# Iterative Calculus of Voting Under Plurality

**Fabricio Vasselai**
University of Michigan
5700 Haven Hall - 505 South State Street, Ann Arbor, MI 48109
vasselai@umich.edu

### Abstract

We formalize a voting model for plurality elections that combines Iterative Voting and Calculus of Voting. Each iteration, autonomous agents simultaneously maximize the utility they expect from candidates. Agents are aware of neither other individuals' preferences or choices, nor of the distribution that preferences come from. They know only of candidates' current expected vote shares (e.g. from a poll) and with that calculate expected rewards from each candidate, pondering the probability that voting for each would alter the final outcome. We define the general form of those pivotal probabilities, then we derive efficient exact and approximated calculations Lastly, we prove formally the model converges with asymptotically large electorates and show via simulations that it nearly always converges even with very few agents.

## 1 Introduction

In many multi-agent systems, agents with diverse preferences have to settle on a single choice or course of action - a challenge usually solved by voting. Yet, it is known from the Gibbard-Satterthwaite theorem that if there are more than two alternatives and no one can dictate the result, intelligent agents' best response may differ from their sincere preference (Gibbard 1973; Satterthwaite 1975).

While, traditionally, the Computational Social Choice literature has treated such a misrepresentation of preferences as a 'manipulation' to be prevented or at least diminished (see Bartholdi, Tovey, and Trick 1989; Conitzer, Sandholdm, and Lang 2007), in the last decade a new approach has emerged. Elkind et al. (2010) posed, instead, that such a strategic voting is "an unavoidable attribute of an electoral system with rational voters" (p.347). Meir et. al.'s (2010) Iterative Voting (IV) allowed agents to sequentially (and deterministically) re-evaluate their choices over multiple iterations, after learning only the most current election score.

Analytical Game-Theory has embraced strategic voting for decades, but their work focus on exactly the opposite. They usually investigate solution concepts of probabilistic non-atomic games where individuals know the distribution of voter types and decide their optimal vote simultaneously in a single one-shot iteration (for a review, see Meir 2018). In its most renowned approach, the 'Calculus of Voting' (CV) model inaugurated by Riker and Ordeshook (1968), each agent maximizes its expected reward via weighting candidates' cardinal utility differences by the probability that voting for a candidate would be pivotal to determine the election result (see Palfrey 1988; Myerson and Weber 1993).

Here we formally derive the Iterative Calculus of Voting (ICV), a voting model that joins the main aspects of IV and CV. Agents re-evaluate their choices over multiple iterations, like in IV, but each time doing a simultaneous expected reward optimization, like in CV. This is naturally akin to existing applications based on simulating CV in discrete-time (e.g. Clough 2007; Tsang and Larson 2016; Fairstein et al. 2019). But ICV encompasses those and offers a rigorously formalized and more general framework. Additionally, ICV deals with how to transition CV assumptions to an iterated setting, it is suitable for both small and large electorates and it jointly addresses limitations of both CV and IV.

More specifically, ICV can be described as follows. In each iteration, agents learn only the current expected vote shares of all available alternatives and use those to estimate the probability that voting for each alternative would alter the final outcome. Agents use those pivotal probabilities to weight the utilities they see in each alternative and then update their choices. Therefore, similarly to IV and CV, agents are unaware of each other's preferences or individual choices. However, like in IV and differently from most CV, agents are even unaware of the distribution from which preferences came. And like in CV but differently from most IV, agents update simultaneously instead of sequentially in electorates that can be either small or arbitrarily large. All in all, ICV agents require low information and are boundedly rational in the strategic sense.[1]

After deriving several generalizations and novel pivotal probability calculations - exact and heuristics - convergence properties of ICV under plurality are studied in two ways. In non-atomic ICV games with pivotal probabilities of standard characteristics, we prove that convergence always results, as well as the bounds on convergence rates. Also, we show that similarly to CV, under weak assumptions ICV usually converges to only two alternatives receiving positive vote

[1] Like in CV, they must be computationally sophisticated enough to estimate their probabilities of being pivotal. As will be discussed, conceptually principled heuristics can be developed.

counts, which is known as Duvergerian equilibria (Palfrey 1988; Meir 2018). In atomic ICV games, while the existence of cycles cannot be theoretically excluded, we argue why, under reasonable specifications, they will be rare. We also show through simulations that empirically, convergence is nearly always achieved even with the smallest electorates.[2]

## 2 Related work

Meir et. al.'s (2010) inaugural IV analyzes convergence to a Nash Equilibrium under plurality voting when agents are allowed to iteratively re-evaluate their choices in light of the current election outcome. Working with ordinal candidate utilities, they show that convergence is only guaranteed if agents update sequentially and under linear ordered tie-breaking rules. Similar convergence properties were shown to apply to veto rule (Lev and Rosenschein 2012), but no convergence is guaranteed under any other scoring rule (Lev and Rosenschein 2016), nor under Single-Transferable Vote (Koolyk et al. 2017). On the other hand, it was shown that even if IV agents operate under uncertainty (Meir, Lev, and Rosenschein 2014) and are boundedly rationally using heuristics (Meir 2015), convergence under plurality holds.

Furthermore, nearly all work on IV focus on atomic games. Meir et.al. (2010) already pointed that their original model was particularly suited for small electorates and that "an analysis in the spirit of Myerson and Weber (1993)" - i.e. of the CV - "would be more suitable when the number of voters increases" (p.828). As a consequence, the subsequent IV literature also focused mostly on sequential agent updates, since it is rarely possible to guarantee that atomic games with simultaneous updating will be free of cycles (Fabrikant, Jaggard, and Schapira 2013). An exception is Meir (2015), where a deterministic model is proposed that allows simultaneous updates in large electorates. The author claims his model is like Myerson and Weber's (1993) plurality CV, without probabilities (and using ordinal utilities) - yet it could not be immediately generalized beyond plurality.

CV has been extensively generalized and is defined for non-atomic electorates. However, due to its one-shot nature, its equilibrium conditions actually rely entirely on asymptotically large electorates and on agents knowing the distribution of preference orderings. While agents are unaware of other individuals' preferences or choices just like in IV, Palfrey (1988) showed it is because they know the distribution of voter types that, simultaneously and at once, they choose an optimal strategic vote under plurality that leads to Duvergerian equilibria. Or, more generally, always to *an* equilibrium - under plurality, approval or Borda (Myerson and Weber 1993), SNTV (Cox 1994), Proportional Representation (Cox and Shugart 1996) or Runoff (Bouton 2013). Of course, that information requirement is as unrealistic in human elections (Myatt 2007) as it is limiting in practical AI applications. But without it, Meir (2018, p.86)'s assertion about CV becomes irreproachable: "it is not clear how [agents] are supposed to reach [equilibrium] in a game that is only played once without some means for coordination".

Despite not satisfying those assumptions, applications implementing CV as computational simulations exist. Clough (2007) uses a simplified simulation of CV to study the effects of information uncertainty on voter's strategic choices, while Tsang and Larson (2016) expand that to agents connected through networks. Fairstein et al. (2019) also employ a similar CV simulation, alongside others, to predict voting behavior observed in human online voting experiments. Being discrete-time simulations, they are allowed to have multiple time iterations, which approximates them to ICV. However, largely non-formalized, those applications do not explore the iterative aspect of vote re-evaluation and do not engage in much of the theoretical details required to bridge IV and CV. Moreover, they consider only the pivotal probability of *breaking* ties (not the probability of *making* ties) for first and disregard the possibility of multi-way ties - both approaches that only make sense with asymptotic electorates.

ICV rigorously formalizes a general iterative CV model, thus encompassing those applications, and with a definition that includes all probabilities and considers multi-way ties, so any electorate size is covered. Besides, ICV is defined such that the only information agents access are candidates' expected vote shares at each iteration, like in IV. For simplicity, here those are assumed to be unbiased public information (e.g. coming from polls Cox 1994; Fey 1997). Hence, differently from CV, in ICV eventual convergence is constructed by strategic choices made over multiple iterations. It neither relies on distributional assumptions nor requires agents aware of the distribution preferences come from.

Meir, Lev, and Rosenschein (2014) claim that another issue with CV models - one that would be inherited by ICV - lays in assuming agents capable of optimizing complex expected utility functions. We do not see that as a problem for general AI applications, where human-realism is not always the goal. Nonetheless, boundedly rational agents in the computational sense may be desired for computational efficiency (Zilberstein 2008) or for purely theoretical reasons (Conlisk 1996). As we will show, in ICV the comparison of utilities by agents amounts to a simple sum of weighted subtractions; all challenge is in calculating pivotal probabilities. This is why, since Black (1978), many heuristics have been proposed for CV pivotal probabilities of breaking two-way ties (Hoffman 1982; Cranor 1996; Myerson 1998; Smith, de Mesquita, and LaGatta 2016). Recently, Tsang, Salehi-Abari, and Larson (2018) proposed sampling simplifications to Palfrey (1988)'s Multinomial pivotal probabilities.

In this regard, we first define pivotal probabilities under plurality generically - a theoretical framework from which principled heuristics can be derived. Next, we derive exact calculations for when electorate size is known (generalizing Palfrey 1988) and for when it is not - thus generalizing Myerson (1998) to multi-candidate plurality, which leads to the novel Poisson pivotal probabilities. Then, we prove the Skellam pivotal probability approximation (generalizing a path explored by Smith, de Mesquita, and LaGatta 2016; Tsang, Salehi-Abari, and Larson 2018; Anonymous 2018). With key pivotal probabilities properly defined for ICV, our idea is to offer a benchmark through which heuristics may be proposed and tested.

[2] Online Appendix and all code necessary to replicate the paper can be found at https://github.com/vasselai/aaai22-icv-plurality.

# 3 Model derivation

Recapitulating, in ICV, in each iteration agents simultaneously maximize the rewards they expect from the election, with respect to their current vote choice. Knowing only of candidates' current expected vote shares, agents ponder the probability that voting for each candidate would alter the final outcome and use those probabilities as weights to calculate the expected rewards. In this section we formalize that.

For consistency, whenever possible we follow the notation from the game theoretical CV literature. Like in IV, however, ICV games have multiple iterations, here indexed by $\delta \in \mathbb{N}_{\geq 0}$, where $\delta = 0$ is the initial condition. Let $i$ be the focal agent (hereafter, elector) on whom all derivations will focus, without loss of generality, and $N \in \mathbb{N}_{>1}$ be the number of electors. Their voting choices (hereafter, candidates) are represented by the set $\mathcal{J}$, s.t. $m = |\mathcal{J}| \in \mathbb{N}_{\geq 3}$ is the number of candidates. Then, let vector $\mathbf{n}^{*,\delta} \in \mathbb{N}^m_{\geq 0}$ hold true candidates' expected votes counts in iteration $\delta$ and vector $\mathbf{n}^{i,\delta} \in \mathbb{N}^m_{\geq 0}$ hold elector $i$'s guesses (hence random variables) about candidates' expected vote counts minus $i$'s vote in $\delta$ - we discuss, later, how those guesses are updated. Finally, vector $\mathbf{u}^i \in \mathbb{Q}^m_{[0,1]}$ holds normalized cardinal utilities that $i$ would accrue in case each candidate won.[3] The final reward $i$ gets from the election is represented by $\mathbf{u}^i_w$, where $w$ is the election winner - with ties resolved randomly:

**Assumption 1.** *$w \in \mathcal{W} \subseteq \mathcal{J}$ s.t. $w$ is equiprobably randomly chosen from $\mathcal{W}$, the nonempty set of candidates that end tied for first.*

Not knowing future $\mathcal{W}$, elector $i$ cannot know final reward $\mathbf{u}^i_w$. Instead, $i$ can only estimate which $\mathbf{u}^i_w$ to currently *expect* in iteration $\delta$, given $i$'s current electoral choice in $\delta$. Formally, this can be written as $E[\mathbf{u}^i_w|\pi^{i,\delta}]$, where $\pi^{i,\delta}$ represents the electoral choice of $i$ in $\delta$. In general, the choice electors are faced with is to either abstain or to choose a candidate in $\mathcal{J}$. Note, however, that like in CV models, rationality imposes that electors never vote for their sincerely least preferred candidate, since it can only lead to $i$'s maximum regret (see Myerson and Weber 1993; Cox 1994). Therefore, if abstention is loosely represented by $\varnothing$, then $\pi^{i,\delta} \in \varnothing \cup \mathcal{J} \setminus \operatorname{argmin}_{h\in\mathcal{J}} (\mathbf{u}^i_h)$.

In other words, $E[\mathbf{u}^i_w|\pi^{i,\delta} = j]$ and $E[\mathbf{u}^i_w|\pi^{i,\delta} = \varnothing]$ are the utility elector $i$ sees in whoever she currently thinks will win if, respectively, she were to vote for $j$ or if she were to abstain. Hence, every iteration $\delta$, $i$ chooses the electoral choice $\pi^{i,\delta}$ that maximizes $E[\mathbf{u}^i_w|\pi^{i,\delta}]$. Crucially, while here we will work with full turnout only,[4] considering abstention in the definition is important not only for generality, but because as we will show shortly, calculating the final formula for $E[\mathbf{u}^i_w|\pi^{i,\delta}=j] - E[\mathbf{u}^i_w|\pi^{i,\delta}=\varnothing]$ turns out to be much easier than the one for $E[\mathbf{u}^i_w|\pi^{i,\delta}]$. The reason those are, in the end, equivalent, is that $i$ abstains iff $E[\mathbf{u}^i_w|\pi^{i,\delta}=\varnothing] \geq E[\mathbf{u}^i_w|\pi^{i,\delta}=j], \forall j \in \mathcal{J}$. In summary, the electoral choice of elector $i$ in iteration $\delta$ is defined as:

**Definition 1.** *Let $\mathcal{J}^i = \mathcal{J} \setminus \operatorname{argmin}_{h\in\mathcal{J}} (\mathbf{u}^i_h)$.*

$$\pi^{i,\delta} = \begin{cases} \varnothing & \text{if } E^{i,\delta}_j - E^{i,\delta}_\varnothing \leq 0, \forall j \in \mathcal{J}^i \\ \underset{j\in\mathcal{J}^i}{\operatorname{argmax}}(E^{i,\delta}_j - E^{i,\delta}_\varnothing) & \text{otherwise} \end{cases}$$

*where $E^{i,\delta}_j \coloneqq E[\mathbf{u}^i_w|\pi^{i,\delta} = j]$ and $E^{i,\delta}_\varnothing \coloneqq E[\mathbf{u}^i_w|\pi^{i,\delta} = \varnothing]$*

Now, to find the formula for $E[\mathbf{u}^i_w|\pi^{i,\delta}]$ we start from partitioning it into three exhaustive election scenarios. The first two correspond to when $i$ believes to be pivotal, i.e. when $i$ voting for $j$ instead of abstaining would make a difference to the election outcome. The third is their complement. Under plurality, the two pivotal scenarios happen in the events when $i$'s vote either creates or breaks a tie for first. Formally, $A^{i,\delta}_{j,\mathcal{T}}$ is the event that, in $\delta$, in $i$'s perception, candidates in the set $\mathcal{T} \in \bigcup_{r=1}^{m-1} \binom{\mathcal{J}\setminus\{j\}}{r}$ are the only tied for first and $j$ is in second with one vote less than those in $\mathcal{T}$ (so a vote for $j$ would create a tie for first). $B^{i,\delta}_{j,\mathcal{T}}$ is the event that, in $\delta$, in $i$'s perception, candidates in $\mathcal{T}$ and $j$ are tied for first, with all others behind (so a vote for $j$ would isolate $j$ in first). To define that in detail, for convenience let $\mathcal{K} = \mathcal{J} \setminus \{j, \mathcal{T}\}$ be the set of remaining trailing candidates. Then:

**Definition 2.** $A^{i,\delta}_{j,\mathcal{T}} \coloneqq \{\mathbf{n}^{i,\delta}_j = \mathbf{n}^{i,\delta}_t - 1, \mathbf{n}^{i,\delta}_t > \mathbf{n}^{i,\delta}_k, \mathbf{n}^{i,\delta}_j \geq \mathbf{n}^{i,\delta}_k\}$ *and* $B^{i,\delta}_{j,\mathcal{T}} \coloneqq \{\mathbf{n}^{i,\delta}_j = \mathbf{n}^{i,\delta}_t, \mathbf{n}^{i,\delta}_t > \mathbf{n}^{i,\delta}_k, \mathbf{n}^{i,\delta}_j > \mathbf{n}^{i,\delta}_k\}$, $\forall t \in \mathcal{T}, \forall k \in \mathcal{K}$ *where* $\mathcal{T} \in \bigcup_{r=1}^{m-1} \binom{\mathcal{J}\setminus\{j\}}{r}$ *and* $\mathcal{K} = \mathcal{J} \setminus \{j, \mathcal{T}\}$.

The probabilities of those events happening, i.e. $\Pr(A^{i,\delta}_{j,\mathcal{T}})$ and $\Pr(B^{i,\delta}_{j,\mathcal{T}})$, are called pivotal probabilities. Respectively representing them by $\alpha^{i,\delta}_{j,\mathcal{T}}$ and $\beta^{i,\delta}_{j,\mathcal{T}}$ to shorten notation, we can finally partition $E[\mathbf{u}^i_w|\pi^i]$ into the three scenarios:

**Lemma 1.** *Elector $i$'s expected reward from choice $\pi^{i,\delta}$ is:*

$$\begin{aligned} E[\mathbf{u}^i_w|\pi^{i,\delta}] =& \alpha^{i,\delta}_{j,\mathcal{T}} E[\mathbf{u}^i_w|\pi^{i,\delta}, A^{i,\delta}_{j,\mathcal{T}}] + \beta^{i,\delta}_{j,\mathcal{T}} E[\mathbf{u}^i_w|\pi^{i,\delta}, B^{i,\delta}_{j,\mathcal{T}}] \\ &+ (1 - \alpha^{i,\delta}_{j,\mathcal{T}} - \beta^{i,\delta}_{j,\mathcal{T}}) E[\mathbf{u}^i_w|(A^{i,\delta}_{j,\mathcal{T}} \cup B^{i,\delta}_{j,\mathcal{T}})^\complement] \end{aligned}$$

*Proof.* From cond. expect. and Definition 2. Note in event $(A^{i,\delta}_{j,\mathcal{T}} \cup B^{i,\delta}_{j,\mathcal{T}})^\complement$, $\mathbf{u}^i_w$ is independent from $\pi^{i,\delta}$ $\forall i, \delta$ since election outcome $\mathcal{W}$ does not change regardless of $i$'s vote. □

The third term in the $E[\mathbf{u}^i_w|\pi^{i,\delta}]$ formula above cannot be easily calculated. But since it is the same regardless of $\pi^{i,\delta}$, it gets canceled out in $E^{i,\delta}_j - E^{i,\delta}_\varnothing$, which is why working with $E^{i,\delta}_j - E^{i,\delta}_\varnothing$ is easier. The other two terms are easily defined using Assumption 1 and Definition 2, then reducing to:

**Proposition 1** (Expected reward in plurality voting)**.**

$$E^{i,\delta}_j - E^{i,\delta}_\varnothing = \sum_{\mathcal{T}} \left( \frac{\alpha^{i,\delta}_{j,\mathcal{T}}}{|\mathcal{T}|} + \beta^{i,\delta}_{j,\mathcal{T}} \right) \left( \frac{\sum_{t\in\mathcal{T}} (\mathbf{u}^i_j - \mathbf{u}^i_t)}{|\mathcal{T}| + 1} \right) \quad (1)$$

*Proof.* See the regular Appendix at the end. □

The only thing left to define is how to calculate the pivotal probabilities - which are, clearly, also the only that may vary per iteration $\delta$. Next we discuss how they can be calculated.

[3] We follow the usual game-theoretical assumption of strict preferences, so $\mathbf{u}^i_j \neq \mathbf{u}^i_h, \forall j, h \in \mathcal{J} : j \neq h$.

[4] Given the bounds on utilities chosen and assuming it never happens that all electors prefer a same single candidate, the condition for abstaining in Definition 1 is never achieved unless a cost to voting is introduced. For simplicity, here it is not, but elsewhere we explore a variation of this model where voting is costly.

## 4 Pivotal Probabilities

Multiple authors have *described* pivotal probabilities, but usually, rigorous definitions are offered only for probability of breaking two-way ties, and in special cases (e.g. Hoffman 1982; Palfrey 1988; Myerson 2000). Here we start by defining a generic formalization (dropping the iteration superscript $\delta$ for readability), required for what comes next:

**Proposition 2** (General pivotal probs. in plurality voting)**.**

$$\begin{aligned} \alpha^i_{j,\mathcal{T}} = Pr(A^i_{j,\mathcal{T}}) &\coloneqq\ Pr(\bigcap_{t\in\mathcal{T}} \mathbf{n}^i_j = \mathbf{n}^i_t - 1 \bigcap_{k\in\mathcal{K}} \mathbf{n}^i_j \geq \mathbf{n}^i_k) \\ \beta^i_{j,\mathcal{T}} = Pr(B^i_{j,\mathcal{T}}) &\coloneqq\ Pr(\bigcap_{t\in\mathcal{T}} \mathbf{n}^i_j = \mathbf{n}^i_t \bigcap_{k\in\mathcal{K}} \mathbf{n}^i_j > \mathbf{n}^i_k) \end{aligned} \tag{2}$$

*Proof.* Immediate from Definition 2, noting that in event $A^i_{j,\mathcal{T}}$, because $\mathbf{n}^i_t = \mathbf{n}^i_j + 1$, then $\mathbf{n}^i_j \geq \mathbf{n}^i_k$ guarantees $\mathbf{n}^i_t > \mathbf{n}^i_k$, $\forall t \in \mathcal{T}$, $\forall k \in \mathcal{K}$. Hence the latter are not considered to avoid double-counting. Analogous for $B^i_{j,\mathcal{T}}$, but then $j$ cannot be tied with those in $\mathcal{K}$, by definition of $\mathcal{T}$. ☐

Therefore, each iteration, elector $i$'s perceived pivotal probabilities are basically functions of $\mathbf{n}^i$, current candidates' expected votes in the eyes of $i$. Those expected votes depend on two things. One is $N^i \in \mathbb{N}_{>1}$, what $i$ thinks is the number of electors. The way electors calculate pivotal probabilities depends on whether they all know true $N$ (s.t. $N^i = N, \forall i$) or, if they don't, what they assume about it.

Another is the vector $\mathbf{s}^\delta \in \mathbb{Q}^m_{[0,1]}$, which holds the probabilities that an elector chosen at random in iteration $\delta - 1$ would vote for each candidate - which for $\delta > 0$ can be seen as candidates' true expected vote shares in $\delta - 1$. When $\delta = 0$, $\mathbf{s}^\delta$ can be initialized in many ways. For instance, it can hold proportions of sincere votes each candidate would get in the absence of strategic voting (truthful initial condition). Or it can be simply randomized (random initial condition). While here we focus on the latter, results qualitatively hold with the former. Finally, it is assumed that $\mathbf{s}^\delta$ is common knowledge (e.g. informed by polls like in Fey (1997)).[5]

### Known electorate size

Suppose all electors know the electorate size, i.e. $N^i = N, \forall i$. Then, in any iteration, they do not need to guess candidates' votes. Instead, they can exhaustively list all vote combinations that form pivotal scenarios and calculate pivotal probabilities exactly. This is known to follow a Multinomial distribution (Palfrey 1988). We generalize it for multi-candidate cases, like in Cox (1994); Fey (1997); Tsang and Larson (2016), but with multi-way ties and both probabilities of breaking and making ties. Let vector $\mathbf{v}$ hold a fictitious final vote counts combination disregarding $i$'s vote. Then, let $\mathbf{V}^\alpha$ be the set of all possible $\mathbf{v}$ that correspond to a scenario where one more vote for $j$ would create a tie with candidates in $\mathcal{T}$, i.e. all possible $\mathbf{v}$ that correspond to an event $A^i_{j,\mathcal{T}}$, in case such an event was defined in terms of $\mathbf{v}$, instead of $\mathbf{n}^i$. Similarly for $\mathbf{V}^\beta$ in relation to $B^i_{j,\mathcal{T}}$:

[5] If expected vote shares come from polls, it is assumed, of course, that elector $i$ treats them *as if* they are the vote shares of the entire population. For an application with non-unique expected vote shares, coming from networks, see Tsang and Larson (2016).

**Definition 3.** $\mathbf{v} \in \mathbb{N}^m$ *s.t.* $\sum_h \mathbf{v}_h = \lfloor \sum_h \mathbf{n}_h \rceil - 1$. *Then from (2):*

$$\begin{aligned} \mathbf{V}^\alpha &= \Big\{ \mathbf{v}\colon \mathbf{v}_t = \mathbf{v}_j + 1\ \forall t \in \mathcal{T} \wedge \mathbf{v}_j \geq \mathbf{v}_k\ \forall k \in \mathcal{K} \Big\} \\ \mathbf{V}^\beta &= \Big\{ \mathbf{v}\colon \mathbf{v}_t = \mathbf{v}_j\ \forall t \in \mathcal{T} \wedge \mathbf{v}_j > \mathbf{v}_k\ \forall k \in \mathcal{K} \Big\} \end{aligned} \tag{3}$$

**Proposition 3** (Multinomial piv. probs. in plurality voting)**.** $\forall i, i', \alpha^i_{j,\mathcal{T}} = \alpha^{i'}_{j,\mathcal{T}} = \alpha_{j,\mathcal{T}}$ *and* $\beta^i_{j,\mathcal{T}} = \beta^{i'}_{j,\mathcal{T}} = \beta_{j,\mathcal{T}}$*, s.t.:*

$$\alpha_{j,\mathcal{T}} = \sum_{\mathbf{v} \in \mathbf{V}^\alpha} \frac{(N-1)!}{\mathbf{v}_1! \mathbf{v}_2! \ldots \mathbf{v}_m!} \prod_{h=1}^{m} (\mathbf{s}_h)^{\mathbf{v}_h} \tag{4}$$

*and same for* $\beta_{j,\mathcal{T}}$*, just summing over all* $\mathbf{v} \in \mathbf{V}^\beta$ *instead.*

*Proof.* $\mathbf{n}^i = \mathbf{n}^{i'}$, $\forall i, i'$ because $N^i = N, \forall i$. Then, from (2), $\alpha^i_{j,\mathcal{T}} = \alpha^{i'}_{j,\mathcal{T}}$ and $\beta^i_{j,\mathcal{T}} = \beta^{i'}_{j,\mathcal{T}}$, $\forall i, i'$. Now, consider any fictitious final outcome $\mathbf{v}$. Unaware of others' votes, in the eyes of $i$ an election is a collection of other $N-1$ stochastic decisions, with $m$ possible results whose probabilities are given by $\mathbf{s}$. That experiment follows a Multinomial distribution with event probabilities $\mathbf{s}$, support $\mathbf{v}$ and number of trials $N-1$. Then, $\alpha_{j,\mathcal{T}}$ is a convolution of Multinomial PMFs evaluated at every $\mathbf{v} \in \mathbf{V}^\alpha$, same for $\beta_{j,\mathcal{T}}$ and $\mathbf{v} \in \mathbf{V}^\beta$. ☐

Clearly, (4) scales poorly - each pair $\alpha_{j,\mathcal{T}}$, $\beta_{j,\mathcal{T}}$ taking $\mathcal{O}\big(m(|\mathbf{V}^\alpha| + |\mathbf{V}^\beta|)\big)$ (see regular Appendix for the equation for $|\mathbf{V}^\alpha|$ and $|\mathbf{V}^\beta|$). That can be ameliorated by memoizing expensive terms, but another critical computational cost is that of merely finding $\mathbf{V}^\alpha$ and $\mathbf{V}^\beta$. The naive approach is to find the $m$-fold cartesian product $\langle 0, 1, \ldots, N \rangle^m$ and drop subvectors that do not follow conditions in (3). Algorithm 1 decreases the problem by realizing that for each fictitious vote of the focal candidate, those tied for first rank can only have same or one extra vote, which limits the max and total votes others can have (see Online Appendix for details).

Algorithm 1: Listing pivotal outcomes $\mathbf{V}^\alpha$ and $\mathbf{V}^\beta$

**function** LISTPIVOTALOUTCOMES($N, \mathcal{J}, \mathcal{T}$)
  $\mathbf{V}^\alpha, \mathbf{V}^\beta \leftarrow \emptyset$
  **for** $x \leftarrow 0$ **to** $(N-1)/(|\mathcal{T}|+1)$ **do**
    $\mathcal{A} \leftarrow \langle 0, \ldots, x \rangle^{N-|\mathcal{T}|-1}$ ▷ cartesian power
    $\mathcal{B} \leftarrow \emptyset$
    **for all** $\boldsymbol{a} \in \mathcal{A}$ **do**
      append $\boldsymbol{a}$ to $\mathcal{B}$ **if** $\max(\boldsymbol{a}) < x$
    **end for**
    $y \leftarrow \langle x+1 \rangle^{|\mathcal{T}|} \oplus \langle x \rangle$ ▷ $\oplus$: concatenation
    $z \leftarrow \langle x \rangle^{|\mathcal{T}|+1}$
    **for all** $\mathbf{v} \in y \times \mathcal{A}$ **do**
      append $\mathbf{v}$ to $\mathbf{V}^\alpha$ **if** $\mathrm{sum}(\mathbf{v}) = N-1$
    **end for**
    **for all** $\mathbf{v} \in z \times \mathcal{B}$ **do**
      append $\mathbf{v}$ to $\mathbf{V}^\beta$ **if** $\mathrm{sum}(\mathbf{v}) = N-1$
    **end for**
  **end for**
  **return** $\mathbf{V}^\alpha, \mathbf{V}^\beta$
**end function**

## Unknown electorate size

Consider now that electors do not know the electorate size, which happens in many applications, from voting in online forums to any larger electorate. Then, following Myerson (1998) it can be assumed that unknown $N$ comes from a Poisson distribution with mean $\lambda$. In which case, the author proved that players' guesses about total number of players would be random variables coming from same distribution:

**Lemma 2.** *$\forall i$, $N^i \sim Pois(\lambda)$ iff $N \sim Pois(\lambda)$. From environmental equivalence, Theorem 2 in Myerson (1998).*

Then in any iteration $\delta$, candidates' expected votes $\mathbf{n}^*_j$ are also distributed Poisson with mean equal to $\lambda$ times the probability a voter chosen at random votes for that candidate:

**Lemma 3.** *$\mathbf{n}^*_j \sim Pois(\mathbf{s}_j\lambda)$, $\forall j \in \mathcal{J}$. From Poisson decomposition property (see Myerson 2000).*

Furthermore, candidates' expected votes are then Poisson random variables independent from each other:

**Lemma 4.** *$\mathbf{n}^*_j \perp\!\!\!\perp \mathbf{n}^*_h$, $\forall j, h \in \mathcal{J} : j \neq h$. From independent action property, Theorem 1 in Myerson (1998).*

With that, Myerson (2000) showed the pivotal probability of breaking a tie in two-candidate plurality to be equivalent to the probability mass function of a Skellam distribution evaluated at zero (Smith, de Mesquita, and LaGatta 2016). This comes from the fact that the subtraction of two Poisson random variables results in a Skellam: if $\mathcal{T} = \{h\}$ and $\mathcal{K} = \emptyset$, because $\mathbf{n}^*_j - \mathbf{n}^*_h \sim \text{Skellam}(\mathbf{s}_j\lambda, \mathbf{s}_h\lambda)$, then $\beta^i_{j,\mathcal{T}} = f_S(0, \mathbf{s}_j\lambda, \mathbf{s}_h\lambda)$. Yet, that connection to the Skellam does not extend naturally to more general scenarios.

As a workaround, recent work have suggested *approximations* to extend that to multi-candidate cases with two-way ties, based on arbitrary and more-or-less defined simplifying assumptions (Tsang, Salehi-Abari, and Larson 2018; Anonymous 2018). However, properties of those approximations have not been studied, and it is unclear how good they are, what their quality depends on or whether they are computationally worth. Differently, below we prove the *exact* generalization of Myerson's pivotal probabilities - to breaking or making ties, in multi-candidate plurality, with multi-way ties - which we call Poisson pivotal probabilities.

**Proposition 4** (Poisson pivotal probs. in plurality voting)**.** *$\forall i, i'$, $\alpha^i_{j,\mathcal{T}} = \alpha^{i'}_{j,\mathcal{T}} = \alpha_{j,\mathcal{T}}$ and $\beta^i_{j,\mathcal{T}} = \beta^{i'}_{j,\mathcal{T}} = \beta_{j,\mathcal{T}}$, s.t.:*

$$\begin{aligned}\alpha_{j,\mathcal{T}} &= \sum_{d=0}^{\infty}\Big(f_P(d, \mathbf{s}_j\lambda)\prod_{t\in\mathcal{T}} f_P(d+1, \mathbf{s}_t\lambda)\prod_{k\in\mathcal{K}} F_P(d, \mathbf{s}_k\lambda)\Big)\\ \beta_{j,\mathcal{T}} &= \sum_{d=0}^{\infty}\Big(f_P(d, \mathbf{s}_j\lambda)\prod_{t\in\mathcal{T}} f_P(d, \mathbf{s}_t\lambda)\prod_{k\in\mathcal{K}} F_P(d-1, \mathbf{s}_k\lambda)\Big)\end{aligned} \tag{5}$$

*where $f_P$ and $F_P$ are Poisson distributions PMF and CDF.*

*Proof.* Since $\mathbf{n}^i_j$ is part of all intersecting events in (2), by Lemma 4 conditioning on $\mathbf{n}^i_j$ makes the events independent:

$$\alpha^i_{j,\mathcal{T}} = \sum_{d=0}^{\infty}\Big(\Pr(\mathbf{n}^i_j = d)\prod_{t\in\mathcal{T}}\Pr(\mathbf{n}^i_t = d+1)\prod_{k\in\mathcal{K}}\Pr(\mathbf{n}^i_k \leq d)\Big)$$

$$\beta^i_{j,\mathcal{T}} = \sum_{d=0}^{\infty}\Big(\Pr(\mathbf{n}^i_j = d)\prod_{t\in\mathcal{T}}\Pr(\mathbf{n}^i_t = d)\prod_{k\in\mathcal{K}}\Pr(\mathbf{n}^i_k < d)\Big)$$

Now, by Lemma 2 and noting that $\mathbf{n}^i_j$ is a partition of $N^i$ as much as $\mathbf{n}^*_j$ is of $N$ in Lemma 3, $\mathbf{n}^i_j \sim \text{Pois}(\mathbf{s}_j\lambda)$, $\forall i, \forall j$; therefore $\alpha^i_{j,\mathcal{T}} = \alpha^{i'}_{j,\mathcal{T}}$ and $\beta^i_{j,\mathcal{T}} = \beta^{i'}_{j,\mathcal{T}}$, $\forall i, i'$. $\square$

Critically, note that the proof also shows that despite each voter holding a different guess $N^i$, pivotal probabilities end being the same for all voters. This is useful to enable agent simulations in large electorates, since it prevents having to otherwise calculate the set of probabilities $N$ times per iteration. Besides, while the inexistence of a closed form for the summations in (5), whose theoretical rate of convergence is unknown, may look challenging, in practice they converge quickly. Furthermore, Algorithm 2 shows how (5) can be simplified such that $\alpha_{j,\mathcal{T}}$ and $\beta_{j,\mathcal{T}}$ are found jointly and efficiently, with no explicit calculation of $f_P$ or $F_P$ (see the regular Appendix at the end for details).

Algorithm 2: Joint calculation of Poisson pivotal probs.

**function** POISSONPIVOTALPR($j$, $\mathcal{T}$, $\mathcal{K}$, $\mathbf{s}$, $\lambda$, maxd)
    $\boldsymbol{\epsilon} \in \mathbb{Q}^{|\mathbf{s}|}_{[0,1]} \leftarrow$ vector with $\exp(-\mathbf{s}_h\lambda)$ $\forall h = 1 \dots |\mathbf{s}|$
    $\boldsymbol{\kappa} \in \mathbb{Q}^{|\mathcal{K}|}_{[0,1]} \leftarrow$ zero vector of length $|\mathcal{K}|$
    $\boldsymbol{\tau} \in \mathbb{Q}^{|\mathcal{T}|}_{[0,1]} \leftarrow$ vector with $\boldsymbol{\epsilon}_t$ $\forall t \in \mathcal{T}$
    a, b $\leftarrow \boldsymbol{\epsilon}_j$
    $\alpha$, $\beta$, $\alpha'$, $\beta'$, d $\leftarrow 0$
    q $\leftarrow 1$
    **while** ($\alpha \neq \alpha'$ **or** $\beta \neq \beta'$ **or** d $= 0$) **and** $d <$ maxd **do**
        a $\leftarrow$ b
        b $\leftarrow$ b $\cdot \frac{\mathbf{s}_j\lambda}{(\mathrm{d}+1)}$
        **for all** $t \in \mathcal{T}, r = 1 \dots |\mathcal{T}|$ **do**
            $\boldsymbol{\tau}_r \leftarrow \boldsymbol{\tau}_r \cdot \frac{\mathbf{s}_t\lambda}{(\mathrm{d}+1)}$
        **end for**
        **for all** $k \in \mathcal{K}, r = 1 \dots |\mathcal{K}|$ **do**
            $\boldsymbol{\kappa}_r \leftarrow \boldsymbol{\kappa}_r + \boldsymbol{\epsilon}_k \cdot \frac{(\mathbf{s}_k\lambda)^{\mathrm{d}}}{\mathrm{q}}$
        **end for**
        $\alpha' \leftarrow \alpha$
        $\beta' \leftarrow \beta$
        z $\leftarrow \prod_{t\in\mathcal{T}}(\boldsymbol{\tau}_t)\prod_{k\in\mathcal{K}}(\boldsymbol{\kappa}_k)$ **if** $|\mathcal{K}| > 0$ **else** $1$
        $\alpha \leftarrow \alpha +$ a $\cdot$ z
        $\beta \leftarrow \beta +$ b $\cdot$ z
        d $\leftarrow$ d $+ 1$
        q $\leftarrow$ q $\cdot\, d$
    **end while**
    **return** ($\alpha$, $\beta$)
**end function**

Next, we specify exactly under which (unrealistic) simplifying assumptions the above can be approximated solely by Skellam PMFs and CDFs like proposed by Tsang, Salehi-Abari, and Larson (2018) or Anonymous (2018), in a principled manner. Then, we derive the proper approximation, also generalized for probabilities of breaking or making ties, with multi-candidates and allowing for multi-way ties:

**Assumption 2.** *Knowing whether expected votes of two candidates are equal, lower or greater than each other does not give information about any other pair of candidates.*

**Proposition 5** (Skellam pivotal probs. in plurality voting)**.** *$\forall i, i'$, $\alpha^{i}_{j,\mathcal{T}} = \alpha^{i'}_{j,\mathcal{T}} = \alpha_{j,\mathcal{T}}$ and $\beta^{i}_{j,\mathcal{T}} = \beta^{i'}_{j,\mathcal{T}} = \beta_{j,\mathcal{T}}$, s.t.:*

$$\begin{aligned}\alpha_{j,\mathcal{T}} &\approx \prod_{t\in\mathcal{T}} f_S(-1, \mathbf{s}_j\lambda, \mathbf{s}_t\lambda) \prod_{k\in\mathcal{K}} 1 - F_S(-1, \mathbf{s}_j\lambda, \mathbf{s}_k\lambda)\\ \beta_{j,\mathcal{T}} &\approx \prod_{t\in\mathcal{T}} f_S(0, \mathbf{s}_j\lambda, \mathbf{s}_t\lambda) \prod_{k\in\mathcal{K}} 1 - F_S(0, \mathbf{s}_j\lambda, \mathbf{s}_k\lambda)\end{aligned} \tag{6}$$

*where $f_S$ and $F_S$ are Skellam distributions' PMF and CDF.*

*Proof.* Given Assumption 2, eq. (2) can be rewritten: $\alpha^{i}_{j,\mathcal{T}} = \prod_{t\in\mathcal{T}} \Pr(\mathbf{n}^i_j = \mathbf{n}^i_t - 1) \prod_{k\in\mathcal{K}} 1 - \Pr(\mathbf{n}^i_j \leq \mathbf{n}^i_k - 1)$ $\beta^{i}_{j,\mathcal{T}} = \prod_{t\in\mathcal{T}} \Pr(\mathbf{n}^i_j = \mathbf{n}^i_t) \prod_{k\in\mathcal{K}} 1 - \Pr(\mathbf{n}^i_j \leq \mathbf{n}^i_k)$. Same as in Proposition 4, $\mathbf{n}^i_j \sim \text{Pois}(\mathbf{s}_j\lambda)$, $\forall i, \forall j$. Then, by the definition of Skellam random variables, regarding $\alpha^{i}_{j,\mathcal{T}}$ we have $\Pr(\mathbf{n}^i_j - \mathbf{n}^i_t = -1) = f_S(-1, \mathbf{s}_j\lambda, \mathbf{s}_t\lambda)$ and similarly $\Pr(\mathbf{n}^i_j - \mathbf{n}^i_k \leq -1) = F_S(-1, \mathbf{s}_j\lambda, \mathbf{s}_k\lambda)$. Analogous for $\beta^{i}_{j,\mathcal{T}}$. Calculating those is possible because $f_S$ is a convolution of two Poisson PMFs (Skellam 1946), and while $F_S$ has no closed formula, following Johnson (1959) one can use as a step-wise function of the CDF of the Non-central Chi-Square distribution. See regular Appendix at the end. $\square$

Later, we will explore through simulations the quality conditions and computational efficiency of this approximation we call Skellam pivotal probabilities *vis-a-vis* the exact solution in (5), which we dub Possion pivotal probabilities.

## 5 Convergence properties

Now that ICV can be fully implemented, we discuss some of its convergence properties. It is known that plurality CV reaches equilibrium in standard one-shot non-atomic games (Palfrey 1988; Myerson and Weber 1993), including CV with polls (Fey 1997). Here we start by showing that also ICV under plurality converges to a Pure Nash Equilibrium (PNE) as electorates become large. Formally, ICV is guaranteed to have converged to stable equilibria whenever no elector changes their electoral choice any longer:

**Definition 4.** *PNE in ICV under plurality is an iteration $\delta^* > 0$ such that: $\pi^{i,\delta^*} = \pi^{i,\delta^*+1} = \cdots = \pi^{i,\delta^*+\infty}, \forall i$.*

However, it can be shown that if no elector changes their electoral choice for two subsequent iterations, that already guarantees none will ever change. Then, even more conveniently, it can be also shown that if the expected vote shares of all candidates remain the same for two subsequent iterations, that already guarantees no elector will change their electoral choice any longer. This comes from the fact that the only element of electors' election reward function that can vary across iterations are the pivotal probabilities and those only vary when candidates' expected shares of votes vary. This is what the next proposition shows:

**Proposition 6.** *$\mathbf{s}^{\delta}_h = \mathbf{s}^{\delta+1}_h$, $\forall h \in \mathcal{J}$ iff $\delta = \delta^*$ ($\delta$ is a PNE).*

*Proof.* Consider what follows $\forall i$, $\forall \delta > 0$ and $\forall h \in \mathcal{J}$. If $\mathbf{s}^{\delta}_h = \mathbf{s}^{\delta+1}_h$, since $N^i$ is fixed and $\mathbf{n}^i_h$ is a function of only $N^i$ and $\mathbf{s}^{\delta}_h$, then also $\mathbf{n}^{i,\delta}_h = \mathbf{n}^{i,\delta+1}_h$. Thus, given (2), pivotal probabilities do not change from $\delta$ to $\delta + 1$ and because they are the only that could change in (1), $E^{i,\delta}_j - E^{i,\delta}_\varnothing = E^{i,\delta+1}_j - E^{i,\delta+1}_\varnothing$. Which results in $\pi^{i,\delta} = \pi^{i,\delta+1}$ and then, by the definition of $\mathbf{s}^\delta$, also in $\mathbf{s}^{\delta+1}_h = \mathbf{s}^{\delta+2}_h$. The same logic follows successively, leading to both $\mathbf{s}^{\delta^*}_h = \mathbf{s}^{\delta^*+1}_h = \cdots = \mathbf{s}^{\delta^*+\infty}_h$ and $\pi^{i,\delta^*} = \pi^{i,\delta^*+1} = \cdots = \pi^{i,\delta^*+\infty}$. $\square$

This is important because, then, proving that ICV results in PNE with large electorates simplifies to proving that there will be two subsequent iterations where candidates' expected vote shares will remain the same. But before we proceed to that, a few details need to be established. Firstly, we will need to impose the (rather innocuous) assumption that, as the electorate size approaches infinity, so do electors' eventual guesses about that size (and vice-versa):

**Assumption 3.** *$N \to \infty$ iff $N^i \to \infty$, $\forall i$.*

Secondly, we need to formally establish the otherwise intuitive fact that, as a candidate's expected vote share approximates zero, events where it appears tied for 1st become nearly impossible[6]:

**Lemma 5.** *$\lim_{\mathbf{s}_t \to 0} \alpha^{i}_{j,\mathcal{T}} = \lim_{\mathbf{s}_t \to 0} \beta^{i}_{j,\mathcal{T}} = 0$, $\forall j, \forall t \in \mathcal{T}, \forall \mathcal{K} \neq \varnothing$.*

*Proof.* Rewrite (2) in term of events $\mathbf{n}^i_t > \mathbf{n}^i_k$, $\forall t \in \mathcal{T}, \forall k \in \mathcal{K}$, instead of $\mathbf{n}^i_j \geq \mathbf{n}^i_k$ or $\mathbf{n}^i_j > \mathbf{n}^i_k$. Then, the lower $\mathbf{s}_t$, the less likely that $\mathbf{n}^i_t > \mathbf{n}^i_k$, $\forall t \in \mathcal{T}, \forall k \in \mathcal{K}$ - until it becomes impossible when $\mathbf{s}_t = 0$, in which case $\alpha^{i}_{j,\mathcal{T}} = \beta^{i}_{j,\mathcal{T}} = 0$. $\square$

Finally, we recall from the CV literature the well known condition that the ratio between the probability of non-leading candidates being pivotal and the probability of leading candidates being pivotal must go to zero as $N \to \infty$:

**Condition 1.** *Let $1, \cdots, g, \cdots, m$ be the rank of $m$ candidates' expected vote shares, where $g \in \mathbb{N}_{[2,m-1]}$, s.t. $\mathbf{s}_1 = \cdots = \mathbf{s}_{g-1}$ and $\mathbf{s}_{g+1} \geq \cdots \geq \mathbf{s}_m \geq 0$, with either $\mathbf{s}_{g-1} > \mathbf{s}_g \geq \mathbf{s}_{g+1}$ or $\mathbf{s}_{g-1} \geq \mathbf{s}_g > \mathbf{s}_{g+1}$. Then, respectively:*

$$\lim_{N\to\infty} \alpha^i_{h,G} / \alpha^i_{r,G'} = 0 \quad \textit{and} \quad \lim_{N\to\infty} \beta^i_{h,G} / \beta^i_{r,G'} = 0$$

*where $r \in \{1, \cdots, g-1\}$, $h \in \{g+1, \cdots, m\}$, $G \subseteq \{1, \cdots, m\} \setminus \{h, r\}$ and $G' \subseteq \{1, \cdots, g\} \setminus \{r\}$.*

Palfrey (1988) proved the condition holds for the case $N$ is constant and Chen and Xia (2011) for cases when it is a random variable, including a Poisson in a Poisson game. For simplicity, both proofs focus on $\beta$ with $|G| = |G'| = 1$, but the logic is identical for $\alpha$ and extends mathematically trivially to $|G| > 1$, $|G'| > 1$. Given the condition holds, we can prove both that $\delta^*$ exists in non-atomic plurality ICV as $N \to \infty$, and also how $\delta^*$ is reached. While not the same, the intuition of the proof resembles Dhillon and Lockwood (2004)'s iterative elimination of dominated strategies in simultaneous voting - in the sense that voting options that are deserted due to strategic voting become progressively less appealing up to full abandonment.

There are four cases to consider. In *Case 1*, all candidates are tied for 1st. Since then all expected vote shares are the

[6] Except in the degenerate circumstance when all candidates are expected to have approximately same vote shares near zero.

same, all pivotal probabilities are identical - thus voters simply vote sincerely. If this leads to same vote shares, we say a trivial PNE is reached; otherwise it leads to Case 2, 3, or 4. In *Case 2*, one candidate is in 1st, all else are tied for 2nd. Again voters will vote sincerely; some because they genuinely prefer the leader, others because probability of being pivotal when voting for any runner-up is identical. This leads to a trivial PNE or to Case 1, 3 or 4.[7] In *Case 3*, two candidates are isolated in the top 2. Condition 1, together with Assumption 3, make it so that as the electorate size approaches infinity, only the top 2 candidates are seen as viable - leading to the abandonment of others. Among the top 2, voters will of course choose sincerely, which guarantees a PNE. In *Case 4*, multiple (but not all) candidates are tied for 1st. Again from Condition 1 and Assumption 3, as the electorate size approaches infinity trailing candidates are abandoned. This leads to Case 3 or Case 4 with shrank set of viable options (and so on up until it becomes Case 3). Formally:

**Proposition 7** (ICV asymptotic convergence)**.**
$\lim_{N\to\infty} Pr(\exists\delta : \mathbf{s}_{j,\delta} = \mathbf{s}_{j,\delta-1}, \forall j) = 1.$

*Proof.* See regular Appendix at the end. □

**Corollary 1.** *As $N \to \infty$, we have the following convergence bounds. If $\delta = 0$ is Case 1 or 2, $1 \leq \delta^* \leq m-1$. If $\delta = 0$ is Case 3, $\delta^* = 1$. If $\delta = 0$ is Case 4, $2 \leq \delta^* \leq g$.*

Meir (2018) points out that reproducing the Duverger's Law (Duverger 1963) is a critical scientific criterion for a plurality voting model. Usually, plurality Duvergerian equilibrium is defined as an equilibrium where only two candidates have positive votes (Palfrey 1988; Cox 1994). We propose treating that as *strong* Duvergerian equilibrium:

**Definition 5.** *A Strong Duvergerian Equilibrium (SDE) is a $\delta^* : \mathbf{s}_{1,\delta^*} \geq \mathbf{s}_{2,\delta^*} > 0$ and $\mathbf{s}_{h,\delta^*} = 0, \forall h \in \mathcal{J} \setminus \{1,2\}$.*

Just like standard CV games, as electorate size approaches infinity, ICV also converges to SDE:

**Corollary 2.** *Because as $N \to \infty$ the prob. of cases 1, 2 or 4 in Proposition 7 happening in $\delta > 0$ becomes infinitesimal (Hoffman 1982), then $\lim_{N\to\infty} Pr(\delta^* \text{is a SDE}) = 1$.*

Nevertheless, a natural question is how large $N$ has to be for convergence to be guaranteed, since real applications have finite agents. In theory, because a candidate can become more pivotal both when it gains or loses votes, depending on the context - differently from Local-Dominance IV models (Meir, Lev, and Rosenschein 2014; Meir 2015) - in ICV voters can move back to past choices and thus cycles are a possibility. Yet, we will show through simulations that convergence nearly always results, even with the smallest electorates. To see why, recall that in (2), in general the lower $\mathbf{s}_j$ is, the less likely $\mathbf{n}_j > \mathbf{n}_k, \forall k$. Then, in practice, voting for $j$ generally becomes more (less) pivotal as $j$ gains (loses) votes more rapidly than as $j$ looses (gains) votes. In other words, once a candidate starts loosing support, it is hard to recover it - which is intensified the larger the electorate is.

[7]Note that a cycle between Cases 1 and 2 is not possible. All electors voting sincerely can lead to only of them.

But such a wasted vote avoidance does not necessarily lead to *full* abandonment of trailing candidates in finite electorates, thus SDE is not guaranteed. Since Duverger's Law merely states that votes *tend to concentrate* in 2 candidates, we propose also a weaker Duvergerian Equilibrium concept. Let $\rho \in \mathbb{Q}_{1,m}$ be the effective number[8] of candidates $\rho = 1/\sum_h^m (\mathbf{s}_h)^2$ (Laakso and Taagepera 1979):

**Definition 6.** *A Weak Duvergerian Equilibrium (WDE) is a $\delta^*$ where $\rho_{\delta^*} \leq 3 - \varepsilon$, for a chosen $0 < \varepsilon \leq 1$.*

**Conjecture 1.** *Let $\rho_{\delta^*} = 2 + x$, where $x \in \mathbb{Q}_{[-1,m-2]}$. Then $N \uparrow \implies x \to 0$. Therefore, $N \uparrow \implies Pr(\delta^* \text{is a WDE}) \uparrow$.*

While that conjecture cannot be proved easily for finite electorates, it does make intuitive sense. As candidates lose significant expected vote shares to others, by definition $\rho$ becomes more concentrated. Our simulations confirm the pattern: the greater the electorate, the less likely $\rho_{\delta^*}$ diverges much from 2. Therefore, even with finite agents, $\delta^*$ becomes most often at least WDE, depending on the chosen $\epsilon$.

## 6 Simulations

We implemented ICV in Python 3.7.3. First, 10,000 simulations were performed using Multinomial pivotal probabilities. Then, their pseudo-random seeds were used to repeat the simulations twice, each time using Poisson or Skellam probabilities - for a total of 30,000 simulations.

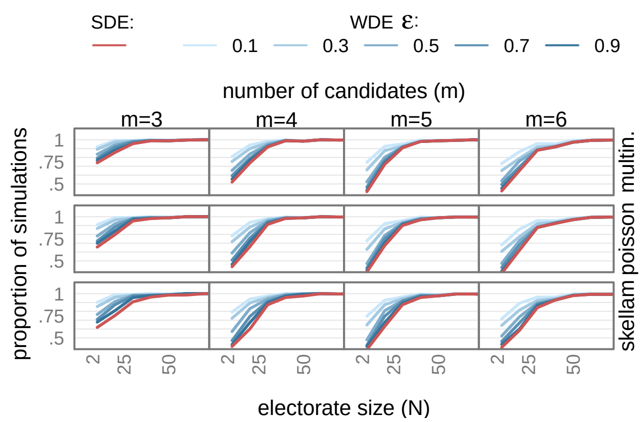


Figure 1: Prop. of simulations converging to Strong (SDE) or to Weak (WDE) Duvergerian Equilibria for a chosen $\varepsilon$.

**Initialization Values.** In those simulations, model hyperparameters were specified as: $\lambda \sim \text{Uniform}(2, 100)$, $N \sim \text{Pois}(\lambda)$ and $m \sim \text{Uniform}(3, 6)$. Electors' candidate utilities were drawn from Beta distributions, with varying parameters: $\mathbf{u}^i \sim \text{Beta}(\text{Uniform}(0.1, 5.0), \text{Uniform}(0.1, 5.0))$. The Beta ensures diversity: it can approximate the Uniform, the Gaussian, the Gamma, a Power Law or be bimodal. Hence, initial votes will be different across different initial seeds.

**Convergence.** Convergence to an equilibrium was established when no electors altered their chosen candidate for

[8]This widely used index in Political Science measures concentration. Suppose 3 candidates: the 1st has 50% of votes, the 2nd has 45% and the 3rd has 5%. Then $\rho \approx 2.2$ effective candidates.

two iterations in a row. All runs converged (in at most 15 iterations, 9th percentile of 5 iterations), except for 63 of $10,000$ runs employing Skellam pivotal probabilities, which resulted in cycles (all $< 55$ electors, 9th perc. 31 electors).

**Duvergerian equilibria.** Around $85.7\%$ simulations converged to a SDE. Red lines in Figure 1 show nearly all with more than 30 agents resulted in SDE. From blue lines, notice that outcomes not SDE are often at least WDE.[9]

**Outcome similarity.** Interestingly, simulations repeated with same starting pseudo-random number generating seed and differing only in types of pivotal probabilities used, generally resulted in same winner and in very similar overall outcomes, when $N > 25$ (see Figure 2). In fact, as the electorate increases, not knowing electorate sizes makes progressively less of a difference (Poisson or Skellam vs. Multinomial) and the Skellam becomes a very good approximation of the Poisson pivotal probabilities (which is the more important the greater the number of candidates).

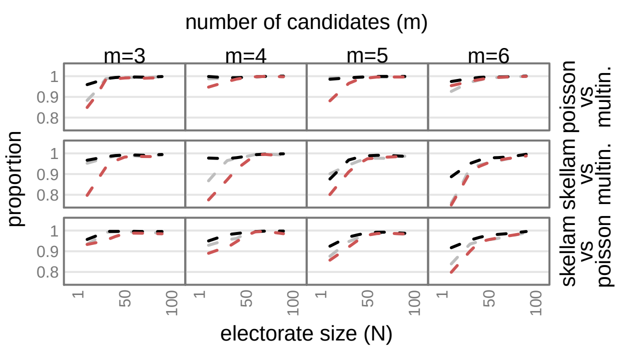


Figure 2: Proportion of simulations that, repeated with same seed using different types of pivotal probabilities, converged to same election winner (black) or to at least 90% similar candidate ranking (red) or final vote counts (gray).

**Runtime.** Figure 3 confirms that simulations using Multinomial pivotal probabilities scale much worse, in particular as $m$ increases. But note that the Skellam pivotal probabilities approximation is just a bit faster than the exact Poisson and both scale similarly. Therefore, the main practical advantage of the Skellam approximation is that it seems to lose numerical precision more slowly, likely because (6) is composed of a few multiplications of infinite sums, while (5) has infinite sums of multiplications.[10]

## 7 Discussion

Differently from IV models, ICV is inherently capable of admitting simultaneous optimization updates and is useful for large electorates. Differently from CV models, agents achieve convergence through repeated iterations, not abstractly, and without knowledge of preference distributions

[9]In their application, Tsang and Larson (2016) show via simulations that if electors see different expected vote shares according to voter homophily, what we call SDE may be harder to achieve.

[10]Comparison with recent applications (e.g.Tsang, Salehi-Abari, and Larson 2018) is hard. They only included prob. of breaking two-way ties and only loosely approximated the Skellam.

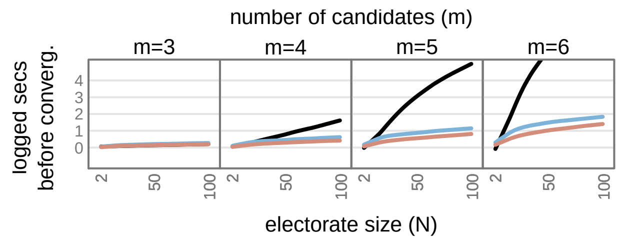


Figure 3: Lowess lines: log. seconds to convergence with Multinm. (black), Poisson (blue) or Skellam pivotal probs.

- knowing only candidates' expected vote shares. This way, ICV addresses limitations of most IV and CV models.

We have shown that with pivotal probabilities that conform to certain characteristics, convergence in ICV is guaranteed for arbitrarily large electorates. Furthermore, in practice, convergence can usually be achieved even with tiny electorates under weak assumptions. Besides, unless electorates are tiny, ICV with proper pivotal probabilities mostly converges to either having only 2 candidates with positive votes, or at least having votes concentrated in 2 candidates - thus passing the critical Duverger's Law test (Meir 2018).

ICV can also be efficiently simulated using either Poisson or Skellam pivotal probabilities. However, a theoretical limitation is that, like the Multinomial, they also require sophisticated voters. Agents must realize that, under the specified circumstances, their probability of being pivotal is a function of one of those distributions. This is in opposition to the lesser sophistication requirements in most IV (see e.g. Meir 2015). Simpler heuristics can and have been proposed. Our hope is that our thorough treatment may serve as a theoretical guide to inform principled heuristics, as well as a benchmark to evaluate them regarding efficiency and outcome similarity. We also have showed characteristics that pivotal probabilities require if heuristics were to aim at guaranteeing convergence and Duvergerian equilibria.

Other limitations of the presented ICV are the costlessness of voting and of strategizing. Exploring ICV with abstention, lazy-voting (Elkind et al. 2010, 2015) and truth-bias (Meir et al. 2010; Obraztsova, Markakis, and Thompson 2013; Elkind et al. 2015) are logical next steps. Introducing uncertainty about states is also desirable, which has been explored through poll uncertainty both in CV (Myatt 2007) and in IV (Reijngoud and Endriss 2012; Wilczynski 2019). Lastly, extending ICV to other election rules seems promising, like SNTV (Cox 1994), Approval and Borda (Myerson and Weber 1993) and Runoff (Bouton and Gratton 2015).

## A Proof of Proposition 1

*Proof.* Recall Assumption 1. Consider first $A^{i,\delta}_{j,\mathcal{T}}$. If $i$ votes for $j$, then a tie between $j$ and candidates in $\mathcal{T}$ is created, so $E[\mathbf{u}^i_w|\pi^{i,\delta}=j, A^{i,\delta}_{j,\mathcal{T}}] = {}^{(\mathbf{u}^i_j+\sum_{t\in\mathcal{T}}\mathbf{u}^i_t)}/_{(|\mathcal{T}|+1)}$. If, instead, $i$ abstains, not creating a tie between $j$ and those in $\mathcal{T}$, $E[\mathbf{u}^i_w|\pi^{i,\delta}=\varnothing, A^{i,\delta}_{j,\mathcal{T}}] = {}^{(\sum_{t\in\mathcal{T}}\mathbf{u}^i_t)}/_{|\mathcal{T}|}$. Now, consider $B^{i,\delta}_{j,\mathcal{T}}$. If $i$ votes for $j$, breaking the tie and isolating $j$ in first, $E[\mathbf{u}^i_w|\pi^{i,\delta}=j, B^{i,\delta}_{j,\mathcal{T}}] = \mathbf{u}^i_j$. If, instead, $i$

abstains, not breaking the tie, $E[\mathbf{u}_w^i|\pi^{i,\delta}=\varnothing, B_{j,\mathcal{T}}^{i,\delta}] = (\mathbf{u}_j^i+\sum_{t\in\mathcal{T}}\mathbf{u}_t^i)/(|\mathcal{T}|+1)$. From Lemma 1, $E_j^{i,\delta} - E_\varnothing^{i,\delta} = \sum_{\mathcal{T}}\left[\alpha_{j,\mathcal{T}}^{i,\delta}\frac{\mathbf{u}_j^i+\sum_{t\in\mathcal{T}}\mathbf{u}_t^i}{|\mathcal{T}|+1} + \beta_{j,\mathcal{T}}^{i,\delta}\mathbf{u}_j^i\right] - \sum_{\mathcal{T}}\left[\alpha_{j,\mathcal{T}}^{i,\delta}\frac{\sum_{t\in\mathcal{T}}\mathbf{u}_t^i}{|\mathcal{T}|} + \beta_{j,\mathcal{T}}^{i,\delta}\frac{\mathbf{u}_j^i+\sum_{t\in\mathcal{T}}\mathbf{u}_t^i}{|\mathcal{T}|+1}\right] = \sum_{\mathcal{T}}\left(\frac{\alpha_{j,\mathcal{T}}^{i,\delta}}{|\mathcal{T}|}+\beta_{j,\mathcal{T}}^{i,\delta}\right)\left(\frac{\sum_{t\in\mathcal{T}}(\mathbf{u}_j^i-\mathbf{u}_t^i)}{|\mathcal{T}|+1}\right)$. □

# B Proof of Proposition 5

*Proof.* To shorten notation, let $\mathbf{\Lambda}\in\mathbb{Q}_{\geq 0}^m$ s.t. $\mathbf{\Lambda}_j=\mathbf{s}_j\lambda$, $\forall j$. Using Assumption 5, rewrite (2) as:

$\alpha_{j,\mathcal{T}}^i = \prod_{t\in\mathcal{T}}\Pr(\mathbf{n}_t^i=\mathbf{n}_j^i+1)\prod_{k\in\mathcal{K}}\Pr(\mathbf{n}_k^i\leq\mathbf{n}_j^i)$

$\beta_{j,\mathcal{T}}^i = \prod_{t\in\mathcal{T}}\Pr(\mathbf{n}_t^i=\mathbf{n}_j^i)\prod_{k\in\mathcal{K}}1-\Pr(\mathbf{n}_j^i\leq\mathbf{n}_k^i)$

Focusing on $\alpha_{j,\mathcal{T}}^i$, sum over all possible values of $\mathbf{\Lambda}_j$ to rewrite $\Pr(\mathbf{n}_t^i=\mathbf{n}_j^i+1)=\sum_{d=0}^{\infty}f_P(d,\mathbf{\Lambda}_j)f_P(d+1,\mathbf{\Lambda}_t)$ and $\Pr(\mathbf{n}_k^i\leq\mathbf{n}_j^i)=\sum_{d'=0}^{\infty}f_P(d',\mathbf{\Lambda}_j)F_P(d',\mathbf{\Lambda}_k)$. From Skellam (1946) we know $f_S(x,\mu_1,\mu_2)=\sum_{d=-\infty}^{\infty}f_P(d+x,\mu_1)f_P(d,\mu_2)$. Plugging it all in:

$$\alpha_{j,\mathcal{T}}^i=\prod_{t\in\mathcal{T}}f_S(-1,\mathbf{\Lambda}_j,\mathbf{\Lambda}_t)\prod_{k\in\mathcal{K}}\sum_{d'=0}^{\infty}f_P(d',\mathbf{\Lambda}_j)F_P(d',\mathbf{\Lambda}_k)$$

Substitute the well known formulae for $f_P$ and $F_P$:

$$\alpha_{j,\mathcal{T}}^i=\prod_{t\in\mathcal{T}}f_S(-1,\mathbf{\Lambda}_j,\mathbf{\Lambda}_t)\prod_{k\in\mathcal{K}}\sum_{d=0}^{\infty}\frac{e^{-\mathbf{\Lambda}_j}\mathbf{\Lambda}_j{}^d}{d!}\frac{\Gamma(d+1,\mathbf{\Lambda}_k)}{\Gamma(d+1)}$$

$$\alpha_{j,\mathcal{T}}^i=\prod_{t\in\mathcal{T}}f_S(-1,\mathbf{\Lambda}_j,\mathbf{\Lambda}_t)\prod_{k\in\mathcal{K}}1-F_{\chi^2}(2\mathbf{\Lambda}_k,-2(-1),2\mathbf{\Lambda}_j)$$

Where $F_{\chi^2}$ is the CDF of the Noncentral Chi-Square distribution. While $F_S$ has no closed formula, following Johnson (1959) it can be calculated as a function of $F_{\chi^2}$:

$$F_S(x,\mu_1,\mu_2)=\begin{cases}F_{\chi^2}(2\mu_2,-2x,2\mu_1) & \text{if } x<0\\ 1-F_{\chi^2}(2\mu_1,2(x+1),2\mu_2) & \text{if } x\geq 0\end{cases}$$

$$\alpha_{j,\mathcal{T}}^i=\prod_{t\in\mathcal{T}}f_S(-1,\mathbf{\Lambda}_j,\mathbf{\Lambda}_t)\prod_{k\in\mathcal{K}}1-F_S(-1,\mathbf{\Lambda}_j,\mathbf{\Lambda}_k)$$

Analogous for $\beta_{j,\mathcal{T}}^i$. See Online Appendix for details. □

# C Proof of Proposition 7

*Proof.* Let $1,\cdots,g,\cdots,m$ represent the rank of $m$ candidates' expected vote shares, with $g\in\mathbb{N}_{[2,m-1]}$.

*Case 1:* (all tied for 1st) Fix $\mathbf{s}_1^\delta=\cdots=\mathbf{s}_m^\delta$. Then, in $\delta$, $\alpha_{1,\mathcal{T}}^i=\cdots=\alpha_{m,\mathcal{T}}^i$ and $\beta_{1,\mathcal{T}}^i=\cdots=\beta_{m,\mathcal{T}}^i$, $\forall i$. Hence, given (1), $\pi^{i,\delta+1}=\operatorname{argmax}_{j\in\mathcal{J}}(\mathbf{u}_j^i)$, $\forall i$. If that leads to same vector $\mathbf{s}$, $\delta$ is a trivial PNE; otherwise it is Case 2, 3 or 4.

*Case 2:* (one in 1st, all else tied for 2nd) Fix $\mathbf{s}_1^\delta>\mathbf{s}_2^\delta=\cdots=\mathbf{s}_m^\delta$. Then, $\forall i$, if $\mathbf{u}_1^i>\mathbf{u}_2^i\geq\cdots\geq\mathbf{u}_m^i$, $E_{1,\delta}^i>E_2^{i,\delta}\geq\cdots\geq E_m^{i,\delta}$ regardless of pivotal probabilities; otherwise, note $\alpha_{2,\mathcal{T}}^i=\cdots=\alpha_{m,\mathcal{T}}^i$ and $\beta_{2,\mathcal{T}}^i=\cdots=\beta_{m,\mathcal{T}}^i$ in $\delta$. So, given (1), again $\pi^{i,\delta+1}=\operatorname{argmax}_{j\in\mathcal{J}}(\mathbf{u}_j^i)$, $\forall i$. Then, if $\mathbf{s}^{\delta+1}=\mathbf{s}^\delta$, $\delta$ is a trivial PNE; if not, it is Case 1, 3 or 4. Note that a cycle between Cases 1 and 2 is not possible. All electors voting sincerely can lead to Case 1 or Case 2.

*Case 3:* (2 candidates in the top 2) Fix $\min(\mathbf{s}_1^\delta,\mathbf{s}_2^\delta)>\mathbf{s}_h^\delta$, s.t. $3\leq h\leq m$. From Condition 1 and Assumption 3, it follows that as $N\to\infty$, given (1), $\min(E_1^{i,\delta},E_2^{i,\delta})>E_h^{i,\delta}$, $\forall i,\forall h$, i.e. $\lim_{N\to\infty}\Pr(\pi^{i,\delta+1}\in\{1,2\})=1$, $\forall i$, hence $\lim_{N\to\infty}\mathbf{s}_h^{\delta+1}=0$, $\forall h$. From Lemma 5, since either $\mathbf{u}_1^i-\mathbf{u}_2^i$ or $\mathbf{u}_2^i-\mathbf{u}_1^i$ is positive, $\lim_{N\to\infty}\Pr(\pi^{i,\delta+1}=\operatorname{argmax}_{1,2}(\mathbf{u}_1^i,\mathbf{u}_2^i))=1$, $\forall i$, which does not vary, thus ensuring a PNE in $\delta+1$.

*Case 4:* (up to $g-1$ tied for 1st) fix $\mathbf{s}_1^\delta=\cdots=\mathbf{s}_{g-1}^\delta$ and $\mathbf{s}_{g+1}^\delta\geq\cdots\geq\mathbf{s}_m^\delta\geq 0$, with either $\mathbf{s}_{g-1}^\delta>\mathbf{s}_g^\delta\geq\mathbf{s}_{g+1}^\delta$ or $\mathbf{s}_{g-1}^\delta\geq\mathbf{s}_g^\delta>\mathbf{s}_{g+1}^\delta$, where $g+1\leq h\leq m$. Again from Condition 1 and Assumption 3, as $N\to\infty$ and given (1), $\min(E_1^{i,\delta},\cdots,E_g^{i,\delta})>E_h^{i,\delta}$, $\forall i,\forall h$, i.e. $\lim_{N\to\infty}\Pr(\pi^{i,\delta+1}\in\{1,\cdots,g\})=1$, $\forall i$, hence $\lim_{N\to\infty}\mathbf{s}_h^{\delta+1}=0$, $\forall h$. Then, from Lemma 5, as $N\to\infty$, $\delta+1$ can be: (i) Case 3; (ii) Case 4 again, with up to $g$–1 candidates tied for 1st and $\mathbf{s}_h^{\delta+1}=0,\forall h$. If $\mathbf{s}^{\delta+2}=\mathbf{s}^{\delta+1}$ a trivial PNE is reached, otherwise $\delta+2$ is either Case 3 or Case 4 with even less candidates tied for 1st, and so forth until Case 3 is reached. (iii) Case 4 again, with $\mathbf{s}_1^{\delta+1}=\cdots=\mathbf{s}_g^{\delta+1}>0$, $\mathbf{s}_h^{\delta+1}=0,\forall h$, in which case $\delta+2$ is either (i), (ii) or a trivial PNE. □

# D Counting possible pivotal outcomes

**Proposition 8.** *Num. of cases in $\mathbf{V}^\alpha$ and $\mathbf{V}^\beta$ are given by:*

$$\begin{cases}\sum_{x=1}^{\hat{x}}\sum_{y=0}^{\hat{y}}\left[(-1)^y\binom{|\mathcal{K}|}{y}\binom{N-x(|\mathcal{T}|+1+y)+d+|\mathcal{K}|-1}{|\mathcal{K}|-1}\right] & \text{if } |\mathcal{T}|+1<m\\ \mathbb{1}(N+d \bmod m=0) & \text{if } |\mathcal{T}|+1=m\end{cases}$$

*where $d=0$ and $\hat{x}=\lfloor N/(|\mathcal{T}|+1)\rfloor$ when calculating $|\mathbf{V}^\beta|$, $d=1$ and $\hat{x}=\lceil N/(|\mathcal{T}|+1)\rceil$ when calculating $|\mathbf{V}^\alpha|$, with $\hat{y}=min(|\mathcal{K}|,\lfloor(N-x(|\mathcal{T}|+1)+d)/x\rfloor)$ and $\mathbb{1}$ being the indicator function.*

*Proof.* See Online Appendix. □

# E Proof of Algorithm 2

**Proposition 9.** Poisson pivotal probabilities can be jointly calculated and with no explicit calculation of $f_P$ or $F_P$.

*Proof.* Re-write $\beta_{j,\mathcal{T}}^i$ in (5) with the infinite summation starting at minus one, so all but one term of $\alpha_{j,\mathcal{T}}^i$ and $\beta_{j,\mathcal{T}}^i$ are the same:

$$\alpha_{j,\mathcal{T}}^i=\sum_{d=0}^{\infty}\left(f_P(d,\mathbf{n}_j^i)\prod_{t\in\mathcal{T}}f_P(d+1,\mathbf{n}_t^i)\prod_{k\in\mathcal{K}}F_P(d,\mathbf{n}_k^i)\right)$$

$$\beta_{j,\mathcal{T}}^i=\sum_{d=-1}^{\infty}\left(f_P(d+1,\mathbf{n}_j^i)\prod_{t\in\mathcal{T}}f_P(d+1,\mathbf{n}_t^i)\prod_{k\in\mathcal{K}}F_P(d,\mathbf{n}_k^i)\right)$$

Let $\forall\mu\in\mathbb{Q}_{\geq 0}$. Begin by noting that $f_P(0,\mu)=F_P(0,\mu)=e^{-\mu}$ and, by definition, $\forall x\in\mathbb{N}_{<0}$, $f_P(x,\mu)=F_P(x,\mu)=0$. Hence, the joint calculation of both infinite summations is initialized by calculating $e^{-\mathbf{n}_h^i}$, $\forall h\in\mathcal{J}$, then the infinite summations start at $d=-1$ and those exponentials are simply updated, up to joint convergence of $\alpha_{j,\mathcal{T}}^i$ and $\beta_{j,\mathcal{T}}^i$, according to the following: $\forall x\in\mathbb{N}_{\geq 0}$, since $f_P(x,\mu)=\frac{(\mu^x e^{-\mu})}{x!}$, then $f_P(x+1,\mu)=\frac{(\mu^{(x+1)}e^{-\mu})}{(x+1)!}=\frac{(\mu^x e^{-\mu})}{x!}\frac{\mu}{x+1}=f_P(x,\mu)\frac{\mu}{x+1}$. Similarly, if $F_P(x,\mu)=e^{-\mu}\sum_{r=0}^{\lfloor x\rfloor}\frac{\mu^r}{r!}$, then $F_P(x+1,\mu)=e^{-\mu}\sum_{r=0}^{\lfloor x+1\rfloor}\frac{\mu^r}{r!}=e^{-\mu}\sum_{r=0}^{\lfloor x\rfloor}\frac{\mu^r}{r!}+e^{-\mu}\frac{\mu^{(x+1)}}{(x+1)!}=F_P(x,\mu)+e^{-\mu}\frac{\mu^{(x+1)}}{(x+1)!}$. □

## Acknowledgments

We are thankful for crucial comments received from Reshef Meir, Ed Durfee, Walter Mebane and anonymous reviewers to earlier versions of this paper. We are also thankful for funding received from the MICDE Fellowship, from the Horowitz Foundation's Irving Louis Howard Award and Joshua Feigenbaum Award and from Walter Mebane's NSF award SES 1523355.

## References

Anonymous. 2018. Omitted title for anonymity .

Bartholdi, J.; Tovey, C.; and Trick, M. 1989. The computational difficulty of manipulating an election. *Social Choice and Welfare* 6(3).

Black, J. H. 1978. The Multicandidate Calculus of Voting: Application to Canadian Federal Elections. *American Journal of Political Science* 22(3): 609–638. ISSN 00925853, 15405907. URL http://www.jstor.org/stable/2110464.

Bouton, L. 2013. A theory of strategic voting in runoff elections. *American Economic Review* (103): 1248–1288.

Bouton, L.; and Gratton, G. 2015. Majority runoff elections: Strategic voting and Duverger's hypothesis. *Theoretical Economics* 10.

Chen, Y.; and Xia, A. 2011. The wasted vote phenomenon with uncertain voter population. *Social Choice Welfare* 37: 471–492.

Clough, E. 2007. Strategic Voting under Conditions of Uncertainty: A Re-Evaluation of Duverger's Law. *British Journal of Political Science* 37(2): 313–332. ISSN 00071234, 14692112. URL http://www.jstor.org/stable/4497293.

Conitzer, V.; Sandholdm, T.; and Lang, J. 2007. When are elections with few candidates hard to manipulate. *JACM* 54.

Conlisk, J. 1996. Why Bounded Rationality? *Journal of Economic Literature* 34(2): 669–700. ISSN 00220515. URL http://www.jstor.org/stable/2729218.

Cox, G. W. 1994. Strategic Voting Equilibria Under the Single Nontransferable Vote. *American Political Science Review* 88(3): 608–621. URL https://doi.org/10.2307/2944798.

Cox, G. W.; and Shugart, M. S. 1996. Strategic Voting Under Proportional Representation. *Journal of Law, Economics & Organization* 12(2): 299–324.

Cranor, L. F. 1996. *Declared-Strategy Voting: An Instrument for Group Decision-Making*. Ph.D. thesis, USA. AAI9719204.

Dhillon, A.; and Lockwood, B. 2004. When are plurality rule voting games dominance-solvable? *Games and Economic Behavior* 46: 55–75.

Duverger, M. 1963. *Political Parties: Their Organization and Activity in the Modern State*. New York: Wiley.

Elkind, E.; Markakis, E.; Obraztsova, S.; and Skowron, P. 2010. Equilibria of plurality voting with abstentions. In *Proceedings of the 11th Conference on Electronic Commerce (ACM-EC)*, 347–356.

Elkind, E.; Markakis, E.; Obraztsova, S.; and Skowron, P. 2015. Equilibria of Plurality Voting: Lazy and Truth-Biased Voters. In *Proceedings of the 8th Symposium on Algorithmic Game Theory (SAGT)*, 110–122.

Fabrikant, A.; Jaggard, A. D.; and Schapira, M. 2013. On the Structure of Weakly Acyclic Games. *Theory of Computing Systems* 53: 107–122.

Fairstein, R.; Lauz, A.; Meir, R.; and Gal, K. 2019. Modeling People's Voting Behavior with Poll Information. In *Proceedings of the 15th International Conference on Autonomous Agets and Multiagent Systems*. Canada: International Foundation for Autonomous Agents and Multiagent Systems.

Fey, M. 1997. Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and Strategic Voting. *American Political Science Review* 91(1): 135–147.

Gibbard, A. 1973. Manipulation of voting schemes: A general result. *Econometrica* 41(4): 587–601.

Hoffman, D. T. 1982. A Model for Strategic Voting. *SIAM Journal on Applied Mathematics* 42(4): 751–761.

Johnson, N. 1959. On an Extension of the Connexion Between Poisson and X2 Distributions. *Biometrika* 46(3/4): 352–363.

Koolyk, A.; Strangway, T.; Lev, O.; and Rosenschein, J. S. 2017. Convergence and Quality of Iterative Voting Under Non-Scoring Rules. In *International Joint Conference on Artificial Intelligence*, 273–279. Melbourne, Australia: International Foundation for Autonomous Agents and Multiagent Systems.

Laakso, M.; and Taagepera, R. 1979. Effective Number of Parties: A Measure with Application to West Europe. *Comparative Political Studies* 12.

Lev, O.; and Rosenschein, J. S. 2012. Convergence of Iterative Voting. In *Proceedings of the 11th International Conference on Autonomous Agets and Multiagent System*. Valencia, Spain: International Foundation for Autonomous Agents and Multiagent Systems.

Lev, O.; and Rosenschein, J. S. 2016. Convergence of Iterative Scoring Rules. *Journal of Artifi-cial Intelligence Research (JAIR)* (57): 573–591.

Meir, R. 2015. Plurality Voting under Uncertainty. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2103–2109. Buenos Aires, Argentina: Association for the Advancement of Artificial-Intelligence.

Meir, R. 2018. *Strategic Voting*. Cambridge, Massachusetts: Morgan & Claypool Publishers.

Meir, R.; Lev, O.; and Rosenschein, J. S. 2014. A local-dominance theory of voting equilibria. In *Proceedings of the 15th ACM conference on Economics and Computation*, 313–330. Palo Alto, California: ACM conference on Economics and Compu-tation.

Meir, R.; Polukarov, M.; Rosenschein, J. S.; and Jennings, N. R. 2010. Convergence to Equilibria in Plurality Voting.

In *Proceedings of The Twenty-Fourth National Conference on Artificial Intelligence*, 823–828. Atlanta, Georgia: International Foundation for Autonomous Agents and Multiagent Systems.

Myatt, D. P. 2007. On the Theory of Strategic Voting. *Review of Economic Studies* (74): 255–281.

Myerson, R. 1998. Population Uncertainty and Poisson games. *International Journal of Game Theory* (27): 375–392.

Myerson, R. 2000. Large Poisson Games. *Journal of Economic Theory* 94(1): 7–45.

Myerson, R.; and Weber, R. J. 1993. A Theory of Voting Equilibria. *American Political Science Review* 87(1): 102–114. URL https://doi.org/10.2307/2944798.

Obraztsova, S.; Markakis, E.; and Thompson, D. R. M. 2013. Plurality Voting with Truth-Biased Agents. In *Proceedings of the 6th Symposium on Algorithmic Game Theory (SAGT)*, 26–37.

Palfrey, T. 1988. *Essays in Contemporary Political Theory*, chapter: A Mathematical Proof of Duverger's Law, 51–74. Ann Arbor, MI, USA: University of Michigan Press. ISBN 1-59140-056-2.

Reijngoud, A.; and Endriss, U. 2012. Voter Response to Iterated Poll Information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*. Valencia, Spain: International Conference on Autonomous Agents and Multiagent Systems.

Riker, W.; and Ordeshook, P. 1968. A Theory of the Calculus of Voting. *American Political Science Review* 62(1): 25–42.

Satterthwaite, M. A. 1975. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory* 10(2): 187–217.

Skellam, J. G. 1946. The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society* 109(3).

Smith, A.; de Mesquita, B. B.; and LaGatta, T. 2016. Group incentives and rational voting. *Journal of Theoretical Politics* 29(2).

Tsang, A.; and Larson, K. 2016. The Echo Chamber: Strategic Voting and Homophily in Social Networks. In *Proceedings of the 15th International Conference on Autonomous Agets and Multiagent Systems*, 368–375. Singapore: International Foundation for Autonomous Agents and Multiagent Systems.

Tsang, A.; Salehi-Abari, A.; and Larson, K. 2018. Boundedly rational voters in large (r) networks. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, 301–308. Singapore: International Foundation for Autonomous Agents and Multiagent Systems.

Wilczynski, A. 2019. Poll-Confident Voters in Iterative Voting. In *Proceedings of The Thirty-Third AAAI Conference on Artificial Intelligence*, 2205–2212. Honolulu, Hawaii: Association for the Advancement of Artificial Intelligence.

Zilberstein, S. 2008. Metareasoning and Bounded Rationality. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence*. Chicago, US: Association for the Advancement of Artificial Intelligence.