The Emergence of Two-Party Systems

Practical Work

Ștefan Malanik
smalanik@ethz.ch

Distributed Computing Group
Computer Engineering and Networks Laboratory
ETH Zürich

Supervisors:
Andrei Constantinescu
Prof. Dr. Roger Wattenhofer

August 30, 2025

Abstract

The emergence of two-party systems has been studied extensively both from a mathematical and social perspective. This phenomenon is closely linked to the famous Duverger's Law, which states that plurality voting systems, such as first-past-the-post voting, tend to be dominated by two parties. This happens because voters are discouraged from supporting smaller parties that are unlikely to win, leading them to choose the "lesser of two evils" among the leading contenders.

Duverger's Law is shown to hold in specific cases. However, no such model sufficiently incorporates the behavior of society at large, where citizens live in a social web and have complex interaction patterns that influence their behavior and, in turn, their political opinion.

We propose a few voting system models that consider the social network as a key contributor to how voters behave, in isolation of strategic voting. We then test whether these models show tendencies of two main parties emerging.

Contents

CHAPTER 1

Introduction

1.1 General Definitions

We define a universe as a set of voters and a set of candidates, coupled with a voting rule. The state of the universe evolves over time in a round-based fashion, where opinions change based on the defined social structure of the model. The voting rule can be applied at any time to elect a candidate based on the state of the universe at that time.

In our analysis, we limit the dynamics of the system to interactions between voters, through which their political opinion is altered based on the employed model. As a consequence, the voting rule and candidate set remain static throughout the evolution of the system.

Moreover, the voting rule that we will focus on is plurality voting (also known as first-past-the-post), where each voter chooses exactly one candidate. The votes of all voters are then tallied up, and the candidate with the most votes wins. In case of equality, we associate a unique ID to each candidate and preferentially pick the one with the lowest ID first.

1.2 Previous Work

The problem of two-party systems and their link to plurality voting is a well-discussed topic in political science [1]. One of the phenomena that is frequently discussed regarding plurality voting systems refers to the prevalence of *strategic voting*. This voter behavior appears when, due to the circumstances imposed by the voting rule, the optimal choice is to sometimes opt for a candidate that does not match the true political preference of a voter.

In [2] and [3], it is shown that only specific dynamics converge to a Nash equilibrium (a voting decision that will never change from this point onward after the system continues to evolve). However, the proposed models will not directly include strategic voting, but instead focus on the influence of the social network of the voters.

A result from [4] proves Duverger's Law in a setting where the voter pool is simulated at random without strategic voting, and the candidate pool dynamically evolves in response to iterative "election" rounds. In our models, the candidate pool and the voting strategy is not of focus and will remain fixed, while the dynamics affects the voters and their opinion.

Most papers presented explicitly focus their analysis of converge and strategic voting behavior on models using plurality voting. Within the scope of plurality voting, a potential issue arises in relation to tie-breaking rules. In some papers ([5], [6]), ties of plurality voting are broken uniformly at random, while in others ([3]) ties are broken deterministically based on a fixed secondary ordering of candidates. In our case, we opt for the latter. In [2], tie-breaking is elevated to a first-order element of convergence analysis in the case of plurality voting, showing separate proofs or counterexamples for randomized and deterministic procedures.

In terms of interactions in a social network, [7] was an essential inspiration for the proposed stubbornness-charisma universe. Previous studies on opinion formation have introduced a variety of models to explain how individual beliefs evolve over time. Foundational approaches include the DeGroot [8] and Friedkin–Johnsen [9] models, which focus on eventual convergence of opinions and further modeling individuals by splitting their opinion in two: innate and expressed. The stubbornness-charisma universe combines the dynamics of the Friedkin-Johnsen model with aspects of Kantian morality ([10]), where voters eventually tend back towards their innate opinion.

Chapter 2

Models

2.1 Stubbornness-Charisma Universe

Definition

In a stubbornness-charisma universe, voters live in a social network represented as $G(V, E)$, and the candidates represent a static set C. Each candidate has an associated policy, defined by a mapping $pol : C \rightarrow P$ on a policy spectrum $P = [0, 1]$.

Voters have an internal opinion $op_{int}(v)$ and an external opinion $op_{ext}(v)$, both in P. Voting works as follows: Voters pick the candidate c closest to them in terms of distance $d(op_{ext}(v), pol(c))$ and vote for them. Then, normal plurality voting ensues and a candidate is selected as winner.

Initially, all voters are seeded with an initial opinion $op_{ext}(v) = op_{int}(v)$ picked at random from P. Moreover, stubbornness $st(v)$ and charisma $ch(v)$ weights are also randomly selected from an interval $\subseteq [0, 1]$. These values remain constant throughout the evolution of the universe.

Dynamics

The dynamics of this system happen after a round of elections, where two random neighbors v_1, v_2 (according to G) meet and interact such that one of their external opinion changes. The changed opinion is always $op_{ext}(v_1)$.

The charisma coefficient affects the impact of v_2's opinion on v_1, while the stubbornness coefficient measures how important v_1's internal opinion is to them as opposed to what their peers believe. The final computed external opinion is a weighted sum applied twice, as follows:

$$op'_{ext}(v_1) = op_{ext}(v1)(1 - ch(v_2)) + op_{ext}(v2)ch(v_2)$$
$$op''_{ext}(v_1) = op'_{ext}(v1)(1 - st(v_1)) + op_{ext}(v1)st(v_1)$$

This approach of computation of expressed opinion is similar to the Friedkin–Johnsen model [9] in terms of updating nodes and inspired by [10] in terms of the weighted sum.

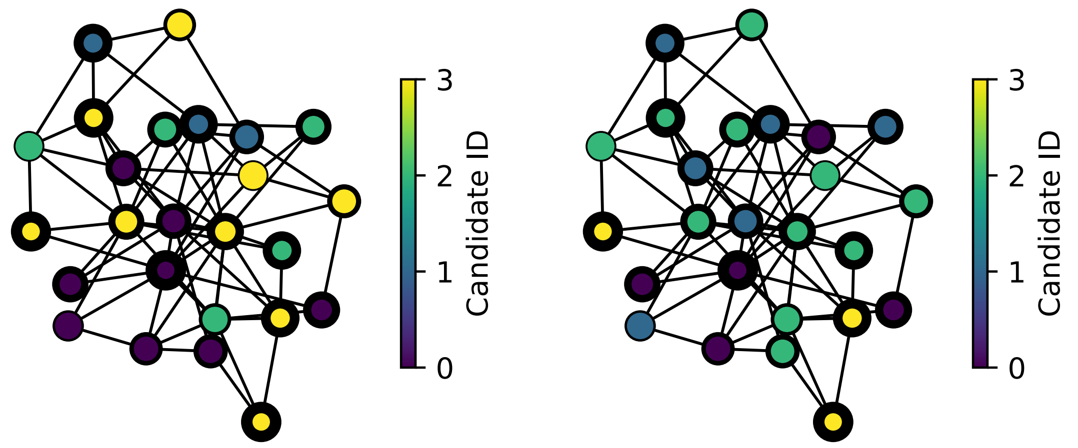


Figure 2.1: Voter graph of 25 nodes before and after 100 rounds of evolution.
Nodes are colored based on the candidate they voted for
(out of 4 candidates uniformly distributed in P).
Node border thickness indicates stubbornness, sampled from the interval [0,0.5].

The argument presented in the latter paper regarding Kantian morality closely matches the second adjustment of the expressed opinion based on stubbornness. In this case, high stubbornness is similar to the morality factor, since in the edge case where $st(v_1) = 1$, the voter is completely truthful and does not consider other opinions.

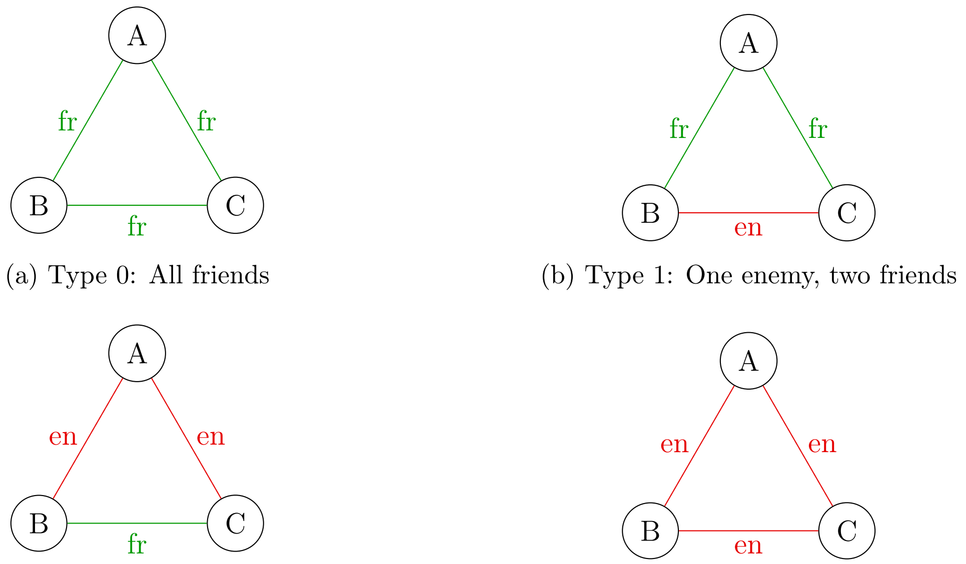
An example of how the dynamics of the system behave can be seen in figure 2.1. Observe how nodes that have high charisma are very likely to keep their initial opinion (resulting in voting for the same candidate).

2.2 Binary Triadic Closure Universe

Definition

In a triadic closure universe, a pair of voters form relations belonging to a discrete set $\mathcal{R}$. These relations evolve over time, potentially reaching a stable state where no further evolution is possible. This network of voters can be naturally represented with a graph $G(V, E)$ where nodes are voters and edges represent relations between them. $\text{rel} : V^2 \to \mathcal{R}$ is the function returning the relation in between two voters in G. Furthermore, G is always complete, with the possibility of having a "missing" edge by associating it with a special value $\mathcal{R}$ (e.g. $\perp$).

The dynamics of this system are defined as *triadic* transformations happening on a tuple of 3 voters (called a *triangle*), where one of the relations is potentially modified



(a) Type 0: All friends

(b) Type 1: One enemy, two friends

(c) Type 2: One friend, two enemies

(d) Type 3: All enemies

Figure 2.2: The four possible triangles in a binary triadic closure

based on the other two relations. This happens in discrete rounds, where only one such transformation can occur. The triangle is chosen uniformly at random from the list of all tuples of 3 distinct voters.

Theorem 2.1 demonstrates that, under additional constraints, the requirement of completeness for the graph G can be asymptotically satisfied through iterative evolution of the network, starting from a G that is not complete.

Binary Restriction

We constrain the model further by only allowing exactly two relations between each pair of voters: **friend** relations (fr) and **enemy** relations(en). That is to say, $\mathcal{R} = \{fr, en\}$. Furthermore, we can define the friend graph F as a spanning subgraph of G where we keep all nodes but only the friend edges.

Observe that under these restrictions, there are 4 possible triangle configurations (ignoring rotational symmetries), illustrated in figure 2.2.

Theorem 2.1. *A triadic closure universe that also permits an absence of a relation (edge) in addition to friend/enemy relations between voters evolves to a complete network graph, given that the initial network graph is connected.*

Proof. Given any triadic transformation, a triangle formed from 3 voters u, w, v, the edge uv can either be created, modified, or remain unchanged (as seen from the binary triadic closure transformations in figure 2.3). Note that the relation uv can never be removed once it exists.

Since the network graph is connected, $\forall$ nodes u, v that are at distance $d(u, v) = d$ apart, $\exists$ shortest path P . Consider only nodes u, v s.t. $d \geq 2$. Let w, t be the nodes following u on path P (note that t may be identical to v if $d = 2$). Assume that triangle uwt is sampled for a transformation in a given round. Since edges $uw, wt \in R$ define exactly one transformation, after the transformation is applied, edge ut must exist. Thus, the new distance between u and v is $d - 1$.

Eventually, under u.a.r. sampling, there must exist no pair of nodes u, v s.t. $d(u, v) \geq 2$. As a result, the graph will be complete with edges from R. □

The proposed dynamics look at any combination of two relations in a triangle to produce the remaining one. A systematic approach of triadic closures and the transformation rules has been explored in social psychology in the past in [11]. In our case, the intuition behind these rules can be summarized to the phrase "the enemy of my enemy is my friend". The 3 possible transformations are illustrated in figure 2.3.

Observe that for some triangles (Type 0 and 2 in figure 2.2), no transformation can change any of the relations between edges. We call these triangles *stable*. For the other triangles (named *unstable*), transformations will alter the triangle such that it becomes *stable*:

- Type 1 - can either turn into Type 0 or 2
- Type 3 - always turns into Type 2

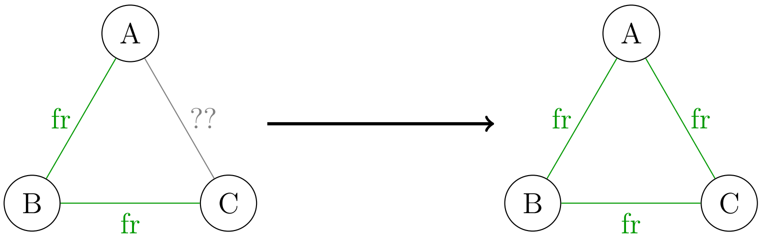
In addition, note that a relation flip (swapping fr with en) between any 2 of the 3 voters corresponds to exactly one transformation in both unstable triangle configurations.

Theorem 2.2. *Any binary triadic closure network eventually converges into a state where the underlying friend graph is a union of two complementary cliques.*

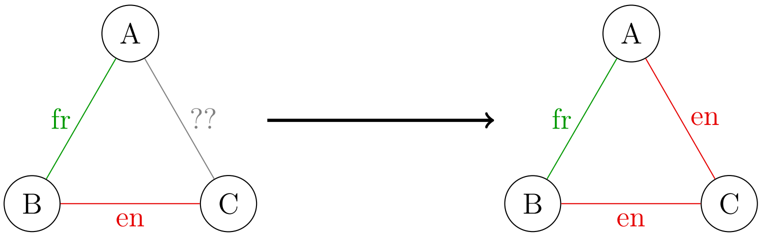
Proof. First, we prove that a friend graph representing a union of two cliques is stable. Then, we show that any initial friend graph can evolve into a union of two cliques after a finite number of transformations are applied. Finally, we claim that eventually all graphs stabilize on a fixed point.

A graph is stable if there is no transformation that can be applied which changes a relation between two nodes. This is also equivalent with the graph being composed of stable triangles (by definition). If the friend graph is a union of two cliques (see example in figure 2.4, there are only two possible types of triangles that can exist:

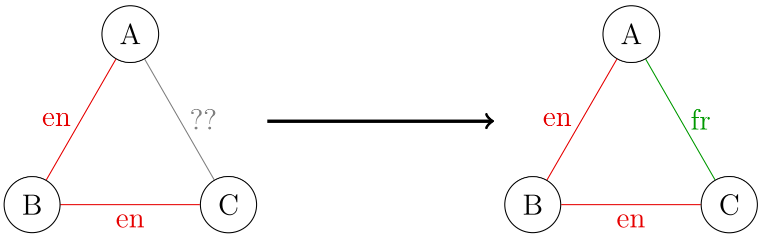
- Type 0: between 3 nodes in the same clique (e.g. B1-B2-B3)
- Type 2: between 2 nodes in one clique and 1 node in the other (e.g. A1-A2-B2)



(a) FFF: Friend of my friend is my friend



(b) FEE: Friend of my enemy is my enemy



(c) EEF: Enemy of my enemy is my friend

Figure 2.3: Possible triangle transformations

No other triangle type is possible since a type 1 implies that one of the cliques is not well-defined (at least one edge is missing in the friend graph), while a type 3 implies the existence of at least 3 cliques.

To produce a union of two cliques from an arbitrary friend graph, we can find a series of transformations from figure 2.3 in the following way:

1. *Remove type 3 triangles*: Select a type 3 triangle and apply transformation EEF, which guarantees that a friend relation gets introduced, so no new type 3 triangles can be created after each individual transformation. This is repeated until no further type 3 triangles are left. In the end, the friend graph will contain at most 2 connected components (otherwise a type 3 triangle must exist).

2. *Separate*: In case the friend graph is connected, we find an enemy edge and use transformation FEE to propagate enemy edges to adjacent relations. This can be

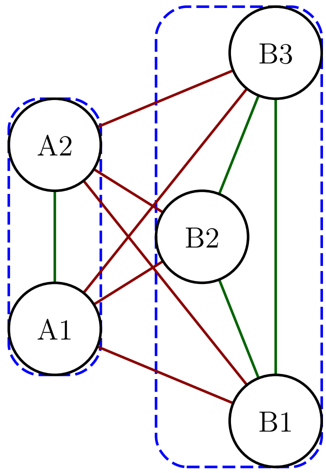


Figure 2.4: Union of two cliques in a stable configuration.
Red represents **enemy** relations, green represents **friend** relations.

used to disconnect the graph into exactly 2 components.

3. *Create cliques*: In each of the 2 components, we use transformation FFF on 3 nodes A-B-C that are on a path to potentially add edge A-C. This can be repeated until each component becomes a clique.

Note that step 2 might fail in case there are no enemy edges (friend graph is already a clique). We consider this to be a **degenerate** case when one of the two cliques is void, since an all-friend graph is already stable (containing only type 0 triangles).

Finally, considering that any graph can reach stability in a finite number of transformations, there is a non-zero probability of such a sequence of transformations to be chosen in a randomized universe. We can treat these transformations as a Markov chain, where states represent the structure of the graph in between transformations. Since a stable graph will not be further altered by transformations, this represents an absorbing state in a Markov chain. Thus, any graph will eventually reach such an absorbing state. □

Sampling

The aim of our universe is fast convergence to a stable voter graph from any initial state. Applying transformations on u.a.r. sampled triangles leads to an exponential expected number of steps to convergence, since the number of states in a Markov chain as described in Theorem 2.2 is $2^{\binom{n}{2}}$, where n is the number of voters.

In every round, 3 voters (u, v, w) are selected and one of the possible transformations is applied. Since only unstable triangle configurations will produce modifications in the graph after a transformation is applied, we use this restricted sampling pool.

The round consists of an evaluation step, followed by deciding whether to initiate a relationship flip (equivalent to a transformation in an unstable triangle) for uv. This will always be possible, since any relation in an unstable triangle can be flipped.

The evaluation step works by analyzing the entourage of u and their opinion about v. This is done by counting the number of friends of u that have the same relation to v as u does.

Let:

$$\mathrm{Fr}(u) := \{t \in V \mid t \neq u \text{ and } e(u,t) = \text{'fr'}\}$$
$$\mathrm{agree}(t,u,v) := \begin{cases} 1 & \text{if } e(t,v) = e(u,v) \\ 0 & \text{otherwise} \end{cases}$$

We then define the probability of flipping uv as:

$$p_{flip} = \frac{\sum_{t \in \mathrm{Fr}(u)} \mathrm{agree}(t,u,v) + 0.5}{|\,\mathrm{Fr}(u)| + 1}$$

The $0.5/1$ constant factors represent the base opinion of the voter u, emulating the initial impartiality. In isolation of an entourage, the relation flip will be dictated by a fair coin toss. Moreover, due to these constant values, p_{flip} can never be exactly 0 or 1, so there is always some amount of randomness involved in the decision of flipping a relation.

This probability might be further altered to reflect a certain polarization of voters by using a function that pushes away from the middle ($p_{flip} = 0.5$ - perfectly undecided voter):

$$\sigma_k(p) = \frac{x^k}{x^k + (1-x)^k}$$

Voting Rule

Note that the proposed model is not a well-defined universe, since it lacks the actual voting rule. This is the essential component in assessing whether a universe converges to a two-party state.

To define a voting rule, consider a fixed set of candidates $\mathcal{C}$. Each voter has a preference for which she votes. We then use simple plurality voting.

We claim that a friend relation between two voters is heavily correlated with their voting preferences being identical. To simplify, we strengthen this claim and arrive at the following assumptions:

(A1) Two friends vote for the exact same candidate.

(A2) Two enemies never vote for the same candidate.

Let F be the friend subgraph of G. According to Theorem 2.2, after stabilization, we have:

$$\exists K_1, K_2 - \text{cliques} \mid K_1 \cap K_2 = \emptyset, K_1 \cup K_2 = F \tag{2.1}$$

Since all voters in K_1 are friends with each other, from (A1) it follows that $\forall v_1 \in K_1, v_1$ votes for $c_1 \in \mathcal{C}$. Similarly, $\forall v_2 \in K_2, v_2$ votes for $c_2 \in \mathcal{C}$.

In addition, since any edge between nodes in K_1 and K_2 is part of $\overline{F}$, it represents an enemy relation. From (A2) it follows that $c_1 \neq c_2$, provided $F \neq G$.

Under these assumptions, we can conclude that any initial state will eventually stabilize on a configuration where k voters vote for c_1, and $n - k$ voters vote for c_2, where $k = |K_1|$. One of these two candidates wins (or there is a tie), but any other candidate $\mathcal{C} \setminus \{c_1, c_2\}$ will get exactly 0 votes, so this functionally becomes a contest between the two candidates.

In the case that the friend subgraph stabilizes on $F = G$, we are in the situation where instead of two candidates competing, every voter agrees on a single candidate that automatically wins the election. This corresponds to the degenerate case mentioned in Theorem 2.2, and ideally, the frequency of its appearance should be minimized.

Chapter 3

Results

This chapter discusses the simulation of the models and presents the results with respect to different parameter choices. The models are implemented in Python[1].

3.1 Stubbornness-Charisma Universe

The model defining the stubbornness-charisma weights has been implemented and simulated with a range of parameters. The state of the graph never converges (voters can always change their opinion after a certain round) for values of charisma $ch(v) \in (0, 1]$ and stubbornness $st(v) \in (0, 1)$, since as long as there are two neighbors with different opinions, an interaction between them will cause one of them to change their opinion. This in turn means that a convergence to a two-party system is never guaranteed.

Based on experimental runs with different values for stubbornness and charisma (see figure 3.1), after many rounds a general pattern appears. As expected, when charisma is 0, no opinion update occurs, since neighbors never affect each other when they meet. When the stubbornness is 0, there is no bounce back to the original internal opinion, which causes the top row to reach consensus on the average preference over time. For low values of stubbornness (0.05) and higher charisma (0.3), it seems that the system tends towards the two centrist candidates out of the 4 initial options, but this is also susceptible to randomness.

3.2 Binary Triadic Closure Universe

In the case of binary triadic closures, convergence to two parties is guaranteed, as shown in Theorem 2.2. However, two aspects of the model remain of particular interest: the time required for convergence and the quality of the resulting split. The latter concerns the final distribution of nodes between the two cliques–ideally balanced in size, or at the very least, very unlikely to result in a degenerate graph.

[1] `https://github.com/stefanmalanik/two-party-emergence`

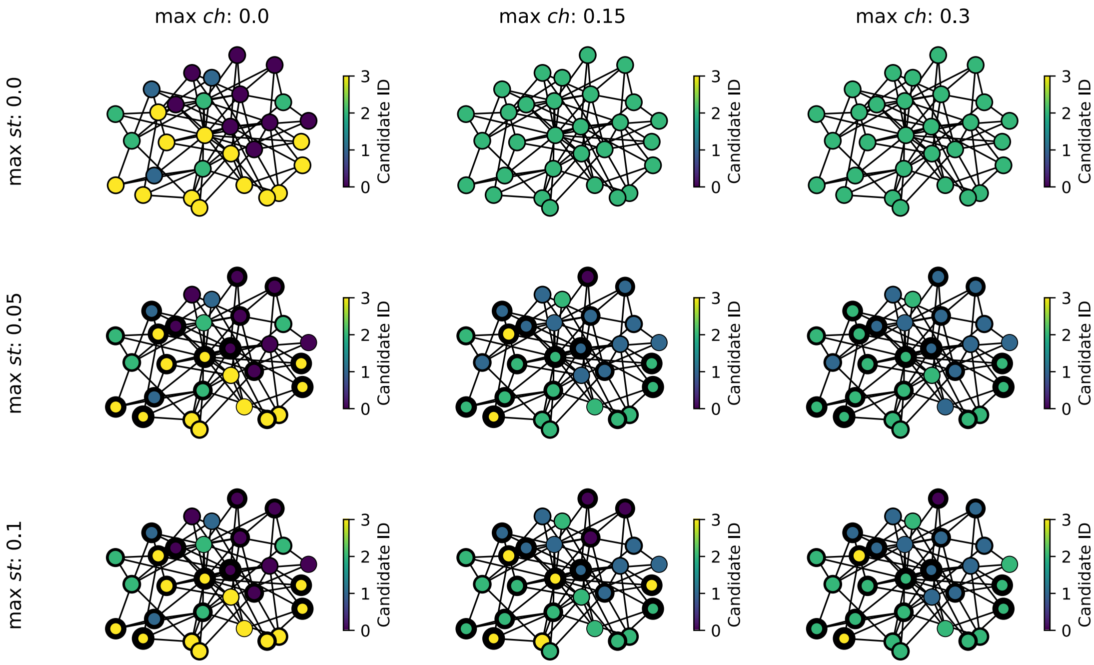


Figure 3.1: Voting results with different coefficients after 1000 rounds of evolution. Top-left graph coincides with the initial state.

The simulations were run on Erdős–Rényi graphs with a probability of a friend edge $p_f = 50\%$. The run statistics can be seen in figures 3.2 and 3.3. It can be observed that both $n = 25$ and $n = 30$ graphs converge to a degenerate graph with very high probability. Moreover, the probability seems to increase with the number of nodes, which indicates that the model in its current state does not achieve the desired properties.

With respect to convergence rate, an increase of 20% in the number of voters led to approximately a $10\times$ increase in the number of rounds necessary to converge. This number is much larger than what we would expect, knowing that the worst-case performance of the procedure outlined in Theorem 2.2 takes $O(n^2)$. This is because all edges need to be flipped at most once to reach a stable configuration.

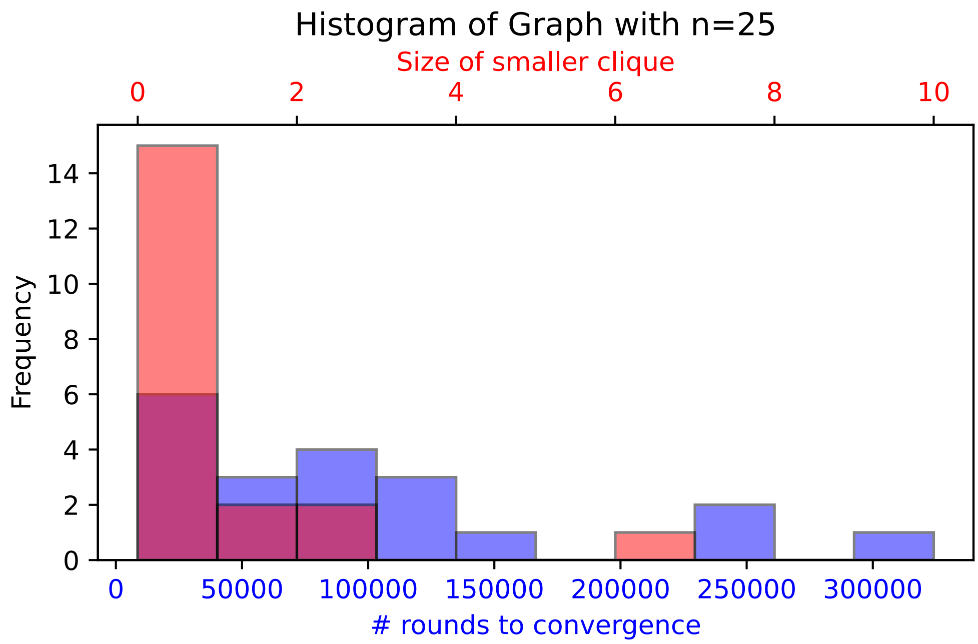


Figure 3.2: Statistics on 20 runs of random graphs with 25 voters evolving until convergence.

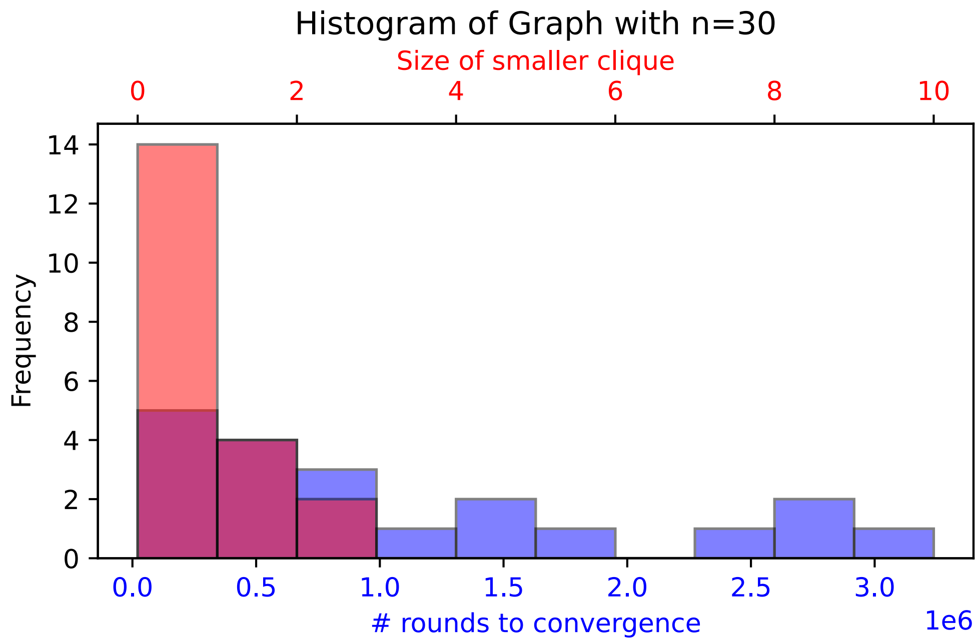


Figure 3.3: Statistics on 20 runs of random graphs with 30 voters evolving until convergence.

Chapter 4

Conclusions and Future Directions

The models presented attempt to prove that a two-party system is an emergent behavior of plurality voting in the context of a simulated social dynamic between voters. While both models are conceptually promising, neither one produces a two-party solution in a reliable fashion.

The *stubbornness-charisma universe* has specific configurations for which two candidates reliably get elected. However, convergence is never guaranteed, and in most cases impossible. A possible avenue for improvement involves incorporating strategic voting into the model's dynamics. This would potentially create a feedback loop that stabilizes the graph if properly modeled.

In contrast, the *binary triadic closure universe* does guarantee convergence, but the quality of that convergence is often poor, considering the high distribution of degenerate one-party results. In addition, the convergence happens significantly more slowly than expected. These limitations could be addressed by further refinement of the model:

1. To prevent degenerate graphs, the model could introduce a bias favoring enemy edges. Encouraging a more balanced distribution between friend and enemy edges tends to lead to cliques (i.e., voter pools) of roughly equal size, which matches intuitive expectations of a two-party system.

2. To accelerate convergence, the edge selection process could prioritize edges that are part of many unstable triangles. If the total number of unstable triangles is used as a monotonic metric that decreases with each transformation, the number of edge flips required to reach a stable state would be bounded by $O(n^3)$, a significant improvement over current behavior.

Bibliography

[1] W. H. Riker, "The Two-party System and Duverger's Law: An Essay on the History of Political Science," *American Political Science Review*, vol. 76, no. 4, pp. 753–766, Dec. 1982.

[2] R. Meir, M. Polukarov, J. Rosenschein, and N. Jennings, "Convergence to Equilibria in Plurality Voting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 24, no. 1, pp. 823–828, Jul. 2010.

[3] G. B. Zvi, E. Leizerovich, and N. Talmon, "Iterative Deliberation via Metric Aggregation," Aug. 2021.

[4] M. Fey, "Duverger's law without strategic voting," *Working Paper*, 2007.

[5] G. Sieg and C. Schulz, "Evolutionary dynamics in the voting game," *Public Choice*, vol. 85, no. 1-2, pp. 157–172, Oct. 1995.

[6] G. W. Cox, "Electoral Equilibrium under Alternative Voting Institutions," *American Journal of Political Science*, vol. 31, no. 1, p. 82, Feb. 1987.

[7] A. Gionis, S. Neumann, and B. Ordozgoiti, "Opinion Formation in Social Networks: Models and Computational Problems," Web Conf 2022, Apr. 2022.

[8] M. H. Degroot, "Reaching a Consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, Mar. 1974.

[9] N. E. Friedkin and E. C. Johnsen, "Social influence and opinions," *The Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193–206, Jan. 1990.

[10] I. Alger and J.-F. Laslier, "Homo moralis goes to the voting booth: Coordination and information aggregation," *Journal of Theoretical Politics*, vol. 34, no. 2, pp. 280–312, Apr. 2022.

[11] F. Heider, "Attitudes and Cognitive Organization," *The Journal of Psychology*, vol. 21, no. 1, pp. 107–112, Jan. 1946.