_{\substack{k_1=j \\ k_2 \neq j}} m_{j k_2, j} p^{(1)}_{i_1,j} p^{(2)}_{i_2,k_2} \\ &\quad - \sum_{\substack{k_2=j \\ k_1 \neq j}} m_{k_1 j, j} p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,j} - m_{jj,j} p^{(1)}_{i_1,j} p^{(2)}_{i_2,j} \\ &= 1 + \sum_{k_1,k_2 \neq j} m_{k_1 k_2, j} p^{(1)}_{i_1,k_1} p^{(2)}_{i_2,k_2}, \end{aligned}$$

where the subtracted terms in the expression above are associated with the nonzero entries of E (i.e., to the columns of P that are set to zero). If all subtracted terms are zero for each $m_{i_1 i_2, j}$, $i_1, i_2 \in \{1, \dots, r\}$, this implies that there exist no $i_1, i_2 \in \{1, \dots, r\}$ such that $p^{(1)}_{i_1,j} > 0$ or $p^{(2)}_{i_2,j} > 0$; that can only be true if there exists no path to node j from any node i_1 or i_2, which is impossible by definition of each $P^{(i)}$. Therefore, for each irreducible row-stochastic component of $\bar{P}$, there is at least one nonzero element E such that a column of that component is set to zero, allowing us to apply Lemma 2.2.

In the N-agent case, similarly to the 2-agent case, a unique solution exists if $(I_{n^{N+1}} - (I_n \otimes P)(I_{n^{N+1}} - E))$ from (3.5) is invertible. For this to hold true it must be that $(I_n \otimes P)(I_{n^{N+1}} - E)$ has a spectral radius less than 1. Similarly to before, given that $I_n \otimes P$ and hence $(I_n \otimes P)(I_{n^{N+1}} - E)$ is a block diagonal matrix, we need only show that each block has a spectral radius less than 1. The proof follows a line of argument parallel to the 2-agent case. More explicitly, each P can be deconstructed into a block upper triangular matrix, as in (3.6), composed of square matrices along the diagonal, each of which is irreducible. Since P is row-stochastic, then each irreducible block is either row-stochastic or substochastic. If the irreducible block is substochastic, then its spectral radius is less than 1. If an irreducible block is row-stochastic, then all we must show is that a column of the row-stochastic block is necessarily set to zero when $I_n \otimes P$ is multiplied by $(I_{n^{N+1}} - E)$. As in the 2-agent case, at least one negative entry in the definition of the $m_{i_1 \dots i_N, j}$ must be nonzero due to the connectivity of the graph. Therefore there is a nonzero entry in E which forces a column of the irreducible block to be zero in $(I_n \otimes P)(I_{n^{N+1}} - E)$. Therefore, by Lemma 2.2, $(I_n \otimes P)(I_{n^{N+1}} - E)$ has a spectral radius less than 1. ∎

Given this representation of the N-agent group hitting time, a natural question is whether one can determine a simplified expression for this quantity which is a function of the eigenvalues of $P^{(h)}$, similar to the expression in Theorem 2.1. As mentioned earlier, proof of that theorem relies on the ability to extort knowledge of M_d. Consider, for example, the 2-agent case; if we try to find the entries of M_d in a similar fashion to the single-agent case by premultiplying (3.2) with $\boldsymbol{\pi}^{(1)} \otimes \boldsymbol{\pi}^{(2)}$, we have that

$$\mathbb{1}_n^T = (\boldsymbol{\pi}^{(1)} \otimes \boldsymbol{\pi}^{(2)})^T M_d.$$

This is a system of n equations and $n(2n-1)$ unknowns, and thus the solution of M_d is underdetermined. Therefore, even though one may be able to express the group hitting time as a function of the eigenvalues, it is currently not well understood how.

3.2. Random walkers covering subgraphs. The group hitting time as stated thus far is of interest in its own right. However, in many applications it is often desirable to have a notion of the same quantity when multiple agents don't have

access to the entire graph. In the following section we tackle this problem by utilizing reducible graphs. We will leverage the framework of the N-agent group hitting time for irreducible Markov chains in order to generalize our results to reducible Markov chains.

Consider $h \in \{1, \dots, N\}$ directed weighted graphs $\mathcal{G}^{(h)} = (V^{(h)}, E^{(h)}, P^{(h)})$ with node sets $V^{(h)} \subset \{1, \dots, n\}$ such that $\cup_{h=1}^{N} V^{(h)} = \{1, \dots, n\}$. The edge sets satisfy $E^{(h)} \subset V^{(h)} \times V^{(h)}$ and have corresponding row-stochastic matrices $P^{(h)} = [p_{i,j}^{(h)}]$ for $h \in \{1, \dots, N\}$ with the property $p_{i,j}^{(h)} \geq 0$ if $\{i, j\} \in E^{(h)}$ and $p_{i,j}^{(h)} = 0$ otherwise. As before, each matrix $P^{(h)}$ is the transition matrix of a Markov chain on $V^{(h)}$.

Let $X_t^{(h)} \in V^{(h)}$ denote the location of N random walkers at time $t \in \{0, 1, 2, \dots\}$. For any $N + 1$ nodes $i_h \in V^{(h)}$ and $j \in \{1, \dots, n\}$, the *first passage time from any* i_h, $h \in \{1, \dots, N\}$ *to* j, denoted by $T_{i_1 \dots i_N, j}$, is the first time that any random walker reaches node j, when starting from nodes i_h, $h \in \{1, \dots, N\}$, and is given by

$$T_{i_1 \dots i_N, j} = \min \big\{ t \geq 1 \big| X_t^{(1)} = j \text{ or } \dots \text{ or } X_t^{(N)} = j \\ \text{given that } X_0^{(h)} = i_h \text{ for all } h \in \{1, \dots, N\} \big\}.$$

Similar to the case where each agent has access to the entire graph, we have the following lemma, whose proof is equivalent to that of Lemma 3.1.

LEMMA 3.3 (recursive formulation of first passage time for multiple random walkers over subgraphs). *Consider the graphs* $\mathcal{G}^{(h)} = (V^{(h)}, E^{(h)}, P^{(h)})$ *satisfying the property* $\cup_{h=1}^{N} V^{(h)} = \{1, \dots, n\}$, *and let* $P^{(h)}$ *be the transition matrix associated with* $\mathcal{G}^{(h)}$. *Also, let* $|V^{(h)}|$ *denote the cardinality of each node set, and let* $m_{i_1 i \dots i_N, j} = \mathbb{E}[T_{i_1 \dots i_N, j}]$ *denote the first passage time from any* $i_h \in V^{(h)}$ *to* $j \in \{1, \dots, n\}$. *Then*

$$m_{i_1 \dots i_N, j} = 1 + \sum_{k_1, \dots, k_N \neq j} m_{k_1 \dots k_N, j} p_{i_1, k_1}^{(1)} \dots p_{i_N, k_N}^{(N)},$$

or, in matrix notation,

$$M = \mathbb{1}_\alpha \mathbb{1}_n^T + (P^{(1)} \otimes \dots \otimes P^{(N)}) M - (P^{(1)} \otimes \dots \otimes P^{(N)}) M_d, \tag{3.7}$$

where $\alpha = \prod_{h=1}^{N} |V^{(h)}|$, $M = [m_{i_1 \dots i_N, j}]$, *and* $M_d = [\delta_{i_1 \dots i_N, j} m_{i_1 \dots i_N, j}]$.

Proof. The formulation of this system follows in a similar fashion the N-agent case, with the exception that now, if node i_h has the property that $i_h \notin V^{(h)} \subset \{1, \dots, n\}$, then the corresponding $m_{i_1 \dots i_N, j}$ value is zero. ∎

Given the formulation for the N-agent first passage time matrix over multiple subgraphs, we now determine the average first visit time to any node in the full graph. Like before, first we calculate the relative frequency of being at any given node. Unlike before, however, each agent operates over a subset of the nodes in the graph. Let $\boldsymbol{\pi}^{(h)}$ be the stationary distribution associated with Markov chain $P^{(h)} \in \mathbb{R}^{r \times r}$, where $r \leq n$, and for convenience assume that $V^{(h)} \subset \{1, \dots, n\}$ is an ordered ascending set (i.e., $V^h = \{3, 9, 20\}$). Then let $\tilde{\boldsymbol{\pi}}^{(h)}$ be the vector whose entries are given by $\tilde{\boldsymbol{\pi}}_{V_i^{(h)}}^{(h)} = \boldsymbol{\pi}_i^{(h)}$ for $i \in \{1, \dots, |V^{(h)}|\}$ and $\tilde{\boldsymbol{\pi}}_i^{(h)} = 0$ otherwise. In other words, $\tilde{\boldsymbol{\pi}}^{(h)}$ corresponds to the stationary distribution of each agent over the entire graph, not just its subgraph. Therefore, if an agent never visits a node, its visit frequency to that node is 0. Given the padded vector $\tilde{\boldsymbol{\pi}}^{(h)}$, we write the average visit frequency as $\tilde{\boldsymbol{\pi}}_{\text{ave}} = \sum_{h=1}^{N} (\tilde{\boldsymbol{\pi}}^{(h)})/N$. Notice that this interpretation of average visit frequency takes into account that multiple Markov chains are running in parallel.

Now, the *N-agent mean first passage time from start nodes i_h, $h \in \{1,\dots,N\}$*, denoted $\boldsymbol{h}_{i_1\dots i_N}$, is given by

$$\boldsymbol{h}_{i_1\dots i_N} = \sum_{j=1}^{n} m_{i_1\dots i_N,j}\tilde{\boldsymbol{\pi}}_{\mathrm{ave},j}.$$

Therefore, the average time to go from any set of N nodes to a single node in a graph is given by

$$\begin{aligned} H_N &= \sum_{i_1=1}^{|V^{(1)}|} \boldsymbol{\pi}_{i_1}^{(1)} \cdots \sum_{i_N=1}^{|V^{(N)}|} \boldsymbol{\pi}_{i_N}^{(N)} \sum_{j=1}^{n} m_{i_1\dots i_N,j}\tilde{\boldsymbol{\pi}}_{\mathrm{ave},j} \\ &= (\boldsymbol{\pi}^{(1)} \otimes \cdots \otimes \boldsymbol{\pi}^{(N)}) M \tilde{\boldsymbol{\pi}}_{\mathrm{ave}} \\ &= (\tilde{\boldsymbol{\pi}}_{\mathrm{ave}} \otimes \boldsymbol{\pi}^{(1)} \otimes \cdots \otimes \boldsymbol{\pi}^{(N)}) \operatorname{vec}(M), \end{aligned} \tag{3.8}$$

where as before we denote H_N as the N-agent group hitting time.

We are now ready to state our main result.

THEOREM 3.4 (group hitting time for irreducible subgraphs). *Consider the N graphs $\mathcal{G}^{(h)} = (V^{(h)}, E^{(h)}, P^{(h)})$ satisfying the property $\cup_{h=1}^{N} V^{(h)} = \{1,\dots,n\}$, and let $P^{(h)} \in \mathbb{R}^{|V^{(h)}|\times|V^{(h)}|}$ be the* irreducible *transition matrices associated with $\mathcal{G}^{(h)}$. Also, let $\tilde{\boldsymbol{\pi}}_{\mathrm{ave}} = (\sum_{i=1}^{N} \tilde{\boldsymbol{\pi}}^{(i)})/N$, and let $E \in \mathbb{R}^{\alpha n\times \alpha n}$ be the diagonal matrix which satisfies the equality $E\operatorname{vec}(M) = \operatorname{vec}(M_d)$. Then the following hold:*

(i) *the group hitting time of the Markov chain is given by*

$$\begin{aligned} H_N &= (\tilde{\boldsymbol{\pi}}_{\mathrm{ave}} \otimes \boldsymbol{\pi}^{(1)} \otimes \cdots \otimes \boldsymbol{\pi}^{(N)})^T \operatorname{vec}(M), \quad \textit{where} \\ \operatorname{vec}(M) &= \left(I_{\alpha n} - (I_n \otimes P^{(1)} \otimes \cdots \otimes P^{(N)})(I_{\alpha n} - E)\right)^{-1} \mathbb{1}_{\alpha n}, \end{aligned} \tag{3.9}$$

and $\alpha = \prod_{h=1}^{N} |V^{(h)}|$; and

(ii) *the hitting time between nodes $h_1,\dots,h_N$ and k, denoted $m_{h_1\dots h_N,k}$, of the Markov chain is given by*

$$m_{h_1\dots h_N,k} = \operatorname{vec}([\mathcal{I}_{i_1\dots i_N,j}^{h_1\dots h_N,k}])^T \operatorname{vec}(M).$$

Proof. The proof of this theorem follows the exact same logic as the proof of Theorem 3.2 ∎

This theorem immediately leads to the following corollary.

COROLLARY 3.5 (group hitting time for reducible Markov chains). *Consider N Markov chains with transition matrices $P^{(h)} \in \mathbb{R}^{n\times n}$, $h \in \{1,\dots,N\}$, defined over the same graph and satisfying the following property: if $P^{(h)}$ is reducible, then there exists a permutation matrix $S^{(i)}$ such that $(S^{(h)})^T P^{(h)} S^{(h)}$ is block upper triangular with exactly one nonzero irreducible block on the diagonal and all other diagonal entries equal to zero. (In other words, each Markov chain has a unique sink in the condensation digraph, and all other strongly connected components contain a single node.) Under this assumption, let $\boldsymbol{\pi}^{(h)}$ be the unique stationary distribution for $P^{(h)}$. Let $\boldsymbol{\pi}_{\mathrm{ave}} = (\sum_{h=1}^{N} \boldsymbol{\pi}^{(h)})/N$, with the property that $\boldsymbol{\pi}_{\mathrm{ave},j} \neq 0$ for all j. Also, let $E \in \mathbb{R}^{n^{N+1}\times n^{N+1}}$ be the diagonal matrix which satisfies the equality $E\operatorname{vec}(M) = \operatorname{vec}(M_d)$. Then the following properties hold:*

(i) *the group hitting time of the Markov chain is given by*

$$H_N = (\boldsymbol{\pi}_{\text{ave}} \otimes \boldsymbol{\pi}^{(1)} \otimes \cdots \otimes \boldsymbol{\pi}^{(N)})^T \operatorname{vec}(M), \quad \textit{where}$$
$$\operatorname{vec}(M) = (I_{n^{N+1}} - (I_n \otimes P^{(1)} \otimes \cdots \otimes P^{(N)})(I_{n^{N+1}} - E))^{-1} \mathbb{1}_{n^{N+1}}; \textit{ and}$$

(ii) *the hitting time between nodes $h_1, \ldots, h_N$ and k, denoted $m_{h_1 \ldots h_N, k}$, of the Markov chain is given by*

$$m_{h_1 \ldots h_N, k} = \operatorname{vec}([\mathcal{I}^{h_1 \ldots h_N, k}_{i_1 \ldots i_N, j}])^T \operatorname{vec}(M).$$

Proof. First, note that the Kronecker product of two block diagonal matrices generates a block diagonal matrix. Second, notice from Definition 1.1(i) that

$$(S^{(1)})^T P^{(i)} S^{(1)} \otimes \cdots \otimes (S^{(N)})^T P^{(i)} S^{(N)}$$
$$= ((S^{(1)})^T \otimes \cdots \otimes (S^{(N)})^T)(P^{(1)} \otimes \cdots \otimes P^{(N)})(S^{(1)} \otimes \cdots \otimes S^{(N)}),$$

so there exists a permutation matrix $(S^{(1)} \otimes \cdots \otimes S^{(N)})$ that makes $(P^{(1)} \otimes \cdots \otimes P^{(N)})$ block diagonal. Since exactly one block from each matrix is not exactly zero, the same is true of $(S^{(1)})^T P^{(i)} S^{(1)} \otimes \cdots \otimes (S^{(N)})^T P^{(i)} S^{(N)}$. This block corresponds to matrix P in Theorem 3.4. The rest of the proof follows by noticing that each node in the graph is reached if and only if $\boldsymbol{\pi}_{\text{ave}}, j \neq 0$ for all j. This is due to the fact that $\boldsymbol{\pi}^{(h)}_k \neq 0$ when k denotes a persistent reachable node in the graph, and $\boldsymbol{\pi}^{(h)}_k = 0$ otherwise. ∎

3.3. Computational complexity. Due to the extensive use of Kronecker products, it is important to verify the memory and computational costs of the group hitting time. Looking at (3.4), we can assert that the group hitting time is affected by the curse of dimensionality; with n nodes and N agents, the matrix $I_n \otimes P^{(1)} \otimes \cdots \otimes P^{(N)}$ contains n^{2N+2} elements. For example, given $n = 100$ nodes and $N = 10$, agents the size of that matrix is $10^{44} \times 10^{44}$. The most intense operation is the inversion of $I_n \otimes P^{(1)} \otimes \cdots \otimes P^{(N)}$, which requires a cost of $O(k^3)$, where k is the number of elements in the matrix [6], and thus in our case this becomes $O(n^{6N+6})$. Noticing that the first Kronecker product in (3.5) is between the identity matrix and the $P = P^{(1)} \otimes \cdots \otimes P^{(N)}$ matrix, this implies $I_n \otimes P$ is block diagonal, and therefore we can store and invert single blocks, reducing memory cost to $O(n^{2N})$ and inversion cost to $O(n^{6N})$. For the more general hitting time formula described by (3.8), the complexity can be further reduced. Given that (3.8) describes random walks on subgraphs, then $|V^{(h)}| = n\beta_h$ for some $\beta_h \in (0, 1]$, and therefore the number of elements in the matrix $I_{\alpha n} \otimes P^{(1)} \otimes \cdots \otimes P^{(N)}$ is $(\prod_{h=1}^N \beta_h^2) n^{2N+2}$. This leads to a computational complexity for the inversion equal to $O\big((\prod_{h=1}^N \beta_h^6) n^{6N+6}\big)$. Similarly to before, we can take advantage of the fact that the first Kronecker product in $I_{\alpha n} \otimes P^{(1)} \otimes \cdots \otimes P^{(N)}$ is the identity matrix, reducing memory and inversion costs to $(\prod_{h=1}^N \beta_h^2) n^{2N}$ and $O\big((\prod_{h=1}^N \beta_h^6) n^{6N}\big)$, respectively.

In special circumstances, the above computational complexity can be dramatically reduced. This happens when the intersection between a subgraph of a single agent does not intersect with any other agent's subgraph. In this case, the disjoint agent's hitting time over its subgraph can be calculated independently of the group hitting time of the other $N-1$ agents. The single agent hitting time can then be averaged with the $N-1$ agent group hitting time to generate the N agent group hitting time. In the case where every agent owns its own disjoint region, the computational complexity scales to $O(\sum_{h=1}^N |V^{(h)}|^{12})$, or to $O(\sum_{h=1}^N |V^{(h)}|^6)$ when exploiting the Kronecker

product between the identity matrix and a transition matrix, as was done previously. It is clear that uniformly partitioning the graph between agents, or partitioning as close to uniform as is possible, we get the lowest computational complexity.

In the next section we compute an optimized group hitting time for various graph topologies.

4. Numerical optimization of the group hitting time. In the following sections we study the transition matrices that arise from the numerical optimization of the group hitting time. In particular, we look for the transition matrices, $P^{(1)}, P^{(2)}, \dots, P^{(N)}$, that minimize (3.9) as described by Problem 1 below. The problem is numerically solved using a *sequential quadratic programming* solver as implemented by MATLAB's `fmincon` optimization algorithm, details of which are discussed in section 4.3.

Problem 1 (group hitting time minimization). Given $h \in \{1, \dots, N\}$ directed weighted graphs $\mathcal{G}^{(h)} = (V^{(h)}, E^{(h)}, P^{(h)})$ with node sets $V^{(h)} \subset \{1, \dots, n\}$, edge sets $E^{(h)} \subset V^{(h)} \times V^{(h)}$, and stationary distributions $\boldsymbol{\pi}^{(h)}$, find the corresponding transitions matrices $P^{(h)}$ solving

$$
\begin{aligned}
\inf \quad & (\tilde{\boldsymbol{\pi}}_{\text{ave}} \otimes \boldsymbol{\pi})^T (I_{\alpha n} - (I_n \otimes (P^{(1)} \otimes P^{(2)} \cdots \otimes P^{(n)}))(I_{\alpha n} - E))^{-1} \mathbb{1}_{\alpha n} \\
\text{subject to} \quad & P^{(i)} \mathbb{1}_{|V^{(i)}|} = \mathbb{1}_{|V^{(i)}|} \text{ for each } i \in \{1, \dots, N\}, \\
& (\boldsymbol{\pi}^{(i)})^T P^{(i)} = (\boldsymbol{\pi}^{(i)})^T \text{ for each } i \in \{1, \dots, N\}, \\
& 0 \le p^{(i)}_{h,k} \le 1 \text{ for each } (h,k) \in E^{(i)} \text{ and } i \in \{1, \dots, N\}, \\
& p^{(i)}_{h,k} = 0 \text{ for each } (h,k) \notin E^{(i)} \text{ and } i \in \{1, \dots, N\}, \\
& P^{(i)} \text{ is irreducible for } i \in \{1, \dots, N\},
\end{aligned}
$$

where $\alpha = \prod_{h=1}^{N} |V^{(h)}|$.

To the best of our knowledge, a known solution to the above problem exists only for the case where $\mathcal{G}$ is a complete graph and $N = 1$ [20].

It is worth noting that Problem 1 is looking for a infimum and not for minimum; for more details we direct interested readers to [19]. The constraints in Problem 1, including the final one on the irreducibility of $P^{(i)}$, guarantee that the conditions of Theorem 3.4 are satisfied. In practice, it is hard to enforce the irreducibility constraint during each step of an iterative optimization algorithm; our approach is to relax the constraint and verify a posteriori that the iteratively computed solution satisfies the irreducibility constraint. In all the computational settings we considered, we never encountered a solution that violated the irreducibility constraint.

In order to build intuition on the Markov chain combinations that generate optimal group hitting time values, we present numerical results for the ring graphs, complete graph, and lattice graph shown in Figure 4.1. We look at two cases in particular: one in which every random walker is required to visit all nodes in the graph, and one in which random walkers are allowed to visit subgraphs. For simplicity, we always restrict the stationary distribution $\tilde{\boldsymbol{\pi}}_{\text{ave}}$ to be uniform.

4.1. Random walkers covering the full graph. In the following we study which Markov chains generate optimal group hitting time values when every random walker must visit every node in the graph. Surprisingly, in fact, we will observe the random walks that generate optimal group hitting times can be different. In each example we define an individual agent's stationary distribution as the uniform

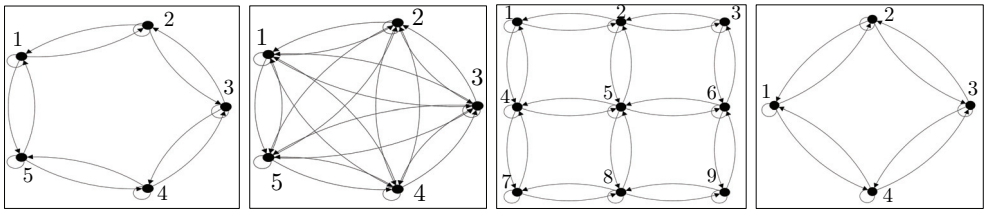

FIG. 4.1. *From left to right, example of a* 5*-node ring,* 5*-node complete,* 9*-node lattice, and* 4*-node ring graphs with self-loops.*

distribution. It is easily verified that with this choice of stationary distribution, the condition $\tilde{\boldsymbol{\pi}}_{\text{ave},j} = \tilde{\boldsymbol{\pi}}_{\text{ave},k}$ for all j, k is always met no matter how many agents are added to the system.

We begin with the ring graph. For a single random walker, the transition matrix which generates the minimal hitting time is simply the one describing a cycle (i.e., moving to a neighboring node with probability 1), and for the 5-node ring graph shown in Figure 4.1. This is as expected, as a cycle describes the fastest time to reach any node from any other node. It turns out, for the multiagent case, the optimal group hitting time occurs when every agent performs its own cycle; an example of this for three random walkers is shown in Figure 4.2. Moreover, since the group hitting time averages over all potential initial conditions, the direction of cycles does not matter. In other words, one agent can go clockwise, while the other goes counterclockwise. A summary of the optimal group and individual random walker hitting times on a ring graph shown in Figure 4.1 is given in Table 4.1.

Given the results for a ring graph, one can imply what will happen for the complete graph for a single agent. As it happens, since a cycle exists, this is also the optimal strategy for the complete graph for both the single and multiple random walker cases. Again, the direction of cycle does not matter, as is seen for the 3-agent case shown in Figure 4.2. A summary of the optimal group and individual random walker hitting times on a complete graph shown in Figure 4.1 is equivalent to the ring graph and thus is given in Table 4.1.

Next, we look at the lattice graph. Figure 4.3 shows the optimal trajectories found ranging from a single random walker up to 3 random walkers. It is interesting to note that individual agent trajectories are quite different, unlike for the ring and complete graphs. This can be more easily seen by observing each agent's individual hitting time as shown in Table 4.2. What is surprising is that the transition matrices that generate the optimal group hitting time for the multiple random walker cases are in fact suboptimal individually. Interestingly enough, if one substitutes the optimal single agent transition matrix from the single random walker case in for any/all of the multiwalker transition matrices, the group hitting time becomes worse for those multiwalker cases.

Thus far we have seen that repeating multiple copies of the fastest random walk is not always the most optimal. In fact, through simulation we've seen that repeating random walks is only optimal when a cycle is the optimal single-agent strategy. In the following section we explore in more detail how the optimal group hitting time is affected when working with subgraphs.

4.2. Random walkers covering subgraphs. In the following section we build intuition on how the group hitting time is affected over subgraphs. We will observe

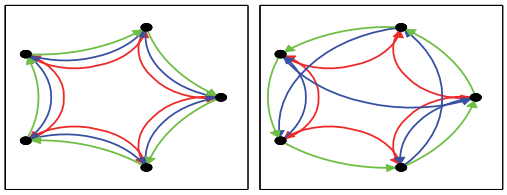

FIG. 4.2. *Probability to move along each edge of a ring graph (left) and complete graph (right) for* 3 *random walkers. In the case above, the probability to move along each edge is* 1*, which indicates a cycle. The group hitting time for the shown trajectories for both ring and complete graph is* $H_3 = 1.8$.

TABLE 4.1

Group hitting time values for random walks shown in Figure 4.3*. The last column indicates the group hitting time for each case, whereas the middle three columns indicate each random walker's individual hitting time. Surprisingly, the ring and complete graphs exhibit equivalent hitting time results. The values reported are exact.*

Random walker(s)	**Red**	**Blue**	**Green**	H_N
One	3.0	–	–	3.0
Two	3.0	3.0	–	2.2
Three	3.0	3.0	3.0	1.8

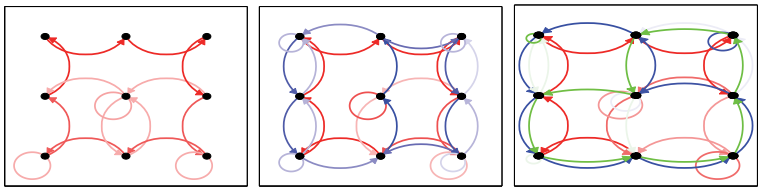

FIG. 4.3. *Probability to move along each edge of a lattice graph for* 1 *random walker (left),* 2 *random walkers (middle), and* 3 *random walkers (right). In each graph, the opacity of a line indicates the probability to move along an edge.*

TABLE 4.2

Group hitting time values for random walks shown in Figure 4.3*. The last column indicates the group hitting time for each case, whereas the middle three columns indicate each random walker's individual hitting time. The values reported are rounded to the second/third significant figures.*

Random walker(s)	**Red**	**Blue**	**Green**	H_N
One	6.8	–	–	6.8
Two	7.7	10.5	–	3.8
Three	15.2	7.0	17.2	2.9

that working with subgraphs sometimes improves the group hitting time, but can also cause the group hitting time to worsen. We look at two cases in particular: when the subgraphs overlap, and when they do not (i.e., the nodes are partitioned among the random walkers). In order to describe the different subgraphs studied, we utilize Figure 4.1 to specify which edges are kept and which are removed for the subgraphs. For example, if we specify that in the 5-node ring an agent can move only between nodes $\{2, 3, 4\}$, then this means that the agent can move only between these nodes (including self-loops), and any edges connected to nodes 1 and 5 are removed. The subgraphs are not necessarily allocated in any optimal way; they are simply chosen so that the stationary distribution is uniform over all nodes. For the simple

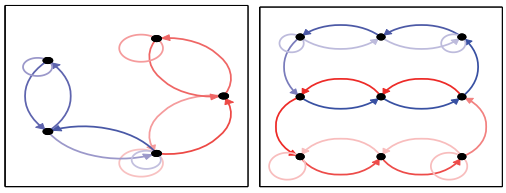

FIG. 4.4. *Probability to move along each edge of a* 5*-node ring graph with two agents (left), and* 9*-node lattice graph with two agents (right). In each graph, the opacity of a line indicates the probability to move along an edge.*

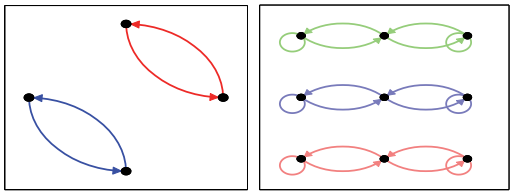

FIG. 4.5. *Probability to move along each edge of a* 4*-node ring graph with two agents (left), and* 9*-node lattice graph with three agents (right). In each graph, the opacity of a line indicates the probability to move along an edge.*

examples shown, the stationary distributions for each random walker are defined as $\boldsymbol{\pi}_j^{(i)} = \tilde{\boldsymbol{\pi}}_{\text{ave},j}/N_j$, where N_j denotes the number of agents who share node j. For example, $\boldsymbol{\pi}_j^{(i)} = \tilde{\boldsymbol{\pi}}_{\text{ave},j}$ if node j is only owned by one agent. For comparison with results presented in the previous section, we work with the ring and lattice graphs shown in Figure 4.1.

The case when subgraphs overlap is studied for the 5-node ring and 9-node lattice graph, the results of which can be seen in Figure 4.4. For the ring graph shown, one agent is restricted to moving between nodes $\{1, 5, 4\}$ and one is restricted to moving between nodes $\{2, 3, 5\}$, and the optimal group hitting time is $H_2 = 2.5$ (the exact value is 2.4655), in contrast to $H_2 = 2.2$ from full graph case (Table 4.1). For the lattice graph, one agent is restricted to moving between nodes $\{1, 2, 3, 4, 5, 6\}$ and one is restricted to moving between nodes $\{4, 5, 6, 7, 8, 9\}$, and the group hitting time is $H_2 = 3.6$ (the exact value is 3.6291), in contrast to $H_2 = 3.8$ (Table 4.2). Therefore, from these two examples, we see that each agent covering fewer nodes is not always indicative of a lower group hitting time.

The case when subgraphs are partitioned is analyzed for the 9-node lattice and 4-node ring graph shown in Figure 4.5. For the ring graph, with one agent restricted to moving between nodes $\{1, 5\}$ and one restricted to moving between nodes $\{2, 4\}$, we see that the group hitting time is $H_2 = 1.5$, whereas for the case where each agent covers the whole graph, $H_2 = 1.9$; we do not show the figure for the latter case, but recall that the optimal full graph trajectory is simply a cyclic tour for each agent. Now, for the partitioned lattice graph, with one agent restricted to moving between nodes $\{1, 2, 3\}$, one agent restricted to nodes $\{4, 5, 6\}$, and one agent restricted to nodes $\{7, 8, 9\}$, the group hitting time is $H_3 = 3.7$ (the exact value is $11/3$), in contrast to $H_3 = 2.9$ (Table 4.2). As before, we see that each agent covering fewer nodes, which are partitioned, is not indicative of a lower group hitting time.

Clearly, the small sample study presented here leaves open many future avenues that can be explored when attempting to optimize the group hitting time. For ex-

ample, notice that one cannot always partition a graph and achieve an arbitrary $\tilde{\boldsymbol{\pi}}_{\text{ave}}$. Also, it is unclear what happens when you allow $\tilde{\boldsymbol{\pi}}_{\text{ave}}$ and therefore individual stationary distributions to vary. We leave these, among other questions, to future work.

4.3. Implementation notes. In order to generate repeatable and accurate results, we utilize a *sequential quadratic programming* (SQP) solver as implemented in MATLAB's `fmincon` optimization algorithm. Several other solvers are available; however, the SQP solver has the most desirable mathematical properties. More specifically, with the SQP solver, given that the maximum number of iterations is not reached, the minimization algorithm stops when the first-order necessary Karush–Kuhn–Tucker conditions are approximately satisfied; conditions are satisfied in the norm sense with a tolerance of 10^{-6}. For all results used, the first-order optimality stopping criteria were satisfied.

The group hitting time results presented were taken as the minimum of 1000 optimization runs, each starting from a random initial conditions. Surprisingly, with the exception of the lattice, every initial condition converged to the minimum group hitting time value (within a 10^{-6} to 10^{-11} tolerance). For the 2- and 3-agent results presented in Table 4.2, we found that solutions converged to the minimum group hitting time value within a 10^{-2} tolerance for 96 and 93 percent of samples, respectively. Therefore, we claim with reasonable confidence that a local minimum for each group hitting time value was reached, if not a global minimum. On a desktop computer with an Intel i7-4790 processor and 8 GB of RAM running MATLAB 2014b, the lattice with 9 nodes and 3 agents shown in Figure 4.2 took the longest time to run, averaging 10 minutes per run, whereas all other simulations took less than a minute to fractions of a second per case run.

5. Conclusions. We have studied the hitting time of multiple random walkers on a graph and have presented the first formulation of this quantity, which we have denoted the *group hitting time*. Moreover, we have presented an alternate closed form solution for calculating the first hitting time between any specified set of nodes for a single random walker, as well as the first closed form solution of this quantity for the multiagent case. Finally, we posed the group hitting time as an optimization problem and provided detailed simulation results which help to build insight into the transition matrices that minimize this quantity.

This work leaves multiple directions open for further research. First, although our method provides a closed-form solution, in general it can be difficult to compute when the size of the graph and number of agents increase. It would be of practical interest to find a formulation which is computationally less expensive or a method in which the group hitting time can be calculated in a distributed way. Second, we provided results for multiple Markov chains in which travel times are homogeneous. A clear extension is to consider the case of heterogeneous travel times similar to what was done in [32]. Finally, given the maximum *pairwise* hitting time of a Markov chain, there exist bounds on the cover time for multiple copies of that Markov chain running in parallel [11, 2]. It would be interesting to see if our results can be leveraged to extend those bounds to multiple heterogeneous Markov chains running in parallel.

REFERENCES

[1] 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.

[2] N. Alon, C. Avin, M. Koucky, G. Kozma, Z. Lotker, and M. R. Tuttle, *Many random walks are faster than one*, Combin. Probab. Comput., 20 (2011), pp. 481–502.
[3] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, *The role of social networks in information diffusion*, in Proceedings of the 21st International Conference on World Wide Web, Lyon, France, 2012, pp. 519–528.
[4] T. H. Blackwell and J. S. Kaufman, *Response time effectiveness: Comparison of response time and survival in an urban emergency medical services system*, Academic Emergency Med., 9 (2002), pp. 288–295.
[5] A. Z. Broder and A. R. Karlin, *Bounds on the cover time*, J. Theoret. Probab., 2 (1989), pp. 101–120.
[6] P. Burgisser, M. Clausen, and A. Shokrollahi, *Algebraic Complexity Theory*, Springer, New York, 1997.
[7] M. Catral, S. J. Kirkland, M. Neumann, and N.-S. Sze, *The Kemeny constant for finite homogeneous ergodic Markov chains*, J. Sci. Comput., 45 (2010), pp. 151–166.
[8] C. Cooper, A. Frieze, and T. Radzik, *Multiple random walks and interacting particle systems*, in Automata, Languages and Programming, S. Albers, A. Marchetti-Spaccamela, Y. Matias, S. Nikoletseas, and W. Thomas, eds., Lecture Notes in Comput. Sci. 5556, Springer, New York, 2009, pp. 399–410.
[9] C. Cooper, A. Frieze, and T. Radzik, *Multiple random walks in random regular graphs*, SIAM J. Discrete Math., 23 (2009), pp. 1738–1761, doi:10.1137/080729542.
[10] P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks*, Mathematical Association of America, Washington, D.C., 1984.
[11] K. Efremenko and O. Reingold, *How well do random walks parallelize?*, in Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, I. Dinur, K. Jansen, J. Naor, and J. Rolim, eds., Lecture Notes in Comput. Sci. 5687, Springer, New York, 2009, pp. 476–489.
[12] Y. Elmaliach, A. Shiloni, and G. A. Kaminka, *A realistic model of frequency-based multi-robot polyline patrolling*, in Proceedings of the International Conference on Autonomous Agents, Estoril, Portugal, 2008, pp. 63–70.
[13] R. Elsässer and T. Sauerwald, *Tight bounds for the cover time of multiple random walks*, Theoret. Comput. Sci., 412 (2011), pp. 2623–2641.
[14] U. Feige, *A tight upper bound on the cover time for random walks on graphs*, Random Structures Algorithms, 6 (1995), pp. 51–54.
[15] C. Gkantsidis, M. Mihail, and A. Saberi, *Random walks in peer-to-peer networks*, in Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies, 2004.
[16] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1994.
[17] J. J. Hunter, *The role of Kemeny's constant in properties of Markov chains*, Comm. Statist. Theory Methods, 43 (2014), pp. 1309–1321.
[18] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, Springer, New York, 1976.
[19] S. Kirkland, *Fastest expected time to mixing for a Markov chain on a directed graph*, Linear Algebra Appl., 433 (2010), pp. 1988–1996.
[20] S. Kirkland, *On the Kemeny constant and stationary distribution vector for a Markov chain*, Electron. J. Linear Algebra, 27 (2014), pp. 354–372.
[21] J. Leskovec, D. Chakrabarti, J. Kleinberg, and C. Faloutsos, *Realistic, mathematically tractable graph generation and evolution, using Kronecker multiplication*, in Knowledge Discovery in Databases (PKDD), Springer, New York, 2005, pp. 133–145.
[22] J. Leskovec and C. Faloutsos, *Scalable modeling of real graphs using Kronecker multiplication*, in Proceedings of the 24th International Conference on Machine Learning, ACM, New York, 2007, pp. 497–504.
[23] M. Levene and G. Loizou, *Kemeny's constant and the random surfer*, Amer. Math. Monthly, 109 (2002), pp. 741–745.
[24] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*, AMS, Providence, RI, 2009.
[25] M. Mahdian and Y. Xu, *Stochastic Kronecker graphs*, Random Structures Algorithms, 38 (2011), pp. 453–466.
[26] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.
[27] Y. Nonaka, H. Ono, K. Sadakane, and M. Yamashita, *The hitting and cover times of Metropolis walks*, Theoret. Comput. Sci., 411 (2010), pp. 1889–1894.
[28] J. L. Palacios, *On hitting times of random walks on trees*, Statist. Probab. Lett., 79 (2009), pp. 234–236.

[29] J. L. Palacios and P. Tetali, *A note on expected hitting times for birth and death chains*, Statist. Probab. Lett., 30 (1996), pp. 119–125.
[30] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, *Review of road traffic control strategies*, Proc. IEEE, 91 (2003), pp. 2043–2067.
[31] F. Pasqualetti, A. Franchi, and F. Bullo, *On cooperative patrolling: Optimal trajectories, complexity analysis and approximation algorithms*, IEEE Trans. Robotics, 28 (2012), pp. 592–606, doi:10.1109/TRO.2011.2179580.
[32] R. Patel, P. Agharkar, and F. Bullo, *Robotic surveillance and Markov chains with minimal weighted Kemeny constant*, IEEE Trans. Automat. Control, 60 (2015), pp. 3156–3157, doi:10.1109/TAC.2015.2426317.
[33] E. Seneta, *Non-negative Matrices and Markov Chains*, 2nd ed., Springer, New York, 1981.
[34] S. D. Servetto and G. Barrenechea, *Constrained random walks on random graphs: Routing algorithms for large scale wireless sensor networks*, in Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications, ACM, New York, 2002, pp. 12–21.
[35] C. Seshadhri, A. Pinar, and T. G. Kolda, *An in-depth analysis of stochastic Kronecker graphs*, J. ACM, 60 (2013), pp. 13:1–13:32.
[36] M. S. Squillante, D. Towsley, and S. Barker, *Improving the scalability of search in networks through multiple random walks*, ACM SIGMETRICS Performance Evaluation Rev., 42 (2014), pp. 73–75.
[37] V. Srivastava, F. Pasqualetti, and F. Bullo, *Stochastic surveillance strategies for spatial quickest detection*, Internat. J. Robotics Res., 32 (2013), pp. 1438–1458, doi:10.1177/0278364913490322.
[38] S. Susca, S. Martínez, and F. Bullo, *Monitoring environmental boundaries with a robotic sensor network*, IEEE Trans. Control Syst. Technol., 16 (2008), pp. 288–296, doi:10.1109/TCST.2007.903395.
[39] P. Tetali, *Random walks and the effective resistance of networks*, J. Theoret. Probab., 4 (1991), pp. 101–109.
[40] P. M. Weichsel, *The Kronecker product of graphs*, Proc. Amer. Math. Soc., 13 (1962), pp. 47–52.
[41] X. Wu and Z. Liu, *How community structure influences epidemic spread in social networks*, Phys. A, 387 (2008), pp. 623–630.