

Quantifying Sensitivity to Vote Flips in Two-Candidate Popular Voting via Empirical and Combinatorial Analysis

Girish Prasad¹

Received January 6, 2026

Accepted August 29, 2026

Electronic access September 30, 2026

This study investigates the sensitivity of popular vote outcomes to random vote flips, aiming to quantify the probability that an election result changes when a fixed number of votes are flipped at random. We hypothesized that there exists an exact combinatorial formula that can be used to calculate this probability. To test this, we conducted Monte Carlo simulations to empirically estimate this probability. We used combinatorial analysis to identify the desired closed-form expression for the probability $P(n,k)$, or in other words, the probability that flipping k votes changes the winner in an election with n total voters. The resulting formula was followed by approximations that can be applied in large- n environments. These results describe the influence of a group of vote changes in election outcomes, expressing this relationship in the form of a probability. Notably, the approximation shows that this probability depends on only one quantity, namely the proportion of votes flipped.

Keywords: Combinatorics, Probability, Elections, Voting

Introduction

Elections are central to democratic decision-making, and understanding how changes in votes affect election outcomes is critical for evaluating the strength of voting systems^{1,2}. Popular voting systems are widely used in both governmental and organizational contexts^{1,3}. While much attention has been given to strategic voting and individual influence^{4,5}, we quantify the impact on outcomes of multiple, randomly chosen vote flips. This impact is measured as the probability that flipping a given number of votes changes the overall winner. This probability directly reflects how stable or unstable a given election result is.

Early work on voting power includes the finding that under majority rule, the probability that a single vote is decisive is proportional to $\frac{1}{\sqrt{n}}$ where n is the number of voters, one of the first quantitative results linking electorate size to voting power⁶. Building on this, the probability was estimated directly for the 2008 U.S. presidential election, finding it to be at most 1 in 10 million in the most competitive states and 1 in 60 million on average nationally². These results highlight how rare individual influence becomes at scale, but leave open the question of how larger groups of votes can affect outcomes. Taking a different angle, Monte Carlo simulation under the impartial culture assumption was used to estimate the probability that all weighted scoring rules pick the same winner. It found that this probability is 0.535 for three-candidate elections but drops sharply as the number of candidates in-

creases⁷. Further analysis of election outcomes as the number of voters grows large, applying the central limit theorem, shows that certain outcome patterns become increasingly unlikely in large electorates, which motivates the large- n scenario that the approximation of $P(n,k)$ operates within³. Exact election outcome probabilities in five-candidate elections have also been computed using polytope volume methods, demonstrating that exact combinatorial approaches remain feasible even in complex settings⁸. Additionally, a closed-form expression has been derived for each candidate's winning probability as a function of current support rates and the rate at which information is revealed to voters⁹, providing precedent for the combinatorial and closed-form approach we take here.

More directly related to the work in this paper is the question of how election outcomes change when multiple votes are altered. It has been shown that for many standard voting rules, changing just one voter's preference ranking can replace an entire elected committee, illustrating how sensitive outcomes can be to small changes¹⁰. Along similar lines, polynomial-time sampling algorithms have been derived to estimate the margin of victory, defined as the smallest number of vote changes needed to flip the winner, showing this quantity can be estimated without examining every ballot¹¹. This is closely related to $P(n,k)$, which quantifies the probability that k randomly chosen vote changes do flip the winner. Studies on the hierarchical voting systems where voters elect representatives in groups of three have found that information accuracy falls off predictably with each added level of hierarchy¹², a striking similarity to our tournament voting extension. Equi-

¹ South Windsor High School, Connecticut, USA

libria in plurality voting has been characterized when voters are lazy or truth-biased, showing that strategic behavior shifts outcome probabilities in predictable ways relative to sincere voting¹³, further motivating a formal probabilistic treatment of vote changes. On the simulation side, Bayesian inference combined with Monte Carlo simulation has been applied to estimate seat distribution probabilities in proportional elections¹⁴, and agent-based simulation across 100 runs per configuration has been applied to forecast vote shares in the 2020 Taiwan and U.S. elections, achieving prediction errors below 1%. Both support the reliability of the simulation-based approach we use to estimate $P(n,k)$ empirically¹⁵.

Consider n voters and 2 candidates, one of which is represented by the value 1 and the other of which is represented by the value -1 . We define the ballots to be the set $V_b = \{x_i | 1 \leq i \leq n\}$, where integer x_i is the value that represents the i th vote. We define a voting system V to be a function whose domain is all such sets V_b and range is $\{-1, 1\}$. The output of the function under a given input is considered the winner V_w ¹.

We define a popular voting system P with n voters, where n is an odd positive integer, such that the input is ballots $P_b = \{x_i | 1 \leq i \leq n\}$ and the winner P_w is $\text{sgn}(\sum_{i=1}^n x_i)$.

Define $P(n,k)$ to be the probability that the result of a popular voting system with n total voters changes when k randomly chosen votes are flipped.

The objective of this study is to express $P(n,k)$ in a closed-form formula, because such a formula would reveal patterns that are useful for close approximations as the total number of voters n grows larger.

We used Monte Carlo simulations to empirically estimate $P(n,k)$, as detailed in Methods. Following this, we used combinatorial analysis to derive a far more exact expression for $P(n,k)$ and proposed efficient approximations in large- n environments.

As an extension, we analyze the tournament voting system to identify patterns and form a particular conjecture.

A tournament voting system T is defined with 3^n voters for a nonnegative integer n (here, n need not be odd; unlike the n used for popular voting systems above, n instead denotes the number of tournament rounds), such that the input is ballots $T_b = \{x_i | 1 \leq i \leq 3^n\}$. The winner T_w of T is x_1 if $n = 0$ and otherwise is the winner of the tournament voting system with ballots $\{y_i | 1 \leq i \leq 3^{n-1}\}$ where y_i is the winner of the popular voting system with ballots $\{x_{3i-2}, x_{3i-1}, x_{3i}\}$ ¹².

Define $T(n,k)$ to be the probability that the result of a tournament voting system with 3^n total voters changes when k randomly chosen votes are flipped.

Furthermore, using the same simulation procedure as the popular voting system, we conducted simulations for pairs (n,k) that ran 100,000 trials of the tournament voting system

of 3^n voters and k vote flips and recorded the proportion of trials in which the result changes as the empirical value of $T(n,k)$.

As a further extension, we disprove a corollary of prior work on coalition voting.

Methods

Monte Carlo Simulation Procedure⁷

A popular voting system trial was run for a given number of voters and number of flipped votes k . In each of these trials, each vote was randomly and independently assigned to one of the two candidates; a uniformly random subset of k votes were flipped, and the outcome was recorded. The proportion of trials in which the result changed was used as the empirical estimate of $P(n,k)$.

To determine the number of trials for each (n,k) combination, we consider each trial as a Bernoulli outcome because the election outcome either changes or it does not. Each empirical estimate $\hat{P}(n,k)$ has a standard error of $\sqrt{\frac{(\hat{P}(n,k))(1-\hat{P}(n,k))}{n_{\text{trials}}}}$. We require our estimates to be accurate to within 0.005 of the true probability at 95% confidence. This precision was selected because it is small enough to meaningfully validate agreement with the exact formula derived in Results while remaining computationally efficient across the full range of (n,k) pairs tested. At 95% confidence, recall that $z^* = 1.96$. In the worst-case scenario ($\hat{P}(n,k) = 0.5$), it follows that the margin of error satisfies

$$\begin{aligned} 1.96 \sqrt{\frac{(\hat{P}(n,k))(1-\hat{P}(n,k))}{n_{\text{trials}}}} &\leq 0.005 \\ \sqrt{\frac{\hat{P}(n,k)(1-\hat{P}(n,k))}{n_{\text{trials}}}} &\leq \frac{0.005}{1.96} = \frac{1}{392} \\ \sqrt{\frac{0.5 \cdot 0.5}{n_{\text{trials}}}} &= \sqrt{\frac{0.25}{n_{\text{trials}}}} \leq \frac{1}{392} \\ \frac{0.25}{n_{\text{trials}}} &\leq \frac{1}{392^2} \\ n_{\text{trials}} &\geq (0.25)(392^2) = 38,416 \end{aligned}$$

Therefore, 100,000 trials were run for each (n,k) pair since $100,000 \geq 38,416$, giving an actual margin of error of

$$1.96 \sqrt{\frac{0.25}{100,000}} \approx 0.003$$

For the popular voting system, this procedure was applied for each $n \in \{11, 21, 101, 1001\}$ and each $1 \leq k < n$. For the tournament voting system, the same procedure was applied with 3^n voters for each $n \in \{2, 3, 4\}$ and each $1 \leq k < 3^n$.

From these results, a pattern emerged that motivated extending the tournament voting simulation to $(n,k)=(5,15)$, $(6,31)$, $(7,63)$, $(8,127)$.

Simulations were implemented in Python 3.11 using NumPy 1.26.2. All code was self-generated and is publicly available on GitHub for verification and reproducibility.

Results

Model Dependent

To begin the investigation, we recorded the empirical estimates of $P(n,k)$ as described in the methodology. Across the tested values of n , the empirical curves for $P(n,k)$ displayed consistent structural patterns. For both $n = 11$ and $n = 21$, the curves suggested that $P(n,k) = P(n,k+1)$ for odd k . This result contradicts the intuitive expectation that $P(n,k)$ increases strictly with k (Figure 1a).

For $n = 101$, the curve still showed this pairing, but the overall shape began to resemble a smooth trigonometric wave, suggesting that such a closed-form expression or approximation for $P(n,k)$ may exist. At $n = 1001$, the sensitivity curve followed a similar trigonometric pattern (Figure 1b).

Rigorous Proof Dependent

Having observed these empirical patterns, we derive $P(n,k)$ mathematically. Since n must be odd in a popular voting system, we may let $n = 2x + 1$ for $x \geq 0$. Consider a 2-candidate popular voting system with $2x + 1$ total voters in which k votes are flipped. Here, x is the integer for which $x + 1$ is the number of votes needed for either candidate to secure a strict majority.

Let the candidates in this election be A and B.

We seek a closed-form expression for $P(2x + 1, k)$. To derive this, we condition on the number of votes each candidate receives before and after the flip.

Assume without loss of generality that the original loser of this election is A. We will multiply the resulting probability by 2 to account for the symmetric case in which B is the original loser.

By the construction in Methods, each ballot is independently and uniformly assigned to A or B, and the set of k flipped votes is chosen uniformly at random from the n votes, independently of the ballots' original values. Let I denote the number of the k flipped votes that were originally cast for A. Let J denote the number of non-flipped votes that were cast for A. Since each of the ballots are independently A or B with probability $\frac{1}{2}$, it follows that $I \sim B(k, \frac{1}{2})$ and $J \sim B(2x + 1 - k, \frac{1}{2})$, where $B(n, p)$ denotes a binomial distribution with n trials and a probability of success p . Hence,

$$P(I = i) = \frac{\binom{k}{i}}{2^k} \quad \text{and} \quad P(J = j) = \frac{\binom{2x+1-k}{j}}{2^{2x+1-k}}$$

Since I of the flipped votes were originally for A and $k - I$ for B, after flipping, A's vote total changes by net amount $(k - I) - I = k - 2I$. Since A was the original loser, A needs this net gain to overtake B. In particular, for A to have any chance of winning, we require

$$k - 2I > 0$$

$$I < \frac{k}{2}$$

$$I \leq \left\lfloor \frac{k-1}{2} \right\rfloor$$

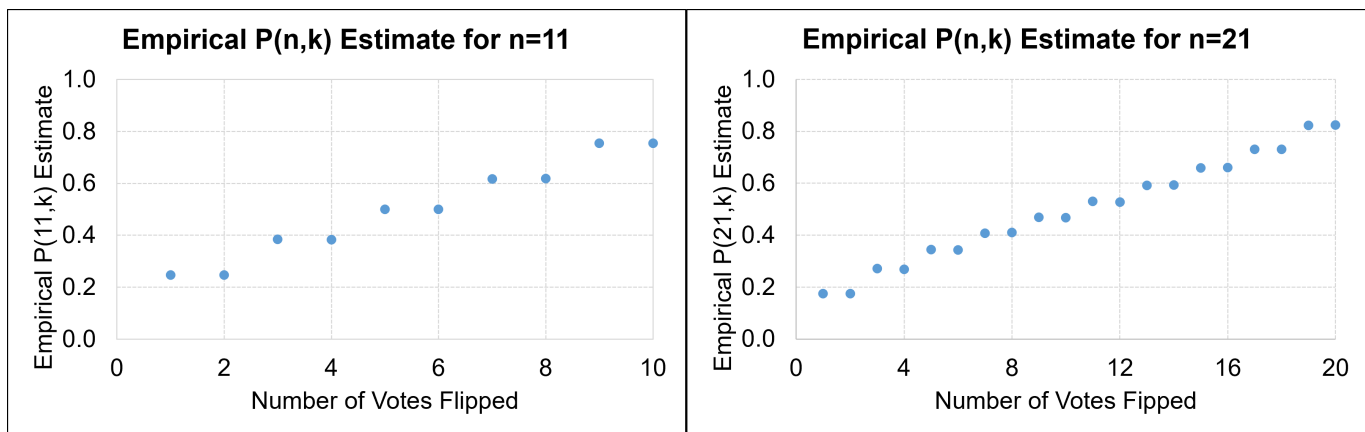


Figure. 1a Empirical estimates of $P(n,k)$ as a function of the number of votes flipped k , for $n = 11$ and $n = 21$, obtained via Monte Carlo simulation (100,000 trials per (n,k) pair).

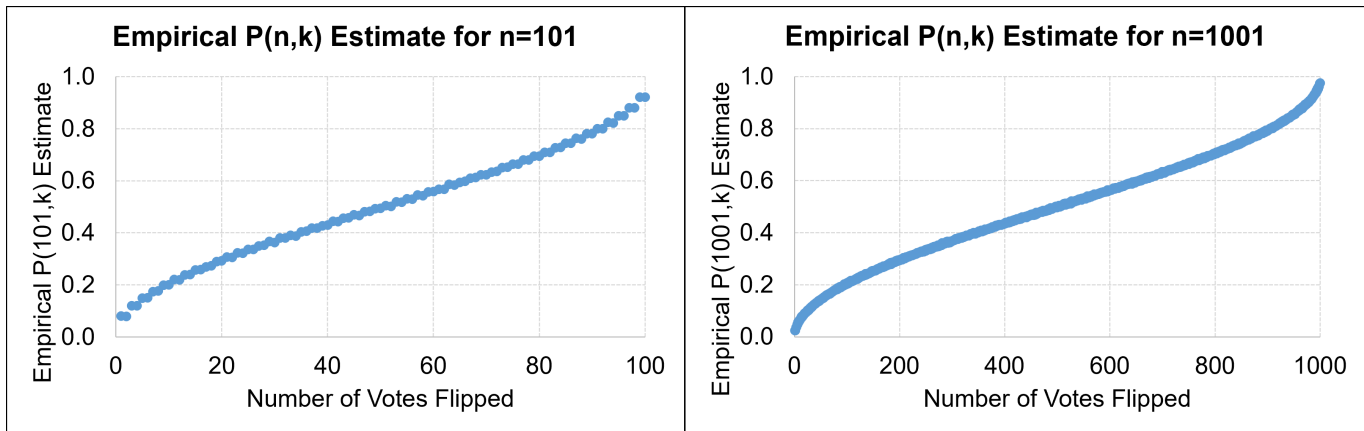


Figure. 1b Empirical estimates of $P(n,k)$ as a function of the number of votes flipped k , for $n = 101$ and $n = 1001$, obtained via Monte Carlo simulation (100,000 trials per (n,k) pair).

A wins post-flip if and only if A's final total exceeds x out of the $2x+1$ votes. A's final total is $J+k-I$, implying that we require

$$\begin{aligned} J+k-I &\geq x+1 \\ J &\geq x-k+I+1 \end{aligned}$$

A loses before votes are flipped if and only if A's original total does not exceed x out of the $2x+1$ votes. A's original total is $J+I$, implying that we require

$$\begin{aligned} J+I &\leq x \\ J &\leq x-I \end{aligned}$$

Hence, it follows that for a fixed value of I , the probability that after k votes are flipped the original loser A wins the election is

$$\sum_{j=x-k+I+1}^{x-I} \frac{\binom{2x+1-k}{j}}{2^{2x+1-k}} = \frac{\sum_{j=x-k+I+1}^{x-I} \binom{2x+1-k}{j}}{2^{2x+1-k}}. \text{ Since } I$$

ranges from 0 to $\lfloor \frac{k-1}{2} \rfloor$, inclusive, as shown above, the overall probability that A loses the original election and wins after k

$$\text{votes are flipped is } \sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} \frac{\binom{k}{i}}{2^k} \frac{\sum_{j=x-k+1+i}^{x-i} \binom{2x+1-k}{j}}{2^{2x+1-k}}.$$

By symmetry, since both A and B are equally likely to be the original loser, and these events are disjoint, the probability $P(2x+1,k)$ is twice the probability computed under the assumption that A is the loser. Hence,

$$P(2x+1,k) = 2 \sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} \frac{\binom{k}{i}}{2^k} \frac{\sum_{j=x-k+1+i}^{x-i} \binom{2x+1-k}{j}}{2^{2x+1-k}}$$

$$= \frac{1}{2^{2x}} \sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} \left(\binom{k}{i} \sum_{j=x-k+1+i}^{x-i} \binom{2x+1-k}{j} \right)$$

We now simplify this sum to obtain the desired closed-form expression.

Claim: For all integers $x \geq 0$ and $k \geq 0$,

$$P(2x+1,k) = \sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} \frac{\binom{2i}{i} \binom{2x-2i}{x-i}}{2^{2x}}.$$

Proof: We proceed to prove this with strong induction on k .

Base Case: When $k = 1$, both formulas give $\frac{\binom{2x}{x}}{2^{2x}}$.

Inductive Hypothesis: Assume that the claim is true for all integers $k \leq m$ for some integer $m \geq 1$. Notice that m denotes the value of k at the current step of induction. We use m rather than k to distinguish the induction variable from the fixed parameter k in the statement being proven.

Inductive Step: We will consider the cases with m odd and m even separately.

Case 1: m is odd

We wish to show that $P(2x+1,m+1) = P(2x+1,m)$

$$\begin{aligned} P(2x+1,m) &= \frac{1}{2^{2x}} \sum_{i=0}^{\lfloor \frac{m-1}{2} \rfloor} \left(\binom{m}{i} \sum_{j=x-m+1+i}^{x-i} \binom{2x+1-m}{j} \right) \\ &= \frac{1}{2^{2x}} \left(\binom{m}{0} \left(\binom{2x-m+1}{x-(m-1)} + \binom{2x-m+1}{x-(m-2)} \right) \right. \end{aligned}$$

$$+ \dots + \binom{2x-m+1}{x} \Big) + \binom{m}{1} \Big(\binom{2x-m+1}{x-(m-2)} + \dots + \binom{2x-m+1}{x-1} \Big) + \dots + \binom{m}{\frac{m-1}{2}} \Big(\binom{2x-m+1}{x-\frac{m-1}{2}} \Big) \Big)$$

Recall that Pascal's Identity states that for integers $n \geq 1$ and $0 \leq r \leq n$,

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

Combinatorially, this reflects partitioning the $\binom{n}{r}$ subsets of size r into those that exclude a fixed element (counted by $\binom{n-1}{r}$) and those that include it (counted by $\binom{n-1}{r-1}$). In the context of this election, this partitioning is determined by whether or not the $(2x-m+1)$ th non-flipped vote was for A or B.

We may now expand each binomial coefficient using Pascal's Identity to achieve the following

$$\begin{aligned} P(2x+1, m) &= \frac{1}{2^{2x}} \left(\binom{m}{0} \left(\binom{2x-m}{x-m} + 2 \binom{2x-m}{x-(m-1)} \right) \right. \\ &\quad + 2 \binom{2x-m}{x-(m-2)} + \dots + 2 \binom{2x-m}{x-1} + \binom{2x-m}{x} \Big) \\ &\quad + \binom{m}{1} \left(\binom{2x-m}{x-(m-1)} + 2 \binom{2x-m}{x-(m-2)} + \dots \right. \\ &\quad \left. + 2 \binom{2x-m}{x-2} + \binom{2x-m}{x-1} \right) \\ &\quad + \dots + \binom{m}{\frac{m-1}{2}} \left(\binom{2x-m}{x-\frac{m+1}{2}} + \binom{2x-m}{x-\frac{m-1}{2}} \right) \Big) \\ &= \frac{1}{2^{2x}} \left(\binom{2x-m}{x-m} \binom{m}{0} \right. \\ &\quad + \binom{2x-m}{x-(m-1)} \left(2 \binom{m}{0} + \binom{m}{1} \right) + \dots \\ &\quad + \binom{2x-m}{x-\frac{m+1}{2}} \left(2 \binom{m}{0} + 2 \binom{m}{1} + \dots + \binom{m}{\frac{m-1}{2}} \right) \\ &\quad + \binom{2x-m}{x-\frac{m-1}{2}} \left(2 \binom{m}{0} + 2 \binom{m}{1} + \dots + \binom{m}{\frac{m-1}{2}} \right) \\ &\quad \left. + \dots + \binom{2x-m}{x} \binom{m}{0} \right). \end{aligned}$$

We now apply Pascal's Identity a second time, in the form $\binom{m}{j} + \binom{m}{j-1} = \binom{m+1}{j}$ to condense the binomial coefficients and achieve the following

$$\begin{aligned} P(2x+1, m) &= \frac{1}{2^{2x}} \left(\binom{2x-m}{x-m} \binom{m+1}{0} \right) \\ &\quad + \binom{2x-m}{x-(m-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} \right) + \dots \\ &\quad + \binom{2x-m}{x-\frac{m+1}{2}} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots + \binom{m+1}{\frac{m-1}{2}} \right) \\ &\quad + \binom{2x-m}{x-\frac{m-1}{2}} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots + \binom{m+1}{\frac{m-1}{2}} \right) \\ &\quad + \dots + \binom{2x-m}{x} \binom{m+1}{0} \Big) \\ &= \frac{1}{2^{2x}} \left(\binom{m+1}{0} \left(\binom{2x-m}{x-m} + \dots + \binom{2x-m}{x} \right) \right. \\ &\quad + \binom{m+1}{1} \left(\binom{2x-m}{x-(m-1)} + \dots + \binom{2x-m}{x-1} \right) \\ &\quad \left. + \dots + \binom{m+1}{\frac{m-1}{2}} \left(\binom{2x-m}{x-\frac{m+1}{2}} + \binom{2x-m}{x-\frac{m-1}{2}} \right) \right) \\ &= P(2x+1, m+1). \end{aligned}$$

Case 2: m is even

We wish to show that

$$\begin{aligned} P(2x+1, m+1) &= P(2x+1, m) + \frac{\binom{m}{\frac{m}{2}} \binom{2x-m}{x-\frac{m}{2}}}{2^{2x}}. \\ P(2x+1, m) &= \frac{1}{2^{2x}} \sum_{i=0}^{\lfloor \frac{m-1}{2} \rfloor} \left(\binom{m}{i} \sum_{j=x-m+1+i}^{x-i} \binom{2x+1-m}{j} \right), \\ &= \frac{1}{2^{2x}} \left(\binom{m}{0} \left(\binom{2x-m+1}{x-(m-1)} + \binom{2x-m+1}{x-(m-2)} \right. \right. \\ &\quad \left. \left. + \dots + \binom{2x-m+1}{x} \right) \right. \\ &\quad + \binom{m}{1} \left(\binom{2x-m+1}{x-(m-2)} + \dots + \binom{2x-m+1}{x-1} \right) \\ &\quad \left. + \dots + \binom{m}{\frac{m}{2}-1} \left(\binom{2x-m+1}{x-\frac{m}{2}} + \binom{2x-m+1}{x-(\frac{m}{2}-1)} \right) \right) \end{aligned}$$

As described in the previous case, we may now expand each binomial coefficient using Pascal's Identity to achieve the following

$$\begin{aligned}
P(2x+1, m) &= \frac{1}{2^{2x}} \left(\binom{m}{0} \left(\binom{2x-m}{x-m} + 2 \binom{2x-m}{x-(m-1)} \right) \right. \\
&\quad + 2 \binom{2x-m}{x-(m-2)} + \dots + 2 \binom{2x-m}{x-1} + \binom{2x-m}{x} \Bigg) \\
&\quad + \binom{m}{1} \left(\binom{2x-m}{x-(m-1)} + 2 \binom{2x-m}{x-(m-2)} + \dots \right. \\
&\quad \left. + 2 \binom{2x-m}{x-2} + \binom{2x-m}{x-1} \right) \\
&\quad + \dots + \binom{m}{\frac{m}{2}-1} \left(\binom{2x-m}{x-(\frac{m}{2}+1)} + 2 \binom{2x-m}{x-\frac{m}{2}} \right. \\
&\quad \left. + \binom{2x-m}{x-(\frac{m}{2}-1)} \right) \Bigg) \\
&= \frac{1}{2^{2x}} \left(\binom{2x-m}{x-m} \binom{m}{0} \right. \\
&\quad + \binom{2x-m}{x-(m-1)} \left(2 \binom{m}{0} + \binom{m}{1} \right) + \dots \\
&\quad + \binom{2x-m}{x-(\frac{m}{2}+1)} \left(2 \binom{m}{0} + 2 \binom{m}{1} + \dots \right. \\
&\quad \left. + \binom{m}{\frac{m}{2}-1} \right) \\
&\quad + \binom{2x-m}{x-\frac{m}{2}} \left(2 \binom{m}{0} + 2 \binom{m}{1} + \dots + 2 \binom{m}{\frac{m}{2}-1} \right) \\
&\quad + \binom{2x-m}{x-(\frac{m}{2}-1)} \left(2 \binom{m}{0} + 2 \binom{m}{1} + \dots \right. \\
&\quad \left. + \binom{m}{\frac{m}{2}-1} \right) + \dots + \binom{2x-m}{x} \binom{m}{0} \Bigg).
\end{aligned}$$

Once again, we may now condense the binomial coefficients using Pascal's Identity to achieve the following

$$\begin{aligned}
P(2x+1, m) &= \frac{1}{2^{2x}} \left(\binom{2x-m}{x-m} \binom{m+1}{0} \right. \\
&\quad + \binom{2x-m}{x-(m-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} \right) + \dots
\end{aligned}$$

$$\begin{aligned}
&\quad + \binom{2x-m}{x-(\frac{m}{2}+1)} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
&\quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) \\
&\quad + \binom{2x-m}{x-\frac{m}{2}} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
&\quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) \\
&\quad + \binom{2x-m}{x-(\frac{m}{2}-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
&\quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) + \dots \\
&\quad + \binom{2x-m}{x} \binom{m+1}{0} + \left[\binom{2x-m}{x-\frac{m}{2}} \binom{m}{\frac{m}{2}-1} \right] \Bigg).
\end{aligned}$$

Hence,

$$\begin{aligned}
P(2x+1, m) &+ \frac{\binom{m}{\frac{m}{2}} \binom{2x-m}{x-\frac{m}{2}}}{2^{2x}} \\
&= \frac{1}{2^{2x}} \left(\binom{2x-m}{x-m} \binom{m+1}{0} \right. \\
&\quad + \binom{2x-m}{x-(m-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} \right) + \dots \\
&\quad + \binom{2x-m}{x-(\frac{m}{2}+1)} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
&\quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) \\
&\quad + \binom{2x-m}{x-\frac{m}{2}} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
&\quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) \\
&\quad + \binom{2x-m}{x-(\frac{m}{2}-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
&\quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) + \dots + \binom{2x-m}{x} \binom{m+1}{0} \Bigg).
\end{aligned}$$

$$\begin{aligned}
& + \left[\binom{2x-m}{x-\frac{m}{2}} \binom{m}{\frac{m}{2}-1} + \binom{m}{\frac{m}{2}} \binom{2x-m}{x-\frac{m}{2}} \right] \\
& = \frac{1}{2^{2x}} \left(\binom{2x-m}{x-m} \binom{m+1}{0} \right. \\
& \quad + \binom{2x-m}{x-(m-1)} \binom{m+1}{0} + \binom{m+1}{1} \Big) + \dots \\
& \quad + \binom{2x-m}{x-(\frac{m}{2}+1)} \left(\binom{m+1}{0} + \binom{m+1}{1} \right) + \dots \\
& \quad + \binom{m+1}{\frac{m}{2}-1} \Big) \\
& \quad + \binom{2x-m}{x-\frac{m}{2}} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
& \quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) \\
& \quad + \binom{2x-m}{x-(\frac{m}{2}-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
& \quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) + \dots \\
& \quad + \binom{2x-m}{x} \binom{m+1}{0} \Big) \\
& \quad + \left[\binom{2x-m}{x-\frac{m}{2}} \binom{m+1}{\frac{m}{2}} \right] \\
& = \frac{1}{2^{2x}} \left(\binom{2x-m}{x-m} \binom{m+1}{0} \right. \\
& \quad + \binom{2x-m}{x-(m-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} \right) + \dots \\
& \quad + \binom{2x-m}{x-(\frac{m}{2}+1)} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
& \quad \left. + \binom{m+1}{\frac{m}{2}-1} \right) \\
& \quad + \binom{2x-m}{x-\frac{m}{2}} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \\
& \quad \left. + \binom{m+1}{\frac{m}{2}-1} + \binom{m+1}{\frac{m}{2}} \right) \\
& \quad \left. + \binom{2x-m}{x-(\frac{m}{2}-1)} \left(\binom{m+1}{0} + \binom{m+1}{1} + \dots \right. \right.
\end{aligned}$$

$$\begin{aligned}
& \left. + \binom{m+1}{\frac{m}{2}-1} \right) + \dots + \binom{2x-m}{x} \left(\binom{m+1}{0} \right) \\
& = \frac{1}{2^{2x}} \left(\binom{m+1}{0} \left(\binom{2x-m}{x-m} + \dots + \binom{2x-m}{x} \right) \right. \\
& \quad + \binom{m+1}{1} \left(\binom{2x-m}{x-(m-1)} + \dots + \binom{2x-m}{x-1} \right) \\
& \quad + \dots + \binom{m+1}{\frac{m}{2}} \left(\binom{2x-m}{x-\frac{m}{2}} \right) \Big) \\
& = P(2x+1, m+1).
\end{aligned}$$

Hence, we conclude $P(2x+1, k) = \sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} \frac{\binom{2i}{i} \binom{2x-2i}{x-i}}{2^{2x}}$. Having established this exact formula, we now derive a computationally efficient approximation that captures its essential behavior for large n . To simplify the complexity of the expression within the summation above, we may find an approximation for $P(2x+1, k)$ as follows. Recall that the Central Binomial Coefficient $\binom{2n}{n}$ can be approximated as shown below¹⁶

$$\binom{2n}{n} \approx \frac{4^n}{\sqrt{\pi n}} = \frac{2^{2n}}{\sqrt{\pi n}}$$

Hence,

$$\begin{aligned}
P(2x+1, k) & = \sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} \frac{\binom{2i}{i} \binom{2x-2i}{x-i}}{2^{2x}} \\
& \approx \frac{\binom{2x}{x}}{2^{2x}} + \sum_{i=1}^{\lfloor \frac{k-1}{2} \rfloor} \frac{\binom{2^{2i}}{\sqrt{\pi i}} \binom{2^{2x-2i}}{\sqrt{\pi(x-i)}}}{2^{2x}} \\
& = \frac{\left(\frac{2^{2x}}{\sqrt{\pi x}}\right)}{2^{2x}} + \frac{1}{\pi} \sum_{i=1}^{\lfloor \frac{k-1}{2} \rfloor} \frac{1}{\sqrt{i(x-i)}} \\
& = \frac{1}{\sqrt{\pi x}} + \frac{1}{\pi} \sum_{i=1}^{\lfloor \frac{k-1}{2} \rfloor} \frac{1}{\sqrt{i(x-i)}}
\end{aligned}$$

By Stirling's series, this approximation has relative error

$$\frac{\left(\frac{\binom{2x}{x}}{2^{2x}}\right)}{\left(\frac{1}{\sqrt{\pi x}}\right)} = 1 - \frac{1}{8x} + O\left(\frac{1}{x^2}\right)$$

Hence, the relative error is $O\left(\frac{1}{x}\right)$ as x grows without bound¹⁶⁻¹⁸. Let $f(i) = \frac{1}{\sqrt{i(x-i)}}$

The following is a corollary of the Euler-Maclaurin Formula¹⁹

$$\sum_{i=1}^{\lfloor \frac{k-1}{2} \rfloor} f(i) \approx \int_1^{\lfloor \frac{k-1}{2} \rfloor} f(i) di$$

The error of this sum-to-integral approximation is bounded by the leading Euler-Maclaurin correction term $\frac{f(1)+f(\lfloor \frac{k-1}{2} \rfloor)}{2}$, which for $f(i) = \frac{1}{\sqrt{i(x-i)}}$ is $O\left(\frac{1}{\sqrt{x}}\right)$ and therefore vanishes as x approaches infinity¹⁹.

Notice that

$$\int f(i)di = 2 \tan^{-1} \left(\frac{\sqrt{i}}{\sqrt{x-i}} \right)$$

Thus,

$$\begin{aligned} \sum_{i=1}^{\lfloor \frac{k-1}{2} \rfloor} f(i) &\approx \int_1^{\lfloor \frac{k-1}{2} \rfloor} f(i)di \\ &= 2 \tan^{-1} \left(\frac{\sqrt{\lfloor \frac{k-1}{2} \rfloor}}{\sqrt{x - \lfloor \frac{k-1}{2} \rfloor}} \right) - 2 \tan^{-1} \left(\frac{1}{\sqrt{x-1}} \right) \end{aligned}$$

Hence, it follows that

$$\begin{aligned} P(2x+1, k) &\approx \frac{1}{\sqrt{\pi x}} + \frac{1}{\pi} \left(2 \tan^{-1} \left(\frac{\sqrt{\lfloor \frac{k-1}{2} \rfloor}}{\sqrt{x - \lfloor \frac{k-1}{2} \rfloor}} \right) \right. \\ &\quad \left. - 2 \tan^{-1} \left(\frac{1}{\sqrt{x-1}} \right) \right) \\ &= \frac{1}{\sqrt{\pi x}} + \frac{2}{\pi} \left(\tan^{-1} \left(\frac{\sqrt{\lfloor \frac{k-1}{2} \rfloor}}{\sqrt{x - \lfloor \frac{k-1}{2} \rfloor}} \right) - \tan^{-1} \left(\frac{1}{\sqrt{x-1}} \right) \right) \end{aligned}$$

For large values of x , both $\frac{1}{\sqrt{\pi x}}$ and $\tan^{-1} \left(\frac{1}{\sqrt{x-1}} \right)$ are extremely close to 0. Thus,

$$\begin{aligned} P(2x+1, k) &\approx \frac{2}{\pi} \tan^{-1} \left(\frac{\sqrt{\lfloor \frac{k-1}{2} \rfloor}}{\sqrt{x - \lfloor \frac{k-1}{2} \rfloor}} \right) \\ &\approx \frac{2}{\pi} \tan^{-1} \left(\frac{\sqrt{\frac{k-1}{2}}}{\sqrt{x - \frac{k-1}{2}}} \right) \approx \frac{2}{\pi} \tan^{-1} \left(\frac{\sqrt{k}}{\sqrt{2x+1-k}} \right) \end{aligned}$$

Therefore,

$$P(n, k) \approx \frac{2}{\pi} \tan^{-1} \left(\sqrt{\frac{k}{n-k}} \right)$$

Consider the expression that is contained within the $\tan^{-1}()$ function of this approximation. This expression may be rewritten as $\sqrt{\frac{k}{n-k}} = \sqrt{\frac{\frac{k}{n}}{1-\frac{k}{n}}}$. Hence, the approximation of $P(n, k)$ may be rewritten as $\frac{2}{\pi} \tan^{-1} \left(\sqrt{\frac{k}{n-k}} \right) =$

$\frac{2}{\pi} \tan^{-1} \left(\sqrt{\frac{\frac{k}{n}}{1-\frac{k}{n}}} \right)$. This modification suggests that $P(n, k)$ can be approximated given just the ratio $\frac{k}{n}$, or in other words, the proportion of votes in a popular voting system that are flipped. Furthermore, it follows that flipping the same proportion of votes in two popular voting systems, each with a sufficiently large, although not the same, number of votes, results in approximately equal values of $P(n, k)$.

To test this interpretation systematically, we plot the empirical estimates of $P(n, k)$ against $\frac{k}{n}$ for every tested pair. The resulting points increasingly collapse onto a single common curve as n increases. The $n = 101$ and $n = 1001$ datasets align closely across nearly the entire domain of $\frac{k}{n}$ while the $n = 11$ and $n = 21$ datasets, corresponding to much smaller elections, show visibly more scatter around this common curve. This is consistent with $P(n, k)$ depending asymptotically only on the ratio $\frac{k}{n}$, with the approximation to a single limiting curve improving as n grows. This decreasing scatter is consistent with the theoretical error bounds derived above: the $O(\frac{1}{x})$ error of the central binomial coefficient approximation and the $O(\frac{1}{\sqrt{x}})$ Euler-Maclaurin correction term become insignificant as n grows large (Figure 2).

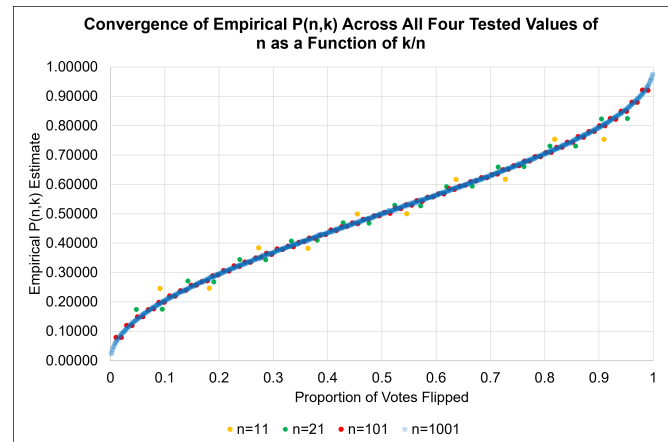


Figure. 2 Empirical $P(n, k)$ plotted against $\frac{k}{n}$ for $n = 11$, $n = 21$, $n = 101$, and $n = 1001$, showing convergence onto a common curve as n increases.

We now turn to the tournament voting system extension. From the tournament voting system simulations, empirical values of $T(n, k)$ for cases with $k = 2^{n-1} - 1$ were $T(2, 1) = 0.25068$, $T(3, 3) = 0.25935$, $T(4, 7) = 0.26032$, $T(5, 15) = 0.25995$, $T(6, 31) = 0.25420$, $T(7, 63) = 0.25255$. Hence, we conjecture that $T(n, 2^{n-1} - 1) = \frac{1}{4}$ for all $n \geq 2$.

$T(1, 1)$ can be found to be $\frac{1}{2}$ by going through all $2^3 * 3 = 24$ possibilities of 3 votes with one of the votes being flipped.

$T(n, 1)$ is the product of two probabilities, namely the probability that a single vote flip changes the result of a tourna-

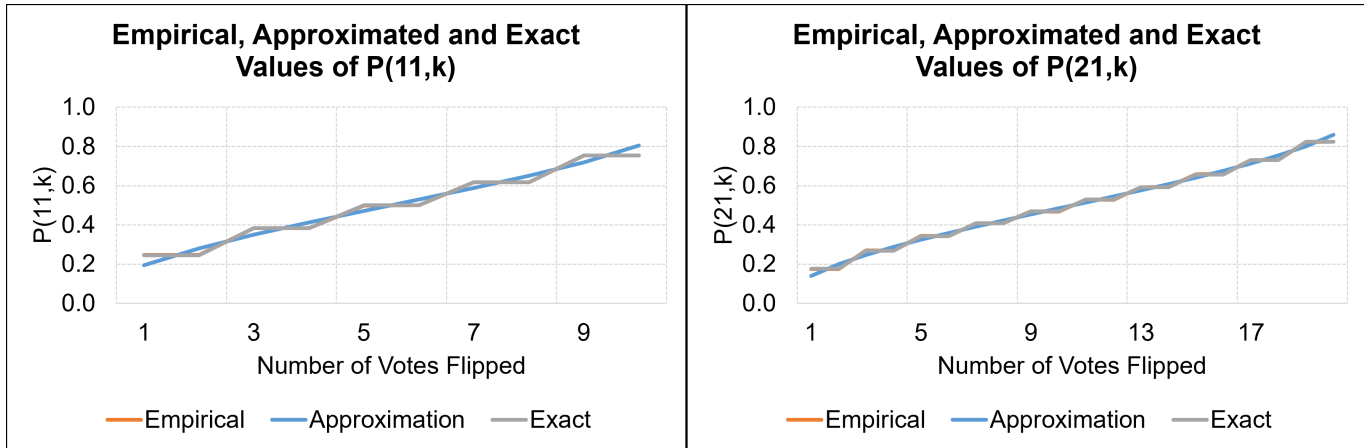


Figure. 3a Empirical, Approximate, and Exact values of $P(n,k)$ plotted together for $n = 11$ and $n = 21$ for visual agreement

ment voting system with 3^{n-1} total votes and the probability that this result change affects the last round of the tournament voting process. It follows that $T(n, 1) = \frac{1}{2}T(n-1, 1)$ for all $n > 1$ and therefore $T(n, 1) = \frac{1}{2^n}$ for all $n \geq 1$.

Furthermore, prior work has considered the concept of coalitions of voters in popular voting systems. A coalition of m voters for an odd integer $m \geq 3$ is a subset of voters $\{v_i | 1 \leq i \leq m\}$ where each v_i is 1 or -1 , such that each of the m votes is replaced with $\text{sgn}(\sum_{i=1}^m v_i)$ before executing the popular voting system^{1,6,20}. Gelman et al. showed that for $k = 1$ and any fixed $n > 1$, the value of $P(n,k)$ is maximum with no coalitions¹.

Finally, we disprove the conjecture that for any fixed $k > 1$ and fixed $n > k$, the value of $P(n,k)$ is maximum with no coalitions. Consider two popular voting systems, each with $n = 9$ and $k = 2$, such that one of them has no coalitions and the other has a single coalition of 5 voters. In the popular voting system with no coalitions, the probability simply comes out to $P(9,2) = P(9,1) = \frac{\binom{8}{4}}{2^8} = 0.2734375$. In the popular voting system with a coalition of 5 voters, the majority within the coalition becomes the overall result of the popular voting system. Hence, we need the probability that the majority within the coalition changes when 2 of the 9 votes are flipped. When neither of the votes are in the coalition, the probability for the result to flip is 0. In the $\frac{4 \cdot 5}{\binom{9}{2}} = \frac{20}{36}$ probability that exactly one of the two votes is in the coalition, there is a $P(5,1)$ probability that the majority within the coalition flips. In the $\frac{\binom{5}{2}}{\binom{9}{2}} = \frac{10}{36}$ probability that both votes are in the coalition, there is a $P(5,2)$ probability that the majority within the coalition flips. Hence, the desired overall probability is $\frac{20}{36}P(5,1) + \frac{10}{36}P(5,2) = \frac{20}{36}P(5,1) + \frac{10}{36}P(5,1) = \frac{30}{36}P(5,1) = \frac{5}{6} * \frac{\binom{4}{2}}{2^4} = \frac{5}{16} = 0.3125 > 0.2734375$ and the conjecture is dis-

proven.

Discussion

We quantified the probability $P(n,k)$ that flipping k votes in a random two-candidate election with n voters changes the outcome. We now examine the implications of these results, assess limitations of the experimental design, and outline directions for further study. We found the exact formula to match the simulation results across all tested values of n and k . We also introduced an approximation to simplify evaluation in large- n regimes.

The exact formula provides an immediate explanation of the empirically observed pairing $P(n,k) = P(n,k+1)$ for odd k . For different values of k , the maximum value of i in the summation of the exact formula varies. For odd integers k , it follows that $\lfloor \frac{k-1}{2} \rfloor = \lfloor \frac{(k+1)-1}{2} \rfloor$. Hence, the exact formula confirms the prior $P(n,k) = P(n,k+1)$ conjecture for odd k .

We plotted the empirical, exact, and approximate values of $P(n,k)$ together on the same graph for each $n \in \{11, 21, 101, 1001\}$. In all graphs, the empirical curve is visually indistinguishable from the exact formula, indicating that simulation-based $P(n,k)$ estimates converge to the evaluated results across the full domain of k . This agreement confirms that the Monte Carlo procedure provides a reliable estimate of the true probability across the full domain of k (Figure 3a).

Furthermore, for both $n = 101$ and $n = 1001$, the approximated curve is also visually indistinguishable from the exact curve. This overlap implies that the error between the approximation and the exact formula is insignificant as n becomes significantly large (Figure 3b).

In fact, the maximum error drops from approximately 0.05 at $n = 11$ to 0.005 at $n = 1001$, indicating that the approxi-

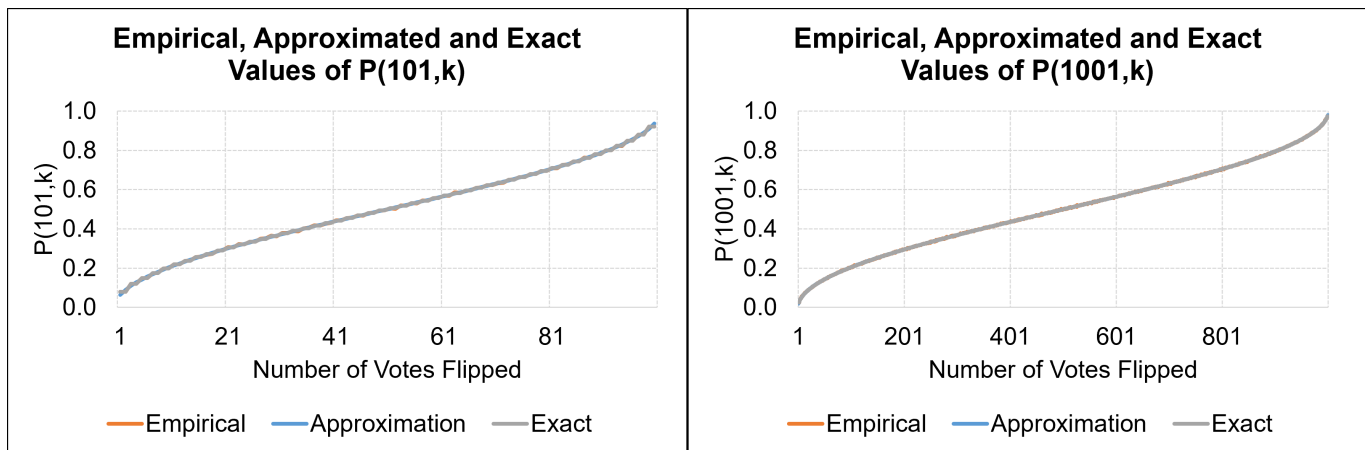


Figure. 3b Empirical, Approximate, and Exact values of $P(n,k)$ plotted together for $n = 101$ and $n = 1001$ for visual agreement

mation becomes far more accurate in larger elections. Hence, these observations validate both the accuracy of the exact formula for $P(n,k)$, as well as the use of the approximation when exact evaluation is inefficient (Figure 4).

These current formulas and approximations are limited to 2-candidate voting systems and model only popular voting, where outcomes are determined by direct aggregation of individual votes. The present model does not accommodate multi-candidate elections or alternate voting systems, each of which would introduce additional combinatorial conclusions and alter the probability for a voting system winner to change under random vote flips. Future work may be extended to alternative voting systems such as the tournament voting system discussed previously or multi-candidate elections.

Furthermore, the exact formula for $P(n,k)$ remains without a combinatorial proof, despite its correctness being supported both algebraically and through simulation. A combinatorial proof may provide an alternative perspective when approaching extensions and future works. This open problem may be approached by considering combinatorial interpretations for the quantities $\binom{2i}{i}$ and $\binom{2x-2i}{x-i}$ for values of i ranging from 0 to $\lfloor \frac{k-1}{2} \rfloor$, inclusive, likely calling for a casework argument on the value of i where the first $2i$ votes are considered separately from $2x - 2i$ other votes.

While this work is purely mathematical, the quantities derived here have implications for both election security and potential misuse, from which ethical considerations arise. On one hand, $P(n,k)$ and its approximation could in principle^{21,22} inform an adversary seeking to manipulate an election outcome by altering or coercing a target number of votes²³, since it quantifies precisely how many votes would need to be flipped to have a given probability of changing the result. Notably, our finding that this probability depends asymptotically only on the ratio $\frac{k}{n}$, not on n itself, implies that large elec-

tions are not inherently more resistant, in relative terms, to a fixed proportion of compromised votes than small elections. This result should inform, rather than discourage, efforts to secure elections at every scale. On the other hand, the same formula is directly useful for election security as it provides a principled way to quantify how many ballots would need to be audited or how small a margin of victory should be considered significant to detect or rule out manipulation at a given confidence level. This connects naturally to existing work on risk-limiting audits in election administration^{24,25}, where the central question, how many ballots must be checked to be statistically confident the reported outcome is correct, is closely related to the quantities we derive here.

We emphasize that this work concerns idealized, randomly-flipped votes under a simplified two-candidate model, and does not constitute a practical attack. Real-world election manipulation involves substantial additional constraints such as detection risk, verification procedures, and physical/digital security that are not modeled here.

References

- 1 A. Gelman, J. N. Katz, F. Tuerlinckx. The Mathematics and Statistics of Voting Power. *Statistical Science*. Vol. 17, pg. 420-425, 2002, <http://doi.org/10.1214/ss/1049993201>.
- 2 A. Gelman, N. Silver, A. Edlin. What is the probability your vote will make a difference? *Economic Inquiry*. Vol. 50, pg. 321-326, 2012, <http://doi.org/10.1111/j.1465-7295.2010.00272.x>.
- 3 M. Harrison-Trainor. An analysis of random elections with large numbers of voters. *Mathematical Social Sciences*. Vol. 116, pg. 68-84, 2022, <https://doi.org/10.1016/j.mathsocsci.2022.01.002>.
- 4 S. Cho, J. Jun. Flip-flopping and valence in two-candidate competition. *Journal of Theoretical Politics*. Vol. 38, pg. 3-27, 2025, <https://doi.org/10.1177/09516298251334320>.
- 5 P. Klimek, A. Aykaç, S. Thurner. Forensic analysis of the Turkey 2023 presidential election reveals extreme vote swings in remote areas. *PLOS*

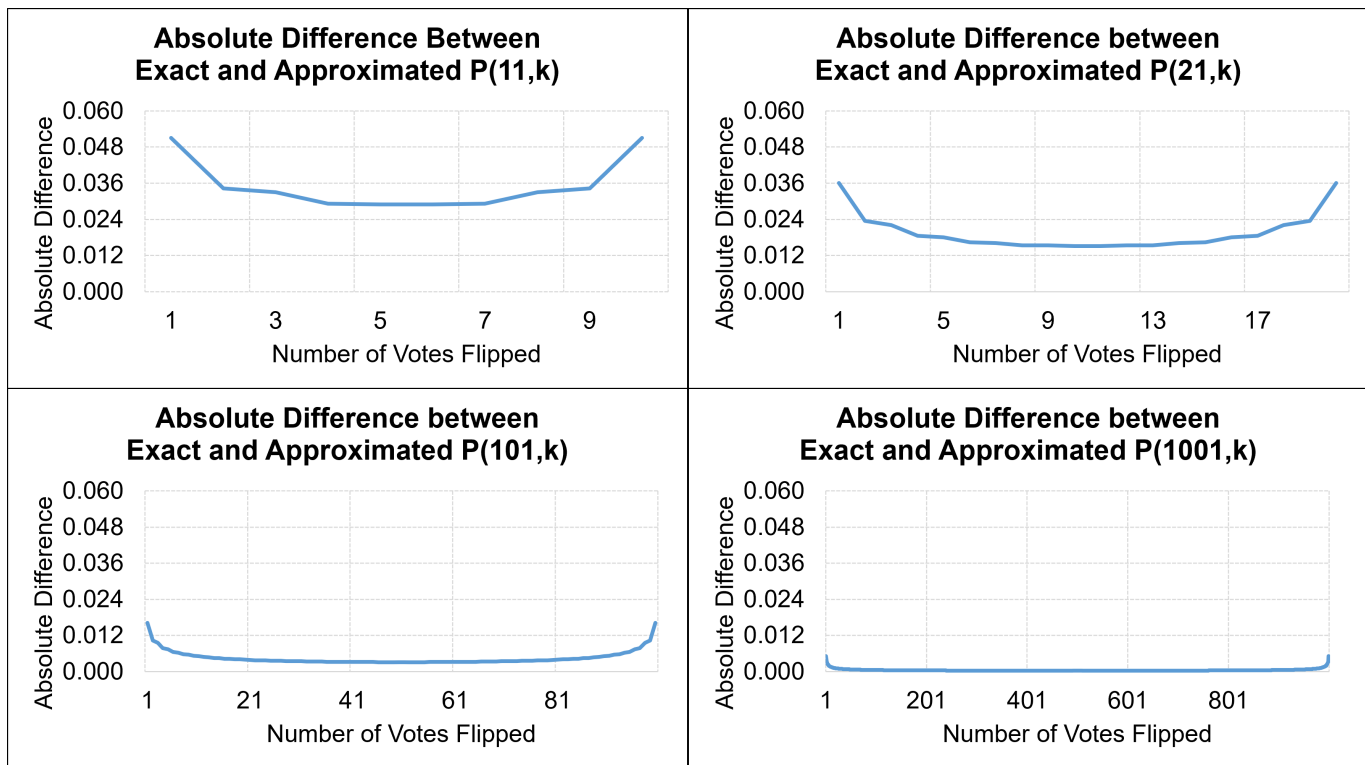


Figure. 4 Absolute difference between the exact value and the large- n approximation of $P(n,k)$, for $n = 11$, $n = 21$, $n = 101$, and $n = 1001$

- ONE. Vol. 18, pg. 1-15, 2023, <https://doi.org/10.1371/journal.pone.0293239>.
- 6 L. S. Penrose. The elementary statistics of majority voting. *Journal of the Royal Statistical Society*. Vol. 109, pg. 53-57, 1946, <https://doi.org/10.1111/j.2397-2335.1946.tb04638.x>.
 - 7 W. V. Gehrlein, D. Lepelley. The probability that all weighted scoring rules elect the same winner. *Economics Letters*. Vol. 66, pg. 191-197, 2000, [https://doi.org/10.1016/S0165-1765\(99\)00224-4](https://doi.org/10.1016/S0165-1765(99)00224-4).
 - 8 W. Bruns, B. Ichim. Computations of volumes in five candidates elections. *Scientific Reports*. Vol. 13, pg. 13266, 2023, <https://doi.org/10.1038/s41598-023-39656-8>.
 - 9 D. C. Brody, T. Yuasa. Three-candidate election strategy. *Royal Society Open Science*. Vol. 10, pg. 230584, 2023, <https://doi.org/10.1098/rsos.230584>.
 - 10 R. Bredereck, P. Faliszewski, A. Kaczmarczyk, R. Niedermeier, P. Skowron, N. Talmon. Robustness among multiwinner voting rules. *Artificial Intelligence*. Vol. 290, pg. 103403, 2021, <https://doi.org/10.1016/j.artint.2020.103403>.
 - 11 A. Bhattacharyya, P. Dey. Predicting winner and estimating margin of victory in elections using sampling. *Artificial Intelligence*. Vol. 296, pg. 1-23, 2021, <https://doi.org/10.1016/j.artint.2021.103476>.
 - 12 L. Böttcher, G. Kernell. Examining the limits of the Condorcet jury theorem: Tradeoffs in hierarchical information aggregation systems. *SAGE Open*. Vol. 12, pg. 1-12, 2022, <https://doi.org/10.1177/26339137221133401>.
 - 13 E. Elkind, E. Markakis, S. Obraztsova, P. Skowron. Equilibria of Plurality Voting: Lazy and Truth-biased Voters. *Symposium on Algorithmic Game Theory*. Vol. 8, pg. 1-27, 2015, https://doi.org/10.1007/978-3-662-48433-3_9.
 - 14 G. H. V. Brunello, E. Y. Nakano. Bayesian inference on proportional elections. *PLOS ONE*. Vol. 10, pg. 1-11, 2015, <https://doi.org/10.1371/journal.pone.0116924>.
 - 15 M. Gao, Z. Wang, K. Wang, C. Liu, S. Tang. Forecasting elections with agent-based modeling: Two live experiments. *PLOS ONE*. Vol. 17, pg. 1-11, 2022, <https://doi.org/10.1371/journal.pone.0270194>.
 - 16 I. Mortici. A substantial improvement of the Stirling formula. *Applied Mathematics Letters*. Vol. 24, pg. 1-4, 2011, <https://doi.org/10.1016/j.aml.2011.03.008>.
 - 17 R. P. Brent, J. H. Osborn, W. D. Smith. Asymptotic approximation of central binomial coefficients with rigorous error bounds. *Open Journal of Mathematical Sciences*. Vol. 4, pg. 261-275, 2021, <https://doi.org/10.30538/oms2021.0173>.
 - 18 N. Elezović. Asymptotic expansions of central binomial coefficients and Catalan numbers. *Journal of Integer Sequences*. Vol. 17, pg. 1-16, 2014, <https://cs.uwaterloo.ca/journals/JIS/VOL17/Elezovic/elezovic5.pdf>.
 - 19 F. Dubeau. On Euler-Maclaurin formula. *Journal of Computational and Applied Mathematics*. Vol. 296, pg. 649-660, 2016, <https://doi.org/10.1016/j.cam.2015.10.023>.
 - 20 S. J. Brams, P. C. Fishburn. Coalition voting. *Mathematical and Computer Modelling*. Vol. 16, pg. 15-26, 1992, [https://doi.org/10.1016/0895-7177\(92\)90084-X](https://doi.org/10.1016/0895-7177(92)90084-X).
 - 21 M. Zhang, R. M. Alvarez, I. Levin. Election forensics: Using machine learning and synthetic data for possible election anomaly detection. *PLOS ONE*. Vol. 14, pg. 1-14, 2019, <https://doi.org/10.1371/journal.pone.0223950>.
 - 22 A. Rozenas. Detecting election fraud from irregularities in vote-share

-
- distributions. *Political Analysis*. Vol. 25, pg. 41-56, 2017, <https://doi.org/10.1017/pan.2016.9>.
- 23 J. Gilbert, I. Laurenceau, J. Louis. A Study of Ballot Anomaly Detection with a Transparent Voting Machine. *Interactions*. Vol. 28, pg. 57-61, 2021, <http://doi.org/10.1145/3484937>.
- 24 J. Jaffe, J. R. Loffredo, S. Baltz, A. Flores, C. Stewart III. Trust in the count: Improving voter confidence with post-election audits. *Public Opinion Quarterly*. Vol. 88, pg. 585-607, 2024, <https://doi.org/10.1093/poq/nfae029>.
- 25 P. Klimek, Y. Yegorov, R. Hanel, S. Thurner. Statistical detection of systematic election irregularities. *Proceedings of the National Academy of Sciences*. Vol. 109, pg. 1-5, 2012, <https://doi.org/10.1073/pnas.1210722109>.