scientific reports

OPEN

Maximal modularity and the optimal size of parliaments

Luca Gamberi, Yanik-Pascal Förster, Evan Tzanis, Alessia Annibale & Pierpaolo Vivo✉

An important question in representative democracies is how to determine the optimal parliament size of a given country. According to an old conjecture, known as the cubic root law, there is a fairly universal power-law relation, with an exponent equal to 1/3, between the size of an elected parliament and the country's population. Empirical data in modern European countries support such universality but are consistent with a larger exponent. In this work, we analyse this intriguing regularity using tools from complex networks theory. We model the population of a democratic country as a random network, drawn from a growth model, where each node is assigned a constituency membership sampled from an available set of size D. We calculate analytically the modularity of the population and find that its functional relation with the number of constituencies is strongly non-monotonic, exhibiting a maximum that depends on the population size. The criterion of maximal modularity allows us to predict that the number of representatives should scale as a power-law in the size of the population, a finding that is qualitatively confirmed by the empirical analysis of real-world data.

In modern times, representative democracies have played a leading role in the advancement of human rights, education, and technology on a global scale. At the heart of every representative democracy is a centralised parliament: an assembly of elected citizens who are delegated by their constituents to exercise the legislative power, and to keep the government in check[1]. This apparatus has an operating cost and, in the shadows of political scandals, economic crises, and social turmoil, people have questioned the effectiveness of their country's costly political and administrative structure and have claimed that a reduction of the number of elected representatives would reduce deviant behaviours and enhance efficiency of parliamentary works[2]. However, there is so far no sound analytical framework to determine the optimal parliament size of a given country, so to ensure an adequate representation and cost-effectiveness, which are both in the public interest. In this paper, we argue that a *principle of maximal modularity* can provide some reliable guidance on how to determine the absolute number of representatives required for efficient public representation in a democratic country. This principle may therefore provide a transparent reference point to inform public policies.

Generally speaking, the ideal number of members of Parliament (MPs) has to strike a balance between *efficiency*, in terms of the share of power held by each MPs and their ability to realise their electoral agenda, and *optimal representativity*, i.e. the ability of the MPs to promote the instances of their voters, in proportion to their number. Both criteria are encoded in the assembly size, as a bigger chamber allows constituencies to be smaller and thus more homogeneous in terms of character, local economic activity, and social needs. On the other hand, it diminishes the influence and resources that each MP can count on to advance their agenda and thus promote their constituents' interests[3]. The "efficiency" paradigm has been at the core of a flourishing line of research amongst political scientists and "electoral engineers", since the late 1980s[4]. Researchers have revealed the effect of different electoral systems on the efficiency and stability of political architecture, in relation to the size of the corresponding assembly[5]. Representation of minority groups, gender quotas, ballot votes, and district sizes are believed to heavily influence the efficiency of a parliament and the relative voting power of political parties[6,7]. Another pressing issue that has been thoroughly studied concerns the distribution of relative weights of votes for delegates in international bodies or for parliaments in federal states such as the United States. The most famous approach is the one proposed by Webster and Sainte–Laguë independently, after which several other quotient rules were adopted[8]. Game theory approaches, such as the Penrose square root law[9], were also proposed later on and are currently used for instance in the Council of the European Union, to implement a "one person, one vote" system[10]. An interesting empirical decision-making model linking participation in elections and electoral college size, at any level (local to national), was proposed in[11].

Both the problems of efficiency and relative representativity have been investigated for a long time in the political science literature and share a common denominator: they depend—directly or indirectly—on the absolute chamber size in a way that is yet to be fully understood[6].

Department of Mathematics and Quantitative and Digital Law Lab, King's College London, London WC2R 2LS, UK. ✉email: pierpaolo.vivo@kcl.ac.uk
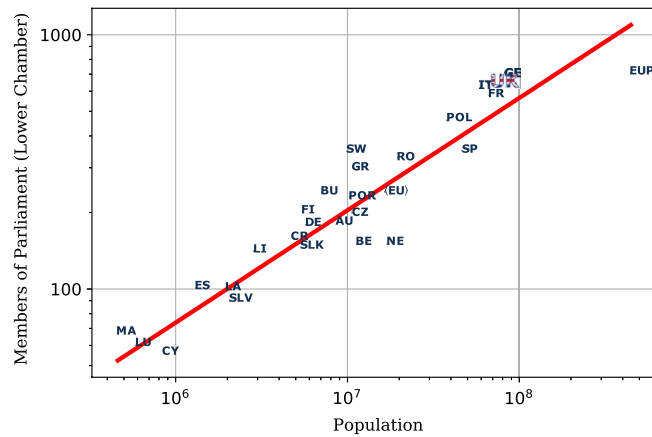
Figure 1. The figure shows a log-log scale plot of the size of lower chamber S vs. population N for some European countries and the EU Parliament. The best fit with a power-law (red line) shows that the size of the chamber grows as $S \approx \alpha N^\gamma$, with $\gamma \approx 0.44$ and $\alpha \approx 0.17$. Demographic data from Eurostat (2017).

In recent times, political scientists and technocrats have heavily relied on the so-called *cubic root law* (CRL) formulated by Taagepera and collaborators in[12–15]. This law follows the realisation that the size of most elected parliaments exhibits a strong statistical regularity with respect to their population. The proposed empirical model optimises the assemblies' representation based on the efficiency of communication between MPs and their constituents. According to Taagepera's arguments, the size of parliaments should follow $S \propto N_0^\gamma$, where $\gamma = 1/3$ and N_0 an "effective" population size, rescaled by considering only the portion of active voters and, among them, the fraction of literate adults[12]. As literacy is believed to be strongly correlated with mobility, the latter rescaling was introduced to account for social mobility in the absence of a reliable direct measure for this parameter[16]. Figure 1 highlights the aforementioned regularity, but when considering the sizes of European lower chambers only, the best fitting curve deviates from the theoretical CRL, resulting in $\gamma \approx 0.44$. A more comprehensive analysis of parliaments' size data can be found in[11]. It is also worth mentioning that the CRL formulated in terms of an effective population was perceived to be in contrast with the spirit of any "good" representation model that should include the entire pool of constituents, regardless of age, political engagement, or education[6]. Moreover, Taagepera's derivation has been criticised in a recent paper[17], whose formulation leads instead to a square-root law (SRL). The same SRL arises from a constitutional game-theoretical model put forward in[18].

In this work, we tackle the democratic representation problem using network theory. Networks have been successfully employed in social science for over 30 years, for their versatility in describing different aspects of political, behavioural, and social interaction between individuals[19]. In social networks, nodes represent social agents (for example, individuals in a population) and links represent their interactions. The network structure contains important information about relational ties in a society and is likely influenced by agents' attributes (e.g. age, occupation, wealth ...)[20]. An important topological feature of real-world social networks is their *scale-free* degree distribution, i.e. the distribution of the number of ties following a power-law[21]. This topology can be reproduced in growing networks using a preferential attachment wiring protocol that was proposed in[22].

The objective of this work is to shed light on the observed statistical regularities in the size of parliaments. We will be focusing on electoral systems in which representatives are elected according to an FPTP ("First-Past-The-Post") principle, i.e. whoever collects the majority of votes within a constituency gets elected. We expect that simple modifications of the model presented here could be devised to account for other scenarios. For example, the case where more than one representative is expressed by the same electoral pool, as in the case of the US Senate, could be accounted for by suitably increasing the number of representatives per constituency.

Taking the United Kingdom as an example, each Member of Parliament is elected to the House of Commons from one of the 650 constituencies. The nature and physical boundaries of the constituencies are regulated by the *House of Commons (Redistribution of Seats) Act* (1944), which prescribes that the MP's role is to "represent the common interest of the residents in a spatially bounded territory". Thus, when constituencies are designed, the legislator should aim to enclose within geographical boundaries areas that share common interests and values[23].

The network model we propose is inspired by this design principle. We build a synthetic scale-free network in which N agents (nodes) represent the entire population of a country that has to be partitioned into constituencies, each electing their MP to the national Parliament. Two citizens are connected if there is a stable social interaction between them. Individuals are therefore arranged into social communities, as proposed, for instance, in[24]. The key result of our paper is a principle for determining the optimal number of constituencies, i.e. for grouping the population into electoral clusters that "best" represent the underlying community structure of the network. We remark that we need to strike a balance between representativity and homogeneity in constituency size. Hence, our approach needs to improve upon the standard "community detection" framework, which would allow the constituency size to fluctuate wildly. We achieve this result by constraining the size of the constituencies at the outset, and determining the number of constituencies that optimises the partitioning of the underlying network.

More specifically, we generate synthetic networks from a growth model with preferential attachment to nodes of higher degrees and within the same constituency. We introduce a *mobility* (or affinity) parameter into our
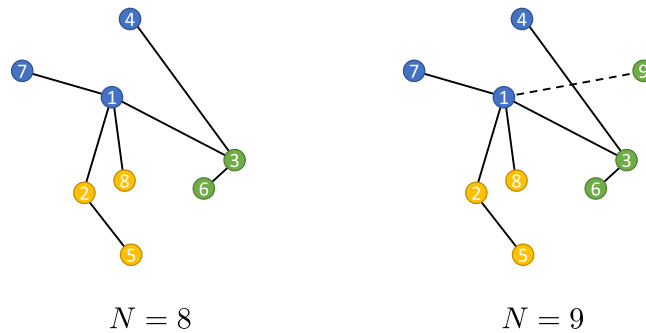
Figure 2. Sketch of a $D = 3$ constituencies network, for the case $m = 1$, at the steps $N = 8$ (left) and $N = 9$ (right) of the growth algorithm, respectively. Membership attributes 1, 2 and 3 represented by the colour blue, yellow and green respectively are allocated sequentially in such a way that $\sigma_i =$ mod $(1 + i, 3)$, hence $\sigma_1 = 1$, $\sigma_2 = 2, \sigma_3 = 3, \sigma_4 = 1$ and so on. The dashed line represents the new connection made at step $N = 9$.

model, of a similar nature to that proposed in[12], which allows us to tune the probability that a node interacts with foreign constituencies. Note that in this approach, each node is assigned a constituency label a priori, and the network topology follows as a result of this assignment. This is in contrast to what usually happens in community detection, where node memberships are determined a posteriori, based on the network topology. In this regard, our approach is based on a generative model for block-structured networks rather than on a cluster detection model.

Working with synthetic networks relieves us from making assumptions over geographical constraints, as constituencies are purely virtual, i.e. designed around groups of people with stronger interpersonal ties. Although this modelling choice may need to be supplemented with more realistic assumptions, non-geographical electoral systems have been proposed in the past with strong supporting arguments in terms of representation of minorities and dispersed communities[25]. Furthermore, geographical constraints would strongly depend on the country at hand, whereas our model aims to be as general as possible.

We adopt the *modularity* as a metric to measure the *goodness* of these partitions, and we derive an exact expression for the average network modularity, in terms of the number D of constituencies for fixed network size N. By maximising the modularity, we are able to determine analytically the optimal number of equally sized constituencies into which networks generated according to our prescription should be partitioned. We show that our findings are robust against the tuning of resolution parameters, by also considering the generalised definition of modularity introduced in[27]. In our investigation, we find that the empirical regularities discussed above arise quite naturally from the topology of the clustered networks that we study here.

The manuscript is organised as follows: in "The model" section, we introduce the network growth model. In "Recursive equations" section, we derive and solve the recursive equation for the expected modularity at generic network size and number of constituencies. In "Approximate solution" section, we present a numerical solution for the maximum modularity as a function of the network size and we construct an approximate scheme to solve the problem analytically. In addition, we show that the size dependence of the maximum modularity is robust when using a generalised definition of modularity which allows to tune the resolution of communities. Finally, we present our main findings in the "Conclusion" and we compare them with empirical evidence. The technical details of our derivation are presented in the Appendix.

Our findings reveal that the optimal partitioning in constituencies for a given population is well approximated by a power-law $S \simeq N^\gamma$, which is in qualitative agreement with the empirical data. Interestingly, we observe that the mobility does not play a significant role in determining the exponent γ, at least for the homogeneous mobility case studied in this work.

The model

We model social interactions within a population by means of simple, undirected networks, which are constructed using a block preferential attachment prescription, a modified version of the Barabási–Albert (BA) algorithm in which the target node's block membership influences the wiring probability[22,28–31].

The network is formed dynamically in such a way that at each time step N, a node N is created, with m stubs, and a constituency membership label $\sigma_N \in \{1, \ldots, D\}$ is assigned, according to a prescribed sequential order such that $\sigma_N = \text{mod}(1 + N, D)$ (see Fig. 2 for an illustration). In this way, any two nodes i and $i + D$ have the same membership and all the constituencies have roughly equal size (their sizes are either identical or differ by one unit). The sequential assignment is a modelling choice that greatly simplifies the analytical treatment presented in this section, however the final outcome does not heavily depend on the way the constituencies σ are assigned, provided that they are on average all equally sized. The network at time step N is represented by an $N \times N$ adjacency matrix $\boldsymbol{A}(N)$, with entries $A_{ij}(N)$ for $i, j \leq N$. We prescribe that the initial configuration of the network be a clique of $m + 1$ nodes. Accordingly, we set the initial time at $m + 1$, so that the growth process starts at $m + 2$.

When a node N is added, each of its m stubs is wired to a random node i of the existing network sampled with probability

3

$$p_{iN} = m \frac{\sum_{\ell=1}^{N-1} A_{i,\ell}(N-1)}{L(N-1)/D} p(\sigma_i|\sigma_N), \tag{1}$$

with $L(N) = \sum_{i=1}^{N} k_i(N) = m(2N - m - 1)$, being the total number of links present in the network, $k_i(N) = \sum_{j=1}^{N} A_{i,j}(N)$ the degree of node i, calculated at $N > m$, and $p(\sigma_i|\sigma_N)$ being the probability that any node with given constituency label σ_N attaches to any of the nodes with constituency σ_i. The denominator in (1) ensures normalization of p_{iN}, such that $\sum_{i=1}^{N} p_{iN} = m$, where we have used $\sum_i k_i(N-1)\delta_{\sigma,\sigma_i} = L(N-1)/D \, \forall \sigma$. As the addition of new nodes cannot modify the links between pre-existing nodes, we have that $A_{ij}(N)$ is the same for any $N \geq \max(i,j)$, so from now on, we will drop the time index from the entries of the adjacency matrix.

The probability $p(\sigma_i|\sigma_N)$ can be parametrised by a mobility parameters μ that controls the likelihood to pick the target constituency, as

$$p(\sigma_i|\sigma_N) = \frac{\mu}{D} + (1-\mu)\delta_{\sigma_i,\sigma_N}, \tag{2}$$

which is normalised $\sum_{\sigma_i=1}^{D} p(\sigma_i|\sigma_N) = 1$, as it should.

Hence, for $\mu = 0$, the new node N will attach necessarily to a member of its own community whereas, for $\mu = 1$, the node N can attach to any community with the same probability $1/D$. Thus, the probability that a new node attaches to a given node of a foreign constituency is μ/D. The contribution $\sum_{\ell=1}^{N-1} A_{i,\ell}/L(N-1)$ in the definition (1) ensures that new nodes attach preferentially to nodes with higher degree. Our attachment prescription realizes a power-law degree distribution with exponent $= 3$, that is typical of social networks[22] and, although the attachment mechanism is extremely simple, it is rich enough for our purposes[32,33]. Using Eq. (1), we can write the probability for the entry $A_{i,N}$, with $i = 1, \ldots, N-1$, of the adjacency matrix, given its previous configuration $\mathbf{A}(N-1)$ and the community membership sequence denoted by $\boldsymbol{\sigma}$, as

$$p(A_{i,N}|\mathbf{A}(N-1), \boldsymbol{\sigma}(N)) = [p_{iN}\delta_{A_{i,N},1} + (1-p_{iN})\delta_{A_{i,N},0}]\delta_{A_{i,N},A_{N,i}}\delta_{k_N(N),m}. \tag{3}$$

Assuming that each of the m stubs is wired independently to a randomly drawn node, the joint distribution for the N-th row and column is

$$p(A_{1,N} \ldots A_{N-1,N}|\mathbf{A}(N-1), \boldsymbol{\sigma}(N)) = \prod_{i=1}^{N-1} p(A_{i,N}|\mathbf{A}(N-1), \boldsymbol{\sigma}(N)), \tag{4}$$

and, by iteration, one can get the full distribution for the configuration $\mathbf{A}(N)$ of the adjacency matrix

$$p(\mathbf{A}(N)|\boldsymbol{\sigma}(N)) = \prod_{i=1}^{N-1} p(A_{i,N}|\mathbf{A}(N-1), \boldsymbol{\sigma}(N)) \prod_{i=1}^{N-2} p(A_{i,N-1}|\mathbf{A}(N-2), \boldsymbol{\sigma}(N-1)) \ldots p(\mathbf{A}(m+1)), \tag{5}$$

with $p(\mathbf{A}(m+1)) = \prod_{i<j}^{m+1} \delta_{A_{i,j},1}\delta_{A_{i,j},A_{j,i}}$, determined by the initial configuration of the growth algorithm.

For the sake of simplicity, we limit our analytical considerations to the case $m = 1$. A different choice of the number of stubs m, or a different initial configuration of the growth model, do not significantly affect the degree distribution, in the large-N limit[34]. Our numerical explorations suggest that this also holds true for our key observable and results. Therefore, we leave the $m > 1$ case for future investigations.

The key observable we will monitor in our model is the modularity, introduced in[35], as a quality factor for a partition of a network in communities. The modularity of a graph is defined as

$$Q_D(N) = \frac{1}{L(N)} \sum_{s,r}^{N} \left[A_{r,s} - \frac{k_r(N)k_s(N)}{L(N)} \right] \delta_{\sigma_r,\sigma_s}. \tag{6}$$

This quantity compares the intra-cluster edge density of a given network (in our case, the clusters are defined by the constituency membership attribute) with the edge density of a null model, i.e. a set of unbiased random graphs that are wired regardless of the community structure but with the same degree sequence as the original network[36]. This comparison mechanism provides a reliable metric to establish the goodness of a network clustering procedure. Moreover, the modularity takes values $Q_D(N) \in [-1, 1]$, with positive values denoting that a graph exhibits a community structure being captured by their assigned memberships[37].

We will use the modularity to assess the cluster structure induced by the sequence $\boldsymbol{\sigma}$ and the underlying social structure originated from the web of connections. We aim to find the number of constituencies that maximises this observable, resulting in the optimal partitioning of the synthetic population created by the growth algorithm. For a network of size N, we have that the expected modularity is given by

$$\langle Q_D(N) \rangle = \frac{1}{L(N)} \left\langle \sum_{s,r}^{N} \left[A_{r,s} - \frac{k_r(N)k_s(N)}{L(N)} \right] \delta_{\sigma_r,\sigma_s} \right\rangle =: \frac{1}{L(N)} a_N - \frac{1}{L(N)^2} b_N, \tag{7}$$

where the expectation is over the distribution (5). At the $(N + 1)$-th step, one row and one column are added to the adjacency matrix as follows
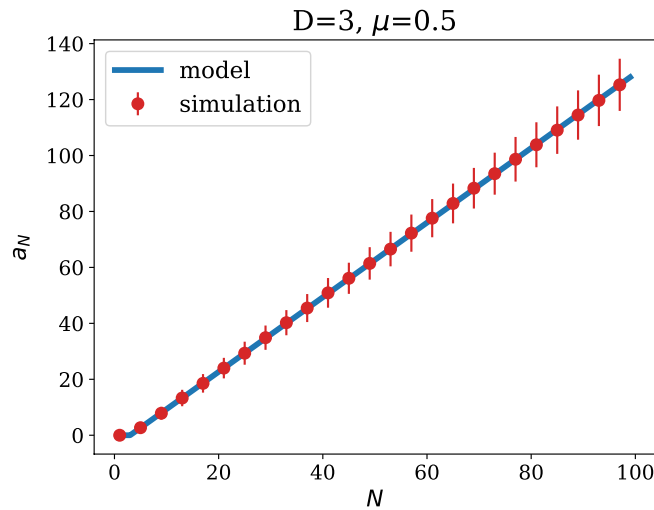
Figure 3. Plots of a_N as a function of the network size N, with parameters $\mu = 0.5$ and $D = 3$. The simulation data were obtained averaging over 200 realisations of the network generative process. Error bars, red vertical segments, are calculated as standard deviation from the mean.

$$A(N+1) = \begin{bmatrix} & & & & 0 \\ & & & & 0 \\ & A(N) & & & \vdots \\ & & & & 1 \\ & & & & 0 \\ \hline 0 & 0 & \cdots & 1 & 0 & 0 \end{bmatrix}.$$

We note that the number of links is deterministic as at each time step $m(= 1)$ links are added to the network. Therefore we argue that an expression for $\langle Q_D(N+1) \rangle$ can be found recursively, in particular by solving recursions for the coefficients a_N and b_N that we present in the following subsection.

Recursive equations. We now construct a recursion for the term a_N. We note that the term a_{N+1} can be split into the contribution from the new row/column and the rest of the matrix as follows

$$a_{N+1} = a_N + 2 \left\langle \sum_{r:\sigma_r = \sigma_{N+1}}^{N} A_{r,N+1} \right\rangle, \tag{8}$$

where the sum runs over all pre-existing nodes (up to N) belonging to the community σ_{N+1}. We recognise that the expectation value represents the probability that node $N+1$ attaches to any node within its own community at step N. Under the assumption that any target constituency is chosen independently of the degrees of its members, and using Eq. (2), one can derive the expression provided in (9), as shown in more detail in the Appendix

$$\left\langle \sum_{r:\sigma_r = \sigma_{N+1}}^{N} A_{r,N+1} \right\rangle = \begin{cases} 0 & \text{when } N < D \\ 1 - \mu \frac{D-1}{D} & \text{when } N \geq D. \end{cases} \tag{9}$$

Using Eq. (9), the solution of the recursion (8) is found as

$$a_N = \begin{cases} 0 & \text{when } N \leq D \\ 2(N-D)\left(1 - \mu \frac{D-1}{D}\right) & \text{when } N > D. \end{cases} \tag{10}$$

A comparison between Eq. (10) and a numerical simulation is shown in Fig. 3. The simulation data in Fig. 3 was obtained by computing the observable $a_N = \sum_{r,s=1}^{N} A_{r,s} \delta_{\sigma_r,\sigma_s}$ on synthetic networks of different sizes, generated using the growth algorithm described in this section. For each size, results were averaged over 200 realisations of the network generative process.

We then consider the recursion for term b_N. Following our definition in Eq. (7),

$$b_{N+1} = \underbrace{\left\langle \sum_{s,r}^{N} k_r(N+1) k_s(N+1) \delta_{\sigma_r,\sigma_s} \right\rangle}_{(i)} + \underbrace{2 \left\langle \sum_{r}^{N} k_r(N+1) \delta_{\sigma_r,\sigma_{N+1}} \right\rangle}_{(ii)} + 1, \tag{11}$$

where we used $k_{N+1}(N+1) = m = 1$. Distinguishing the two cases

- **case N ≥ D**: the expectation (i) in Eq. (11) yields

$$\left\langle \sum_{s,r}^{N} k_r(N+1)k_s(N+1)\delta_{\sigma_r,\sigma_s} \right\rangle = b_N + 4\frac{\mu}{D}(N-1) + 2(1-\mu)C_{N+1}(N) + 1, \qquad (12)$$

as discussed in the Appendix, and for the expectation (ii) in Eq. (11) we get

$$\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} k_r(N+1) \right\rangle = C_{N+1}(N) + 1 - \mu\frac{D-1}{D}, \qquad (13)$$

also in the Appendix, where

$$C_{N+1}(N) = \left\langle \sum_{r:\sigma_r=\sigma_{N+1}} k_r(N) \right\rangle \qquad (14)$$

represents the average number of intra-cluster connections for constituency σ_{N+1} at time N. When evaluating the expectation in Eq. (14) one gets

$$C_{N+1}(N) = \left[\sum_{x=\mathrm{mod}(N+1,D)+1}^{D} \frac{1}{x-1} + \frac{\mu}{D}(\mathrm{mod}(N+1,D)-1) \right](1-\delta_{\mathrm{mod}(N+1,D),0})+ \\ + \mu\frac{D-1}{D}\delta_{\mathrm{mod}(N+1,D),0} + 2\left\lfloor \frac{N}{D} \right\rfloor - 1, \qquad (15)$$

where $\mathrm{mod}(\cdot,D)$ is the modulus operator with divisor D and $\lfloor\cdot\rfloor$ denotes the floor operation.
- **case N < D**: this case is characterised by only N constituencies being yet populated and a uniform probability of wiring, leading to the following expectation for (i) in Eq. (11)

$$\left\langle \sum_{s,r}^{N} k_r(N+1)k_s(N+1)\delta_{\sigma_r,\sigma_s} \right\rangle = b_N + 4\frac{N-1}{N} + 1, \qquad (16)$$

as shown in the Appendix, while the expectation (ii) in Eq. (11) reads

$$\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} k_r(N+1) \right\rangle = 0, \qquad (17)$$

since constituency σ_{N+1} is populated by one node at time step $N+1$, which is however excluded from the sum.

Gathering all the terms, the recursive equation for b_N is found to be

$$b_{N+1} = \begin{cases} b_N + 4\frac{N-1}{N} + 2 & \text{when } 2 < N < D \\ b_N + 4\mu\frac{N-1}{D} + 2(2-\mu)C_{N+1}(N) + 2\left(2-\mu\frac{D-1}{D}\right) & \text{when } N \geq D, \end{cases} \qquad (18)$$

with initial condition

$$b_2 = 2 + 2\delta_{D,1}. \qquad (19)$$

Solving the recursion in Eq. (18), for general N, is not an easy task. In the next Section, we provide an analytical expression for $C_{N+1}(N)$ in the limit $N \gg D$, which turns out to be a good approximation for the exact solution, even at small N. In Fig. 4 we plot the approximate solution for b_N against numerical simulations of the growing process.

Approximate solution. To make analytical progress, we assume that the edges are uniformly distributed between the D communities so that $C_{N+1}(N) \simeq L(N)/D$. This holds true in the limit $N \gg D$, however, uniformity is not expected in the regime $N \sim D$. As shown in Fig. 5, though, the exact expression for $C_{N+1}(N)$ (red dots), defined in Eq. (14), is well approximated by $L(N)/D$ (blue line) in the whole range. The exact calculation for $C_{N+1}(N)$ in the regime $N < D$ is carried out in the Appendix. Combining the results in the two regimes, the analytical bottleneck arising in Eq. (11) - (i) simplifies as follows (see Appendix, eq. (38))

$$\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1s}\delta_{\sigma_r,\sigma_s} \right\rangle = \begin{cases} 2\frac{(N-1)}{N} & \text{when } N < D \\ 2\frac{(N-1)}{D} & \text{when } N \gg D. \end{cases} \qquad (20)$$

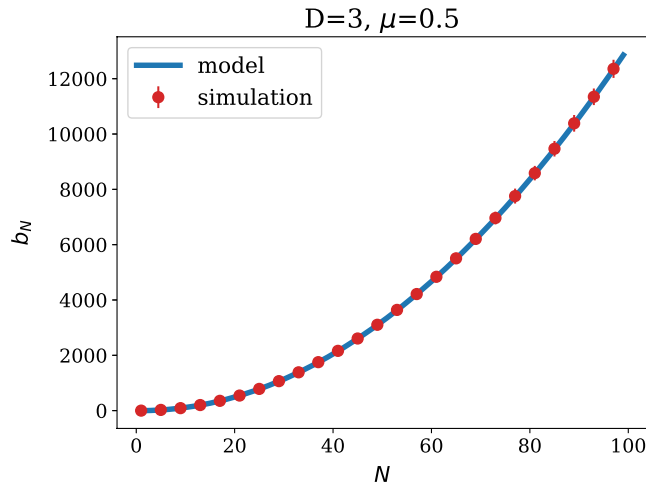Under this approximation, the recursion in Eq. (11), now simplifies to

Figure 4. b_N as a function of the network size N for the parameters $D = 3$ and $\mu = 0.5$. The simulation data were obtained averaging over 200 realisations of the network generative process. Error bars, red vertical segments, are calculated as standard deviation from the mean. Note that error bars are smaller than symbols for most values of N.
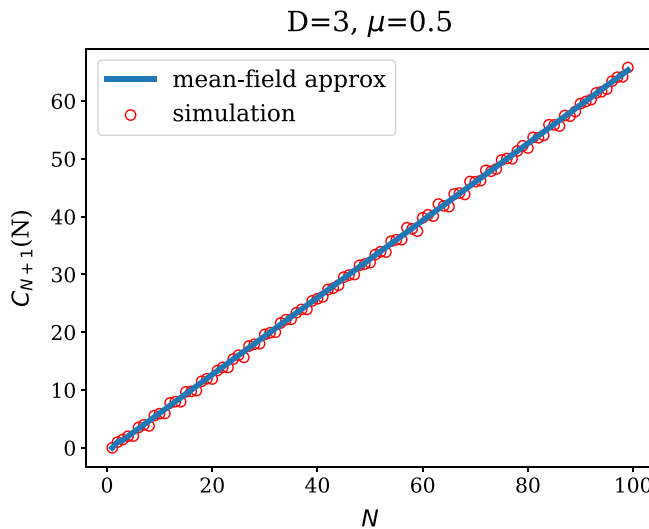

Figure 5. Mean-field approximation of $C_{N+1}(N)$ plotted against numerical simulations for $D = 3$ and $\mu = 0.5$. The simulation data were obtained averaging over 200 realisations of the same experiment.

$$b_{N+1} \simeq \begin{cases} b_N + 4\frac{N-1}{N} + 2 & \text{when } 2 < N < D \\ b_N + 4\frac{2N-1}{D} + 2 & \text{when } N \geq D, \end{cases} \tag{21}$$

with the initial condition (19). The solution is given by

$$b_N = \begin{cases} -2(3 + 2\gamma - 3N + 2\psi^{(0)}(N)) & \text{when } N < D \\ \frac{2}{D}(D - 4N + DN + 2N^2 - 2D(\gamma + \psi^{(0)}(D))) & \text{when } N \geq D, \end{cases} \tag{22}$$

with γ being the Euler–Mascheroni constant and $\psi^{(0)}(x)$ the digamma function, arising from summing the first inverse integers series $\sum_{k=1}^{x-1} \frac{1}{k}$. This approximate solution is in very good agreement with numerical simulations, as shown in Fig. 6.

Inserting in Eq. (7) the expressions for a_N and b_N, provided in Eq. (10) and in Eq. (22) respectively, we obtain the solution for $\langle Q_D(N) \rangle$ in the regime $N \geq D$

$$\langle Q_D(N) \rangle = \frac{N - D}{N - 1}\left(1 - \mu\frac{D-1}{D}\right) - \frac{1}{2D(N-1)^2}(D - 4N + DN + 2N^2 - 2D(\gamma + \psi^{(0)}(D))). \tag{23}$$
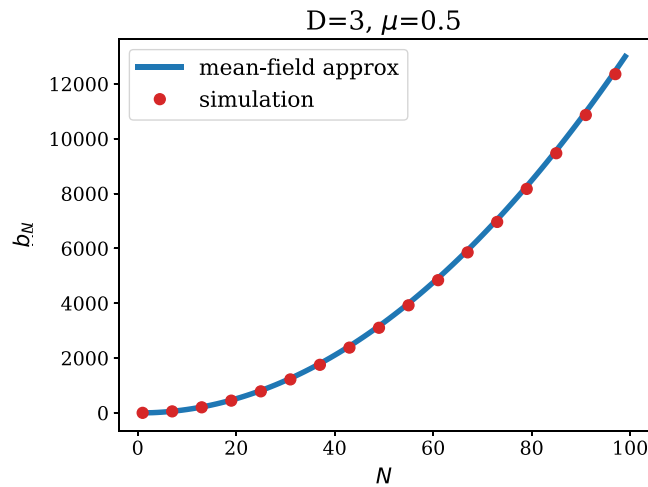
Figure 6. Approximate solution for b_N plotted against numerical simulations of the generative process for $D = 3$ and $\mu = 0.5$. The simulation data were obtained averaging over 200 realisations of the same experiment.
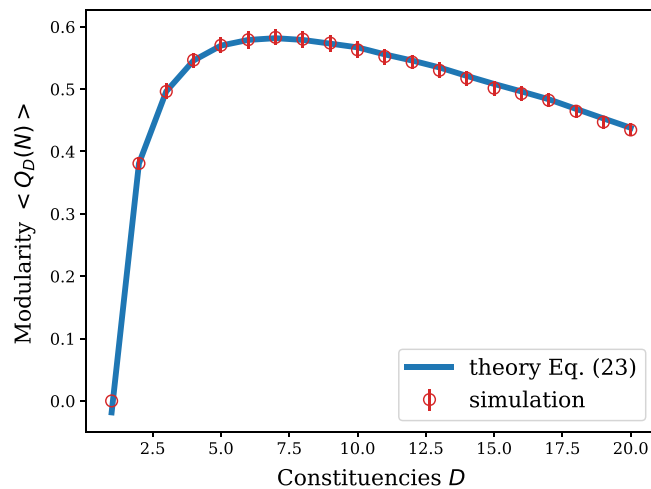

Figure 7. Expression for the average modularity $\langle Q_D(N) \rangle$ in Eq. (23) as a function of the number of constituencies D, with parameters $\mu = 0.2$ and $N = 50$. The simulation data was obtained averaging over 350 realisations of the same experiment. Error bars, red vertical segments, are calculated as standard deviation from the mean.

The numerical simulation in Fig. 7 shows a perfect agreement with the average modularity given by Eq. (23). We observe that the expected modularity is strongly non-monotonic in the number of constituencies. This is in agreement with the following observations about the limiting cases: for $D = 1$, the two terms of the sum in Eq. (6) cancel out, and for $D = N$ the concept of community is lost and the Kronecker delta is always zero, thus both cases result in $Q_D(N) = 0$. Note that in the intermediate regime $1 < D < N$ and provided that $\mu \in [0, 1)$ we have that the probability for a node to link to any of its fellow constituents is higher than the "rest of the population". This ensures that, on average, the modularity is positive. This non-monotonic behaviour was also observed empirically in[38].

Values for $\mu \in [0, 1)$ are consistent with the constituencies design process according to which boundaries should be drawn around local communities. The regime $\mu \geq 1$ would result in an equal or lower intra-constituency edge density compared to the density of outgoing edges, suggesting that the imposed partitions would not capture the real community structure of the network and thus would not be interesting for our purpose. Moreover, $\mu = 1$ is the physiological upper bound to ensure that probability (2) is non-negative.

Furthermore, we observe that the mobility parameter μ dampens the modularity without producing a pronounced shift of its maximum, as shown in Fig. 8. This effect is due to a tightening of the community structures within each constituency as the effect of a decreasing μ is to increase their average intra-cluster density and thus to increase the overall modularity. Conversely, when $\mu \to 1$, nodes attach randomly to any constituency resulting in an average modularity $Q_D(N)$ that tends to zero.
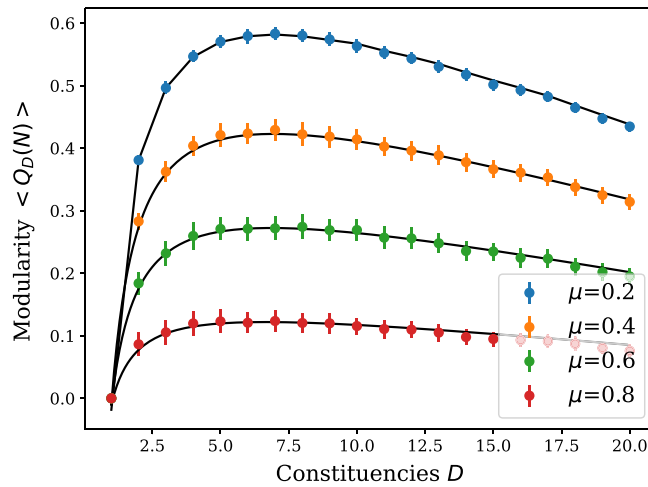
Figure 8. The plot shows the effect of the mobility parameter μ on the modularity $\langle Q_D(N) \rangle$, with $N = 50$. The simulation data were obtained averaging over 200 realisations of the same experiment for $\mu = \{0.4, 0.6, 0.8\}$ and 350 realisations for $\mu = 0.2$. Error bars, vertical segments, are calculated as standard deviation from the mean.

Finally, we find an expression for the maximum value of the modularity. Indeed, it is our key objective to find an optimal way to partition our synthetic population into constituencies. We argue that this optimal way of partitioning is realised when the modularity reaches its maximum and the imposed partitions best capture the underlying community structure of the network. We derive then the location $D^*(N) = \arg\max_D \langle Q_D(N) \rangle$, in the regime $N \gg D$, from Eq. (23)

$$\frac{\partial \langle Q_D(N) \rangle}{\partial D} = \frac{1}{(N-1)^2}[(N-1)(\mu-1) + \frac{1}{D^{*2}}N(N + \mu(1-N) - 2) + \psi^{(1)}(D^*)] = 0, \qquad (24)$$

with $\psi^{(1)}(D)$ being the first order polygamma function, defined as the first derivative of the digamma function. This expression constitutes our main result, as it gives a recipe to pick the optimal number of constituencies, for a given population size N. The implicit Eq. (24) can be solved numerically for D^*. Interestingly, we find that $D^*(N)$ has a clear power-law behaviour similar to the one observed in demographic data. Figure 9 shows $D^*(N)$ for small networks, the numerical data being fitted by a power-law $D^* = \alpha N^\gamma$, in the interval $N \in [100, 1000]$, with exponent $\gamma \approx 0.53$ and $\alpha \approx 0.77$. In this range of N, the real value of D^* leads to the same integer constituency size for any values of μ. In this sense, μ does not significantly affect the behaviour of D^*. The value of the exponent can be also determined by the following analytical considerations in the more interesting large network limit. Expression (24) can be rewritten as follows

$$\alpha_N + \frac{1}{D^{*2}}\beta_N + \psi^{(1)}(D^*) = 0, \qquad (25)$$

with the coefficients $\alpha_N = (N-1)(\mu-1)$ and $\beta_N = N(N + \mu(1-N) - 2)$. By setting $\psi^{(1)}(D^*) = T$ and extracting D^* from Eq. (25)

$$T = \psi^{(1)}\left(\sqrt{\frac{-\beta_N}{\alpha_N + T}}\right). \qquad (26)$$

Now, since we are evaluating this quantities in the large network limit, we may use the polygamma asymptotic behaviour $\psi^{(1)}(x) \sim \frac{1}{x}$ in Eq. (26), and solve for T obtaining

$$T \approx \frac{1}{2}\left(-\frac{1}{\beta_N} + \frac{\sqrt{1 - 4\alpha_N\beta_N}}{\beta_N}\right). \qquad (27)$$

Inserting Eq. (27) in the original Eq. (25), we obtain the asymptotic optimal number of constituencies

$$D^* \approx \sqrt{\frac{-\beta_N}{\alpha_N + \frac{1}{2}\left(-\frac{1}{\beta_N} + \frac{\sqrt{1-4\alpha_N\beta_N}}{\beta_N}\right)}} \sim \sqrt{N}, \qquad (28)$$

which is consistent with the numerical solution of the implicit Eq. (24) shown in Fig. 9.

We conclude this section by noting that generalised definitions of modularity have been introduced in the literature[26,27,39], which allow to control the resolution of communities by tuning a control parameter. The purpose of such generalizations is to overcome the so-called resolution limit of the original definition of modularity, that is known to prefer merging two well-defined clusters into larger ones, when these are smaller (in terms of
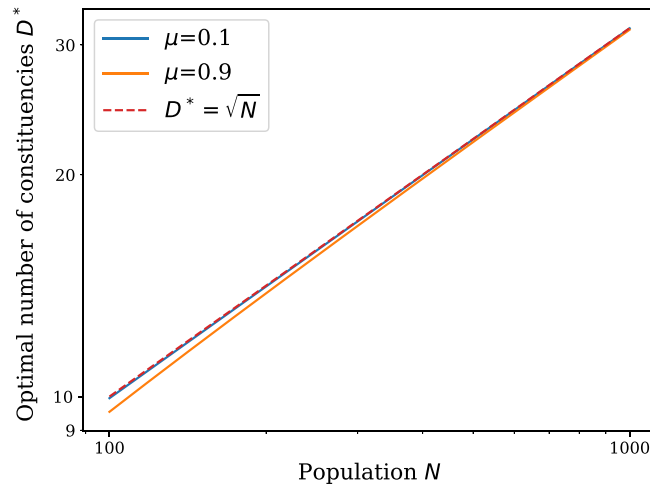
Figure 9. The log-log plot shows the maximum modularity $D^* = \alpha N^\gamma$, in the approximate regime, for $\mu = 0.9$ and $\mu = 0.1$ compared with $D^* = N^{1/2}$. A fit for the curve $\mu = 0.9$ results in $\alpha \approx 0.77$ and $\gamma \approx 0.53$. The fit was performed in the interval $N \in [100, 1000]$.

internal links) than a certain threshold. Denoting with η the resolution parameter and considering the generalised modularity defined in[27] (other definitions have been shown to be equivalent[39])

$$Q_D(N) = \frac{1}{L(N)} \sum_{s,r}^{N} \left[A_{r,s} - \eta \frac{k_r(N)k_s(N)}{L(N)} \right] \delta_{\sigma_r,\sigma_s}, \tag{29}$$

we can repeat the analysis carried out earlier to obtain the value of D that maximises the generalised modularity. In the limit of large network size, this is given by the η-dependent value

$$D^*(\eta) \sim \sqrt{N} \sqrt{\frac{\eta - \mu}{1 - \mu}}, \tag{30}$$

which correctly reproduces the result obtained in (28), for $\eta = 1$. Values of $\eta > 1$ shift the maximum of the average modularity to values larger than $D^* = \sqrt{N}$, whereas for $\eta < 1$ the maximum is shifted to smaller values than D^*, consistently with the interpretation of η as a fine-tuning parameter. Interestingly, (30) shows that there is a physical lower bound on the values that η is allowed to take, given by $\eta \geq \mu$. Notably, the introduction of the control parameter η does not affect the overarching behaviour in N of D^*, as long as $\eta = \mu + \mathcal{O}(N^0)$, showing robustness of our results at different resolutions. While there is no general recipe to choose η a priori, the choice $\eta = 1$, for which the modularity (29) weighs equally the network **A** under consideration and the null model, allows for a kinetic interpretation of the modularity as the autocovariance function of the cluster occupancy of a random walk on the network between two successive time steps[39], and it has the additional advantage of making the optimal number of constituencies, as given in (30), independent of the mobility μ. This intriguing independence may constitute an interesting subject for further investigations, as well as the interplay between mobility and resolution parameters in generalised definitions of network modularity.

Conclusion

The problem of democratic representation is of primary importance for modern societies. In this work, we proposed a network model representing a growing population of final size N that has to be partitioned into D equally sized constituencies. The underlying network community structure can be tuned by the mobility parameter μ that controls the interaction probability between nodes belonging to different constituencies.

We adopted the average modularity as a measure for the goodness of the resulting partitioning and showed that it displays a strong non-monotonic behaviour, as a function of D, in the regime $\mu \in [0, 1)$. By solving the recurrence equations for the modularity in the regime $N \gg D$, we found an analytical expression for the optimal number of constituencies D^* that maximises the modularity w.r.t. the number of induced partitions.

The approximate regime in which the problem is solved corresponds to one MP accounting for a large fraction of the population. This is arguably a reasonable assumption when considering democratically elected parliaments for which the condition above is always satisfied. Nevertheless, a numerical solution was also attainable for any value of N and D. Our main finding concerns the functional form for the optimal size of a Parliament $D^* \sim N^{1/2}$ that is found to be in reasonable agreement with what is observed in real-world data, thus providing further support for a revision of Tageepera's arguments in the direction of a "square root" law[17,18]. While a larger mobility parameter induces a more efficient mixing of the population and therefore reduces its average modularity for a fixed number of available constituencies (see Fig. 8), quite interestingly it does not influence the position of the maximum D^* as a function of N to leading order.

It is worth noting that, when considering a generalized definition of the modularity, which explicitly depends on a resolution parameter η, the position of the maximum retains the same dependence on the population size, i.e. $D^* \sim N^{1/2} f(\mu, \eta)$, but acquires a prefactor $f(\eta, \mu)$ that depends on the mobility μ for all values of the resolution parameter η except $\eta = 1$, for which the original definition of the modularity is retrieved. Hence, the intriguing independence of the position of the maximum on the mobility emerges as a peculiarity of the modularity (in its original definition). Given the lack of a general criterion to fix η a priori, explorations of the interplay between η and parameters that control preferential attachment in growth models, such as μ in our model, can provide an interesting pathway for future work.

As a further pathway for future work one could consider introducing geographical constraints in the model and including a mobility parameter that depends on the population density of each constituency. This is expected to generate a richer behaviour for D^*. Other, more complex, network topologies may also achieve a similar result. In future research, adopting different topologies, or studying small network scenarios, may broaden the applicability of our model to other absolute representation problems, from management (optimal proportion of managers vs. employees in a company) to determining the optimal number of parishes in a community.

Data availability

According to UK research councils' Common Principles on Data Policy, all codes used to produce and analyse numerical data supporting this study will be available upon request.

Appendix

In this section, we discuss the recursion for a_N and we present a more detailed version of the exact calculation for b_N.

a_N: We consider the term $\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} A_{r,N+1} \right\rangle$, appearing in Eq. (8). The expectation over the distribution of Eq. (5) is given by

$$
\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} A_{r,N+1} \right\rangle_{p(\mathbf{A}(N+1)|\boldsymbol{\sigma}(N+1))} = \left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} \sum_{A_{r,N+1}=\{0,1\}} A_{r,N+1} p(A_{r,N+1}|\mathbf{A}(N), \boldsymbol{\sigma}(N)) \right\rangle_{p(\mathbf{A}(N)|\boldsymbol{\sigma}(N))}
$$

$$
= \left\langle \sum_{r=1}^{N} \delta_{\sigma_r,\sigma_{N+1}} \frac{k_r(N)}{L(N)/D} p(\sigma_r|\sigma_{N+1}) \right\rangle_{p(\mathbf{A}(N)|\boldsymbol{\sigma}(N))}
$$

$$
= p(\sigma_{N+1}|\sigma_{N+1}) \frac{1}{L(N)/D} \left\langle \sum_{r=1}^{N} k_r(N) \delta_{\sigma_r,\sigma_{N+1}} \right\rangle_{p(\mathbf{A}(N)|\boldsymbol{\sigma}(N))},
$$

$$(31)$$

where we have used Eq. (1). Note that the term within angle brackets is zero for $N < D$, as in this case no node is present in the network with the relevant label. For N large, the average of the sum in (31) converges to $N \sum_k k p(k, \sigma_{N+1})$, in terms of the joint probability $p(k, \sigma_{N+1})$. This in turn factorises in the product of marginals due to the independence of k and σ, leading eventually to the result

$$
\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} A_{r,N+1} \right\rangle_{p(\mathbf{A}(N+1)|\boldsymbol{\sigma}(N+1))} \simeq p(\sigma_{N+1}|\sigma_{N+1}),
$$

$$(32)$$

where one uses $L(N) = N\langle k(N) \rangle$, and $p(\sigma_{N+1}) = 1/D$. For $N \geq D$ but not too large, fluctuations of $\mathcal{O}(1/N)$ are expected.

b_N: Recall the expression from Eq. (11)

$$
b_{N+1} = \underbrace{\left\langle \sum_{s,r}^{N} k_r(N+1) k_s(N+1) \delta_{\sigma_r,\sigma_s} \right\rangle}_{(i)} + 2 \underbrace{\left\langle \sum_{r}^{N} k_r(N+1) \delta_{\sigma_r,\sigma_{N+1}} \right\rangle}_{(ii)} + 1,
$$

$$(33)$$

where the factor 2 is due to the symmetry of the adjacency matrix. We analyse each average in Eq. (11), labelled (i) and (ii), separately:

(i) Starting with the case $N \geq D$, we have observed that, when $m = 1$, the degree of a given node can only increase by one at each time step and only if the newly created node connects to it, i.e. $k_r(N+1) = k_r(N) + A_{r,N+1}$. This leads to

$$\left\langle \sum_{s,r}^{N} k_r(N+1)k_s(N+1)\delta_{\sigma_r,\sigma_s} \right\rangle = \left\langle \sum_{s,r}^{N} (k_r(N) + A_{r,N+1})(k_s(N) + A_{s,N+1})\delta_{\sigma_r,\sigma_s} \right\rangle$$

$$= \left\langle \sum_{s,r}^{N} k_r(N)k_s(N)\delta_{\sigma_r,\sigma_s} \right\rangle + 2\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle$$

$$+ \left\langle \sum_{s,r}^{N} A_{r,N+1}A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle$$

$$= b_N + 2\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle + \left\langle \sum_{s,r}^{N} A_{r,N+1}A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle. \tag{34}$$

The first expectation may now be rewritten in the following way

$$\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle = \left\langle \sum_{s}^{N} A_{N+1,s} \sum_{r:\sigma_r=\sigma_s}^{N} k_r(N) \right\rangle$$

$$= \left\langle \sum_{s:\sigma_s=\sigma_{N+1}} A_{N+1,s} \sum_{r:\sigma_r=\sigma_{N+1}} k_r(N) \right\rangle + \left\langle \sum_{s:\sigma_s\neq\sigma_{N+1}} A_{N+1,s} \sum_{r:\sigma_r\neq\sigma_{N+1}} k_r(N) \right\rangle. \tag{35}$$

The $D-1$ constituencies such that $\sigma_r \neq \sigma_{N+1}$ are equivalent in the sum above as the probability of wiring $N+1 \sim r$ is μ/D for all, thus, considering the contribution to Eq. (35) from one constituency $\sigma_{\bar{s}} \neq \sigma_{N+1}$ we may rewrite

$$\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle = \left\langle \sum_{s:\sigma_s=\sigma_{N+1}} A_{N+1,s} \sum_{r:\sigma_r=\sigma_{N+1}} k_r(N) \right\rangle + (D-1)\left\langle \sum_{s:\sigma_s=\sigma_{\bar{s}}} A_{N+1,s} \sum_{r:\sigma_r=\sigma_{\bar{s}}} k_r(N) \right\rangle. \tag{36}$$

Now performing the expectation value of $\sum_{s:\sigma_s=\sigma_{N+1}} A_{N+1,s}$, using Eq. (9), we get

$$\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle = \left(1 - \mu\frac{D-1}{D}\right)\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} k_r(N) \right\rangle + \mu\frac{D-1}{D}\left\langle \sum_{r:\sigma_r=\sigma_{\bar{s}}} k_r(N) \right\rangle. \tag{37}$$

We are left with the term $\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} k_r(N) \right\rangle = C_{N+1}(N)$ to calculate, being the expected number of links in community σ_{N+1} at time N and its complementary $\left\langle \sum_{r:\sigma_r\neq\sigma_{N+1}} k_r(N) \right\rangle = \bar{C}_{N+1}(N)$. When the first round of communities has been assigned, i.e. $N \geq D$, the expression becomes

$$\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1s}\delta_{\sigma_r,\sigma_s} \right\rangle = \left(1 - \mu\frac{D-1}{D}\right)C_{N+1}(N) + \mu\frac{D-1}{D}\frac{\bar{C}_{N+1}(N)}{D-1}$$

$$= \left(1 - \mu\frac{D-1}{D}\right)C_{N+1}(N) + \frac{\mu}{D}(2(N-1) - C_{N+1}(N)) \tag{38}$$

$$= 2\frac{\mu}{D}(N-1) + (1-\mu)C_{N+1}(N),$$

where we have used $\bar{C}_{N+1}(N) + C_{N+1}(N) = 2(N-1)$. The expression for $C_{N+1}(N)$ is found to be

$$C_{N+1}(N) = \left[\sum_{x=\text{mod}(N+1,D)+1}^{D} \frac{1}{x-1} + \frac{\mu}{D}(\text{mod}(N+1,D)-1) \right](1 - \delta_{\text{mod}(N+1,D),0}) +$$

$$+ \mu\frac{D-1}{D}\delta_{\text{mod}(N+1,D),0} + 2\left\lfloor \frac{N}{D} \right\rfloor - 1, \tag{39}$$

where $\text{mod}(\cdot, D)$ is the modulus operator with divisor D and $\lfloor \cdot \rfloor$ denotes the floor operator. This builds up by summing the contribution to the expected links in community σ_{N+1} from every node. For the case $N \geq D$, a node j such that its membership $\sigma_j = \sigma_{N+1}$ always contributes with one link (its own stub) plus another link with probability $1 - \mu\frac{D-1}{D}$, in case the target node at the other end of the stub has the same membership σ_{N+1}. Conversely, if $\sigma_j \neq \sigma_{N+1}$, the new node j can only add one link to constituency σ_{N+1} with probability μ/D. In the first round of membership assignment, a contribution to the number of links can come from nodes generated following the first node in σ_{N+1}, with a probability $1/j$. The expression can be intuitively checked following Fig. 10. The remaining expectation of Eq. (34), is non-zero only if $s = r$, so

$$\left\langle \sum_{s,r}^{N} A_{r,N+1}A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle = \left\langle \sum_{s}^{N} A_{N+1,s} \right\rangle = 1. \tag{40}$$
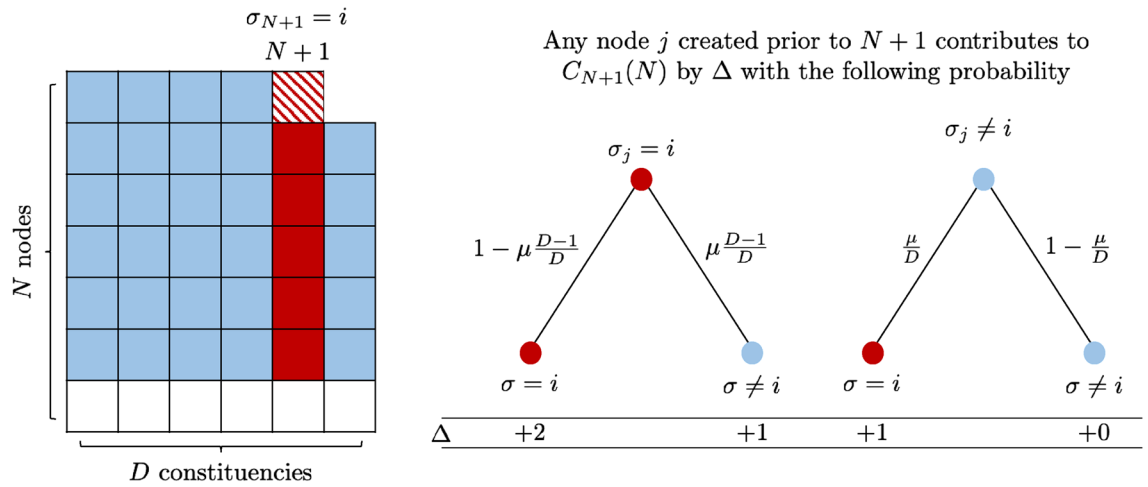
Figure 10. Illustration showing the possible increment for the number of links $C_{N+1}(N)$ in community $\sigma_{N+1} = i$ at time N with the corresponding probability, for all nodes $j \geq D$ in case of sequential assignment of constituency memberships.

We finally note that, in the case $N < D$, the expectations in Eq. (35) $\left\langle \sum_{s:\sigma_s=\sigma_{N+1}} A_{N+1,s} \right\rangle = 0$ and $\left\langle \sum_{s:\sigma_s\neq\sigma_{N+1}} A_{N+1,s} \right\rangle = 1$. This is due to constituency σ_{N+1} not being populated before time $N + 1$. As a result the average number of links in any given community is given by

$$\left\langle \sum_{s,r}^{N} k_r(N)A_{N+1,s}\delta_{\sigma_r,\sigma_s} \right\rangle = \left\langle \sum_{r:\sigma_r\neq\sigma_{N+1}} k_r(N) \right\rangle = 2\frac{N-1}{N}. \tag{41}$$

(ii) Now, this expectation value is similar to $C_{N+1}(N) = \left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} k_r(N) \right\rangle$. The difference lies in the time step at which this is being calculated. Thus it can be given a similar interpretation as the expected number of links in community $N + 1$, at time step $N + 1$, excluding the last node. This implies that if at time $N + 1$ the newly created node connects to one of its own community, the quantity in question increases by one, in a similar way to what is shown in Fig. 10. We note that, since the last node is excluded, its degree shall not be counted. So, the relation with $C_{N+1}(N)$ can be made explicit as follows

$$\left\langle \sum_{r:\sigma_r=\sigma_{N+1}}^{N} k_r(N+1) \right\rangle = C_{N+1}(N) + 1 - \mu\frac{D-1}{D}. \tag{42}$$

Back to Eq. (9), we finally get the recursive relation we are after for the coefficient b_N,

$$b_{N+1} = b_N + 4\mu\frac{N-1}{D} + 2(2-\mu)C_{N+1}(N) + 2\left(2 - \mu\frac{D-1}{D}\right). \tag{43}$$

References

1. Rush, M. *Parliament Today* (Manchester University Press, 2005).
2. Leston-Bandeira, C. Studying the relationship between parliament and citizens. *J. Legis. Stud.* **18**, 265–274. https://doi.org/10.1080/13572334.2012.706044 (2012).
3. Rogers, R. & Walters, R. *How Parliament Works* (Taylor & Francis, 2015).
4. Farrell, D. *Electoral Systems: A Comparative Introduction* (Palgrave Macmillan, 2011).
5. Taagepera, R. & Shugart, M. S. Designing electoral systems. *Elect. Stud.* **8**, 49–58. https://doi.org/10.1016/0261-3794(89)90021-8 (1989).
6. Jacobs, K. & Otjes, S. Explaining the size of assemblies. A longitudinal analysis of the design and reform of assembly sizes in democracies around the world. *Elect. Stud.* **40**, 280–292. https://doi.org/10.1016/j.electstud.2015.10.001 (2015).
7. Lijphart, A. Democracies: Forms, performance, and constitutional engineering. *Eur. J. Polit. Res.* **25**, 1–17. https://doi.org/10.1111/j.1475-6765.1994.tb01198.x (1994).
8. McLean, I. Don't let the lawyers do the math: Some problems of legislative districting in the UK and the USA. Mathematical Modeling of Voting Systems and Elections: Theory and Applications. *Math. Comput. Model.* **48**, 1446–1454. https://doi.org/10.1016/j.mcm.2008.05.025 (2008).
9. Penrose, L. S. The elementary statistics of majority voting. *J. R. Stat. Soc.* **109**, 53–57 (1946).
10. Slomczynski, W. & Życzkowski, K. Penrose voting system and optimal quota. *Acta Phys. Polon. B* **37**, 3133–3143 (2006).

11. Borghesi, C., Hernández, L., Louf, R. & Caparros, F. Universal size effects for populations in group-outcome decision-making problems. *Phys. Rev. E* **88**, 062813. https://doi.org/10.1103/PhysRevE.88.062813 (2013).
12. Taagepera, R. The size of national assemblies. *Soc. Sci. Res.* **1**, 385–401. https://doi.org/10.1016/0049-089X(72)90084-1 (1972).
13. Taagepera, R. & Shugart, M. S. *Seats and Votes: The Effects and Determinants of Electoral Systems* (Yale University Press, 1989).
14. Taagepera, R. & Recchia, S. P. The size of second chambers and European assemblies. *Eur. J. Polit. Res.* **41**, 165–185. https://doi.org/10.1111/1475-6765.00008 (2002).
15. Taagepera, R. *Predicting Party Sizes: The Logic of Simple Electoral Systems. Oxford scholarship online* (Oxford University Press, 2007).
16. Russett, B. M., Alker, H. R., Deutsch, K. W. & Laswell, H. *World Handbook of Political and Social Indicators* (Yale University Press, 1964).
17. Margaritondo , G. Size of National Assemblies: The Classic Derivation of the Cube-Root Law is Conceptually Flawed. *Front. Phys.* **8, 606.** https://doi.org/10.3389/fphy.2020.614596 **(2021).**
18. Auriol, E. & Gary-Bobo, R. J. On the optimal number of representatives. *Public Choice* **153**, 419-445. https://doi.org/10.1007/s11127-011-9801-3 (2012).
19. Newman, M., Barabási, A. L. & Watts, D. J. *The Structure and Dynamics of Networks. Princeton Studies in Complexity* (Princeton University Press, 2011).
20. Wasserman, S., Faust, K., Granovetter, M. & Iacobucci, D. *Social Network Analysis: Methods and Applications. Structural Analysis in the Social Sciences* (Cambridge University Press, 1994).
21. Caldarelli, G. *Scale-Free Networks: Complex Webs in Nature and Technology. Oxford Finance Series* (Oxford University Press, 2007).
22. Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512. https://doi.org/10.1126/science.286.5439.509 (1999).
23. Rossiter, D., Johnston, R. & Pattie, C. Representing people and representing places: Community, continuity and the current redistribution of parliamentary constituencies in the UK. *Parliam. Aff.* **66**, 856–886. https://doi.org/10.1093/pa/gss037 (2012).
24. White, H. C., Boorman, S. A. & Breiger, R. L. Social structure from multiple networks. I. Blockmodels of roles and positions. *Am. J. Sociol.* **81**, 730–780. https://doi.org/10.1086/226141 (1976).
25. Rehfeld, A. *The Concept of Constituency: Political Representation, Democratic Legitimacy, and Institutional Design* (Cambridge University Press, 2005).
26. Fortunato, S. & Barthélemy, M. Resolution limit in community detection. *Proc. Natl. Acad. Sci.* **104**, 36–41 (2007).
27. Reichardt, J. & Bornholdt, S. Statistical mechanics of community detection. *Phys. Rev. E* **74**, 016110 (2006).
28. Dorogovtsev, S. N. & Mendes, J. F. F. Evolution of networks. *Adv. Phys.* **51**, 1079–1187 (2002).
29. Tang, W., Guo, X. & Tang, F. The Buckley–Osthus model and the block preferential attachment model: Statistical analysis and application. *Proceedings of the 37th International Conference on Machine Learning*, Vol. 119, 9377–9386 (2020).
30. Hajek, B. & Sankagiri, S. Community recovery in a preferential attachment graph. *IEEE Trans. Inform. Theory* **65**, 6853–6874 (2019).
31. Holland, P. W., Laskey, K. B. & Leinhardt, S. Stochastic blockmodels: First steps. *Soc. Netw.* **5**, 109–137. https://doi.org/10.1016/0378-8733(83)90021-7 (1983).
32. Annibale, A., Coolen, A. C. C., Fernandes, L. P., Fraternali, F. & Kleinjung, J. Tailored graph ensembles as proxies or null models for real networks. I: Tools for quantifying structure. *J. Phys. A: Math. Theor.* **42**, 485001. https://doi.org/10.1088/1751-8113/42/48/485001 (2009).
33. Buckley, P. G. & Osthus, D. Popularity-based random graph models leading to a scale-free degree sequence. *Discrete Math.* **282**, 53–68 (2004).
34. Wacław, B. & Sokolov, I. M. Finite-size effects in Barabási–Albert growing networks. *Phys. Rev. E* **75**, 056114 (2007).
35. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113. https://doi.org/10.1103/PhysRevE.69.026113 (2004).
36. Newman, M. E. J. Modularity and community structure in networks. *Proc. Natl. Acad. Sci.* **103**, 8577–8582. https://doi.org/10.1073/pnas.0601602103 (2006).
37. Fortunato, S. Community detection in graphs. *Phys. Rep.* **486**, 75–174. https://doi.org/10.1016/j.physrep.2009.11.002 (2010).
38. Holmstrm, E., Bock, N. & Brännlund, J. Modularity density of network community divisions. *Phys. D: Nonlinear Phenom.* **238**, 1161–1167. https://doi.org/10.1016/j.physd.2009.03.015 (2009).
39. Lambiotte, R., Delvenne, J. & Barahona, M. Random walks, Markov processes and the multiscale modular organization of complex networks. *IEEE Trans. Netw. Sci. Eng.* **1**, 76–90. https://doi.org/10.1109/TNSE.2015.2391998 (2014).

Acknowledgements

Author contributions

L.G. designed the model, conducted numerical simulations, and wrote a draft of the manuscript. Y.-P.F., E.T., A.A., and P.V. analysed the results and contributed to drafting the manuscript. All authors reviewed the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.