Corvinus University of Budapest

Doctoral School of Economics, Business and Informatics

THESIS COLLECTION

Szádoczki Zsombor

Ph.D. dissertation titled

Preference modelling with a graph theoretic approach

Thesis supervisor:

Bozóki Sándor
professor

Budapest, 2024

Department of Operations Research and Actuarial Sciences

THESIS COLLECTION

Szádoczki Zsombor

Ph.D. dissertation titled

Preference modelling with a graph theoretic approach

Thesis supervisor:

Bozóki Sándor
professor

Contents

List of Figures

I. Research background and relevance of the chosen field

1. Research background

The doctoral thesis focuses on multicriteria decision making (MCDM) problems (Ishizaka and Nemery, 2013), especially methods connected to pairwise comparisons. The aim of MCDM is to select the best, the best few, or to provide a whole ranking of a finite set of alternatives based on a finite number of (usually conflicting) quantitative and/or qualitative criteria.

One of the most popular MCDM methodologies is the Analytic Hierarchy Process (AHP), proposed by Saaty (1977, 1980). It is based on a hierarchical system of criteria, subcriteria, etc., while uses pairwise comparison matrices (PCMs) to evaluate the alternatives according to each criterion separately and to determine the importance weights of the criteria as well. An element of a (multiplicative, ratio scale) PCM shows how many times the alternative (criterion) corresponding to the given row of the PCM is better/stronger/larger/more important than the alternative (criterion) corresponding to the given column of the matrix.

Besides decision modelling, pairwise comparisons are used in many other areas as well, e.g., preference measurement, ranking, sports, and psychometrics (Thurstone, 1927; Davidson and Farquhar, 1976; Csató, 2021). These types of comparisons are placed into a matrix in the case of a PCM. The main idea behind this process is that the decision makers cannot provide their preferences accurately for a complex problem; however, they can estimate their real preferences well between a pair of alternatives according to a single criterion.

The focus of our research is the case when some of the comparisons are missing, thus, we have to deal with a not complete data set, an incomplete pairwise comparison matrix (IPCM). Although we apply the decision modelling point of view throughout the dissertation, as both pairwise comparisons and missing data are common in many different research fields, our results can be useful in a much wider range.

There can be many reasons behind the incompleteness of a pairwise comparison matrix. Some data may have been lost, certain comparisons can be simply impossible, or the decision maker might have no time or willingness to provide all the comparisons, which is a lingering task.

Harker (1987) was among the first to propose IPCMs in order to reduce the number of questions asked from the decision maker in the Analytic Hierarchy Process. It is especially useful in the case of group decision making, when the choice should be made based on the preferences of several decision makers, and all of them have to fill in all PCMs.

If we are dealing with incomplete data, the result—i.e., the ranking of the alternatives calculated from the IPCM—is heavily dependent on the number of known elements and their arrangement. The latter one, the structure of the comparisons can be suitably handled by the representing graph of the IPCM (Gass, 1998). In the representing graph, the vertices correspond to the alternatives (criteria), while there is an edge between two vertices if and only if the comparison between the appropriate two alternatives (criteria) is known.

Although the literature of IPCMs is relatively limited compared to other areas connected to pairwise comparisons, there are many recent studies on theoretical results (Zhou et al., 2018; Kułakowski and Talaga, 2020; Szybowski et al., 2020; Ágoston and Csató, 2022), as well as applications (Bozóki et al., 2016).

A large portion of our analysis is centered around recommended filling in patterns for incomplete pairwise comparison matrices. What kind of designs of comparisons ensure that the computed results are close to the ones that would be calculated from the complete PCM? This and similar questions can be answered using graph theoretical properties of the representing graphs. The results are not just important from a theoretical point of view, but they can be easily applied in the practice of multicriteria decision making problems as well.

All of the research included in the Ph.D. thesis have been developed as academic journal articles. We compiled the following studies in the dissertation without re-editing, and present them in the form as they were—or are planned to be—published.

I. Szádoczki, Zs., Bozóki, S., and Tekile, H. A. (2022). Filling in pattern designs for incomplete pairwise comparison matrices: (Quasi-)regular graphs with minimal diameter. *Omega*, 107:102557. https://doi.org/10.1016/j.omega.2021.102557.

II. Szádoczki, Zs., Bozóki, S., Juhász, P., Kadenko, S. V., and Tsyganok, V. (2023). Incomplete pairwise comparison matrices based on graphs with average degree approximately 3. *Annals of Operations Research*, 326(2):783-807. https://doi.org/10.1007/s10479-022-04819-9.

III. Szádoczki, Zs., and Bozóki, S. (2023). Optimal sequences for pairwise comparisons: the graph of graphs approach. *Working paper.* https://doi.org/10.48550/arXiv.2205.08673.

IV. Temesi, J., **Szádoczki, Zs.** and Bozóki, S. (2024). Incomplete pairwise comparison matrices: Ranking top women tennis players. *Journal of the Operational Research Society*, 75(1):145-

157. https://doi.org/10.1080/01605682.2023.2180447.

2. Research frame

Our dissertation belongs to the same academic research group as Bozóki (2006), Csató (2015), Ábele-Nagy (2019), and Poesz (2022). As mentioned before, the main questions and studied problems are focused around the topic of the graph theoretic properties of representing graphs of incomplete pairwise comparison matrices. The relation of research questions, publications, and results of the studies included in the Ph.D. thesis can be seen in Figure 1.

Studies I., II., and III. are natural (and in some sense linear) continuations of each other. From a methodological point of view, all of them rely on different simulations. Study IV. uses the same tools from the literature of multicriteria decision making and graph theory as the previous publications, however, it focuses on the ranking aspect of pairwise comparisons instead of the question of the optimal filling in designs.

All research presented in this Ph.D. thesis started with a conjecture that the diameter (the longest shortest path) of the representing graph of incomplete pairwise comparison matrices can be important to get a reliable and good estimation of the decision maker's preferences. This conjecture came from the master thesis of one of the co-authors of Study I. (Tekile, 2017). It showed an example where the graph generated from the table tennis players' matches included a long shortest path between two vertices (players), and the calculated result appeared to be misleading because of that.

We carried out an extensive literature review on the filling designs of incomplete pairwise comparison matrices, and found that some sense of regularity of the representing graph was detected as an important property (Wang and Takahashi, 1998; Kułakowski et al., 2019), but the diameter was almost entirely missing from the relevant papers.

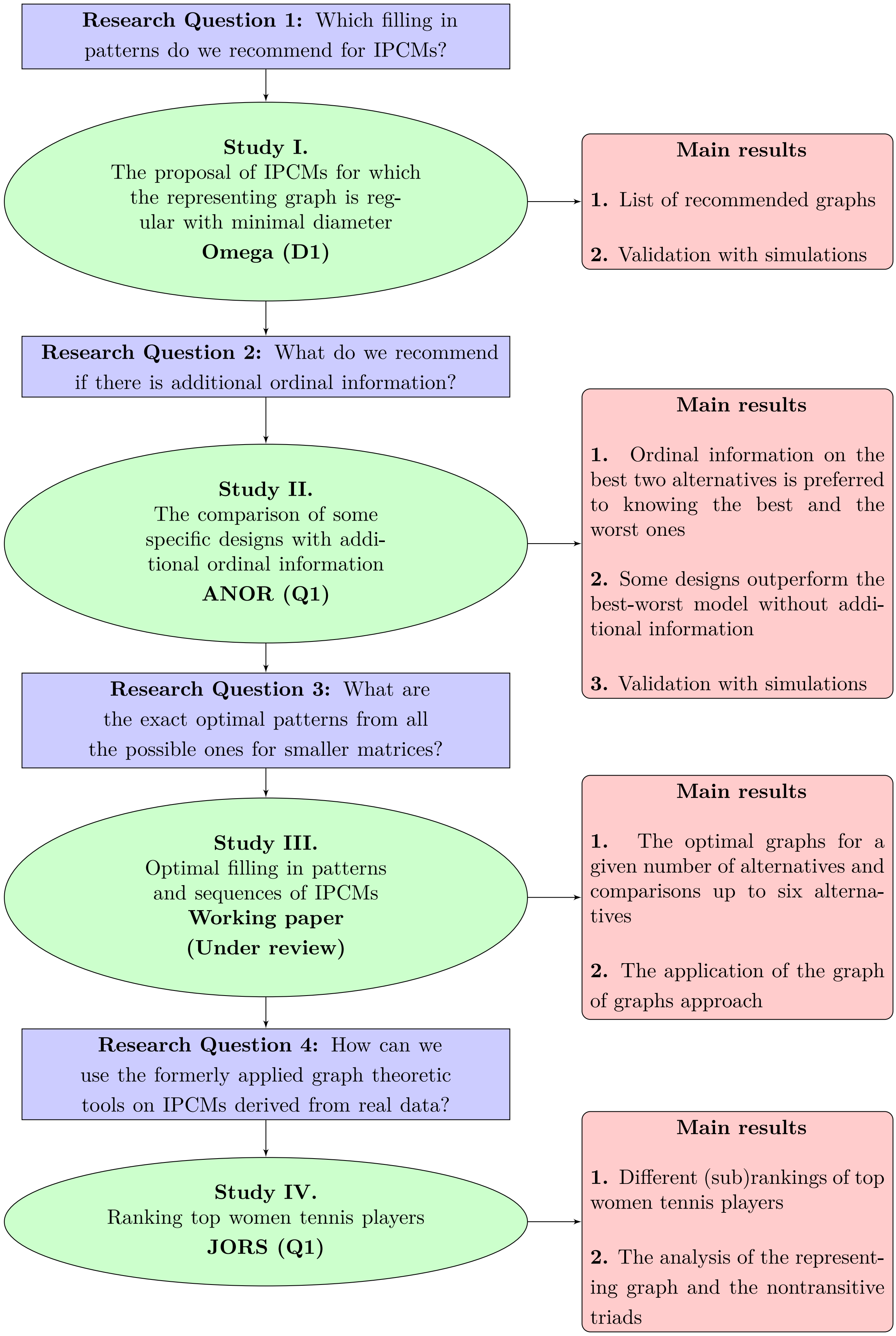

Figure 1: The flow chart of research questions, publications, and main results

In Study I., we proposed regular graphs with minimal diameter as a new design of filling in patterns for incomplete pairwise comparison matrices, created a list of proposed graphs and validated our recommendations by simulations. These results led to numerous new research questions.

First, in Study II., we dealt with the most common cases with a few number of alternatives in more detail, as well as focused on the inclusion of additional a priori ordinal information that is often used in multicriteria decision making methods, such as the best-worst method (Rezaei, 2015). It turned out that a Ukrainian research group found the importance of the diameter of the representing graph more or less at the same time, independently from us (Kadenko and Tsyganok, 2020). Thus, we continued the research together in Study II. We were able to compare some popular designs with our proposals, and the usefulness of additional ordinal information was also evaluated.

In the case of Study III., we continued to examine small matrices (with at most six alternatives), which are the most common in multicriteria decision making problems. One of the most important limitations of Study II. was that in some instances the examined designs used a different number of comparisons, thus, the effect of the filling structure and the effect of the number of known comparisons were inseparable. Based on that, in Study III., all possible filling in patterns for incomplete pairwise comparison matrices have been compared with a given number of comparisons, thus, it was possible to select the best one among them. According to Gyarmati et al. (2023), our results seem to be more general and not specific for the domain of pairwise comparison matrices.

As mentioned before, Study IV. uses the same tools as the other included papers, however, it focuses on the ranking aspect of incomplete pairwise comparisons as top women tennis players are ranked with this method. We revisited a former research of Bozóki et al. (2016) with a similar real-world database, however, we extended the results with a deeper analysis of the graph representation of the matches between the players, as well as the detailed investigation of nontransitive triads.

II. Applied methods: the main concepts connected to (incomplete) pairwise comparisons

In this section, the most important concepts connected to (incomplete) pairwise comparison matrices and their graph representations are defined formally, as these are the main applied methods in the dissertation. Most of the definitions listed here are also included in the four original studies of the thesis.

From now on, let us denote the number of criteria (alternatives) in a multicriteria decision making problem by n.

Definition 1 (Pairwise comparison matrix (PCM)) *The $n \times n$ matrix $A = [a_{ij}]$ is called a pairwise comparison matrix if it is positive ($a_{ij} > 0$ for all i and j) and reciprocal ($1/a_{ij} = a_{ji}$ for all i and j).*

When a decision maker fills in a PCM, there are usually some kind of inconsistency among the elements of the matrix. It can occur that alternative A is 2 times better than alternative B, and alternative B is 3 times better than alternative C, but alternative A is not $(2 \times 3 =)6$ times better than alternative C.

Definition 2 (Consistent PCM) *A PCM is said to be consistent if $a_{ik} = a_{ij}a_{jk}$ for all i, j, k. If a PCM is not consistent, then it is called inconsistent.*

There are several ways to measure the level of inconsistency (Brunelli, 2018; Kułakowski and Talaga, 2020), however, in practice the most often applied metric is still Saaty's Consistency Ratio (CR) (Saaty, 1977).

Definition 3 (Consistency Ratio (CR)) *The CR of an $n \times n$ PCM A is defined as follows:*

$$CR = \frac{CI}{RI}, \tag{1}$$

where CI stands for Consistency Index, that is:

$$CI = \frac{\lambda_{\max} - n}{n - 1}, \tag{2}$$

where $\lambda_{\max}$ is the principal eigenvalue of matrix A, and RI is the Random Index, which is the average CI obtained from a sufficiently large set of randomly generated PCMs of size n.

To determine the ranking of the alternatives, we need a weight (priority) vector calculation technique. The elements of the computed vector show the performance of the given alternatives. In the case of consistent PCMs, all of these techniques provide the same vector, however, for inconsistent matrices the result can be different.

Probably the two most commonly used techniques to calculate a weight vector are the logarithmic least squares (LLSM) (Crawford and Williams, 1985) and the eigenvector (EV) (Saaty, 1977) methods.

Definition 4 (Logarithmic Least Squares Method (LLSM)) *Let A be an $n \times n$ PCM. The positive weight vector w of A determined by the LLSM is the optimal solution of the following problem:*

$$\min_{w} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\ln(a_{ij}) - \ln\left(\frac{w_i}{w_j}\right) \right)^2, \tag{3}$$

where w_i is the ith coordinate of w.

Definition 5 (Eigenvector (EV) Method) *Let A be an $n \times n$ PCM. The positive weight vector w of A determined by the EV method is defined as follows:*

$$A \cdot w = \lambda_{\max} \cdot w, \tag{4}$$

where λ_{max} is the principal eigenvalue of matrix A.

The solutions of (3) and (4) are only unique up to scalar multiplication, thus, the sum of the components (the weights) are usually normalized to one.

Most of our research focuses on the representing graphs of incomplete pairwise comparison matrices.

Definition 6 (Incomplete pairwise comparison matrix (IPCM)) *An $n \times n$ matrix $A = [a_{ij}]$ is an incomplete pairwise comparison matrix (IPCM) if:*

- $a_{ij} \in \mathbb{R}_+ \cup \{*\}\ \forall\ 1 \leq i, j \leq n$ *and*
  - $a_{ji} = 1/a_{ij}$ *if* $a_{ij} \in \mathbb{R}_+$,
  - $a_{ji} = *$ *if* $a_{ij} = *$,

where $*$ denotes the missing elements, and $\mathbb{R}_+$ is the set of positive real numbers.

Definition 7 (Representing graph) *An incomplete pairwise comparison matrix A can be represented by an undirected graph $G = (V, E)$, where:*

- *the vertices $V = \{1, 2, \ldots, n\}$ correspond to the alternatives,*

- *while the edge set E represents the known elements of A outside the main diagonal:*

$$e_{ij} \in E \iff a_{ij} \neq * \text{ and } i \neq j.$$

Definition 8 (Path) *A path is a finite (or infinite) sequence of edges which joins a sequence of vertices such that all vertices (and therefore all edges) are distinct.*

Both the logarithmic least squares and eigenvector weight calculation methods can be used for incomplete pairwise comparison matrices as well. In the case of LLSM, optimization problem (3) is only applied to the known elements of the matrix, while for the EV method, the priority vector is calculated from the complete matrix where all missing comparisons are replaced by variables such that the Consistency Ratio (CR) is minimized (this technique is sometimes denoted as the CREV method). Both approaches provide a unique weight vector if and only if the representing graph of the IPCM is connected (Bozóki et al., 2010).

Definition 9 (Connected graph) *In an undirected graph, two vertices u and v are called connected if the graph contains a path from u to v. A graph is said to be connected if every pair of vertices in the graph is connected.*

The smallest connected graphs are the so-called spanning trees, which contain $n-1$ edges for n vertices, and play a special role in our research.

Definition 10 (Spanning tree) *Let $G = (V, E)$ be a connected graph. $G' = (V, E')$ is a spanning tree of G if $E' \subseteq E$ is a minimal set of edges that connects all vertices of G.*

Probably the two most important graph properties that we examine in our studies are related to the regularity and the diameter of the representing graph.

Definition 11 (k-regularity) *A graph is called k-regular if every vertex has k neighbours, which means that the degree of every vertex is k.*

Definition 12 (The diameter of a graph) *The diameter (denoted by d) of a graph G is the length of the longest shortest path between any two vertices:*

$$d = \max_{u,v \in V(G)} \ell(u, v),$$

where $V(G)$ denotes the set of vertices of G and $\ell(.,.)$ is the graph distance between two vertices, namely the length of the shortest path between them.

The concepts defined in this section are used in all of the studies included in the Ph.D. thesis. Besides this, Study I., II. and III. apply different simulations, while Study IV. uses different tests on the nontransitivity of pairwise comparisons. These methods are detailed in the next section, as they form an integral part of the scientific results of the thesis.

III. Scientific results of the thesis

1. Study I. Filling in pattern designs for incomplete pairwise comparison matrices: (Quasi-)regular graphs with minimal diameter

Study I. (Szádoczki et al., 2022) served as the basis and inspiration for most of the later research included in the dissertation. Our goal was to propose a new approach of filling in patterns for incomplete pairwise comparison matrices that shows which comparisons should be made. We heavily relied on the graph representations of these matrices (Gass, 1998), and their graph theoretical properties.

We had the following crucial assumptions applied in this article:

- it can be chosen which comparisons should be made (they are not given);
- there is no prior information available on the items that should be compared, we can handle them in a symmetric way and focus on unlabelled graphs (the isomorphic graphs are considered to be equivalent);
- the whole set of comparisons should be determined before the start of the decision making process (the questions are not adaptive, they do not depend on the previous answers).

As mentioned before, we had a conjecture based on a previous result (Tekile, 2017) that the diameter of the representing graph can be an important property to find a filling in pattern that estimates the results of the complete matrix well.

The diameter measures the longest shortest path in a graph. With the minimization of this metric, we can ensure that there is not a pair of alternatives that can only be compared through a long indirect path. In that case, the small errors along this path could sum up causing a significant error in the results.

After an extensive literature review, we found that although some sense of regularity of the representing graph was proposed in several designs (McCormick and Bachus, 1952; McCormick and Roberts, 1952; Wang and Takahashi, 1998), the diameter was not studied broadly.

The regularity of the representing graph is also important as it provides some kind of symmetry to the comparisons, all the alternatives are compared to the same number of items, and there is no pivotal element in the system of comparisons.

It is also true that the regularity definitions used by former designs were rather restrictive. We used a more general concept, and provided a more systematic list of graphs in the online Appendix of the study, which can serve as a 'recipe' for practitioners to determine which comparisons should be asked from the decision maker in a given problem.

To find the interesting graphs for our research, the three parameters were the number of vertices (alternatives) n, the level of regularity k, and the diameter of the given pattern d.

It is important to mention that k-regular graphs with n vertices do not exist if both k and n are odd, thus, we defined k-quasi-regular graphs, where one vertex has degree $k+1$, while all the other vertices have k neighbours.

We aimed to keep the number of comparisons (and so parameter k) as low as possible, and minimize the diameter at the same time. From a mathematical point of view, we were looking for the minimal d for a given (n, k) pair.

It turned out that this problem has a strong relation to the degree/diameter problem (see, for instance, Miller and Širáň (2013)), which looks for the largest n for a given (k, d) pair.

Based on the known results of this field and the characteristics of multicriteria decision making problems, it turned out that the parameter values of $k = 3, 4, 5$ and $d = 2, 3$ can be interesting for us ($d = 1$ means a complete PCM), while the number of alternatives were examined between 5 and 24. It is true that finding graphs and running simulations would become difficult above 24 vertices. However, in practical problems the number of alternatives is usually much below 24, and the largest 5-regular graph with diameter 2 has 24 vertices, thus, it is a natural theoretical bound as well. At the same time the problem is not too relevant below five alternatives.

Although some graphs were known from the degree/diameter literature, as mentioned before, most of them had to be determined by us. We used several different methods, for smaller cases the enumeration of all the graphs and the selection of the ones with minimal diameter, while for larger cases different graph extending techniques, graph products, integer linear programming, and so on. This was a long, time-consuming process, the same idea rarely worked twice, and we determined 34 types of graph instances for different parameter combinations.

After the finding of the graphs, we had to validate that these designs provide better results, namely, closer weight vectors to the ones calculated from the complete matrix than other filling patterns. In order to test this, we used simulations, where the CREV and the incomplete LLSM weight calculation techniques were used, while the closeness measures were the Euclidean and Chebyshev (maximum absolute) distances. We determined the means and standard deviations of the distances between the weight vectors calculated from a given design and the complete matrix for all combinations of the parameters.

We generated 1000 random, consistent PCMs for every parameter combination, and perturbed those using three different levels of perturbation (inconsistency) to get inconsistent matrices. The weight vectors were determined from these matrices using only the elements included in the examined patterns.

The following designs were compared in our simulations (all of the graphs were connected):

- (i) k-(quasi-)regular graphs with minimal diameter;
- (ii) Random connected graphs with the same number of edges as our recommendations;
- (iii) k-(quasi-)regular graphs, but not of minimal diameter;
- (iv) Randomly generated, minimal diameter, but not regular graphs;
- (v) Modified star graphs with the same number of edges (and minimal diameter).

The simulations confirmed that our recommendations outperformed all other designs in the case of almost all examined parameter combinations, for both measures, both weight calculation techniques, and every level of inconsistency. It was also validated that both properties—regularity and the minimal diameter—are necessary. The latter was also highlighted with a motivational example, where a regular graph had a long diameter, which resulted in large distances in the weight vectors from the complete case.

As mentioned before, the results of this study can be instantly applied in practical problems, while it also generated many further research questions, some of which later on have been addressed in Studies II. and III.

Our individual contributions to this study and joint results with our co-authors are detailed below.

Individual contributions:

- Implementation of the validating simulations in Scilab;
- Implementation of graph searching methods in Wolfram Mathematica and Python;
- Creating the systematic list of recommended graphs in several formats (adjacency matrix, edge list, graph, and 'Graph6' format);
- Creating the (more than 100 pages long) online Appendix;
- Implementing LaTeX codes to create illustrative figures.

Joint (inseparable) **results with our co-authors** (*Sándor Bozóki* and *Hailemariam Abebe Tekile*)**:**

- Extensive literature review connected to the filling designs for incomplete pairwise comparison matrices and the degree/diameter graph theoretical problem;
- Searching for the representing graphs with the proposed properties;
- Finding a motivational example to show the importance of the diameter;
- Running the simulations;
- Editing and writing the article.

2. Study II. Incomplete pairwise comparison matrices based on graphs with average degree approximately 3

It occurs in many different multicriteria decision making methods, such as the SMART (simple multi-attribute rating technique) (Edwards, 1977), the Swing (von Winterfeldt and Edwards, 1986), and their generalizations, that some additional ordinal information—usually the best, the worst or both alternatives—are used.

Accordingly, we continued our research started in Study I., and resolved the assumption that no prior information is available. We were particularly interested in the additional information used by the best-worst method (Rezaei, 2015), which generated an extensive literature in the past few years (Mi et al., 2019). The study (Szádoczki et al., 2023) also focuses on the smaller matrices (with at most 10 alternatives), which are more common in MCDM problems.

We applied a similar simulation-based approach as in Study I. with three different levels of inconsistency, however, several aspects of the methods were improved. Instead of the matrix-wise solution, we used element-wise perturbations, and also improved the handling of different scales. The elements of the consistent PCMs were randomly generated from the interval $[1/9, 9]$, and we used uniformly distributed errors for the perturbation. However, the new elements were uniformly distributed not on the original scale, but around the original value on the scale presented in Figure 2. The reasoning behind this is as follows. If a decision maker is hesitant whether item A is 2 or 3 times as good as item B, then that is the same problem as if the decision maker would be hesitant whether B is 1/2 or 1/3 as good as A. Thus, we used a scale where the distance between 1/9 and 1 is the same as between 1 and 9.

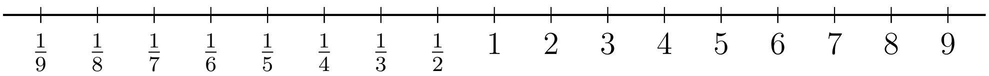

Figure 2: The scale on which the perturbed elements are uniformly distributed around the original value

The list of applied measures was also extended in the simulations. The following types of metrics were included:

- distances (Euclidean, Chebyshev, and Manhattan);
- rank correlations (Kendall's tau and Spearman's rho);
- and compatibility (similarity) indices (Garuti's compatibility index (Garuti, 2017), cosine similarity index (Kou et al., 2021), and Dice similarity index (Ye, 2012)).

We used a sample of 10 000 PCMs for the simulations, and aimed to keep the number of comparisons relatively low, thus, the following designs (representing graphs) were compared:

- (i) Best-worst graph;
- (ii) TOP2 graph;

- (iii) Best-random graph;
- (iv) Random-random graph;
- (v) 3-(quasi-)regular graph with minimal diameter;
- (vi) Union of two random edge-disjoint spanning trees;
- (vii) Random 2-edge-connected graph.

The first four designs are based on the same graph representation (the union of two star graphs) with different additional ordinal information. Thus, there are two pivotal elements, which are different in these models, and all of the other alternatives are only compared to them. We were interested in these kinds of structures because of the best-worst method, where every alternative is compared to the best and the worst items. This method also assumes that it is always possible to determine the best and the worst elements, however, this implies that it is always possible to determine the ranking of the alternatives according to one criterion (i.e., if we remove the former best and worst, we can determine the new best and worst again). Thus, it can make sense to test other pairs of highlighted alternatives.

In TOP2, the pivotal elements are the best and the second best alternatives, for the Best-random graph they are the best and a randomly chosen element, while for the Random-random case both of them are chosen randomly. This way we were able to evaluate these additional information as well, whether comparing to those elements provide weight vectors closer to the one calculated from the complete case. The TOP2 design was proposed by us, as usually if there are more comparisons for a given alternative, then its weight is estimated more accurately, and probably the most important items are the two best ones.

The 3-(quasi-)regular graph design was also proposed by us in Study I., but it does not use any additional ordinal information. Nor do the Union of two random edge-disjoint spanning trees and the Random 2-edge-connected graphs. The latter two are generalizations of the union of two star graphs, as a star graph is a certain type of spanning tree, and a 2-edge connected graph is a graph that remains connected after deleting an edge, which also holds for the union of two star graphs.

According to the results of the simulations, basically every measure that we used provided the same findings, thus, the outcomes do not heavily depend on the used metrics. It turned out that the TOP2 graph is preferred to the Best-worst one, which means that if we have an opportunity in this kind of problems to ask for additional ordinal information, on average it is better to ask the two best alternatives compared to the popular best and worst ones. However, some of the designs that do not use any additional information outperformed all the models using the union of two star graphs. For smaller cases, this is true for the 3-(quasi-)regular graphs, and for basically all examined cases, this is observed for the Union of two random edge-disjoint spanning trees.

Nonetheless, it is important to keep in mind one of the most important limitations of this research. That is, some of these designs use a different number of comparisons in certain cases, which does not make it possible to separate the effect of the filling pattern from the effect of the number of comparisons. We tried to resolve this problem in Study III.
In this study, all of the results are equally joint (inseparable) with our co-authors (*Sándor Bozóki*, *Patrik Juhász*, *Sergii V. Kadenko*, and *Vitaliy Tsyganok*).

3. Study III. Optimal sequences for pairwise comparisons: the graph of graphs approach

In Study III. (Szádoczki and Bozóki, 2022), we aimed to elaborate on the smallest cases that most commonly occur in MCDM problems, and also address the limitation of Study II. that the number of comparisons were different, thus, in some cases it was difficult to draw strong conclusions.

We were able to automize our simulations in order to examine larger samples and check all the possible patterns of comparisons (representing graphs) instead of some special designs as before, only analyzing the structures with the same number of comparisons (edges) this time, up to six alternatives (vertices). At the same time, we also went back to our original assumption that no prior information is available about the elements, as the labelling of graphs results in even more possibilities.

As measures, the Euclidean distance and the Kendall's tau metrics were used, with a sample size of 1 million matrices for a given number of alternatives (n) and a given number of comparisons (edges – e), which resulted in certain margins of error and significance levels for the different means.

It turned out that basically always the same representing graph provided the closest weight vector to the complete case for a given (n, e) pair according to the Euclidean distance and Kendall's tau, and the results neither depended on the level of inconsistency nor on the weight calculation technique. Thus, in that sense we were able to determine the optimal filling in patterns for a given number of alternatives (up until six) and a given number of comparisons, which is a key contribution of this study to the literature.

Furthermore, we also found that many optimal cases are reachable from each other, namely, if we add an additional comparison to the optimal pattern with e comparisons, it results in the optimal case with $e + 1$ comparisons. This way, we can determine optimal filling in sequences for pairwise comparison matrices, which can be especially useful in the case of group decision making based on online questionnaires, where the number of comparisons provided by the decision makers is uncertain. If we follow an optimal filling in sequence, then we can ensure that the preferences of the decision makers are estimated on average as well as possible whenever they stop making further comparisons.

After that, we conducted a literature review in order to find out what kind of visualization can be used to present our results. Based on that, we applied the graph of graphs approach (see, for instance, Lovász (1977)), where there is a large graph and its vertices are also graphs, in our case the representing graphs of filling patterns. There is an edge between two small graphs if and only if they are reachable from each other, we can get one of them from the other one by adding (or deleting) exactly one comparison (edge).

As an example, the graph of graphs on five vertices can be seen in Figure 3. The number of comparisons (edges) are denoted on the left side of the chart, while the optimal graphs and partial optimal sequences are highlighted with blue color and boldness. It is important to note that the pattern of results is similar for different number of alternatives as well.

Among the spanning trees (in the first row), the star graph proved to be optimal for all cases examined, and it is expected that this structure retains its optimality for larger cases. Regular and quasi-regular graphs were also optimal, moreover, regularity was important in an even more general way, namely, the degrees of the vertices were always as close to each other as possible in the optimal graphs. The bipartiteness of the graphs also turned out to be important, i.e., the vertices were divisible to two independent sets A and B, and all edges were between a vertex from A and one from B.

As one can see in Figure 3, there is not a total optimal filling in sequence for the PCM in this case, however, we can determine a path that contains as many optimal cases as possible, and all the other graphs are close to optimal as well.

Similarly to Study I., our results were presented in different formats that make it possible for practitioners to instantly apply them in multicriteria decision making problems. Moreover, some of our findings can be used to determine the minimal number of questions to be asked from the

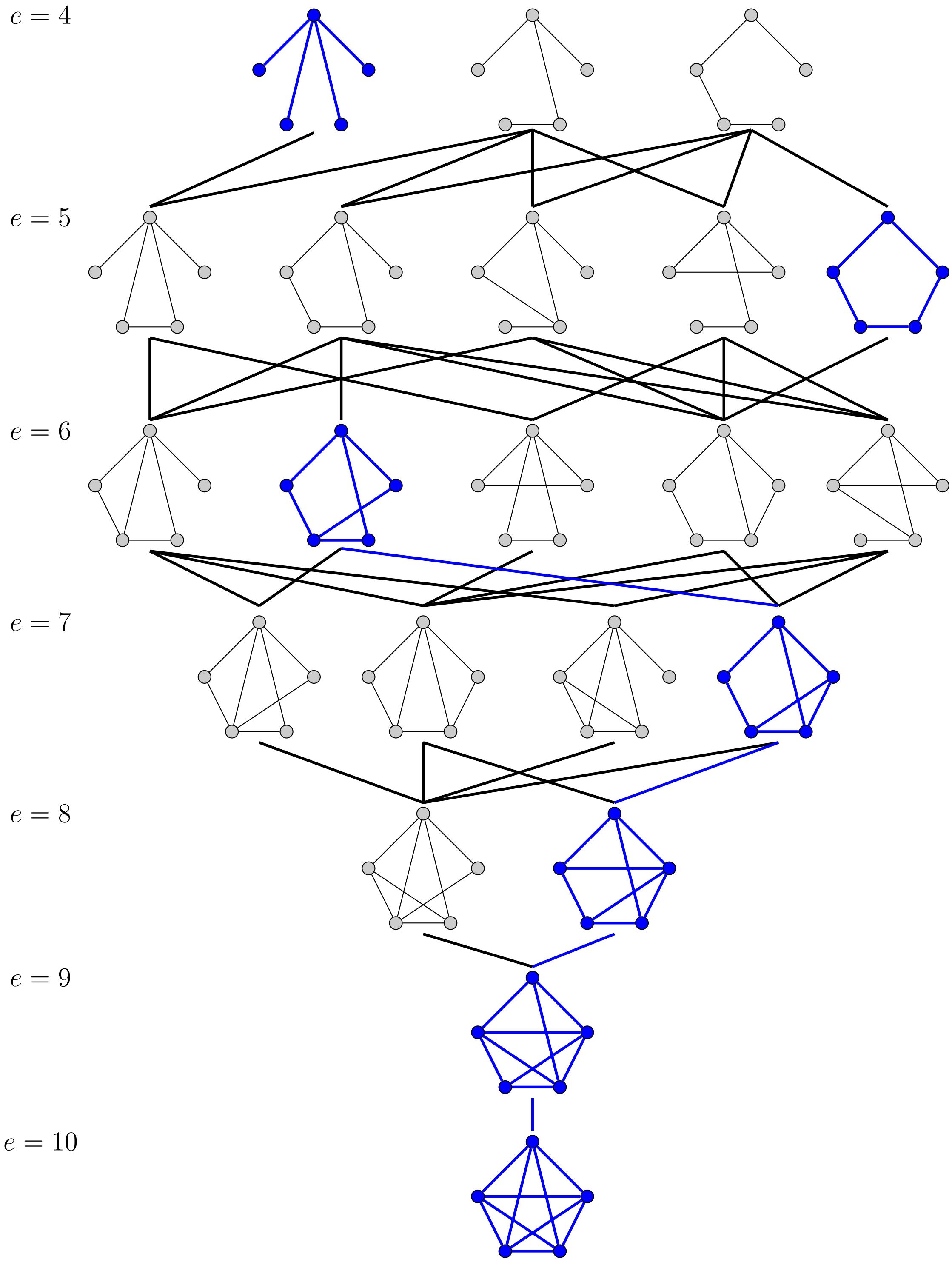

Figure 3: The graph of graphs for $n = 5$, optimal graphs are highlighted by blue and bold

decision makers in the case of a given problem, if we would like to have a threshold for the metrics examined. Based on our recent results (Gyarmati et al., 2023), which largely rely on Study III., the findings (the proposed graphs and sequences) also seem to be more general, not specific for PCMs, but optimal in the case of other pairwise comparison-based models as well.

Our individual contributions to this study and joint results with our co-author are detailed below.

Individual contributions:

- Implementation of the methods that determine all the possible representing graphs in Wolfram Mathematica;
- Implementation of the simulations in Scilab, with which we can determine the optimal cases;
- Implementing LaTeX codes to create visualizations (creating the graphs of graphs).

Joint (inseparable) **results with our co-author** (*Sándor Bozóki*)**:**

- Literature review connected to the filling sequences for pairwise comparison matrices and the graph of graphs concept;
- Running the simulations;
- Editing and writing the article.

4. Study IV. Incomplete pairwise comparison matrices: Ranking top women tennis players

Study IV. (Temesi et al., 2024) applies the same graph theoretical tools as the previously presented articles on a real database, while in some sense continues the research started by Bozóki et al. (2016).

Our aim was to create all-time rankings of top women tennis players based on an incomplete pairwise comparison matrix derived from their Win/Loss ratios. The alternatives were those players who were at some time between 1973 (the foundation of the Women's Tennis Association (WTA)) and 2022 number one in the official ranking of the WTA. We found 28 such players, and there were, of course, missing elements in our data, as some of them never played with each other during their career.

In order to apply the weight calculation techniques (besides checking whether the representing graph is connected), we had to make several technical adjustments in the data, in which we followed Bozóki et al. (2016).

If only one of the players won all the matches in a pair, then there should have been a 0 in the denominator of the given Win/Loss ratio. In this case, we used the number of wins and added 2 to it as the value in the matrix. After that, we also modified all the values by a power, in order to account for the fact that some pairs only played a few games against each other, while others clashed in several tournaments. The point of this modification was that if a pair had more matches with each other, than their comparison was considered to be more reliable.

We used both the eigenvector and the LLSM weight calculation methods with different parameters in the aforementioned adjustments, as well as applied another ranking technique as a benchmark, the Bradley–Terry model (Bradley and Terry, 1952). The rankings turned out to be similar to each other, the calculated Spearman rank correlations were convincingly high (0.860 was the smallest value) and robust to the modifications as well.

The first two places were occupied by Serena Williams and Steffi Graf, respectively, according to all methods. They were followed by Navratilova, Hingis, Clijsters, and Henin with slight modifications. Somewhat surprisingly, the main difference between the results of the Bradley–Terry model and the IPCM-based models was that some of the earliest players (Goolagong, Evert, and Austin) performed better according to the Bradley–Terry method.

Besides demonstrating that the MCDM methods connected to IPCMs can be applied to create rankings based on large real data sets, the goal of this research was also to thoroughly investigate the structure of the pairwise comparisons. We focused on several submatrices of the data as well as the case when the players entered to the data set one by one.

We found that it is possible to determine relatively large subsets of players with their most active career in the same time period, so that their results against each other practically determine their positions in the overall ranking. Entering players one by one to the data set showed that there are usually only small changes due to an entry, and every player affects the most those ones against whom she played the most.

Looking into the properties of the representing graph, it turned out that if we eliminate the four earliest players (Goolagong, Evert, Navratilova and Austin) along with the most recent world number one (Swiatek), then we get a graph, for which the different connectivity measures are really strong. The diameter (longest shortest path) of this modified graph is only 2, and it also can be interpreted as a union of two star graphs centered around the Williams sisters (i.e., they played with everyone else) complemented with further edges. This kind of structure and the strong connectivity might be the reason behind the robustness of the rankings.

We also analyzed the ordinally nontransitive triads in the data set. In sports, it can happen that A won against B in the majority of their matches, and generally B was successful against C as well. We would expect that A also beat C, however, in reality A lost against C in most occasions.

These kind of triads also have a significant literature, Kendall and Babington Smith (1940) determined the distribution of them for low number of elements and proposed a significance test, which was later extended by several researchers. Connected to PCMs and decision making problems, Iida (2009) investigated different tests and indices connected to the question.

In order to perform the necessary chi-square nontransitivity test, we had to modify the originally used IPCM, as it can be applied only on complete data without ties. The adjusted matrix contained W/L set ratios instead of match ratios for every pair. When we had ties even for the set ratios, the original LLSM ranking was used as a reference to make a precedence relation. If two players have never played against each other during their career (no edge exists between the two vertices), then we used the same LLSM ranking to determine the winner of the pair.

Applying these adjustments, we were able to construct a directed graph, on which the test itself is based on. The number of ordinally nontransitive triads turned out to be insignificant in our database, which can also explain the robustness of the rankings.

Our results can be interesting not only for tennis experts and fans, but also provide empirical evidence that the method of incomplete pairwise comparison matrices is appropriate for producing well-understandable rankings, as well as connect the research of representing graphs to a large real-world database and sports.

Our individual contributions to this study and joint results with our co-authors are detailed below.

Individual contributions:

- Calculating the rankings of top women tennis players based on different methods and submatrices;
- Implementing R, Scilab, and Wolfram Mathematica codes to apply different methods and tests to the data;

- Implementing LaTeX codes to create illustrative figures.

Joint (inseparable) **results with our co-authors** (*József Temesi* and *Sándor Bozóki*)**:**

- Extensive literature review connected to ranking in sports, nontransitive triads, and ranking with incomplete pairwise comparisons;
- Editing and writing the article.

IV. Main references

Ábele-Nagy, K. (2019). *Pairwise comparison matrices in multi-criteria decision making*. PhD thesis, Corvinus University of Budapest. In Hungarian. https://doi.org/10.14267/phd.2019036.

Ágoston, K. Cs. and Csató, L. (2022). Inconsistency thresholds for incomplete pairwise comparison matrices. *Omega*, 108:102576. https://doi.org/10.1016/j.omega.2021.102576.

Bozóki, S., Csató, L., and Temesi, J. (2016). An application of incomplete pairwise comparison matrices for ranking top tennis players. *European Journal of Operational Research*, 248(1):211–218. https://doi.org/10.1016/j.ejor.2015.06.069.

Bozóki, S. (2006). *Weighting with pairwise comparisons and evaluation with utility functions in multicriteria decision making problems*. PhD thesis, Corvinus University of Budapest. In Hungarian. https://phd.lib.uni-corvinus.hu/245/1/bozoki_sandor.pdf.

Bozóki, S., Fülöp, J., and Rónyai, L. (2010). On optimal completion of incomplete pairwise comparison matrices. *Mathematical and Computer Modelling*, 52(1):318–333. https://doi.org/10.1016/j.mcm.2010.02.047.

Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345. https://doi.org/10.2307/2334029.

Brunelli, M. (2018). A survey of inconsistency indices for pairwise comparisons. *International Journal of General Systems*, 47(8):751–771. https://doi.org/10.1080/03081079.2018.1523156.

Crawford, G. and Williams, C. (1985). A note on the analysis of subjective judgment matrices. *Journal of Mathematical Psychology*, 29(4):387–405. https://doi.org/10.1016/0022-2496(85)90002-1.

Csató, L. (2015). *Methodological and applicational issues of paired comparison based ranking*. PhD thesis, Corvinus University of Budapest. In Hungarian. https://doi.org/10.14267/phd.2015022.

Csató, L. (2021). *Tournament Design: How Operations Research Can Improve Sports Rules*. Palgrave Pivots in Sports Economics, Palgrave Macmillan. https://doi.org/10.1007/978-3-030-59844-0.

Davidson, R. and Farquhar, P. (1976). A bibliography on the method of paired comparisons. *Biometrics*, 32(2):241–252. https://www.jstor.org/stable/2529495.

Edwards, W. (1977). How to use multiattribute utility measurement for social decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics*, 7(5):326–340. https://doi.org/10.1109/TSMC.1977.4309720.

Garuti, C. E. (2017). Reflections on scales of measurement, not measurement of scales. *International Journal of the Analytic Hierarchy Process*, 9(3). https://doi.org/10.13033/ijahp.v9i3.522.

Gass, S. (1998). Tournaments, transitivity and pairwise comparison matrices. *Journal of the Operational Research Society*, 49(6):616–624. https://www.tandfonline.com/doi/abs/10.1057/palgrave.jors.2600572.

Gyarmati, L., Orbán-Mihálykó, É., Mihálykó, Cs., Szádoczki, Zs., and Bozóki, S. (2023). The incomplete analytic hierarchy process and Bradley–Terry model: (In)consistency and information retrieval. *Expert Systems with Applications*, 229(Part B):120522. https://doi.org/10.1016/j.eswa.2023.120522.

Harker, P. T. (1987). Incomplete pairwise comparisons in the analytic hierarchy process. *Mathematical Modelling*, 9(11):837–848. https://doi.org/10.1016/0270-0255(87)90503-3.

Iida, Y. (2009). The number of circular triads in a pairwise comparison matrix and a consistency test in the AHP. *Journal of the Operations Research Society of Japan*, 52(2):174–185. https://doi.org/10.15807/jorsj.52.174.

Ishizaka, A. and Nemery, P. (2013). *Multi-Criteria Decision Analysis: Methods and Software*. Wiley. https://doi.org/10.1002/9781118644898.

Kadenko, S. and Tsyganok, V. (2020). An update on combinatorial method for aggregation of expert judgments in AHP. Proceedings of the International Symposium on the Analytic Hierarchy Process, ISAHP-2020. https://doi.org/10.13033/isahp.y2020.012.

Kendall, M. G. and Babington Smith, B. (1940). On the method of paired comparisons. *Biometrika*, 31(3/4):324–345. https://doi.org/10.2307/2332613.

Kou, G., Peng, Y., Chao, X., Herrera-Viedma, E., and Alsaadi, F. E. (2021). A geometrical method for consensus building in GDM with incomplete heterogeneous preference information. *Applied Soft Computing*, 105:107224. https://doi.org/10.1016/j.asoc.2021.107224.

Kułakowski, K., Szybowski, J., and Prusak, A. (2019). Towards quantification of incompleteness in the pairwise comparisons methods. *International Journal of Approximate Reasoning*, 115:221–234. https://doi.org/10.1016/j.ijar.2019.10.002.

Kułakowski, K. and Talaga, D. (2020). Inconsistency indices for incomplete pairwise comparisons matrices. *International Journal of General Systems*, 49(2):174–200. https://doi.org/10.1080/03081079.2020.1713116.

Lovász, L. (1977). A homology theory for spanning trees of a graph. *Acta Mathematica Academiae Scientiarum Hungaricae*, 30(3-4):241–251. https://doi.org/10.1007/bf01896190.

McCormick, E. and Bachus, J. (1952). Paired comparison ratings: 1. The effect on ratings of reductions in the number of pairs. *Journal of Applied Psychology*, 36(3):123–127. https://doi.org/10.1037/h0054842.

McCormick, E. and Roberts, W. (1952). Paired comparison ratings: 2. The reliability of ratings based on partial pairings. *Journal of Applied Psychology*, 36(3):188–192. https://doi.org/10.1037/h0055956.

Mi, X., Tang, M., Liao, H., Shen, W., and Lev, B. (2019). The state-of-the-art survey on integrations and applications of the best worst method in decision making: Why, what, what for and what's next? *Omega*, 87:205–225. https://doi.org/10.1016/j.omega.2019.01.009.

Miller, M. and Širáň, J. (2013). Moore graphs and beyond: A survey of the degree/diameter problem. *Electronic Journal of Combinatorics*, 20(2):1–92. https://doi.org/10.37236/35.

Poesz, A. (2022). *Inconsistency in multi-attribute decision problems*. PhD thesis, Corvinus University of Budapest. In Hungarian. https://doi.org/10.14267/phd.2022053.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, 53:49–57. https://doi.org/10.1016/j.omega.2014.11.009.

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3):234–281. https://doi.org/10.1016/0022-2496(77)90033-5.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill, New York.

Szádoczki, Zs. and Bozóki, S. (2022). Optimal sequences for pairwise comparisons: the graph of graphs approach. *Working paper*. https://doi.org/10.48550/arXiv.2205.08673.

Szádoczki, Zs., Bozóki, S., Juhász, P., Kadenko, S. V., and Tsyganok, V. (2023). Incomplete pairwise comparison matrices based on graphs with average degree approximately 3. *Annals of Operations Research*, 326(2):783–807. https://doi.org/10.1007/s10479-022-04819-9.

Szádoczki, Zs., Bozóki, S., and Tekile, H. A. (2022). Filling in pattern designs for incomplete pairwise comparison matrices: (Quasi-)regular graphs with minimal diameter. *Omega*, 107:102557. https://doi.org/10.1016/j.omega.2021.102557.

Szybowski, J., Kułakowski, K., and Prusak, A. (2020). New inconsistency indicators for incomplete pairwise comparisons matrices. *Mathematical Social Sciences*, 108:138–145. https://doi.org/10.1016/j.mathsocsci.2020.05.002.

Tekile, H. A. (2017). Incomplete pairwise comparison matrices in multi-criteria decision making and ranking. Master's thesis, Central European University.

Temesi, J., Szádoczki, Zs., and Bozóki, S. (2024). Incomplete pairwise comparison matrices: Ranking top women tennis players. *Journal of the Operational Research Society*, 75(1):145–157. https://doi.org/10.1080/01605682.2023.2180447.

Thurstone, L. (1927). A law of comparative judgment. *Psychological Review*, 34(4):273–286. https://doi.org/10.1037/h0070288.

von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioral Research*. Cambridge University Press, Cambridge.

Wang, K. and Takahashi, I. (1998). How to select paired comparisons in AHP of incomplete information – strongly regular graph design. *Journal of the Operations Research Society of Japan*, 41(2):311–328. https://doi.org/10.15807/jorsj.41.311.

Ye, J. (2012). Multicriteria decision-making method using the Dice similarity measure based on the reduct intuitionistic fuzzy sets of interval-valued intuitionistic fuzzy sets. *Applied Mathematical Modelling*, 36(9):4466–4472. https://doi.org/10.1016/j.apm.2011.11.075.

Zhou, X., Hu, Y., Deng, Y., Chan, F. T. S., and Ishizaka, A. (2018). A DEMATEL-based completion method for incomplete pairwise comparison matrix in AHP. *Annals of Operations Research*, 271:1045–1066. https://doi.org/10.1007/s10479-018-2769-3.

V. List of publications

1. Publications in English

Publications included in the Ph.D. thesis

1) Szádoczki, Zs. and Bozóki, S. (2022). Optimal sequences for pairwise comparisons: the graph of graphs approach. *Working paper, under review.* https://doi.org/10.48550/arXiv.2205.08673.

2) Szádoczki, Zs., Bozóki, S., Juhász, P., Kadenko, S. V., and Tsyganok, V. (2023). Incomplete pairwise comparison matrices based on graphs with average degree approximately 3. *Annals of Operations Research*, 326(2):783–807. https://doi.org/10.1007/s10479-022-04819-9.

3) Szádoczki, Zs., Bozóki, S., and Tekile, H. A. (2022). Filling in pattern designs for incomplete pairwise comparison matrices: (Quasi-)regular graphs with minimal diameter. *Omega*, 107:102557. https://doi.org/10.1016/j.omega.2021.102557.

4) Temesi, J., Szádoczki, Zs., and Bozóki, S. (2024). Incomplete pairwise comparison matrices: Ranking top women tennis players. *Journal of the Operational Research Society*, 75(1):145-157. https://doi.org/10.1080/01605682.2023.2180447.

Peer-reviewed journal articles that are not included in the Ph.D. thesis

5) Duleba, Sz. and Szádoczki, Zs. (2022). Comparing aggregation methods in large-scale group AHP: Time for the shift to distance-based aggregation. *Expert Systems with Applications*, 196:116667. https://doi.org/10.1016/j.eswa.2022.116667.

6) Gyarmati, L., Orbán-Mihálykó, É., Mihálykó, Cs., Szádoczki, Zs., and Bozóki, S. (2023). The incomplete analytic hierarchy process and Bradley–Terry model: (In)consistency and information retrieval. *Expert Systems with Applications*, 229(Part B):120522. https://doi.org/10.1016/j.eswa.2023.120522.

7) Kadenko, S., Tsyganok, V., Szádoczki, Zs., and Bozóki, S. (2021). An update on combinatorial method for aggregation of expert judgments in AHP. *Production*, 31:e20210045. https://doi.org/10.1590/0103-6513.20210045.

8) Szabo, Zs. K., Szádoczki, Zs., Bozóki, S., Stănciulescu, G. C., and Szabo, D. (2021). An Analytic Hierarchy Process approach for prioritisation of strategic objectives of sustainable development. *Sustainability*, 13(4). https://doi.org/10.3390/su13042254.

9) Szádoczki, Zs. and Duleba, Sz. (2022). Distance-based aggregation in group AHP. *Journal of Decision Systems*, 31(sup1):98–106. https://doi.org/10.1080/12460125.2022.2070952.

Working papers that are not included in the Ph.D. thesis

10) Csató, L., Kiss, L. M., and Szádoczki, Zs. (2024). The allocation of FIFA World Cup slots based on the ranking of confederations. *Working paper, under review.* https://doi.org/10.48550/arXiv.2310.19100.

11) Kelemen, A., Szabo, Zs. K., Szádoczki, Zs., and Bozóki, S. (2024). A sensitivity analysis of composite indicators: Min/max thresholds. *Working paper, under review.*

2. Publications in Hungarian

Peer-reviewed journal article that is not included in the Ph.D. thesis

12) Temesi, J., Szádoczki, Zs., and Bozóki, S. (2022). Nem teljes páros összehasonlítások: A női teniszezők világrangsorának példája. *Szigma*, 53(1):1–32. https://journals.lib.pte.hu/index.php/szigma/article/view/5709/5492.

Book review that is not included in the Ph.D. thesis

13) Szádoczki, Zs. (2022). Operációkutatás a sportok profitabilitásáért. Könyvismertetés, László Csató: *Tournament Design. How Operations Research Can Improve Sports Rules? Közgazdasági Szemle*, 69(2):283–288. https://doi.org/10.18414/KSZ.2022.2.283.