IEICE TRANSACTIONS

on Communications

VOL. E104-B NO. 6
JUNE 2021



A PUBLICATION OF THE COMMUNICATIONS SOCIETY

The Institute of Electronics, Information and Communication Engineers
Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3chome, Minato-ku, TOKYO, 105-0011 JAPAN

IEICE TRANS. COMMUN., VOL.E104–B, NO.6 JUNE 2021

PAPER

# Graph Degree Heterogeneity Facilitates Random Walker Meetings*

**Yusuke SAKUMOTO**[†a)] ***and*** **Hiroyuki OHSAKI**[†b)]**,** ***Members***

**SUMMARY** Various graph algorithms have been developed with multiple random walks, the movement of several independent random walkers on a graph. Designing an efficient graph algorithm based on multiple random walks requires investigating multiple random walks theoretically to attain a deep understanding of their characteristics. The first meeting time is one of the important metrics for multiple random walks. The first meeting time on a graph is defined by the time it takes for multiple random walkers to meet at the same node in a graph. This time is closely related to the rendezvous problem, a fundamental problem in computer science. The first meeting time of multiple random walks has been analyzed previously, but many of these analyses focused on regular graphs. In this paper, we analyze the first meeting time of multiple random walks in arbitrary graphs and clarify the effects of graph structures on expected values. First, we derive the spectral formula of the expected first meeting time on the basis of spectral graph theory. Then, we examine the principal component of the expected first meeting time using the derived spectral formula. The clarified principal component reveals that (a) the expected first meeting time is almost dominated by $n/(1 + d_{\mathrm{std}}^2/d_{\mathrm{avg}}^2)$ and (b) the expected first meeting time is independent of the starting nodes of random walkers, where $n$ is the number of nodes of the graph. $d_{\mathrm{avg}}$ and $d_{\mathrm{std}}$ are the average and the standard deviation of weighted node degrees, respectively. Characteristic (a) is useful for understanding the effect of the graph structure on the first meeting time. According to the revealed effect of graph structures, the variance of the coefficient $d_{\mathrm{std}}/d_{\mathrm{avg}}$ (degree heterogeneity) for weighted degrees facilitates the meeting of random walkers.

***key words:*** *first meeting time, random walk, spectral graph theory*

## 1. Introduction

Various graph algorithms have been developed with multiple random walks, the movement of several independent random walkers on a graph, as a result of graph algorithms offer ease of analysis and light-weight processing. Notable applications include (a) a search algorithm for finding a particular node on a graph [1], [2], (b) an algorithm for spreading information across graphs by exchanging information only between adjacent nodes [3], and (c) the rendezvous algorithm for efficient meeting of multiple random walkers at the same node [4]. Designing an efficient graph algorithm based on multiple random walks requires studying multiple random walks theoretically in order to understand their characteristics at a deep level.

Manuscript received June 16, 2020.
Manuscript revised November 8, 2020.
Manuscript publicized December 14, 2020.

[†]The authors are with Kwansei Gakuin University, Sanda-shi, 669-1337 Japan.

*The preprint of this paper is available at https://arxiv.org/abs/2005.11161.

a) E-mail: sakumoto@kwansei.ac.jp
b) E-mail: ohsaki@kwansei.ac.jp

DOI: 10.1587/transcom.2020EBP3093

Several important metrics (e.g., first hitting time, recurrence time, cover time, re-encountering time, and first meeting time) have been used for investigating the characteristics of multiple random walks. The first hitting time is the time it takes for any random walker to arrive at a specified node, and it is important for evaluating the performance of relevant search algorithms. The recurrence time is the time required to return any one of the random walkers to the starting node, and it is thus a particular case of the first hitting time. The cover time is the time it takes for any random walker to reach all of the nodes and corresponds to the maximum value of the first hitting times. The cover time strongly affects the information dissemination speed in the graph. The re-encountering time and the first meeting time are the times it takes for multiple random walkers to meet at the same node. The re-encountering time relates to random walkers starting from the same node, and the first meeting time relates to those starting from different nodes. In particular, the first meeting time is closely related to the rendezvous problem, a fundamental problem in computer science. The rendezvous problem occurs in a number of engineering problems (e.g., the self-stabilizing token management system problem [5], [6] and the $k$-server problem [7]). Designing efficient algorithms for the rendezvous problem requires clarification of the characteristics of the first meeting time.

The first meeting time of multiple random walks was analyzed in [8]–[12]. However, many of these previous studies focus on regular graphs. In [12], George et al. did pioneering work on multiple random walks on non-regular graphs and derived a closed-form formula for calculating the expected value of the first meeting time in arbitrary graphs. However, the effects of graph structures on the expected first meeting time remain unclear. Designing effective algorithms using multiple random walks for realistic graphs (e.g., social networks and communication networks) benefits from understanding the effects of graph structures on the expected first meeting time. Since it is difficult to clarify these effects numerically using the closed-form formula derived in [12], the effects must be examined using analysis of multiple random walks.

In this paper, we analyze the first meeting time of multiple random walks in arbitrary graphs and clarify the effects of graph structures on its expected value. First, we derive the spectral formula of the expected first meeting time on the basis of spectral graph theory that is used to analyze the characteristics of graphs. Then, we examine the prin-

cipal component of the expected first meeting time using the derived spectral formula. The clarified principal component reveals that (a) the expected first meeting time is almost dominated by $n/(1 + d_{\rm std}^2/d_{\rm avg}^2)$, where $n$ is the number of nodes in the graph and $d_{\rm avg}$ and $d_{\rm std}$ are the average and standard deviation of weighted node degrees, respectively, and (b) the expected first meeting time is independent of the starting nodes of random walkers. Characteristic (a) provides understanding of the effect of the graph structure on the first meeting time. In addition, we verify the validity of the analysis results through numerical examples using popular random graphs.

The contributions in this paper are summarized as follows.

- We extend the analysis of a single random walks to multiple random walks using spectral graph theory.
- We derive the spectral formula (40) of the expected first meeting time.
- We clarify the principal component of the expected first meeting time from Eqs. (62) and (63).
- We reveal the effect of graph structures on the expected first meeting time from Eq. (63).
- We confirm the validity of the derived spectral formula (40) and the clarified principal component for popular random graphs with different scales and different structures.

The remainder of this paper is organized as follows. In Sect. 2, we describe the definition of graphs and random walks and introduce the previous analysis of a single random walk using spectral graph theory. In Sect. 3, we derive the spectral formula of the expected first meeting time using spectral graph theory and clarify the principal component of the expected first meeting time on the basis of the derived spectral formula. Section 4 confirms the validity of the analysis results through numerical examples using popular random graphs. Section 5 concludes the paper and discusses future work.

## 2. Preliminary

In this section, we provide the definition of graphs and random walks that we use in our analysis. In addition, we review existing analysis results of a single random walk based on spectral graph theory.

A graph is given by $G = (V, E)$, where $V$ and $E$ are a set of nodes and a set of links, respectively. Self-loop links $(i, i)$ for $i \in V$ are not included in $E$. The weight for link $(i, j) \in E$ is $w_{ij}$, where $w_{ij} > 0$ and $w_{ij} = w_{ji}$. We denote the set of adjacent nodes of node $i \in V$ by $\partial i$. Letting $d_i$ be the weighted degree of node $i \in V$, we define $d_i$ as

$$d_i := \sum_{k \in \partial i} w_{ik}. \tag{1}$$

We describe a random walk starting from node $a \in V$. In this random walk, the random walker at node $i \in V$ moves to adjacent node $j \in \partial i$ using transition probability $p_{i\to j}$ given by

$$p_{i\to j} = \frac{w_{ij}}{d_i}. \tag{2}$$

Let $x_{a:i}(t)$ be the probability that a random walker starting from node $a \in V$ is at node $i \in V$ at time $t$, where $\sum_{i\in V} x_{a:i}(t) = 1$. If Eq. (2) is used, then $x_{a:i}(t + 1)$ is

$$x_{a:i}(t + 1) = \sum_{j\in\partial i} x_{a:j}(t)\, p_{j\to i}. \tag{3}$$

Using column vector $\boldsymbol{x}_a(t) = (x_{a:i}(t))_{i\in V}$, Eq. (3) for all nodes $\forall i \in V$ can be written simultaneously as

$$\boldsymbol{x}_a(t + 1) = \boldsymbol{A}\,\boldsymbol{D}^{-1}\,\boldsymbol{x}_a(t), \tag{4}$$

where $\boldsymbol{D}$ and $\boldsymbol{A}$ are the degree and adjacency matrices defined as

$$\boldsymbol{D} := \mathrm{diag}(d_i)_{i\in V}, \tag{5}$$

$$\boldsymbol{A} := \begin{cases} w_{ij} & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

respectively. $\boldsymbol{A}\,\boldsymbol{D}^{-1}$ is the matrix whose $(i, j)$th element is the transition probability $p_{j\to i}$. Random walks defined by Eqs. (3) through (6) are equivalent to those used in [8]–[12], and are applied to search in unstructured P2P networks [1], [2]. Hence, it is worth analyzing these random walks in term of engineering applications. Equation (4) describes the behavior of the random walk. Since $\boldsymbol{A}\,\boldsymbol{D}^{-1}$ is an asymmetric matrix, Eq. (4) is not easy to handle analytically using linear algebra. Consequently, we modify Eq. (4) to

$$\begin{aligned} \boldsymbol{D}^{-1/2}\,\boldsymbol{x}_a(t + 1) &= \boldsymbol{D}^{-1/2}\,\boldsymbol{A}\,\boldsymbol{D}^{-1/2}\,\boldsymbol{D}^{-1/2}\,\boldsymbol{x}_a(t) \\ &= \boldsymbol{W}\,\boldsymbol{D}^{-1/2}\,\boldsymbol{x}_a(t) \\ \hat{\boldsymbol{x}}_a(t + 1) &= \boldsymbol{W}\,\hat{\boldsymbol{x}}_a(t), \end{aligned} \tag{7}$$

where $\boldsymbol{W} = \boldsymbol{D}^{-1/2}\,\boldsymbol{A}\,\boldsymbol{D}^{-1/2}$ and $\hat{\boldsymbol{x}}_{\boldsymbol{a}}(t) = \boldsymbol{D}^{-1/2}\,\boldsymbol{x}_a(t)$. Since $\boldsymbol{W}$ is a symmetric matrix, Eq. (7) is easier to handle than Eq. (4). In general, in spectral graph theory, the behavior of an analysis target is expressed in terms of a matrix such as Eq. (7), and the characteristics of the target are analyzed using the eigenvalues and eigenvectors of the matrix on the basis of linear algebra.

$\boldsymbol{W}$ can always be diagonalized using the orthogonal matrix $\boldsymbol{Q}$ that satisfies ${}^t\boldsymbol{Q} = \boldsymbol{Q}^{-1}$. Let $\lambda_k$ be the $k$th largest eigenvalue of $\boldsymbol{W}$. Note that the maximum eigenvalue $\lambda_1$ is always 1. In this paper, we assume that $G$ is connected and not a bipartite graph. In this case, the eigenvalues $\lambda_k$ for $2 \le k \le n$ satisfy

$$-1 < \lambda_n < \cdots < \lambda_2 < 1. \tag{8}$$

We let $\boldsymbol{q}_k$ be the eigenvector for eigenvalue $\lambda_k$ by, with the consequence that $\boldsymbol{Q}$ is $\boldsymbol{Q} = (\boldsymbol{q}_k)_{1\le k\le n}$. Since $\boldsymbol{Q}$ is an orthogonal matrix, $\boldsymbol{q}_k$ and $\boldsymbol{q}_l$ for $1 \le k, l \le n$ satisfy

$${}^t\boldsymbol{q}_k\,\boldsymbol{q}_l = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

In particular, the maximum eigenvector $\boldsymbol{q}_1$ is

$$\boldsymbol{q}_1 = \frac{1}{\sqrt{s_1}}{}^t(\sqrt{d_1}, \sqrt{d_2}, \cdots, \sqrt{d_n}), \tag{10}$$

where $s_1 = \sum_{i \in V} d_i$. $s_1$ is related to a statistic of the graph structure of $G$ and can be written by $s_1 = n\, d_{\text{avg}}$, where $d_{\text{avg}}$ is the average weighted degree.

In [13], Lovász analyzed a single random walk on graph $G$ on the basis of spectral graph theory. Solving Eq. (7) in [13] results in the probability $x_{a:i}(t)$ being

$$x_{a:i}(t) = \frac{\sqrt{d_i}}{\sqrt{d_a}} \sum_{k=1}^{n} q_k(a)\, q_k(i)\, \lambda_k^t. \tag{11}$$

According to this equation, $x_{a:i}(t)$ can be calculated using eigenvalues $\lambda_k$ and eigenvectors $\boldsymbol{q}_k$. A closed-form formula using eigenvalues and eigenvectors such as Eq. (11) is referred to as a *spectral formula*.

Let $x^*_{a:i}$ be the limit value of $x_{a:i}(t)$ for $t \to \infty$. From Eq. (11), $x^*_{a:i}$ can be derived as

$$\begin{aligned} x^*_{a:i} &= \lim_{t\to\infty} x_{a:i}(t) \\ &= \lim_{t\to\infty} \left[ \frac{\sqrt{d_i}}{\sqrt{d_a}} q_1(i)\, q_1(a) + \frac{\sqrt{d_i}}{\sqrt{d_a}} \sum_{k=2}^{n} q_k(a)\, q_k(i)\, \lambda_k^t \right] \\ &= \frac{\sqrt{d_i}}{\sqrt{d_a}} q_1(a)\, q_1(i) = \frac{d_i}{s_1}. \end{aligned} \tag{12}$$

In this derivation process, we used $|\lambda_k| < 1$ for $k \geq 2$. According to Eq. (12), $x_{a:i}(t)$ is roughly proportional to the weighted degree $d_i$ if sufficient time has elapsed since the random walker started.

The analysis in [13] derived the expected first hitting time $\mu_{a:i}$, the expected time it takes for a random walker starting from node $a$ to arrive at node $i$. From Eq. (11), the spectral formula of expected first hitting time $\mu_{a:i}$ is derived as

$$\mu_{a:i} = s_1 \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left( \frac{q_k^2(i)}{d_i} - \frac{q_k(a)\, q_k(i)}{\sqrt{d_a\, d_i}} \right). \tag{13}$$

In [14], the effect of the graph structure on $\mu_{a:i}$ was clarified using the spectral formula of the expected first hitting time $\mu_{a:i}$. According to [14], $\mu_{a:i}$ satisfies

$$\left| \frac{1}{s_1}\mu_{a:i} - \frac{1}{d_i} \right| \leq \frac{2\, w_{\max}}{d_{\min}^2} \left( \frac{1}{1-\lambda_2} + 1 \right), \tag{14}$$

where $w_{\max}$ and $d_{\min}$ are the maximum of link weights and the minimum of weighted degrees, respectively. If the right-hand side of Eq. (14) is sufficiently small, the expected first hitting time $\mu_{a:i}$ is approximated by

$$\mu_{a:i} \approx \frac{s_1}{d_i}. \tag{15}$$

In this case, $\mu_{a:i}$ is almost dominated by $s_1/d_i$, with the consequence that $s_1/d_i$ can be expected to be the principal component of $\mu_{a:i}$. According to Eq. (15), $\mu_{a:i}$ is roughly proportional to $s_1$, which is a statistic of the graph structure. In other words, $\mu_{a:i}$ corresponds to the search time of node $i$ using the random walk. Therefore, Eq. (15) is also important for understanding the characteristics of the search algorithm using a random walk.

According to [14], when $n$ is sufficiently large, the right-hand side of Eq. (14) is negligibly small in several graphs (e.g., $\epsilon$-graphs, k-nearest neighbor graphs, Gaussian similarity graphs, and Erdös-Rényi graphs). Considering the result of [14], we expect that the right-hand side of Eq. (14) is negligibly small in various graph classes.

## 3. Analysis

In this section, we analyze the expected first meeting time $\mu_{a,b}$ of two random walkers starting from node $a \in V$ and $b \in V$ in graph $G$ on the basis of spectral graph theory. We first derive the spectral formula of $\mu_{a,b}$. Then, we clarify the principal component of $\mu_{a,b}$ using the derived spectral formula. Finally, we reveal the effect of the graph structure on $\mu_{a,b}$ on the basis of the clarified principal component.

Our analysis results are important also for understanding the first meeting time $T_{\boldsymbol{s}}^{(n_{\text{RW}})}$ of $n_{\text{RW}}$ random walkers, where $n_{\text{RW}} > 2$ and $\boldsymbol{s}$ is the vector of starting nodes of $n_{\text{RW}}$ random walkers. This is because $T_{\boldsymbol{s}}^{(n_{\text{RW}})}$ is strongly affected by the first meeting time $T_{\boldsymbol{s}}^{(2)}$ of two random walkers. Note that the expectation of $T_{\boldsymbol{s}}^{(2)}$ with $\boldsymbol{s} = (a, b)$ is correspond to $\mu_{a,b}$. To attain an efficient meeting of $n_{\text{RW}}$ random walkers, two of the $n_{\text{RW}}$ random walkers must move together after meeting at the same node. Namely, when two random walkers meet at the same node, they are combined into a single random walker. Finally, if all of $n_{\text{RW}}$ random walkers become a single random walker, the first meeting of them are accomplished. In this case, the first meeting time $T_{\boldsymbol{s}}^{(n_{\text{RW}})}$ of $n_{\text{RW}}$ random walkers is obtained as the sum of the first meeting times $T_{\boldsymbol{s}}^{(2)}$ of two random walkers. Consequently, the characteristics of the first meeting time $T_{\boldsymbol{s}}^{(n_{\text{RW}})}$ for $n_{\text{RW}}$ random walkers can be expected to be strongly associated with that of two random walkers.

### 3.1 Spectral Formula of Expected First Meeting Time $\mu_{a,b}$

We derive the spectral formula of the expected first meeting time $\mu_{a,b}$ using the same method as is used to derive that of the expected first hitting time $\mu_{a:i}$ in [13]. In [13], the spectral formula was derived using the generating function of the existing probability $x_{a:i}(t)$. In general, the generating function $F(z)$ of the probability $f(t)$ is

$$F(z) := \sum_{t=0}^{\infty} f(t) z^t, \tag{16}$$

Note that $f(t)$ is the probability that an event occurs at time $t$. The existing probability $x_{a:i}(t)$ is an example of $f(t)$. If the probability $f(t)$ is satisfied with $\sum_{t=0}^{\infty} f(t) = 1$, using the generating function $F(z)$, the expectation $\mathrm{E}(t)$ with the probability $f(t)$ is

$$\mathrm{E}(t) = \sum_{t=1}^{\infty} t\, f(t) = \left.\frac{\mathrm{d}F(z)}{\mathrm{d}z}\right|_{z=1}. \tag{17}$$

Importantly, even if we do not know the closed-form formula of the probability $f(t)$, we can still derive the expectation $\mathrm{E}(t)$ using the generating function $F(z)$ on the basis of the above equation. We first discuss the generating function of the first meeting probability. Without the value of the first meeting probability, we then derive the spectral formula of the expected first meeting time $\mu_{a,b}$ by substituting the generating function obtained into Eq. (17).

Let $r_{a,b:c}(t)$ be the probability that two random walkers first meet at node $c$ at time $t$. Since two random walkers can meet at any node, the first meeting probability $r_{a,b:*}(t)$ is

$$r_{a,b:*}(t) = \sum_{c\in V} r_{a,b:c}(t). \tag{18}$$

In $r_{a,b:*}(t)$, the symbol $*$ designates any node in $V$. Deriving the spectral formula of the expected first meeting time $\mu_{a,b}$ using Eq. (17) requires the generating function $R_{a,b:*}(z)$ of $r_{a,b:*}(t)$.

The probabilities that the two random walkers are at node $i$ at time $t$ are $x_{a:i}(t)$ and $x_{b:i}(t)$, respectively. Since the spectral formulas of $x_{a:i}(t)$ and $x_{b:i}(t)$ are given by Eq. (11), the generating functions of $x_{a:i}(t)$ and $x_{b:i}(t)$ can be derived using Eq. (16). However, since it is not easy to obtain the spectral formula of $r_{a,b:*}(t)$, we obtain the generating function $R_{a,b:*}(z)$ of $r_{a,b:*}(t)$ from those of $x_{a:i}(t)$ and $x_{b:i}(t)$ and then derive the spectral formula of the expected first meeting time $\mu_{a,b}$ using Eq. (17).

With the aim of obtaining the generating function $R_{a,b:*}(z)$ of the first meeting probability $r_{a,b:*}(t)$, we discuss the relationship between $r_{a,b:*}(t)$, $x_{a:i}(t)$ and $x_{b:i}(t)$. Let $x_{a,b:c}(t)$ be the probability that the two random walker meet at the same node $c \in V$ at time $t$. The meeting probability $x_{a,b:*}(t)$ at any node is

$$x_{a,b:*}(t) = \sum_{c\in V} x_{a,b:c}(t). \tag{19}$$

Since each random walker moves independently, $x_{a,b:*}(t)$ is

$$x_{a,b:*}(t) = \sum_{c\in V} x_{a,b:c}(t) = \sum_{c\in V} x_{a:c}(t)\, x_{b:c}(t). \tag{20}$$

$x_{a,b:*}(t)$ includes both the first meeting probability $r_{a,b:*}(t)$ and also the probability of the second and subsequent meetings. Hence, as shown in Fig. 1, we divide the transition of the two random walks from time 0 to time $t$ into two transitions, (a) the transition until they first meet at time $s$, and (b) the rest transition. The probability for the former transition is the first meeting probability $r_{a,b:*}(t)$. The probability for the latter transition is the probability that the two random walkers starting from same node $c' \in V$ at time $s$ meet again at the same node $c \in V$ at time $t$. Since node $c'$ and node $c$ can be any node, we denote such a probability by $x_{*',*':*}(t-s)$. With these probabilities, $x_{a,b:*}(t)$ is

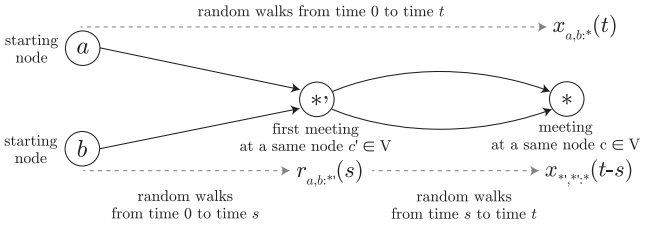


**Fig. 1** Random walks starting from nodes $a$ and $b$ from time 0 to time $t$.

$$x_{a,b:*}(t) = \sum_{s=0}^{t} r_{a,b:*'}(s)\, x_{*',*':*}(t-s). \tag{21}$$

Using the probability $x_{c',c':c}(t)$ that the two random walkers starting at node $c' \in V$ at time 0 meet again at node $c \in V$ at time $t$, we approximately set $x_{*',*':*}(t)$ as

$$x_{*',*':*}(t) \approx \sum_{c'\in V} \frac{d_{c'}^2}{s_2} \sum_{c\in V} x_{c',c':c}(t), \tag{22}$$

where $s_2 = \sum_{i\in V} d_i^2$. The reason that $x_{*',*':*}(t)$ is not set as a simple sum of values of $x_{c',c':c}(t)$ in Eq. (22) is as follows. According to Eq. (12), if $t$ is sufficient large, the existing probability $x_{a:i}(t)$ is approximately proportional to the weighted degree $d_i$ of node $i$. Therefore, the probability of the first meeting of the two random walkers at node $c'$ can be expected to be proportional to $d_{c'}^2$. Consequently, in the sum of Eq. (22), $x_{c',c':c}(t)$ is weighted by $d_{c'}^2/s_2$. In Sect. 4, the validity of Eq. (22) will be confirmed through numerical examples.

The right-hand side of Eq. (21) is a convolutional sum, with the consequence that the generating function $X_{a,b:*}(z)$ of $x_{a,b:*}(t)$ is

$$X_{a,b:*}(z) = R_{a,b:*'}(z)\, X_{*',*':*}(z). \tag{23}$$

From this equation, the generating function $R_{a,b:*'}(z)$ of the first meeting probability $r_{a,b:*}$ is

$$R_{a,b:*'}(z) = \frac{X_{a,b:*}(z)}{X_{*',*':*}(z)}. \tag{24}$$

Substituting the spectral formulas of $x_{a:i}(t)$ and $x_{b:i}(t)$ given by Eq. (11) into Eq. (16) yields the following spectral formula for $X_{a,b:*}(z)$:

$$\begin{aligned} X_{a,b:*}(z) &= \sum_{c\in V} \sum_{t=0}^{\infty} x_{a:c}(t)\, x_{b:c}(t)\, z^t \\ &= \sum_{c\in V} \frac{d_c}{\sqrt{d_a d_b}} \sum_{k=1}^{n}\sum_{k'=1}^{n} q_k(a) q_k(c) q_{k'}(b) q_{k'}(c) \sum_{t=0}^{\infty} (\lambda_k \lambda_{k'} z)^t \\ &= \sum_{c\in V} \frac{d_c}{\sqrt{d_a\, d_b}} \sum_{k=1}^{n}\sum_{k'=1}^{n} \frac{q_k(a)\, q_k(c)\, q_{k'}(b)\, q_{k'}(c)}{1-\lambda_k\, \lambda_{k'}\, z}. \end{aligned} \tag{25}$$

$z$ takes a value within the range of $|\lambda_k\, \lambda_{k'}\, z| < 1$ with the result that the sum of the infinite geometric series converges. From Eq. (22), the generating function $X_{*',*':*}(z)$ is

$$X_{*',*':*}(z) \approx \sum_{c' \in V} \frac{d_{c'}^2}{s_2} \sum_{c \in V} X_{c',c':c}(z). \tag{26}$$

Substituting the generating function $R_{a,b:*'}(z)$ into Eq. (17) results in the expected first meeting time $\mu_{a,b}$ being

$$\mu_{a,b} = \sum_{s=0}^{\infty} s\, r_{a,b:*'}(s) = \left. \frac{\mathrm{d}R_{a,b:*'}(z)}{\mathrm{d}z} \right|_{z=1}. \tag{27}$$

In this equation, $\mathrm{d}R_{a,b:*'}(z)/\mathrm{d}z$ is

$$\begin{aligned} \frac{\mathrm{d}R_{a,b:*'}(z)}{\mathrm{d}z} &= \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{X_{a,b:*}(z)}{X_{*',*':*}(z)}\right) \\ &= \frac{\dfrac{\mathrm{d}X_{a,b:*}(z)}{\mathrm{d}z} X_{*',*':*}(z) - X_{a,b:*}(z) \dfrac{\mathrm{d}X_{*',*':*}(z)}{\mathrm{d}z}}{X_{*',*':*}^2(z)} \\ &= \frac{A(z) - B(z)}{C(z)}, \end{aligned} \tag{28}$$

where

$$A(z) = \frac{\mathrm{d}X_{a,b:*}(z)}{\mathrm{d}z} X_{*',*':*}(z), \tag{29}$$

$$B(z) = X_{a,b:*}(z) \frac{\mathrm{d}X_{*',*':*}(z)}{\mathrm{d}z}, \tag{30}$$

$$C(z) = X_{*',*':*}^2(z). \tag{31}$$

According to Eqs. (25) and (26), $A(z)$, $B(z)$, and $C(z)$ can be written as polynomials for $(1-z)$ because of $1-\lambda_1\lambda_1 z = 1-z$. Hence, we also obtain $A(z)$ as

$$A(z) = \frac{A_3(z)}{(1-z)^3} + \frac{A_2(z)}{(1-z)^2} + \frac{A_1(z)}{1-z} + A_0(z). \tag{32}$$

Substituting Eqs. (25) and (26) into the right-hand side of Eq. (29) yields $A_3(z)$ and $A_2(z)$ as

$$\begin{aligned} A_3(z) = &\left[\sum_{c \in V} \frac{d_c}{\sqrt{d_a d_b}} q_1(a)\, q_1^2(c)\, q_1(b)\right] \\ &\left[\sum_{c \in V} \sum_{c' \in V} \frac{d_{c'}\, d_c}{s_2} q_1^2(c)\, q_1^2(c')\right], \end{aligned} \tag{33}$$

$$\begin{aligned} A_2(z) = &\left[\sum_{c \in V} \frac{d_c}{\sqrt{d_a d_b}} q_1(a)\, q_1^2(c)\, q_1(b)\right] \\ &\left[\sum_{c \in V} \sum_{c' \in V} \frac{d_{c'} d_c}{s_2} \sum_{\substack{1 \le k,k' \le n \\ (k,k') \ne (1,1)}} \frac{q_k(c) q_k(c') q_{k'}(c) q_{k'}(c')}{1 - \lambda_k \lambda_{k'} z}\right] \\ = &\frac{1}{s_1^2}\left[\sum_{c \in V} \sum_{c' \in V} d_{c'} d_c \sum_{\substack{1 \le k,k' \le n \\ (k,k') \ne (1,1)}} \frac{q_k(c) q_k(c') q_{k'}(c) q_{k'}(c')}{1 - \lambda_k \lambda_{k'} z}\right]. \end{aligned} \tag{34}$$

We do not provide $A_1(z)$ and $A_0(z)$ in this paper because $A_1(z)/(1-z)$ and $A_0(z)$ disappear when deriving $\mu_{a,b}$ using Eq. (27). Similarly, $B(z)$ is

$$B(z) = \frac{B_3(z)}{(1-z)^3} + \frac{B_2(z)}{(1-z)^2} + \frac{B_1(z)}{1-z} + B_0(z). \tag{35}$$

Substituting Eqs. (25) and (26) into the right-hand side of Eq. (30) yields $B_3(z)$ and $B_2(z)$ as

$$\begin{aligned} B_3(z) = &\left[\sum_{c \in V} \frac{d_c}{\sqrt{d_a d_b}} q_1(a)\, q_1^2(c)\, q_1(b)\right] \\ &\left[\sum_{c \in V} \sum_{c' \in V} \frac{d_{c'}\, d_c}{s_2} q_1^2(c)\, q_1^2(c')\right], \end{aligned} \tag{36}$$

$$\begin{aligned} B_2(z) = &\left[\sum_{c \in V} \sum_{\substack{1 \le k,k' \le n \\ (k,k') \ne (1,1)}} \frac{d_c}{\sqrt{d_a\, d_b}} \frac{q_k(a)\, q_k(c)\, q_{k'}(b)\, q_{k'}(c)}{1 - \lambda_k \lambda_{k'}\, z}\right] \\ &\left[\sum_{c \in V} \sum_{c' \in V} \frac{d_{c'}\, d_c}{s_2} q_1^2(c)\, q_1^2(c')\right] \\ = &\frac{1}{s_1^2}\left[\sum_{c \in V} \frac{d_c}{\sqrt{d_a\, d_b}} \sum_{1 \le k,k' \le 1 \atop (k,k') \ne (1,1)} \frac{q_k(a)\, q_k(c)\, q_{k'}(b)\, q_{k'}(c)}{1 - \lambda_k \lambda_{k'}\, z}\right]. \end{aligned} \tag{37}$$

Using these equations, we have found that $A_3(z) = B_3(z)$. Consequently, $A(z) - B(z)$ in the numerator of Eq. (28) does not contain the term $(1-z)^{-3}$, with the result that the term $(1-z)^{-2}$ becomes the highest-order term in the polynomials for $(1-z)$ in $A(z) - B(z)$. Thus, $C(z)$ is

$$C(z) = \frac{C_2(z)}{(1-z)^2} + \frac{C_1(z)}{1-z} + C_0(z). \tag{38}$$

Solving this equation in the same manner yields $C_2(z)$ as

$$\begin{aligned} C_2(z) &= \left[\sum_{c \in V} \sum_{c' \in V} \frac{d_{c'}\, d_c}{s_2} q_1^2(c)\, q_1^2(c')\right]^2 \\ &= \left[\frac{1}{s_1^2} \sum_{c \in V} d_c^2\right]^2 = \frac{(s_2)^2}{s_1^4}. \end{aligned} \tag{39}$$

Since $(A(z) - B(z))/C(z)$ is an indeterminate form at $z = 1$, we discuss $\lim_{z \to 1}(A(z) - B(z))/C(z)$ to derive the spectral formula of the expected first meeting time $\mu_{a,b}$ using Eqs. (27) and (28). As the limit of $z \to 1$, $\mu_{a,b}$ is

$$\begin{aligned} \mu_{a,b} &= \lim_{z \to 1} \frac{A(z) - B(z)}{C(z)} = \lim_{z \to 1} \frac{(1-z)^2 (A(z) - B(z))}{(1-z)^2\, C(z)} \\ &= \frac{A_2(1) - B_2(1)}{C_2(1)} \\ &= \frac{s_1^2}{(s_2)^2}\left[\sum_{c \in V} \sum_{c' \in V} d_{c'} d_c \sum_{\substack{1 \le k,k' \le n \\ (k,k') \ne (1,1)}} \frac{q_k(c) q_k(c') q_{k'}(c) q_{k'}(c')}{1 - \lambda_k \lambda_{k'}} \right. \\ &\qquad \left. - \sum_{c \in V} \frac{d_c}{\sqrt{d_a d_b}} \sum_{\substack{1 \le k,k' \le 1 \\ (k,k') \ne (1,1)}} \frac{q_k(a)\, q_k(c)\, q_{k'}(b)\, q_{k'}(c)}{1 - \lambda_k \lambda_{k'}}\right] \end{aligned}$$

$$= \frac{1}{(s_2)^2} \sum_{c\in V} d_c^2 \sum_{c'\in V} d_{c'}^2 \Bigg[ s_1 \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left[ \frac{2q_k(c)q_k(c')}{\sqrt{d_c d_{c'}}} - \frac{q_k(c)}{\sqrt{d_c}} \left( \frac{q_k(a)}{\sqrt{d_a}} + \frac{q_k(b)}{\sqrt{d_b}} \right) \right] + s_1^2 \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \frac{q_k(c)q_{k'}(c)}{(1-\lambda_k\lambda_{k'})d_c} \left( \frac{q_k(c')q_{k'}(c')}{d_{c'}} - \frac{q_k(a)q_{k'}(b)}{\sqrt{d_a\, d_b}} \right) \Bigg]. \tag{40}$$

Since this equation is expressed by the eigenvalues and eigenvectors of $\boldsymbol{W}$, it is the spectral formula of $\mu_{a,b}$.

Equation (40) appears to be complicated, but if we use the expected first meeting time $\mu_{a,b:c}$, the time until the two random walkers first meet at node $c \in V$, then $\mu_{a,b}$ is

$$\mu_{a,b} = \frac{1}{s_2} \sum_{c\in V} d_c^2\, \mu_{a,b:c} - \frac{1}{s_2^2} \sum_{c\in V} d_c^2 \sum_{c'\in V} d_{c'}^2\, \mu_{c',c':c}, \tag{41}$$

where the spectral formula of $\mu_{a,b:c}$ is

$$\mu_{a,b:c} = \mu_{a:c} + \mu_{b:c} + s_1^2 \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \frac{q_k(c)q_{k'}(c)}{(1-\lambda_k\lambda_{k'})d_c} \left( \frac{q_k(c)q_{k'}(c)}{d_c} - \frac{q_k(a)q_{k'}(b)}{\sqrt{d_a\, d_b}} \right). \tag{42}$$

Note that Eq. (41) is derived by substituting Eq. (42) into Eq. (40). Equation (42) can be obtained from the similar derivation process of Eq. (40). Owing to space limitations, we do not provide the detailed deviation process of Eq. (42) in this paper. In the derivation process of Eq. (42), we discuss the first meeting probability $r_{a,b:c}(t)$ of two random walkers at node $c \in V$, and obtain $\mu_{a,b:c}$ as the expectation with $r_{a,b:c}(t)$. As with the probability $r_{a,b:*}(t)$, the first meeting probability $r_{a,b:c}(t)$ is satisfied with

$$x_{a,b:c}(t) = \sum_{s=0}^{t} r_{a,b:c}(s)\, x_{c,c:c}(t-s). \tag{43}$$

Hence, the generating function $R_{a,b:c}(z)$ for $r_{a,b:c}(t)$ is given by

$$R_{a,b:c}(z) = \frac{X_{a,b:c}(z)}{X_{c,c:c}(z)}. \tag{44}$$

Since $\mu_{a,b:c} = \mathrm{d}R_{a,b:c}(z)/\mathrm{d}z|_{z=1}$, we can derive the spectral formula (42) of $\mu_{a,b:c}$ from $R_{a,b:c}(z)$ like the spectral formula (40) of $\mu_{a,b}$.

### 3.2 Principal Component of the Expected First Meeting Time $\mu_{a,b}$

We examine the principal component of $\mu_{a,b}$ with the spectral formula of the expected first meeting time $\mu_{a,b}$ and reveal mathematically the effect of the graph structure on the expected first meeting time $\mu_{a,b}$ on the basis of the clarified principal component. We use the method for examining the first hitting time $\mu_{a:i}$ used in [14] to derive the principal component of $\mu_{a,b}$.

First, we introduce

$$\hat{\boldsymbol{N}} := \boldsymbol{I} \otimes \boldsymbol{I} - \boldsymbol{W} \otimes \boldsymbol{W} = \hat{\boldsymbol{I}} - \hat{\boldsymbol{W}}, \tag{45}$$

where $\boldsymbol{I}$ is the $n \times n$ unit matrix and $\otimes$ is the Kronecker product. According to the definition of the Kronecker product, $\hat{\boldsymbol{I}}$ and $\hat{\boldsymbol{W}}$ are $n^2 \times n^2$ matrices. Let $\hat{\boldsymbol{N}}^\dagger$ be the pseudo-inverse matrix of $\hat{\boldsymbol{N}}$ with the result that $\hat{\boldsymbol{N}}\hat{\boldsymbol{N}}^\dagger\hat{\boldsymbol{N}} = \hat{\boldsymbol{N}}$,

$$\hat{\boldsymbol{N}}^\dagger = \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \frac{\hat{\boldsymbol{q}}_{kk'}\, {}^t\hat{\boldsymbol{q}}_{kk'}}{1-\lambda_k\lambda_{k'}}, \tag{46}$$

where $\hat{\boldsymbol{q}}_{kk'}$ is the following column vector with $n^2$ elements:

$$\hat{\boldsymbol{q}}_{kk'} := \boldsymbol{q}_k \otimes \boldsymbol{q}_{k'}. \tag{47}$$

Substituting $\hat{\boldsymbol{N}}^\dagger$ into Eq. (40) yields the following as the expected first meeting time $\mu_{a,b}$:

$$\mu_{a,b} = \frac{s_1^2}{s_2^2} \sum_{c\in V} \sum_{c'=1}^{n} d_c^2\, d_{c'}^2\, {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{N}}^\dagger (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}). \tag{48}$$

In this equation, $\hat{\boldsymbol{u}}_{ab}$ is

$$\hat{\boldsymbol{u}}_{ab} := \boldsymbol{u}_a \otimes \boldsymbol{u}_b, \tag{49}$$

where $\boldsymbol{u}_a$ is the column vector whose $i$th element $\boldsymbol{u}_a(i)$ is

$$\boldsymbol{u}_a(i) = \begin{cases} \dfrac{1}{\sqrt{d_a}} & \text{if} \quad i = a \\ 0 & \text{otherwise} \end{cases}. \tag{50}$$

The pseudo-inverse matrix $\hat{\boldsymbol{N}}^\dagger$ of $\hat{\boldsymbol{N}}$ in Eq. (46) is also

$$\begin{aligned} \hat{\boldsymbol{N}}^\dagger &= \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \frac{\hat{\boldsymbol{q}}_{kk'}\, {}^t\hat{\boldsymbol{q}}_{kk'}}{1-\lambda_k\lambda_{k'}} \\ &= \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \frac{(1-\lambda_k\lambda_{k'}+\lambda_k\lambda_{k'})\, \hat{\boldsymbol{q}}_{kk'}\, {}^t\hat{\boldsymbol{q}}_{kk'}}{1-\lambda_k\lambda_{k'}} \\ &= \hat{\boldsymbol{I}} - \hat{\boldsymbol{q}}_{11}\, {}^t\hat{\boldsymbol{q}}_{11} + \hat{\boldsymbol{M}}, \end{aligned} \tag{51}$$

where $\hat{\boldsymbol{M}}$ is

$$\begin{aligned} \hat{\boldsymbol{M}} &= \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \frac{\lambda_k\, \lambda_{k'}\, \hat{\boldsymbol{q}}_{kk'}\, {}^t\hat{\boldsymbol{q}}_{kk'}}{1-\lambda_k\lambda_{k'}} \\ &= \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \sum_{s=1}^{\infty} \left( \lambda_k\, \lambda_{k'}\, \hat{\boldsymbol{q}}_{kk'}\, {}^t\hat{\boldsymbol{q}}_{kk'} \right)^s \\ &= \sum_{s=1}^{\infty} \left( \sum_{\substack{1\le k,k'\le n \\ (k,k')\neq(1,1)}} \lambda_k\, \lambda_{k'}\, \hat{\boldsymbol{q}}_{kk'}\, {}^t\hat{\boldsymbol{q}}_{kk'} \right)^s \end{aligned}$$

$$= \sum_{s=1}^{\infty} \left( \sum_{k=1}^{n} \sum_{k'=1}^{n} \lambda_k \lambda_{k'} \, \hat{\boldsymbol{q}}_{kk'} \, {}^t\hat{\boldsymbol{q}}_{kk'} - \hat{\boldsymbol{q}}_{11} \, {}^t\hat{\boldsymbol{q}}_{11} \right)^s$$

$$= \sum_{s=1}^{\infty} \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} \, {}^t\hat{\boldsymbol{q}}_{11} \right)^s$$

$$= \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} + \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right)^2 \sum_{s=0}^{\infty} \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right)^s$$

$$= \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} + \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right)^2 \sum_{\substack{1 \le k,k' \le n \\ (k,k') \ne (1,1)}} \frac{\hat{\boldsymbol{q}}_{kk'} {}^t\hat{\boldsymbol{q}}_{kk'}}{1 - \lambda_k \lambda_{k'}}. \quad (52)$$

This derivation process involved the use of

$${}^t\hat{\boldsymbol{q}}_{ij} \, \hat{\boldsymbol{q}}_{kl} = \begin{cases} 1 & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise} \end{cases}, \quad (53)$$

$$(\hat{\boldsymbol{q}}_{kk'} \, {}^t\hat{\boldsymbol{q}}_{kk'})^s = \hat{\boldsymbol{q}}_{kk'} \, {}^t\hat{\boldsymbol{q}}_{kk'}. \quad (54)$$

Substituting Eq. (51) into Eq. (48) yields the following as $\mu_{a,b}$:

$$\mu_{a,b} = \frac{s_1^2}{s_2} + \frac{s_1^2}{s_2^2} \sum_{c \in V} \sum_{c'=1}^{n} d_c^2 \, d_{c'}^2 \, {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{M}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}). \quad (55)$$

The following was used to obtain this equation:

$$\begin{aligned} & {}^t\hat{\boldsymbol{u}}_{cc} \left( \hat{\boldsymbol{N}}^{\dagger} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right) \\ &= {}^t\hat{\boldsymbol{u}}_{cc} \left[ \left( \hat{\boldsymbol{I}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} + \hat{\boldsymbol{M}} \right) (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right] \\ &= \begin{cases} \dfrac{1}{d_c^2} + {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{M}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) & \text{if } c = c' \\ {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{M}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) & \text{otherwise} \end{cases}. \end{aligned} \quad (56)$$

The first term on the right-hand side of Eq. (55) corresponds to the principal component of the expected first meeting time $\mu_{a,b}$.

To confirm that $s_1^2/s_2$ is the principal component of the expected first meeting time $\mu_{a,b}$, we discuss

$$\left| \frac{\mu_{a,b}}{s_1^2} - \frac{1}{s_2} \right| = \frac{1}{s_2^2} \sum_{c \in V} \sum_{c'=1}^{n} d_c^2 \, d_{c'}^2 \left| {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{M}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right|. \quad (57)$$

The right-hand side of this equation expresses the error between $\mu_{a,b}$ and the principal component $s_1^2/s_2$.

We examine the upper bound on the right-hand side of Eq. (57) using

$$\frac{1}{1 - \lambda_k \lambda_{k'}} \le \frac{1}{1 - \lambda_2}, \quad (58)$$

for $2 \le k, k' \le n$. Using the above equation, we obtain

$$\begin{aligned} & \left| {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{M}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| \\ &= \left| {}^t\hat{\boldsymbol{u}}_{cc} \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right) (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| \\ &\quad + \left| {}^t\hat{\boldsymbol{u}}_{cc} \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right)^2 \sum_{\substack{1 \le k,k' \le n \\ (k,k') \ne (1,1)}} \frac{\hat{\boldsymbol{q}}_{kk'} {}^t\hat{\boldsymbol{q}}_{kk'}}{1 - \lambda_k \lambda_{k'}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| \\ &\le \left| {}^t\hat{\boldsymbol{u}}_{cc} \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right) (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| \\ &\quad + \frac{1}{1 - \lambda_2} \left| {}^t\hat{\boldsymbol{u}}_{cc} \left( \hat{\boldsymbol{W}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right)^2 \left( \hat{\boldsymbol{I}} - \hat{\boldsymbol{q}}_{11} {}^t\hat{\boldsymbol{q}}_{11} \right) (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| \\ &= \left| {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{W}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| + \frac{1}{1 - \lambda_2} \left| {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{W}}^2 (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| \\ &= \left| {}^t\hat{\boldsymbol{u}}_{cc} \hat{\boldsymbol{W}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| + \frac{1}{1 - \lambda_2} \left| \hat{\boldsymbol{W}} \hat{\boldsymbol{u}}_{cc} \right| \left| \hat{\boldsymbol{W}} (\hat{\boldsymbol{u}}_{c'c'} - \hat{\boldsymbol{u}}_{ab}) \right| \\ &\le \frac{2 \, w_{\max}^2}{d_{\min}^4} + \frac{1}{1 - \lambda_2} \frac{w_{\max}}{d_{\min}^2} \frac{\sqrt{2} \, w_{\max}}{d_{\min}^2} \\ &\le \frac{2 \, w_{\max}^2}{d_{\min}^4} \left( \frac{1}{1 - \lambda_2} + 1 \right). \end{aligned} \quad (59)$$

The following was used in this derivation process:

$$\begin{aligned} \left| {}^t\hat{\boldsymbol{u}}_{ij} \hat{\boldsymbol{W}} \hat{\boldsymbol{u}}_{kl} \right| &= {}^t\boldsymbol{u}_i \boldsymbol{W} \boldsymbol{u}_k \otimes {}^t\boldsymbol{u}_j \boldsymbol{W} \boldsymbol{u}_l \\ &= \frac{w_{ik}}{d_i \, d_k} \otimes \frac{w_{jl}}{d_j \, d_l} \le \frac{w_{\max}^2}{d_{\min}^4}, \end{aligned} \quad (60)$$

$$\begin{aligned} \left| \hat{\boldsymbol{W}} \hat{\boldsymbol{u}}_{ij} \right|^2 &= {}^t \left( \hat{\boldsymbol{W}} \hat{\boldsymbol{u}}_{ij} \right) \left( \hat{\boldsymbol{W}} \hat{\boldsymbol{u}}_{ij} \right) \\ &= \sum_{k \in \partial i}^{n} \sum_{l \in \partial j}^{n} \frac{w_{ki}^2 \, w_{lj}^2}{d_k \, d_i^2 \, d_l \, d_j^2} \\ &\le \frac{w_{\max}^2}{d_i^2 \, d_j^2 \, d_{\min}^2} \sum_{k \in \partial i}^{n} \sum_{l \in \partial j}^{n} w_{ki} \, w_{lj} \\ &= \frac{w_{\max}^2}{d_i \, d_j \, d_{\min}^2} \le \frac{w_{\max}^2}{d_{\min}^4}. \end{aligned} \quad (61)$$

Substituting Eq. (59) into Eq. (57) yields the following as the upper bound of the error between $\mu_{a,b}$ and the principal component $s_1^2/s_2$:

$$\left| \frac{\mu_{a,b}}{s_1^2} - \frac{1}{s_2} \right| \le \frac{2 \, w_{\max}^2}{d_{\min}^4} \left( \frac{1}{1 - \lambda_2} + 1 \right). \quad (62)$$

According to this equation, the error can be expected to be small for graph $G$, where $\lambda_2$ and $w_{\max}$ are small but $d_{\min}$ is large. In this case, the expected first meeting time $\mu_{a,b}$ is approximated as

$$\mu_{a,b} \approx \frac{s_1^2}{s_2} = \frac{n}{1 + d_{\text{std}}^2 / d_{\text{avg}}^2}, \quad (63)$$

where $d_{\text{avg}}$ and $d_{\text{std}}$ are the average and standard deviation of weighted degrees, respectively. Equations (62) and (63) are coincidentally similar to Eqs. (14) and (15), respectively. Since the analysis target for Eqs. (62) and (63) is different from that of Eqs. (14) and (15), these equations are new findings. In Sect. 4, we will show that the right hand side of Eq. (62) is not large enough to be a problem in popular random graphs (Barabási-Albert graphs and Erdös-Rényi graphs). In this paper, the graphs where the right-hand side of Eq. (62) is negligibly small are not yet clear. However, since Eq. (14) based on the result of [14] and Eq. (62) are coincidentally similar, we may prove that the right hand side of

Eq. (62) is negligibly small in several graphs (e.g., $\epsilon$-graphs, k-nearest neighbor graphs, Gaussian similarity graphs, and Erdös-Rényi graphs) on the basis of the similar discussion in [14].

If the approximation formula (63) holds for the expected first meeting time $\mu_{a,b}$, we derive the following characteristics: (a) $\mu_{a,b}$ is small when the coefficient of variation $d_{\rm std}/d_{\rm avg}$ is large and (b) $\mu_{a,b}$ does not depend on the starting nodes $a$ and $b$. Characteristic (a) is useful for understanding the effect of the graph structure on $\mu_{a,b}$. The reason why characteristic (b) emerges can be explained as follows. First, a random walk is a Markov chain, and thus the information of the initial state is lost over time. Since the time to lose the information is shorter than the first meeting time of two random walks, the expected first meeting time would be independent of the starting nodes. As the number of nodes in a graph increases, the first meeting time becomes longer. Hence, in order for the independency of the expected first meeting time to emerge, the graph should be large enough.

## 4. Numerical Example

In this section, we confirm the validity of the spectral formula (40) and the principal component of the expected first meeting time $\mu_{a,b}$ revealed in Sect. 3. We also examine the error in the approximation formula (63) obtained when $\mu_{a,b}$ is replaced by its principal component.

### 4.1 Setting

In this subsection, we use BA (Barabási-Albert) graphs [15] and ER (Erdös-Rényi) graphs [16]. The spectral formula (40) and the approximation formula (63) depend on the degree distribution of a graph. Since the degree distribution of a BA graph is different from that of an ER graph, these graphs are useful for clarifying the effects of the degree distribution on these formulas. Owing to space limitations, we provide the results for unweighted graphs, where $w_{ij} = 1$ is provided for all links $\forall(i, j) \in E$. In unweighted graphs, the weighted degree $d_i$ of node $i$ corresponds to the degree $m_i$, the number of links of node $i$. According to the spectral characteristics of ER graphs and BA graphs shown in [17], our analysis results may be valid in these graphs. Hence, the scope of the numerical example seems to be biased. However, these graphs are popular random graphs, and thus this numerical example is worthwhile as the first step to confirm the validity of our analysis results.

The BA model [15] is a typical model for scale-free random graphs. A BA graph is generated using the following procedure. First, a complete graph with $n_0$ nodes is created. We assume that $n_0 = m$ for the sake of simplicity. Next, nodes are inserted one by one until the number of nodes in the BA graph is equal to $n$. When adding the $t$th node ($t = m + 1, m + 2, \cdots, n$), $m_{\rm new}$ links are created from node $t$ to nodes $j \in \{1, 2, \cdots, t - 1\}$ with the connection probability $p_j^{\rm BA}(t)$. The connection probability $p_j^{\rm BA}(t)$ is

$$p_j^{\rm BA}(t) = \frac{m_j(t)}{\sum_{l=1}^{t-1} m_l(t)}, \tag{64}$$

where $m_j(t)$ is the degree of node $j$ when the insertion of the $t - 1$th node is completed. BA graphs have the power-law degree distribution (i.e, Prob $[m_i = m] \sim m^{-3}$). If $G$ is unweighted, then the average weighted degree $d_{\rm avg}$ is equal to the average degree $k_{\rm avg}$. Hence, the average weighted degree $d_{\rm avg}$ of a BA graph is approximated as

$$d_{\rm avg} = \frac{m\,(m-1) + 2\,m\,(n-m)}{n} \approx 2\,m, \tag{65}$$

where we assume that $n \gg m$.

In contrast, the ER model [16] is a classical random graph model. An ER graph is generated through the following procedure. First, $n$ nodes are created. Next, links are created between any pair of nodes with probability $p^{\rm ER}$. If the graph is not connected, then the link creation process is begun again. The average weighted degree $d_{\rm avg}$ of an ER graph is

$$d_{\rm avg} = (n-1)\,p^{\rm ER}. \tag{66}$$

The degree distribution Prob $[m_i = m]$ of an ER graph follows the binomial distribution. According to the difference between the power law and the binomial distribution, the standard deviation $d_{\rm std}$ of weighted degrees in a BA graph is greater than that in an ER graph.

To focus on the difference in the standard deviation $d_{\rm std}$ of weighted degrees, we set $m$ and $p^{\rm ER}$ as

$$m = \left\lfloor \frac{d_{\rm avg}}{2} \right\rfloor, \tag{67}$$

$$p_{\rm ER} = \frac{d_{\rm avg}}{n-1}, \tag{68}$$

with the result that the average weighted degree $d_{\rm avg}$ of an ER graph and a BA graph are roughly equal. The minimum weighted degree $d_{\rm min}$ in both graphs also increases as $d_{\rm avg}$ increases.

To examine the validity and the error of the spectral formula (40) and the approximation formula (63), we measure the average of the first meeting times in simulation using the following procedure.

1. Generate a BA graph or an ER graph using the above procedures.
2. Put random walkers on nodes $a \in V$ and $b \in V$.
3. Move each random walker with the transition probability $p_{i\to j}$ in accordance with Eq. (2).
4. Repeat step 3 until the two random walks meet at the same node.
5. Repeat step 2 through step 4 10,000 times to calculate the average of the first meeting times.

We use the parameter configuration shown in Table 1 as a default parameter configuration.

**Table 1** Parameter configuration.

| | | |
|---|---|---|
| Number of nodes | $n$ | 1,000 |
| Weight of link $(i, j)$ | $w_{ij}$ | 1 |
| Average weighted degree | $d_{\text{avg}}$ | 6 |
| Random walker's starting node | $a$ | 1 |

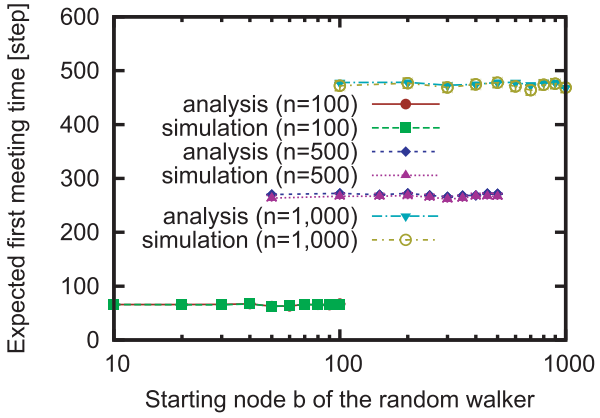


**Fig. 2** Expected first meeting time $\mu_{a,b}$ for different settings of the starting node $b$ of the random walker in BA graphs.

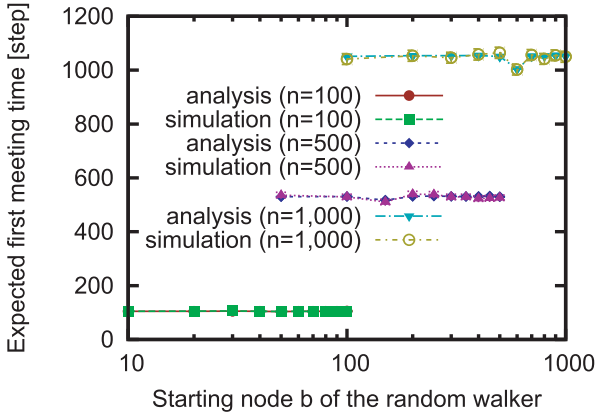


**Fig. 3** Expected first meeting time $\mu_{a,b}$ for different settings of the starting node $b$ of the random walker in ER graphs.

## 4.2 Validity of the Spectral Formula for the Expected First Meeting Time $\mu_{a,b}$

We confirm the validity of the spectral formula of the expected first meeting time $\mu_{a,b}$ given by Eq. (40).

Figures 2 and 3 show the first meeting times obtained from the simulation and the analysis (i.e., the spectral formula (40)) for different settings of the random walker's starting node $b$ in the BA and ER graphs, respectively. According to Figs. 2 and 3, the analysis results are almost the same as the simulation results regardless of the choice of $n$ and $b$.

We then evaluate the error in the spectral formula (40). In this evaluation, we use the relative error $\epsilon_{a,b}$ of the expected first meeting time $\mu_{a,b}$. The relative error $\epsilon_{a,b}$ is defined as

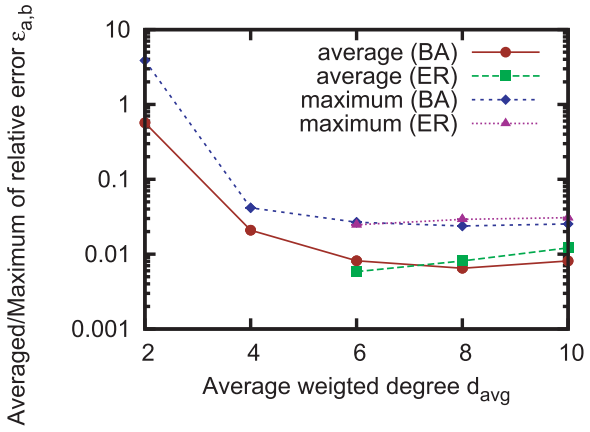


**Fig. 4** Average weighted degree $d_{\text{avg}}$ vs. the average and the maximum of the relative error $\epsilon_{a,b}$ for the expected first meeting $\mu_{a,b}$.

$$\epsilon_{a,b} := \frac{\left|\mu_{a,b} - \mu_{a,b}^{\text{sim}}\right|}{\mu_{a,b}^{\text{sim}}}, \tag{69}$$

where $\mu_{a,b}^{\text{sim}}$ is the average of the first meeting times obtained from the simulation. We examine the average and the maximum of the relative errors $\epsilon_{a,b}$ when changing starting node $b$ while the starting node $a$ is fixed.

Figure 4 shows the average and the maximum of the relative errors $\epsilon_{a,b}$ of the expected first meeting time $\mu_{a,b}$ in the BA and ER graphs with different settings of the average weighted degree $d_{\text{avg}}$. In this figure, we do not plot the results for the ER graphs with $d_{\text{avg}} = 2$ and 4, because a connected ER graph cannot be generated. According to Fig. 4, if $d_{\text{avg}} \geq 4$, then the maximum of relative errors $\epsilon_{a,b}$ is only a few percent. Therefore, the spectral formula (40) is valid for the graphs with $d_{\text{avg}} \geq 4$.

We discuss the reason that the relative error $\epsilon_{a,b}$ is large when we use a BA graph with a small-average weighted degree (i.e., $d_{\text{avg}} = 2$). In Sect. 3, we use Eq. (22) to derive the spectral formula (40). Equation (22) assumes that the first meeting probability of two random walkers at node $c$ is proportional to $d_c^2$. Hence, we confirm the acceptance of this assumption to clarify the reason for the large relative error.

Figures 5(a) through (c) show scatter plots of the first meeting frequency of two random walkers at node $c$ in BA graphs with different settings of the average weighted degree $d_{\text{avg}}$. The first meeting frequency at each node was obtained from the simulation, where the starting nodes $a$ and $b$ are fixed. In order to confirm easily the correctness of the assumption, we plot the fitting curve of $d_c^2$ in these figures. According to Figs. 5(a) through (c), the first meeting frequencies with $d_{\text{avg}} = 2$ differ only largely from the fitting curve, with the consequence that the assumption must not be accepted for the cases with $d_{\text{avg}} = 2$. Therefore, we conclude that the large relative error shown in Fig. 4 is caused by the assumption for Eq. (22).

According to the results, the spectral formula of first meeting time $\mu_{a,b}$ will be valid if the average weighted degree $d_{\text{avg}}$ is sufficiently large (i.e., $d_{\text{avg}} \geq 4$).

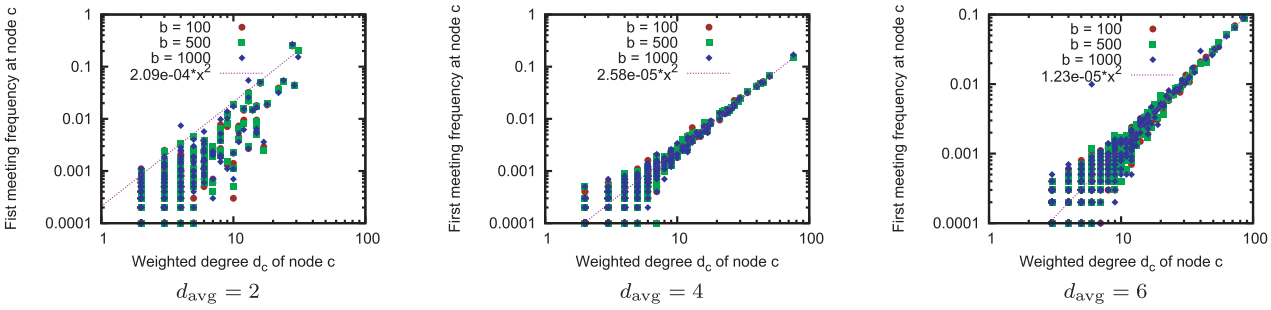


**Fig. 5** Weighted degree $d_c$ of node $c$ vs. the first meeting frequency at node $c$ in BA graphs.

### 4.3 Validity for the Principal Component of the Expected First Meeting Time $\mu_{a,b}$

We clarify the validity for the principal component of the expected first meeting time $\mu_{a,b}$ derived in Sect. 3. Specifically, we examine the relative error $\epsilon'_{a,b}$ of the approximation formula (63) obtained when the expected first meeting time $\mu_{a,b}$ is given by the principal component (i.e., $s_1^2/s_2$). The relative error $\epsilon'_{a,b}$ is defined as

$$\epsilon'_{a,b} := \frac{\left|\frac{s_1^2}{s_2} - \mu_{a,b}^{\mathrm{sim}}\right|}{\mu_{a,b}^{\mathrm{sim}}}. \tag{70}$$

Figures 6 and 7 show the averages of the relative errors $\epsilon'_{a,b}$ of the approximation formula (63) for BA and ER graphs with different numbers of nodes, $n$, respectively. The average of the relative errors $\epsilon'_{a,b}$ was calculated from 10,000 simulations, where the starting nodes $a$ and $b$ are selected randomly. In Fig. 7, we do not plot the result for $n = 10,000$ and $d_{\mathrm{avg}} = 6$, because a connected ER graph cannot be generated. According to the results, if the average weighted degree $d_{\mathrm{avg}}$ is sufficiently large, the relative error $\epsilon'_{a,b}$ is small, and the derived principal component is valid. This can also be explained by Eq. (62). The right-hand side of Eq. (62) represents the upper bound of the error in the approximation formula (63). If the average weighted degree $d_{\mathrm{avg}}$ is large, the minimum weighted degree $d_{\mathrm{min}}$ is also large. As the minimum weighted degree $d_{\mathrm{min}}$ increases, the upper bound becomes small, and the relative error $\epsilon'_{a,b}$ of the approximation formula (63) can be expected to decrease. Moreover, according to Figs. 6 and 7, the average of the relative errors $\epsilon'_{a,b}$ is constant or becomes smaller as $n$ increases, and hence the approximation formula (63) is also effective for large-scale graphs.

From the above results, the derived principal component is valid if the average weighted degree $d_{\mathrm{avg}}$ is sufficiently large (i.e., $d_{\mathrm{avg}} \geq 4$). According to the site [18], which collecting statistical information (e.g., average degree) of various existing graphs, the average degree of a typical graph is greater than four. Hence, our analysis results are expected to be useful for many real graphs.

Finally, we confirm the effect of the graph structure on the expected first meeting time $\mu_{a,b}$ revealed in Sect. 3. Ac-

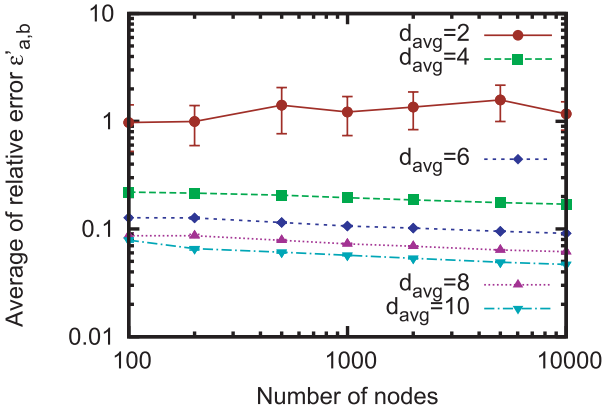


**Fig. 6** Number of nodes, $n$, vs. the average of the relative errors $\epsilon'_{a,b}$ of the approximation formula (63) in the BA graphs.

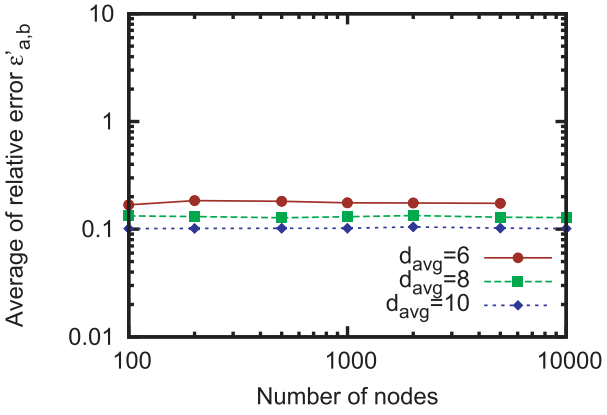


**Fig. 7** Number of nodes, $n$, vs. the average of the relative errors $\epsilon'_{a,b}$ of the approximation formula (63) in the ER graphs.

cording to the approximation formula (63), $\mu_{a,b}$ increases as $s_1^2/s_2$ increases. To confirm the effect from the numerical example, we compare $s_1^2/s_2$ and the average of the first meeting times obtained in the simulation.

Figures 8 and 9 show the averages of the first meeting times obtained from the simulation with different settings of $s_1^2/s_2$ in the BA and ER graphs, respectively. To calculate the average of the first meeting times, we conduct 10,000 simulations, where the starting nodes $a$ and $b$ are selected randomly. In these figures, we plot the straight line for $y = x$ to confirm the effect easily. According to the results, the

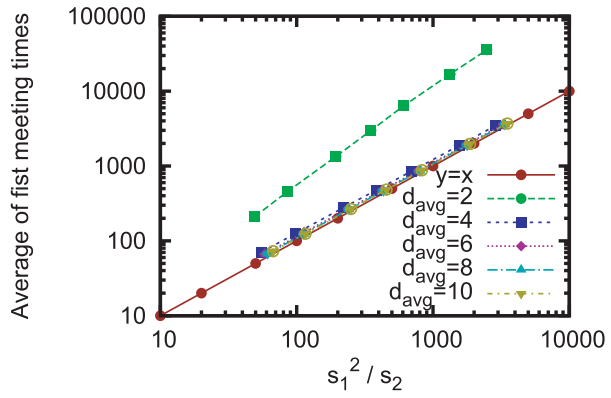


**Fig. 8** Principal component $s_1^2/s_2$ vs. the average of the first meeting times in the BA graphs.

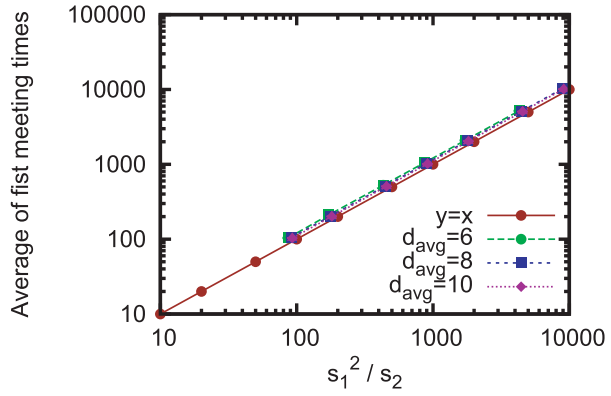


**Fig. 9** Principal component $s_1^2/s_2$ vs. the average of the first meeting times in the ER graphs.

average of the first meeting times is approximately along the $y = x$ line, except for the result for BA graphs with average weighted degree $d_{\text{avg}} = 2$. Therefore, the effect is also confirmed from the numerical example if the average weighted degree $d_{\text{avg}}$ is sufficiently large.

## 5. Conclusion and Future Work

In this paper, we analyzed the first meeting time of multiple random walks in arbitrary graphs and clarified the effects of graph structures on its expected value. First, we derived the spectral formula of the expected first meeting time for two random walkers using spectral graph theory. Then, we examined the principal component of the expected first meeting time using the derived spectral formula. The clarified principal component reveals that (a) the expected first meeting time is almost dominated by $n/(1 + d_{\text{std}}^2/d_{\text{avg}}^2)$, and (b) the expected first meeting time is independent of the starting nodes of random walkers, where $n$ is the number of nodes. $d_{\text{avg}}$ and $d_{\text{std}}$ are the average and the standard deviation of the weighted degree, respectively. Characteristic (a) is useful for understanding the effect of the graph structure on the first meeting time. In addition, we confirmed the validity of the analysis results through numerical examples using popular random graphs. According to the revealed effects of the graph structures, the variance of the coefficient for weighted degrees, $d_{\text{std}}/d_{\text{avg}}$ (degree heterogeneity), facilitates the meeting of random walkers.

As future work, we plan to examine the validity of the analysis results with real graphs and apply them to the development of efficient graph algorithms.

## Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 19K11927.

### References

[1] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker, "Search and replication in unstructured peer-to-peer networks," Proc. 16th ACM international conference on Supercomputing (ICS'02), pp.84–95, June 2002.
[2] C. Gkantsidis, M. Mihail, and A. Saberi, "Random walks in peer-to-peer networks," Proc. 23rd Conference of the IEEE Communications Society (INFOCOM 2004), pp.120–130, March 2004.
[3] C. Dutta, G. Pandurangan, R. Rajaraman, and S. Roche, "Coalescing-branching random walks on graphs," ACM Trans. Parallel Comput. (TOPC), vol.2, no.3, pp.1–29, Nov. 2015.
[4] Y. Metivier, N. Saheb, and A. Zemmari, "Randomized rendezvous," Mathematics and Computer Science, pp.183–194, Birkhäuser, Basel, 2000.
[5] A. Israeli and M. Jalfon, "Token management schemes and random walks yield self-stabilizing mutual exclusion," Proc. 9th Annual ACM Symposium on Principles of Distributed Computing (PODC'90), pp.119–131, ACM, Aug. 1990.
[6] P. Tetali and P. Winkler, "On a random walk problem arising in self-stabilizing token management," Proc. 10th annual ACM symposium on Principles of distributed computing (PODC'91), pp.273–280, ACM, Aug. 1991.
[7] D. Coppersmith, P. Doyle, P. Raghavan, and M. Snir, "Random walks on weighted graphs and applications to on-line algorithms," J. ACM, vol.40, no.3, pp.421–453, July 1993.
[8] D.J. Aldous, "Meeting times for independent Markov chains," Stochastic Processes and their Applications, vol.38, no.2, pp.185–193, Aug. 1991.
[9] N.H. Bshouty, L. Higham, and J. Warpechowska-Gruca, "Meeting times of random walks on graphs," Inform. Process. Lett., vol.69, no.5, pp.259–265, 1999.
[10] C. Cooper, A. Frieze, and T. Radzik, "Multiple random walks in random regular graphs," SIAM J. Discrete Math., vol.23, no.4, pp.1738–1761, June 2009.
[11] Y. Zhang, Z. Tan, and B. Krishnamachari, "On the meeting time for two random walks on a regular graph," arXiv preprint arXiv:1408.2005, 2014.
[12] M. George, R. Patel, and F. Bullo, "The meeting time of multiple random walks," Preprint submitted to Linear Algebra and Its Applications, March 2017, available at http://motion.me.ucsb.edu/pdf/2014l-gpb.pdf
[13] L. Lovász, "Random walks on graphs: A survey," Combinatorics, Paul Erdős is eighty, vol.2, pp.353–398, 1996.
[14] U. Von Luxburg, A. Radl, and M. Hein, "Hitting and commute times in large random neighborhood graphs," The Journal of Machine Learning Research, vol.15, no.1, pp.1751–1798, Jan. 2014.
[15] A.L. Bárabasi and R. Albert, "Emergence of scaling in random networks," Science, vol.286, no.5439, pp.509–512, Oct. 1999.
[16] P. Erdös and A. Rényi, "On random graphs," Mathematicae, vol.6, no.26, pp.290–297, 1959.
[17] Y. Sakumoto and M. Aida, "Wigner's semicircle law of weighted random networks," IEICE Trans. Commun., vol.E104-B, no.3,

pp.251–261, March 2021.
[18] J. Kunegis, "The Koblenz network collection (KONECT)," http://konect.uni-koblenz.de/ (accessed on May 29, 2020).

**Yusuke Sakumoto** received M.E. and Ph.D. degrees in the Information and Computer Sciences from Osaka University in 2008 and 2010, respectively. From 2010 to 2019, he was a associate professor of Tokyo Metropolitan University. He is currently an associate professor at Kwansei Gakuin University. His research work is in the area of analysis of communication network and social network. He is a member of the IEEE, IEICE and IPSJ.

**Hiroyuki Ohsaki** received the M.E. degree in the Information and Computer Sciences from Osaka University, Osaka, Japan, in 1995. He also received the Ph.D. degree from Osaka University, Osaka, Japan, in 1997. He is currently a professor at Department of Informatics, School of Science and Technology, Kwansei Gakuin University, Japan. His research work is in the area of design, modeling, and control of large-scale communication networks. He is a member of IEEE and Institute of Electronics, Information, and Computer Engineers of Japan (IEICE).