

Structure and evolution of a European Parliament via a network and correlation analysis

Elena Puccio^{a,1}, Antti Pajala^b, Jyrki Piilo^c, Michele Tumminello^d

^a*Dipartimento di Fisica e Chimica, Università di Palermo, Viale delle Scienze, 90128 Palermo, Italy*

^b*Department of Political Science, University of Turku, FI-20014 Turun yliopisto, Finland*

^c*Turku Centre for Quantum Physics, Department of Physics and Astronomy, University of Turku, FI-20014 Turun yliopisto, Finland*

^d*Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università di Palermo, Viale delle Scienze, 90128 Palermo, Italy*

Abstract

We present a study of the network of relationships among elected members of the Finnish parliament, based on a quantitative analysis of initiative co-signatures, and its evolution over 16 years. To understand the structure of the parliament, we constructed a statistically validated network of members, based on the similarity between the patterns of initiatives they signed. We looked for communities within the network and characterized them in terms of members' attributes, such as electoral district and party. To gain insight on the nested structure of communities, we constructed a hierarchical tree of members from the correlation matrix. Afterwards, we studied parliament dynamics yearly, with a focus on correlations within and between parties, by also distinguishing between government and opposition. Finally, we investigated the role played by specific individuals, at a local level. In particular, whether they act as proponents who gather consensus, or as signers. Our results provide a quantitative background to current theories in political science. From a methodological point of view, our network approach has proven able to highlight both local and global features of a complex social system.

Keywords: Complex Systems, Networks, Bipartite System, Community Detection, Social Systems, Correlation Analysis.

PACS: 89.75.Hc, 89.75.Fb, 89.65.Ef

1. Introduction

Since the seminal papers by Watts and Strogatz [1], Barabasi and Albert [2], and Newman, Watts and Strogatz [3], the use of networks to describe the structure and evolution of complex social systems has become a standard approach in the scientific literature. Emergent structures have been studied in social networks [4] as well as their evolution [5, 6].

Email address: elena.puccio@unipa.it (Elena Puccio)

Political systems such as the European parliaments and the US congress represent a class of social systems. The structure of similarity of the US congress has been investigated by using network theory methodologies in [7], where the authors built a hierarchical tree of congress members based on law initiatives they cosponsored and attempted to characterize the two biggest communities found in terms of parties (Republicans and Democrats), by using the concept of modularity.

Although being placed in this framework, our work focuses on a typical European parliament, it takes into account more than one kind of initiatives (legislative and budget ones), and quantitatively addresses the problem of noise in a dense network, as well as the statistical significance of attributes, such as party or district, which characterize communities [8] in the network.

The political system under scrutiny is the Finnish parliament, which is a typical parliamentary system among its peers in Europe. In addition to the government, the parliament also has the right to initiate legislation. The thousands of private initiatives by members of the parliament have a primary signer, however, they are rather often supported by a multitude of co-signers, thus forming a network of co-sponsorships. We are interested in studying the network of these relationships among parliament members and the evolution of its structure over the years.

Our goal is to gain a quantitative understanding of the system by observing it with different techniques and by following the dynamics of its structure over time. In order to do so, we take a network approach to discover the internal ordering of the system in communities and to characterize them, and perform a correlation analysis of members, in order to study the hierarchical structure of communities [9, 10, 11] and investigate its evolution. Specifically, we study how correlations inside each party and between parties, as well as within and between government and opposition, vary over time, by using the Frobenius distance to measure how similar each year of a parliament term is to the next one.

We look at the database of initiatives as a bipartite network in which there are two sets of nodes—parliament members on the one side, and initiatives on the other—where a parliament member is linked to an initiative if she/he signed it. According to that description, two parliament members would show a “similar” profile if they had signed “several” initiatives together. To provide a quantitative meaning to the expressions “similar” and “several”, recently, a method for filtering out statistically significant links in bipartite networks has been proposed [12]. This approach has already been used to investigate the structure of several systems, including stock returns [13] and investors’ activity [14] in a financial market, the specializations of criminal activity [15], the inter-bank market [16, 17], a mobile communication network [18], and a large survey on aging [19].

Our results indicate that the methodologies we employ are able to single out both local and global characteristics of a social system such as a parliament, which appear consistent with pre-existent theories in political science. Although our conclusions pertain to the Finnish legislature, the techniques we used are exportable to any similar systems, paying due attention to any necessary adjustments when carrying them over to a different political context.

Dataset summary statistics				
	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
N	2,467	3,163	3,143	1,808
h_i	2-144	2-175	2-136	2-150
M	179	186	183	199
h_m	2-445	4-524	2-696	2-793

Table 1: Summary statistics for each parliament: N is the number of initiatives signed by at least 2 members, h_i is the heterogeneity of initiatives, that is, the range (min-max) of signatures initiatives receive, M is the number of members who signed at least 2 initiatives, h_m is the heterogeneity of members, that is, the range of signatures members affix (min-max).

2. Results and Discussion

2.1. Data

The database consists of Finnish initiatives, submitted between 1999 and 2014 by members of the parliament. The parliament has 200 members. Complementary information stored in the database comprises who submitted each initiative, who signed it and the year it was submitted, along with the following general attributes regarding members: their gender, party and electoral district of origin.

Since a parliament term lasts four years, our data encompasses four different parliaments: Dataset I corresponds to the 1999-2002 parliament, Dataset II to 2003-2006, Dataset III to 2007-2010 and Dataset IV to 2011-2014. Summary statistics for each dataset are shown in TABLE 1. See the Appendix for TABLE A.1 of political parties with their abbreviations, number of seats and political position, TABLE A.2 for the electoral districts and TABLE A.3 for government/opposition coalitions.

According to [20], parliament members have the right to introduce a legislative initiative containing a proposal for the enactment of an Act. As the final decision in the State’s annual budget lies in the hands of the parliament, parliament members can also propose a budget amendment containing a proposal for an appropriation to be included in the budget or for other budgetary decision. In addition, parliament members can propose a petitionary motion containing a proposal for the government for drafting a law or for taking other governmental measures. Finally, a parliament member may ask the Speaker’s Council that a topical debate be held in a plenary session. All the initiatives have to be in writing and must be signed together with possible co-signatures before the introduction. Among its peers, Finland belongs to the most liberal parliaments in what comes to parliament members’ private initiative introduction and there are very few restrictions regarding this activity [21].

The Finnish election system is proportional with open candidate lists. The country is divided in 15 electoral districts. The southern and more densely populated districts are geographically smaller than the northern ones. For example, the geographically small (capitol) Helsinki district has many more parliament members as compared to the very large northern Lapland district. Finland is a parliamentary system with eight political parties represented in the parliament. During 1999-2010 two out of the three largest parties formed a majority coalition together with certain small party groups while the third large party became the main opposition party. General elections in 2011 changed the number of large parties to four. The three largest parties formed a majority coalition

leaving the fourth party in the opposition together with the smaller parties. The cabinet (majority) coalition governs while the opposition has very limited power.

2.2. Network Description

Here, we construct statistically validated networks (SVNs) of parliament members which point towards the presence of preferential relationships within each parliament. After having revealed the informative structure of the parliament system, we detect members' communities and find out whether these can be characterized by attributes such as party and electoral district. The aim is to identify key features that drive the formation of communities made of parliament members.

Network construction. Our system is a bipartite network, consisting of parliament members on the one side and initiatives they either proposed or signed on the other. Henceforth we treat first signers, that is, those who initially propose the initiative, simply as signers. When dealing with bipartite networks, one can obtain a network made of nodes of the same type by projecting on the corresponding set of nodes. In our case, by projecting on the parliament members set, we obtain a network made of connections (links) between members (nodes). The projection establishes a link between each pair of members who sign the same initiative.

Unfortunately, it's well known that such projection often produces a highly connected network, due to the heterogeneity inherent in both sets. This effectively hides the most informative structure of the system, therefore we apply the validation method developed in [12]. Ultimately, this resolves the validated and informative links from the random ones. Our system is highly connected, indeed the density of links, defined as the fraction of links present in the network, m , over the number of maximum possible links between the M nodes, $D = 2m/M(M - 1)$, ranges between 0.93 and 0.99.

Before projecting on the members' set, the heterogeneity of the initiatives' set needs to be given due consideration. Indeed, initiatives display a number of signatures ranging from just a few (low degree), up to three quarters of the parliament (high degree). To account for this effect, we follow Ref. [12] and divide initiatives in bins B_k according to their degree, so that each subset of the bipartite network now comprises initiatives with a specific degree range $d_{min}^i - d_{max}^i$ (which highly reduces heterogeneity on the initiatives' side), and only those members who actually signed them. Considering a range of degrees proves necessary for higher degrees, to keep statistical resolution high in the resulting bins. We set the binning intervals equal for all four datasets, assuming that the final results do not strongly depend on the choice made. The set of bins is defined $B_1 = \{2\}$, $B_2 = \{3\}$, $B_3 = \{4\}$, $B_4 = \{5\}$, $B_5 = \{6\}$, $B_6 = \{7\}$, $B_7 = \{8\}$, $B_8 = \{9\}$, $B_9 = \{10\}$, $B_{10} = \{11-13\}$, $B_{11} = \{14-20\}$, $B_{12} = \{21-40\}$, $B_{13} = \{41-100\}$, $B_{14} = \{> 100\}$.

If we now consider a pair of members (i, j) in bin B_k and each signed respectively n_i^k and n_j^k initiatives, out of the total number N^k of initiatives within the bin, we expect to find a number X of initiatives they co-signed, purely at random, that follows the hypergeometric distribution

$$H(X|N_k, n_i^k, n_j^k) = \frac{\binom{n_i^k}{X} \binom{N^k - n_i^k}{n_j^k - X}}{\binom{N^k}{n_j^k}}.$$

Bonferroni Network statistics

	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
M_B	172	177	161	153
m_B	1633	1839	1162	1811
f	10.4%	10.9%	7.2%	9.9%
$\mu(w)$	4.1±2.0	3.6±1.9	3.3±1.9	2.9±1.5
$\mu(w_B)$	1.4±0.8	1.4±0.7	1.6±1.0	1.3±0.7
w -range	1-8	1-6	1-9	1-8

Table 2: Bonferroni network statistics for parliament terms *I – IV*. The number of members involved in the validated network M_B , the number of validated links m_B , the percentage of validated links f out of the total number of original links, the mean weight $\mu(w)$ of the original network links and its s.d., the mean weight of validated links $\mu(w_B)$ along with its s.d., and links’ weight range in the Bonferroni Network.

It’s straightforward then to assign a p-value to the link between i and j within the bin, under the hypergeometric distribution null hypothesis

$$p_{ij}^k = 1 - \sum_{X=0}^{n_{ij}^k-1} \frac{\binom{n_i^k}{X} \binom{N^k-n_i^k}{n_j^k-X}}{\binom{N^k}{n_j^k}},$$

where n_{ij}^k is the number of initiatives i and j co-signed, in B_k . This represents the probability of randomly obtaining a value equal to or greater than what was actually observed, n_{ij}^k . The univariate significance level of the test, or threshold value, is usually set at either 5% or 1%. In our case, we perform multiple tests, by validating the full set of links over all bins simultaneously. For this very reason, we need to control for false positives, or familywise error rate, by introducing a multiple test corrected threshold. The most conservative choice, in terms of type I errors, is applying Bonferroni correction [22, 23] $\alpha_B = \alpha/m$, where m is the total number of tests performed. From here on, we choose a univariate threshold of 1% significance and validate each link (i, j) if $p_{ij}^k < \alpha_B$, that is, if the p-value associated with the link is smaller than $\frac{0.01}{m}$, according to the Bonferroni correction for multiple comparisons.

The last step is obtaining the weighted network. In fact, a link between the same pair of members could be validated over different bins. We account for this effect by assigning a weight to each link equal to the number of bins the link was validated in. Since we aim at building a validated network (Bonferroni network hereafter), members who signed less than 2 initiatives, as well as initiatives with less than 2 signatures, have no relevance to our analysis and have thus been omitted. The Bonferroni network’s general statistics are shown in TABLE 2, along with some properties of the full network, that show the advantage accorded by this filtering method.

Basically, we’re interested in studying how members cluster together on the basis of the initiatives they co-sign, thus the null hypothesis and validation allow us to distinguish information from noise.

Community detection. After building the network, our main interest lies in finding out whether it’s internally organized in communities, and ultimately, which attributes, the additional information on members, characterize each community.

Modularity and Mutual Information

	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
Q	0.51	0.56	0.61	0.53
$\mu(MI)$	0.87	0.91	0.92	0.90

Table 3: Modularity for the best partition Q and overall mean of the mutual information between the best partition and each of the other 5 partitions, $\mu(MI)$.

However, partitioning a network is neither straightforward nor simple, there are a variety of algorithms that attempt to do so, each with its benefits and shortcomings [24]. Nonetheless, the majority of techniques relies on the concept of modularity [25, 26]:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j),$$

where m is the total number of links in the network, A_{ij} is the adjacency matrix, (k_i, k_j) are node i 's and j 's degrees, the sum is carried out over all the nodes and the delta function returns 0 if i 's and j 's communities are different $C_i \neq C_j$ and 1 otherwise. The goal of every algorithm is to maximize the modularity over all partitions found, iteratively.

In our case, we chose a community detection software Radatools ¹, which employs a combination of different algorithms [27, 28, 29, 30], allows for weighted networks and multiple repetitions of each algorithm, thus producing high modularities. We tried different combinations of algorithms, run one after the other in cascade, as detailed in the software's manual ²

For each cascade, we ran several repetitions and usually had the best results by using 200 repetitions of (vi) e b r f b r. Furthermore, in order to check the stability of our best partition, we evaluated the similarity between the former and each of the other partitions found with the various cascades, by using the mutual information (MI) [31, 32]. As TABLE 3 indicates, we found a high mean similarity, which shows a good stability of the best partition, regardless of the algorithm cascade chosen.

Characterization of communities. After finding a stable partition of the system, we look for each community's characterizing attributes, by validating any additional information on members contained in the database. This is accomplished in a similar fashion as what done in Section III A when validating links, following the method in [8]. The method makes use of the hypergeometric distribution to assess the probability that a given attribute is over-expressed in the elements of a community in respect to all the elements of the investigated set. Again, each attribute is validated, community-wise, if the associated p-value falls below the Bonferroni threshold at 1% of significativity.

The attributes we consider are both specific, such as a member's party, district and gender, and more generic, as whether the district was in a rural or metropolitan area

¹Radatools 3.2, Copyright (c) 2011 by S. Gomez, A. Fernandez, J. Borge-Holthoefer and A. Arenas, all rights reserved.

²(i) t (tabu search) r (fine-tuning by reposition) f (fine-tuning by fast algorithm) r; (ii) e (extremal optimization) r f r; (iii) t r f r e r; (iv) e r f r t r; (v) t b (fine-tuning by bootstrapping based on tabu search) r f b r; (vi) e b r f b r.

(Helsinki and Uusimaa), whether the party was in the government or in the opposition and its political position (right, centre or left).

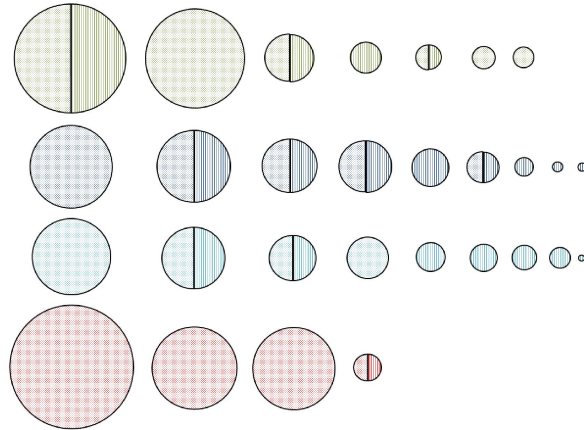


Figure 1: Qualitative illustration of communities. Top row with 7 communities is the I parliament term (green), II term has 9 communities (blue), III has 9 (cyan) and IV has 4 (red). A dotted pattern indicates characterization by party, a striped pattern characterization by district, the presence of both patterns indicates characterization by both. The radius of each circle is proportional to the number of members within the community. For example, parliament II in the second row has a large number of modest size communities where all except one has district characterization. In contrast, for parliament IV (fourth line), the number of communities is limited but their size is large, with only one community characterized by district whilst party drives community formation.

Overall, we found a satisfactory characterization of communities by party, district or both, as shown in Fig. 1, and some characterization by area, coalition and political position (details in the Appendix, TABLES A.5, A.6, A.7 and A.8). Surprisingly, gender doesn't appear characterizing, indicating there is no influence of a member's gender over whom he collaborates with.

Still, the most interesting result is that, although party appears more explanatory for the first dataset, showing up in all communities, both party and district concur to characterize communities in the second and third datasets, with district growing in importance. Finally, party reverts to being the most characterizing attribute for the last dataset.

Indeed, it appears that collaboration moved from being among members of the same party to being among members from the same district, as time passed by, with the last parliament going against the trend. The Bonferroni Network for the third parliament is shown in Fig. 2, this is the best example of districts playing a major role in how members cooperate. Broadly speaking, the power of our method lies in that it allows us to quantitatively characterize communities in the network and is further able to reveal changes occurring in the structure of the collaboration web among members, from term to term.

2.3. Correlation analysis

In this section, we investigate the structure of the correlation matrix with the aim of building a hierarchical tree of members. The advantage that the correlation matrix

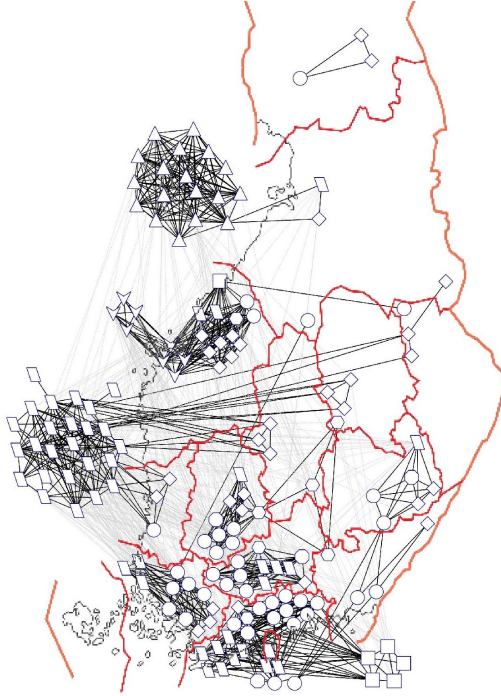


Figure 2: Bonferroni Network for the III parliament. The map in red lines shows electoral districts in Finland, node shapes indicate each member’s party: hexagon=KD, diamond=KESK, ellipse=KOK, rectangle=PS, V=RKP, parallelogram=SDP, triangle=VAS, octagon=VIHR. Nodes are located either on the map, according to their district of origin or outside the map, when party characterization dominates (the network reflects what shown in Fig. 1, 3rd row). To emphasize the network partition, links within each community are displayed in black, links between communities are in light gray. All communities are well characterized by district, with the exception of parties VAS and SDP, on the left-hand side.

offers over the network approach, is the hierarchical ordering of members. This allows to look for substructure within each cluster. The drawback is that we no longer use a quantitative tool to discriminate the random part of correlations present in the system.

The first step to obtain a correlation matrix between members, is to define a correlation coefficient, that hinges on the similarity between patterns of signatures. Specifically, we’d like to evaluate the similarity between a pair of members (i, j) by using a correlation coefficient $s_{i,j} \in [-1, 1]$ [33] that takes into account the number of joint signatures (n_{ij}), the number of individual signatures (n_i, n_j) and the total number of initiatives (M). Basically, when considering binary vectors, that is, vectors whose entries can be either 1 or 0, this turns out to be just a straightforward derivation of the Pearson correlation coefficient

$$s_{i,j} = \frac{n_{ij} - n_i n_j / M}{\sqrt{n_i(1 - n_i/M)n_j(1 - n_j/M)}}. \quad (1)$$

Hierarchical trees. Hierarchical trees are built by using the average linkage method, that is, by successively merging pairs of clusters A and B according to a mean distance between

elements $x \in A$ and $y \in B$ given by

$$\mu(d_{x,y}) = \frac{1}{|A| \cdot |B|} \sum_{x \in A} \sum_{y \in B} d_{x,y},$$

where $|A|$ and $|B|$ indicate the number of elements belonging to cluster A and B , respectively. The quantity $d_{x,y}$ is a measure of distance between members x and y [10], which is a function of the correlation coefficient given in Eq. 1:

$$d_{x,y} = \sqrt{2(1 - s_{x,y})}.$$

By looking at all trees, we again find that both parties and districts play a role in how members cluster together, within each parliament term. Such a strong characterization by party, district and coalition gives an insight on what actually determines members affiliation to a specific cluster.

The main result lies in the hierarchical structure revealed by this method: the opposition tends to cluster strongly by party, while the government shows a collaboration among parties and a further subclustering by district. We chose to display figures just for the last two terms, where the change in the system's structure is more striking, see Fig. 3 and Fig. 4.

Political position doesn't appear to be crucial, in general left-wing parties are somewhat closer together. On the other hand, coalition plays a major role, with government and opposition parties neatly separated. Specifically, opposition parties show closed ranks, while the government cluster doesn't show much of a party structure. Finally, some districts seem to work together, in particular, the metropolitan ones of Uusimaa and Helsinki. This behaviour is common to the first three parliaments, see Fig. 3 for the hierarchical tree of the third term. An exception to this overall behaviour is represented by the last term, where we found the government coalition clustering mainly by party against the trend of the previous years, as shown in Fig. 4.

Structure of the correlation matrix. Now we focus on the correlation matrices for the last two parliaments (see Fig. 5), where the subclusters are easily seen, along with the change of structure from one term to the next. Members are arranged according to the hierarchical trees in the previous subsection.

Apparently, during the III term there's a higher overall correlation and small district-clusters in the government, along with 4 bigger party-clusters in the opposition. On the other hand, during the last term, a collaborative government is clearly seen as a yellow huge cluster with a left and right subdivision and some party-subclusters. For the opposition, there are now only 2 very compact parties, who slightly collaborate with the government and anti-correlate with each other.

By comparing the general structure of the correlation matrices, we draw the conclusion that party counts more than district during the last parliament and opposition parties make a statement of behaving in opposite ways, which is noticeably different than previous terms. Our approach is thus able to detect the change that occurs in the system's structure when the parliament changes, and to quantitatively assess the different webs of collaboration/anti-collaboration that arise within it.

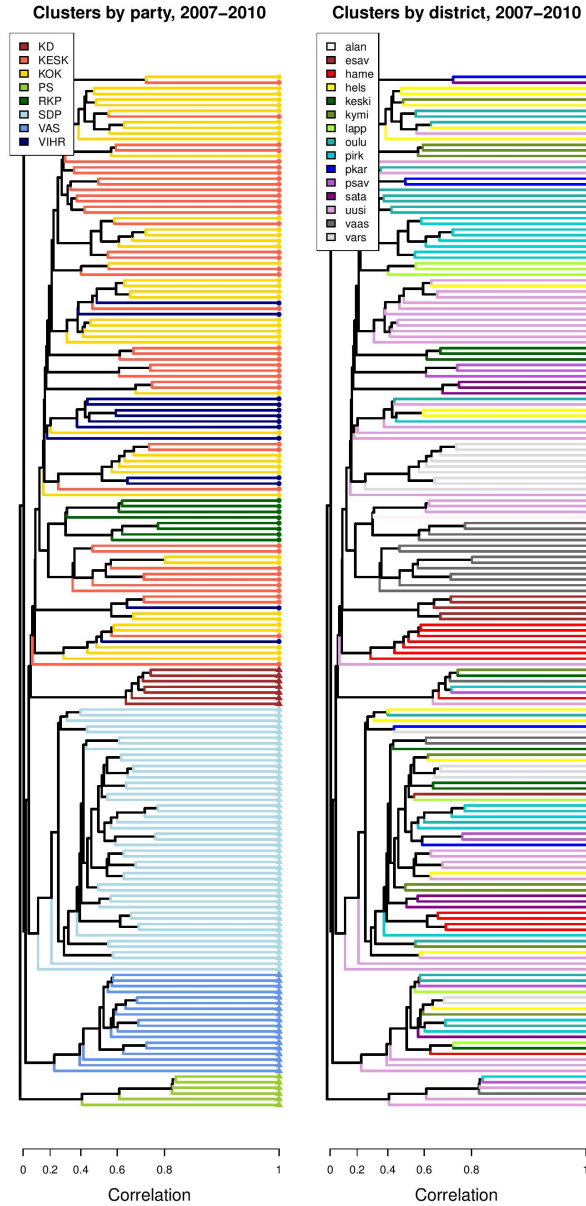


Figure 3: Clustering of members during the III parliament term. On the left, leaves and branches are colored according to parties, circle-shaped leaves indicate nodes in the government (top half of the tree), while triangles indicate nodes in the opposition (bottom half). Coloring on the right-side tree is according to district. The x-axis shows correlation values, going from 0 (no correlation) to 1 (maximum correlation), the spacing is not linear due to the measure of distance chosen. Party dominates in the opposition (bottom-left), while district sub-clustering occurs in the government (top-right).

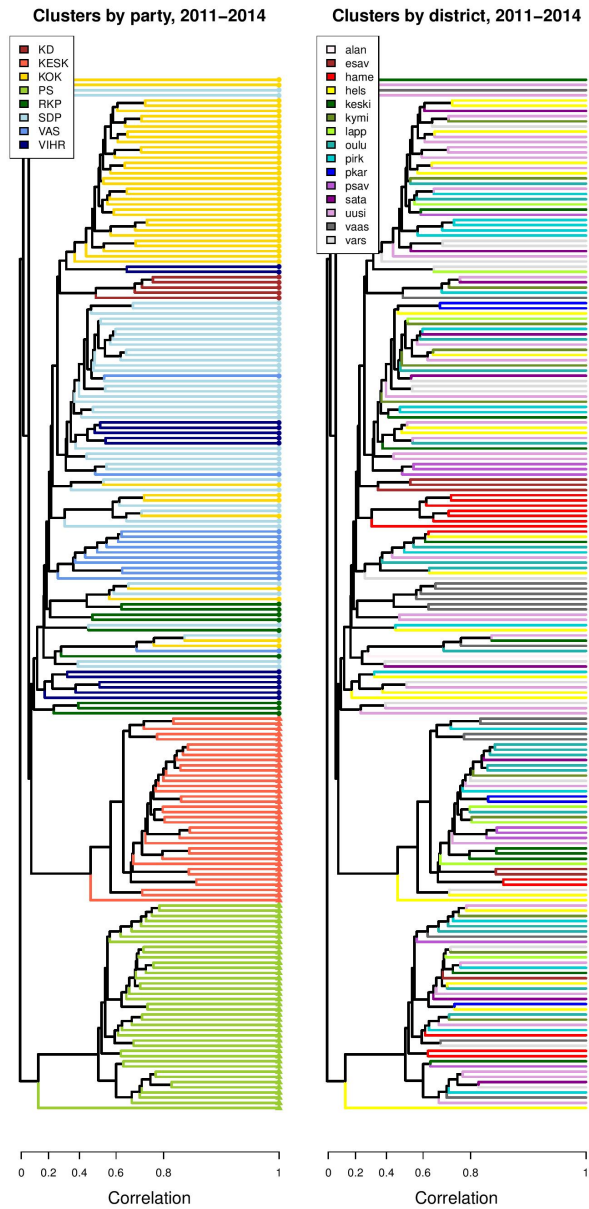


Figure 4: Clustering of members during the IV parliament term. On the left, leaves and branches are colored according to parties, circle-shaped leaves indicate nodes in the government (top half of the tree), while triangles indicate nodes in the opposition (bottom half). Coloring on the right-side tree is according to district. The x-axis shows correlation values, going from 0 (no correlation) to 1 (maximum correlation), the spacing is not linear due to the measure of distance chosen. Here party dominates strongly both in the opposition and in the government, while district ceases to matter.

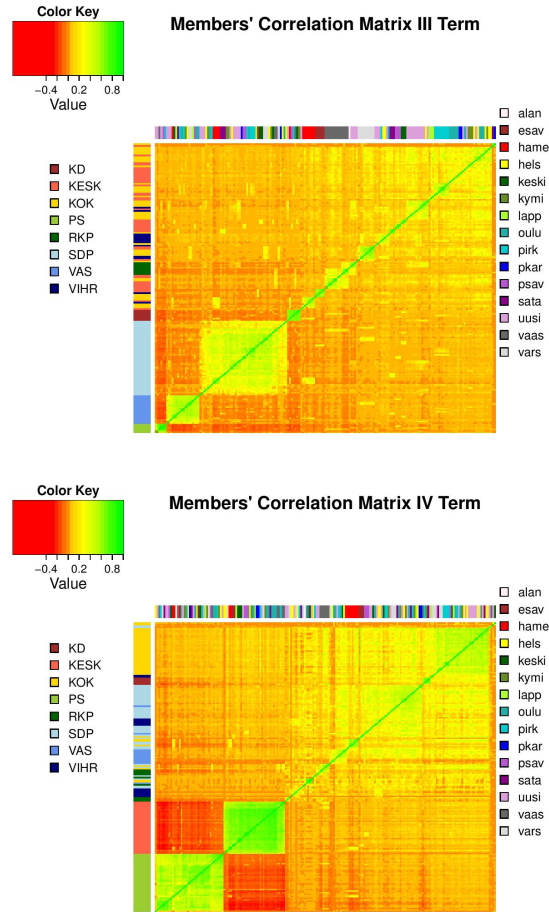


Figure 5: Correlation matrices of members for III (top panel) and IV (bottom panel) parliament terms. The color key displays negative correlations in red, near-zero values in orange, mild correlations in yellow and high correlations in green. Colored bars on the side indicate either members' party (left bar) or district (top bar), as displayed by side legends. In each matrix, members are ordered according to the corresponding hierarchical tree. In the III term, district subclustering can be clearly discerned within the government (top-right square). Party clustering is present in the opposition (bottom-left 4 clusters). In the IV term, we find party subclustering within the government (top-right square), with right-wing parties at the top and left-wing parties at the bottom. The 2 opposition parties are on the lower-left corner and are negatively correlated. Here we notice how district subclustering disappears.

2.4. Some Dynamical Features within and over the parliamentary terms

We focus here on a more detailed look on the dynamical features, in particular by looking at the annual evolution of the correlation matrix. Correlation matrices are henceforth computed year by year, and arranged in blocks representing parties, in order to have a closer look at how interactions evolve within each party (diagonal blocks) and between them (off-diagonal blocks).

Mean correlations. We are interested in how correlations within each party and among parties evolve year by year. The purpose here is obtaining a closer look at the dynamics of interactions between members depending on the party they belong to. We consider the mean correlation calculated over members belonging to a specific party (intra-party), and the mean correlation calculated between members of that party and each of the others (inter-party).

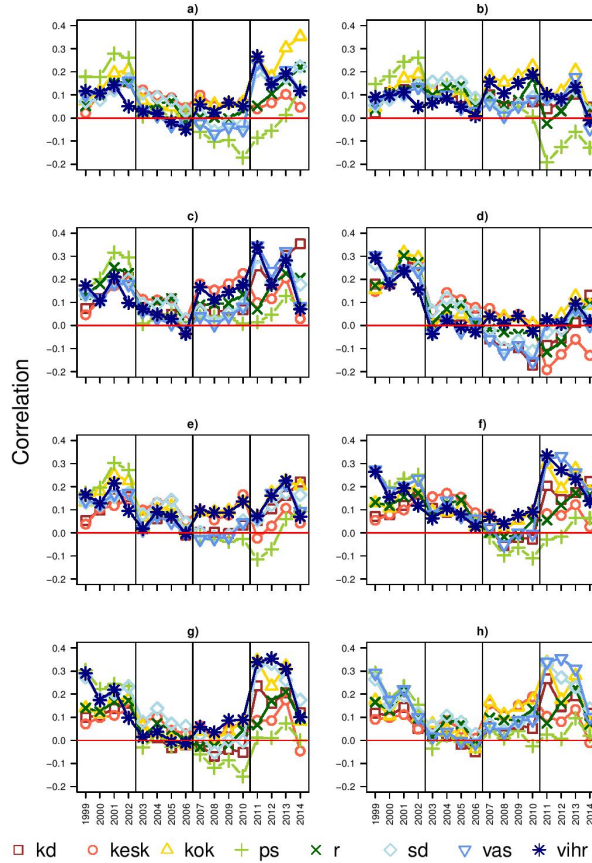


Figure 6: Mean correlation between parties, over time. Panel a) Inter-party mean correlation between KD and all the others; b) KESK-others; c) KOK-others; d) PS-others; e) RKP-others; f) SDP-others; g) VAS-others; h) VIHR-others. Red horizontal lines mark the 0-value. Vertical lines separate parliament terms. The general trend shows higher inter-party mean correlation during I term, decreasing, packed together, going below zero, mean correlation values during II and III term and increasing, broader values during last term (with the sole exception of the 2 opposition parties).

From Fig. 6 and 7, we can see how each party presents its own pattern of collaboration with every other party, which is not constant over the years, not even within a single parliament term. Nonetheless, there are some general trends, for example it's apparent how the first and even more so the last term are unusual in the sense of having spread out of correlations, while second and third terms (in the middle) display a narrower, more similar pattern. Moreover, mean inter-party correlation tend to be positive during

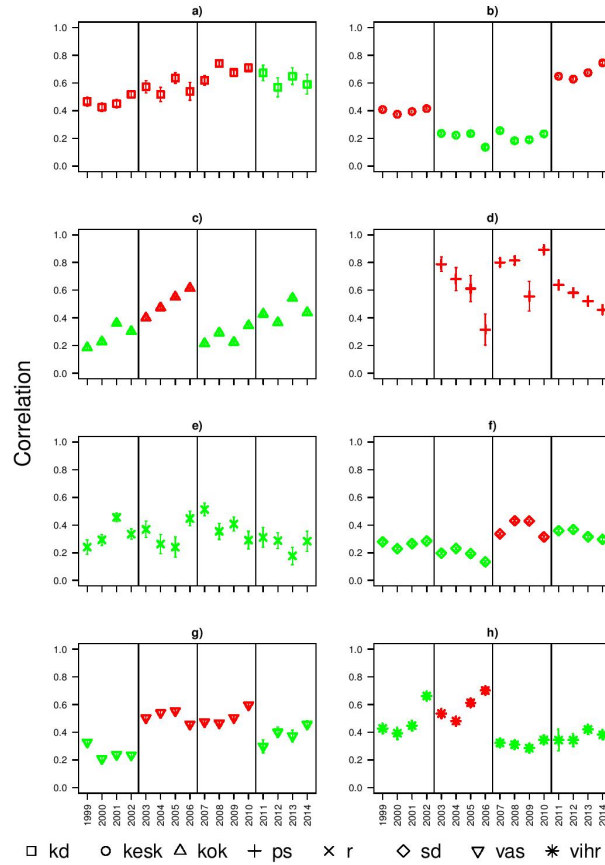


Figure 7: Mean correlation within each party, over time. Panel a) Intra-party mean correlation for party KD; b) KESK; c) KOK; d) PS; e) RKP; f) SDP; g) VAS; h) VIHR. Vertical lines help distinguishing each parliament term. In general, the trend indicates that when parties are at the government (shown in green) they have lower mean correlation than when at the opposition (shown in red). In any case, intra-party mean correlation values are on average higher than inter-party ones, as expected. Error bars are 6 standard deviations of the mean.

the I term, still positive but closer to zero during the II, they turn slightly negative for opposition parties during the III (kd, ps, sdp, vas) and plunge even more for opposition parties during last term (kesk, ps). Error bars were calculated by bootstrapping data 1000 times and choosing 5-95% confidence intervals, and are all of the order of 10^{-2} .

The overall mean correlation between a party and all the others, all parliaments considered, is strictly positive, ranging from 0.06 to 0.12. So, on average, ps is the least collaborative, while kok, sdp and vihr seem the most willing to collaborate, as expected from parties often at the government. Therefore, we can see the behavioural change of parties, especially in correspondance to a change of parliament.

To get a closer look at how correlations vary within government and opposition coalitions, we repeated the whole analysis by grouping parties according to their coalition. Our main finding is that mean correlation tends to be higher within parties in the op-

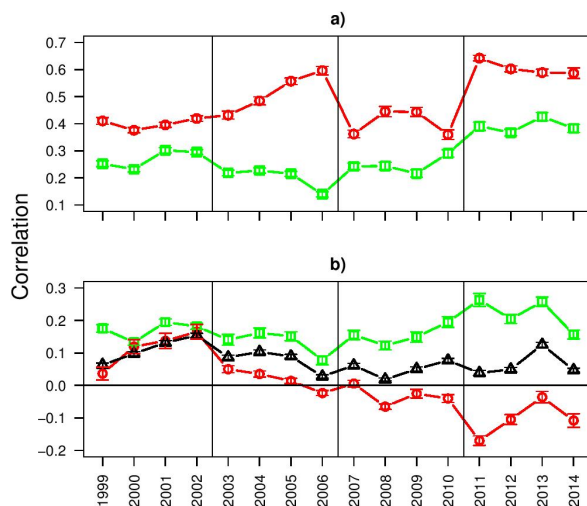


Figure 8: Panel a): Mean intra-party correlation for government parties (green squares), and for opposition parties (red circles), over time. Opposition parties display a higher mean correlation throughout the parliament terms. Panel b): mean inter-party correlation between all government parties (green squares), between all opposition parties (red circles) and between government and opposition parties (black triangles), over time. Here correlations between government parties become noticeably higher over time than those between government and opposition, which still best those between opposition parties. The 3 curves tend to spread out over time, and again collaboration between government parties is high, whilst that between opposition parties grows scarcer and scarcer. Error bars are 6 standard deviations of the mean. Vertical lines separate terms, and the 0-value is clearly indicated.

position than parties in the government, while collaboration is higher between parties in the government than between those in the opposition, as expected (see Fig. 8). Collaboration between government and opposition is low, but still better than collaboration within the opposition, which becomes negative by the end of the second parliament, stays slightly negative during the III and takes a dive during the fourth. When looking at mean inter-party correlation we also notice how the trend is for the 3 curves to spread out over time.

In order to stress how high values of the mean correlation within the opposition do not reflect a collaboration between parties, we also looked at the distribution of correlations within the opposition. What we found is that these high values within the opposition are only due to high intra-party correlations, but do not point towards a strong collaboration between parties, as shown in Fig. 9. In fact, the distribution of correlations is not homogeneous, but displays 2 peaks. By coloring the area due to intra-party correlations, we can explicitly see how indeed these fall on the high side (right peak), while correlations between opposition parties (left peak) center on near-zero values for the first 3 terms and turn negative during the last term.

On the other hand, if we look at the distribution of correlations within the government, we find that it is rather homogeneous, which is an indication of how government parties present more of a united stance.

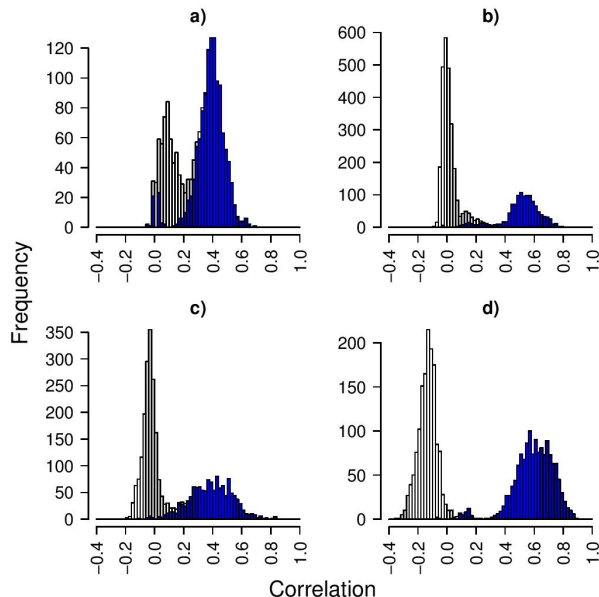


Figure 9: Distribution of correlations between members of opposition parties. Panel a) I parliament; b) II parliament; c) III parliament; d) IV parliament. The shaded area (in blue) represents correlations within each party, the rest are correlations between opposition parties. All distributions are two-peaked, showing how collaboration is strong within each party (right peak), but weak between different opposition parties (left peak), this feature grows stronger from term to term and is enhanced during the last one.

Annual distance within each parliament. We also study the Frobenius distance between the correlation matrices of members corresponding to the 4 different years composing each parliament. We aim to understand whether there are similar years within a parliament term and unusual ones, in terms of how members cooperate as expressed by the correlation matrix relative to each year.

The Frobenius distance between two matrices Σ_1 and Σ_2 is defined as:

$$F(\Sigma_1, \Sigma_2) = \sqrt{\text{Tr}[(\Sigma_1 - \Sigma_2)(\Sigma_1 - \Sigma_2)^T]}.$$

In order to compare two matrices using this measure, their dimension must be the same. For this reason, we restricted the analysis to members who signed at least 1 initiative per year (and thus are active every year).

The results are shown in Fig. 10 for all terms, although it should be borne in mind that it's not possible to compare parliaments, because the corresponding correlation matrices involve different members, are in general of different dimensions, and the Frobenius distance depends on the dimension of the matrices involved.

From Fig. 10, we can see an initial stabilising trend, followed by a diverging one. Indeed, for the first term, the most anomalous year is the first, when the parliament has just formed, while the last 2 years are the most similar. The second term is in a sort of stable equilibrium, while for the last two terms, the trend is diverging, with the 3rd and 4th years most dissimilar and the first two being more or less alike.

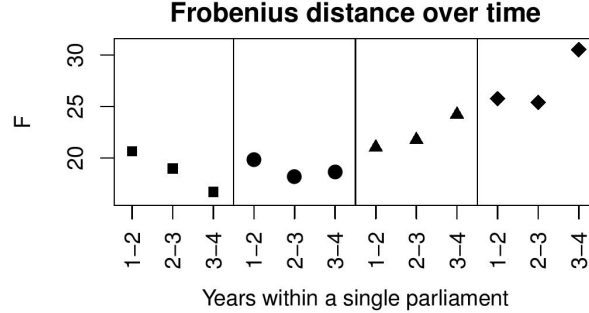


Figure 10: Frobenius distance between correlation matrices corresponding to a year and the one immediately after, within each parliament. Different symbols indicate different parliaments, separated by vertical black lines: squares for the I term, circles for the II, triangles for the III and diamonds for the last one.

2.5. Internal Structure: Reciprocity and Disparity

We now turn our attention to individual parliament members. Specifically, we are looking for special nodes, as for example members of relevance for their own party (local leaders), or members who gain favour from many others (influential people), as well as members who mostly sign other’s initiatives (followers).

Reciprocity. A feature of our social system we’re interested in is reciprocity, or in other words the tendency of a member to reciprocate a signature he received.

As a measure of reciprocity between two people, we suggest the following:

$$r_{i,j} = \frac{n_{i \rightarrow j} - n_{j \rightarrow i}}{n_{i \rightarrow j} + n_{j \rightarrow i}}, \quad i < j$$

where $n_{i \rightarrow j}$ stands for the number of signatures by i on initiatives proposed by j and $n_{j \rightarrow i}$ is vice-versa. The two extreme categories are the one with $r = 0$, implying full reciprocity, and the one with $r = \pm 1$, implying unreciprocation of signatures (either from i to j or from j to i). Interestingly, most members fall in these two categories, meaning that either they fully reciprocate all signatures they receive ($r = 0$), or they just gain signatures from others without ever once returning them ($r = \pm 1$).

In our data we can single out a few members in each term who receive an outstanding amount of signatures (from several hundreds, to a few thousands) from all the others, that they do not reciprocate. TABLE A.4 in the Appendix shows the top three scorers in reciprocity during last term.

Disparity. An interesting feature of the internal structure of our network is, more in general, the presence of net givers and net receivers of signatures. We introduced four centrality measures, that aim to weigh the tendency of a member to sign others’ initiatives against his tendency to receive signatures. Basically, we’re introducing “disparity” measures of out-degree/in-degree imbalance in our bipartite network.

The measures we consider take into account the number of initiatives proposed and signed by member i , the number of signatures member i received, the number of different

members who signed i 's initiatives and the number of different members whose initiatives i signed. Unfortunately, given the several dimensions of heterogeneity involved, there is no unique way to calculate out-degree/in-degree imbalance for our system.

The first measure we consider is the number of initiatives n_p proposed by member i minus the total number of initiatives i signed, n_s , normalized to the sum:

$$d_1 = \frac{n_p - n_s}{n_p + n_s}.$$

If d_1 is positive, it indicates that member i tends to propose more initiatives than he signs: he's more focused on submitting proposals than on signing any.

The second measure is the number of signatures i received on average per initiative proposed n_r/n_p , minus the number of initiatives i signed, normalized to the sum:

$$d_2 = \frac{n_r/n_p - n_s}{n_r/n_p + n_s}.$$

If d_2 is positive, it means that member i receives more signatures per initiative than he affixes. Whoever scores high here, proposed just a few very successful initiatives, without signing many initiatives proposed by others.

The third measure is the absolute number n_r of signatures i received minus the number of initiatives i signed, normalized to the sum:

$$d_3 = \frac{n_r - n_s}{n_r + n_s}.$$

A positive value of d_3 implies a member is a net receiver overall, regardless of the number of initiatives he proposed.

The last measure is a bit different from the previous three in that it focuses on members instead of signatures. It counts the number of different members m_r who signed i 's initiatives minus the number of different members m_s whose initiatives i signed, normalized to the sum:

$$d_4 = \frac{m_r - m_s}{m_r + m_s}.$$

This measure, d_4 , is a direct proxy of whether a member i has a wide network of collaboration: a high (low) score here indicates a member is a global receiver (signer), linking members from different clusters.

TABLE A.9 and A.10 in the Appendix show the top and bottom 10 scores, respectively, for all these measures. Looking at who comes out on top, in terms of the number of unreciprocated signatures received and the scores on the four disparity measures, and considering also the lowest scores, we get the following picture, consistent over all terms: there are few highly active members who rank top, the leaders, and a bulk of members who rank low and are mainly signers (followers).

Opposition parties are the most active, directing initiatives through one or two selected members who are not party leaders, but often are veterans or chairmen, while the rest of the party, the signers, simply sign their initiatives. This is easily seen by looking at the maximum number of unreciprocated signatures received, at measures d_1 (high number of proposals) and d_3 (high absolute number of signatures received), where we

find the same couple of members from opposition parties, showing a huge level of activity, along with many members of government parties who display a moderate activity.

Government parties dominate measures d_2 and d_4 , where the first number represents the mean number of signatures received per proposal, while the second number measures the number of signatures by different members. This is probably due to the way government parties collaborate among them by both proposing and signing each other's initiatives, and in so doing they display a larger area of influence.

Although the top scores appear dominated by the government, with just a few very active members in the opposition, acting for their own parties, the lowest scores are completely ruled by the opposition. The reason for this is, as we stated, that the whole party focuses support for its spokesmen, trying its best to push their plans through, since they can't count on any substantial collaboration outside their own party.

The scenario revealed by this analysis is one of a government with a strong collaboration web between parties, with some moderately active members who split the job of proposing and signing initiatives among themselves, thus widening their collaboration network across party lines. On the other hand, each opposition party presents a united front, channeling proposals through one or two members and raising consensus within the party.

3. Conclusions

The statistically validated network method has been used to retrieve the informative structure of a social system, the Finnish parliament. After reducing the density of links in our network to the fundamental core, we detect communities of parliament members, and characterize them in terms of gender, party and electoral district. The result is that collaboration between parliament members is driven by party, especially for members belonging to the opposition, and by the district of election, especially for members belonging to the parties that support the government. At difference with other studies on Finnish parliament members and candidates [34], gender does not play any major role in the collaboration of parliament members. This result indicates that the observed polarization by gender (and across party) of some political views [34] does not influence the actual behavior of parliament members when it comes to propose and sign initiatives. The characterization of communities in terms of electoral district and party is rather stable across the first three parliaments under scrutiny. However, a major change of the collaboration structure occurs between the third and fourth parliament, due to the outburst of the PS party in 2011 election. As a consequence the number of large parties, in terms of parliament representatives, increased from three to four. Furthermore, such an outburst triggered a polarization of members behavior in terms of party, even within the government coalition.

The statistically validated network analysis does not allow one to gain deep insight about the behavior of members who show peculiar patterns of signatures, and the internal structure of revealed communities. To gain a closer insight into the nested structure of communities, how they emerge and how collaboration profiles change from parliament to parliament, we performed a correlation analysis and constructed hierarchical trees of parliament members, based on pairwise similarity. We show that the structure of similarity of members changes from a term to the next, due to the varying composition of the parliament itself. A major change of the system's structure occurs between the third

an fourth parliament. Indeed the increased number of parliament members from the PS party in the last term, not only changed dramatically the composition of the parliament in terms of the relative weight of parties, but it also polarized the behavior of members, who became less prone to cooperation outside of their own party. Such an inter-party collaboration, though, is more pronounced among parties that support the government, and such a stylized fact is persistent throughout the sixteen years analyzed. The result is that government and opposition show an intrinsically different behavior in that, within the former, collaboration happens across party lines, with distinct sub-clusters characterized by district of election, whilst, within the opposition, parties close ranks and tend not to collaborate with each other. As a result, different parties in the opposition can even become negatively correlated, as is the case during the last parliament. Such a behavior of opposition parties is so extreme that they tend to collaborate more with parties that support the government than within the opposition.

The correlation analysis also shows how the parliament inner workings change, according to the Frobenius distance between correlation matrices of one year and the next within a single term. Specifically during the first term we observe a negative trend of the Frobenius distance over the years, indicating that the correlation structure of the parliament tends to stabilize over time. Such a stabilizing trend almost disappears in the second parliament, where the similarity between the correlation matrices is stable over time. In the third parliament, we observe an inversion of the trend, and correlation matrices at the end of the term are more different than those at the beginning. Finally, such a de-stabilizing trend becomes very pronounced in the last parliament.

To study the system from the perspective of influential parliament members, we develop ad hoc measures, and look at parliament members as leaders and followers. The result is again a different behavioral pattern between opposition, which channels all initiatives through one or two leaders and a bulk of followers, and government, which develops a wider collaboration web, displaying a variety of influential members who split the load between them.

The presented quantitative analysis of initiative co-signature has revealed the structure of the collaboration network among parliament members and its evolution over time, by also revealing how deeply such a structure changed in 2011, that is, in correspondence to the rise of the nationalist party (PS) which changed the number of major players within the parliament from 3 to 4.

4. Acknowledgements

J.P. acknowledges financial support from Magnus Ehrnrooth Foundation. E.P. thanks the Turku groups for hospitality during research visits and acknowledges financial support from Quantum Optics Lab (Turun Yliopisto).

Appendix A. Tables

Parties, political position and number of seats

	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
<i>KD</i> (<i>Right</i>)	10 (4.2%)	7 (5.3%)	7 (4.9%)	6 (4.0%)
<i>KESK</i> (<i>Centre</i>)	48 (22.4%)	55 (24.7%)	51 (23.1%)	35 (15.8%)
<i>KOK</i> (<i>Right</i>)	46 (21.0%)	40 (18.6%)	50 (22.3%)	44 (20.4%)
<i>PS</i> (<i>Right</i>)	1 (1.0%)	3 (1.6%)	5 (4.1%)	39 (19.1%)
<i>RKP</i> (<i>Centre</i>)	12 (5.1%)	9 (4.6%)	10 (4.6%)	10 (4.3%)
<i>SDP</i> (<i>Left</i>)	51 (22.9%)	53 (24.5%)	45 (21.4%)	42 (19.1%)
<i>VAS</i> (<i>Left</i>)	20 (10.9%)	19 (9.9%)	17 (8.8%)	14 (8.1%)
<i>VIHR</i> (<i>Left</i>)	11 (7.3%)	14 (8.0%)	15 (8.5%)	10 (7.3%)

Table A.1: Political party groups, with party’s political position (in parenthesis) and seats for each parliament term. The numbers in parenthesis give percentage from total number of votes. KD=Christian Democrats, KESK=Center Party, KOK=National Coalition Party, PS=True Finns, RKP=Swedish People’s Party, SDP=Social Democratic Party, VAS=Left Alliance, VIHR=Green League.

Electoral districts with seats

Åland	alan	1
Etelä-savo	esav	6
Häme	hame	14
Helsinki	hels	21
Central Finland	keski	10
Kymi	kymi	12
Lapland	lapp	7
Oulu	oulu	18
Pirkanmaa	pirk	18
North Karelia	pkar	6
Pohjois-Savo	psav	10
Satakunta	sata	9
Uusimaa	uusi	34
Vaasa	vaas	17
Varsinais-Suomi	vars	17

Table A.2: Electoral districts, their abbreviations, and number of seats.

Government and Opposition				
	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
Gov.	KOK	KESK	KESK	KD
	RKP	RKP	KOK	KOK
	SDP	SDP	RKP	RKP
	VAS		VIHR	SDP
	VIHR			VAS VIHR
Opp.	KD	KD	KD	KESK
	KESK	KOK	PS	PS
	PS	PS	SDP	
		VAS	VAS	
		VIHR		

Table A.3: Parties in the Government (top panel) and Opposition (bottom panel) coalitions for each term.

Reciprocity top 3 scores		
<i>Surname</i>	<i>Party</i>	<i>Signatures</i>
<i>Tiilikainen</i>	KESK	3,002
<i>Ruohonen</i>	PS	1,412
<i>Kalmari</i>	KESK	292

Table A.4: Reciprocity top 3 scorers (in italics). All are from opposition parties (in bold letters) and receive the highest number of unreciprocated signatures.

Communities in the I parliament					
N_C	<i>Party</i>	<i>District</i>	<i>Area</i>	<i>Coalition</i>	<i>Position</i>
53	SDP,VAS	pirk	rural	gov	left
48	KESK	-	-	opp	centre
23	KOK	uusi	metro	gov	right
15	-	vars	-	-	-
12	RKP	vaas	-	-	-
11	VIHR	-	-	-	left
10	KD	-	-	opp	right

Table A.5: Community characterization for the I term. N_C is the number of members in each community, characterizing attributes are indicated on the top row: Party, District, Area, Party's Coalition and Party's Political Position. Communities are ordered by decreasing size.

Communities in the II parliament

N_C	<i>Party</i>	<i>District</i>	<i>Area</i>	<i>Coalition</i>	<i>Position</i>
40	KOK	-	-	opp	right
35	VIHR	uusi	metro	-	left
26	VAS	psav	-	-	left
25	KESK	pkar,kymi	rural	gov	centre
18	-	hame,vars	-	-	-
15	PS	vaas	-	-	-
9	-	pirk	-	-	-
5	-	esav	-	-	-
4	-	sata	-	-	-

Table A.6: Community characterization for the II term.

Communities in the III parliament

N_C	<i>Party</i>	<i>District</i>	<i>Area</i>	<i>Coalition</i>	<i>Position</i>
37	SDP	-	rural	opp	left
30	PS	uusi	metro	-	-
22	RKP	vaas	-	-	centre
20	VAS	-	-	opp	left
14	-	pirk	-	-	-
13	-	vars	-	-	-
12	-	hame	-	-	-
10	-	esav	-	-	-
3	-	lapp	-	-	-

Table A.7: Community characterization for the III term.

Communities in the IV parliament

N_C	<i>Party</i>	<i>District</i>	<i>Area</i>	<i>Coalition</i>	<i>Position</i>
60	SDP	-	-	gov	left
40	PS	-	-	opp	right
40	KESK	-	-	opp	centre
13	RKP	vaas	-	-	-

Table A.8: Community characterization for the IV term.

Disparity top 10 Scorers

Surname	Party	n_p	n_s	n_r	m_s	m_r
<i>Tiilikainen</i>	KESK	147	40	4750	26	39
Thors	RKP	22	15	31	10	4
<i>Ruohonen</i>	PS	195	168	6174	45	51
Arhinmaki	VAS	3	6	30	5	11
Nylander	RKP	7	17	7	11	3
Nylund	RKP	23	59	83	21	40
Gestrin	RKP	6	17	20	13	7
Hanninen	VAS	3	13	12	9	8
Kanerva	KOK	9	41	102	34	26
Eloranta	SDP	9	41	136	26	75
Harkimo	KOK	1	41	103	31	103
Kopra	KOK	1	44	101	33	101
Sarkomaa	KOK	1	27	56	19	56
Makela	KOK	4	36	256	29	133
Kymalainen	SDP	2	55	188	39	135
Mantymaa	KOK	1	72	120	33	120
Arhinmaki	VAS	3	6	30	5	11
Sinnemaki	VIHR	2	13	39	10	34
Kataja	KOK	3	35	143	26	132
Lapintie	VAS	1	27	31	13	31
<i>Tiilikainen</i>	KESK	147	40	4750	26	39
<i>Ruohonen</i>	PS	195	168	6174	45	51
Makela	KOK	4	36	256	29	133
Arhinmaki	VAS	3	6	30	5	11
Satonen	KOK	5	50	247	31	129
Pelkonen	KOK	5	45	190	35	130
Kataja	KOK	3	35	143	26	132
Autto	KOK	4	57	225	39	136
Kauma	KOK	4	37	131	28	89
Tolvanen	KOK	4	52	183	32	109
Kataja	KOK	3	35	143	26	132
Toivakka	KOK	4	46	145	26	133
Makela	KOK	4	36	256	29	133
Satonen	KOK	5	50	247	31	129
<i>Kalmari</i>	KESK	11	229	427	40	162
Kalliorinne	VAS	6	46	151	30	117
Pelkonen	KOK	5	45	190	35	130
Mantymaa	KOK	1	72	120	33	120
Viitamies	SDP	6	65	174	33	122
Autto	KOK	4	57	225	39	136

Table A.9: Disparity top scorers for the IV term, we report the top 10, that is, 5% right tail of the distribution. Each panel corresponds to a disparity measure, in order from top to bottom: d_1 , d_2 , d_3 , d_4 . Opposition parties are in bold letters. The top 3 scorers in Reciprocity, all from opposition parties, also score high in disparity measures (in italics). The trend is similar over all parliaments.

Disparity bottom 10 Scorers

Surname	Party	n_p	n_s	n_r	m_s	m_r
Puumala	KESK	1	215	101	36	101
Koskela	PS	3	528	53	56	28
Vaatainen	PS	1	368	25	54	25
Vahamaki	PS	3	455	30	58	10
Alatalo	KESK	2	208	14	47	11
Lintila	KESK	2	229	35	32	34
Lohi	KESK	2	239	2	43	1
Maijala	KESK	2	205	110	48	104
Eerola	PS	4	460	143	66	76
Virtanen	PS	2	244	21	37	13
Lohi	KESK	2	239	2	43	1
Korhonen	KESK	13	226	33	47	14
Pirttilahti	KESK	27	265	79	50	23
Torniainen	KESK	18	245	34	43	11
Hautala	KESK	49	278	207	48	13
Kaikkonen	KESK	24	215	79	46	41
Rundgren	KESK	3	239	10	44	8
Vehkapera	KESK	6	223	18	43	10
Niikko	PS	31	565	296	63	38
Yrttiaho	VIHR	3	56	3	39	1
Lohi	KESK	2	239	2	43	1
Jaskari	kok	1	50	1	31	1
Peltokorpi	KESK	1	71	2	5	2
Rundgren	KESK	3	239	10	44	8
Yrttiaho	VIHR	3	56	3	39	1
Vahamaki	PS	3	455	30	58	10
Alatalo	KESK	2	208	14	47	11
Vaatainen	PS	1	368	25	54	25
Tiainen	VAS	1	64	5	44	5
Vehkapera	KESK	6	223	18	43	10
Lohi	KESK	2	239	2	43	1
Yrttiaho	VIHR	3	56	3	39	1
Jaskari	KOK	1	50	1	31	1
Feldt	SDP	2	23	2	18	1
Ojala	SDP	4	43	7	29	2
Saarinen	SDP	6	47	11	32	3
Palm	KD	7	74	28	36	4
Tiainen	VAS	1	64	5	44	5
Mustajarvi	VIHR	13	65	16	46	5
Jaaskelainen	KD	6	59	24	25	4

Table A.10: Disparity worst 10 scorers for the IV term, we report the bottom 10, that is, 5% left tail of the distribution. Each panel corresponds to a disparity measure, in order from top to bottom: d_1 , d_2 , d_3 , d_4 . Opposition parties are in bold letters. Lowest scores are dominated by opposition members who mainly support their spokesmen. The trend is similar over all parliaments.

- [1] D. J. Watts and S. H. Strogatz, Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440-442 (1998).
- [2] A. L. Barabasi and R. Albert, Emergence of scaling in random networks. *Science* **15**, 509-512 (1999).
- [3] M.E.J. Newman, D. J. Watts and S. H. Strogatz, Random graph models of social networks. *PNAS* **99**, 2566-2572 (2002).
- [4] R. Guimerà, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas, Self-similar community structure in a network of human interactions. *Phys. Rev. E* **68**, 065103 (2003).
- [5] P. Holme and M. E. J. Newman, Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* **74**, 056108 (2006).
- [6] B. Kozma and A. Barrat, Consensus formation on adaptive networks. *Phys. Rev. E* **77**, 016102 (2008).
- [7] Y. Zhang, A. J. Friend, A.L. Traud, M. A. Porter, J. H. Fowler and P. J. Mucha, Community structure in congressional cosponsorship networks. *Physica A* **387**, 1705-1712 (2008).
- [8] M. Tumminello, S. Miccichè, F. Lillo, J. Varho, J. Piilo and R. N. Mantegna, Statistically validated networks in bipartite complex systems. *J. Stat. Mech.* **2011**, P01019 (2011).
- [9] M. R. Anderberg, *Cluster Analysis for Applications* (Academic, New York, 1973).
- [10] R. N. Mantegna, Hierarchical structure in financial markets. *Eur. Phys. J. B* **11**(1), 193-197 (1999).
- [11] M. Tumminello, F. Lillo, and R. N. Mantegna, Correlation, hierarchies, and networks in financial markets. *J. Econ. Behav. Organ.* **75**(1), 40-58 (2010).
- [12] M. Tumminello, S. Miccichè, F. Lillo, J. Piilo and R. N. Mantegna, Statistically Validated Networks in Bipartite Complex Systems. *PLoS ONE* **6** (3), e17994 (2011).
- [13] C. Curme, M. Tumminello, R. N. Mantegna, H. E. Stanley and D. Y. Kenett, Emergence of statistically validated financial intraday lead-lag relationships. *Quant. Financ.* **15**, 1-12 (2015).
- [14] M. Tumminello, F. Lillo, J. Piilo and R. N. Mantegna, Identification of clusters of investors from their real trading activity in a financial market. *New J. Phys.* **14**(1), 013041 (2012).
- [15] M. Tumminello, C. Edling, F. Lijeros, R. N. Mantegna and J. Sarnecki, The phenomenology of specialization of criminal suspects. *PLoS One* **8**(5), e64703 (2013).
- [16] V. Hatzopoulos, G. Iori, R. N. Mantegna, S. Miccichè and M. Tumminello, Quantifying preferential trading in the e-MID interbank market. *Quant. Financ.* **15**(4), 693-710 (2015).
- [17] G. Iori, R. N. Mantegna, L. Marotta, S. Miccichè, J. Porter and M. Tumminello, Networked relationships in the e-MID Interbank market: A trading model with memory. *J. Econ. Dyn. Control* **50**, 98-116 (2015).
- [18] M.-X. Li, Z.-Q. Jiang, W.-J. Xie, S. Miccichè, M. Tumminello, W.-X. Zhou and R. N. Mantegna, A comparative analysis of the statistical properties of large mobile phone calling networks. *Sci. Rep.* **4**, 5132 (2014).
- [19] M. Tumminello, S. Miccichè, L. J. Dominguez, G. Lamura, M. G. Melchiorre, M. Barbagallo and R. N. Mantegna, Happy aged people are all alike, while every unhappy aged person is unhappy in its own way. *PLoS One* **6**(9), e23377 (2011).
- [20] A.Pajala, Lonely riders and cooperation networks: member of parliaments’ private motions in the finnish parliament 1999-2010. *World Political Science Review* **10**(1), 91-110 (2014).
- [21] I. Mattson, in *Parliaments and Majority Rule in Western Europe*, edited by H. Döring (Campus Verlag/St Martins Press, Frankfurt/New York, 1995).
- [22] C. E. Bonferroni, *Teoria statistica delle classi e calcolo delle probabilità* (Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 1936).
- [23] H. Abdi, *Bonferroni and Šidák corrections for multiple comparisons* (ed. N. J. Salkind, Encyclopedia of Measurement and Statistics, Thousand Oaks, CA: Sage, 2007).
- [24] S. Fortunato, Community detection in graphs. *Phys. Rep.* **486**(3), 75-174 (2010).
- [25] M.E.J. Newman, Mixing patterns in networks. *Phys. Rev. E* **67**(2), 026126 (2003).
- [26] M.E.J. Newman and M. Girvan, Finding and evaluating community structure in networks. *Phys. Rev. E* **69**(2), 026113 (2004).
- [27] A. Arenas, A. Fernandez and S. Gomez, Analysis of the structure of complex networks at different resolution levels. *New J. Phys.* **10**(5), 053039 (2008).
- [28] J. Duch and A. Arenas, Community identification using Extremal Optimization. *Phys. Rev. E* **72**(2), 027104 (2005).
- [29] M.E.J. Newman, Modularity and community structure in networks. *PNAS* **103**(23), 8577-8582 (2006).
- [30] M.E.J. Newman, Fast algorithm for detecting community structure in networks. *Phys. Rev. E* **69**(6), 066133 (2004).

- [31] A. Strehl, Ph.D. thesis, University of Texas at Austin, 2002.
- [32] Y.-Y. Yao, *Entropy Measures, Maximum Entropy Principle, and Emerging Applications* (ed. Karmeshu) 115-136 (Springer, 2003).
- [33] A. Fiasconaro, M. Tumminello, V. Nicosia, V. Latora, and R. N. Mantegna, Hybrid recommendation methods in complex networks. *Phys. Rev. E* **92**, 012811 (2015).
- [34] M. Tumminello, S. Miccichè, J. Varho, J. Piilo, and R. N. Mantegna, Quantitative Analysis of Gender Stereotypes and Information Aggregation in a National Election. *PLoS ONE* **8(3)**, e58910 (2013).