

## Article

# Pairing Optimization via Statistics: Algebraic Structure in Pairing Problems and Its Application to Performance Enhancement

Naoki Fujita <sup>1,\*</sup>, André Röhm <sup>1</sup>, Takatomo Mihana <sup>1</sup>, Ryoichi Horisaki <sup>1</sup>, Aohan Li <sup>2</sup>, Mikio Hasegawa <sup>3</sup> and Makoto Naruse <sup>1</sup>

<sup>1</sup> Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

<sup>2</sup> Graduate School of Informatics and Engineering, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan

<sup>3</sup> Department of Electrical Engineering, Faculty of Engineering, Tokyo University of Science, 6-3-1 Niijuku, Katsushika-ku, Tokyo 125-8585, Japan

\* Correspondence: fujitanaoki6322@gmail.com

**Abstract:** Fully pairing all elements of a set while attempting to maximize the total benefit is a combinatorically difficult problem. Such pairing problems naturally appear in various situations in science, technology, economics, and other fields. In our previous study, we proposed an efficient method to infer the underlying compatibilities among the entities, under the constraint that only the total compatibility is observable. Furthermore, by transforming the pairing problem into a traveling salesman problem with a multi-layer architecture, a pairing optimization algorithm was successfully demonstrated to derive a high-total-compatibility pairing. However, there is substantial room for further performance enhancement by further exploiting the underlying mathematical properties. In this study, we prove the existence of algebraic structures in the pairing problem. We transform the initially estimated compatibility information into an equivalent form where the variance of the individual compatibilities is minimized. We then demonstrate that the total compatibility obtained when using the heuristic pairing algorithm on the transformed problem is significantly higher compared to the previous method. With this improved perspective on the pairing problem using fundamental mathematical properties, we can contribute to practical applications such as wireless communications beyond 5G, where efficient pairing is of critical importance. As the pairing problem is a special case of the maximum weighted matching problem, our findings may also have implications for other algorithms on fully connected graphs.

**Keywords:** pairing; optimization; matching; maximum weighted matching; heuristic algorithm



**Citation:** Fujita, N.; Röhm, A.; Mihana, T.; Horisaki, R.; Li, A.; Hasegawa, M.; Naruse, M. Pairing Optimization via Statistics: Algebraic Structure in Pairing Problems and Its Application to Performance Enhancement. *Entropy* **2023**, *25*, 146. <https://doi.org/10.3390/e25010146>

Academic Editors: Christian H. Weiss and Serge M. Prudhomme

Received: 2 November 2022

Revised: 14 December 2022

Accepted: 10 January 2023

Published: 11 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The procedure of generating pairs of elements among all entries of a given system often arises in various situations in science, technology, and economy [1–7]. Here, we call such a process pairing, and the number of elements is considered to be an even number for simplicity. One immediately obvious problem is that the number of pairing configurations grows rapidly with the number of elements. The number of possible pairings is given by  $(n - 1)!!$ , where  $n$  indicates the number of elements in the system and  $!!$  is the double factorial operator. For example, when  $n$  is 100, the total number of possible pairings is on the order of  $10^{78}$ . Hence, finding the pairing that maximizes the benefit of the total system is difficult.

Notably, the pairing problem corresponds to the maximum weighted matching (MWM) problem on the complete graph. Multiple algorithms exist for solving the MWM problem [8–15]. In contrast to these conventional methods, we propose a heuristic and fast algorithm at the

cost of some performance. The advantage of a fast heuristic algorithm is that it can be useful in environments where weights change dynamically or a quick pairing is required, such as in communications technology. A heuristic algorithm for the MWM problem using deep reinforcement learning was recently proposed by [16] with a similar goal. Furthermore, our research proposes algorithms that work under the limited observation constraint, which is explained later. In our previous study, we proposed an algorithm with a computational complexity of  $\mathcal{O}(n^2)$  [17].

To the best of our knowledge, there is no exact algorithm that works on the order of  $\mathcal{O}(n^2)$  for arbitrary weights. For example, Gabow [9] proposed a MWM algorithm with a computation time of  $|E||V| + |V|^2 \log |V|$ , where  $V$  is a set of vertices and  $E$  is a set of edges. However, randomized or approximate algorithms can reduce computational time for some cases. For example, Cygan et al. [12] developed a randomized algorithm with a computation time of  $L|V|^\omega$  for graphs with integer weights ( $\omega < 2.373$  is the exponent of  $n \times n$  matrix multiplication [18] and  $L$  is the maximum integer edge weight). Duan et al. [15] proposed an approximate algorithm achieving an approximation ratio of  $(1 - \epsilon)M$  with a computation time of  $|E|\epsilon^{-1} \log \epsilon^{-1}$  for arbitrary weights and  $|E|\epsilon^{-1} \log N$  for integer weights ( $\epsilon$  is a positive arbitrary value and  $M$  is the maximum possible weight matching value). Here,  $|V| = n$ ,  $|E| = n(n - 1)/2$ . Here, we aim to improve our previous pairing problem result, i.e., to determine a higher-accuracy heuristic algorithm that works with  $\mathcal{O}(n^2)$  computational complexity.

Note that the pairing problem should not be confused with the assignment problem, which is another special case of the MWM setting. The assignment problem requires the graph to be a weighted bipartite graph. Furthermore, in the assignment problem there are two classes of objects, where it is the goal to always match an object from the first class with an object from the second. However, in the pairing problem, there is only a single class of objects, and we allow any of them to be potentially paired with any other. The assignment problem is also related to the single-source shortest paths problem. Several well-known assignment algorithms [19–21] or single-source shortest paths algorithms [22] are known. For example, the Hungarian algorithm [19] solves the assignment problem  $\mathcal{O}(n^3)$ , the auction algorithm [20] works with parallelism and the Bellman–Ford algorithm runs with  $\mathcal{O}(|V||E|)$  [22]. However, in this study, we consider a fully connected graph with an even number of elements, where the MWM problem cannot be solved by assignment problem algorithms.

An example of a pairing problem is found in a recent communication technology called non-orthogonal multiple access (NOMA) [23–29]. In NOMA, multiple terminals simultaneously share a common frequency band to improve the efficiency of frequency usage. The simultaneous use of the same frequency band causes interference in the signals from the base station to each terminal. To overcome this problem, NOMA uses a signal processing method called successive interference cancellation (SIC) [30] to distinguish individual channel information in the power domain, allowing multiple terminals to rely on the same frequency band. For simplicity, here we consider that the number of terminals that can share a frequency is given by two. Herein, the usefulness of the whole system can be measured by the total communication quality, such as high data throughput and low error rate, which depends crucially on the method of pairing.

The most fundamental parameter of the pairing problem is the merit between any two given elements, which we call individual compatibility, while the summation of compatibilities for a given pairing is called its total compatibility. The detailed definition is introduced below. Our goal is to derive pairings yielding high total compatibility.

In general, we do not need to assume that the individual compatibility of a pair is observable, i.e., only the total compatibility of a given pairing may be observed. Our previous study [17] divided the pairing problem into two phases. The first is the observation phase, where we observe total compatibilities for several pairings and estimate the individual compatibilities. The second is the combining phase, in which a search is performed for a pairing that provides high total compatibility. This procedure is referred to as pairing

optimization. The search is based on the compatibility information obtained in the first phase. In [17], we show that the pairing optimization problem can be transformed into a travelling salesman problem (TSP) [31] with a three-layer structure, allowing us to benefit from a variety of known heuristics.

However, we consider that there is substantial room for further performance optimization. This study sheds new light on the pairing problem from two perspectives. The first is to clarify the algebraic structure of the pairing optimization problem. Because we care only about the total compatibility when all elements are paired, there are many compatibility matrices (defined in Section 2) that share the same total compatibilities. In other words, we can consider an equivalence class of compatibility matrices that yield the same total compatibilities and that cannot be distinguished if individual compatibilities are not measurable. We show that the compatibility matrices in each equivalence class have an invariant value.

Second, although any compatibility matrices in the same equivalence class theoretically provide the same total compatibility, the heuristic pairing optimization process can result in different total compatibility values. These differences are not caused by incomplete or noisy observations, but are due to the convergence properties of the heuristic pairing algorithms, which yield better results on some distributions than others. We examine how the statistics of the compatibility matrix affect the pairing optimization problem and propose a compatibility matrix that yields higher total compatibility after optimization. More specifically, we propose a transformation to the compatibility matrix that minimizes the variance of the elements therein, which we call the variance optimization. We confirmed numerically that enhanced total compatibility is achieved via the compatibility matrix after variance optimization. Furthermore, the proposed variance optimization algorithm may also be applicable when no observation phase is required, i.e., when the individual compatibilities are directly observable. In other words, there are cases where a compatibility matrix unsuitable for a heuristic combining algorithm can be converted to one that is easily combinable.

The remainder of this paper is organized as follows. In Section 2, we define the pairing optimization problem mathematically. Section 3 describes the mathematical properties of the equivalence class. Section 4 explains the concept of variance optimization and presents a solution by which it can be achieved. Section 5 presents results of numerical simulations of the proposed variance optimization. Furthermore, there are two optimization problems in this paper. The first is the pairing problem we aim to solve in Section 2.1. Second is the variance optimization which enables us to enhance the performance of the PNN+p2-opt algorithm in Section 4.2. Finally, Section 6 concludes the paper.

## 2. Problem Setting

In this section, we provide a mathematical definition of the pairing optimization problem that we address in this study, and define some of the mathematical symbols used in the following discussion. In addition, we explain the constraints applied to the pairing optimization problem.

### 2.1. Pairing Optimization Problem

Here, we assume that the number of elements is an even natural integer  $n$ , while the index of each element is a natural number between 1 and  $n$ . Parts of the pairing problem can be described elegantly in set theory, while others benefit from using matrix representations. We will use either, where appropriate. Here we use  $\mathbb{U}(n)$  to denote the set of  $n$  elements:

$$\mathbb{U}(n) \equiv \{i \mid i \in \mathbb{Z}, 1 \leq i \leq n\}. \quad (1)$$

Then, we define the set of all possible pairs for  $\mathbb{U}(n)$  as  $\mathbb{P}(n)$ , which contains  $N(N - 1)/2$  pairs:

$$\mathbb{P}(n) \equiv \{\{i, j\} \mid i, j \in \mathbb{U}(n), i < j\}. \quad (2)$$

To describe the compatibilities of these pairs, we now define a “compatibility matrix”  $C$  as follows:

$$\begin{aligned} C &\in \mathbb{R}^{n \times n}, \\ \forall \{i, j\} \in \mathbb{P}(n), C_{i,j} &= C_{j,i}, \\ 1 \leq i \leq n, C_{i,i} &= 0. \end{aligned}$$

The compatibility between elements  $i$  and  $j$  is denoted by  $C_{i,j} \in \mathbb{R}$ . The matrix  $C$  is always symmetric and the major diagonal is zero, because pairing  $i$  and  $j$  does not depend on the order of elements and an element cannot be paired with itself. The set of all possible compatibility matrices is denoted as  $\Omega_n$  when the number of elements is  $n$ . In other words,  $\Omega_n$  is the set of all  $n \times n$  symmetric distance matrices, or symmetric hollow matrices. To describe a pairing, i.e., which elements are paired together, we now define a pairing matrix  $S \in \mathbb{R}^{n \times n}$ :

$$\begin{aligned} \forall \{i, j\} \in \mathbb{P}(n), S_{i,j} &= S_{j,i} \text{ and } S_{i,j} \in \{0, 1\}, \\ 1 \leq i \leq n, S_{i,i} &= 0, \\ \forall i, \sum_{j=1}^n S_{i,j} &= 1. \end{aligned}$$

$S$  is symmetric, because pairing element  $i$  with  $j$  is equivalent to pairing  $j$  with  $i$ . The pairing matrix  $S$  is also hollow, because pairing  $i$  with itself is not allowed. Each row and column contains only a single non-zero element, as each element  $i$  can only be paired once. Therefore, a pairing matrix  $S$  is an  $n \times n$  symmetric and hollow permutation matrix. We define the set of all pairing matrices  $\mathbb{S}(n) \equiv \{S\}$  when the number of elements is  $n$ :

$$S \in \mathbb{S}(n). \tag{3}$$

To derive the set representation of a pairing, we introduce the map  $f_{\text{set}}$  as follows:

$$f_{\text{set}}(S) \equiv \{\{i, j\} \mid i < j \text{ and } S_{i,j} = 1\}. \tag{4}$$

A function denoted by  $\langle X, C \rangle$  is then defined as follows, using the Frobenius inner product  $\langle \cdot \rangle_{\text{F}}$ :

$$\begin{aligned} C &\in \Omega_n, \\ X &\in \mathbb{R}^{n \times n}, \\ \langle X, C \rangle &= \frac{1}{2} \langle X, C \rangle_{\text{F}}. \end{aligned}$$

For a given compatibility matrix  $C$ , we call  $\langle S, C \rangle$  for  $S \in \mathbb{S}(n)$  the “total compatibility” for pairing  $S$ . This formulation is equivalent to the one used in our previous work [17], and corresponds to summing the individual compatibilities  $C_{i,j}$  of the pairs defined by  $S$ :

$$\langle S, C \rangle = \sum_{\{i,j\} \in f_{\text{set}}(S)} C_{i,j}.$$

For any given compatibility matrix  $C$ , the pairing optimization problem can then be formulated as follows:

$$\begin{aligned} \text{max: } &\langle S, C \rangle, \\ \text{subject to: } &S \in \mathbb{S}(n). \end{aligned}$$

### 2.2. Limited Observation Constraint

As briefly mentioned in Section 1, in practice there may often exist one more constraint on the pairing optimization problem. We will assume that initially we do not know each compatibility value. Moreover, we assume that only the value of total compatibility  $\langle S, C \rangle$  for any pairing  $S \in \mathbb{S}(n)$  is observable. We call this condition the “limited observation constraint”.

Under this constraint, we must execute two phases, the “observation phase” and the “combining phase”, as introduced in our previous study [17]. First, we estimate the ground-truth compatibility matrix  $C^g$  through observations of the total compatibilities of several pairings in the observation phase. We denote the estimated compatibility matrix by  $C^e$ . Our previous work [17] calculated the minimum number of observations that are necessary for deducing  $C^e$  and presents a simple algorithm for doing so efficiently.

### 3. Mathematical Properties of the Pairing Problem

In this section, we consider algebraic structures in the pairing problem. An equivalence relation is defined among compatibility matrices to construct equivalence classes. Then we show a conserved quantity within the equivalence class and that all members of the class yield the same total compatibility for any given pairing. Furthermore, the statistical properties of compatibility matrices are examined, forming the mathematical foundation of the variance optimization to be discussed in Section 4.

#### 3.1. Adjacent Set

We define the adjacent set matrix  $R_i (1 \leq i \leq n)$  as follows:

$$R_i \in \mathbb{R}^{n \times n},$$

$$(R_i)_{k,l} = \begin{cases} 1 & \text{if } i \in \{k, l\} \text{ and } k \neq l \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

We can also describe  $f_{\text{set}}(R_i)$  as follows:

$$f_{\text{set}}(R_i) = \{\{i, j\} \mid 1 \leq j \leq n, j \neq i\}. \quad (6)$$

With these adjacent sets, the following theorem holds.

**Theorem 1.**  $C \in \Omega_n$  is fully determined by  $\{\langle S, C \rangle \mid S \in \mathbb{S}(n)\}$  and  $\{\langle R_i, C \rangle \mid 1 \leq i \leq n - 1\}$ .

Note that  $\langle R_n, C \rangle$  is not included, i.e., only  $n - 1$  terms involving  $R_i$  are needed. Here, we have chosen to exclude index  $n$  without loss of generality.

**Proof of Theorem 1.** Our strategy to prove this involves calculating the dimension of the involved subspaces. First, we prove the equation

$$\text{span}\{S\}_{S \in \mathbb{S}(n)} \cap \text{span}\{R_i\}_{1 \leq i \leq n-1} = \{O_n\} \quad (7)$$

where  $O_n$  denotes the  $n \times n$  zero matrix. Then, we focus on the following equation to check linear independence. Here, we number all pairings such as  $S_1, S_2, \dots, S_u \dots S_{(N-1)!!}$ . We introduce the coefficients  $a_u$  and  $b_v$  and calculate the overlap of the spans:

$$1 \leq u \leq (n - 1)!!, a_u \in \mathbb{R},$$

$$1 \leq v \leq n - 1, b_v \in \mathbb{R},$$

$$\sum_{u=1}^{(n-1)!!} a_u S_u = \sum_{v=1}^{n-1} b_v R_v. \quad (8)$$

We focus on the summation of the  $k$ th-column on both sides. Note that for every  $S_u$  there is exactly one non-zero element in column  $k$ , while for  $R_v$  there may be more than one if

$v = k$  and  $1 \leq k \leq n - 1$ , or exactly one non-zero element otherwise. Then, the following equations hold:

When  $1 \leq k \leq n - 1$

$$(n - 2)b_k + \sum_{l=1}^{n-1} b_l - \sum_{l=1}^{(n-1)!!} a_l = 0. \tag{9}$$

When  $k = n$  (because of our choice in formulating Theorem 1)

$$\sum_{l=1}^{n-1} b_l - \sum_{l=1}^{(n-1)!!} a_l = 0. \tag{10}$$

With Equations (9) and (10),  $b_k = 0$  ( $1 \leq k \leq n - 1$ ) holds. This means that

$$\text{span}\{S\}_{S \in \mathbb{S}(n)} \cap \text{span}\{R_i\}_{1 \leq i \leq n-1} = \{O_n\}, \tag{11}$$

$$\dim \text{span}\{R_i\}_{1 \leq i \leq n-1} = n - 1. \tag{12}$$

By our previous study [17],

$$\dim \text{span}\{S\}_{S \in \mathbb{S}(n)} = L_{\min}(n). \tag{13}$$

Here, we denote  $L_{\min}(n) \equiv (n - 1)(n - 2)/2$ . By Equations (12) and (13), the following equation holds:

$$\dim \text{span}\{S\}_{S \in \mathbb{S}(n)} + \dim \text{span}\{R_i\}_{1 \leq i \leq n-1} = \dim \Omega_n. \tag{14}$$

Therefore, by Equations (11) and (14),

$$\dim \text{span}\{S\}_{S \in \mathbb{S}(n)} \cup \text{span}\{R_i\}_{1 \leq i \leq n-1} = \dim \Omega_n. \tag{15}$$

The pairing matrices  $S$  are a subset of  $\Omega_n$ . In addition, the adjacent set matrices  $R_i$  are also a subset of  $\Omega_n$ . Therefore, the following equation holds:

$$\text{span}\{S\}_{S \in \mathbb{S}(n)} \cup \text{span}\{R_i\}_{1 \leq i \leq n-1} \subseteq \Omega_n. \tag{16}$$

With Equations (15) and (16),

$$\text{span}\{S\}_{S \in \mathbb{S}(n)} \cup \text{span}\{R_i\}_{1 \leq i \leq n-1} = \Omega_n. \tag{17}$$

That is,  $\{S\}_{S \in \mathbb{S}(n)}$  plus  $\{R_i\}_{1 \leq i \leq n-1}$  can construct  $\Omega_n$ . Finally,  $\langle S, C \rangle$  is a linear transformation of  $S$  which comes from the property of the Frobenius inner product. Therefore,  $C \in \Omega_n$  can be constructed as a linear combination of  $\{\langle S, C \rangle \mid S \in \mathbb{S}(n)\}$  and  $\{\langle R_i, C \rangle \mid 1 \leq i \leq n - 1\}$ . Therefore, the theorem holds.  $\square$

**Corollary 1.**

$$\begin{aligned} &A, B \in \Omega_n, \\ &A = B \quad \text{if and only if} \\ &\forall S \in \mathbb{S}(n), \\ &\langle S, A \rangle = \langle S, B \rangle \text{ and } 1 \leq i \leq n, \langle R_i, A \rangle = \langle R_i, B \rangle. \end{aligned} \tag{18}$$

This corollary is a special case of Theorem 1 because Equation (18) means that  $A$  and  $B$  have the same total compatibilities for all pairings and all adjacent sets.

Here, we present an example for Theorem 1 for the  $n = 4$  case to illustrate the relationship of the involved subspaces. We define the following  $H_i$ :

$$H_i = \begin{cases} \text{span}\{S\}_{S \in \mathbb{S}(n)} & \text{if } i = 0, \\ \text{span}\{R_i\} & \text{if } 1 \leq i \leq n - 1. \end{cases} \tag{19}$$

We represent  $H_i$  as follows, where  $D_{i,j} \in \Omega_n$  is defined as the  $n \times n$  matrix whose  $(i, j)$ th element is 1 and all other elements are 0:

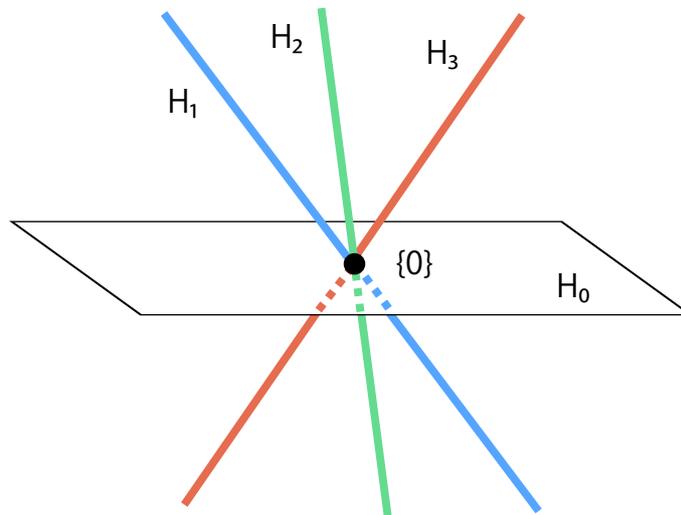
$$H_i = \begin{cases} \text{if } i = 0, \\ \{k_1(D_{1,2} + D_{3,4}) + k_2(D_{1,3} + D_{2,4}) + k_3(D_{1,4} + D_{2,3}) \mid k_1, k_2, k_3 \in \mathbb{R}\} \\ \text{if } i = 1, \\ \{k_4(D_{1,2} + D_{1,3} + D_{1,4}) \mid k_4 \in \mathbb{R}\} \\ \text{if } i = 2, \\ \{k_5(D_{2,1} + D_{2,3} + D_{2,4}) \mid k_5 \in \mathbb{R}\} \\ \text{if } i = 3, \\ \{k_6(D_{3,1} + D_{3,2} + D_{3,4}) \mid k_6 \in \mathbb{R}\}, \end{cases} \tag{20}$$

$$\tilde{H} = \{l_{i,j}D_{i,j} \mid 1 \leq i < j \leq n, l_{i,j} \in \mathbb{R}\}. \tag{21}$$

The image of these spaces is represented in Figure 1. That is,

$$0 \leq i < j \leq n - 1, i \neq j, H_i \cap H_j = \{O_n\}, \tag{22}$$

$$\tilde{H} = H_0 \cup H_1 \cup H_2 \cup H_3. \tag{23}$$



**Figure 1.** A schematic illustration of the relationship among  $H_0, H_1, H_2$  and  $H_3$ .

### 3.2. Equivalence Class

We define the relation  $\sim$  as follows:

$$A, B \in \Omega_n, \\ A \sim B \text{ if and only if } \forall S \in \mathbb{S}(n), \langle S, A \rangle = \langle S, B \rangle. \tag{24}$$

This represents an equivalence relationship between  $A$  and  $B$ , leading to the construction of an equivalence class.

Regarding this equivalence class, the following theorem holds:

**Theorem 2.**

$$\begin{aligned}
 &A, B \in \Omega_n, \\
 &A \sim B \text{ if and only if} \\
 &\forall \{i, j\} \in \mathbb{P}(n), \\
 &A_{i,j} - \frac{1}{n-2} (\langle R_i, A \rangle + \langle R_j, A \rangle) = B_{i,j} - \frac{1}{n-2} (\langle R_i, B \rangle + \langle R_j, B \rangle). \tag{25}
 \end{aligned}$$

That is, for any matrix  $C$  in the equivalence class, the values given by the following are conserved.

$$\forall \{i, j\} \in \mathbb{P}(n), C_{i,j} - \frac{1}{n-2} (\langle R_i, C \rangle + \langle R_j, C \rangle). \tag{26}$$

The matrix form of the conserved values is described in Appendix A.

**Proof of Theorem 2.** First, we prove sufficiency. We assume that the following equation holds:

$$\forall \{i, j\} \in \mathbb{P}(n), A_{i,j} - \frac{1}{n-2} (\langle R_i, A \rangle + \langle R_j, A \rangle) = B_{i,j} - \frac{1}{n-2} (\langle R_i, B \rangle + \langle R_j, B \rangle). \tag{27}$$

With Equation (27), the following equation holds:

$$\begin{aligned}
 &\sum_{\{i,j\} \in \mathbb{P}(n)} \left\{ A_{i,j} - \frac{1}{n-2} (\langle R_i, A \rangle + \langle R_j, A \rangle) \right\} \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} \left\{ B_{i,j} - \frac{1}{n-2} (\langle R_i, B \rangle + \langle R_j, B \rangle) \right\}. \tag{28}
 \end{aligned}$$

Here, the left side can be calculated as follows because the number of pairs including element  $k$  in  $\mathbb{P}(n)$  is  $n - 1$ :

$$\begin{aligned}
 &\sum_{\{i,j\} \in \mathbb{P}(n)} \left\{ A_{i,j} - \frac{1}{n-2} (\langle R_i, A \rangle + \langle R_j, A \rangle) \right\} \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j} - \frac{n-1}{n-2} \sum_{k=1}^n \langle R_k, A \rangle \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j} - \frac{n-1}{n-2} \sum_{k=1}^n \sum_{l \neq k} A_{k,l} \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j} - \frac{2(n-1)}{n-2} \sum_{\{k,l\} \in \mathbb{P}(n)} A_{k,l} \\
 &= -\frac{n}{n-2} \sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j}. \tag{29}
 \end{aligned}$$

Using Equation (29), Equation (28) is transformed into the following:

$$-\frac{n}{n-2} \sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j} = -\frac{n}{n-2} \sum_{\{i,j\} \in \mathbb{P}(n)} B_{i,j}. \tag{30}$$

Therefore,

$$\sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j} = \sum_{\{i,j\} \in \mathbb{P}(n)} B_{i,j}. \tag{31}$$

The following equation holds for any pairing  $S$  by Equation (27):

$$\begin{aligned} & \sum_{\{i,j\} \in f_{\text{set}}(S)} \left\{ A_{i,j} - \frac{1}{n-2} (\langle R_i, A \rangle + \langle R_j, A \rangle) \right\} \\ &= \sum_{\{i,j\} \in f_{\text{set}}(S)} \left\{ B_{i,j} - \frac{1}{n-2} (\langle R_i, B \rangle + \langle R_j, B \rangle) \right\}. \end{aligned} \tag{32}$$

Here, the following equation holds. Note that  $\{i, j\}$  belongs to  $f_{\text{set}}(S)$ ; hence,  $\langle R_k, A \rangle$  appears only once and all indexes  $k$  ranging from 1 to  $n$  appear over the summation:

$$\sum_{\{i,j\} \in f_{\text{set}}(S)} (\langle R_i, A \rangle + \langle R_j, A \rangle) = \sum_{k=1}^n \langle R_k, A \rangle \tag{33}$$

$$\begin{aligned} &= \sum_{k=1}^n \sum_{l, l \neq k} A_{k,l} \\ &= 2 \sum_{\{k,l\} \in \mathbb{P}(n)} A_{k,l}. \end{aligned} \tag{34}$$

For  $B$ , the following equation also holds:

$$\sum_{\{i,j\} \in f_{\text{set}}(S)} (\langle R_i, B \rangle + \langle R_j, B \rangle) = \sum_{k=1}^n \langle R_k, B \rangle \tag{35}$$

$$= 2 \sum_{\{k,l\} \in \mathbb{P}(n)} B_{k,l}. \tag{36}$$

Using these transformations, Equation (32) is transformed as follows:

$$\langle S, A \rangle - \frac{2}{n-2} \sum_{\{k,l\} \in \mathbb{P}(n)} A_{k,l} = \langle S, B \rangle - \frac{2}{n-2} \sum_{\{k,l\} \in \mathbb{P}(n)} B_{k,l}. \tag{37}$$

With Equation (31),

$$\langle S, A \rangle = \langle S, B \rangle. \tag{38}$$

Then,  $A \sim B$  holds.

Second, we prove the necessity. We assume that  $A \sim B$  holds. We define  $A^* \in \Omega_n$  as follows:

$$A^*_{i,j} \equiv \frac{1}{n-2} (\langle R_i, A \rangle + \langle R_j, A \rangle) + B_{i,j} - \frac{1}{n-2} (\langle R_i, B \rangle + \langle R_j, B \rangle). \tag{39}$$

By Equations (33), (35) and (39),

$$\begin{aligned} \forall S \in \mathbb{S}(n), \langle S, A^* \rangle &= \sum_{\{i,j\} \in f_{\text{set}}(S)} A^*_{i,j} \\ &= \langle S, B \rangle + \frac{1}{n-2} \sum_{i=1}^n \langle R_i, A \rangle - \frac{1}{n-2} \sum_{i=1}^n \langle R_i, B \rangle. \end{aligned} \tag{40}$$

We derive the relationship between  $\sum_{i=1}^n \langle R_i, A \rangle$  and  $\sum_{S \in \mathbb{S}(n)} \langle S, A \rangle$  here in order to transform Equation (40). By Equation (34),

$$\sum_{i=1}^n \langle R_i, A \rangle = 2 \sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j}. \tag{41}$$

For  $\sum_{S \in \mathbb{S}(n)} \langle S, A \rangle$ , we focus on the fact that the number of appearances of  $A_{i,j}$  is  $(n - 3)!!$ ,

$$\sum_{S \in \mathbb{S}(n)} \langle S, A \rangle = (n - 3)!! \sum_{\{i,j\} \in \mathbb{P}(n)} A_{i,j}. \tag{42}$$

With Equations (41) and (42), the following relationship holds:

$$\sum_{i=1}^n \langle R_i, A \rangle = \frac{2}{(n - 3)!!} \sum_{S \in \mathbb{S}(n)} \langle S, A \rangle. \tag{43}$$

Therefore, the following holds by  $A \sim B$  and Equation (43):

$$\begin{aligned} \sum_{i=1}^n \langle R_i, A \rangle &= \frac{2}{(n - 3)!!} \sum_{S \in \mathbb{S}(n)} \langle S, A \rangle \\ &= \frac{2}{(n - 3)!!} \sum_{S \in \mathbb{S}(n)} \langle S, B \rangle \\ &= \sum_{i=1}^n \langle R_i, B \rangle. \end{aligned} \tag{44}$$

By Equation (44), we can cancel the second and third terms of (40),

$$\langle S, A^* \rangle = \langle S, B \rangle. \tag{45}$$

In addition,  $A \sim B$  holds. Therefore,

$$\forall S \in \mathbb{S}(n), \langle S, A^* \rangle = \langle S, B \rangle = \langle S, A \rangle. \tag{46}$$

Additionally, the following also holds by  $A \sim B$  and Equation (44):

$$\begin{aligned} \sum_{j \neq i} A_{i,j}^* &= \frac{n - 1}{n - 2} \langle R_i, A \rangle + \frac{1}{n - 2} \sum_{j \neq i} \langle R_j, A \rangle + \sum_{j \neq i} B_{i,j} \\ &\quad - \frac{n - 1}{n - 2} \langle R_i, B \rangle - \frac{1}{n - 2} \sum_{j \neq i} \langle R_j, B \rangle \\ &= \frac{1}{n - 2} \left( \sum_{j=1}^n \langle R_j, A \rangle - \sum_{j=1}^n \langle R_j, B \rangle \right) + \langle R_i, A \rangle \\ &= \langle R_i, A \rangle. \end{aligned} \tag{47}$$

By Equation (47),

$$1 \leq i \leq n, \langle R_i, A^* \rangle = \langle R_i, A \rangle. \tag{48}$$

Therefore, by Equations (46) and (48) and Corollary 1,

$$A = A^* \tag{49}$$

is valid. That is to say, the following equation holds:

$$\{i, j\} \in \mathbb{P}(n), A_{i,j} - \frac{1}{n - 2} (\langle R_i, A \rangle + \langle R_j, A \rangle) = B_{i,j} - \frac{1}{n - 2} (\langle R_i, B \rangle + \langle R_j, B \rangle). \tag{50}$$

□

### 3.3. Mean and Covariance

Here, we analyze statistical properties associated with the compatibility matrix and the total compatibility.

We define the mean values of compatibilities and total compatibilities as

$$\begin{aligned}
 C &\in \Omega_n, \\
 \mu_{\text{element}}(C) &\equiv \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} C_{i,j}, \\
 \mu_{\text{sum}}(C) &\equiv \frac{1}{(n-1)!!} \sum_{S \in \mathbb{S}(n)} \langle S, C \rangle.
 \end{aligned}$$

By Equation (42),  $\mu_{\text{sum}}(C)$  is transformed into

$$\begin{aligned}
 \mu_{\text{sum}}(C) &\equiv \frac{1}{(n-1)!!} \sum_{S \in \mathbb{S}(n)} \langle S, C \rangle \\
 &= \frac{1}{n-1} \sum_{1 \leq i < j \leq n} C_{i,j} \\
 &= \frac{n}{2} \mu_{\text{element}}(C)
 \end{aligned} \tag{51}$$

where  $\mu_{\text{element}}(C)$  indicates the mean value of the elements of the compatibility matrix  $C$  and  $\mu_{\text{sum}}(C)$  is the mean of the total compatibility across all possible pairing with respect to the compatibility matrix  $C$ .

We define the square root of the covariance values for compatibilities and total compatibilities as follows:

$$\begin{aligned}
 \sigma_{\text{element}}(A, B) &\equiv \sqrt{\sum_{1 \leq i < j \leq n} \frac{2}{n(n-1)} (A_{i,j} - \mu_{\text{element}}(A))(B_{i,j} - \mu_{\text{element}}(B))}, \\
 \sigma_{\text{sum}}(A, B) &\equiv \sqrt{\frac{1}{(n-1)!!} \sum_{S \in \mathbb{S}(n)} (\langle S, A \rangle - \mu_{\text{sum}}(A))(\langle S, B \rangle - \mu_{\text{sum}}(B))}.
 \end{aligned} \tag{52}$$

Clearly,  $\sigma_{\text{element}}^2(C, C)$  and  $\sigma_{\text{sum}}^2(C, C)$  are variance values for compatibilities and total compatibilities when the compatibility matrix is  $C$ .

Regarding  $\sigma_{\text{sum}}^2(C, C)$ , the following theorem holds.

**Theorem 3.** Let  $I_n$  be the  $n \times n$  identity matrix,  $J_n$  the  $n \times n$  matrix where all elements are 1, and  $C \in \Omega_n$ ,  $\hat{C} \equiv C - \mu_{\text{element}}(C)(J_n - I_n)$ . Then, the following equation holds:

$$\sigma_{\text{sum}}^2(C, C) = \frac{n(n-2)}{2(n-3)} \sigma_{\text{element}}^2(C, C) - \frac{1}{(n-1)(n-3)} \sum_{k=1}^n \langle R_k, \hat{C} \rangle^2. \tag{53}$$

**Proof of Theorem 3.** By definition,

$$\sigma_{\text{sum}}^2(C, C) = \frac{1}{(n-1)!!} \sum_{S \in \mathbb{S}(n)} \{ \langle S, C \rangle - \mu_{\text{sum}}(C) \}^2$$

Using Equation (51),

$$\sigma_{\text{sum}}^2(C, C) = \frac{1}{(n-1)!!} \sum_{S \in \mathbb{S}(n)} \left\{ \langle S, C \rangle - \frac{n}{2} \mu_{\text{element}}(C) \right\}^2 \tag{54}$$

Here, the following equation holds:

$$\begin{aligned}
 \langle S, \hat{C} \rangle &= \frac{1}{2} \langle S, \hat{C} \rangle_F \\
 &= \frac{1}{2} \langle S, C \rangle_F - \frac{1}{2} \mu_{\text{element}}(C) \langle S, J_n - I_n \rangle \\
 &= \frac{1}{2} \langle S, C \rangle_F - \frac{n}{2} \mu_{\text{element}}(C) \\
 &= \langle S, C \rangle - \frac{n}{2} \mu_{\text{element}}(C)
 \end{aligned} \tag{55}$$

Therefore, by Equations (54) and (55),

$$\begin{aligned}
 \sigma_{\text{sum}}^2(C, C) &= \frac{1}{(n-1)!!} \sum_{S \in \mathbb{S}(n)} \left\{ \langle S, C \rangle - \frac{n}{2} \mu_{\text{element}}(C) \right\}^2 \\
 &= \frac{1}{(n-1)!!} \sum_{S \in \mathbb{S}(n)} \langle S, \hat{C} \rangle^2 \\
 &= \frac{1}{(n-1)!!} \cdot (n-3)!! \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j}^2 \\
 &\quad + \frac{1}{(n-1)!!} \cdot (n-5)!! \sum_{\{i,j\} \in \mathbb{P}(n)} \sum_{\substack{\{k,l\} \in \mathbb{P}(n) \\ \{k,l\} \cap \{i,j\} = \emptyset}} \hat{C}_{i,j} \hat{C}_{k,l} \\
 &= \frac{1}{n-1} \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j}^2 + \frac{1}{(n-1)(n-3)} \sum_{\{i,j\} \in \mathbb{P}(n)} \sum_{\substack{\{k,l\} \in \mathbb{P}(n) \\ \{k,l\} \cap \{i,j\} = \emptyset}} \hat{C}_{i,j} \hat{C}_{k,l} \\
 &= \frac{1}{n-1} \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j}^2 + \frac{1}{(n-1)(n-3)} \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j} \sum_{\substack{\{k,l\} \in \mathbb{P}(n) \\ \{k,l\} \cap \{i,j\} = \emptyset}} \hat{C}_{k,l}.
 \end{aligned} \tag{56}$$

Here, we focus on  $\sum_{\substack{\{k,l\} \in \mathbb{P}(n) \\ \{k,l\} \cap \{i,j\} = \emptyset}} \hat{C}_{k,l}$ . This term is transformed as follows:

$$\begin{aligned}
 \sum_{\substack{\{k,l\} \in \mathbb{P}(n) \\ \{k,l\} \cap \{i,j\} = \emptyset}} \hat{C}_{k,l} &= \hat{C}_{i,j} + \sum_{\{k,l\} \in \mathbb{P}(n)} \hat{C}_{k,l} - \sum_{k,k \neq i} \hat{C}_{i,k} - \sum_{k,k \neq j} \hat{C}_{j,k} \\
 &= \hat{C}_{i,j} - \langle R_i, \hat{C} \rangle - \langle R_j, \hat{C} \rangle + \sum_{\{k,l\} \in \mathbb{P}(n)} \hat{C}_{k,l} \\
 &= \hat{C}_{i,j} - \langle R_i, \hat{C} \rangle - \langle R_j, \hat{C} \rangle + \sum_{\{k,l\} \in \mathbb{P}(n)} (C_{k,l} - \mu_{\text{element}}(C)) \\
 &= \hat{C}_{i,j} - \langle R_i, \hat{C} \rangle - \langle R_j, \hat{C} \rangle + \left( \sum_{\{k,l\} \in \mathbb{P}(n)} C_{k,l} \right) - \frac{n(n-1)}{2} \mu_{\text{element}}(C) \\
 &= \hat{C}_{i,j} - \langle R_i, \hat{C} \rangle - \langle R_j, \hat{C} \rangle.
 \end{aligned} \tag{57}$$

Then, using this formula,

$$\begin{aligned}
 & \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j} \sum_{\substack{\{k,l\} \in \mathbb{P}(n) \\ \{k,l\} \neq \{i,j\}}} \hat{C}_{k,l} \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j} (\hat{C}_{i,j} - \langle R_i, \hat{C} \rangle - \langle R_j, \hat{C} \rangle) \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j}^2 - \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j} (\langle R_i, \hat{C} \rangle + \langle R_j, \hat{C} \rangle) \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j}^2 - \sum_{i=1}^n \sum_{j \neq i} \hat{C}_{i,j} \langle R_i, \hat{C} \rangle \\
 &= \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j}^2 - \sum_{i=1}^n \langle R_i, \hat{C} \rangle^2. \tag{58}
 \end{aligned}$$

By Equations (56) and (58), the following equation holds:

$$\begin{aligned}
 \sigma_{\text{sum}}^2(C, C) &= \frac{n-2}{(n-1)(n-3)} \sum_{\{i,j\} \in \mathbb{P}(n)} \hat{C}_{i,j}^2 - \frac{1}{(n-1)(n-3)} \sum_{k=1}^n \langle R_k, \hat{C} \rangle^2 \\
 &= \frac{n(n-2)}{2(n-3)} \sigma_{\text{element}}^2(C, C) - \frac{1}{(n-1)(n-3)} \sum_{k=1}^n \langle R_k, \hat{C} \rangle^2. \tag{59}
 \end{aligned}$$

Therefore, the theorem holds. □

#### 4. Variance Optimization

This section examines the performance enhancement from deriving a pairing that yields higher total compatibility by exploiting the algebraic structures identified in the previous section. We first show that the variance of the elements in a compatibility matrix affects the performance of the heuristic algorithm proposed in our previous study. Then we propose the transformation of a compatibility matrix to another one that minimizes the variance while ensuring that the total compatibility is maintained.

##### 4.1. Performance Degradation through the Observation Phase

In our previous study [17], we proposed an algorithm for recognizing the compatibilities among elements through multiple measurements of total compatibility. To summarize, we estimate the compatibility matrix denoted by  $\tilde{C} \in \Omega_n$ , which is given by

$$\begin{aligned}
 & C \in \Omega_n, \\
 & \tilde{C}_{i,j} = \begin{cases} 0 & \text{if } 1 \in \{i, j\} \\ C_{i,j} - C_{1,i} - C_{1,j} + \frac{2}{n-2} \sum_{k=2}^n C_{1,k} & \text{otherwise.} \end{cases} \tag{60}
 \end{aligned}$$

This  $\tilde{C} \in \Omega_n$  is one of the elements in the equivalence class. That is,  $C \sim \tilde{C}$  holds. By this property and Equation (60), the dimension of  $\{S\}_{S \in \mathbb{S}(n)}$  is given by  $(n-1)(n-2)/2$ , which we refer to as  $L_{\text{min}}(n)$ . This means that the number of observations required to grasp the compatibilities through an observation phase is  $L_{\text{min}}(n)$ .

Indeed, our previous study proposed an observation algorithm which needs  $\mathcal{O}(n^2)$  measurements. We have also confirmed numerically that the observation strategy provides a compatibility matrix, which is in the equivalence class of the ground-truth compatibility matrix  $C^g$ . In the numerical studies, the elements of the ground-truth compatibility matrix,  $C_{i,j}^g$ , were specified by uniformly distributed random numbers in the range of  $[0, 1]$ .

However, finding a pairing yielding a greater total compatibility becomes difficult based on  $C^e$ , including the above-mentioned  $\tilde{C}$ , even though  $C^e$  is in the equivalence class where the ground-truth compatibility  $C^g$  is included. In searching for a better pairing,

we use a heuristic algorithm, which is named Pairing-2-opt [17]. We consider the difficulty comes from the fact that the variance of the elements of the compatibility matrix  $\sigma_{\text{element}}^2(C^e, C^e)$  would be larger than those of  $\sigma_{\text{element}}^2(C^g, C^g)$ , which is highly likely to cause the combining algorithm to become stuck in a local minimum.

Hence, our idea is to find a compatibility matrix  $X$  which is in the same equivalence class of matrix  $C$

$$\forall S \in \mathbb{S}(n), \langle S, X \rangle = \langle S, C \rangle \tag{61}$$

while simultaneously minimizing the variance of the elements of  $\sigma_{\text{element}}^2(X, X)$ .

#### 4.2. Transforming the Compatibility Matrix with Minimized Variance

We solve the following optimization problem:

$$\begin{aligned} \min : & \quad \sigma_{\text{element}}^2(X, X), \\ \text{subject to :} & \quad X, C \in \Omega_n, C \text{ is fixed,} \\ & \quad X \sim C. \end{aligned} \tag{62}$$

By Theorem 3 and  $\sigma_{\text{sum}}^2(X, X) = \sigma_{\text{sum}}^2(C, C)$ , we transform this problem into the following form:

$$\begin{aligned} \min : & \quad \sum_{k=1}^n \langle R_k, \hat{X} \rangle^2, \\ \text{subject to :} & \quad X, C \in \Omega_n, C \text{ is fixed,} \\ & \quad X \sim C, \\ & \quad \hat{X} \equiv X - \mu_{\text{element}}(X)(J_n - I_n). \end{aligned} \tag{63}$$

The optimal solution for this problem holds because the sum of squares is minimized when all values are 0:

$$1 \leq k \leq n, \langle R_k, \hat{X} \rangle = 0. \tag{64}$$

Hence, the following equation is derived:

$$1 \leq k \leq n, \langle R_k, X \rangle = (n - 1)\mu_{\text{element}}(C). \tag{65}$$

By Equation (65) and Theorem 2, the optimal solution is represented as follows:

$$X_{i,j} = \frac{2(n - 1)}{n - 2}\mu_{\text{element}}(C) + C_{i,j} - \frac{1}{n - 2}(\langle R_i, C \rangle + \langle R_j, C \rangle). \tag{66}$$

Thus, the compatibility matrix with minimal variance is derived. In addition, this discussion and solution mean that the optimal-variance solution is unique with respect to the equivalence class.

### 5. Simulation

In this section, we evaluate the performance of the proposed method on the pairing optimization problem. There are two important points that should be clarified through the simulations. One is to quantitatively evaluate the performance reduction of the combining algorithm proposed in the previous study, based on the observation phase. The other is to demonstrate the performance enhancement due to the variance optimization discussed in Section 4.

### 5.1. Setting

We configure the ground-truth compatibility matrix  $C^g \in \Omega_n$  with two different distributions. The first is the uniform distribution:

$$\forall \{i, j\} \in \mathbb{P}, C_{i,j}^g \sim U(0, 1). \tag{67}$$

Here, we denote the uniform distribution between 0 and 1 as  $U(0, 1)$ . The second distribution is the Poisson distribution:

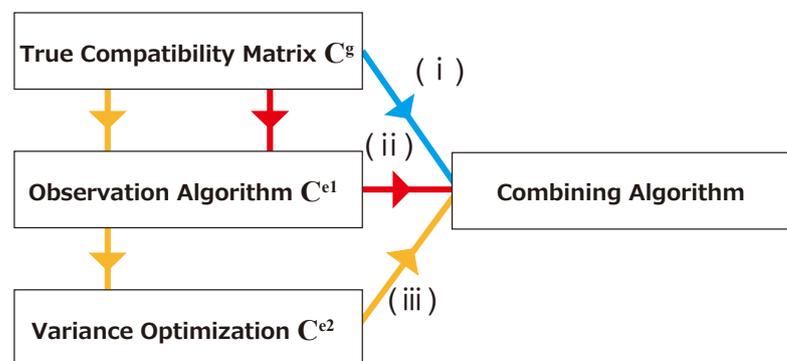
$$\forall \{i, j\} \in \mathbb{P}, C_{i,j}^g \sim \text{Poisson}(1). \tag{68}$$

Here, we denote the Poisson distribution whose mean is  $\lambda$  as  $\text{Poisson}(\lambda)$ . In the numerical simulation, the number of elements in the system  $n$  varied from 100 to 1000 in intervals of 100. For each  $n$ , we conducted 100 trials with different randomly generated ground-truth compatibility matrices  $C^g$  based on Equation (67) or Equation (68). We quantified the performance for each derived pairing  $S \in \mathbb{S}(n)$  by  $2\langle S, C^g \rangle / n$  and evaluated its average over 100 trials for each value of  $n$ .

### 5.2. Simulation Flow

The ground-truth compatibility matrix  $C^g$  is transformed into  $C^{e1}$  by the observation algorithm based on Equation (60). The variance optimization transforms  $C^{e1}$  into  $C^{e2}$ . The combining algorithm, which is called PNN+p2-opt [17], yields a pairing with the intention of achieving higher total compatibility. The exchange limit  $l$  is an internal parameter in PNN+p2-opt. This determines the number of maximum trials, and is set to 600 in the present study.

We evaluated the performance on the basis of  $C^g$ ,  $C^{e1}$ , and  $C^{e2}$ , as shown in flows (i), (ii), and (iii), respectively, in Figure 2.

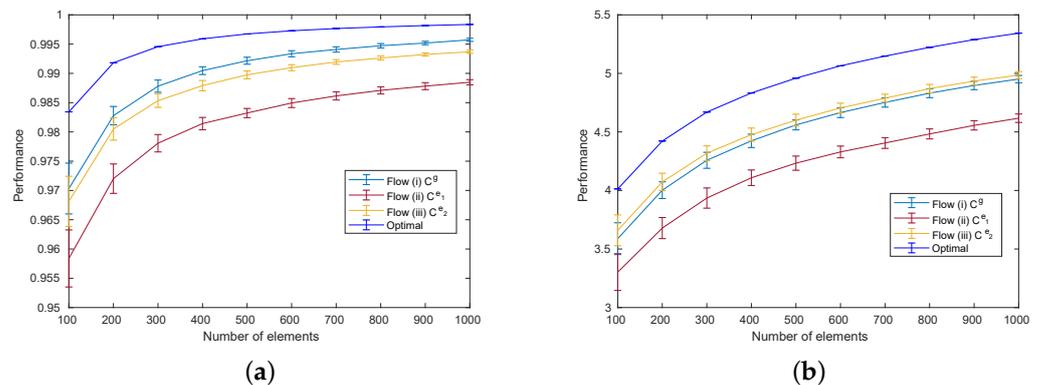


**Figure 2.** Schematic illustration of the three heuristic pairing optimization algorithms tested in the simulation. Case (i) (blue) applies the combining algorithm directly to the ground-truth compatibility matrix  $C^g$ . Case (ii) (red) first applies the observation algorithm to obtain an estimated compatibility matrix  $C^{e1}$ , followed by the combining algorithm. Case (iii) (yellow) first estimates the compatibility from observation ( $C^{e1}$ ), followed by the variance optimization ( $C^{e2}$ ), and then executes the combining algorithm.

### 5.3. Performance

The blue, red, and yellow curves in Figure 3 demonstrate the performance of cases (i), (ii), and (iii), respectively, as a function of the number of elements for the uniform distribution (Figure 3a) and the Poisson distribution (Figure 3b). For the uniform distributed ground-truth we observe that the performance of case (ii) is inferior to that of case (i), demonstrating the performance degradation by the transformation from  $C^g$  to  $C^{e1}$  through observation. Furthermore, the performance of case (iii) is enhanced compared with that of case (ii), which confirms the performance gain from variance optimization. The results

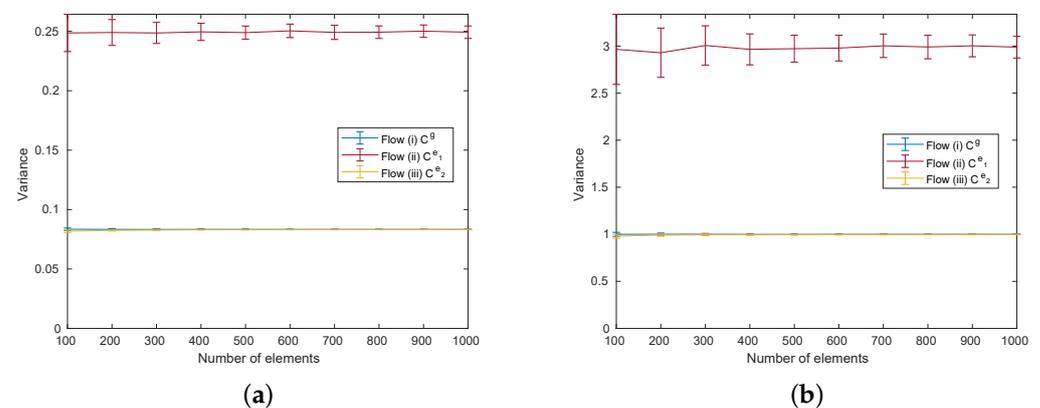
differ for the Poisson distribution. Here, the performance of case (iii) is higher than case (i). That is, for the Poisson case the variance optimization (Flow (iii)) not only counteracted the performance loss of the observation algorithm (Flow (ii)), but actually enhanced the performance compared to the ground truth matrix  $C^g$  (Flow (i)). Further numerical tests revealed that the relationship of performances for a Gaussian distribution are similar to those for the uniform distribution. Conversely, the performance for a binary distribution hardly differed between any of the algorithms.



**Figure 3.** Comparison of the achieved total compatibility for Flows (i), (ii), and (iii), as described in the caption for Figure 2. Each graph shows the mean and standard deviation of the performance of 100 different compatibility matrices with each given number of elements, simulated under (a) uniform distributions and (b) Poisson distributions.

The variance of  $C^g$ ,  $C^{e1}$ , and  $C^{e2}$  are evaluated as shown in Figure 4 as a function of the number of elements. We clearly observe that the variance of  $C^{e1}$  is higher than  $C^g$  while the variance of  $C^{e2}$  becomes comparable to the ground-truth case  $C^g$  for both the uniform and Poisson distributions.

From these numerical results, we can conclude that the variance optimization minimizes the variance and enhances the performance of the achieved total compatibility. It is worth noting that the performance with the uniform distribution after variance optimization is still lower than the case based on the ground-truth matrix  $C^g$ , as observed in Figure 3a. This occurs because the variance optimization algorithm does not transform  $C^{e1}$  to the original compatibility matrix  $C^g$ . In other words, there exist additional factors that influence the performance of the combining algorithm that are related to the compatibility distribution. The distribution of the original compatibility  $C^g$  (uniform distribution) is seemingly beneficial for the performance of the heuristic combining algorithm, even when compared to the compatibility matrix with minimum variance  $C^{e2}$ .



**Figure 4.** Comparison of the variance of the compatibility matrices of  $C^g$ ,  $C^{e1}$ ,  $C^{e2}$  as a function of the number of elements in the system under (a) uniform distributions and (b) Poisson distributions.

## 6. Conclusions

One of the most challenging issues in the pairing problem is how to understand the underlying compatibilities among the elements under study. An accurate and efficient approach is essential for practical applications such as wireless communications and online social networks. This study reveals several algebraic structures in the pairing optimization problem.

We introduce an equivalence class in the compatibility matrices, containing matrices that yield the same total compatibility although the matrices themselves differ. This can also be expressed through a conserved value or invariance in the equivalence class. Based on such insights, we propose a transformation of the initially estimated compatibility matrix to another form that minimizes the variance of the elements. We demonstrate that the highest total compatibility found heuristically is improved significantly with the proposed transformation relative to the direct approach.

In the future, the proposed algorithm may be applied to bipartite matching and assignment problems, for example. Therefore, if the compatibility between elements that should not be paired is set to a negative value with a relatively large absolute value, we may solve the problem heuristically. Hence, the variance optimization proposed in this study may aid in performance enhancement.

**Author Contributions:** Conceptualization, N.F., M.H. and M.N.; methodology, N.F.; software, N.F.; validation, N.F., A.R. and T.M.; formal analysis, N.F. and A.R.; investigation, N.F., A.R., T.M., R.H., A.L., M.H. and M.N.; resources, T.M., R.H. and M.N.; data curation, N.F.; writing—original draft preparation, N.F., A.R. and M.N.; writing—review and editing, N.F., A.R., T.M., R.H., A.L., M.H. and M.N.; visualization, N.F.; supervision, M.N.; project administration, M.H. and M.N.; funding acquisition, M.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in part by the Japan Science and Technology Agency through the Core Research for Evolutionary Science and Technology (CREST) Project (JPMJCR17N2), and in part by the Japan Society for the Promotion of Science through the Grants-in-Aid for Scientific Research (A) (JP20H00233) and Transformative Research Areas (A) (JP22H05197). AR is a JSPS International Research Fellow.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** We would like to thank the editors of this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Matrix Form of Conserved Quantities

In Theorem 2, the following values are conserved in the same equivalence class.

$$\forall \{i, j\} \in \mathbb{P}(n), C_{i,j} - \frac{1}{n-2} (\langle R_i, C \rangle + \langle R_j, C \rangle). \tag{A1}$$

We can transform Equation (A1) into the following form using the Hadamard product  $\circ$ .

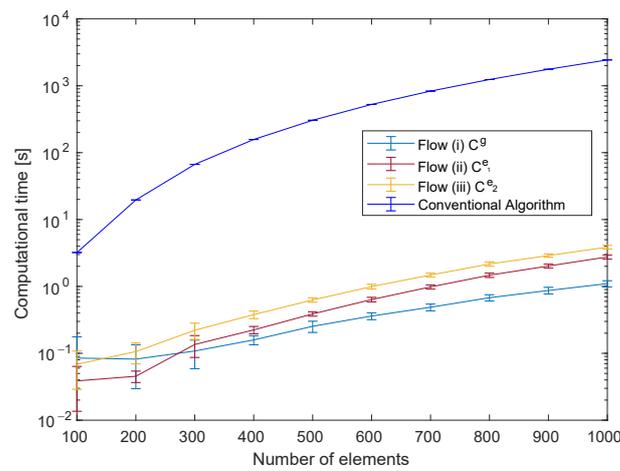
$$C - \frac{1}{n-2} (J_n - I_n) \circ (J_n C + C J_n). \tag{A2}$$

Therefore, the following equation holds:

$$A \sim B \quad \text{if and only if} \\ A - \frac{1}{n-2} (J_n - I_n) \circ (J_n A + A J_n) = B - \frac{1}{n-2} (J_n - I_n) \circ (J_n B + B J_n). \tag{A3}$$

## Appendix B. Computational Time

We compared the computational time of four different algorithms in the new Figure A1 in the new version of the manuscript. Three of them are for the cases from Figure 2. For cases (ii) and (iii), the computational time also includes the time needed for variance optimization. We compared them to a conventional MWM algorithm whose code (<https://jp.mathworks.com/matlabcentral/fileexchange/42827-weighted-maximum-matching-in-general-graphs> (accessed on 10 January 2023)) was developed and distributed by Daniel R. Saunders (<http://danielrsaunders.com> (accessed on 10 January 2023)). This conventional algorithm is based on “Efficient algorithms for finding maximum matching in graphs” by Zvi Galil [32]. The number of elements changes from 100 to 1000. One hundred different compatibility matrices were simulated and averaged to obtain the computational time.



**Figure A1.** Comparison of computational time between different four algorithms, which are case (i), (ii), (iii) and the conventional MWM algorithm.

Figure A1 shows that, as expected, the PNN+p2-opt algorithm is significantly faster than the conventional algorithm, and Flow (i), Flow (ii), Flow (iii) work faster in this order. These computational times can be explained as follows: First, PNN+p2-opt is heuristic and a  $\mathcal{O}(n^2)$  algorithm. Therefore, PNN+p2-opt is significantly faster than the conventional MWM algorithm that aims to find the absolute best solution. Second, we count the computational time, including the variance optimization procedure. The variance optimization takes some time, so the computational time of flow (ii) and flow (iii) is longer than flow (i). Third, flow (ii) has a tendency to become stuck in local minima, resulting in less computational time than flow (iii), due to the faster termination of the p2-opt algorithm.

In the future, the comparison to machine-learning-based methods such as the one proposed in Ref. [16] is of great interest. However, at this point, it is unclear how to conduct a fair comparison, as the ML-based algorithm requires extensive training on multiple examples before it is able to solve the problem. Nevertheless, as machine learning is a rapidly evolving field, it is possible that ML-based algorithms specialized for the pairing problem could be developed in the near future.

## References

1. Gale, D.; Shapley, L.S. College admissions and the stability of marriage. *Am. Math. Mon.* **1962**, *69*, 9–15. [[CrossRef](#)]
2. Roth, A.E. The economics of matching: Stability and incentives. *Math. Oper. Res.* **1982**, *7*, 617–628. [[CrossRef](#)]
3. Ergin, H.; Sönmez, T.; Ünver, M.U. Dual-Donor Organ Exchange. *Econometrica* **2017**, *85*, 1645–1671. [[CrossRef](#)]
4. Kohl, N.; Karisch, S.E. Airline crew rostering: Problem types, modeling, and optimization. *Ann. Oper. Res.* **2004**, *127*, 223–257. [[CrossRef](#)]
5. Gambetta, J.M.; Chow, J.M.; Steffen, M. Building logical qubits in a superconducting quantum computing system. *npj Quantum Inf.* **2017**, *3*, 1–7. [[CrossRef](#)]
6. Gao, Y.; Dai, Q.; Wang, M.; Zhang, N. 3D model retrieval using weighted bipartite graph matching. *Signal Process. Image Commun.* **2011**, *26*, 39–47. [[CrossRef](#)]

7. Bellur, U.; Kulkarni, R. Improved matchmaking algorithm for semantic web services based on bipartite graph matching. In Proceedings of the IEEE International Conference on Web Services (ICWS 2007), Salt Lake City, UT, USA, 9–13 July 2007; pp. 86–93.
8. Edmonds, J. Paths, trees, and flowers. *Can. J. Math.* **1965**, *17*, 449–467. [[CrossRef](#)]
9. Gabow, H.N. Data structures for weighted matching and nearest common ancestors with linking. In Proceedings of the First Annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco, CA, USA, 22–24 January 1990; pp. 434–443.
10. Huang, C.C.; Kavitha, T. Efficient algorithms for maximum weight matchings in general graphs with small edge weights. In Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, Kyoto, Japan, 17–19 January 2012; pp. 1400–1412.
11. Pettie, S. A simple reduction from maximum weight matching to maximum cardinality matching. *Inf. Process. Lett.* **2012**, *112*, 893–898. [[CrossRef](#)]
12. Cygan, M.; Gabow, H.N.; Sankowski, P. Algorithmic applications of baur-strassen’s theorem: Shortest cycles, diameter, and matchings. *J. ACM (JACM)* **2015**, *62*, 1–30. [[CrossRef](#)]
13. Duan, R.; Pettie, S. Approximating maximum weight matching in near-linear time. In Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science, Las Vegas, NV, USA, 23–26 October 2010; pp. 673–682.
14. Hanke, S.; Hougardy, S. New Approximation Algorithms for the Weighted Matching Problem. Citeseer. 2010. Available online: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=f6bc65fe193c8afd779f2867831a869c59554661> (accessed on 10 January 2023).
15. Duan, R.; Pettie, S. Linear-time approximation for maximum weight matching. *J. Acm (JACM)* **2014**, *61*, 1–23. [[CrossRef](#)]
16. Wu, B.; Li, L. Solving maximum weighted matching on large graphs with deep reinforcement learning. *Inf. Sci.* **2022**, *614*, 400–415. [[CrossRef](#)]
17. Fujita, N.; Chauvet, N.; Röhm, A.; Horisaki, R.; Li, A.; Hasegawa, M.; Naruse, M. Efficient Pairing in Unknown Environments: Minimal Observations and TSP-based Optimization. *IEEE Access* **2022**, *10*, 57630–57640. [[CrossRef](#)]
18. Williams, V.V. Multiplying matrices faster than Coppersmith-Winograd. In Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing, New York, NY, USA, 20–22 May 2012; pp. 887–898.
19. Kuhn, H.W. The Hungarian method for the assignment problem. *Nav. Res. Logist. Q.* **1955**, *2*, 83–97. [[CrossRef](#)]
20. Bertsekas, D.P. The auction algorithm: A distributed relaxation method for the assignment problem. *Ann. Oper. Res.* **1988**, *14*, 105–123. [[CrossRef](#)]
21. Munkres, J. Algorithms for the assignment and transportation problems. *J. Soc. Ind. Appl. Math.* **1957**, *5*, 32–38. [[CrossRef](#)]
22. Goldberg, A.; Radzik, T. *A Heuristic Improvement of the Bellman-Ford Algorithm*; Technical Report; STANFORD UNIV CA DEPT OF COMPUTER SCIENCE: Stanford, CA, USA, 1993.
23. Aldababsa, M.; Toka, M.; Gökçeli, S.; Kurt, G.K.; Kucur, O. A tutorial on nonorthogonal multiple access for 5G and beyond. *Wirel. Commun. Mob. Comput.* **2018**, *2018*, 9713450. [[CrossRef](#)]
24. Ding, Z.; Fan, P.; Poor, H.V. Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions. *IEEE Trans. Veh. Technol.* **2015**, *65*, 6010–6023. [[CrossRef](#)]
25. Chen, L.; Ma, L.; Xu, Y. Proportional fairness-based user pairing and power allocation algorithm for non-orthogonal multiple access system. *IEEE Access* **2019**, *7*, 19602–19615. [[CrossRef](#)]
26. Ali, Z.; Khan, W.U.; Ihsan, A.; Waqar, O.; Sidhu, G.A.S.; Kumar, N. Optimizing resource allocation for 6G NOMA-enabled cooperative vehicular networks. *IEEE Open J. Intell. Transp. Syst.* **2021**, *2*, 269–281. [[CrossRef](#)]
27. Zhang, H.; Duan, Y.; Long, K.; Leung, V.C. Energy efficient resource allocation in terahertz downlink NOMA systems. *IEEE Trans. Commun.* **2020**, *69*, 1375–1384. [[CrossRef](#)]
28. Shahab, M.B.; Irfan, M.; Kader, M.F.; Young Shin, S. User pairing schemes for capacity maximization in non-orthogonal multiple access systems. *Wirel. Commun. Mob. Comput.* **2016**, *16*, 2884–2894. [[CrossRef](#)]
29. Zhu, L.; Zhang, J.; Xiao, Z.; Cao, X.; Wu, D.O. Optimal user pairing for downlink non-orthogonal multiple access (NOMA). *IEEE Wirel. Commun. Lett.* **2018**, *8*, 328–331. [[CrossRef](#)]
30. Higuchi, K.; Benjebbour, A. Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access. *IEICE Trans. Commun.* **2015**, *98*, 403–414. [[CrossRef](#)]
31. Halim, A.H.; Ismail, I. Combinatorial optimization: Comparison of heuristic algorithms in travelling salesman problem. *Arch. Comput. Methods Eng.* **2019**, *26*, 367–380. [[CrossRef](#)]
32. Galil, Z. Efficient algorithms for finding maximum matching in graphs. *ACM Comput. Surv. (CSUR)* **1986**, *18*, 23–38. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.