# Probability of a Condorcet Winner for Large Electorates: An Analytic Combinatorics Approach

Emma Caizergues, Élie de Panafieu, François Durand, Marc Noy, Vlady Ravelomanana

**Abstract**

We study the probability that a given candidate is an $\boldsymbol{\alpha}$-*winner*, *i.e.,* a candidate preferred to each other candidate $j$ by a fraction $\alpha_j$ of the voters. This extends the classical notion of *Condorcet winner*, which corresponds to the case $\boldsymbol{\alpha} = (\frac{1}{2}, \dots, \frac{1}{2})$. Our analysis is conducted under the general assumption that voters have independent preferences, illustrated through applications to well-known models such as Impartial Culture and the Mallows model. While previous works use probabilistic arguments to derive the limiting probability as the number of voters tends to infinity, we employ techniques from the field of analytic combinatorics to compute convergence rates and provide a method for obtaining higher-order terms in the asymptotic expansion. In particular, we establish that the probability of a given candidate being the Condorcet winner in Impartial Culture is $a_0 + a_{1,n} n^{-1/2} + \mathcal{O}(n^{-1})$, where we explicitly provide the values of the constant $a_0$ and the coefficient $a_{1,n}$, which depends solely on the parity of the number of voters $n$. Along the way, we derive technical results in multivariate analytic combinatorics that may be of independent interest.

## 1 Introduction

**Motivation** A long-standing tradition in social choice theory is to compute the probability that a *Condorcet winner* exists, *i.e.,* a candidate preferred to every other candidate by a majority of voters. A comprehensive overview of this research is presented in Gehrlein [12], which is entirely dedicated to this topic. Niemi and Weisberg [26] derived the limit of this probability as the number of voters tends to infinity, under the only assumption of independent preferences. More recently, Krishnamoorthy and Raghavachari [20] revisited and modernized this result.

However, to the best of our knowledge, all previous studies have relied on classical combinatorial or probabilistic methods. For instance, the results by Niemi and Weisberg [26] and Krishnamoorthy and Raghavachari [20] are primarily based on a Gaussian approximation. In this paper, we introduce techniques from *analytic combinatorics* into this domain, drawing on the approach outlined in the reference book by Flajolet and Sedgewick [8]. These methods encode combinatorial problems into analytic objects, allowing the application of powerful analytical tools to extract information, especially limits, but also speeds of convergence, and more generally asymptotic expansions as certain parameters tend to infinity.

**Contributions** Similarly to Niemi and Weisberg [26] and Krishnamoorthy and Raghavachari [20], we consider a general probability distribution over voter preferences which we call *General Independent Culture (GIC)*, with the primary assumption that individual preferences are independent. Additionally, we assume that the distribution is *generic*, in the sense that every possible ranking over the candidates has a positive probability.

To obtain results as general as possible, we introduce the notion of an $\boldsymbol{\alpha}$-*winner*, a generalization of the Condorcet winner, where the required victory threshold for each pairwise comparison is not necessarily one-half of the voters.

We analyze the probability that a given candidate is an $\boldsymbol{\alpha}$-winner, with a particular focus on its asymptotic behavior as the number of candidates $m$ remains fixed while the number of voters $n$ tends to infinity. Our main contributions lie in providing a method that not only recovers previous results on limiting probabilities using a different approach but also enables the computation of the rate of convergence and higher-order terms in the asymptotic expansion.

These general theoretical results are illustrated through applications to the usual notion of Condorcet winner under probabilistic models known as the *Impartial Culture* and the *Mallows model*. In particular, we establish that the probability of a given candidate being the Condorcet winner in Impartial Culture is $a_0 + a_{1,n} n^{-1/2} + \mathcal{O}(n^{-1})$, where we explicitly provide the values of the constant $a_0$ and the coefficient $a_{1,n}$, which depends solely on the parity of the number of voters $n$.

In the process, we also establish results in multivariate analytic combinatorics that are of independent interest.

This paper is a shortened version of a manuscript currently under submission to the Twenty-Sixth ACM Conference on Economics and Computation (EC'25), also available on arXiv [4].

**Related Work** Regarding the probability of a Condorcet winner, the standard reference is Gehrlein [12]. In the General Independent Culture (GIC), which assumes independent voter preferences, the cases for finite values of voters $n$ and candidates $m$ were explored by Gehrlein and Fishburn [14], Gillett [16], and Gehrlein [10]. The most relevant works to ours are Niemi and Weisberg [26] and Krishnamoorthy and Raghavachari [20], which derive the limiting probability as $n$ tends to infinity in this general setting. Our approach differs in methodology and results: we use analytic combinatorics and provide more precise asymptotic behaviors.

Among the specific cases of GIC, the most studied is *Impartial Culture (IC)*, where all rankings are equally likely [12]. For finite values of $n$ and $m$ or for the asymptotic regime where $m \to \infty$, see Sauermann [27]. Other works, like ours, focus on large electorates (*i.e.,* $n \to \infty$). An explicit formula exists for $m = 3$ [17, 26, 9], while the limit for $m = 4$ follows from a recurrence relation [25, 7]. Explicit formulas for $m \in \{5, 6, 7\}$ have been derived in [15, 10]. The general case of arbitrary $m$ with $n \to \infty$ was addressed by Niemi and Weisberg [26] and Krishnamoorthy and Raghavachari [20]. In IC, as in the broader GIC framework, our main contribution is to provide the rate of convergence and a method for computing higher-order terms in the asymptotic expansion.

Other specific cases of GIC have been studied, notably the *Dual Culture* [12] and the *Perturbed Culture* [29, 12]. Although the *Mallows model* [24] is popular, the probability of a Condorcet winner in this framework has received little attention. This is likely because, for large electorates, the limiting behavior is trivial: the candidate favored by the culture becomes the Condorcet winner with probability tending to 1. The question of the convergence rate, which we explore here, is much more interesting.

More distant from our work, other models of culture fall outside the GIC framework as they do not assume voter independence. Notable examples include *Impartial Anonymous Culture (IAC)* [13, 21, 12] and, more generally, *Pólya-Eggenberger models* [2, 12].

In Impartial Anonymous Culture, many voting questions [11, 18, 22, 30] have been studied using Ehrhart polynomials [1], reducing the enumeration of voting configurations to counting integral points in a polyhedron. Maassen and Bezembinder [23] use generating functions to express the probability of a Condorcet Winner in Impartial Culture or its variant with weak orders, obtaining exact expressions and asymptotics for several cases. However, none of these works apply complex analysis to derive asymptotics. May [25] uses the saddle-point method to study the asymptotics as $m \to \infty$ of the limit as $n \to \infty$ of the probability that a candidate is the Condorcet Winner in Impartial Culture. This double limit problem is not addressed in this paper and leads to very different computations, involving only a univariate real integral.

**Limitations** A limitation of our work is that we focus on the probability that a *given* candidate is an $\boldsymbol{\alpha}$-winner, rather than the probability that at least one exists. For some values of $\boldsymbol{\alpha}$, several candidates may satisfy this condition simultaneously, requiring inclusion-exclusion techniques to compute the total probability. We leave this issue out of scope, as it does not arise for the Condorcet winner, who is unique when they exist.

Moreover, our main results concern the asymptotic behavior as $n$ tends to infinity. While our approach also provides exact expressions for finite $n$, these expressions are essentially the same as those obtained using classical combinatorial methods. The main advantage of our method, therefore, lies in its ability to analyze asymptotic behavior.

To avoid degenerate cases, we also assume that the preference distribution is *generic*, meaning that all rankings have positive probability. While analytic combinatorics methods could be applied without this assumption, doing so would introduce additional subcases that we prefer to avoid, where the *saddle point* (which we will define shortly) has null or infinite coordinates.

**Roadmap** Section 2 presents the necessary preliminaries. Section 3 translates our combinatorial problem into an analytic question and introduces the notion of *saddle point*. Section 4 presents our main findings, and Section 5 outlines directions for future research. The full version of the paper [4] also includes appendices that provide complete technical details. The paper is accompanied by the Python package `Actinvoting` (Analytic Combinatorics Tools In Voting), which was used for symbolic computations and numerical simulations [6].

## 2 Preliminaries

In this section, we begin by introducing the main definitions and notations, and we formulate our research question. We then briefly present the field of analytic combinatorics.

### 2.1 Definitions, Notations, and Research Question

The number of voters is denoted by $n$. We define the set of candidates as $\{1, \dots, m\}$, where $m$ denotes the number of candidates. A *profile* represents the preferences of the voters and is defined as a list of $n$ rankings over the set of candidates. A *culture* refers to a probability distribution over the $(m!)^n$ possible profiles. We assume that voters have independent preferences, meaning the culture is characterized by a probability $p_r$ for each preference ranking $r$. Surprisingly, there is no standard terminology for this common assumption, which we refer to as a *General Independent Culture (GIC)*. We also assume that the culture is *generic*, in the sense that $p_r > 0$ for every ranking $r$. The notation $\mathbb{P}$ refers to the probability under the given culture.

A *Condorcet winner* is a candidate who is preferred to every other candidate by more than $\frac{n}{2}$ voters. Without loss of generality, we study the probability $\mathbb{P}(m \text{ is CW})$ that candidate $m$ is the Condorcet winner. Our theoretical results extend this question to a more general notion, which we call an $\boldsymbol{\alpha}$*-winner*, defined as follows. Let $\mathcal{A} := \{1, \dots, m-1\}$ be the set of adversaries of candidate $m$. Consider a vector $\boldsymbol{\alpha} := (\alpha_1, \dots, \alpha_{m-1}) \in (0,1)^{m-1}$, where each $\alpha_j$ represents the proportion of voters that candidate $m$ needs on their side to win the pairwise comparison against candidate $j$. For an adversary $j \in \mathcal{A}$, we write $m \succ_{\boldsymbol{\alpha}} j$, and we say that $m$ wins against $j$ in the sense of $\boldsymbol{\alpha}$, if candidate $m$ is preferred to $j$ by more than $\alpha_j n$ voters. For a subset of adversaries $\mathcal{X} = \{j_1, \dots, j_k\} \subseteq \mathcal{A}$, we write $m \succ_{\boldsymbol{\alpha}} \mathcal{X}$, or equivalently $m \succ_{\boldsymbol{\alpha}} j_1, \dots, j_k$, if $m \succ_{\boldsymbol{\alpha}} j$ holds for every $j \in \mathcal{X}$. Similarly, we write $m \preccurlyeq_{\boldsymbol{\alpha}} j$, and we say that $m$ does not win against $j$ in the sense of $\boldsymbol{\alpha}$, if candidate $m$ is preferred to $j$ by at most $\alpha_j n$ voters. We say that candidate $m$ is an $\boldsymbol{\alpha}$*-winner* if $m \succ_{\boldsymbol{\alpha}} \mathcal{A}$, *i.e.,* $m$ is preferred to each adversary $j$ by more than $\alpha_j n$ voters. The standard notion of a Condorcet winner corresponds to the special case

$\boldsymbol{\alpha} = \left(\frac{1}{2}, \ldots, \frac{1}{2}\right)$. The notion of a *generalized Condorcet winner*, as introduced by Sertel and Sanver [28], is recovered by considering a vector $\boldsymbol{\alpha}$ whose coordinates are equal.

The goal of this paper is to study $\mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner})$, the probability that candidate $m$ is an $\boldsymbol{\alpha}$-winner, with a particular focus on its asymptotic behavior as the number of voters $n$ tends to infinity. Note that the probability of the existence of a Condorcet winner can be obtained as the sum of the probabilities for all $m$ candidates.

Our theoretical results will be illustrated using the *Mallows model* [24]. This model is characterized by a reference ranking $r_0$ over the candidates and a concentration parameter $\rho \in \mathbb{R}_{\geq 0}$. The probability of a ranking $r$ is given by

$$p_r := \gamma e^{-\rho d(r, r_0)},$$

where $\gamma$ is a normalization constant ensuring $\sum_r p_r = 1$, and $d(r, r_0)$ denotes the Kendall-tau distance [19], which counts the minimum number of adjacent swaps needed to transform $r$ into $r_0$. The Mallows model is commonly used in the field of *epistemic democracy* [3, Chapters 8 and 10], where noisy evaluations from voters are collected with the aim of uncovering a hidden truth. The reference ranking models the hidden truth, and the concentration parameter $\rho$ indicates the skill level of the voters. When $\rho = 0$, all rankings are equally probable, recovering the classical *Impartial Culture* model. We denote by $\mathcal{M}_{m \text{ last}}$ (resp. $\mathcal{M}_{m \text{ first}}$) a Mallows culture with parameter $\rho$, where the reference ranking $r_0$ places candidate $m$ last (resp. first). We can assume, without loss of generality, that the reference ranking is $(1, \ldots, m)$ (resp. $(m, \ldots, 1)$).

Finally, we typeset vectors in boldface, *e.g.*, $\boldsymbol{x} := (x_1, \ldots, x_{m-1})$. If $\mathcal{X}$ is a set of indices, we define $\boldsymbol{x}_{\mathcal{X}} := (x_j)_{j \in \mathcal{X}}$. We define $\log(\boldsymbol{x}) := (\log(x_1), \ldots, \log(x_{m-1}))$, and similarly for the exponential. We write $\boldsymbol{u} \leq \boldsymbol{v}$ if $u_j \leq v_j$ for every coordinate $j$, with similar notation for strict inequalities. We denote $\mathbf{0} := (0, \ldots, 0)$ and $\mathbf{1} := (1, \ldots, 1)$, where the vector's size is understood from context. The diagonal matrix with diagonal elements $u_1, u_2, \ldots$ is denoted $\operatorname{diag}(\boldsymbol{u})$. Finally, for convenience, we set $\boldsymbol{\beta} := \mathbf{1} - \boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is the vector of victory thresholds in the definition of an $\boldsymbol{\alpha}$-winner.

### 2.2 A Short Introduction to Analytic Combinatorics

The techniques employed in this paper come from the field of analytic combinatorics [8]. Within this framework, formal variables are introduced and linked to specific parameters that define a particular counting problem. Once these variables are properly introduced, a formal power series is defined as a function of these variables, called the *generating function*. It serves as a formal notation that succinctly represents the combinatorial structure of the problem. Furthermore, it can be interpreted as a function of complex variables, which can be studied with the powerful tools of complex analysis, hence giving insights into the original combinatorial problem.

To illustrate this, consider a classical combinatorial example: counting binary trees with $k$ vertices. The problem is first encoded as a formal power series by introducing a variable $z$ marking the number of vertices. The generating function $T(z)$ is then defined as

$$T(z) := \sum_{k=0}^{\infty} T_k z^k,$$

where $T_k$ is the number of binary trees with $k$ vertices. The *symbolic method* then translates the combinatorial structure into analytic properties. In this example, whose detailed analysis is beyond the scope of this paper, the combinatorial property that a binary tree is either an isolated leaf or a root with two binary subtrees translates into the equation $T(z) = z + zT(z)^2$. To determine $T_k$, we solve this equation and extract the coefficient of $z^k$ in the Taylor expansion of $T(z)$, which is denoted by the *coefficient extraction* $[z^k]T(z)$. While this example is solved using an algebraic equation, other cases may involve different analytical techniques, such as differential equations or complex calculus.

## 3 From Our Combinatorial Problem to Its Analytic Formulation

As for binary trees, we first encode the problem of determining the probability that a candidate is an $\boldsymbol{\alpha}$-winner. However, instead of a univariate infinite series, we now work with a multivariate polynomial. We then show that the coefficient extractions can be achieved through *Cauchy integrals*. Finally, to analyze their asymptotic behavior for large electorates, we introduce the general principle of the *saddle-point method*. Its different sub-cases are later developed in Section 4.

### 3.1 Symbolic Method

As a warm-up, consider the case $m = 3$. To determine whether candidate 3 is the Condorcet winner, or more generally an $\boldsymbol{\alpha}$-winner, it suffices to know whether each voter ranks candidate 1 and/or 2 above candidate 3, rather than their full ranking. Thus, we introduce formal variables $x_1$ (respectively $x_2$) that indicates when a voter prefers candidate 1 (respectively 2) over candidate 3.

To represent the probability distribution governing the preferences of a single voter, we introduce the characteristic polynomial $P(x_1, x_2)$, as illustrated in Figure 1. In this polynomial, the probability of each ranking is multiplied by the appropriate formal variables, depending on whether candidate 1 and/or 2 is preferred over candidate 3. At this stage, the characteristic polynomial may appear to be merely a compact way of representing the probability distribution of a single voter.

| $p_{123}$ | $p_{132}$ | $p_{213}$ | $p_{231}$ | $p_{312}$ | $p_{321}$ |
|---|---|---|---|---|---|
| $1$<br>$2$<br>$\boxed{3}$ | $1$<br>$\boxed{3}$<br>$2$ | $2$<br>$1$<br>$\boxed{3}$ | $2$<br>$\boxed{3}$<br>$1$ | $\boxed{3}$<br>$1$<br>$2$ | $\boxed{3}$<br>$2$<br>$1$ |
| $x_1 x_2$ | $x_1$ | $x_1 x_2$ | $x_2$ | $1$ | $1$ |

$$P(x_1, x_2) := p_{123} \cdot x_1 x_2 + p_{132} \cdot x_1 + p_{213} \cdot x_1 x_2 + p_{231} \cdot x_2 + p_{312} \cdot 1 + p_{321} \cdot 1$$

**Figure 1:** Definition of the characteristic polynomial $P(x_1, x_2)$ encoding the probability distribution for a single voter when $m = 3$. The notation $p_{123}$, for example, is a shorthand for $p_{(1,2,3)}$. The formal variable $x_j$ marks rankings where candidate $j$ is preferred to candidate 3.

Note that the characteristic polynomial can be rewritten as

$$P(x_1, x_2) = p_{\emptyset} + p_{\{1\}} x_1 + p_{\{2\}} x_2 + p_{\{1,2\}} x_1 x_2, \tag{1}$$

where, for example, $p_{\{1,2\}} := p_{123} + p_{213}$ is the probability that a voter ranks both candidates 1 and 2 above candidate 3. This probability can also be expressed as a multivariate coefficient extraction, denoted by $[{x_1}^1 {x_2}^1] P(x_1, x_2)$.

More generally, for an arbitrary value of $m$ and a subset of adversaries $\mathcal{X} \subseteq \mathcal{A}$, we overload the notation $p$ by defining $p_{\mathcal{X}}$ as the total probability that in a random ranking, the set of adversaries above $m$ is exactly $\mathcal{X}$. We then encode the probability distribution for a single voter as follows.

**Definition 1.** *The* characteristic polynomial *is defined as*

$$P(\boldsymbol{x}) := \sum_{\mathcal{X} \subseteq \mathcal{A}} \Big( p_{\mathcal{X}} \prod_{j \in \mathcal{X}} x_j \Big).$$

Note that $\sum_r p_r = 1$ implies that $P(\mathbf{1}) = 1$.

Now, let us examine what happens with several voters, beginning with the case with $m = 3$ candidates and $n = 2$ voters. The core observation, illustrated in Figure 2, is that the algebraic operation of developing $P(x_1, x_2)^2$ is isomorphic to the probability tree associated with the preferences of two voters. For example, the probability that, among the two voters, exactly two prefer candidate 1 over candidate 3, and exactly one prefers candidate 2 over candidate 3 is given by the coefficient extraction $[{x_1}^2 {x_2}^1]P(x_1, x_2)^2 = 2p_{\{1\}}p_{\{1,2\}}$.

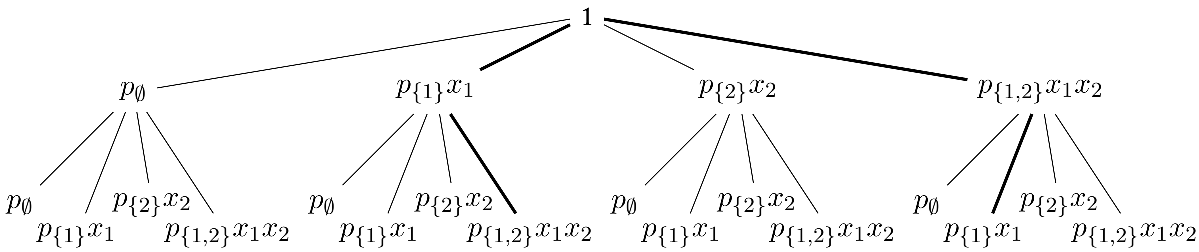


**Figure 2:** Tree representing the algebraic expansion of $P(x_1, x_2)^2$. The edges correspond to multiplications, and each path from the root to a leaf represents a term in the expansion. For example, the highlighted paths correspond to the terms involving ${x_1}^2{x_2}^1$, with a total coefficient given by $[{x_1}^2{x_2}^1]P(x_1, x_2)^2 = 2p_{\{1\}}p_{\{1,2\}}$. This coefficient extraction corresponds to the probability that, among the two voters, exactly two prefer candidate 1 over candidate 3, and exactly one prefers candidate 2 over candidate 3.

More generally, when $m$ and $n$ are arbitrary, for a vector $\boldsymbol{\ell} \in \mathbb{N}^{m-1}$, the coefficient extraction

$$[\boldsymbol{x}^{\boldsymbol{\ell}}]P(\boldsymbol{x})^n := [{x_1}^{\ell_1} \cdots {x_{m-1}}^{\ell_{m-1}}]P(x_1, \ldots, x_{m-1})^n$$

corresponds to the probability that each adversary $j$ is preferred to candidate $m$ by exactly $\ell_j$ voters.

Now, for $m$ to be an $\boldsymbol{\alpha}$-winner, every adversary $j$ must be preferred to $m$ by less than $(1 - \alpha_j)n = \beta_j n$ voters. Summing over all such cases, the probability that candidate $m$ is an $\boldsymbol{\alpha}$-winner is given by

$$\mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner}) = [\boldsymbol{x}^{<\boldsymbol{\beta} n}]P(\boldsymbol{x})^n := \sum_{\boldsymbol{\ell} < \boldsymbol{\beta} n} [\boldsymbol{x}^{\boldsymbol{\ell}}]P(\boldsymbol{x})^n.$$

However, the coordinates of $\boldsymbol{\beta} n$ are not necessarily integers, and we prefer to express the summation bounds in terms of integers:

$$\mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner}) = [\boldsymbol{x}^{\leq \lceil \boldsymbol{\beta} n \rceil - 1}]P(\boldsymbol{x})^n := \sum_{\boldsymbol{\ell} \leq \lceil \boldsymbol{\beta} n \rceil - 1} [\boldsymbol{x}^{\boldsymbol{\ell}}]P(\boldsymbol{x})^n. \tag{2}$$

Note that if we define a *weak $\boldsymbol{\alpha}$-winner* as a candidate preferred to any other candidate $j$ by *at least* $\alpha_j n$ voters, then all the results of this paper extend to weak $\boldsymbol{\alpha}$-winners by replacing $\lceil \boldsymbol{\beta} n \rceil - 1$ with $\lfloor \boldsymbol{\beta} n \rfloor$. In particular, our findings for the Condorcet winner naturally extend to the *weak Condorcet winner* (defined analogously) by replacing all occurrences of $\lceil n/2 \rceil - 1$ with $\lfloor n/2 \rfloor$.

### 3.2 Coefficient Extraction via Cauchy Integrals

In analytic combinatorics, it is common to extract the coefficient $f_\ell$ of a series $F(z) = \sum_k f_k z^k$ with a positive radius of convergence by expressing it as a Cauchy integral [8, Th. IV.4]:

$$[z^\ell]F(z) = \frac{1}{2i\pi} \oint F(z) \frac{dz}{z^{\ell+1}}. \tag{3}$$

The symbol $\oint$ indicates that the integral is taken over a positively oriented loop around 0 in the complex plane. To intuitively understand this equality, one can expand $F(z)$ into its series, interchange sum and integral, and integrate over a circle of small radius centered at 0. The residue theorem (see *e.g.*, Flajolet and Sedgewick [8, Th. IV.3]) then implies that all terms of the form $z^{k-\ell-1}$ result in a zero integral unless $k=\ell$. Now, to extract $[z^{\leq L}]F(z) := f_0 + \cdots + f_L$, we apply

$$[z^{\leq L}]F(z) = [z^L]\frac{F(z)}{1-z}, \tag{4}$$

which is obtained by remarking that $\frac{F(z)}{1-z} = \left(\sum_k f_k z^k\right)\left(\sum_k z^k\right) = \sum_k \left(\sum_{\ell=0}^{k} f_\ell\right) z^k$.

Equations (3) and (4) extend to the multivariate case. Applying them to Equation (2) gives

$$\mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner}) = \frac{1}{(2i\pi)^{m-1}} \oint \frac{P(\boldsymbol{x})^n}{\prod_{j=1}^{m-1}(1-x_j)} \frac{d\boldsymbol{x}}{\prod_{j=1}^{m-1} x_j^{\lceil \beta_j n \rceil}}, \tag{5}$$

where $\oint d\boldsymbol{x}$ is a shorthand for $\oint \cdots \oint dx_1 \dots dx_{m-1}$.

### 3.3 Saddle Point Method

We study the asymptotic behavior of our complex integrals as $n \to +\infty$. To build intuition, let us momentarily disregard the ceiling function in Equation (5) and observe that

$$\oint \frac{P(\boldsymbol{x})^n}{\prod_{j=1}^{m-1}(1-x_j)} \frac{d\boldsymbol{x}}{\prod_{j=1}^{m-1} x_j^{\beta_j n}} = \oint \frac{e^{-n(-\log(P(\boldsymbol{x}))+\boldsymbol{\beta}^\intercal \log(\boldsymbol{x}))}}{\prod_{j=1}^{m-1}(1-x_j)} d\boldsymbol{x} = \oint A(\boldsymbol{x}) e^{-n\psi(\log(\boldsymbol{x}))} d\boldsymbol{x},$$

for an appropriately defined function $A(\boldsymbol{x})$ and $\psi(\boldsymbol{t}) = -\log(P(e^{\boldsymbol{t}})) + \boldsymbol{\beta}^T \boldsymbol{t}$, where $\boldsymbol{t} = \log(\boldsymbol{x})$.

To analyze the asymptotic behavior of such integrals, it is standard to apply the *saddle-point method* [8, Chapter VIII]. The key idea is to find a contour of integration where, as $n$ approaches infinity, the integral's dominant contribution arises from the neighborhood of a specific point, while the contribution from the rest of the contour becomes negligible in comparison.

For this approach to be valid, the chosen point must satisfy the condition that the gradient of the function $\psi$ vanishes. In the univariate case, since a holomorphic function cannot exhibit local extrema in modulus, the graph of the function's modulus at such a point resembles a saddle. This is why it is called a *saddle point*, even in the general multivariate case.

In our case, the presence of the integer parts $\lceil \beta_j n \rceil$ does not fundamentally alter the method. The necessary adaptations are detailed in the appendices of the full paper [4]. The absolute value of $e^{-n\psi(\log(\boldsymbol{x}))}$ at the saddle point is minimal on the real line and maximal on the integration circle. These considerations lead to define the following objects, which we will use extensively in the rest of the paper.

**Definition 2.** *The* cumulant generating function *of $P$ is*

$$K : \boldsymbol{t} \in \mathbb{R}^{m-1} \mapsto \log\left(P(e^{\boldsymbol{t}})\right).$$

*Observing that the function $\psi : \boldsymbol{t} \mapsto -K(\boldsymbol{t}) + \boldsymbol{\beta}^\intercal \boldsymbol{t}$ is strictly concave on $\mathbb{R}^{m-1}$ (cf. appendices of the full paper [4]), we define the* log saddle point *$\boldsymbol{\tau}$ and the* saddle point *$\boldsymbol{\zeta}$ as*

$$\boldsymbol{\tau} := \underset{\boldsymbol{t} \in \mathbb{R}^{m-1}}{\arg\max} \left(-K(\boldsymbol{t}) + \boldsymbol{\beta}^\intercal \boldsymbol{t}\right), \qquad \boldsymbol{\zeta} = e^{\boldsymbol{\tau}}.$$

*Remark that this implies that $\boldsymbol{\zeta}$ minimizes the function $\boldsymbol{x} \in (\mathbb{R}_{>0})^{m-1} \mapsto \frac{P(\boldsymbol{x})}{\prod_j {x_j}^{\beta_j}}$.*

*To approximate the integral in Equation (5), a key ingredient is the Taylor expansion of $\psi$ at its saddle point. Consequently, we will need the Hessian of $\psi$ at $\boldsymbol{\tau}$. Since the term $\boldsymbol{\beta}^\intercal \boldsymbol{t}$ is linear, this Hessian simplifies to $\mathcal{H}_K(\boldsymbol{\tau})$, the Hessian of $K(\boldsymbol{t})$ at $\tau$, which can be computed either directly or as*

$$\mathcal{H}_K(\boldsymbol{\tau}) = \operatorname{diag}(\boldsymbol{\zeta}) \frac{\mathcal{H}_P(\boldsymbol{\zeta})}{P(\boldsymbol{\zeta})} \operatorname{diag}(\boldsymbol{\zeta}) + \operatorname{diag}(\boldsymbol{\beta}) - \boldsymbol{\beta}\boldsymbol{\beta}^T. \tag{6}$$

**Probabilistic interpretation.**

There is a nice probabilistic interpretation for the cumulant generating function. For any vector parameter $\boldsymbol{s} \in \mathbb{R}^{m-1}$, we define the random vector $\boldsymbol{X}^{(\boldsymbol{s})}$ as follows. For any subset $\mathcal{X} \subseteq \mathcal{A}$ of adversaries whose indicator vector is denoted $\boldsymbol{\ell} \in \{0,1\}^{m-1}$:

$$\mathbb{P}(\boldsymbol{X}^{(\boldsymbol{s})} = \boldsymbol{\ell}) := \frac{[\boldsymbol{x}^{\boldsymbol{\ell}}]P(\boldsymbol{x})e^{\boldsymbol{\ell}^T\boldsymbol{s}}}{P(e^{\boldsymbol{s}})} = \frac{p_{\mathcal{X}}e^{\boldsymbol{\ell}^T\boldsymbol{s}}}{P(e^{\boldsymbol{s}})}.$$

Note that $\boldsymbol{X}^{(\boldsymbol{0})}$ corresponds to the original probability distribution, as encoded by $P$. For $\boldsymbol{s} \neq \boldsymbol{0}$, the distribution of $\boldsymbol{X}^{(\boldsymbol{s})}$ represents a perturbation of the original culture.

The *cumulant generating function* of $\boldsymbol{X}^{(\boldsymbol{s})}$ is classically defined as

$$K_{\boldsymbol{s}} : \boldsymbol{t} \mapsto \log(\mathbb{E}(e^{\boldsymbol{t}^T\boldsymbol{X}^{(\boldsymbol{s})}})).$$

Remark that $K = K_{\boldsymbol{0}}$. It is well known that the gradient and Hessian matrix of this function at $\boldsymbol{t} = \boldsymbol{0}$ are equal to the mean and covariance matrix of $\boldsymbol{X}^{(\boldsymbol{s})}$. Actually, all the information contained in $K_{\boldsymbol{s}}$ can be retrieved through $K$ thanks to the following observation.

$$\begin{aligned} K_{\boldsymbol{s}}(\boldsymbol{t}) &= \log\left(\sum_{\boldsymbol{\ell}} \frac{([\boldsymbol{x}^{\boldsymbol{\ell}}]P(\boldsymbol{x}))e^{\boldsymbol{\ell}^T\boldsymbol{s}}}{P(e^{\boldsymbol{s}})} e^{\boldsymbol{\ell}^T\boldsymbol{t}}\right) = \log\left(\sum_{\boldsymbol{\ell}} ([\boldsymbol{x}^{\boldsymbol{\ell}}]P(\boldsymbol{x}))e^{\boldsymbol{\ell}^T(\boldsymbol{s}+\boldsymbol{t})}\right) - \log(P(e^{\boldsymbol{s}})) \\ &= \log(P(e^{\boldsymbol{s}+\boldsymbol{t}})) - \log(P(e^{\boldsymbol{s}})) = K(\boldsymbol{s}+\boldsymbol{t}) - \log(P(e^{\boldsymbol{s}})), \end{aligned}$$

so the mean and covariance matrix of $\boldsymbol{X}^{(\boldsymbol{s})}$ are respectively equal to the gradient and the Hessian of $K(\boldsymbol{t})$ at $\boldsymbol{t} = \boldsymbol{s}$.

Since $\boldsymbol{\tau}$ is the vector where the gradient of $\psi : \boldsymbol{t} \mapsto -K(\boldsymbol{t}) + \boldsymbol{\beta}^T\boldsymbol{t}$ vanishes, it follows that the gradient of $K$ at $\boldsymbol{\tau}$, *i.e.*, the mean of $X^{(\boldsymbol{\tau})}$, is equal to $\boldsymbol{\beta}$. In other words, the distribution of $X^{(\boldsymbol{\tau})}$, *i.e.*, the perturbation of the original culture induced by $\boldsymbol{\tau}$, is such that, in expectation, candidate $m$ is precisely at the threshold for being an $\boldsymbol{\alpha}$-winner.

## 4 Results

In this section, we provide an overview of our results, along with a brief intuition about the specific features of each case.

In Equation (5), terms of the form $\frac{1}{1-x_j}$ require special care. If $\boldsymbol{\zeta}_j < 1$, there is no issue, as we can integrate over a circle of radius $\zeta_j$ that loops around 0 without enclosing the singularity at 1. In this case, we say that the coordinate $\boldsymbol{\zeta}_j$ is *subcritical*. However, if $\boldsymbol{\zeta}_j$ is *critical*, i.e., if $\boldsymbol{\zeta}_j = 1$, then the integration path must be adjusted to slightly bypass the singularity. Finally, if $\boldsymbol{\zeta}_j$ is *supercritical*, i.e., if $\boldsymbol{\zeta}_j > 1$, any such circle necessarily encloses a singularity, and the issue remains even with slight deformations of the integration path, requiring the singularity at 1 to be explicitly accounted for in the analysis.

In Section 4.1, we analyze the case where all coordinates are subcritical, while Section 4.2 focuses on the case where all coordinates are critical. Section 4.3 extends the analysis to the mixed case, where each coordinate is either subcritical or critical. Supercritical coordinates are addressed in Section 4.4. Finally, Section 4.5 shows that our techniques can be used to derive higher-order terms in the asymptotic expansion of the previous formulas, using the Impartial Culture model as an example.

### 4.1 Subcritical Case

In this section, we study the case where all coordinates of the saddle point are subcritical. We can then apply the "vanilla" saddle point method [8, Chapter VIII]. In the following theorem, recall that $P(\boldsymbol{x})$ is introduced in Definition 1, and that $\boldsymbol{\tau}$, $\boldsymbol{\zeta}$, and $\mathcal{H}_K(\boldsymbol{\tau})$ are defined in Definition 2.

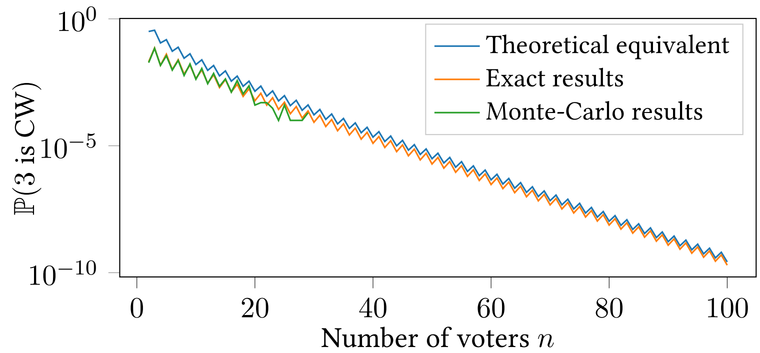


**Figure 3:** Probability that candidate 3 is the Condorcet winner as a function of $n$ in a culture $\mathcal{M}_{3\text{ last}}$ with $\rho = \log(2)$, shown on a semilog scale. The theoretical equivalent is based on Theorem 2, while exact results rely on Equation (2). Monte Carlo simulations use 10,000 profiles per point, yielding an error of order $10^{-2}$. For $n \geq 30$, they return zero, which is not visible due to the logarithmic scale.

**Theorem 1.** *Assume that all coordinates of the saddle point are subcritical, i.e., $\zeta_j < 1$ for every adversary $j \in \mathcal{A}$. Then:*

$$\mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner}) \underset{n\to+\infty}{\sim} \frac{P(\boldsymbol{\zeta})^n}{\prod\limits_{j\in\mathcal{A}} \left((1-\zeta_j){\zeta_j}^{\lceil \beta_j n\rceil - 1}\right)} \frac{1}{\sqrt{(2\pi n)^{m-1}\det(\mathcal{H}_K(\boldsymbol{\tau}))}}.$$

To interpret this asymptotic behavior, recall that $\boldsymbol{\zeta}$ is the unique global minimizer of $\boldsymbol{x} \mapsto \frac{P(\boldsymbol{x})}{\prod_j {x_j}^{\beta_j}}$. Hence, $\frac{P(\boldsymbol{\zeta})}{\prod_j {\zeta_j}^{\beta_j}} < \frac{P(\mathbf{1})}{\prod_j 1^{\beta_j}} = 1$. Thus, Theorem 1 indicates that, in the subcritical case, $\mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner})$ tends to 0 exponentially fast.

Let us illustrate this scenario with $\mathcal{M}_{3\text{ last}}$, defined in Section 2.1, a Mallows culture where the reference ranking is $(1, 2, 3)$, making candidate 3 particularly unlikely to be the Condorcet winner. In that case, we show that $\boldsymbol{\tau} = (\frac{-3\rho}{2}, \frac{-\rho}{2})$, ensuring that both coordinates of the saddle point are subcritical and allowing the application of Theorem 1. Standard algebraic calculations then lead to the following result.

**Theorem 2.** *Under the Mallows culture $\mathcal{M}_{3\text{ last}}$, the probability that candidate 3 is the Condorcet winner has asymptotic behavior*

$$\mathbb{P}(3 \text{ is CW}) \underset{n\to+\infty}{\sim} \frac{P(\boldsymbol{\zeta})^n}{\prod\limits_{j\in\mathcal{A}} \left((1-\zeta_j){\zeta_j}^{\lceil n/2\rceil - 1}\right)} \frac{1}{2\pi n\sqrt{\det(\mathcal{H}_K(\boldsymbol{\tau}))}},$$

*where* $\boldsymbol{\zeta} = \left(e^{-3\rho/2}, e^{-\rho/2}\right)$ , $P(\boldsymbol{\zeta}) = 2\frac{e^{-2\rho}(1+e^{-\rho/2}+e^{-\rho})}{(1+e^{-\rho})(1+e^{-\rho}+e^{-2\rho})}$ *and* $\det(\mathcal{H}_K(\boldsymbol{\tau})) = \frac{1}{4}\frac{e^{-\rho/2}(1+e^{-\rho})}{(1+e^{-\rho/2}+e^{-\rho})^2}$.

Since this probability tends exponentially fast to 0, Theorem 2 can be viewed as a generalization of Condorcet's Jury Theorem [5], which establishes the same conclusion for 2 candidates.

Figure 3 uses our Python package `Actinvoting` [6] to illustrate the case of $\mathcal{M}_{3\text{ last}}$, showing the probability that candidate 3 is the Condorcet winner as a function of the number of voters $n$. We compare three estimations: the theoretical approximation from Theorem 2, the exact result via coefficient extraction from Equation (2), and a Monte Carlo estimate based on 10,000 profiles per value of $n$. The concentration parameter is set to $\rho = \log(2)$ (see [4]). The computational cost differs greatly: for $n = 100$, Monte Carlo takes 46 s, the exact computation 17 s, and the theoretical approximation just 338 $\mu$s.[1] As expected, the approximation converges quickly to the exact value as $n$ grows, while being far more efficient.

[1] All simulations were run on an 11th Gen Intel Core i9-11980HK (8 cores, 16 logical processors) with 64 GB of RAM.

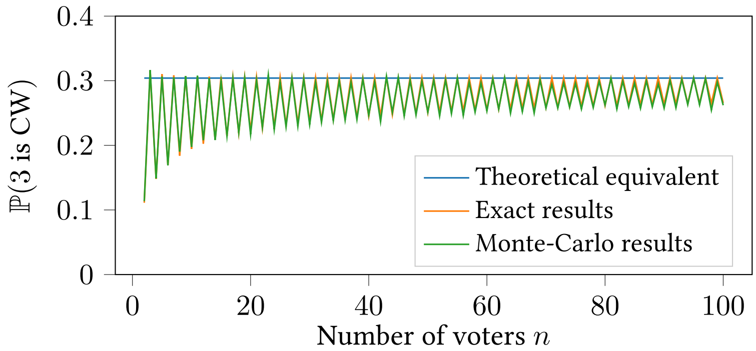


**Figure 4:** Probability that a given candidate is the Condorcet winner as a function of $n$ in the Impartial Culture with three candidates. The theoretical equivalent is based on Theorem 4, while exact results rely on Equation (2). Monte Carlo simulations use 10,000 profiles per point, yielding an error of order $10^{-2}$. When the exact results curve is not visible, it is overlapped by the Monte Carlo results curve.

### 4.2 Critical Case

When $\boldsymbol{\zeta} = \mathbf{1}$, it means that the expected proportion of voters who prefer candidate $m$ to candidate $j$ is exactly $\alpha_j$, as explained in the probabilistic interpretation of Section 3.3. In other words, in expectation, candidate $m$ is precisely at the threshold of being an $\boldsymbol{\alpha}$-winner. The problem is that, as soon as one coordinate $\zeta_j$ is equal to 1, the term $\frac{1}{1-x_j}$ in Equation (5) prevents integration along a circle of radius $\zeta_j$. We circumvent this difficulty by choosing an integration path that slightly deforms around the singularity at 1, leading to the following theorem.

**Theorem 3.** *Assume that all coordinates of the saddle point are critical, i.e., $\zeta_j = 1$ for every adversary $j \in \mathcal{A}$. Then:*

$$\lim_{n\to+\infty} \mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner}) = \frac{1}{\sqrt{(2\pi)^{m-1}\det(\mathcal{H}_K(\boldsymbol{\tau}))}} \int_{(0,+\infty)^{m-1}} e^{-\boldsymbol{u}^\intercal \mathcal{H}_K(\boldsymbol{\tau})^{-1}\boldsymbol{u}/2} d\boldsymbol{u}.$$

In this case, Equation (6) simplifies to $\mathcal{H}_K(\boldsymbol{\tau}) = \mathcal{H}_P(\mathbf{1}) + \operatorname{diag}(\boldsymbol{\beta}) - \boldsymbol{\beta}\boldsymbol{\beta}^\intercal$ because $\boldsymbol{\zeta} = \mathbf{1} = P(\mathbf{1})$.

For the critical case, Theorem 3 generalizes to $\boldsymbol{\alpha}$-winners the result of Niemi and Weisberg [26] and Krishnamoorthy and Raghavachari [20, Equation (12)] on Condorcet winners (see the full paper [4] for a detailed comparison of these results and ours). Applied to Impartial Culture, this leads to the following result.

**Theorem 4.** *Under Impartial Culture with $m$ candidates, the probability that candidate $m$ is the Condorcet winner has the following limit:*

$$\lim_{n\to+\infty} \mathbb{P}(m \text{ is CW}) = \frac{1}{\sqrt{(2\pi)^{m-1}\det(\mathcal{H}_K(\mathbf{0}))}} \int_{(0,+\infty)^{m-1}} e^{-\boldsymbol{u}\mathcal{H}_K(\mathbf{0})^{-1}\boldsymbol{u}/2} d\boldsymbol{u}$$

*where, letting $I$ denote the identity matrix and $J$ the matrix where all elements are 1s, both of dimension $(m-1)\times(m-1)$,*

$$\mathcal{H}_K(\mathbf{0}) = \frac{1}{6}\left(I + \frac{1}{2}J\right), \qquad \mathcal{H}_K(\mathbf{0})^{-1} = 6\left(I - \frac{1}{m+1}J\right), \qquad \det(\mathcal{H}_K(\mathbf{0})) = \frac{m+1}{2}\frac{1}{6^{m-1}}. \tag{7}$$

In the full paper [4, Theorem 5], we provide an alternative expression for this probability and discuss its connection to the existing literature.

Figure 4 illustrates our results for the Impartial Culture with three candidates. The theoretical approximation, which is just a numerical limit here, is simply represented as a horizontal line. Notably, for odd $n$, the exact values closely match the limit, whereas for even $n$, they are significantly lower, although they converge to the same value. In Section 4.5, we show that our techniques not only yield this limiting value, but also provide the convergence rate, thereby offering greater precision than the constant theoretical equivalent presented here.

### 4.3 Mixed Case: Subcritical and Critical

Combining the ideas developed in Sections 4.1 and 4.2, we can deal with the case where some of the coordinates are subcritical and some are critical.

**Theorem 5.** *Let $\mathcal{S} := \{j \in \mathcal{A},\ \zeta_j < 1\}$ and $\mathcal{C} := \{j \in \mathcal{A},\ \zeta_j = 1\}$ respectively denote the subset of subcritical and critical coordinates of the saddle point. Assume $\mathcal{A} = \mathcal{S} \cup \mathcal{C}$. Then:*

$$\mathbb{P}(m \text{ is } \boldsymbol{\alpha}\text{-winner}) \underset{n\to+\infty}{\sim} \frac{P(\boldsymbol{\zeta})^n}{\prod_{j\in\mathcal{S}} \left((1-\zeta_j)\zeta_j^{\lceil \beta_j n\rceil - 1}\right)} \frac{1}{\sqrt{(2\pi)^{m-1} n^{|\mathcal{S}|} \det(\mathcal{H}_K(\boldsymbol{\tau}))}} \int_{(0,+\infty)^{|\mathcal{C}|}} e^{-\boldsymbol{u}^\intercal M \boldsymbol{u}/2} d\boldsymbol{u},$$

*where $M$ is the submatrix of $\mathcal{H}_K(\boldsymbol{\tau})^{-1}$ that corresponds to the rows and columns of $\mathcal{C}$.*

As in the subcritical case, this result establishes that the probability converges exponentially fast to zero whenever at least one coordinate is subcritical. The Gaussian integral associated with the critical coordinates is just a multiplicative constant in this asymptotic behavior.

In the full paper [4], we exemplify this result through the Mallows Culture $\mathcal{M}_{4\,\text{last}}$, showing that the first two coordinates of the saddle point are subcritical, while the last one is critical, and giving a closed-form expression for the theoretical equivalent of $\mathbb{P}(4 \text{ is CW})$.

### 4.4 Dealing With Supercriticality

We now consider the case where at least one coordinate $\zeta_j$ of the saddle point is supercritical, *i.e.*, $\zeta_j > 1$. In this scenario, the integration contour, which must pass sufficiently close to $\zeta_j$, inevitably encloses the singularity at 1. In general, the contribution from this singularity is non-negligible and can be determined via a residue calculation. However, rather than relying solely on calculus, we provide in the full paper [4] an equivalent, more intuitive interpretation based on probabilistic arguments. Applied to the Mallows Culture $\mathcal{M}_{3\,\text{first}}$, this technique leads to the following result.

**Theorem 6.** *Under $\mathcal{M}_{3\,\text{first}}$, the probability that candidate 3 is the Condorcet Winner has asymptotic behavior*

$$1 - \mathbb{P}(3 \text{ is CW}) \underset{n\to+\infty}{\sim} \sqrt{\frac{2}{\pi n}} \frac{e^{-\rho\lceil n/2\rceil}}{1-e^{-\rho}} \left(\frac{2}{1+e^{-\rho}}\right)^n.$$

### 4.5 Asymptotic Expansion

So far, we have derived an asymptotic equivalent for the probability of interest. In the full paper [4], we show how the same techniques can be used to obtain a full asymptotic expansion. Applying this method to the Impartial Culture setting sharpens the result stated in Theorem 4.

**Theorem 7.** *Under Impartial Culture with $m$ candidates, the probability that candidate $m$ is the Condorcet winner has asymptotic behavior*

$$\mathbb{P}(m \text{ is CW}) = a_0 + a_{1,n} n^{-1/2} + \mathcal{O}(n^{-1}),$$

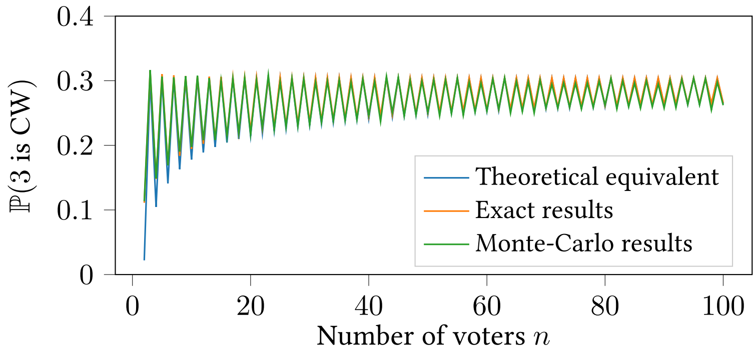


**Figure 5:** Probability that a given candidate is the Condorcet winner as a function of $n$ in the Impartial Culture with three candidates. The theoretical equivalent is based on Theorem 7, while exact results rely on Equation (2). Monte Carlo simulations use 10,000 profiles per point, yielding an error of order $10^{-2}$. Not all curves being visible mean that they overlap.

*where $a_0$ and $a_{1,n}$ are given by*

$$a_0 = \frac{1}{\sqrt{(2\pi)^{m-1}\det(\mathcal{H}_K(\mathbf{0}))}} \int_{(0,+\infty)^{m-1}} e^{-\boldsymbol{u}\mathcal{H}_K(\mathbf{0})^{-1}\boldsymbol{u}/2} d\boldsymbol{u},$$
$$a_{1,n} = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ -\frac{6}{m+1}\frac{1}{\sqrt{(2\pi)^{m-1}\det(\mathcal{H}_K(\mathbf{0}))}} \int_{(0,+\infty)^{m-1}} \mathbf{1}^T \boldsymbol{u} e^{-\boldsymbol{u}^T\mathcal{H}_K(\mathbf{0})^{-1}\boldsymbol{u}/2} d\boldsymbol{u}. & \text{if } n \text{ is even,} \end{cases}$$

*and $\mathcal{H}_K(\mathbf{0})$, its inverse and determinant are given in Equation (7).*

**Numerical simulations** In Figure 4, we observed that the approximation by the limit value was highly accurate for odd values of $n$. Theorem 7 provides an explanation for this: the term of the asymptotic expansion in $n^{-1/2}$ vanishes in this case. In contrast, for even $n$, the term is non-zero and negative, which is consistent with the previous observation that the exact values are significantly lower than the limit in these cases. To illustrate this, Figure 5 shows the same curves as Figure 4 for the Monte Carlo simulations and exact values. However, this time, the theoretical approximation includes the term in $n^{-1/2}$. Now, the three curves become nearly indistinguishable visually: the addition of this term in $n^{-1/2}$ significantly improves the accuracy of the approximation, the error term being $\mathcal{O}(n^{-1})$.

## 5 Future Work

A natural direction for future work is to investigate cases where the saddle point has null or infinite coordinates, which may arise in non-generic cultures. Additionally, it would be valuable to perform a finer analysis of the complexity of the algorithm presented in Section 4.4 for handling supercritical coordinates, particularly to identify conditions under which it runs in polynomial time. Another avenue is to examine the limiting behavior of the probability as the number of candidates $m$ tends to infinity. Finally, it would be insightful to extend our method to other events, such as the transitivity of the majority relation, different kinds of monotonicity failures, or the manipulability, *i.e.,* the susceptibility to strategic voting.

## References


[1] Alexander I Barvinok. A polynomial time algorithm for counting integral points in polyhedra when the dimension is fixed. *Mathematics of Operations Research*, 19(4):769–779, 1994.

[2] Sven Berg and Dominique Lepelley. Note sur le calcul de la probabilité des paradoxes du vote. *Mathématiques et Sciences humaines*, 120:33–48, 1992.

[3] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of computational social choice*. Cambridge University Press, 2016.

[4] Emma Caizergues, Élie de Panafieu, François Durand, Marc Noy, and Vlady Ravelomanana. Probability of a Condorcet winner for large electorates: An analytic combinatorics approach. https://arxiv.org/abs/2505.06028, 2025.

[5] Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie royale, 1785.

[6] François Durand. Actinvoting: Analytic combinatorics tools in voting. https://github.com/francois-durand/actinvoting, 2025.

[7] Peter Fishburn. A proof of May's theorem P(m, 4) = 2 P(m, 3). *Behavioral Science*, 18(3):212, 1973.

[8] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009.

[9] Mark Garman and Morton Kamien. The paradox of voting: Probability calculations. *Behavioral Science*, 13:306–316, 1968.

[10] William Gehrlein. Condorcet's paradox. *Theory and Decision*, 15(2):161–197, 1983.

[11] William Gehrlein. Obtaining representations for probabilities of voting outcomes with effectively unlimited precision integer arithmetic. *Social Choice and Welfare*, 19:503–512, 2002.

[12] William Gehrlein. *Condorcet's Paradox*. Springer, 2006.

[13] William Gehrlein and Peter Fishburn. Condorcet's paradox and anonymous preference profiles. *Public Choice*, pages 1–18, 1976.

[14] William Gehrlein and Peter Fishburn. The probability of the paradox of voting: A computable solution. *Journal of Economic Theory*, 13(1):14–25, 1976.

[15] William Gehrlein and Peter Fishburn. Probabilities of election outcomes for large electorates. *Journal of Economic Theory*, 19(1):38–49, 1978.

[16] Raphael Gillett. A recursion relation for the probability of the paradox of voting. *Journal of Economic Theory*, 18(2):318–327, 1978.

[17] Georges-Théodule Guilbaud. Les théories de l'intérêt général et le problème logique de l'agrégation. *Économie appliquée*, 5:501–584, 1952.

[18] HC Huang and Vincent CH Chua. Analytical representation of probabilities under the IAC condition. *Social Choice and Welfare*, 17(1):143–155, 2000.

[19] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, 1938.

[20] Mukkai Krishnamoorthy and Madabhushi Raghavachari. Condorcet winner probabilities - A statistical perspective. https://arxiv.org/abs/math/0511140, 2005.

[21] Dominique Lepelley. *Contribution à l'analyse des procédures de décision collective*. PhD thesis, Caen, 1989.

[22] Dominique Lepelley, Ahmed Louichi, and Hatem Smaoui. On Ehrhart polynomials and probability calculations in voting theory. *Social Choice and Welfare*, 30(3):363–383, 2008.

[23] Hans Maassen and Thom Bezembinder. Generating random weak orders and the probability of a Condorcet winner. *Social Choice and Welfare*, 19(3):517–532, 2002.

[24] Colin L Mallows. Non-null ranking models. i. *Biometrika*, 44(1/2):114–130, 1957.

[25] Robert May. Some mathematical remarks on the paradox of voting. *Behavioral Science*, 16(2):143–151, 1971.

[26] Richard Niemi and Herbert Weisberg. A mathematical solution for the probability of the paradox of voting. *Behavioral Science*, 13(4):317–323, 1968.

[27] Lisa Sauermann. On the probability of a Condorcet winner among a large number of alternatives. https://arxiv.org/abs/2203.13713, 2022.

[28] Murat R Sertel and M Remzi Sanver. Strong equilibrium outcomes of voting games are the generalized Condorcet winners. *Social Choice and Welfare*, 22(2):331–347, 2004.

[29] Oliver Williamson and Thomas Sargent. Social choice: A probabilistic approach. *The Economic Journal*, 77(308):797–813, 1967.

[30] Mark C Wilson and Geoffrey Pritchard. Probability calculations under the IAC hypothesis. *Mathematical Social Sciences*, 54(3):244–256, 2007.

Emma Caizergues
Université Paris Dauphine and Nokia Bell Labs France
Paris, France
Email: emma.caizergues@nokia-bell-labs.com

Élie de Panafieu
Nokia Bell Labs France
Massy, France
Email: elie.de_panafieu@nokia-bell-labs.com

François Durand
Nokia Bell Labs France
Massy, France
Email: francois.durand@nokia-bell-labs.com

Marc Noy
Universitat Politècnica de Catalunya
Barcelona, Spain
Email: marc.noy@upc.edu

Vlady Ravelomanana
IRIF – UMR CNRS 8243, Université Paris Denis Diderot
Paris, France
Email: vlad@liafa.univ-paris-diderot.fr