LMU

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

INSTITUT FÜR STATISTIK

Georg Schollmeyer

Application of lower quantiles for complete lattices to ranking data: Analyzing outlyingness of preference orderings

Technical Report Number 208, 2017
Department of Statistics
University of Munich

http://www.stat.uni-muenchen.de

Application of lower quantiles for complete lattices to ranking data: Analyzing outlyingness of preference orderings

Georg Schollmeyer

Abstract

The aim of the present paper is to apply a recently developed quantile approach for lattice-valued data to the special case of ranking data. We show how to analyze profiles of total orders by means of lattice-valued quantiles and thereby develop new methods of descriptive data analysis for ranking data beyond known methods like permutation polytopes or multidimensional scaling. We furthermore develop an aggregation rule for social profiles (, called "commonality sharing", here,) that selects from a given profile that ordering(s) that is (are) most centered in the profile, where the notion of centrality and outlyingness are based on purely order-theoretic concepts. Finally, we sketch, how one can use the quantile approach to establish tests of model fit for statistical models of ranking data.

Keywords: complete lattice, quantile, outlyingness, descriptive data analysis, social profile, social choice theory, consensus rule, commonality sharing

1 Introduction

The descriptive analysis of ranking data comes along with some natural difficulties due to the discrete character and the typically high dimensionality of ranking data. Most descriptive methods like for example multidimensional scaling (see, e.g., Kidwell et al. [2008]) are either based on a more or less geometric understanding or rely on an explicit statistical modeling (see, e.g., Biernacki and Jacques [2013], Jacques and Biernacki [2014], Lee et al. [2014]). Other techniques like the use of permutation polytopes (see, e.g., Thompson [1993]) or projections thereof (like, e.g., the marginal matrix, cf., Marden [1996]) are either limited to a small number of items ranked or captures only small parts of the high dimensional data structure. The aim of this paper is to analyze ranking data in a purely order-theoretic manner without the need of introducing some notion of a metric and with a clear emphasis on the structure already inherent in the data. The basic tool we use is a recently developed notion of quantiles and outlyingness for data points that are elements of a complete lattice. The general notion of (lower) quantiles for arbitrary complete lattices and a short analysis of properties are sketched in Schollmeyer [2017]. Here we apply these notions to the complete lattice of all binary relations on a fixed finite set of items.

The paper is structured as follows: Section 2 informally introduces the quantile concept developed in Schollmeyer [2017] and shows, how it could be applied to ranking data. Section 3 indicates, how one can apply the outlyingness concept to get some aggregation function for social profiles of Social Choice Theory by mapping a social profile to that ordering(s) that is (are) the most central. In Section 4 we apply the outlyingness concept to data of the wisdom of the crowds phenomena. We use the outlyingness notion for descriptive data analysis as well as for sketching a way for assessing the fit of two exemplary statistical models for ranking data, namely the Insertion Sorting Rank model ([Biernacki and Jacques, 2013]) and a Thurstonian model used in Lee et al. [2014].

2 A lattice-valued conceptualization of quantiles and outlyingness

First, let us briefly introduce lattices and lattice-valued quantiles. A lattice $\mathbb{L} = (L, \leq)$ is a partially ordered set (that is, a set L with a transitive, reflexive and antisymmetric binary relation $\leq$) such that for every two elements $x, y \in \mathbb{L}$ there always exists a least upper bound (called supremum or join) and a greatest lower bound (called infimum or meet) of these elements. If furthermore there exists suprema and infima for every arbitrary set $M \subseteq L$ (including also the empty set) then $\mathbb{L}$ is called a complete lattice[1]. Note that existing infima and suprema are always unique. For a set $M \subseteq L$ we denote with $\bigwedge M$ the infimum and with $\bigvee M$ the supremum of the set M, respectively.

2.1 A short informal introduction to lattice-valued quantiles

Let now $(x_1, \ldots, x_n)$ be a n-tupel of elements of a complete lattice $\mathbb{L}$ representing an i.i.d-sample of a lattice valued random variable X. Firstly, fix some proportion $\alpha \in [0, 1]$. Then define the set

$$M_\alpha := \{x \in \mathbb{L} \mid \#\{y \in \{x_1, \ldots . x_n\} \mid y \leq x\} \geq \alpha \cdot n\}$$

of all elements x such that at least $\alpha \cdot 100\%$ of the data points $(x_1, \ldots, x_n)$ lie below x. If $\mathbb{L}$ would be totally ordered (meaning that for every $x, y \in \mathbb{L}$ we have always $x \leq y$ or $y \leq x$) then the set M_α would have a smallest element that would be usually called a lower α-quantile. Generally, the set M_α need not to have a smallest element, but one can choose the unique element $q := \bigwedge M_\alpha$ as a lower **prequantile**. Note that below the element q there need not to lie $\alpha \cdot 100\%$ of the data points anymore. Thus, we call α the **prelevel** of q and we call the proportion β of data points that actually lie below q the **level** of q. Since q is furthermore generally not minimal with respect to the property of lying above

[1]Note, that this implies in particular, that every complete lattice necessarily has to have a greatest element as the infimum of the empty set and a least element as the supremum of the empty set.

$\beta \cdot 100\%$ of the data points, we finally take $q' := \bigvee\{x \in (x_1, \ldots, x_n\} \mid x \leq q\}$ as the final **(lower) quantile** that now lies above $\beta \cdot 100\%$ of the data and is actually minimal w.r.t. this property (see Schollmeyer [2017]). Note that the above definition of a lower quantile is actually an asymmetric one. If one uses the relation $\geq$ instead of $\leq$ in the given lattice, then generally one gets a different dual notion of upper quantiles. However, in the case of the application of lattice-valued quantiles for ranking data, the lower and upper quantile constructions are basically isomorphic to each other if we restrict attention to the analysis of total orders, which will be the case in this paper.

2.2 Application of lattice-valued quantiles to ranking data

We now want to apply lattice-valued quantiles to ranking data. Let $C = \{C_1, \ldots, C_q\}$ be a finite set of q items. Now, assume we have data from n persons that rank these q items. The rankings the persons supply could have different meanings, for example they could represent the persons preferences of different candidates in a ballot in the context of Social Choice Theory or the persons opinion about the chronological order of some historical events in the context of the analysis of the wisdom of crowd phenomena (cf., Surowiecki [2005]), etc. In this paper, we assume that the persons supply complete orderings of the items, meaning that the relations given by persons are a total order relations (i.e., reflexive transitive and antisymmetric relations where every two elements are comparable). The vector $R = (R_1, \ldots, R_n)$ of the n total orderings the persons supply is called a **profile** here. To analyze a given profile R, we look at every ordering in this profile as an element of the complete lattice $\mathbb{L} := (2^{C \times C}, \subseteq)$ of binary relations on the fixed item set C equipped with the set inclusion $\subseteq$ as the underlying partial order. In this complete lattice the infimum is the set intersection and the supremum is the set union. In this context, the elements of M_α are binary relations on C that, treated as subsets of $C \times C$, lie above the rankings of at least $\alpha \cdot 100\%$ of the whole population, meaning that they are supersets of at least $\alpha \cdot 100\%$ of the orderings in the profile. The derived quantile q of prelevel α is then the infimum of all the elements of the set M_α, i.e., it is simply the intersection of all binary relations on C that lie above at least $\alpha \cdot 100\%$ of the rankings of the population. This intersection can be understood as the commonalities of all binary relations that represent all sub-populations of a given minimum size $\alpha \cdot n$. The finally derived quantile with actual level β could be understood as the union of all the $\beta \cdot n$ rankings of persons that are represented by this intersection. The higher the quantile q, the bigger and thus more diverse is the sub-population that is represented by this quantile q. Note that different quantiles are always pairwise comparable and thus form a chain. If we map the ranking R of a person to the smallest quantile $\Phi(R)$ that still lies above this ranking R, then we have some kind of qualitative measurement mapping that "measures" the outlyingness of the ranking R of a person: The more "extreme" the ranking R is, the bigger a quantile needs to be, to represent this ranking in the sense of lying above this ranking. Thus, more "extreme" rankings are tendentially mapped to higher quantiles. So, with the mapping Φ, that maps every ranking R to the smallest quantile above R, we have a qualitative measure of outlyingness. If we furthermore map this quantile $\Phi(R)$ to its associated actual level $\beta(\Phi(R))$ we have additionally a quantitative measure of outlyingness that we call **level function** (denoted by λ), here. These two measures are now illustrated by an example. We use the **words** data set (cf., Fligner and Verducci [1986]), collected under the auspices of the Graduate Record Examination Board. A total amount of 98 college students were asked to rank the five words "Thought" (1), "Play" (2),"Theory" (3),"Dream" (4) and "Attention" (5) according to the strength of association with the word 'Idea'. The raw data are given in Table 1. The rankings are given in ranking notation where e.g., the vector $(1, 3, 4, 5, 2)$ in the first column and the first row means that e.g., the second word "Play" got the rank 3, where rank 1 means least associated and rank 5 means most associated with the target word idea.

Ranking	f	λ	Ranking	f	λ	Ranking	f	λ
(1 3 4 5 2)	1	0.99	(4 2 3 5 1)	2	0.96	(5 1 4 2 3)	6	0.84
(1 4 2 3 5)	1	1	(4 3 5 2 1)	1	0.98	(5 1 4 3 2)	33	0.34
(3 2 5 4 1)	2	0.96	(5 1 2 4 3)	5	0.91	(5 2 3 4 1)	8	0.72
(4 1 2 5 3)	1	0.96	(5 1 3 2 4)	2	0.91	(5 2 4 1 3)	1	0.98
(4 1 5 3 2)	5	0.84	(5 1 3 4 2)	18	0.52	(5 2 4 3 1)	12	0.72

Table 1: Rankings of the words data with their frequencies f and their associated level λ.

The first two rankings (1 3 4 5 2) and (1 4 3 2 5) indicate that here the students possibly misunderstood the ranking and gave rank 5 for the least and rank 1 for the most associated word. Since this is clearly only a guess, we decided to not exclude these two rankings. The levels of these two rankings are 0.99 and 1, respectively, these are the two greatest values in the profile, so these rankings are the most outlying as one would also expect from the statements above. Figure 1 shows the Hasse diagram of the semi-lattice of all arbitrary unions of the orderings in the profile together with the altogether 9 quantiles. The orderings itself are colored red and the id's of the persons with the corresponding ordering are listed below. In Figures 2 and 3 the orderings (and unions of orderings) lying below the 9 different quantiles are shaded blue. Here one can also see that the first 2 rankings are most outlying. The most central orderings are the 33 identical orderings (with associated ranking (5 1 4 3 2)) that rank "Thought" before "Theory" before "Dream" before "Attention" before "Play" and have level 0.34. These orderings coincide with the mode of the profile, here.

Figure 4 shows the commonalities of the orderings below the 9 quantiles. More precisely, for given quantile q the intersection of all orderings of the profile below q is represented via its Hasse diagram. (The quantile q itself would be the union of the orderings below q but this union is generally not transitive and is thus more difficult to represent graphically, but note that because we deal only with total orders, the union can also be reconstructed from the intersection: the edge (C_i, C_j) is in the union iff (C_j, C_i) is not in the intersection.) One can see that the lowest quantile is the 34%-quantile consisting only of the modal orderings. The intersection of orderings below this quantile is thus still a chain. The next quantile is the 52%-quantile that is not a chain anymore, the words "Dream" and "Theory" are now incomparable, meaning that the orderings below the quantile contain both orderings that rank “Theory” before “Dream” (these are exactly the mode orderings) and also orderings that rank “Dream” before “Theory”, but the other words are ranked identically by the 52% of the population represented by the quantile. Here, one can **speculate** that these 52% of the students are consisting both of more scientific-minded students that rank “Theory” before “Dream” and more literary-minded students that rank “Dream” before “Thought”. The red colored edges in the Hasse diagrams indicate, which edges would disappear if we would go to the next higher quantile.

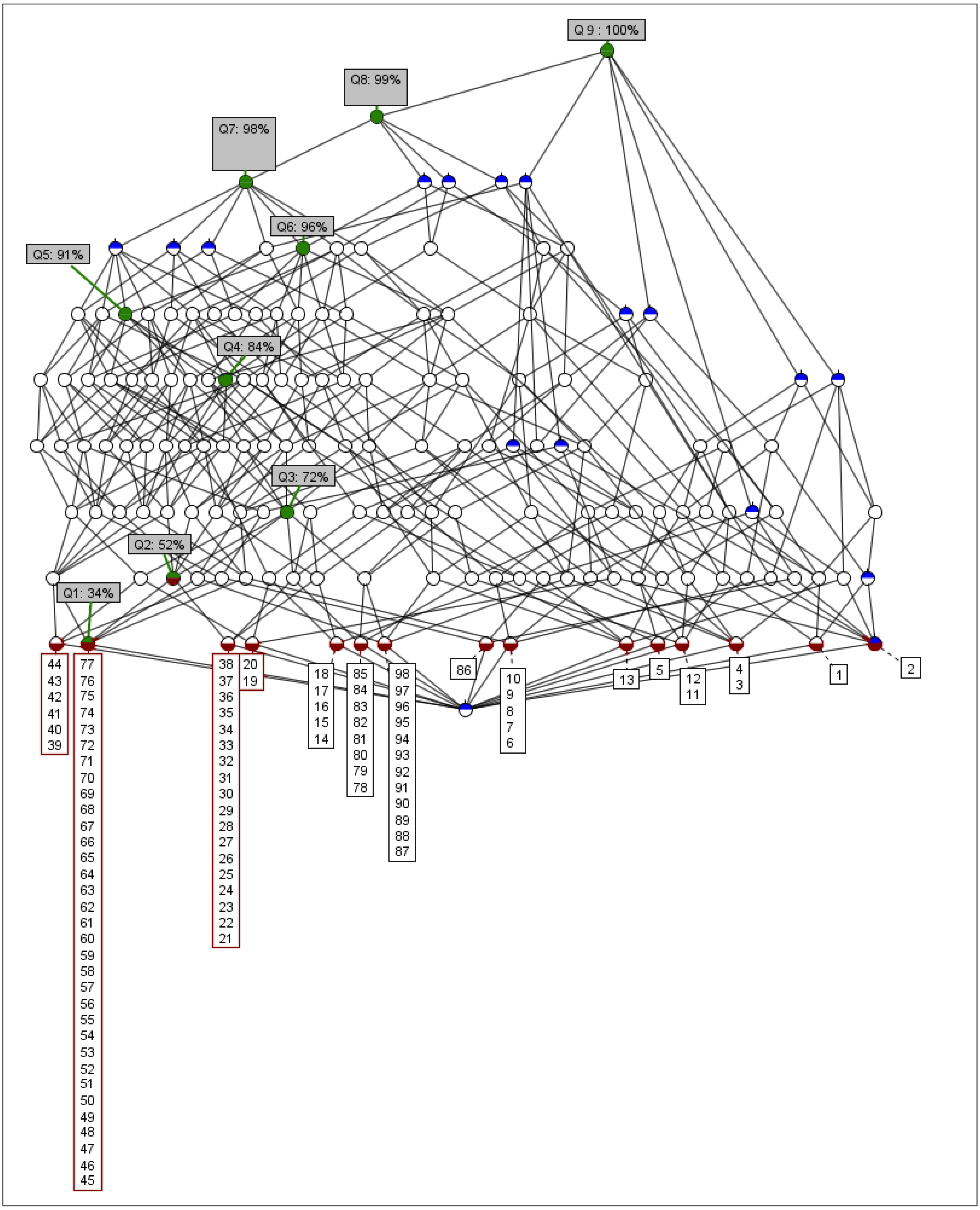

Figure 1: The semilattice of binary relations generated by all unions of the rankings of 98 students of the words dataset, including the 9 quantiles $Q_1, \ldots, Q_9$.

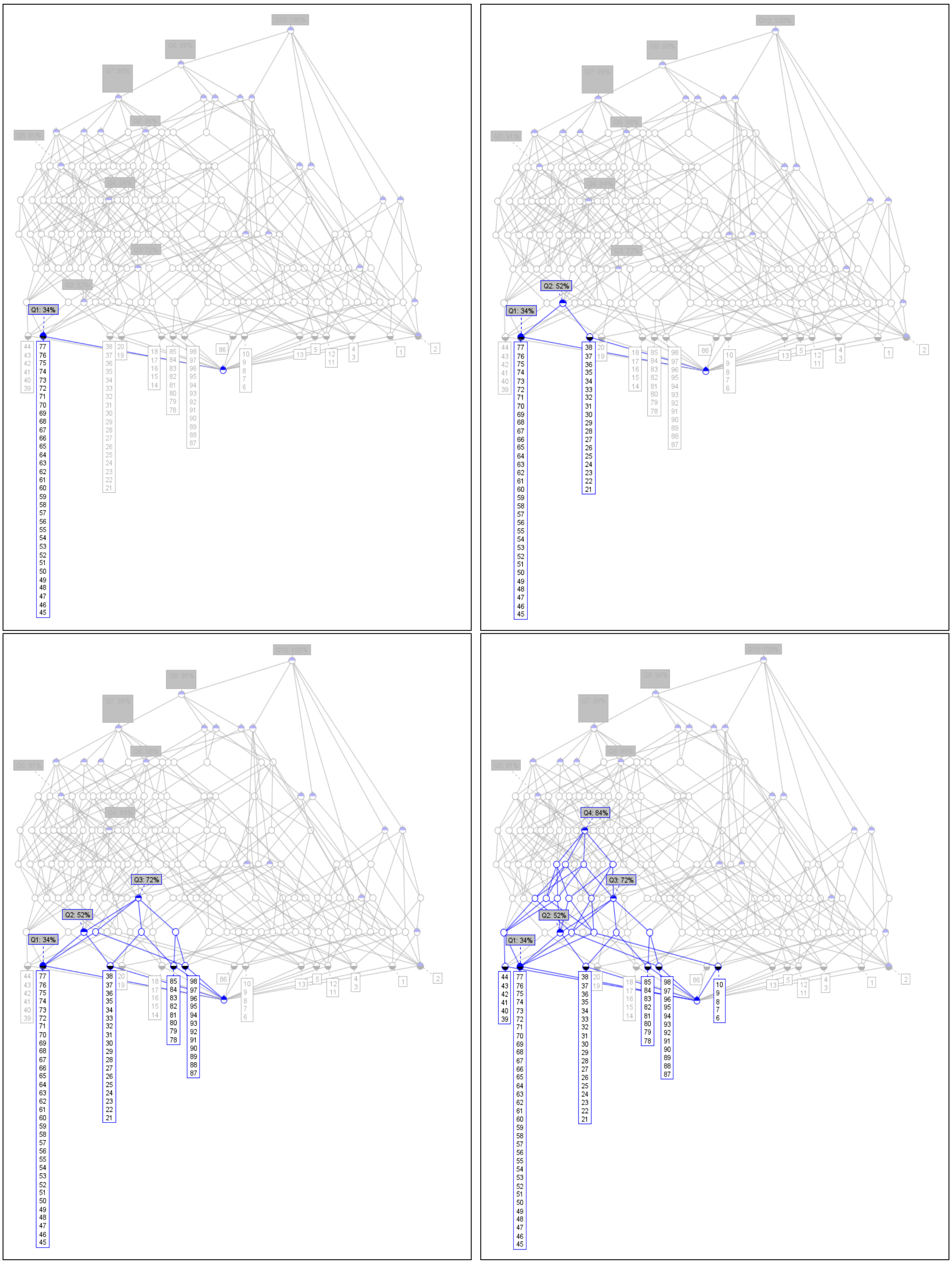

Figure 2: The 34%-, 52%-, 72%- and the 84%-quantile for the words dataset.

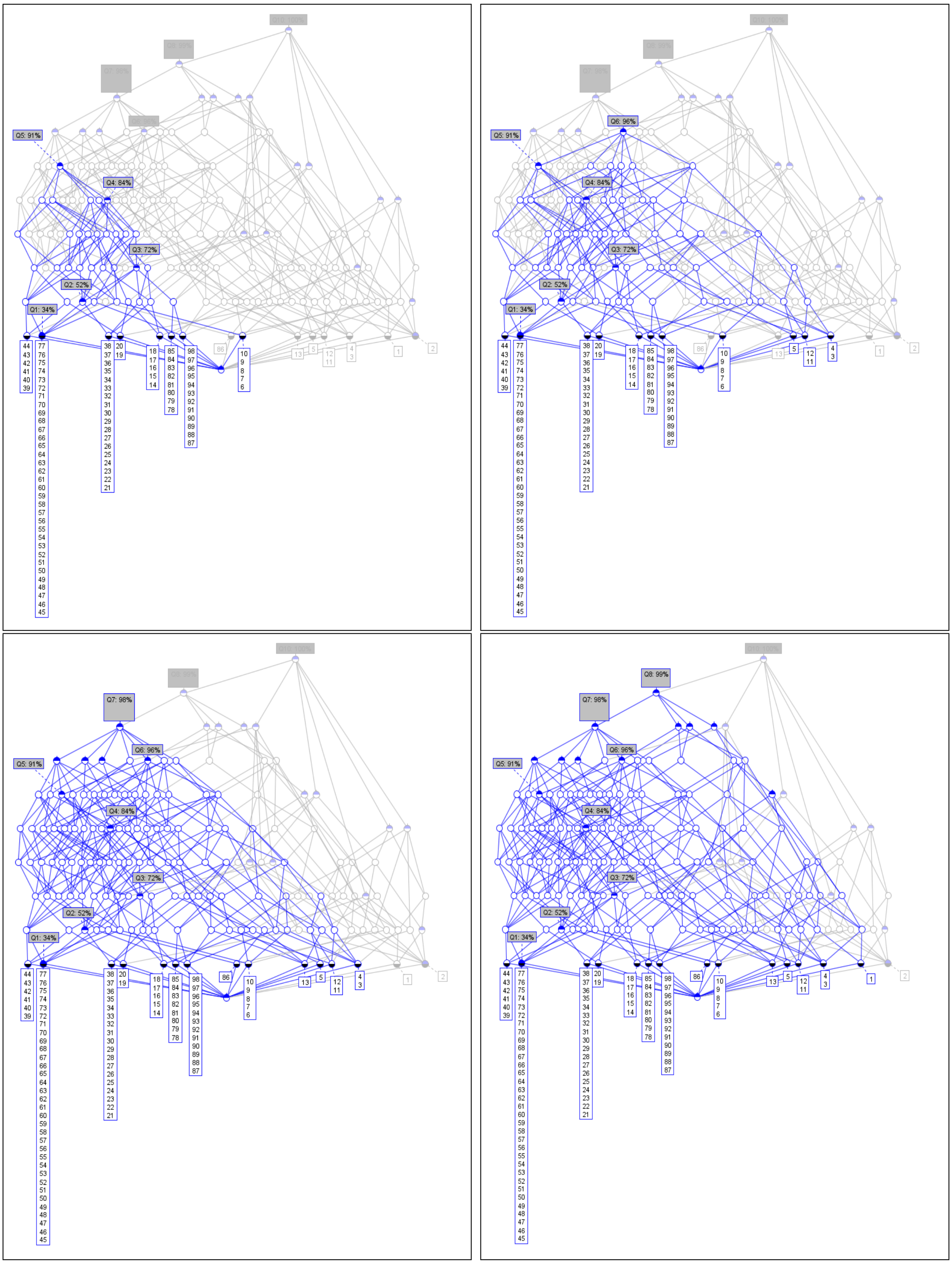

Figure 3: The 91%-, 96%-, 98%- and the 99%-quantile for the words dataset.

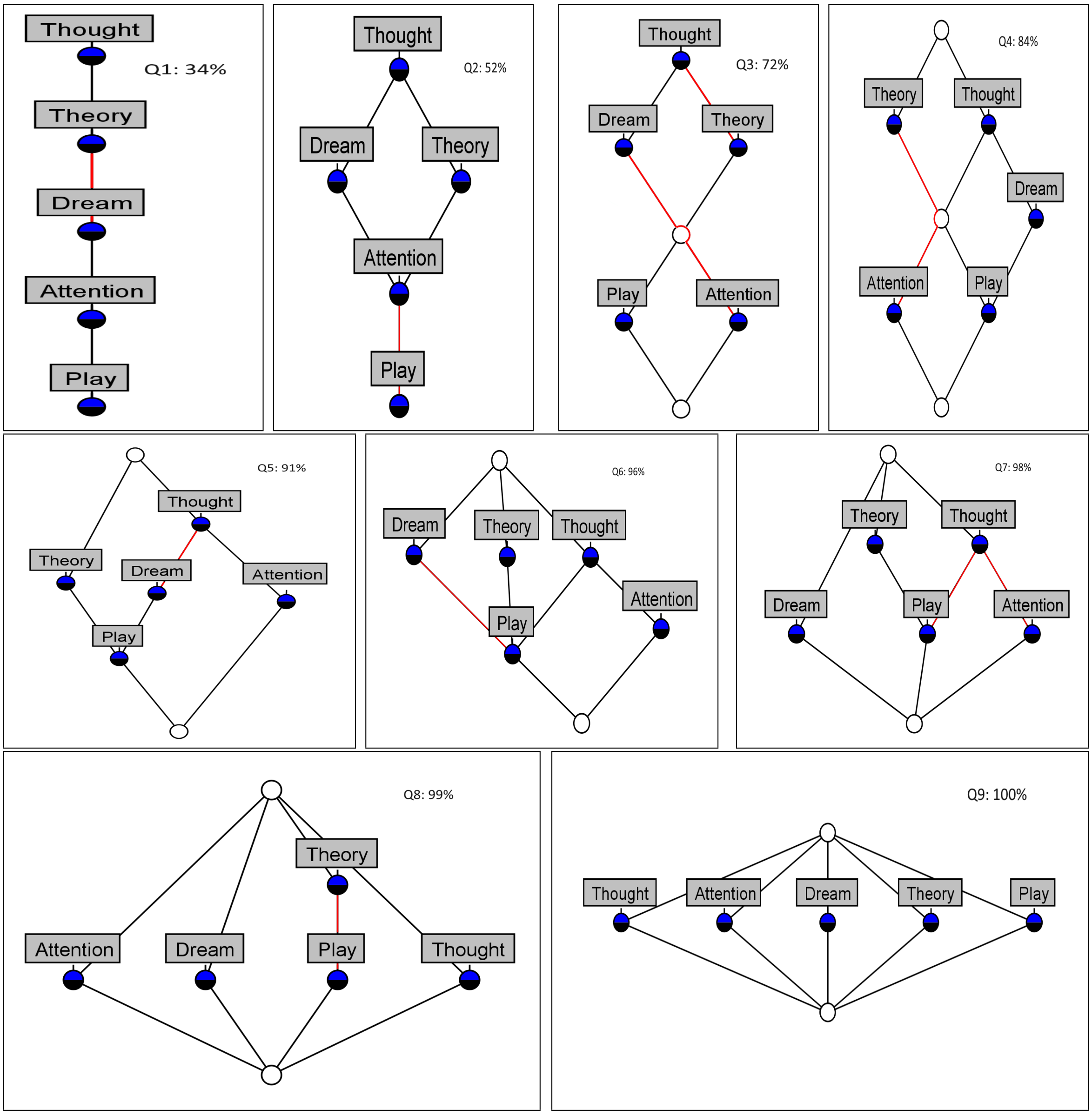

Figure 4: The commonalities of the 9 subpopulations that lie below the 9 quantiles $Q_1, \ldots, Q_9$ for the words dataset.

3 Commonality sharing: An aggregation function for Social Choice Theory based on lattice-valued quantiles

In Social Choice Theory one prominent problem is to aggregate the preference orderings $R_1, \ldots, R_n$ of n voters concerning different candidates $C_1, \ldots, C_q$ in an election to one consensus preference order S that reflects in a fair way the preferences of all voters. The most prominent aggregation rules are the Condorcet method (cf. de Condorcet [1785]) where the consensus relation S prefers candidate C_i to candidate C_j if more than the half of the n voters prefer candidate C_i to candidate j and the Borda count (cf., de Borda [1781]), where the ranks of the consensus order S are defined as the ranks of the mean of the ranks that all persons assign to the corresponding candidates. The herein developed concept of outlyingness for ranking data can also be used to define an aggregation function that we will call "commonality sharing" in the sequel: Given a profile $R = (R_1, \ldots, R_n)$ one can look at the level λ of every ordering in the profile and take the ordering(s) with the smallest level, i.e. the least extreme ordering(s) in the profile. If there is more than one ordering with a minimal level than one could choose either arbitrary one ordering from the set of orderings with minimal level or one could apply another aggregation rule to the profile consisting of the orderings with minimal level. (In the sequel, we will always apply the first approach.) Before we try to further understand, what the so designed aggregation rule actually does and why we call the rule "commonality sharing", we would like to contrast it with a standard way of treating aggregation problems, especially in the field of statistics, and an aggregation rule that is more or less the immediate result of such a way of proceeding.

3.1 A note on mathematical culture

A standard problem in statistics is to aggregate a set of n data points in some space into one single point that in some sense represents to some extent this set. One of the simplest cases is the univariate location problem, where one aggregates a vector of n real numbers $x = (x_1, \ldots, x_n)$ to a single number, for example the mean value of all the n real numbers. This mean value represents the data set to some extent, but there are different possibilities of motivating the use of the mean value. One possibility is to understand the mean value $\bar{x}$ as the unique value, a vector y of n equal real numbers $y_1 = y_2 = \ldots = y_n$ has to have such that the total sum of this numbers is the same as the sum of the given numbers $x_1, \ldots, x_n$. In this understanding, one compares a vector x of n possibly different data points with a vector y of n equal data points that are in some sense comparable to x (i.e. having the same sum) and are more easily to describe (because they are all equal) and thus represent to some extent the original vector x.

Another possibility of getting the mean is to find a real number y that is close to each point x_i in the vector x. For this one has to concretize, what the term "close" exactly means. The usual way to proceed here is to introduce a notion of distance between two points and then to define the distance from y to the data vector x as the sum of the squared distances from y to every point $x_i, i = 1, \ldots, n$. The usage of the squared distances instead of the actual distances is not motivated in this understanding, the usage of the non-squared distances would actually lead to the median, but this is not our point here. This distance-based understanding is often used in statistics to treat more complex situations: One introduces a notion of distance in the data space and then aggregates a vector x to the point y that minimizes the sum of (possibly squared) distances from y to all $x_i, i = 1, \ldots, n$. For the simple case of univariate real valued data the distance to use is very clear and furthermore, from the distances, almost the whole structure of $\mathbb{R}$ can be reconstructed (e.g.: $x = \pm d(0, x), x + y = \pm d(x, -y), -y = \pm w$ with w s.t. $d(y, w) = 2 \cdot d(y, 0)$, ...) and thus the use of the metric is not very problematic. In more complex situations, the introduced distance often does not reflect all the structure in the original space and if only the distance is used for aggregation, useful "information" inherent in the additional structure could possibly be lost. This aspect will now be discussed for the case of preference aggregation functions:

A popular preference aggregation function is the Kemeny-Young rule (cf., Kemeny [1959]). For two ordering R_1, R_2 a distance is introduced as $d(R_1, R_2) = |R_1 \triangle R_2| = \frac{q(q+1)}{2} - |R_1 \cap R_2|$. This distance measure is well motivated if one is willing to accept some axioms of an axiomatic characterization of this distance given in Kemeny [1959]. Note that in the original space of tuples of binary relations there is no obvious reason to understand this space as a space that is equipped with some metric, so the axiom that d is a metric and thus satisfies e.g. the triangle inequality is not necessarily intuitively appealing. The Kemeny-Young aggregation rule now selects as the aggregated ordering the total ordering that minimizes the sum of the (not squared) distances to all orderings in the profile. It is important to note that the distance is introduced and not a priori there and that at the same time, the structure of the space of binary relations that is given more a priori, is ignored to some extent. We will concretize this with a simple example. Consider the candidate set $\{C_1, \ldots, C_6\}$ and the profile $R = (R_1, \ldots, R_5)$ with $R_1 = (6, 2, 1, 5, 4, 3)$; $R_2 = (1, 4, 3, 6, 2, 5)$; $R_3 = (1, 4, 3, 2, 5, 6)$; $R_4 = (4, 5, 1, 3, 2, 6)$ and $R_5 = (2, 1, 3, 5, 6, 4)$ where the relations are given in rank notation. This means that for example preference R_1 gives candidate C_6 rank 3 because the sixth entry in the representing vector for R_1 is 3. The profile R is chosen such that there is a perfect symmetry between R_4 and R_5 if one only looks at pairwise distances between the other orderings in the profile and R_4 or R_5 respectively: As illustrated in Figure 5, the distance between R_4 and all other orderings except R_5 is 6 and the distance between R_5 and all other orderings except R_4 is 6, too. So from a purely metric description one cannot distinguish R_4 and R_5. But if one looks at the profile as a collection of binary relations than there is a clear difference between R_4 and R_5: The distance between R_4 and R_1, R_2 and R_3 is 6 which means in particular that R_4 has $\frac{6 \cdot 7}{2} - 6 = 15$ edges in common with R_1, R_2 and R_3, respectively, the same for R_5. But beyond the distance of pairs of orderings one can also analyze the commonalities of more than two orderings in the profile. Orderings R_4, R_1 and R_2 share altogether 5 edges, ordering R_4 shares with R_1 and R_3 commonly 4 edges and R_4 has with R_2 and R_3 altogether 6 edges in common. For R_5 exactly the same numbers of common edges are obtained, so, until now, if one thinks only in terms of counting common edges, the situation is perfectly symmetric between R_4 and R_5. If we look now at sets of 4 different orderings, the situation changes: R_4, R_1, R_2 and R_3 are sharing 3 common edges whereas R_5, R_1, R_2 and R_3 have only 2 edges in common, so if one had the task of choosing between R_4 and R_5 as a representative ordering of the whole profile, one could say that R_4 has more in common with R_1, R_2 and R_3 than R_5 and since for other subsets of profiles the situation is perfectly symmetric (at least in terms of counting common edges) one could argue that R_4 should be preferred to R_5. This "directed asymmetry" could not be observed by a purely metric analysis and thus a purely distance based approach seems to be missing something, here. If one would actually apply the Kemeny-Young rule to the given profile, then if one would minimize the distance only over the orderings given in the profile, then one would get R_2 as the aggregated order. If one minimizes over all arbitrary total orders than one would obtain the aggregated order $S = (1, 3, 2, 6, 4, 5)$ which is not a member of the profile and is closer to R_5 than to R_4 since it shares 17 edges with R_5 but only 15 edges with R_4. Compared to this, the commonality sharing aggregation method chooses the ordering R_4 as the unique aggregated order. So, there seems to be a conceptual difference between distance based approaches and the commonality sharing approach that tries to use the original structure that is given more naturally when dealing with ranking data. From the given example, distance based approaches seem to miss something of the original structure, but what about the quantile based approach? At this point it is worth to re-translate, what the commonality sharing method, which was based on an outlyingness concept developed in an abstract lattice-theoretic context, actually does in the concrete context of ranking data:

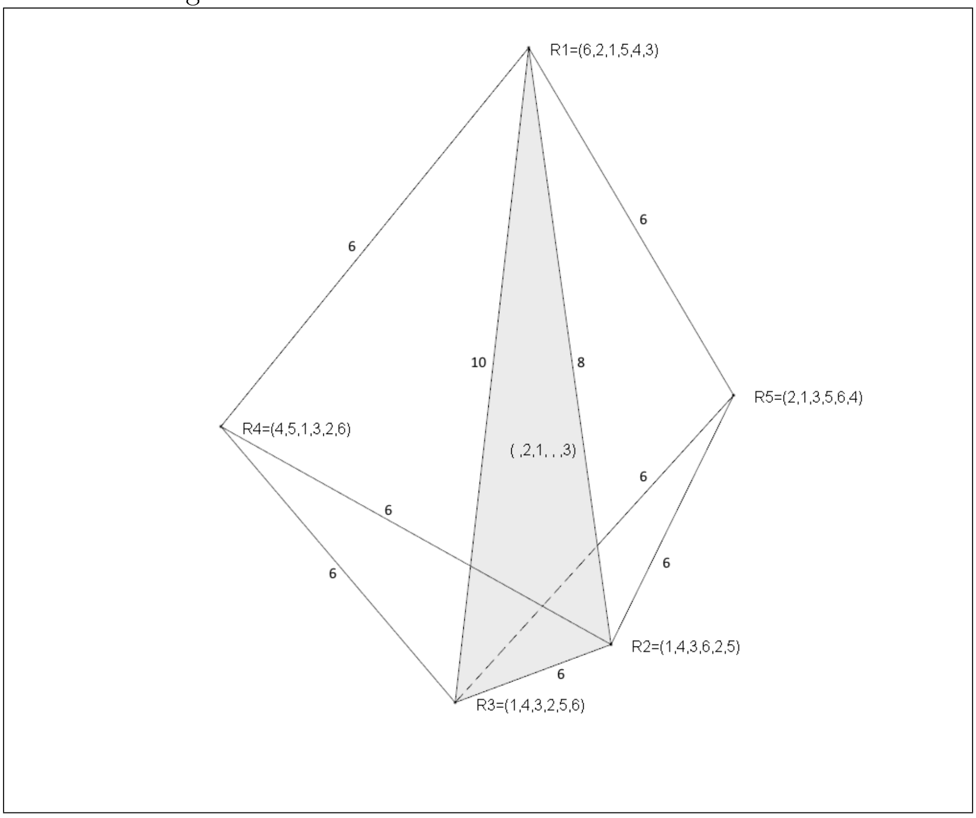

Figure 5: Illustration of the geometric understanding of a profile of 5 rankings.

3.2 A reconstruction of the lattice-valued quantile approach for the case of ranking data

To analyze, how the commonality sharing rule works, we first have to analyze what it means for an order R to lie below a given prequantile q with prelevel α and actual level β. The prequantile q was constructed as the infimum of the set M_α of all elements with at least a proportion α of data points below. Now, note that an element R is below the infimum of a set M of other elements if and only if it is below every element of this set, so R is below q if and only if it is below all elements of M_α. Now, what are the elements in the set M_α? The elements of M_α are binary relations that lie above at least $\alpha \cdot 100\%$ of the orderings in the profile. Lying above here means simply being a superset. The orderings of the profile that lie below an element $z \in M_\alpha$ have some edges in common and other edges are not shared by all the orderings. What means this for z? For pairs C_i, C_j of alternatives where both the edge (C_i, C_j) and the edge (C_j, C_i) are in some of the orderings below z, the element z has to contain these both edges. If all orderings below z agree to share e.g. the edge (C_i, C_j) and not the edge (C_j, C_i) then z also needs only to contain the edge (C_i, C_j). Now, for every z, a further naturally associated element $z' \subseteq z$ is given as the union of all relations in the profile that lie below z. The element z' still lies above all these elements and by construction consists of edges for all pairs of items except the edges that are in none of the orderings below z. Thus, the transposed relation of the complement of z consists exactly of the common edges of all orderings below[2] z. Now, we have to look at all elements of M_α at the same time. To analyze, what it means for an ordering R of the profile to lie below all $z \in M_\alpha$ it suffices to look only at all $z' \in M_\alpha$ that are associated to all $z \in M_\alpha$, because an ordering R of the profile lies below z if and only if it lies below the associated z'. Since z' is constructed as the union of elements of the profile one can directly reinterpret all possible z' as all arbitrary unions of orderings of the profile of a given minimum size $\alpha \cdot n$. (Every z' is a union of at least $\alpha \cdot n$ orderings and for every arbitrary union z of at least $\alpha \cdot n$ orderings we have $z' = z$.) Lying below a prequantile of given prelevel α then means to share all the "non-edges" with every arbitrary union of a minimum size of $\alpha \cdot n$ orderings of the profile. This further means exactly sharing with all sub-populations of minimum size $\alpha \cdot n$ the common edges of these sub-populations.
So, finally, the commonality sharing ordering is the ordering with the smallest level and therefore it lies below the maximal number of quantiles. Thus the commonality sharing ordering can be reconstructed with the following verbalization:

> "The commonality sharing ordering is (are) that ordering(s) that share with every sub-population of minimum-size k, where k is choosen as small as possible, the commonalities of the corresponding sub-population."

[2]If one would have chosen the "$\supseteq$"-relation instead of the "$\subseteq$"-relation as the relation $\leq$ in the lattice, the construction would appear more natural at this concrete point, but this would have "imperfections" at other points, however, the resulting concepts would be equivalent because we are dealing only with total orders in the profile that have for one edge (C_i, C_j) exactly one "non-edge" (C_j, C_i) and thus we have perfect symmetry between edges and "non-edges".

In the concrete context of Social Choice Theory, a still more informal way of reinterpreting the commonality sharing rule would be to state that

> "In many possible subgroups, including very small and thus very specific ones, the commonality sharer(s) would appear as the least extreme in the sense of sharing with the rest of the group much edges that the rest of the group shares with each other."

Thus, the commonality sharer(s) could also be called mediator(s).[3]
Since not only the commonalities of pairs of orderings but also the commonalities of larger sets of orderings are considered in the commonality sharing rule, we in fact did account for the conceptual subtleties of Section 3.1 to some (maybe still very marginal) extent. Conceptually, the commonality sharing rule seems to be an appropriate profile aggregation rule, at least in some contexts (this depends on the assessment, if the above verbalization corresponds to a desirable job the commonality sharing rule would do in the concrete substance matter context). The next section discusses the computational complexity of computing the commonality sharing rule.

3.3 Computational Issues

From the conceptual description of the commonality sharing rule one could expect that its actual computation is rather difficult. However, it turns out that the computation of the commonality sharing rule is actually simple, more precisely, it can be computed in $\mathcal{O}(n \cdot q^2)$ time. To see this, we have to firstly analyze, how the elements of the set M_α look like. For every possible edge $e = (C_i, C_j) \in C \times C$ let α_e denote the proportion of orderings in the profile that actually contain this edge and let z_e denote the special element $z_e = C \times C \backslash \{(C_j, C_i)\}$. Note that different prelevels α can lead to the same quantile with the same actual level β. Actually, it suffices to look only at the special values $\alpha \in A := \{\alpha_e \mid e \in C \times C\} \cap (0.5, 1]$. (Note in particular that for edges $e = (C_i, C_j)$ with $\alpha_e \leq 0.5$ there is no ordering below $\bigwedge M_{\alpha_e}$ because both z_e and $z_{e'}$ with $e' = (C_j, C_i)$ are in M_{α_e}.) For some given $\alpha \in A$ and an edge $e = (C_i, C_j)$ with $0.5 \leq \alpha_e < \alpha$, all elements of M_α contain both the edge $e = (C_i, C_j)$ and the edge $e = (C_j, C_i)$. For edges with $\alpha_e \geq \alpha$ the associated element z_e is a special element of M_α. Thus, we have $q := \bigwedge M_\alpha = \bigcap \{z_e \mid e \in C \times C, \alpha_e \geq \alpha\}$. An ordering R_i of the profile is then below a prequantile q with prelevel α iff all edges $e \in R_i$ satisfy $\alpha_{e'} \ngeq \alpha$ which is equivalent to $1 - \alpha_e \ngeq \alpha$ or $\alpha > 1 - \alpha_e$. The largest prelevel α such that the associated quantile q does not lie above R anymore can thus be calculated as $\alpha = \max\{1 - \alpha_e \mid e \in R_i\}$. Since the ordering of the largest such prelevels $\alpha \in A$ and the corresponding actual levels is identical it suffices to compute these largest prelevels to determine, which ordering(s) of the profile

[3] Note also, that there is a neat relationship between the median and the least outlying data point in classical oytlyingness concepts of multivariate analysis. Especially, for univariate data, the least outlying data point and the median are essentially the same.

has (have) the smallest actual level. The calculation would then consist in computing the q^2 values for the α_e's and the n border-case prelevels $\alpha = \max\{1 - \alpha_e \mid e \in R_i\}$ which would take all in all $\mathcal{O}(n \cdot q^2 + n \cdot q^2) = \mathcal{O}(n \cdot q^2)$ time. An interesting point is here, that if we would not compute the ordering in the profile minimizing the level function, but instead that order of the set of all possible total orders that minimizes the level function, than this computation could still be done in polynomial time. This is a further difference to the computation of the Kemeny-Young consensus ordering(s) (i.e., the ordering(s) with the smallest sum of distances to all orderings in the profile), which is NP-hard still for $n \geq 4$ (cf., e.g., Bartholdi et al. [1989]).

4 Application example: Analysis of wisdom of the crowd data

In this section we analyze a data set from the analysis of the "wisdom of the crowd" phenomena (cf., Surowiecki [2005], Galton [1907]), analyzed in Lee et al. [2014][4]. The wisdom of the crowd phenomena is present if, roughly speaking, the aggregated judgment of a group of individuals about a substance matter question results in an estimate that is closer to the truth than most (or all) individual estimates from which the aggregate estimate was based. One of the first instances of this phenomena was described in Galton [1907], who surveyed English fair-goers in 1906. The estimates of the dressed weight of an ox given by the fair-goers, when aggregated via the median, yielded an estimate of 1207 lb, which was very close to the true weight of 1198 lb, the relative error was actually only 0.8 percent. In Lee et al. [2014] the wisdom of the crowd phenomena was studied for the case of ranking data. Here, the participants of the experiments had to rank for example the 44 presidents of the United States of Amerika according to their chronological order of presidency. The analysis involved altogether 23 experiments including also prediction tasks. An important difference to the aggregation of orderings in Social Choice Theory is that in the context of the wisdom of the crowd phenomena, a ground truth is underlying and one can compare the aggregated orderings to this ground truth.

4.1 The presidents data set

We will now analyze the experiment involving the ordering of the former 44 US presidents. This experiment is especially difficult to analyze, because we have a large number of $q = 44$ items to order and at the same time only a very small number of $n = 26$ participants of the experiment. Because of this, we have only a few number of 4 quantiles compared to a very high number of up to $2^{26} = 67108864$ elements in the semi-lattice induced by all possible unions of orderings. The 4 obtained quantiles are depicted in Figure 8 till Figure 11. Because of the high number of 44 presidents, the corresponding Hasse graphs become quickly hard to read out (cf., Figure 7). Thus, we decided to not represent the partial

[4] The data are available under http://webfiles.uci.edu/mdlee/LeeSteyversMiller2014Data.zip.

intersection of the orderings below the corresponding quantiles via its Hasse graph but represent the strict part of the intersection (this is simply the intersection without the diagonal $\Delta C := \{(a,a) \mid a \in C\}$) with methods of formal concept analysis. Therefore we very briefly sketch the basics of formal concept analysis.[5]:

The basic structure in formal concept analysis is a formal context $\mathbb{K} = (G, M, I)$ consisting of a set G of objects, a set M of attributes and a binary relation $I \subseteq G \times M$. For an object $g \in G$ and an attribute $m \in M$ we interpret $(g,m) \in I$ as object g has attribute m. The theory of formal concept analysis formalizes the notion of a concept as a special pair (A, B) of sets, where $A \subseteq G$ is the extent of the concept consisting of all objects that belong to the concept and $B \subseteq M$ is the intent of the concept consisting of all attributes that all objects of the concept have in common. More precisely, a formal concept is a pair (A, B) with $A \subseteq G$ and $B \subseteq M$ satisfying:

1. $\forall g \in A \forall m \in B : (g,m) \in I$
2. $\forall m \in M : (\forall g \in A : (g,m) \in I) \Longrightarrow m \in B$
3. $\forall g \in G : (\forall m \in B : (g,m) \in I) \Longrightarrow g \in A$.

On the set of all formal concepts of a given formal context one can introduce the sub-concept relation via $(A, B) \sqsubseteq (C, D) \iff A \subseteq C \& C \supseteq B$. The set of all formal concepts together with the sub-concept relation forms a complete lattice that is called the formal concept lattice. From the formal concept lattice one can reconstruct the given formal context. Since the formal concept lattice is an ordered set, we can display it via its Hasse graph. A usual way to label the formal concepts is to not write the whole extent and intent at the concepts but to write the name of an object only at the most specific concept containing it and the name of an attribute only at the most general concept with this attribute. For a given concept displayed in the Hasse graph one can then read out the extent as the objects written on all sub-concepts of the given concepts (including the given concept itself) and the intent as the attributes written on all super-concepts of the given context. The usage of the Hasse graph of the formal concept lattice will help us in displaying the quantiles in a more readable way. In our concrete situation we take both as the objects and as the attributes the set of the 44 presidents. The incidence is the strict part of the intersection of the orderings below a given quantile. In the diagrams showing the concept lattices of the corresponding quantiles, the objects are the white-boxed names of the 44 presidents and the attributes are the gray boxed names of the presidents. In our context, the formal concepts could be understood as some kind of vague "virtual time cutting points". The attributes above such a given "cutting point" are the presidents that, due to the common opinion of the persons below the given quantile presided before this "virtual cutting point". Analogously, the objects below a given "cutting point" are the presidents that are assumed to have presided after the "virtual cutting point" due to the

[5] A brief introduction to formal concept analysis is given, e.g., in Davey and Priestley [2002], a more comprehensive introduction is Ganter and Wille [2012].

common opinion of the persons below the given quantile. A president x is then commonly ranked before another president x' (meaning that x presided before x') if and only if the formal concept with a white box containing the name of president x' is (directly or indirectly) connected via an increasing path to the concept with the gray boxed name of president x. For a given president x the presidents assumed to have presided before president x are given by the gray boxed presidents above the concept with a white box containing president x, the presidents assumed to have lived after president x are the white-boxed presidents below the concept with a gray boxed containing president x. Additionally, if the gray-boxed and the white-boxed label of a given president are closer to each other, the ambiguity for the common localization of this president is more or less smaller. Presidents that are labeled in the same (gray or white) box are incomparable w.r.t. their common localization. The motivation for using the formal concept lattice of the "$<$"-relation instead of displaying the "$\leq$"-relation via its Hasse graph lies in the following considerations:

One could presume that the persons rank the presidents by either directly comparing pairs of presidents or by relating the time of presidency of presidents to external historical events. For the latter, both the time of presidency and the related historical events are no precise time points, but time spans whose exact start- and endpoints are often only known imprecisely. If only the later way of ranking would be present, the underlying orderings of the persons would actually be only partial orderings that are furthermore so-called interval orderings. An interval order is simply an order relation $(X, \leq)$ that can be represented by real-valued intervals via a mapping $f : X \longrightarrow I(\mathbb{R}) : x \mapsto [l(x), u(x)]$ satisfying $x < y \iff u(x) < l(y)$, where $I(\mathbb{R})$ denotes the set of all real-valued compact intervals of $\mathbb{R}$.

Clearly, if in fact such an interval order is underlying, one still cannot observe it directly because the persons are forced to give a total order. However, if for example all persons would have the same interval order underlying and would provide a total order from the underlying interval order by picking random points in the underlying intervals, then one could asymptotically rediscover the interval order as the intersection of all given orders of the persons. Of course, this is a very schematic understanding and we do not expect that all persons have the same underlying interval order and that they pick randomly points from the intervals. Additionally, we do expect that some parts of some of the orderings are obtained by a direct comparison of two presidents. For example, it is expectable that the persons could guess that Grover Cleveland 1 has to be ranked before Grover Cleveland 2 without any further knowledge (including the knowledge hat Cleveland 1 and Cleveland 2 is actually the same president), but simply because of the numbers 1 and 2. (In fact, all persons rank Grover Cleveland 1 before Grover Cleveland 2 in this data set.)

We use the formal concept lattice approach instead of displaying directly the Hasse graph, because if we would actually have an interval order, then the associated formal concept lattice has a very simple structure whereas the Hasse graph could be still very complex. More precisely, the formal concept lattice of the strict part $<$ of an order $\leq$ is a chain if and only if the order $\leq$ is an interval order (see, e.g., [Ganter and Wille, 2012,

p. 237, propostion 103], where the more general notion of a Ferrers relation is used instead of the notion of an interval order, cf., [Ganter and Wille, 2012, p. 244]). Opposed to this, the complexity of the Hasse graph of the original $\leq$ relation, measured e.g. by the order dimension, could become arbitrary high (see Bogart et al. [1976]). This is illustrated in Figure 6, where the interval order of all intervals with endpoints in the set $\{0, 1, \ldots, 10\}$ is displayed. If now the quantiles of the profile are not interval orders but maybe very close to interval orders, then one could expect that the associated formal concept lattice is still easy enough to read out and this is the reason why we decided to display the formal context lattices that actually look easy enough to read out for the presidents data set.

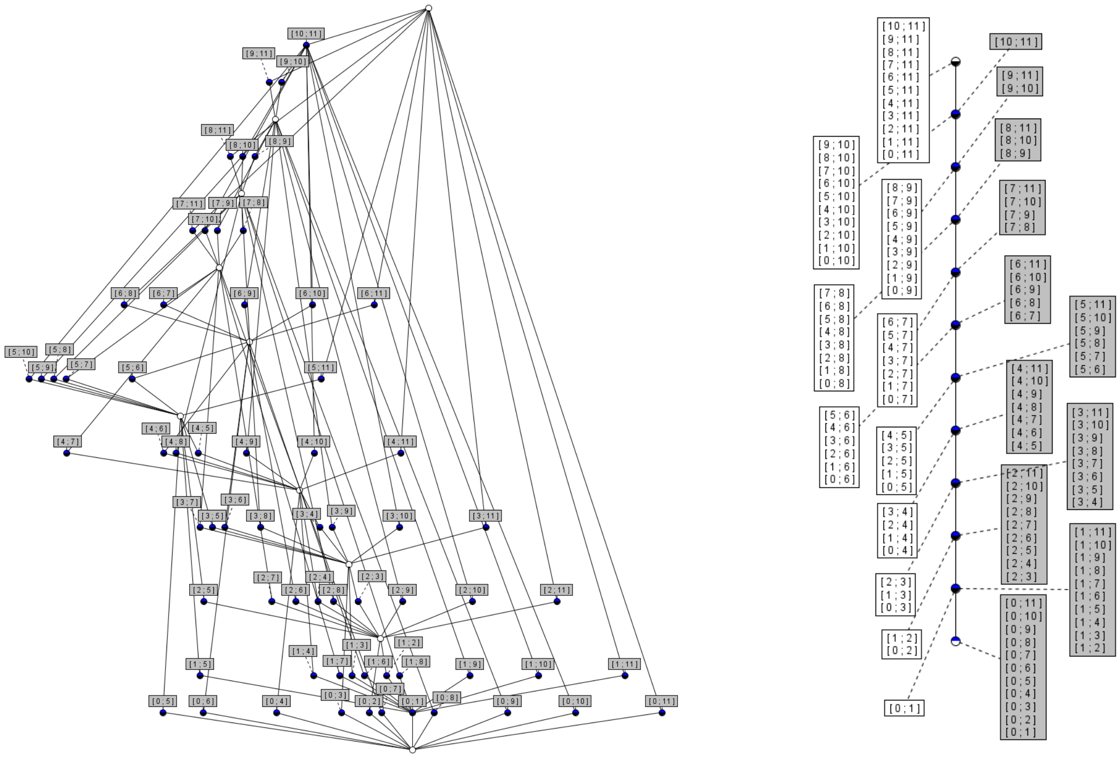

Figure 6: The interval order $(I(\{0, 1, \ldots, n\}), \leq)$ of all intervals with endpoints in the set $\{0, 1, \ldots, 11\}$ represented via its Hasse graph (left) and via the formal context lattice of the strict part $<$ (right).

Now we want to describe some aspects of the presidents data set that become visible by the 4 quantiles:

i) The most central ordering is not unique, there are two orderings with smallest level $\lambda = \frac{2}{26}$. While the two most central orderings have some edges in common that are not in accordance with the true chronological ordering of the presidents, the orderings below the second smallest quantile with level $\lambda = 27\%$ have already no "wrong" common edges anymore. This can be easily seen by looking at the incidence matrices of the intersection of the orderings below the given quantiles given in Figure 12.

ii) The strict part of the intersection of the orderings in the whole profile still contains altogether 337 edges (the strict part of a total order would contain 946 edges) and all these edges are in accordance with the true chronological order of the presidents.

iii) The already mentioned speculation that some presidents are ranked via a direct comparison of pairs of presidents is also suggested by the 4 formal concept lattices of the 4 quantiles: In every concept lattice there is exactly one formal concept, a "virtual cutting point" that could be interpreted as the "time after Cleveland 1 and before Cleveland 2" since it has exactly only Cleveland 1 as a grey-boxed attribute and Cleveland 2 as a white-boxed object and no further attributes or objects.

iv) All partial orderings induced by the quantiles are no interval orders since the corresponding concept lattices are no chains. One can characterize an interval order as an order where there exist no quadruple of elements (x, y, u, v) with $x < y$, $u < v$ and no other comparabilities between these four elements. So, we can analyze, which quadruples (x, y, u, v) are counterexamples that show that our orderings are no interval orderings. Firstly, let us think about how such a quadruple can possibly arise. Such a quadruple (x, y, u, v) consists of a pair of pairs of elements where within each pair the elements are comparable and the elements of different pairs are incomparable. This situation can possibly arise, if we have a quadrupel (x, y, u, v) of presidents, where both x and y, as well as u and v are simple to compare, and all other pairs of presidents are not easy to compare. For example, it is easy to compare Cleveland 1 and Cleveland 2 because of the numbers. But it could be the case that there is another pair of simply to compare presidents that is nevertheless difficult to compare to Cleveland 1 and Cleveland 2, because one possibly knows nothing about Cleveland 1 and Cleveland 2 beyond the numbers. Table 2 shows for all 4 quantiles the number of (unordered) quadruples that are counterexamples showing that the given quantile is not an interval order. Additionally, the number of counterexamples every president is involved in is given for every quantile. One can see that in fact Cleveland 1 and Cleveland 2 are involved in counterexamples for every quantile, especially for the 42% and the 100% quantile they are actually involved in the most counterexamples.

v) All quantiles are very heterogeneous w.r.t. time: The last 5 presidents Ronald Reagon, George H.W. Bush, William Clinton, George W. Bush and Barack Obama are identically ranked correctly by all the participants of the experiment. This is maybe due to the fact that the participants are more aware of presidents of the immediate past. Furthermore, all participants rank George Washington as the first president which may be due to historical knowledge. Opposed to this, presidents that did preside neither in the immediate past nor in the early phase starting with the declaration of independence and the first president George Washington are ranked very differently by the participants leading to vanishing edges in the intersections of orderings below the 4 quantiles. Also the deviation of the quantiles from an interval order is heterogeneous, only for the very early and very late presidents the quantiles are interval orders.

no.	president	7.7%	27%	42%	100%
1	George Washington				
2	John Adams				
3	Thomas Jefferson				
4	James Madison			17	12
5	James Monroe			17	2
6	John Quincy Adams			31	11
7	Andrew Jackson	6	1	25	2
8	Martin Van Buren	16	1	6	
9	William Henry Harrison	14	27	4	
10	John Tyler	2	1	6	
11	James Knox Polk	14	4	4	2
12	Zachary Taylor	1	27	7	
13	Millard Fillmore	6	36	15	
14	Franklin Pierce	15	26	12	2
15	James Buchanan		1	4	2
16	Abraham Lincoln	19	2	14	2
17	Andrew Johnson	1	45	18	2
18	Ulysses S. Grant	36	24		2
19	Rutherford B. Hayes	29	14	16	2
20	James Garfield		31	4	
21	Chester Arthur	14	21	1	
22	Grover Cleveland 1	9	14	46	25
23	Benjamin Harrison	2	45	19	2
24	Grover Cleveland 2	6	49	55	25
25	William McKinley	5	78	16	8
26	Theodore Roosevelt	14	41	28	
27	William Howard Taft	1	44	18	3
28	Woodrow Wilson	3	6	7	3
29	Warren Harding	19	16	2	
30	Calvin Coolidge	39	1 6	25	
31	Herbert Hoover	8	15	41	2
32	Franklin D. Roosevelt	17	17	25	
33	Harry S. Truman		3	22	3
34	Dwight Eisenhower	4	9	11	13
35	John F. Kennedy	1	7		19
36	Lyndon B. Johnson	8	15	21	4
37	Richard Nixon		11		
38	Gerald Ford	1	92	24	4
39	James Carter	1	33	24	
40	Ronald Reagan		4		
41	George H.W. Bush				
42	William Clinton				
43	George W. Bush				
44	Barack Obama				
	overall number of counterexamples	89	230	153	38

Table 2: Overall number of counterexamples and number of counterexamples in which every president is involved in for all 4 quantiles (zeros are omitted).

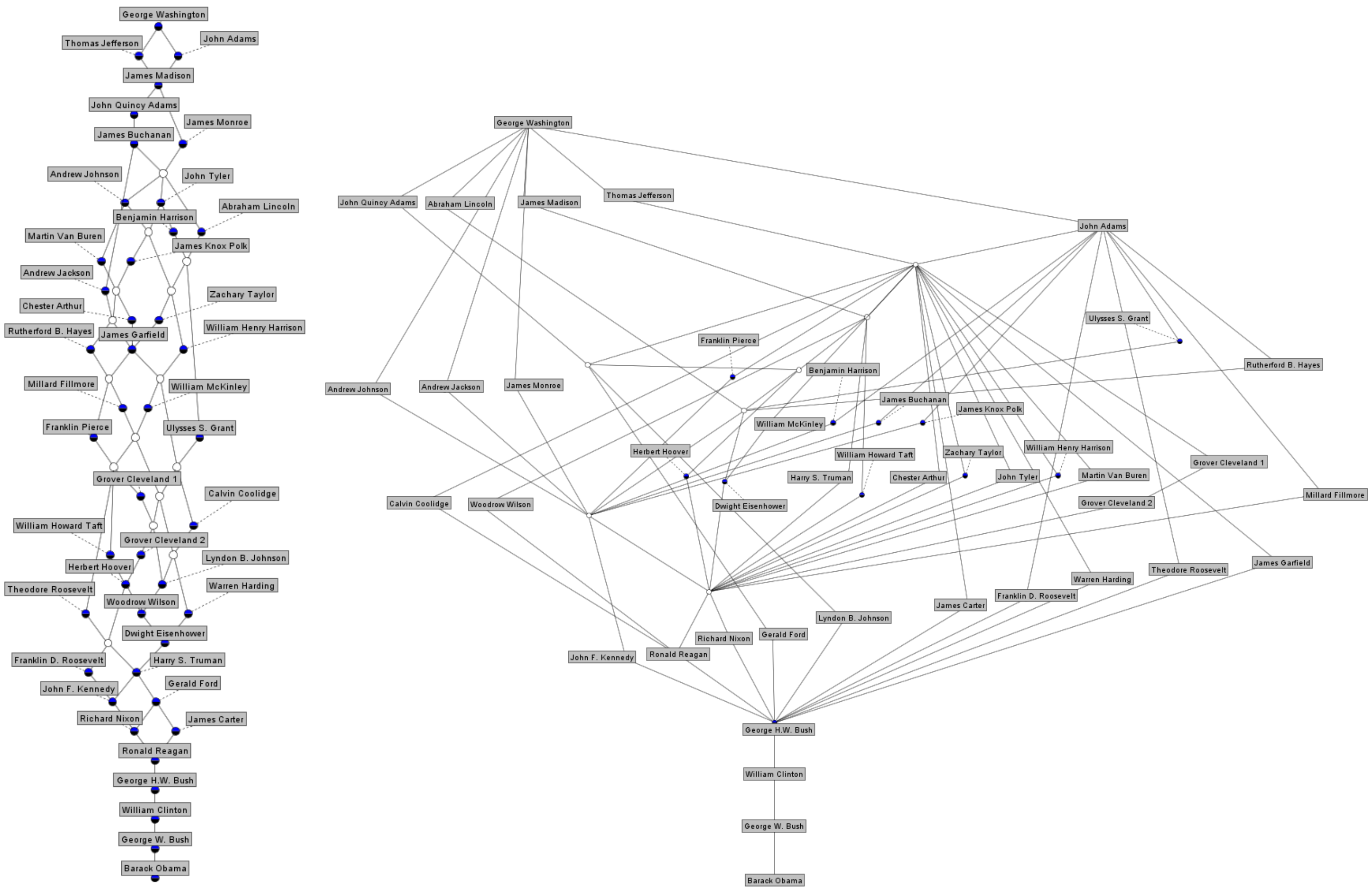

Figure 7: The 7.7% quantile representing the 2 most central orderings of the 44 presidents (left). The 100% quantile (right) represents the commonalities of all 26 persons of the population. Both quantiles are represented via the Hasse graph of the intersection of the orderings below the corresponding quantile.

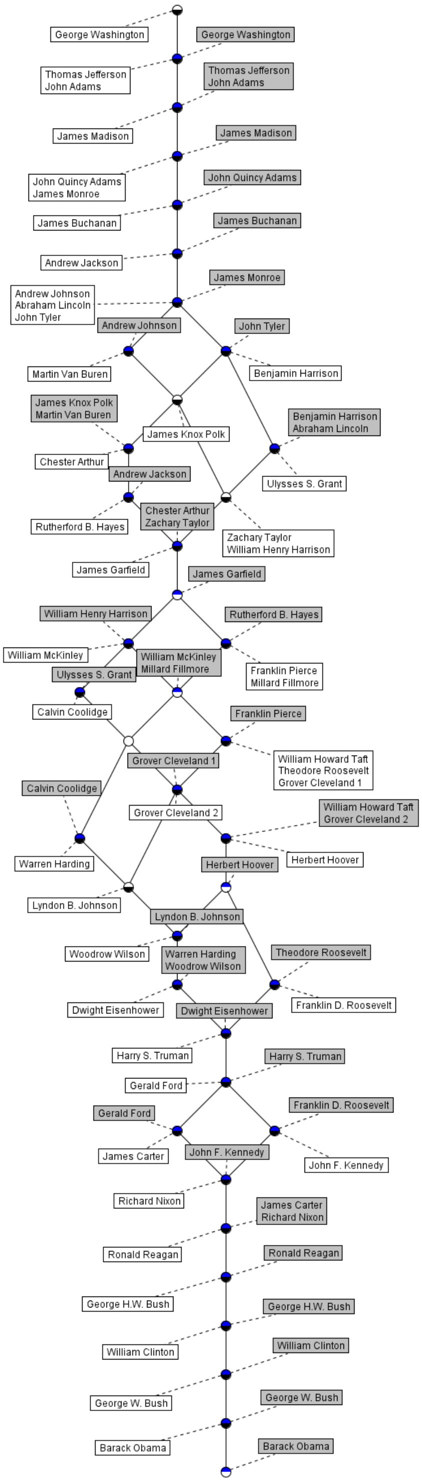

Figure 8: The 7.7% quantile for the president data, represented via the formal concept lattice generated by the strict part of the intersection of the orderings below the 7.7% quantile.

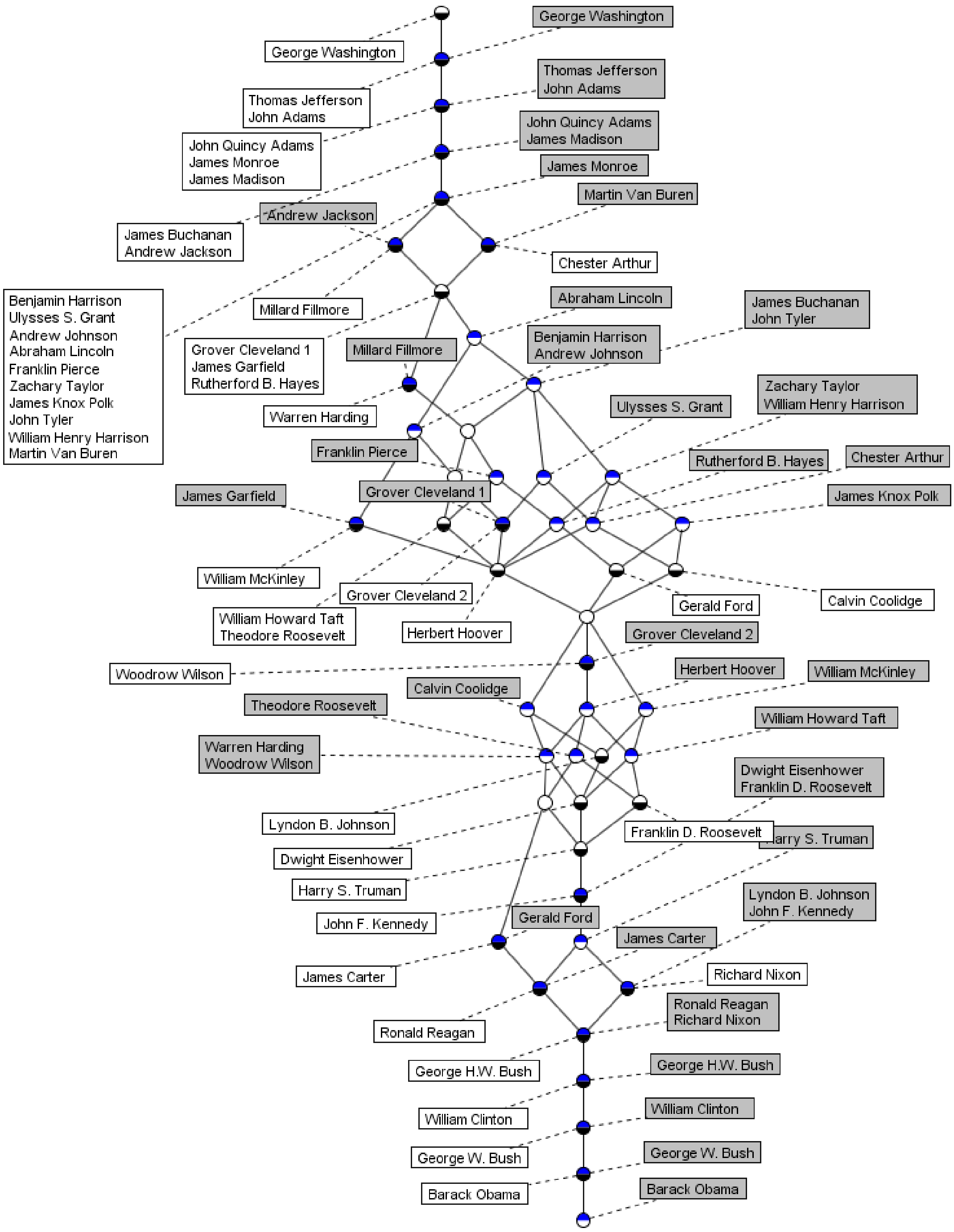

Figure 9: The 27% quantile for the president data, represented via the formal concept lattice generated by the strict part of the intersection of the orderings below the 27% quantile.

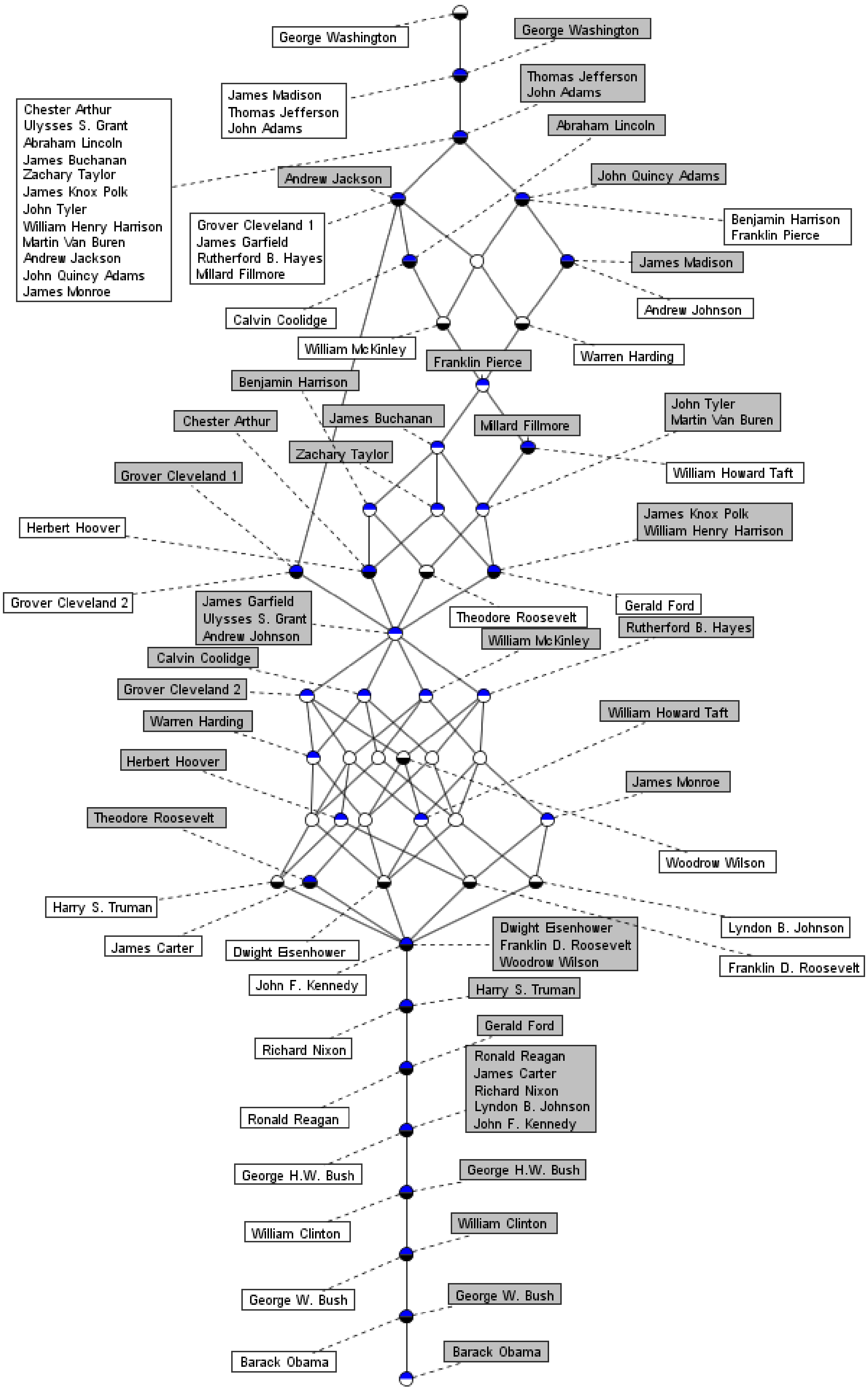

Figure 10: The 42% quantile for the president data, represented via the formal concept lattice generated by the strict part of the intersection of the orderings below the 42% quantile.

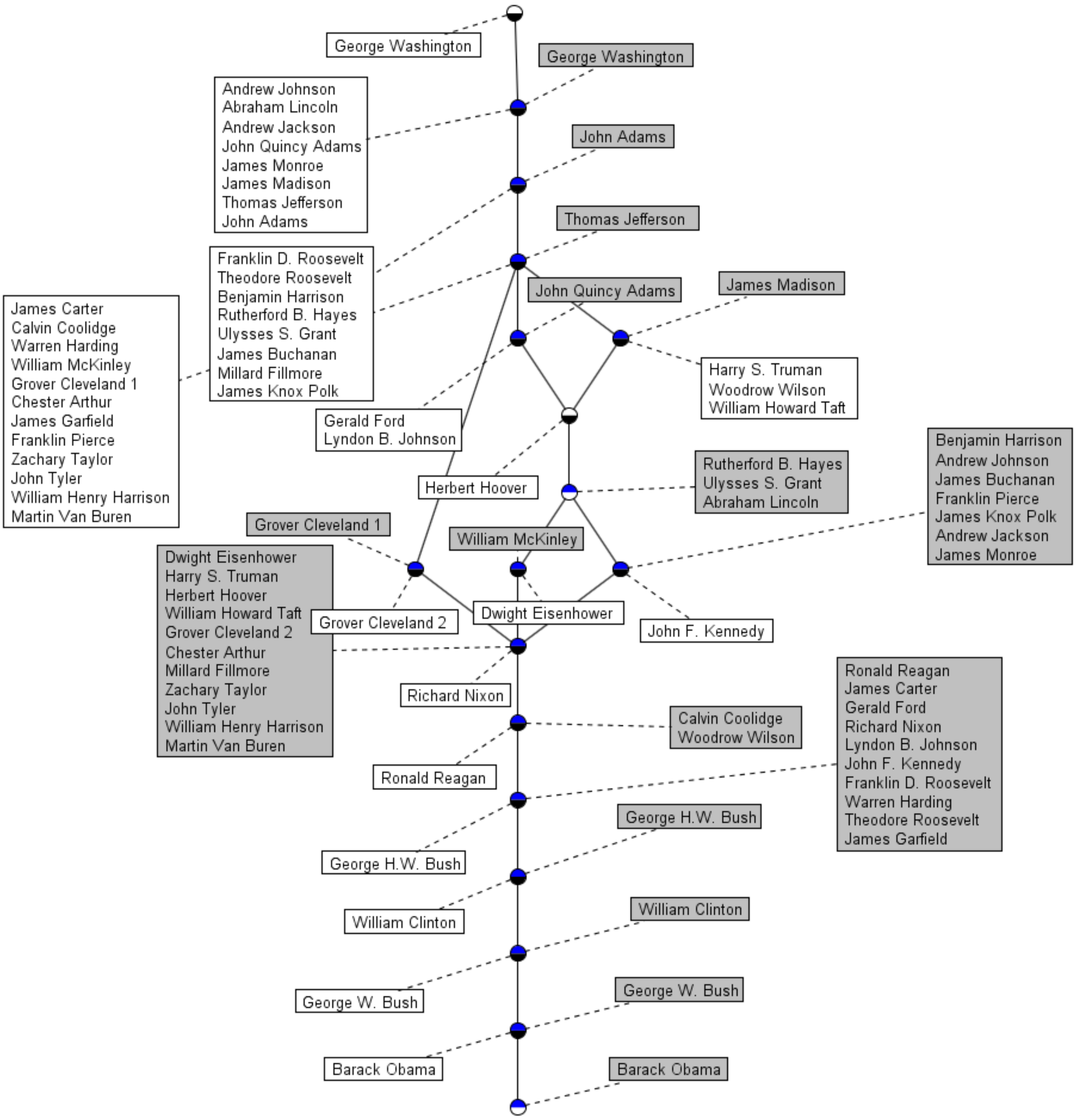

Figure 11: The 100% quantile for the president data, representing the commonalities of the whole population via the formal concept lattice generated by the strict part of the intersection all orderings.

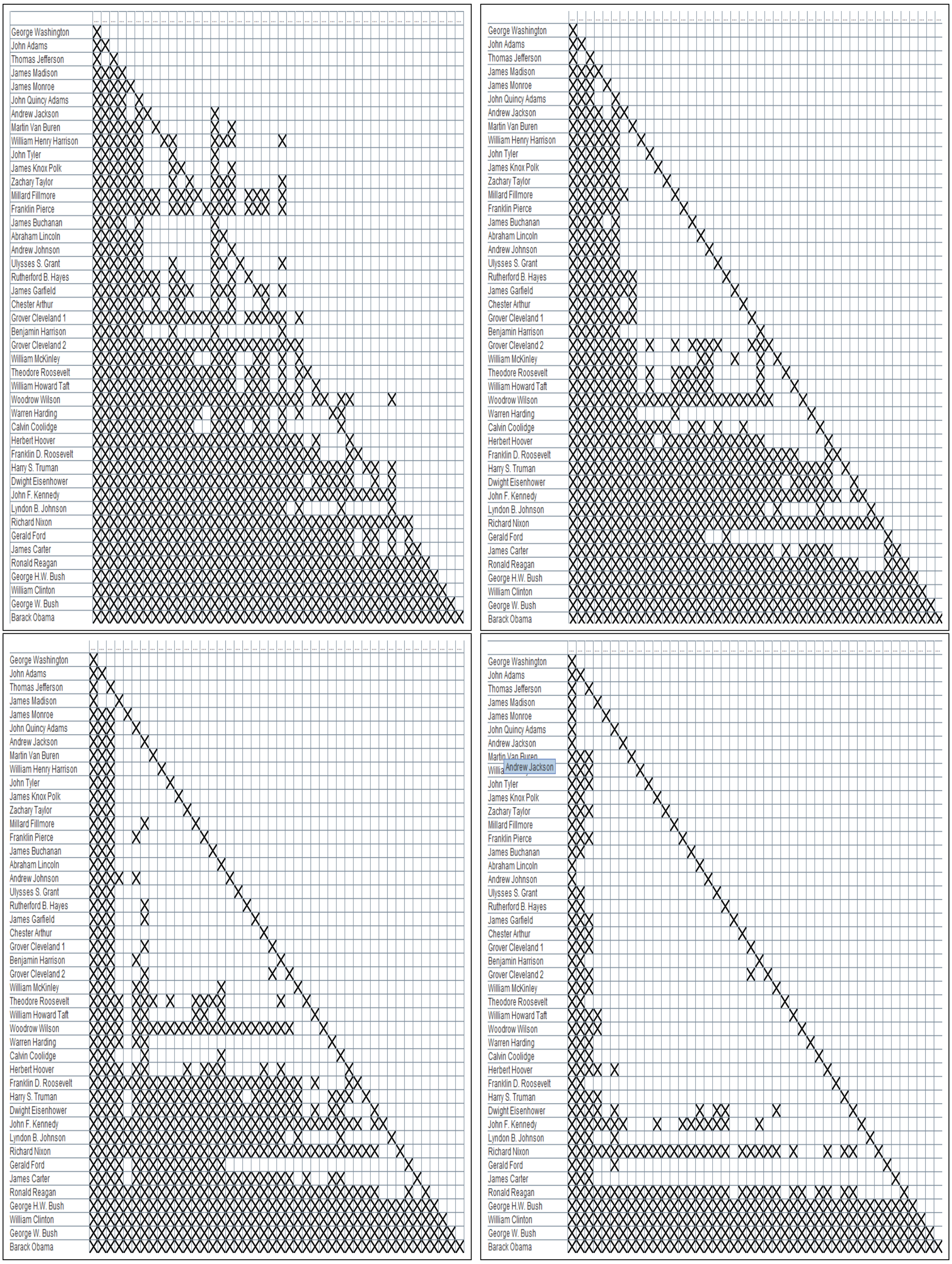

Figure 12: The incidence matrices of the intersection of the orderings below the 7.7%-, 27%-, 42%- and the 100%-quantile. The rows and columns are already ordered by the true chronological order, so a lower triangular matrix of crosses corresponds to a total order that is in full agreement with the actual true order.

4.2 A short analysis of statistical models for ranking data

We now want to relate the results of our descriptive analysis of the presidents data set to two selected statistical models for ranking data. We focus our analysis on the points ii) and iv) of Section 4.1 and are interested in the question, how capable the statistical models are in capturing the structural characteristics that were obtained in the descriptive analysis of the presidents data set.

In the following two sections, for two statistical models we determine via simulations, how probable it is, to observe a profile with specific characteristics like observed for the presidents data set. We look here only at characteristics of the intersection of the whole profile, specifically we analyze the cardinality of the intersection, which gives a rough impression about the homogeneity of the profile, and the proximity of the intersection to an interval order.

4.2.1 Insertion Sorting Rank model

The first statistical model we analyze is the Insertion Sorting Rank model (ISR) described in Biernacki and Jacques [2013]. In this model, it is assumed that the participants order the items by applying the (straight) insertion sorting algorithm to order the items. It is furthermore assumed that every paired comparison involved in the sorting algorithm is exposed to the possibility of making a mistake which will occur with a fixed probability ε, independently for every paired comparison. The ISR model has two parameters: the true ordering μ of the items and the error-probability ε. For a profile randomly drawn from the ISR model we analyze the intersection of all orderings in the profile, we inspect the cardinality $N_\cap$ of (the strict part of) the intersection and the number N_c of counterexamples in the intersection showing that it is no interval order. Figure 13 shows for different values of ε the expected values (solid), the 0.5% and the 99.5% quantile (dashed) and the 0.05% and the 99.95% quantile (dotted) of the distribution of the cardinality $N_\cap$ (black) and the number N_c of counterexamples (grey). Since because of symmetry the distributions of $N_\cap$ and N_c do not depend on the true ordering μ, we only need to do the analysis in dependence on the error probability ε. In the presidents data set we actually observed a cardinality of $N_\cap = 337$ and $N_c = 38$ counterexamples. To make Figure 13 easy to read out, we display $\frac{N_\cap}{337}$ and $\frac{N_c}{38}$ in the same figure. Thus, ε- values with corresponding values of $\frac{N_\cap}{337}$ and $\frac{N_c}{38}$ near to 1 would correspond to a good fit of the model to the actually observed data. We analyzed the distributions by simulating 50000 random profiles consisting of 26 orderings respectively. For the simulation as well as for the estimation of the parameter ε for the presidents data set we used the R-package `Rankcluster` (Jacques et al. [2014]). One can see that the expectation of $N_\cap$ is decreasing in ε and that for ε larger than 0.005 the expectation of N_c is also decreasing in ε which is not surprising, because for very high values of ε it is very probable that the intersection of all orderings in the profile is very sparse and thus there are not many quadruples that can provide counterexamples. Furthermore, there is actually no parameter value for ε such that the $[0.5\%; 99.5\%]$-quantile-range of both statistics $\frac{N_\cap}{337}$ and $\frac{N_c}{38}$ simultaneously covers the value 1. In this sense, the actually observed data do not fit very well to the ISR-model, statistically speaking, a statistical test based on both statistics would significantly reject the ISR model on a 2% significance level (, but note that we analyzed the model after a comprehensive descriptive analysis and did not state a concrete procedure for testing the model in advance, so a frequentist compliance of a nominal level $\alpha = 2\%$ could clearly not be guaranteed). If we look more conservatively at the two statistics, then for values of ε around 0.001 the model fits the data good enough in the sense that the $[0.05\%; 99.95\%]$-quantile-range of the both statistics covers the value one, but note that the actually estimated[6] $\hat{\varepsilon} = 0.064$ is far away from values around 0.001

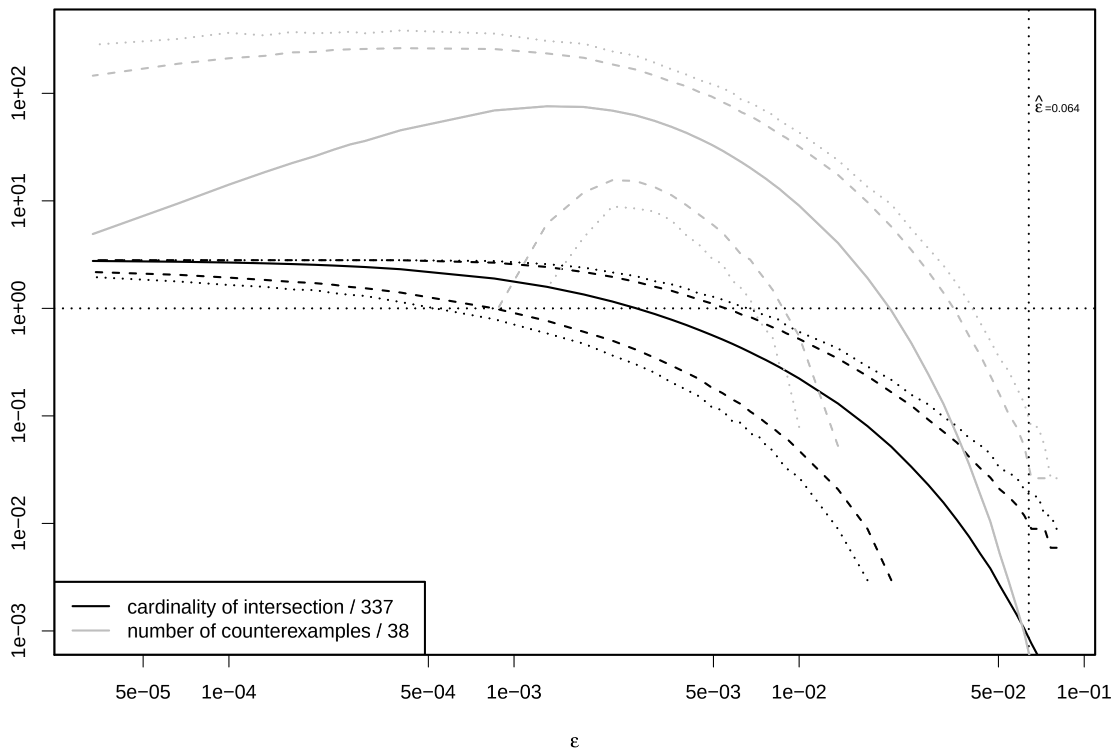

Figure 13: The cardinality of the intersection divided by 337 and number of counterexamples divided by 38 for different parameters ε in the ISR model.

All in all, it seems that the ISR model is more or less unrealistic, it seems to be not able to model situations where a high degree of homogeneity reflected by a high value $N_\cap$ appears together with a profile intersection that is very close to an interval order in the sense that N_c is relatively small.

[6]To get a rough impression about the statistical uncertainty associated with the estimate ε we bootstrapped 10 times and got a bootstrap-standard deviation of around 0.01 (The computation time for the estimation of ε was actually very long (some hours per estimate)).

For the presidents data set, it appeared natural to look at the interval order aspect of the profile, for other data examples it may be not so natural to look at this aspect and the question arises, how one could construct some general type of a model test. One generally applicable (but possible not always meaningful) and a little bit more "abstract" form of a model test could be established by utilizing the lattice valued outlyingness-concept: The level function λ can also be applied to probability distributions on a complete lattice instead of a sample of data points in a complete lattice (cf., Schollmeyer [2017]). To test a model for ranking data one can simply look at the distribution P of the intersection $\bigcap R$ of a profile R randomly generated by the ranking model.[7] If one takes as a test statistic T the level $\lambda_P(\bigcap R)$ of the intersection $\bigcap R$ of the profile R w.r.t. the distribution P then T is discretely uniformly distributed in the sense that $\forall c \in \mathrm{supp}(T) : P(T \leq c) = c$ (cf, Schollmeyer [2017]). A statistical model test with confidence level $\gamma \in \mathrm{supp}(T)$ can thus be simply established by rejecting the model if the statistic for the actually observed data is greater than γ.

Figure 14 shows for different values of ε the level $\lambda_\varepsilon(x)$ of the intersection x of the actually observed profile for the presidents data sets with respect to the distribution of the intersections under the Insertion Sorting Rank model. The computation of this distribution was made by a simulation of 40000 randomly drawn profiles. Furthermore Additionally, horizontal lines for critical values $c = 0.95, c = 0.99$ and $c = 0.995$ are added. For values of ε around the estimated ε of 0.064 one can see that the intersection of the actually observed profile has a level of around $\lambda_\varepsilon(x) \approx 0.999$ which would indicate a rejection of a model test on a significance level of around 0.1% for such ε-values.

[7] If the intersection is too often empty one can alternatively look at the intersection of the $\alpha \cdot 100\%$ most central orderings in the profile, where $\alpha \in [0, 1]$ is appropriately chosen.

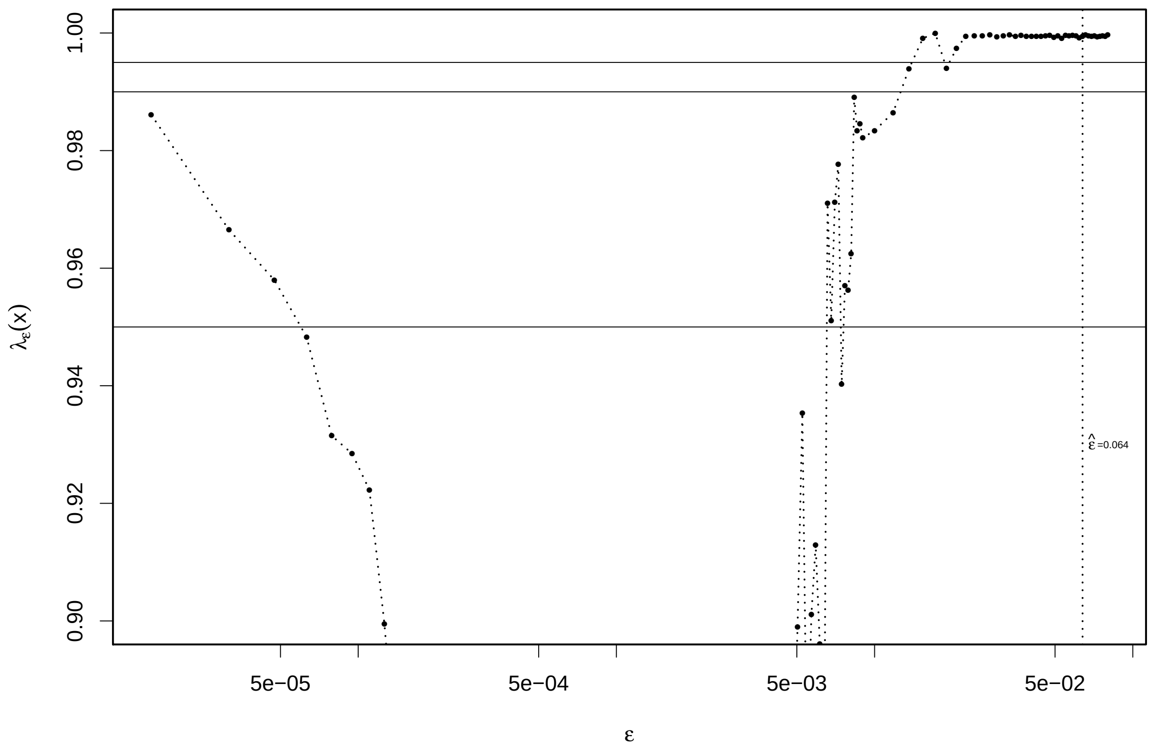

Figure 14: The level $\lambda_\varepsilon(x)$ of the intersection of the presidents profile with respect to the distribution of the intersections of randomly drawn profiles from the Insertion Sorting Rank model for different parameters ε.

4.2.2 Thurstonian model

We finally also very briefly analyze the fit of the Thurstonian model that was used in Lee et al. [2014]. A Thurstonian model is a latent variable model for ranking data. One has a set of q location parameters $\mu_i, i = 1, \ldots, q$, in our case the location of the 44 presidents within a latent ground truth. Every participant $j, j = 1, \ldots, 26$ of the experiment has some access to the ground truth that is modeled by a random draw of a random variable $z_{ij} \sim \mathcal{N}(\mu_i, \sigma_j^2)$. The ranks that person j supplies are then simply the ranks of the values $z_{1j}, \ldots, z_{qj}$. The variance σ_j^2 somehow models the "expertise" of person j. In our case, the model thus has $44 + 26 = 70$ parameters to estimate. Since Maximum likelihood estimation is computationally very hard, one can use Bayesian methods like that used in Lee et al. [2014], for a more detailed description of the Thurstonian model and the used estimation technique, see Lee et al. [2014].

We tried to reproduce the results of Lee et al. [2014] by using the R package `rjags`

(Plummer [2016]) and the JAGS model whose source code is given in the supplementary material of Lee et al. [2014]. (For a general introduction to the implementation of Bayesian Thurstonian models for ranking data in JAGS, see [Johnson and Kuhn, 2013].) The results we got were very similar to the results in Lee et al. [2014]. The violin plot depicted in Figure 15 looks roughly identical to that given in [Lee et al., 2014, p. 3]. The violin plot shows the posterior distribution of the latent location parameters $\mu_i, i = 1, \ldots, 44$ for the 44 presidents.

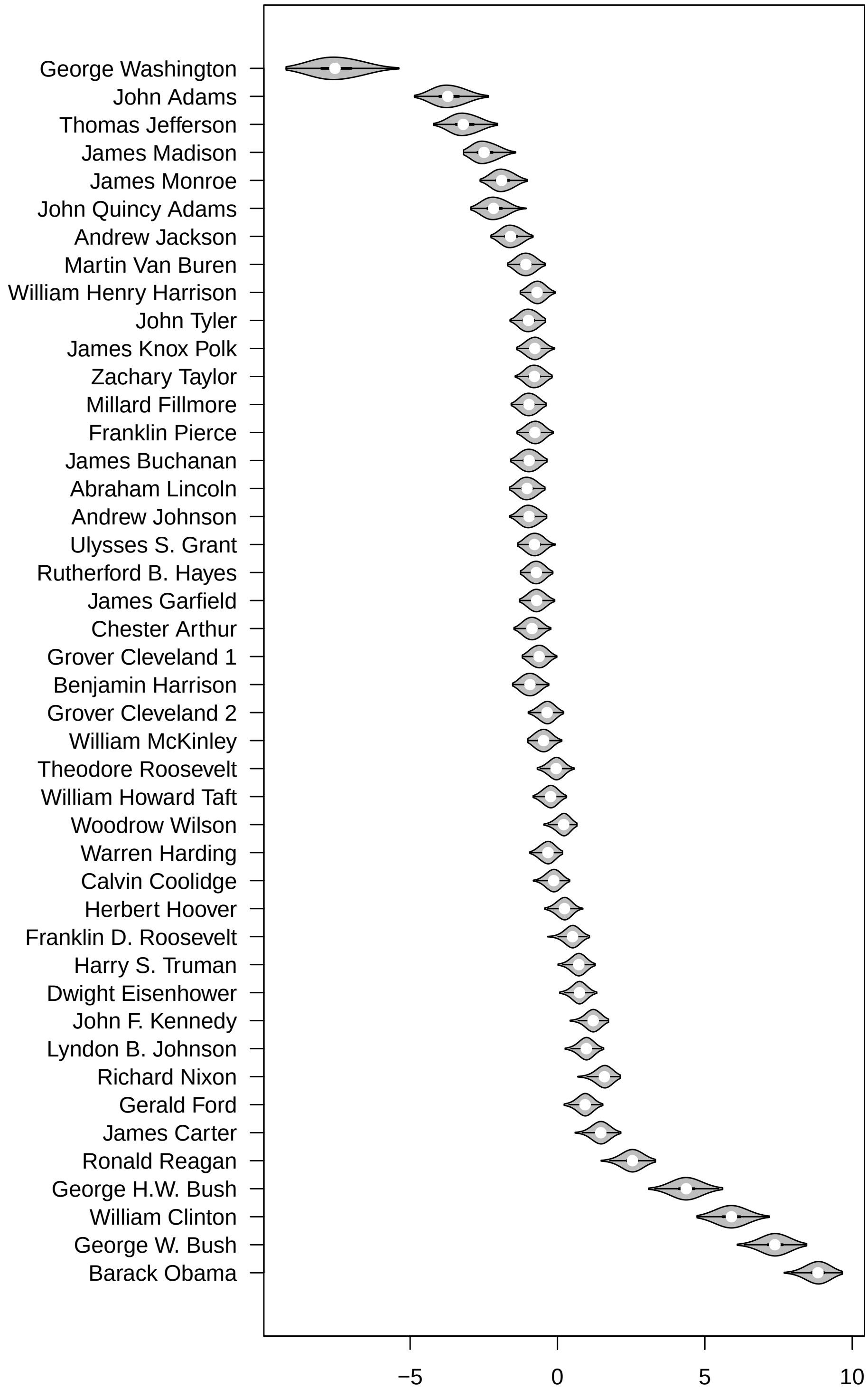

Figure 15: Violinplot for the reproduced Thurstonian model analyzed in Lee et al. [2014]. The violin plot shows the posterior distribution of the latent location parameters $\mu_i, i = 1, \ldots, 44$ for the 44 presidents.

Since the Thurstonian model has 44 parameters of the location of the presidents in the latent continuum and 26 parameters modeling the students "expertise", it is very hard to analyze the models ability to capture the specific properties of the president data set. In particular, since one does not know the true parameters, one has to account for the statistical uncertainty of estimates of the unknown true parameters to get a fair picture of how good the model is in principal able to capture the specific characteristics of the president data set: The estimated parameters differ from the true parameters and one has to account for the fact that the distribution of the profiles under the estimated and under the true parameters can be very different in such a high dimensional model. For the altogether 70 parameters of the Thurstonian model one could get at least credibility regions, but it seems practically not manageable to analyze the distribution of the statistics $N_\cap$ and N_c for all parameter constellations that are lying in a reasonable credibility region, at least if one does the analysis simulation-based. (A theoretical insight into the distribution of $N_\cap$ and N_c would be very helpful, but the author is not aware of theoretical treatments in the context of this very special situation.)

So, we have to live here with a very sketchy analysis, where we only analyze the model under the estimated model parameters, which is of course still of some interest, because for example if the task is that of prediction of a future profile, then one would usually base the prediction on the estimated model parameters. Of course, if the task is only that of prediction, then it is not so important that the model is (approximately) true, it only has to be good in prediction. However, it could be very insightful to know, in which concrete sense the model is "not true" to get a hint how one could possibly modify the model to get better predictions.

A simulation of 100000 profiles under the Thurstonian model and the estimated model parameters[8] led to observed values for $N_\cap$ with a range between 279 and 436 and values for N_c with a range between 1912 and 4410. This indicates that the Thurstonian model may capture the homogeneity of the profiles, quantified by $N_\cap$ acceptably well, but the interval order type character discovered in the president data set seems to be not captured adequately with the Thurstonian model. Also for the abstract model test based on the level function, we got a level of $\lambda_\varepsilon(x) \approx 0.99956$ which also indicates a bad model fit.

The bad ability of the Thurstonian model to capture interval order type characteristics indicates, that possibly, one may get a far more appropriate model if one explicitly models the interval structure, for example by using for every president not one precise location at some latent continuum, but instead a latent interval.

[8]As estimates we used the modes of the posterior distribution of the model parameters.

5 Conclusion

The present paper illustrated the usefulness of the lattice-valued quantile approach of Schollmeyer [2017] for analyzing ranking data. We were able to use the quantile approach for getting purely descriptive insights into ranking data sets, as well as for doing a statistical analysis of model fit of statistical models for ranking data, based on seemingly descriptively meaningful characteristics, like the proximity of profile intersections to interval orderings. Furthermore, we indicated a possibility of using the order theoretic outlyingness concept for ranking data to construct an aggregation function for social profiles that is very similar to Kemeny's rule in the sense that is also some kind of a generalization of the median to ranking data: While Kemeny's rule generalizes the median understood as a minimizer of a $\mathcal{L}^1$-norm, the herein developed commonality sharing rule generalizes the median in a completely order theoretic manner as the least outlying data point(s). The usefulness of this aggregation function has of course to be evaluated in further studies.

References

Bartholdi, J., Tovey, C. A., and Trick, M. A. (1989). Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165.

Biernacki, C. and Jacques, J. (2013). A generative model for rank data based on insertion sort algorithm. *Computational Statistics & Data Analysis*, 58:162–176.

Bogart, K. P., Rabinovich, I., and Trotter, W. T. (1976). A bound on the dimension of interval orders. *Journal of Combinatorial Theory, Series A*, 21(3):319–328.

Davey, B. A. and Priestley, H. A. (2002). *Introduction to Lattices and Order.* Cambridge University Press.

de Borda, J. C. (1781). *Memoire sur les elections au scrutin.* Historie de l'Academie Royale des Sciences.

de Condorcet, N. (1785). *Essai sur l'application de l'analyse a la probabilite des decisions rendues a la pluralite des voix.* Paris.

Fligner, M. A. and Verducci, J. S. (1986). Distance based ranking models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3):359–369.

Galton, F. (1907). Vox populi (the wisdom of crowds). *Nature*, 75(7):450–451.

Ganter, B. and Wille, R. (2012). *Formal Concept Analysis: Mathematical Foundations.* Springer.

Jacques, J. and Biernacki, C. (2014). Model-based clustering for multivariate partial ranking data. *Journal of Statistical Planning and Inference*, 149:201–217.

Jacques, J., Grimonprez, Q., and Biernacki, C. (2014). Rankcluster: An R package for clustering multivariate partial rankings. *The R Journal*, 6(1):10.

Johnson, T. R. and Kuhn, K. M. (2013). Bayesian Thurstonian models for ranking data using JAGS. *Behavior Research Methods*, 45(3):857–872.

Kemeny, J. G. (1959). Mathematics without numbers. *Quantity and Quality*, 27:577–591.

Kidwell, P., Lebanon, G., and Cleveland, W. (2008). Visualizing incomplete and partially ranked data. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1356–1363.

Lee, M. D., Steyvers, M., and Miller, B. (2014). A cognitive model for aggregating people's rankings. *PloS one*, 9(5):e96431.

Marden, J. I. (1996). *Analyzing and Modeling Rank Data*. CRC Press.

Plummer, M. (2016). *RJAGS: Bayesian Graphical Models using MCMC*. R package version 4-6.

Schollmeyer, G. (2017). Lower quantiles for complete lattices. Technical Report 207, Department of Statistics, LMU Munich. URL `http://www.statistik.uni-muenchen.de/forschung/technical_reports/index.html`.

Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor.

Thompson, G. L. (1993). Generalized permutation polytopes and exploratory graphical methods for ranked data. *The Annals of Statistics*, 21(3):1401–1430.