A Quantitative Approach to Incentives: Application to Voting Rules

Online Appendices

Gabriel Carroll, Microsoft Research and Stanford University
gdc@alum.mit.edu

May 25, 2013

This file contains the omitted proofs of results from the main paper, "A Quantitative Approach to Incentives: Application to Voting Rules." Numbering of results and equations here follows the convention from the main text. Bibliography entries appear at the end of this file.

C Computational tools

The present section gathers a collection of technical tools used in subsequent calculations. It includes proofs of the preliminary results stated in Subsection 2.2 of the main paper.

The following notation, not introduced in the main paper, will be useful here and subsequently. For a vector $x = (x_1, \ldots, x_r)$, we will write x_{-k} for the vector of all components except x_k, and x_{-jk} for the vector of all components except x_j and x_k; and x_{-ijk} similarly. We will write x_{j+k} for the sum of components x_j and x_k. Notice that if x_{-jk} and $\sum_l x_l$ are given then x_{j+k} is uniquely determined.

One other useful bit of notation: if f is a function of N and c a constant, we write $f(N) \overset{e}{\sim} c$ to say that $f(N)$ converges exponentially fast to c, i.e. $|f(N) - c| \lesssim e^{-\lambda N}$ for some $\lambda > 0$ (as in the statement of Lemma 2.2(a)).

Lemma C.1 (Stirling's approximation) *For any positive integer* K*,*

$$K! = \sqrt{2\pi K}(K/e)^K \iota \qquad \text{with } 1 < \iota < e^{1/12K}.$$

We cite this without proof; see e.g. [1, eq. 6.1.38].

Proof of Lemma 2.1: Expand the probability explicitly, and apply Lemma C.1 to the factorials. Since the ι factors all tend to 1 as $N \to \infty$, we get

$$\begin{aligned}
\mathbf{P}\left(\begin{array}{c} x_N \\ N - x_N \end{array} \middle| N; \begin{array}{c} \beta_N \\ 1 - \beta_N \end{array} \right) &\sim \frac{\sqrt{2\pi N}\left(\frac{N}{e}\right)^N \cdot \beta_N^{x_N}(1-\beta_N)^{N-x_N}}{\sqrt{2\pi x_N}\left(\frac{x_N}{e}\right)^{x_N} \cdot \sqrt{2\pi(N-x_N)}\left(\frac{N-x_N}{e}\right)^{N-x_N}} \\
&\sim \left(\frac{\beta_N N}{x_N}\right)^{x_N} \left(\frac{(1-\beta_N)N}{N-x_N}\right)^{N-x_N} \frac{1}{\sqrt{2\pi N \beta_N(1-\beta_N)}}.
\end{aligned}$$

We know $\beta_N \to \beta$ (since $|(\beta - \beta_N)N| < 2c$), so the result will follow if we can show

$$\left(\frac{\beta_N N}{x_N}\right)^{x_N} \left(\frac{(1-\beta_N)N}{N-x_N}\right)^{N-x_N} \to 1. \tag{C.1}$$

Now, the logarithm of the left-hand side of (C.1) is $Nh(x_N/N, \beta_N)$, where

$$h(\gamma, \delta) = \gamma(\ln \delta - \ln \gamma) + (1-\gamma)(\ln(1-\delta) - \ln(1-\gamma)).$$

The derivative of h with respect to its first argument is

$$\frac{\partial h}{\partial \gamma} = \ln \frac{\delta}{\gamma} - \ln \frac{1-\delta}{1-\gamma}.$$

In particular, $\partial h/\partial \gamma$ is continuous on $(0,1) \times (0,1)$ and is zero when $\gamma = \delta$. We also have $h(\beta_N, \beta_N) = 0$, and so

$$\begin{aligned}
\left| Nh\left(\frac{x_N}{N}, \beta_N\right)\right| &\leq N\left|\frac{x_N}{N} - \beta_N\right| \cdot \max_{\gamma \in \left[\frac{x_N}{N}, \beta_N\right]} \left|\frac{\partial h}{\partial \gamma}(\gamma, \beta_N)\right| \\
&< c \cdot \max_{\gamma \in \left[\frac{x_N}{N}, \beta_N\right]} \left|\frac{\partial h}{\partial \gamma}(\gamma, \beta_N)\right| \\
&\to c \cdot \frac{\partial h}{\partial \gamma}(\beta, \beta) \\
&= 0.
\end{aligned}$$

(The notation assumes $x_N/N \leq \beta_N$, but of course an identical argument applies when

$\beta_N < x_N/N$.) Then (C.1) follows. $\square$

Proof of Lemma 2.3: Taking logs and ignoring the constant, we see the problem is to maximize $\sum_i x_i \ln \alpha_i$ subject to $\sum_i \alpha_i = 1$. This is a concave maximization problem; the solution is given by the first-order condition $x_i/\alpha_i = \lambda$ for all i, where λ is the Lagrange multiplier on the constraint. Hence the α_i must be proportional to the x_i at the maximum. $\square$

Lemma C.2 *For $1 \leq q < r$, we have*

$$\mathbf{P}\left(\begin{array}{c|cc} x_1 & & \alpha_1 \\ \vdots & K; & \vdots \\ x_r & & \alpha_r \end{array}\right) = \mathbf{P}\left(\begin{array}{c|cc} x_1 & & \alpha_1 \\ \vdots & & \vdots \\ x_q & K; & \alpha_q \\ x & & \alpha \end{array}\right) \cdot \mathbf{P}\left(\begin{array}{c|cc} x_{q+1} & & \alpha_{q+1}/\alpha \\ \vdots & x; & \vdots \\ x_r & & \alpha_r/\alpha \end{array}\right)$$

where $x = x_{q+1} + \cdots + x_r$ and $\alpha = \alpha_{q+1} + \cdots + \alpha_r$ (assuming $\alpha > 0$).

This is the familiar decomposition property of the multinomial distribution: given that $K - x$ voters are of the first q types and the remaining x voters are of the remaining $r - q$ types, the distribution of types among the last x voters is independent of the distribution among the first $K - x$ voters (and in particular is again multinomial $\mathbf{M}(x; \alpha_{q+1}/\alpha, \ldots, \alpha_r/\alpha)$.

Proof: Immediate from the definitions. $\square$

Lemma C.3

$$\sum_{x \text{ even}} \mathbf{P}(x, N - x \mid N; \alpha, 1 - \alpha) = (1 + (1 - 2\alpha)^N)/2.$$

Proof: Write the right-hand side as $(((1-\alpha)+\alpha)^N + ((1-\alpha)-\alpha)^N)/2$; expanding by the binomial theorem, the terms with odd powers of α cancel and we get $\sum_{x \text{ even}} \binom{N}{x}(1-\alpha)^{N-x}\alpha^x$ which is the left-hand side. $\square$

Proof of Lemma 2.2:

(a) Choose ϵ sufficiently small such that if $(\alpha_1, \ldots, \alpha_r) \in J$ and $|\beta_j - \alpha_j| < \epsilon$ for each index j, then $\beta_1, \ldots, \beta_r$ must still satisfy the inequalities $\mathcal{I}$. (We can do this since J is compact and the inequalities $\mathcal{I}$ carve out an open set.) We can find $\kappa < 1$ such that

$$\frac{\alpha^\beta (1-\alpha)^{1-\beta}}{\beta^\beta (1-\beta)^{1-\beta}} < \kappa \qquad \text{for all } \alpha, \beta \in [0,1] \text{ with } \qquad |\beta - \alpha| \geq \epsilon, \tag{C.2}$$

where we interpret 0^0 as 1. Indeed, the denominator of the left side of (C.2) is bounded away from 0, whereas as $\alpha \to 0$ the numerator is $\leq \alpha^\epsilon$ and so converges uniformly to 0 for $\beta \in [\epsilon, 1]$; likewise as $\alpha \to 1$ the numerator is $\leq (1-\alpha)^{1-\epsilon}$ and so converges uniformly to 0 for $\beta \in [0, 1-\epsilon]$. This shows that for some $\eta > 0$, we can choose $\kappa < 1$ to ensure that (C.2) holds when $\alpha \leq \eta$ or $\alpha \geq 1-\eta$. Otherwise, use the fact that the logarithm of the left side of (C.2) is $\beta(\ln\alpha - \ln\beta) + (1-\beta)(\ln(1-\alpha) - \ln(1-\beta))$. This expression is continuous on the rectangle $[\alpha, \beta] \in [\eta, 1-\eta] \times [0, 1]$, and takes its maximum value of zero only at $\alpha = \beta$ (by Lemma 2.3), and therefore is bounded strictly below 0 for $|\alpha - \beta| \geq \epsilon$. Statement (C.2) follows.

Now take any $(\alpha_1, \ldots, \alpha_r) \in J$. Consider any given index j, and any value x_j with $|x_j/N - \alpha_j| > \epsilon$. Let $\beta_j = x_j/N$. The probability that the realized j-th component is x_j is

$$\begin{aligned} \mathbf{P}\left(\begin{array}{c} x_j \\ N - x_j \end{array} \middle| \, N; \begin{array}{c} \alpha_j \\ 1 - \alpha_j \end{array} \right) &= \mathbf{P}\left(\begin{array}{c} x_j \\ N - x_j \end{array} \middle| \, N; \begin{array}{c} \beta_j \\ 1 - \beta_j \end{array} \right) \times \\ &\qquad \left(\frac{\alpha_j^{\beta_j}(1-\alpha_j)^{1-\beta_j}}{\beta_j^{\beta_j}(1-\beta_j)^{1-\beta_j}} \right)^N \\ &\leq \kappa^N. \end{aligned}$$

There are r possible choices of index j and at most $N+1$ values x_j to consider for any given j, so the total probability that some event $|x_j/N - \alpha_j| > \epsilon$ occurs is at most $r(N+1)\kappa^N$. This bound still decays exponentially in N, and is independent of the choice of $(\alpha_1, \ldots, \alpha_r) \in J$.

(b) Fix arbitrarily small $\epsilon > 0$. We will show that

$$\frac{1}{2}\sqrt{\frac{2}{\pi(2\alpha_i + \epsilon)N}} \lesssim \mathbf{P}(S_N^{\mathcal{I}} \cap T_{ij,y} \mid N; \alpha_1, \ldots, \alpha_r) \lesssim \frac{1}{2}\sqrt{\frac{2}{\pi(2\alpha_i - \epsilon)N}} \tag{C.3}$$

and the conclusion will follow by taking $\epsilon \to 0$.

Let $S'_N = \{(x_1, \ldots, x_r) \mid (2\alpha_i - \epsilon)N < x_i + x_j < (2\alpha_i + \epsilon)N\}$. By (a), the probability of drawing a profile in $S_N^{\mathcal{I}}$ and the probability of drawing a profile in S'_N both go to 1 exponentially as $N \to \infty$.

Let S_N^{par} be the set of profiles such that $x_i - x_j - y$ is even, or equivalently $x_i + x_j - y$ is even. Certainly $T_{ij,y} \subseteq S_N^{par}$. From Lemma C.2, $(x_i + x_j, \sum_{k \neq i,j} x_k)$ is multinomial with parameters $N; \alpha_i + \alpha_j, 1 - (\alpha_i + \alpha_j)$. So the probability of drawing a profile in S_N^{par} is $(1 \pm (1 - 2(\alpha_i + \alpha_j))^N)/2$, by Lemma C.3. This converges exponentially to $1/2$ as $N \to \infty$.

Write p_N for the probability that $P \in T_{ij,y}$, *conditional* on $P \in S'_N \cap S_N^{par}$. Because the probabilities of drawing profiles in $S_N^{\mathcal{I}}, S'_N, S_N^{par}$ converge exponentially to $1, 1, 1/2$ respectively, it suffices to show that p_N satisfies

$$\sqrt{\frac{2}{\pi(2\alpha_i + \epsilon)N}} < p_N < \sqrt{\frac{2}{\pi(2\alpha_i - \epsilon)N}} \tag{C.4}$$

and then (C.3) will follow.

For any given N, fix any value of the subvector x_{-ij}, such that

$$(1 - 2\alpha_i - \epsilon)N < \sum_{k \neq i,j} x_k < (1 - 2\alpha_i + \epsilon)N$$

and

$$x_{i+j} = N - \sum_{k \neq i,j} x_k \text{ is the same parity as } y.$$

Also write x_{i+j}^{max} and x_{i+j}^{min} for the maximum and minimum possible values of x_{i+j} subject to these conditions. Note that whether or not $P \in S'_N \cap S_N^{par}$ depends only on x_{-ij}.

Conditional on the values x_{-ij}, the remaining coordinates (x_i, x_j) are distributed $\mathbf{M}(x_{i+j}; 1/2, 1/2)$ by Lemma C.2. Moreover $x \in T_{ij,y}$ if and only if $x_i - x_j = y$, or equivalently $x_i = (x_{i+j} + y)/2$ (which is an integer). Hence, conditional on x_{-ij}, the probability that $x \in T_{ij,y}$ is

$$h_y(x_{i+j}) = \mathbf{P}\left(\begin{array}{c} (x_{i+j} + y)/2 \\ (x_{i+j} - y)/2 \end{array} \middle| \, x_{i+j}; \begin{array}{c} 1/2 \\ 1/2 \end{array} \right).$$

Applying Lemma 2.1 together with $x_{i+j}^{min} \sim (2\alpha_i - \epsilon)N, x_{i+j}^{max} \sim (2\alpha_i + \epsilon)N$ gives

$$\sqrt{\frac{2}{\pi(2\alpha_i + \epsilon)N}} \lesssim \min_{x_{i+j}} h_y(x_{i+j}) \leq \max_{x_{i+j}} h_y(x_{i+j}) \lesssim \sqrt{\frac{2}{\pi(2\alpha_i - \epsilon)N}},$$

where the maxima are taken over $x_{i+j} \in [x_{i+j}^{min}, x_{i+j}^{max}]$. For each realization of x_{-ij}, the conditional probability of $x \in T_{ij,y}$ lies between $\min h_y(x_{i+j})$ and $\max h_y(x_{i+j})$, so the overall probability of $x \in T_{ij,y}$ also lies in between these bounds. At this point (C.4) follows.

As already shown, this in turn implies (C.3), and the proof of part (b) is complete.

□

Proof of Lemma 2.4: For any K, the maximum is attained by $\alpha = K/N$ by Lemma 2.3. Hence it suffices to study the behavior with respect to K of the expression $\binom{N}{K}(K/N)^K((N-K)/N)^{N-K}$, or equivalently of $b(K) = K^K(N-K)^{N-K}/K!(N-K)!$. In particular, by symmetry it suffices to show that $b(K)$ is strictly increasing for $K \geq N/2$.

Put $c(K) = K^K/K!$. Notice that $c(K+1)/c(K) = (1+1/K)^K$ which is increasing in K (this can be verified directly by taking the logarithm and differentiating). Hence for $K \geq N/2$ we have

$$\frac{b(K+1)}{b(K)} = \frac{c(K+1)c(N-K-1)}{c(K)c(N-K)} = \frac{c(K+1)}{c(K)} \Big/ \frac{c(N-K)}{c(N-K-1)} > 1$$

because $K > N-K-1$. □

Next we give a simple bound on the probability of large deviations under multinomial distributions.

Lemma C.4 *For all N, K, α,*

$$\mathbf{P}(K, N-K \mid N; \alpha, 1-\alpha) \leq e^{-N\cdot\frac{(\alpha-K/N)^2}{2}}.$$

(One can obtain a slightly stronger bound from Hoeffding's Inequality [3], but the proof here is self-contained.)

Proof: Consider the function $h(\alpha) = \ln \mathbf{P}(K, N-K \mid N; \alpha, 1-\alpha)$, whose maximum is at $\alpha = K/N$ by Lemma 2.3, and its value there is certainly at most 0. Moreover $d^2h/d\alpha^2 = -(K/\alpha^2 + (N-K)/(1-\alpha)^2)$. Now by Cauchy-Schwarz,

$$\left(\frac{K}{\alpha^2} + \frac{N-K}{(1-\alpha)^2}\right)\left(\alpha^2 + (1-\alpha)^2\right) \geq (\sqrt{K} + \sqrt{N-K})^2 \geq N.$$

Then $d^2h/d\alpha^2 \leq -N$ so $h(\alpha) \leq -N(\alpha - K/N)^2/2$. □

The next two results concern the quantity σ_N^*, defined in Subsection 3.1.

Lemma C.5

$$e^{-\frac{1}{3(N-1)}}\sqrt{\frac{2}{\pi N}} < \sigma_N^* < e^{\frac{1}{12N}}\sqrt{\frac{2N}{\pi(N^2-1)}}.$$

Proof: Put $\alpha = 1/2$ if N is even and $(N-1)/2N$ if N is odd. By Lemma C.1, write

$$\sigma_N^* = \binom{N}{\alpha N}\alpha^{\alpha N}(1-\alpha)^{(1-\alpha)N}$$

$$= \frac{\iota_N (N/e)^N\sqrt{2\pi N}}{[\iota_{\alpha N}(\alpha N/e)^{\alpha N}\sqrt{2\pi\alpha N}]\cdot[\iota_{(1-\alpha)N}((1-\alpha)N/e)^{(1-\alpha)N}\sqrt{2\pi(1-\alpha)N}]}\alpha^{\alpha N}(1-\alpha)^{(1-\alpha)N}$$

where the three ι_x terms satisfy $1 < \iota_x < e^{1/12x}$. Cancelling common factors reduces to

$$\frac{\iota_N}{\iota_{\alpha N}\iota_{(1-\alpha)N}}\cdot\sqrt{\frac{1}{2\pi N\alpha(1-\alpha)}}.$$

Both αN and $(1-\alpha)N$ are at least $(N-1)/2$, hence $e^{-1/3(N-1)} < \iota_N/\iota_{\alpha N}\iota_{(1-\alpha)N} < e^{1/12N}$; and $\alpha(1-\alpha) \in \{(N-1)^2/4N^2, 1/4\}$, hence the square-root term is either $\sqrt{2/\pi N}$ or $\sqrt{2N/\pi(N^2-1)}$. □

Corollary C.6 *σ_N^* is decreasing in N.*

Proof: For $N < 15$, $\sigma_N^* < \sigma_{N-1}^*$ can be verified by direct computation. For $N \geq 15$, Lemma C.5 implies that it is sufficient to check that

$$e^{\frac{1}{12N}}\sqrt{\frac{N}{N^2-1}} < e^{-\frac{1}{3(N-2)}}\sqrt{\frac{1}{N-1}} \tag{C.5}$$

or equivalently

$$e^{\frac{1}{12N}+\frac{1}{3(N-2)}}\sqrt{\frac{N}{N+1}} < 1. \tag{C.6}$$

Since $((N+1)/N)^{N+1} > e$, we have $\sqrt{N/(N+1)} < e^{-1/2(N+1)}$, so (C.6) follows from the inequality $1/12N + 1/3(N-2) < 1/2(N+1)$ which holds for $N \geq 15$. □

We provide a few more useful bounds.

Lemma C.7 *If $x, y \geq K > 0$, then for all α we have*

$$\mathbf{P}(x, y \mid x+y; \alpha, 1-\alpha) \leq \frac{e^{1/12}}{\sqrt{\pi K}}.$$

Proof: By Lemma 2.3, the probability is maximized by taking $\alpha = x/(x+y)$. In this case, we can write the probability explicitly using Lemma C.1 and simplify as in Lemma C.5 to obtain

$$\mathbf{P}\left(\begin{array}{c} x \\ y \end{array} \middle| \, x+y; \begin{array}{c} x/(x+y) \\ y/(x+y) \end{array}\right) \leq \frac{e^{1/12}\sqrt{2\pi(x+y)}}{\sqrt{2\pi x}\sqrt{2\pi y}}.$$

Either $(x+y)/x \leq 2$ or $(x+y)/y \leq 2$, so we can cancel the numerator radical with one of the denominator radicals and a $\sqrt{2}$ factor, and the result follows. □

Lemma C.8 *There exists an absolute constant $c > 0$ with the following property. For every positive integer N and every nonempty subset $S \subseteq \{0, \ldots, N\}$, there exists $\alpha \geq \max(S)/N$ such that*

$$\sum_{K \in S}\left[\mathbf{P}\left(\begin{array}{c} K \\ N-K \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array}\right) - \mathbf{P}\left(\begin{array}{c} K-1 \\ N-K+1 \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array}\right)\right] \geq \frac{c}{N}.$$

Proof: It suffices to prove the lemma when $S = \{K\}$. Indeed, since $\mathbf{P}(K, N-K \mid N; \alpha, 1-\alpha)$ is increasing in K when $K \leq \alpha(N+1)$, every term on the left-hand side of the inequality in the lemma is nonnegative as long as $\alpha \geq \max(S)/N$, so it suffices to show that the term corresponding to $K = \max(S)$ is at least c/N.

So let $S = \{K\}$. If $K = N$ then take $\alpha = 1$. If $K = 0$ then take $\alpha = 0$. Otherwise, let $L = K + \lfloor\sqrt{K(N-K)/N}\rfloor$; we will show that $\alpha = L/N$ does the job. (Note that $L \leq N$, i.e. $\alpha \leq 1$.) We have

$$\begin{aligned}
&\mathbf{P}\left(\begin{array}{c} K \\ N-K \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array}\right) - \mathbf{P}\left(\begin{array}{c} K-1 \\ N-K+1 \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array}\right) \\
&\qquad = \mathbf{P}\left(\begin{array}{c} K \\ N-K \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array}\right) \cdot \left[1 - \frac{K}{N-K+1} \cdot \frac{1-\alpha}{\alpha}\right] \\
&\qquad = \mathbf{P}\left(\begin{array}{c} L \\ N-L \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array}\right) \cdot \\
&\qquad\qquad \left[\prod_{k=K}^{L-1} \frac{k+1}{N-k}\frac{1-\alpha}{\alpha}\right] \cdot \left[1 - \frac{K}{N-K+1} \cdot \frac{1-\alpha}{\alpha}\right].
\end{aligned}$$

Now, the middle bracketed expression is a product consisting of $L - K$ factors, each of which is greater than

$$\frac{K}{N-K} \cdot \frac{1-\alpha}{\alpha} = \frac{K(N-L)}{L(N-K)} \geq 1 - \frac{1}{L-K}$$

(to verify the last inequality, cross-multiply and rearrange terms to find that it is equivalent to $(L-K)^2 N \le L(N-K)$, which is true). Hence this product is

$$> \left(1 - \frac{1}{L-K}\right)^{L-K} \ge \frac{1}{4}$$

as long as $L - K \ge 2$. Otherwise, $L - K = 0$ and the middle product is empty, or else $L - K = 1$ and the middle product equals $(N-K-1)/(N-K) \ge 1/2$ (notice that if $K = N-1$ then $L = K$). Hence in every case the middle bracketed expression is $\ge 1/4$.

It therefore suffices to show that there is some constant c' such that the bound

$$\mathbf{P}\left(\begin{array}{c} L \\ N-L \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array}\right) \cdot \left[1 - \frac{K}{N-K+1} \cdot \frac{1-\alpha}{\alpha}\right] \ge \frac{c'}{N} \tag{C.7}$$

always holds. We split into three cases.

- Suppose $K \le N/2$ and $L > K$. The $\mathbf{P}(\cdots)$ factor is bounded below by $\sigma_N^* \gtrsim \sqrt{2/\pi N}$, by Lemma 2.4. Also, $L - K > 0$ implies $(1/2)\sqrt{K(N-K)/N} \le L - K \le K$, so

$$\begin{aligned} 1 - \frac{K}{N-K+1} \cdot \frac{1-\alpha}{\alpha} &\ge 1 - \frac{K(N-L)}{L(N-K)} \\ &= \frac{(L-K)N}{L(N-K)} \\ &\ge \frac{L-K}{L} \\ &\ge \frac{L-K}{2K} \\ &\ge \frac{1}{4}\sqrt{\frac{N-K}{NK}} \\ &= \frac{1}{4}\sqrt{\frac{1}{K} - \frac{1}{N}} \\ &\ge \frac{1}{4}\sqrt{\frac{1}{N}} \end{aligned}$$

  where there last step uses the assumption $K \le N/2$. So each of the two factors on the left side of (C.7) is bounded below by a constant times $\sqrt{1/N}$.

- Suppose $K > N/2$ and $L > K$. In this case, we apply Stirling's approximation (C.1) as usual to observe that $\mathbf{P}(L, N-L \mid N; \alpha, 1-\alpha)$ is bounded below by a constant times $\sqrt{N/L(N-L)}$. Combining with the chain of inequalities from the previous case, we see that the left side of (C.7) is bounded below by a constant times

$$\sqrt{\frac{N}{L(N-L)}} \cdot \frac{1}{4}\sqrt{\frac{N-K}{NK}} \geq \frac{1}{4}\sqrt{\frac{N}{K(N-K)}} \cdot \sqrt{\frac{N-K}{NK}} = \frac{1}{4K} \geq \frac{1}{4N}.$$

- Finally suppose $L = K$. This can only happen for $K = 1$ or $N - 1$, or for small N (which we can ignore since the result is asymptotic), and so we verify (C.7) directly in these cases. We have $\mathbf{P}(L, N-L \mid N; \alpha, 1-\alpha) = ((N-1)/N)^{N-1} \geq 1/e$, a constant. If $K = 1$ then the second factor in (C.7) is $1/N$; if $K = N-1$ then this factor is $1/2$.

This verifies that (C.7) holds in every case.

□

Lemma C.9 *Fix any positive constant c. If N is taken large enough and $\alpha \leq c/\sqrt{N}$ then*

$$\sum_{K=\lceil 3c\sqrt{N}\rceil}^{N} \mathbf{P}\left(\begin{array}{c} K \\ N-K \end{array} \middle| \, N; \begin{array}{c} \alpha \\ 1-\alpha \end{array} \right) \leq \frac{1}{N}.$$

(Actually the left side goes to zero exponentially fast in $\sqrt{N}$, but this very crude bound is all we will need.)

Proof: Put $p(K) = \mathbf{P}(K, N-K \mid N; \alpha, 1-\alpha)$. We have

$$\frac{p(K+1)}{p(K)} = \frac{N-K}{K+1} \cdot \frac{\alpha}{1-\alpha} \leq \frac{N-K}{K} \cdot \frac{\alpha}{1-\alpha} \leq \frac{1}{2}$$

whenever $K \geq 2N\alpha$. Since $p(K) \leq 1$ for $K = \lceil 2N\alpha \rceil$, we have by induction $p(K) \leq 1/2^{K-\lceil 2N\alpha \rceil}$ for $K \geq 2N\alpha$, and therefore by the expression in the lemma statement is at most

$$\sum_{K=\lceil 3N\alpha \rceil}^{\infty} \frac{1}{2^{K-\lceil 2N\alpha \rceil}} = \frac{1}{2^{\lceil 3N\alpha \rceil - \lceil 2N\alpha \rceil - 1}} \leq \frac{1}{2^{c\sqrt{N}-2}} \lesssim \frac{1}{N}.$$

□

The remaining lemmas in this section are bounds on certain alternating sums of multinomial probabilities. These bounds are useful for the construction in Appendix H.

If S is a set of positive integers, let $\sigma(S)$ and $\pi(S)$ denote, respectively, the sum and the product of elements of S (with $\sigma(\emptyset) = 0, \pi(\emptyset) = 1$). This use of σ of course overlaps with the notation for susceptibility, but there should be no ambiguity.

Lemma C.10 *Fix $\epsilon > 0$ and $\underline{\alpha} \in (0, 1/2)$, and fix a positive integer d. There exists a threshold N_0 with the following property: For all $N > N_0$, all $\alpha \in [\underline{\alpha}, 1 - \underline{\alpha}]$, all integers K, and all sets S of positive integers with $|S| = d$,*

$$\left|\sum_{T\subseteq S}(-1)^{|T|}\mathbf{P}\left(\begin{array}{c} K-\sigma(T) \\ N-K+\sigma(T)\end{array}\;\middle|\; N;\; \begin{array}{c}\alpha \\ 1-\alpha\end{array}\right)\right| \leq \pi(S)N^{-d\left(\frac{1}{2}-\epsilon\right)}.$$

Proof: The expression inside the absolute value is (up to a sign) the coefficient of z^K in the polynomial

$$Q_{\alpha,S}(z) = \left[\prod_{s\in S}(z^s - 1)\right] \cdot (\alpha z + (1-\alpha))^N.$$

However, the standard formula for coefficient extraction using complex roots of unity tells us that this coefficient also equals

$$\frac{1}{L}\sum_{l=1}^{L}\zeta^{-Kl}Q_{\alpha,S}(\zeta^l),$$

where L is any integer greater than the degree of $Q_{\alpha,S}$ and ζ is a primitive Lth root of unity. Therefore, it suffices to show that for some N_0 the following holds: whenever $N > N_0$, for all choices of S and α and every complex number z with $|z| = 1$,

$$|Q_{\alpha,S}(z)| \leq \pi(S)N^{-d\left(\frac{1}{2}-\epsilon\right)}. \tag{C.8}$$

We consider two cases for z. Let $\theta = \arg z$.

- Suppose $|\theta| < N^{-\left(\frac{1}{2}-\epsilon\right)}$. Then $|z - 1| < N^{-\left(\frac{1}{2}-\epsilon\right)}$, from which

  $$|z^s - 1| = \left|\sum_{t=0}^{s-1} z^t(z-1)\right| \leq s|z-1| < sN^{-\left(\frac{1}{2}-\epsilon\right)}$$

  and then multiplying across all $s \in S$, together with $|\alpha z + (1-\alpha)| \leq 1$, gives (C.8).

- Otherwise, $|\theta| \geq N^{-\left(\frac{1}{2}-\epsilon\right)}$. As long as N is not too small, this implies

  $$\begin{aligned} |\alpha z + (1-\alpha)|^2 &= (1-\alpha+\alpha\cos\theta)^2 + (\alpha\sin\theta)^2 \\ &= (1-\alpha)^2 + \alpha^2 + 2(1-\alpha)\alpha\cos\theta \\ &< (1-\alpha)^2 + \alpha^2 + 2(1-\alpha)\alpha\sqrt{1-\frac{1}{4N^{1-2\epsilon}}} \end{aligned}$$

(this follows from $\cos^2 N^{-\left(\frac{1}{2}-\epsilon\right)} = 1 - \sin^2 N^{-\left(\frac{1}{2}-\epsilon\right)} < 1 - 1/4N^{1-2\epsilon}$)

$$\begin{aligned} &< (1-\alpha)^2 + \alpha^2 + 2(1-\alpha)\alpha\left(1 - \frac{1}{8N^{1-2\epsilon}}\right) \\ &= 1 - \frac{(1-\alpha)\alpha}{4N^{1-2\epsilon}} \\ &\leq 1 - \frac{c'}{N^{1-2\epsilon}} \end{aligned}$$

where $c' = (1-\underline{\alpha})\underline{\alpha}/4$. Hence

$$\begin{aligned} |\alpha z + (1-\alpha)|^N &< \left(1 - c'N^{-(1-2\epsilon)}\right)^{N/2} \\ &= \left[\left(1 - c'N^{-(1-2\epsilon)}\right)^{N^{1-2\epsilon}/2}\right]^{N^{2\epsilon}} \\ &< [\exp(-c'/2)]^{N^{2\epsilon}} \\ &\leq N^{-d\left(\frac{1}{2}-\epsilon\right)}/2^d \end{aligned}$$

as long as N is larger than some threshold that depends only on $\underline{\alpha}, \epsilon, d$. Since also $|z^s - 1| \leq 2$ for each $s \in S$, the bound (C.8) holds in this case also.

□

The next lemma will depend on the following notation. For N, K integers, $\alpha \in [0,1]$, and Z a set of integers, put

$$\Sigma(\alpha, Z, N, K) = \sum_{x \in Z} \mathbf{P}\left(\begin{array}{c} K - x \\ N - K + x \end{array} \middle| N; \begin{array}{c} \alpha \\ 1-\alpha \end{array} \right).$$

Lemma C.11 *Let d be a given positive integer. For every positive integer h, let Z^h denote the set $\{0, 1, \ldots, 2^{hd} - 1\}$. Then it is possible to partition each set Z^h into 2^h subsets $Z_0^h, Z_1^h, \ldots, Z_{2^h-1}^h$, of size $2^{h(d-1)}$ each, so that the following property is satisfied:*

For any $\epsilon > 0$ and $\underline{\alpha} \in (0, 1/2)$, there exists a threshold N_0 such that for all $N > N_0$, all $\alpha \in [\underline{\alpha}, 1-\underline{\alpha}]$, all h, and all integers K,

$$\left|\Sigma(\alpha, Z_i^h, N, K) - \Sigma(\alpha, Z_j^h, N, K)\right| \leq 2^{h(d^2+d-1)} h N^{-d\left(\frac{1}{2}-\epsilon\right)} \tag{C.9}$$

for any two sets Z_i^h, Z_j^h in the partition of Z^h.

Proof: We first describe the partition of Z^h. Consider each of the numbers $0, 1, \ldots, 2^{hd} - 1$ written out as a binary string with hd digits. We assign each such number x to a subset Z_i^h as follows:

- Divide the hd digits of x into h segments of d digits each;
- next, replace each segment with a 0 or a 1, depending whether the number of 1's in that segment is even or odd;
- finally, read the resulting h-digit string as a binary number $i \in \{0, 1, \ldots, 2^h - 1\}$, and assign x to Z_i^h.

It should be clear that each Z_i^h consists of exactly $2^{h(d-1)}$ values x.

Now let N_0 be the threshold given by Lemma C.10, with the same $\epsilon, \underline{\alpha}, d$ as in the current lemma. Clearly this threshold does not depend on h, so henceforth we will consider any fixed h, and drop the superscripts on the Z_i's. Assume $N > N_0$, and let $\alpha \in [\underline{\alpha}, 1 - \underline{\alpha}]$ be arbitrary.

It suffices to show that if the binary representations of i and j differ by just one digit, then for all K,

$$|\Sigma(\alpha, Z_i, N, K) - \Sigma(\alpha, Z_j, N, K)| \leq 2^{h(d^2+d-1)} N^{-d\left(\frac{1}{2}-\epsilon\right)}. \tag{C.10}$$

Indeed, since one can get from any i to any j by at most h single-digit changes, applying (C.10) repeatedly will then imply (C.9).

Without loss of generality, i has a 0 in the $(r+1)$th position from the right, while j has a 1 in that position; all other digits in the binary representations of i and j are the same. Then define three sets $Z'_\emptyset, Z'_i, Z'_j$:

- $Z'_\emptyset$ consists of all values of $x \in Z_i$ such that the $(dr+1)$th, $(dr+2)$th, $\ldots$, $(dr+d)$th digits from the right are all 0;
- Z'_i consists of all numbers that can be represented as a sum of an even number of elements of the set $\{2^{dr}, 2^{dr+1}, \ldots, 2^{dr+d-1}\}$;
- Z'_j consists of all numbers that can be represented as a sum of an odd number of elements of $\{2^{dr}, 2^{dr+1}, \ldots, 2^{dr+d-1}\}$.

Then, Z_i consists of all numbers that can be represented as a sum of an element of $Z'_\emptyset$ and one of Z'_i, and for each such number, the representation is unique. Likewise Z_j consists of numbers that can be represented (uniquely) as a sum of an element of $Z'_\emptyset$ and one of Z'_j.

Applying the conclusion of Lemma C.10 with $S = \{2^{dr}, 2^{dr+1}, \ldots, 2^{dr+d-1}\}$, and using the easy bound $\pi(S) \le 2^{d(dr+d-1)}$, gives the following: for any K,

$$\left| \sum_{x \in Z_i'} \mathbf{P}\left(\begin{array}{c} K-x \\ N-K+x \end{array} \middle| N; \begin{array}{c} \alpha \\ 1-\alpha \end{array} \right) - \sum_{x \in Z_j'} \mathbf{P}\left(\begin{array}{c} K-x \\ N-K+x \end{array} \middle| N; \begin{array}{c} \alpha \\ 1-\alpha \end{array} \right) \right| \le 2^{d(dr+d-1)} N^{-d\left(\frac{1}{2}-\epsilon\right)}. \tag{C.11}$$

Now replace K by $K-y$ for each possible $y \in Z_\emptyset'$, and sum over all y. We have

$$\sum_{y \in Z_\emptyset'} \sum_{x \in Z_i'} \mathbf{P}\left(\begin{array}{c} K-y-x \\ N-K+y+x \end{array} \middle| N; \begin{array}{c} \alpha \\ 1-\alpha \end{array} \right) = \sum_{x \in Z_i} \mathbf{P}\left(\begin{array}{c} K-x \\ N-K+x \end{array} \middle| N; \begin{array}{c} \alpha \\ 1-\alpha \end{array} \right),$$

and likewise for Z_j' and Z_j. Thus, summing (C.11) over the $2^{(d-1)(h-1)}$ choices of $y \in Z_\emptyset'$ and applying the triangle inequality gives

$$\begin{aligned} |\Sigma(\alpha, Z_i, N, K) - \Sigma(\alpha, Z_j, N, K)| &\le 2^{(d-1)(h-1)} \cdot 2^{d(dr+d-1)} N^{-d\left(\frac{1}{2}-\epsilon\right)} \\ &\le 2^{(d-1)(h-1)+d(hd-1)} N^{-d\left(\frac{1}{2}-\epsilon\right)}. \end{aligned}$$

Since $(d-1)(h-1) + d(hd-1) \le h(d^2+d-1)$, (C.10) follows. □

D Assorted shorter proofs

Proof of Proposition 3.2:

(a) The argument is actually slightly more complex than that given in the main text, because the alphabetical tie-breaking leads to different cases depending on the parity of N.

If N is even, let the manipulator's preferences be $ACB\ldots$, and let the opponent-profile P be distributed according to $\phi = \left(\frac{1}{2}\ B, \frac{1}{2}\ C\right)$ (only voters' top choices matter). Then the manipulator cannot change the outcome unless $P = \left(\frac{N}{2}\ B, \frac{N}{2}\ C\right)$, in which case strategically voting for C instead of A beneficially changes the outcome from B to C. If N is odd, let the preferences be $ABC\ldots$, and let $\phi = \left(\frac{N-1}{2N}\ B, \frac{N+1}{2N}\ C\right)$. Then the manipulator is pivotal precisely when the opponent-profile is $P = \left(\frac{N-1}{2}\ B, \frac{N+1}{2}\ C\right)$, in which case voting for B changes the outcome from C to B. In both cases, the probability of being pivotal (2.3) is σ_N^*.

(b) First we prove the lower bound. Consider any small $\epsilon > 0$. Let the manipulator's preference be $ABC\ldots$, and consider a distribution $\phi \in \Delta(\mathcal{C})$ of the other voters' first-place votes such that B and C are each chosen with probability $\frac{1}{M} + \epsilon$, and every other candidate is chosen with probability $\frac{1}{M} - \frac{2\epsilon}{M-2}$.

Consider susceptibility as formulated in (2.3), where the proposed manipulation $\succ'$ is one that ranks B first, and the set $\mathcal{C}^+$ of desirable candidates is $\{A, B\}$. Write the relevant expectation as

$$\sigma \geq \sum_P \left[\mathbf{I}(f(\succ', P) \in \mathcal{C}^+) - \mathbf{I}(f(\succ, P) \in \mathcal{C}^+)\right]\mathbf{P}(P \mid N; \phi). \qquad \text{(D.1)}$$

(We write $\geq$ rather than $=$, since we are considering a specific distribution ϕ rather than the max.) Say that an opponent-profile P is *relevant* if B and C both receive a vote share between $1/M + \epsilon/2$ and $1/M + 3\epsilon/2$, and every other candidate receives less than $1/M$ of the vote. By Lemma 2.2(a), the probability that the realized profile is relevant is $\overset{e}{\sim} 1$, so we need only consider the contribution of the relevant profiles to (D.1). For any such profile (assuming N is large enough), no matter what the manipulator does, the outcome will be either B or C. The relevant profiles that contribute to (D.1) are exactly the ones where the manipulator is pivotal in changing the outcome from C to B — that is, the ones for which B receives exactly one less vote than C. It follows from Lemma 2.2(b) that the total probability of these profiles is $\sim (1/2)\sqrt{1/\pi\left(\frac{1}{M} + \epsilon\right)N}$. (Here the lemma applies with B, C corresponding to the indices i, j, and $y = -1$. Note that the definition of a relevant profile is a set of linear inequalities on the vote shares.)

Thus we have

$$\sigma \gtrsim \frac{1}{2}\sqrt{\frac{1}{\pi\left(\frac{1}{M} + \epsilon\right)N}}.$$

Taking $\epsilon \to 0$ gives the lower bound in Proposition 3.2(b).

Now we prove the upper bound. For each value of N, consider the true preference, manipulation, and belief ϕ that attain the maximum in (2.3). (These may vary depending on N, but we will not bother to make this dependence explicit in the notation.) Suppose that, for a given N, the manipulator's true first choice is A_i and the reported first choice is A_j. This manipulation can be beneficial only if it changes the outcome from A_k, for some $k \neq i, j$, to A_j. For each k, let S_{kj} be the set of all N-profiles P such that $f(A_i, P) = A_k$ and $f(A_j, P) = A_j$; and let $S_{\to j} = \cup_{k \neq i,j} S_{kj}$.

We wish to show that $\mathbf{P}(S_{\to j} \mid N; \phi) \lesssim \sqrt{M/\pi N}$.

Now, consider again any fixed $\epsilon > 0$. For each $k \neq i, j$, we have

$$\max_{\phi:\ \phi_j \geq (1+\epsilon)\phi_k} \mathbf{P}(S_{kj} \mid N; \phi) \overset{e}{\sim} 0. \tag{D.2}$$

Indeed, each opponent-profile $P = (x_1, \ldots, x_M) \in S_{kj}$ has $x_j + 1, x_k \geq N/M$, and also $x_j = x_k$ or $x_j = x_k - 1$. Consider such a profile P. Let $p(x_{-jk})$ be the conditional probability of realizing P, given that the components x_{-jk} are realized. By Lemmas C.2 and C.4,

$$p(x_{-jk}) = \mathbf{P}\left(\begin{array}{c} x_j \\ x_k \end{array} \middle| \; x_j + x_k; \; \begin{array}{c} \phi_j/(\phi_j + \phi_k) \\ \phi_k/(\phi_j + \phi_k) \end{array} \right) \leq e^{-(x_j + x_k)\cdot\left(\frac{x_j}{x_j + x_k} - \frac{\phi_j}{\phi_j + \phi_k}\right)^2/2}.$$

The squared expression in the exponent is bounded away from zero, while the $x_j + x_k$ factor is $\geq N/M$, so the upper bound goes to zero exponentially in N. So, given any value of x_{-jk}, the conditional probability of realizing values of x_j and x_k for which the resulting profile is in S_{kj} is bounded above by an expression that decays exponentially in N. Hence the *unconditional* probability of S_{kj} satisfies this same exponential bound, and (D.2) holds.

On the other hand, the worst-case belief ϕ cannot have $\mathbf{P}(S_{\to j} \mid N; \phi) \overset{e}{\sim} 0$, since we already proved this probability satisfies a lower bound on the order of $\sqrt{1/N}$. Thus, as long as N is large enough, there must be some k^* such that $\phi_j < (1+\epsilon)\phi_{k^*}$. (This k^* may not be unique, and may vary depending on N.)

Next, we claim that for any value of x_{-jk^*} there is at most one way of choosing x_j, x_{k^*} (given the additional constraint $\sum_l x_l = N$) so that the resulting N-profile lies in $S_{\to j}$. Indeed, suppose for a contradiction that $(x_j, x_{k^*}, x_{-jk^*}) \in S_{\to j}$, and also $(x_j + s, x_{k^*} - s, x_{-jk^*}) \in S_{\to j}$ for some positive integer s. Then, in particular,

$$f(x_i, x_j + 1, x_{k^*}, x_{-ijk^*}) \;=\; A_j; \tag{D.3}$$

$$f(x_i + 1, x_j + s, x_{k^*} - s, x_{-ijk^*}) \;=\; A_l \neq A_i, A_j. \tag{D.4}$$

If $s \geq 2$, then the profile in (D.4) gives a (weakly) greater advantage for j relative to l than the profile in (D.3) does, so if plurality rule chooses A_j in (D.3) it should choose A_j in (D.4) also, a contradiction. And if $s = 1$, then the profile in (D.4) differs from that in (D.3) by a vote shift from A_{k^*} to A_i, which cannot change the winner from A_j to A_l — a contradiction again. Thus the claim holds.

Consider any x_{-jk^*} such that there exist x_j, x_{k^*} for which the resulting profile lies in $S_{\to j}$. We will again bound the probability of realizing this profile, conditional on x_{-jk*}. For this pivotal profile, we must have $x_{j+k^*} \geq x_j \geq (N+1)/M - 1 \geq N(1/M - \epsilon)$ (as long as N is large). The conditional probability of realizing (x_j, x_{k^*}) given x_{-jk^*} is

$$p(x_{-jk^*}) = \mathbf{P}\left(\begin{array}{c} x_j \\ x_{k^*} \end{array} \middle| \; x_{j+k^*}; \; \begin{array}{c} \phi_j/(\phi_j + \phi_{k^*}) \\ \phi_{k^*}/(\phi_j + \phi_{k^*}) \end{array} \right).$$

(i) If $x_j > (1+2\epsilon)x_{k^*}$ then this probability $p(x_{-jk^*})$ is bounded above by an expression that decays exponentially in x_{j+k^*} (by Lemma 2.2(a) and $\phi_j < (1+\epsilon)\phi_{k^*}$). In particular, across all choices of x_{-jk^*} such that the corresponding profile in $S_{\to j}$ satisfies $x_j > (1+2\epsilon)x_{k^*}$, the probability $p(x_{-jk^*})$ is bounded above uniformly by a quantity that decays exponentially in N.

(ii) If $x_j \leq (1+2\epsilon)x_{k^*}$, then (since we also have $x_j + 1 \geq x_{k^*}$) we get $x_{j+k^*} \geq N(2/M - 3\epsilon)$. Hence

$$p(x_{-jk^*}) \leq \max_{\substack{x+y \geq N(2/M-3\epsilon) \\ x \leq (1+2\epsilon)y \\ y \leq (1+2\epsilon)x}} \mathbf{P}(x, y \mid x+y; \alpha_x, \alpha_y).$$

For given $x+y$, the choices of x, y, α_x, α_y that attain the max are given by Lemmas 2.3 and 2.4, and we obtain

$$p(x_{-jk^*}) \leq \max_{K \geq N(2/M-3\epsilon)} \mathbf{P}\left(x, y \;\middle|\; K; \frac{x}{K}, \frac{y}{K}\right) \text{ with } x = \left\lceil \frac{K}{2+2\epsilon} \right\rceil, y = K - x.$$

Denote the expression inside this maximum by $\widetilde{p}(K)$.

We have thus shown that the conditional probability of realizing (x_j, x_{k^*}) forming a profile in $S_{\to j}$, given x_{-jk^*}, satisfies

$$p(x_{-jk^*}) \leq \max\{ce^{-\lambda N}, \max_{K \geq N(2/M-3\epsilon)} \widetilde{p}(K)\}.$$

(Here c, λ are some positive values.) This inequality applies to the conditional probability of obtaining a profile $x \in S_{\to j}$, given x_{-jk^*}. So it also applies to the unconditional probability of drawing a profile in $S_{\to j}$:

$$\mathbf{P}(S_{\to j} \mid N; \phi) \leq \max\{ce^{-\lambda N}, \max_{K \geq N(2/M - 3\epsilon)} \widetilde{p}(K)\}.$$

Now, Lemma 2.1 gives

$$\widetilde{p}(K) \sim \sqrt{\frac{1}{2\pi K \left(\frac{1}{2+2\epsilon}\right)\left(\frac{1+2\epsilon}{2+2\epsilon}\right)}}.$$

Hence

$$\mathbf{P}(S_{\to j} \mid N; \phi) \lesssim \max\left\{ce^{-\lambda N}, \sqrt{\frac{1}{2\pi N\left(\frac{2}{M} - 3\epsilon\right)\left(\frac{1}{2+2\epsilon}\right)\left(\frac{1+2\epsilon}{2+2\epsilon}\right)}}\right\}.$$

Clearly, for N large enough the square-root term dominates.

Finally, taking $\epsilon \to 0$ gives us the simpler asymptotic upper bound $\sqrt{M/\pi N}$, which is what we wanted to show.

□

Proof of Proposition 3.3:

Given Proposition 3.2(a), we need only show $\sigma_N^{plur} \leq \sigma_N^*$. Consider any true preference for the manipulator and proposed manipulation. For this proof only, label the candidates so that the manipulator's preference is ABC, not necessarily corresponding to the tie-breaking order. A manipulation from A to C can never be beneficial; manipulation to B can be beneficial only when it changes the winner from C to B. So we need to show that the probability of being pivotal from C to B is at most σ_N^*. Let

$$S_0 = \{(x_A, x_B, x_C) \mid x_B = x_C - 1 \geq x_A\},$$

$$S_1 = \{(x_A, x_B, x_C) \mid x_B = x_C \geq x_A + 1\}.$$

The relevant set of pivotal profiles is contained either in S_0 or S_1 (depending on which of B, C wins a tiebreaker), so we just need to show that for any ϕ, both S_0 and S_1 are events of total probability at most σ_N^*.

Consider the ϕ that maximizes $\mathbf{P}(S_0 \mid N; \phi)$. Write $\phi = (\phi_A, \phi_B, \phi_C)$. We then have

$\phi_C \geq \phi_A$. Proof: Suppose not. Then

$$\begin{aligned}
&\frac{d}{d\epsilon}\left[\mathbf{P}(S_0 \mid N; \phi_A - \epsilon, \phi_B, \phi_C + \epsilon)\right] \\
&\qquad = \frac{d}{d\epsilon}\left[\sum_{(x_A,x_B,x_C)\in S_0} \frac{N!}{x_A!x_B!x_C!}(\phi_A-\epsilon)^{x_A}\phi_B^{x_B}(\phi_C+\epsilon)^{x_C}\right] \\
&\qquad = \sum_{(x_A,x_B,x_C)\in S_0} \frac{N!}{x_A!x_B!x_C!}(\phi_A-\epsilon)^{x_A}\phi_B^{x_B}(\phi_C+\epsilon)^{x_C}\cdot\left(\frac{x_C}{\phi_C+\epsilon}-\frac{x_A}{\phi_A-\epsilon}\right).
\end{aligned}$$

For ϵ close to 0, the last factor in parentheses is always positive (since $x_C \geq x_A$ throughout S_0). So changing the belief from (ϕ_A, ϕ_B, ϕ_C) to $(\phi_A - \epsilon, \phi_B, \phi_C + \epsilon)$ increases the probability of drawing a profile in S_0, contrary to the assumption that the belief was chosen to maximize this probability.

Exactly the same reasoning applies for S_1. Thus it suffices to show that each of S_0, S_1 has probability at most σ_N^*, assuming that the belief $\phi = (\phi_A, \phi_B, \phi_C)$ satisfies $\phi_A \leq \phi_C$. In particular, we may assume $\phi_A \leq 1/2$.

We need to show four things:

(i) when N is odd, the probability of drawing a profile in S_0 is at most σ_N^*;

(ii) when N is odd, the probability of S_1 is at most σ_N^*;

(iii) when N is even, the probability of S_0 is at most σ_N^*;

(iv) when N is even, the probability of S_1 is at most σ_N^*.

First consider (i), so N is odd. Then, for $(x_A, x_B, x_C) \in S_0$, we have x_A even and at most $x_A^{max} = 2\lfloor N/6 \rfloor$, so

$$\mathbf{P}(S_0 \mid N; \phi) = \sum_{\substack{x_A \text{ even} \\ 0\leq x_A \leq x_A^{max}}} \mathbf{P}\left(\begin{array}{c} x_A \\ N-x_A \end{array}\middle|\, N; \begin{array}{c} \phi_A \\ 1-\phi_A \end{array}\right)\mathbf{P}\left(\begin{array}{c} x_B \\ x_C \end{array}\middle|\, N-x_A; \begin{array}{c} \phi'_B \\ \phi'_C \end{array}\right)$$

by Lemma C.2 (where $\phi'_B = \frac{\phi_B}{\phi_B+\phi_C}$, $\phi'_C = \frac{\phi_C}{\phi_B+\phi_C}$). Since the relevant x_B, x_C are equal or differ by 1, Lemma 2.3 gives $\mathbf{P}(x_B, x_C \mid N - x_A; \phi'_B, \phi'_C) \leq \sigma^*_{N-x_A}$, which in turn is at most $\sigma^*_{N-x_A^{max}}$ by Corollary C.6. Hence, the above sum is at most

$$\begin{aligned}
&\mathbf{P}\left(\begin{array}{c}0\\N\end{array}\,\middle|\, N;\ \begin{array}{c}\phi_A\\1-\phi_A\end{array}\right)\sigma^*_N + \\
&\qquad\sum_{\substack{x_A \text{ even}\\ 2\le x_A\le x_A^{max}}} \mathbf{P}\left(\begin{array}{c}x_A\\N-x_A\end{array}\,\middle|\, N;\ \begin{array}{c}\phi_A\\1-\phi_A\end{array}\right)\sigma^*_{N-x_A^{max}} \\
\le\ &\mathbf{P}\left(\begin{array}{c}0\\N\end{array}\,\middle|\, N;\ \begin{array}{c}\phi_A\\1-\phi_A\end{array}\right)(\sigma^*_N - \sigma^*_{N-x_A^{max}}) + \\
&\qquad\left[\sum_{x_A \text{ even}} \mathbf{P}\left(\begin{array}{c}x_A\\N-x_A\end{array}\,\middle|\, N;\ \begin{array}{c}\phi_A\\1-\phi_A\end{array}\right)\right]\sigma^*_{N-x_A^{max}}.
\end{aligned}$$

In this last line, the first probability is $(1-\phi_A)^N$, and the bracketed sum is the probability that a binomial distribution with parameters $N;\phi_A$ produces an even number of successes, which is $(1+(1-2\phi_A)^N)/2$ (Lemma C.3). Thus, the probability of drawing a profile in S_0 is at most

$$h(\phi_A) = (1-\phi_A)^N(\sigma^*_N - \sigma^*_{N-x_A^{max}}) + \frac{1+(1-2\phi_A)^N}{2}\sigma^*_{N-x_A^{max}}.$$

Let us find the maximum of h on $[0,1/2]$ (since by assumption ϕ_A lies in this interval). Differentiating gives

$$\frac{dh}{d\phi_A} = -N\left[(1-\phi_A)^{N-1}(\sigma^*_N - \sigma^*_{N-x_A^{max}}) + (1-2\phi_A)^{N-1}\sigma^*_{N-x_A^{max}}\right].$$

This is negative if

$$\left(\frac{1-2\phi_A}{1-\phi_A}\right)^{N-1} > \frac{\sigma^*_{N-x_A^{max}} - \sigma^*_N}{\sigma^*_{N-x_A^{max}}},$$

which holds precisely when ϕ_A is sufficiently small. Therefore h is initially decreasing and then increasing, so the maximum occurs at one of the endpoints of the interval,

$$h(0) = \sigma^*_N \qquad \text{or} \qquad h\left(\frac{1}{2}\right) = \frac{1}{2^N}\sigma^*_N + \left(\frac{1}{2} - \frac{1}{2^N}\right)\sigma^*_{N-x_A^{max}}.$$

The first of these is larger as long as $\sigma^*_N \ge \sigma^*_{N-x_A^{max}}/2$. Using the fact that $N - x_A^{max} \ge 2N/3$ and the bounds in Lemma C.5, we can verify that this always holds. Thus, we have shown that the probability of drawing a profile in S_0 is

$$\mathbf{P}(S_0 \mid N; \phi) \leq h(\phi_A) \leq h(0) = \sigma_N^*.$$

That takes care of (i).

Next we turn to (ii), where we consider the probability of drawing a profile in S_1. In this case, each such profile has x_A odd and at most $x_A^{max} = 2\lfloor N/6 \rfloor + 1$. Hence, by similar calculations, the relevant probability is

$$\sum_{\substack{x_A \text{ odd} \\ 1 \leq x_A \leq x_A^{max}}} \mathbf{P}\left(\begin{array}{c} x_A \\ N - x_A \end{array} \middle| N; \begin{array}{c} \phi_A \\ 1 - \phi_A \end{array} \right) \mathbf{P}\left(\begin{array}{c} x_B \\ x_C \end{array} \middle| N - x_A; \begin{array}{c} \phi_B' \\ \phi_C' \end{array} \right)$$

$$\leq \left[\sum_{x_A \text{ odd}} \mathbf{P}\left(\begin{array}{c} x_A \\ N - x_A \end{array} \middle| N; \begin{array}{c} \phi_A \\ 1 - \phi_A \end{array} \right) \right] \sigma^*_{N - x_A^{max}}.$$

The bracketed expression is the probability that a binomial distribution with parameters $N; \phi_A$ produces an odd number of successes, which is $(1-(1-2\phi_A)^N)/2 \leq 1/2$. (Remember that $\phi_A \leq 1/2$.) Therefore the probability of drawing a profile in S_1 is at most $\sigma^*_{N-x_A^{max}}/2$. This is less than σ_N^*, again by straightforward use of the bounds from Lemma C.5.

In case (iii), the relevant set of profiles again has x_A odd and at most $x_A^{max} = 2\lfloor N/6 \rfloor + 1$, so the reasoning used for (ii) applies again word for word.

Finally, in (iv), the relevant set of profiles has x_A even and at most $x_A^{max} = 2\lfloor N/6 \rfloor$. In this case the reasoning used for (i) applies again.

This covers all four cases (i)-(iv), so the probability that the manipulator is pivotal is never more than σ_N^*.

□

Proof of Proposition 3.5: Again, the tie-breaking assumption leads us to split into cases depending on parity. First suppose M is even. Let the manipulator's preferences be $A_1 A_2 \ldots A_M$. Suppose the belief ϕ is

$$\begin{array}{l} \frac{1}{2}\ A_1 A_2 A_3 \ldots A_M \\ \frac{1}{2}\ A_2 A_1 A_3 \ldots A_M \end{array}.$$

That is, all the other voters prefer A_1 and A_2, then the remaining candidates in numerical order, but are evenly split between ranking A_1 first or A_2 first. The manipulator considers manipulating by moving A_2 to the bottom, thus reporting $A_1 A_3 \ldots A_M A_2$.

Regardless of whether the manipulator tells the truth or lies, A_1 will have a higher score than $A_3, \ldots, A_M$, so the winner must be A_1 or A_2. Suppose x of the other voters rank A_1 first, and the remaining $N - x$ rank A_2 first. The difference in scores between A_1 and A_2 is $(x+1) - (N-x)$ if the manipulator tells the truth and $(x+M-1)-(N-x)$ if he lies. Therefore, manipulation improves the outcome from A_2 to A_1 if

$$2x - N + 1 < 0 \leq 2x - N + M - 1$$

or equivalently

$$\frac{N-(M-1)}{2} \leq x < \frac{N-1}{2}.$$

Otherwise, manipulation has no effect on the outcome.

Given that x has to be an integer, the possible values of x in this range are $\lfloor N/2 - K \rfloor$ for $K = 1, 2, \ldots, (M-2)/2$. For each such K, Lemma 2.1 tells us that the probability that $x = \lfloor N/2 - K \rfloor$ is $\sim \sqrt{2/\pi N}$. Therefore, the total probability of being pivotal is $\sim \frac{M-2}{2}\sqrt{2/\pi N}$, and the result follows via (2.3).

Now suppose M is odd. The argument is essentially the same, except that we have to consider different cases depending on the parity of N. If N is even, then we consider exactly the same preferences, the same manipulation, and the same belief as before. Again, the manipulator is pivotal if $(N-(M-1))/2 \leq x < (N-1)/2$. The integer values of x in this range are $N/2 - K$ for $K = 1, 2, \ldots, (M-1)/2$.

If N is odd, then we reverse the roles of A_1 and A_2 throughout. Thus, the manipulator's belief is the same as before, but his true preference is $A_2 A_1 A_3 \ldots A_M$, and the proposed manipulation is $A_2 A_3 \ldots A_M A_1$. Let x now denote the number of other voters who rank A_2 first. Then the score of A_2 minus the score of A_1 is $(x+1)-(N-x)$ if the manipulator tells the truth and $(x+M-1)-(N-x)$ if he lies; in view of alphabetical tie-breaking, the manipulator is pivotal if

$$2x - N + 1 \leq 0 < 2x - N + M - 1.$$

The integer values of x satisfying these inequalities are $x = (N+1)/2 - K$ for $K = 1, 2, \ldots, (M-1)/2$.

So for both N even and N odd, the manipulator is pivotal when $x = \lceil N/2 \rceil - K$ for some $K = 1, 2, \ldots, (M-1)/2$. The total probability of this event is $\sim \frac{M-1}{2}\sqrt{2/\pi N}$.

□

We next prove Lemma 4.8, the ancillary result en route to the local average lemma.

We use the notation $\overline{f}(\phi)$, $\overline{f}_{A_i}(\phi)$ developed in Subsection 4.3 of the main paper.

Proof of Lemma 4.8: Put $g(x) = \overline{f}_{A_i}(\phi^x)$.

The proof is based on the following observation. Consider the definition (2.4) of $\mathbf{P}$, and take the partial derivative with respect to a parameter α_i (ignoring the fact that our interpretation of (2.4) required $\alpha_1 + \cdots + \alpha_r = 1$). We obtain

$$\frac{\partial}{\partial \alpha_i}\mathbf{P}\left(\begin{array}{c|c} x_1 & \alpha_1 \\ \vdots & \vdots \\ x_i & K;\ \alpha_i \\ \vdots & \vdots \\ x_r & \alpha_r \end{array}\right) = K \cdot \mathbf{P}\left(\begin{array}{c|c} x_1 & \alpha_1 \\ \vdots & \vdots \\ x_i - 1 & K-1;\ \alpha_i \\ \vdots & \vdots \\ x_r & \alpha_r \end{array}\right). \tag{D.5}$$

On the right-hand side, x_i has been replaced by $x_i - 1$ and all other x_j are unchanged.

Now consider the function of x,

$$\overline{f}_{A_i}(\phi^x) = \sum_{f(P)=A_i} \mathbf{P}\left(\begin{array}{c|cc} & \alpha(1-x) & \succ \\ P & N+1;\ \alpha x & \succ' \\ & 1-\alpha & \phi \end{array}\right).$$

The sum is over all $(N+1)$-profiles P such that $f(P) = A_i$. Differentiating this sum term-by-term with respect to x, and applying (D.5), we obtain

$$\frac{d}{dx}\left(\overline{f}_{A_i}(\phi^x)\right) = \sum_{f(P)=A_i} \left(\alpha\widetilde{N} \cdot \mathbf{P}(P - \succ' \mid N; \phi^x) - \alpha\widetilde{N} \cdot \mathbf{P}(P - \succ \mid N; \phi^x)\right). \tag{D.6}$$

The interpretation of the $\mathbf{P}(P - \succ' \mid \cdots)$ term is that if P contains at least one $\succ'$ vote, then $P - \succ'$ is the N-profile consisting of P with a $\succ'$ removed, and otherwise we simply interpret the whole term to be zero; similarly for the $\mathbf{P}(P - \succ \mid \cdots)$ term.

Now (D.6) can be rewritten

$$\frac{d}{dx}\left(\overline{f}_{A_i}(\phi^x)\right) = \alpha\widetilde{N}\left[\sum_{f(\succ',P)=A_i} \mathbf{P}(P \mid N; \phi^x) - \sum_{f(\succ,P)=A_i} \mathbf{P}(P \mid N; \phi^x)\right].$$

Here the first sum is over N-profiles P with $f(\succ', P) = A_i$, and the second is over P with $f(\succ, P) = A_i$. This in turn is equivalent to the difference given in the lemma statement. □

We also include here the proof of the result in Appendix A. It is basically a routine unwinding of definitions.

Proof of Proposition A.1: It suffices to show that for any symmetric equilibrium strategies of the voters, the following holds:

(a) if the planner chooses a voting rule f with $\overline{\sigma}(f) \leq \underline{\epsilon}$ then her utility is given by $\min_{P \in \mathcal{L}^{N+1}} V(f(P), P)$;

(b) if she chooses f with $\overline{\sigma}(f) > \underline{\epsilon}$, then her utility is $\underline{V}$.

Statement (a) holds because the voters will never manipulate. Specifically, suppose the state is $\omega \in \Omega^*$. Then, $\sigma_\omega(f) < \overline{\sigma}(f) \leq \underline{\epsilon}$. Consider a voter with utility function u, manipulation cost ϵ, and belief ψ about the types of the other voters. Composing the strategy τ of the other voters with ψ gives a probability distribution $\phi \in \Delta(\mathcal{L})$, so that other voters' actual reports are expected to be independent draws from ϕ. Consider any manipulation $\succ' \in \mathcal{L}$. From the definition of $\sigma_\omega(f)$ we have

$$u(\omega(f, \succ', \phi)) - u(\omega(f, \succ^* (u), \phi)) \leq \sigma_\omega(f) < \underline{\epsilon} \leq \epsilon.$$

Equivalently,

$$u(\omega(f, \succ', \phi)) - \epsilon < u(\omega(f, \succ^* (u), \phi)).$$

So the voter will choose to simply report the true preference $\succ^* (u)$. Thus, in all possible states $\omega \in \Omega^*$, each equilibrium strategy τ of the voters will specify that they always tell the truth. Then, whenever the voters' true preferences realize the (ordinal) profile P, the planner's utility is $V(f(P), P)$, regardless of the state. From the maxmin specification of the planner's utility, claim (a) follows.

For (b), consider any f with $\overline{\sigma}(f) > \underline{\epsilon}$. We know that there exist some preferences $\succ_1, \ldots, \succ_{N+1}$ and reports $\widehat{\succ}_1, \ldots, \widehat{\succ}_{N+1}$ such that

$$V(f(\widehat{\succ}_1, \ldots, \widehat{\succ}_{N+1}); \succ_1, \ldots, \succ_{N+1}) = \underline{V}.$$

(This follows from the definition of $\underline{V}$ as the minimum value of V, and the fact that f is surjective.) So our strategy will be to construct some state $\omega \in \Omega^*$, and some types $t_i \in \mathcal{T}$ for the voters, such that each voter i has true preference $\succ_i$ but reports $\widehat{\succ}_i$ in any equilibrium.

First we construct the state ω, as follows. Fix a number $\widetilde{\sigma}$ with $\underline{\epsilon} < \widetilde{\sigma} < \min\{1, \overline{\sigma}(f)\}$. We first define $\xi : \mathcal{L} \times \Delta(\mathcal{L}) \to \Delta(\mathcal{C})$ to be any continuous function such that for all preferences $\succ, \succ', \succ'' \in \mathcal{L}$,

$$\xi\left(\succ, \frac{2}{3}\succ' + \frac{1}{3}\succ''\right) = \begin{cases} \text{the candidate ranked first by } \succ'\text{, with certainty} & \text{if } \succ = \succ''; \\ \text{the candidate ranked last by } \succ'\text{, with certainty} & \text{otherwise.} \end{cases}$$

This can be done, since we have only specified the values of ξ at finitely many points.

Now, for the given voting rule f, we define $\omega(f, \succ, \phi) \in \Delta(\mathcal{C})$ for all preferences $\succ \in \mathcal{L}$ and all beliefs $\phi \in \Delta(\mathcal{L})$, by

$$\omega(f, \succ, \phi) = \widetilde{\sigma}\xi(\succ, \phi) + (1 - \widetilde{\sigma})\ A_1.$$

That is, if the voting rule is f, then ω chooses the output of ξ with probability $\widetilde{\sigma}$, and otherwise just chooses the fixed candidate A_1 as winner.

For every other voting rule $f' \neq f$, any $\succ \in \mathcal{L}$ and any $\phi \in \Delta(\mathcal{L})$, put

$$\omega(f', \succ, \phi) = \omega_0(f', \succ, \phi).$$

This completes the definition of ω. It is straightforward to check that ω is indeed a continuous function: we need $\omega(f, \succ, \phi)$ to be continuous in ϕ, but this follows from continuity of ξ; and for each $f' \neq f$ we need $\omega(f', \succ, \phi)$ to be continuous in ϕ, but this follows from continuity for ω_0.

We check that $\omega \in \Omega^*$. Notice that under voting rule f in state ω, each voter cannot affect more than $\widetilde{\sigma}$ probability mass of the outcome by changing his vote. It immediately follows that

$$\sigma_\omega(f) \leq \widetilde{\sigma} < \overline{\sigma}(f).$$

And for any other voting rule f', we have

$$\sigma_\omega(f') = \sigma_{\omega_0}(f') < \overline{\sigma}(f')$$

by the assumption $\omega_0 \in \Omega^*$. Thus, the susceptibility bounds are satisfied, and $\omega \in \Omega^*$.

Next, for each voter i, we construct a type t_i as follows:

- the utility function u_i represents the preference $\succ_i$, and values the most-preferred candidate at 1 and the least-preferred candidate at 0;
- the manipulation cost is $\underline{\epsilon}$;
- the first-order belief about others' preferences is that every other voter

– with probability 2/3, has a utility function that represents $\succ_i$ and has range smaller than $\underline{\epsilon}$; and

– with remaining probability 1/3, has a utility function that represents $\widehat{\succ}_i$ and has range smaller than $\underline{\epsilon}$.

(The first-order belief about others' manipulation costs may be arbitrary.)

By the richness assumption, there exists a type $t_i \in \mathcal{T}$ having this basic type and first-order belief.

Now we consider t_i's equilibrium behavior in state ω. First, in any equilibrium, any voter whose utility function has range smaller than $\underline{\epsilon}$ always votes truthfully (since his material gain from lying is less than $\underline{\epsilon}$). Therefore, voter i's induced belief ϕ about others' behavior is that each other voter will report $\succ_i$ with probability 2/3 and report $\widehat{\succ}_i$ with probability 1/3. Then:

- $\omega(f, \widehat{\succ}_i, \phi)$ is the distribution that chooses the candidate ranked first by $\succ_i$ with probability $\widetilde{\sigma}$, and chooses A_1 with remaining probability $1 - \widetilde{\sigma}$. Thereforefore, if voter i reports $\widehat{\succ}_i$, his expected material utility is $\widetilde{\sigma} + (1 - \widetilde{\sigma})u_i(A_1)$.

- For any $\succ' \neq \widehat{\succ}_i$, $\omega(f, \succ', \phi)$ chooses the candidate ranked last by $\succ_i$ with probability $\widetilde{\sigma}$, and A_1 with remaining probability $1 - \widetilde{\sigma}$. Therefore, i's expected material utility from reporting any such $\succ'$ is $(1 - \widetilde{\sigma})u_i(A_1)$.

Since $\widetilde{\sigma} > \underline{\epsilon}$, voter i's unique best reply is to report $\widehat{\succ}_i$.

Thus, in state $\omega \in \Omega^*$, the types $t_1, \ldots, t_{N+1}$ of the voters have true preferences $\succ_1, \ldots, \succ_{N+1}$ but necessarily report $\widehat{\succ}_1, \ldots, \widehat{\succ}_{N+1}$. This leaves the planner with utility

$$V(f(\widehat{\succ}_1, \ldots, \widehat{\succ}_{N+1}); \succ_1, \ldots, \succ_{N+1}) = \underline{V},$$

her worst possible. Statement (b) follows. □

E Proofs for comparison of voting systems

Here we prove Proposition 3.6, giving lower bounds on the susceptibility of five voting systems from [2].

Proof of Proposition 3.6: We give the proofs for the voting systems one by one in order.

Black's system. This is just an embellishment of the construction given for the Borda count, performed so as to ensure the nonexistence of a Condorcet winner (with probability close to 1). We first present the construction for $M = 5$. For readability we refer to the candidates using letters A, B, C, D, E. Take small $\epsilon > 0$. Consider the following belief of the manipulator: the other voters report

$$\left.\begin{array}{ccc} CDABE & DEABC & ECABD \\ CDBAE & DEBAC & ECBAD \end{array}\right\} \quad \text{each with probability} \qquad 1/12 + \epsilon;$$

$$\left.\begin{array}{ccc} CABDE & ABDEC & ABECD \\ CBADE & BADEC & BAECD \end{array}\right\} \quad \text{each with probability} \qquad 1/12 - \epsilon.$$

Each other voter then:

- prefers C over D with probability $2/3$;
- prefers D over E with probability $2/3$;
- prefers E over C with probability $2/3$;
- prefers C over A and B with probability $1/2 + 2\epsilon$.

By Lemma 2.2(a), with probability converging exponentially to 1, each of these pairwise preferences will be held by a share at least $1/2 + \epsilon$ of opposing voters, so no matter what the manipulator does, we will end up with $C \to A, B, D$; $D \to E$; and $E \to C$. In particular, no candidate can then be a Condorcet winner.

Also, each other voter awards, on average,

- $40/12 - 10\epsilon$ points each to A and B;
- $36/12 + 4\epsilon$ points to C;
- $32/12 + 8\epsilon$ points each to D and E.

Using Lemma 2.2(a) again, we see that with probability converging exponentially to 1, candidates A and B will end up with higher scores than C, D or E, no matter what the manipulator does.

So, neglecting events of exponentially small probability, we can assume that there is no Condorcet winner and, however the manipulator votes, only A or B can possibly win. Since each of the other voters' contribution to the *difference* between A's and B's Borda scores is either $+1$ or -1 with probability $1/2$ each, the same analysis that was used to prove Proposition 3.5 applies here. Explicitly, if N is even, we let the manipulator's true preference be $ABCDE$ and consider the manipulation $ACDEB$; if N is odd, we let the true preference be $BACDE$ and the manipulation be $BCDEA$. The manipulation improves the outcome from the manipulator's second-ranked to his first-ranked candidate with probability $\sim 2\sqrt{2/\pi N}$, and has no effect otherwise.

This covers the case $M = 5$. For $M > 5$, construct a belief by supposing each other voter ranks the first five candidates $A_1, \ldots, A_5$ $(= A, \ldots, E)$ at the top according to the distribution above, and then has all remaining candidates in numerical order after them. Then none of the extra candidates can ever be a Condorcet winner, nor a Borda winner, since they receive lower scores than (say) A_1. So again, with probability converging exponentially to 1, the winner will be either A_1 or A_2 no matter what the manipulator does. Let the manipulator's preferences and proposed manipulation be as for Proposition 3.5; then manipulation succeeds with probability $\sim \lceil ((M-2)/2) \rceil \sqrt{2/\pi N}$ by the same argument as before.

Copeland's system. We will give a construction supposing that $M = 3K - 1$, where $K \geq 3$. If $M \geq 9$ is instead of the form $3K$ or $3K+1$, then we can modify the construction by the usual method of appending the extra one or two candidates at the end of everyone's preferences, and the same argument will apply. At the end of the proof we will also show how to modify the construction for the remaining cases $M = 3, 4, 6, 7$.

It will be convenient to depart from our usual notation for candidates and instead let the candidates be called $A, B, C_1, \ldots, C_K, D_1, \ldots, D_{2K-3}$, where ties are broken in that order. We will also let the D-candidates be numbered cyclically, so that $D_{i+(2K-3)} = D_i$.

Let the manipulator's true preference be $C_1 \ldots C_K D_1 \ldots D_{2K-3} AB$. To describe the belief ϕ, we will not list out all the preferences that other voters may have, as there are too many to list individually. Instead, we describe a randomized procedure to construct a preference ranking, and let ϕ be the resulting distribution over preferences. In this description, we will refer to choosing a *random cyclic permutation* of the D_i, which means an ordering of the form $D_j D_{j+1} \ldots D_{j+2K-4}$, where each possible value of $j \in \{1, 2, \ldots, 2K-3\}$ is chosen with probability $1/(2K-3)$.

- With probability $1/3$, do the following: Begin with BA, then, for each $i = 1, \ldots, K$ in succession, append C_i either at the beginning or at the end, independently each with probability $1/2$. Finally, attach a random cyclic permutation of the D_i at the beginning of the preference order.

- With complementary probability $2/3$, do the following: Begin with BA, immediately followed by a random cyclic permutation of the D_i; then successively append each C_i either at the beginning or at the end, each with probability 1/2.

Whenever one candidate is preferred to another candidate with probability strictly greater than 1/2 under this distribution, the usual application of Lemma 2.2(a) ensures that the former candidate majority-defeats the latter (regardless of what the manipulator does) with probability $\stackrel{e}{\sim} 1$. Thus, we can see that with probability $\stackrel{e}{\sim} 1$, all of the following majority-defeat relations hold:

- $B \to A$;
- $D_i \to D_{i+1}, D_{i+2}, \ldots, D_{i+K-2}$, for each i;
- $B, A \to D_i$ for each i;
- $D_i \to C_j$, for all i and j.

We henceforth assume that these relations hold. Moreover, for each C_j, each of the other voters either prefers both A and B over C_j or prefers C_j over both A and B; each case occurs with probability 1/2, and they are independent across different j's.

Each candidate D_i majority-defeats exactly half of the other D-candidates and all of the C-candidates, for a Copeland score of $2K-2$. Each of the C-candidates is majority-defeated by all of the D-candidates and so has a score of no more than $K+1 \leq 2K-2$. On the other hand, B defeats all of the D-candidates and A, and so has a score of at least $2K-2$. So by alphabetical tie-breaking, no matter what the manipulator does, either A or B must win.

Call a candidate C_j *defeated* if there are at least $\lfloor N/2 \rfloor + 1$ other voters ranking A, B above C_j. Let d be the number of defeated candidates. If the manipulator tells the truth, then A majority-defeats all the D_i and the defeated C_j, for a score of $2K-3+d$; B majority-defeats all the D_i, the defeated C_j, and A, for a score of $2K-2+d$. So B wins.

Now suppose the manipulator reports the ranking $AC_1 \ldots C_K D_1 \ldots D_{2K-3} B$. Say that the manipulator is *pivotal for* C_j if there are exactly $\lfloor N/2 \rfloor$ other voters ranking A, B above C_j. If the manipulator is pivotal for c candidates, then B still has a score of $2K-2+d$, but A now majority-defeats all the candidates for which the manipulator is pivotal and so has score $2K-3+d+c$. Thus, A wins if $c \geq 1$.

The probability of being pivotal for any given C_j is $\sim \sqrt{2/\pi N}$. Then the probability of being pivotal for at least one C_j is asymptotically K times this quantity, since the overlaps between these K events are negligible in comparison (pivotality for C_j is independent of pivotality for C_k for $j \neq k$, so the probability of being pivotal for C_j, C_k simultaneously is $\sim 2/\pi N$). That is, the probability of being pivotal for at least one C_j is $\sim K\sqrt{2/\pi N}$. The lower bound for susceptibility follows.

We still need to give the construction for the cases $M = 3, 4, 6, 7$. For $M = 6$, call the candidates A, B, C_1, C_2, D, E. Let the true preference be C_1C_2DEAB, and the proposed manipulation AC_1C_2DEB. The belief ϕ is given as follows:

- With probability 1/3, do the following: Begin with BAE; successively append C_1 and then C_2 either at the beginning or the end each with probability 1/2; finally, append D at the beginning.
- With probability 1/3, do the following: Begin with $BADE$; then successively append C_1 and then C_2 either at the beginning or the end each with probability 1/2.
- With probability 1/3, do the following: Begin with BAD; successively append C_1 and then C_2 either at the beginning or at the end each with probability 1/2; finally, add E at the beginning.

Now with probability $\stackrel{e}{\sim} 1$ we have the following majority-defeat relations:

- $A \to D, E$;
- $B \to A, D, E$;
- $D \to C_1, C_2, E$;
- $E \to C_1, C_2$.

Then C_1, C_2 both have score at most 3 since they are majority-defeated by D and E. Define d as before. Under truth-telling, A, B, D, E have respective scores $2+d, 3+d, 3, 2$, so that B wins. Under the proposed manipulation, A, B, D, E have scores $2+d+c, 3+d, 3, 2$, so that A wins if the manipulator is pivotal for either C_1 or C_2. The same argument as before shows that this occurs with probability $\sim 2\sqrt{2/\pi N}$.

If $M = 3$, let the manipulator's true preference be CBA, and the belief ϕ be

$$1/4\ ACB, \qquad 1/2\ BAC, \qquad 1/4\ CBA.$$

With probability $\stackrel{e}{\sim} 1$, the resulting profile will have $B \to A \to C$. If exactly $\lfloor N/2 \rfloor$ of the other voters have $B \succ C$, then the manipulator is pivotal for this pair: Telling the truth leads to $C \to B$, in which case A is the winner; manipulation leads to $B \to C$, so that B

wins, a more preferred outcome. If the manipulator is not pivotal, then the manipulation has no effect. So the manipulation is successful when the manipulator is pivotal, which happens with probability $\sim \sqrt{2/\pi N}$.

Finally, for $M = 4$ or 7, we take the construction for 3 or 6, respectively, and add an extra candidate at the end of everyone's preference ranking.

Fishburn's system. Assume $M \geq 4$, since the statement is trivial for $M = 3$. We return to the usual numerical labeling of the candidates. Let the manipulator's true preferences be $A_1 A_2 \ldots A_M$. As with the Copeland system, in order to describe the belief ϕ, we give a randomized procedure for generating a preference, and let ϕ denote the resulting distribution over $\mathcal{L}$.

- With probability 2/3, we construct a preference as follows: Begin with $A_2 A_1 A_3$, and then for each $i = 4, \ldots, M$ in succession, randomly append A_i either at the beginning of the existing ordering or at the end, independently with probability 1/2.
- With complementary probability 1/3, we instead do the following: Begin with $A_2 A_1$, then for each $i = 4, \ldots, M$ in succession, append A_i either at the beginning or the end, independently with probability 1/2; finally, append A_3 at the beginning.

A preference $\succ$ drawn according to this distribution has the following properties:

- With probability 1, $A_2 \succ A_1$.
- With probability 2/3, $A_1, A_2 \succ A_3$.
- For each $i \geq 4$, with probability 2/3, $A_3 \succ A_i$.
- For each $i \geq 4$, with probability 1/2, $A_1, A_2 \succ A_i$; and with probability 1/2, $A_2, A_1 \succ A_i$.

Let the proposed manipulation consist of moving A_2 to the bottom of the ranking, thus reporting $A_1 A_3 \ldots A_M A_2$. Let the set $\mathcal{C}^+$ in (2.3) be A_1 — so the manipulator is concerned only with the probability of A_1 winning.

As usual, with probability $\overset{e}{\sim} 1$, we have $A_2 \to A_1$, $A_1 \to A_3$, $A_2 \to A_3$, and $A_3 \to A_i$ for all $i \geq 4$. We may assume these relations hold.

If the manipulator tells the truth, then for each $i \geq 4$, either A_1, A_2 both majority-defeat A_i, or both are majority-defeated by A_i. It follows that A_1 is covered by A_2, and so cannot win.

Now consider manipulation. For each $i \geq 4$, if there are exactly $\lfloor N/2 \rfloor$ other voters who report $A_2 \succ A_i$, then the manipulation leads to $A_1 \to A_i \to A_2$ in the resulting profile. Say that the manipulator is *pivotal for* A_i if this occurs. In this case, A_2 no longer covers A_1. Notice that A_3 also cannot cover A_1, nor can any A_i for $i \geq 4$, since $A_1 \to A_3 \to A_i$. Hence, A_1 is uncovered and so wins.

So the manipulation is successful whenever the manipulator is pivotal for any A_i, $i \geq 4$. For each such A_i, the probability of being pivotal is $\sim \sqrt{2/\pi N}$. Moreover, as in the argument for the Copeland system above, since pivotality for A_i is independent of pivotality for A_j (for distinct $i, j \geq 4$), the probability of being pivotal for at least one A_i is $\sim (M-3)\sqrt{2/\pi N}$. This gives the claimed bound.

Minimax system. For this voting system, we will vary the beliefs ϕ as N varies. Doing so allows us to obtain a lower bound on susceptibility that converges more slowly than $N^{-1/2}$, although at the cost of requiring some additional computation.

For any given candidate A_i, we will use the term *defeater of* A_i to refer to any candidate A_j achieving the maximum, over $j \neq i$, of the number of voters preferring A_j to A_i.

We prove the bound for $M = 4$; the construction for higher M is identical with the extra candidates added at the end of everyone's preference. Let the four candidates be labeled A, B, C, D. Let the manipulator's preference be $ABCD$, and take the set of desirable candidates in (2.3) to be $\mathcal{C}^+ = \{A\}$. Consider the following belief ϕ:

$$\begin{array}{ll} \frac{1}{\sqrt{N}} & ACBD, \\ \frac{1}{2} - \frac{1}{\sqrt{N}} & ADBC, \\ \frac{1}{2} - \frac{1}{\sqrt{N}} & CBAD, \\ \frac{1}{\sqrt{N}} & DBAC. \end{array}$$

Let the proposed manipulation be $ACBD$.

In order to keep track of the consequences of manipulation, let the number of other voters reporting each of the four preferences be w, x, y, z, respectively. Then, the score of each candidate under truth-telling and under manipulation are as follows:

	Truth	Manipulation
A	$y+z$	$y+z$
B	$\max\{w+x+1, w+y, x+z\}$	$\max\{w+x+1, w+y+1, x+z\}$
C	$w+x+z+1$	$w+x+z+1$
D	$w+x+y+1$	$w+x+y+1$

(To obtain these values, note that B is always a defeater for A, since every voter who does not rank A first ranks B above A. Likewise, A is always a defeater for C and D.)

By Lemma 2.2(a), with probability $\overset{e}{\sim} 1$, D has score $> 3N/4$ whereas A has score $< 2N/3$ (whatever the manipulator does). So D cannot win. In particular, we can assume that the winner is whichever of A, B, C has the lowest score; this introduces negligible error.

We see from the table that the only possible effect of manipulation is to increase the score of B from $w+y$ to $w+y+1$. Hence, since we are ruling out D winning, manipulation can change the outcome of the vote in only two situations:

- Manipulation can change the outcome from B to A if it causes A and B to have equal scores, and C's score is at least as high. This requires $w+y \geq w+x+1, x+z$; $w = z-1$; and $y \leq w+x+1$.

  If $w = z-1$ then $w+x+1 = x+z$, so we actually need only $x+1 \leq y \leq w+x+1$ and $w = z-1$.

- Manipulation can also change the outcome from B to C. However, since both B and C are undesirable outcomes, this case contributes nothing to the expectation in (2.3).

Thus, we are left to estimate the probability that $x+1 \leq y \leq w+x+1$ and $w = z-1$.

Write s for the sum $w+z$, and t for $N-s = x+y$. We use Lemma C.2 to decompose the probability of a profile (w, x, y, z) into the probability of given values of s, t, times the probabilities of w conditional on s and of y conditional on t. Thus, the probability we want becomes

$$\sum_{\substack{s \text{ odd} \\ t=N-s}} \left[\mathbf{P}\left(\begin{matrix} s \\ t \end{matrix} \;\middle|\; N; \begin{matrix} \frac{2}{\sqrt{N}} \\ 1-\frac{2}{\sqrt{N}} \end{matrix} \right) \times \mathbf{P}\left(\begin{matrix} \frac{s-1}{2} \\ \frac{s+1}{2} \end{matrix} \;\middle|\; s; \begin{matrix} \frac{1}{2} \\ \frac{1}{2} \end{matrix} \right) \times \sum_{\frac{t+1}{2} \leq y \leq \frac{2t+s+1}{4}} \mathbf{P}\left(\begin{matrix} y \\ t-y \end{matrix} \;\middle|\; t; \begin{matrix} \frac{1}{2} \\ \frac{1}{2} \end{matrix} \right) \right]. \tag{E.1}$$

A lower bound for this outer sum is given by considering only the terms where $\sqrt{N} < s < 3\sqrt{N}$. In this case, by Lemma 2.1, $\min_s \mathbf{P}((s-1)/2, (s+1)/2 \mid s; 1/2, 1/2) \sim \sqrt{2/3\pi\sqrt{N}}$. Also, the probability in the inner sum of (E.1) is decreasing as a function of y for $y > t/2$, so each such term is at least

$$
\begin{aligned}
\mathbf{P}&\left(\begin{array}{c} \lceil (2t+s+1)/4 \rceil \\ t - \lceil (2t+s+1)/4 \rceil \end{array} \middle| \, t; \begin{array}{c} \frac{1}{2} \\ \frac{1}{2} \end{array} \right) \\
&= \mathbf{P}\left(\begin{array}{c} \lceil t/2 \rceil \\ \lfloor t/2 \rfloor \end{array} \middle| \, t; \begin{array}{c} \frac{1}{2} \\ \frac{1}{2} \end{array} \right) \times \prod_{k=1}^{\lceil (2t+s+1)/4 \rceil - \lceil t/2 \rceil} \frac{\lfloor t/2 \rfloor + 1 - k}{\lceil t/2 \rceil + k} \\
&\geq \min_{\substack{\sqrt{N} < s < 3\sqrt{N} \\ t = N-s}} \left[\mathbf{P}\left(\begin{array}{c} \lceil t/2 \rceil \\ \lfloor t/2 \rfloor \end{array} \middle| \, t; \begin{array}{c} \frac{1}{2} \\ \frac{1}{2} \end{array} \right) \times \left(\frac{t/2 - \lceil \frac{s+1}{4} \rceil}{t/2 + 1 + \lceil \frac{s+1}{4} \rceil} \right)^{\lceil \frac{s+1}{4} \rceil} \right] \\
&\gtrsim \min_{N - 3\sqrt{N} < t < N - \sqrt{N}} \left[\mathbf{P}\left(\begin{array}{c} \lceil t/2 \rceil \\ \lfloor t/2 \rfloor \end{array} \middle| \, t; \begin{array}{c} \frac{1}{2} \\ \frac{1}{2} \end{array} \right) \times \left(1 - \frac{6}{\sqrt{N}} \right)^{\lceil \sqrt{N} \rceil} \right] \\
&\sim \sqrt{\frac{2}{\pi N}} \times e^{-6}.
\end{aligned}
$$

Hence the expression in (E.1) is

$$
\gtrsim \sum_{\substack{s \text{ odd} \\ \sqrt{N} < s < 3\sqrt{N} \\ t = N-s}} \left[\mathbf{P}\left(\begin{array}{c} s \\ t \end{array} \middle| \, N; \begin{array}{c} \frac{2}{\sqrt{N}} \\ 1 - \frac{2}{\sqrt{N}} \end{array} \right) \times \sqrt{\frac{2}{3\pi\sqrt{N}}} \times \left\lfloor \frac{s+1}{4} \right\rfloor e^{-6} \sqrt{\frac{2}{\pi N}} \right]
$$

$$
\gtrsim \left[\sum_{\substack{s \text{ odd} \\ \sqrt{N} < s < 3\sqrt{N}}} \mathbf{P}\left(\begin{array}{c} s \\ N-s \end{array} \middle| \, N; \begin{array}{c} \frac{2}{\sqrt{N}} \\ 1 - \frac{2}{\sqrt{N}} \end{array} \right) \right] \times \frac{e^{-6}}{2\pi\sqrt{3}\sqrt[4]{N}}.
$$

To evaluate the bracketed probability sum, notice that if $(s, N-s)$ follows a binomial distribution $\mathbf{M}(N; 2/\sqrt{N}, 1-2/\sqrt{N})$ then s has mean $2\sqrt{N}$ and variance $2\sqrt{N}-4 < 2\sqrt{N}$; by Chebyshev's inequality, the probability that it differs from its mean by more than $\sqrt{N}$ is less than $2/\sqrt{N}$. Hence this probability sum is ~ 1. Consequently, the probability that the manipulator is pivotal, given by (E.1), is

$$
\gtrsim \frac{e^{-6}}{2\pi\sqrt{3}} \cdot \frac{1}{\sqrt[4]{N}}.
$$

Thus, the susceptibility is asymptotically bounded below by this quantity, as claimed.

Single transferable vote system. Fix small $\epsilon > 0$. Put $\epsilon' = 2\epsilon/(M-2)$. Let the manipulator's true preferences be $A_1 A_2 \ldots A_M$, and let the belief ϕ be as follows:

$$\begin{array}{ll} \frac{1}{2^{M-1}} + \epsilon & A_1 A_2 A_3 \ldots \\ \frac{1}{2^{M-1}} + \epsilon & A_2 A_3 \ldots \\ \frac{1}{2^{M-2}} - \epsilon' & A_3 A_2 \ldots \\ \frac{1}{2^{M-3}} - \epsilon' & A_4 A_2 A_3 \ldots \\ \frac{1}{2^{M-4}} - \epsilon' & A_5 A_2 A_3 \ldots \\ & \vdots \\ \frac{1}{2} - \epsilon' & A_M A_2 A_3 \ldots \end{array} .$$

(In this list, the first 3 preference types are $A_1 A_2 A_3 \ldots$, $A_2 A_3 \ldots$, and $A_3 A_2 \ldots$, and then the remaining preferences — if $M > 3$ — are all of the form $A_i A_2 A_3 \ldots$.) The part of each preference rankings denoted by $\ldots$ may be filled in arbitrarily with the remaining candidates.

By the usual application of Lemma 2.2(a), for N large, we can focus on the realizations such that for each preference ordering $\succ$, the share of the population reporting $\succ$ is within $\epsilon'' = \epsilon/M^2$ of the weight put on $\succ$ by distribution ϕ.

In this case, the single transferable vote procedure follows one of two possible execution paths. Either A_1 or A_2 is eliminated in the first round.

- Suppose A_1 is eliminated first. Then the candidates $A_1, A_3, A_4, A_5, \ldots, A_M$ are eliminated in succession. Indeed, we can show by induction that at the beginning of the kth round of elimination ($k > 1$) that candidates A_2 and $A_{k+1}, \ldots, A_M$ remain. If this holds for some k, then A_2 receives the votes of the first k preference types of voters, thus getting a vote share of at least

  $$\frac{1}{2^{M-k}} + 2\epsilon - (k-2)\epsilon' - k\epsilon'' > \frac{1}{2^{M-k}}.$$

  A_{k+1} has a vote share at most

  $$\frac{1}{2^{M-k}} - \epsilon' + \epsilon'' < \frac{1}{2^{M-k}},$$

  and each of the other remaining candidates has vote share at least

  $$\frac{1}{2^{M-k-1}} - \epsilon' - \epsilon'' > \frac{1}{2^{M-k}}.$$

Thus, A_{k+1} is eliminated next, and the voters who ranked A_{k+1} first have their votes transferred to A_2, giving the induction step.

Thus, in this case, A_2 ends up winning.

- Suppose A_2 is eliminated first. Then the voters who ranked A_2 first have their votes transferrred to A_3. In the second round, A_1 is eliminated, and the voters who ranked A_1 first have their votes transferred to A_3. An induction identical to the previous case now shows that $A_4, A_5, \ldots, A_M$ are eliminated in successive rounds. Thus A_3 ends up winning.

Consider a proposed manipulation of the form $A_2A_3\ldots$. The manipulator can potentially influence the first round of elimination, but conditional on the outcome of that round, the manipulator cannot affect subsequent eliminations. In the first round, pivotality occurs either when A_1 and A_2 receive the same number of first-place votes among the other voters, or when A_2 receives one more first-place vote than A_1. In both of these cases, if the manipulator tells the truth then A_2 is eliminated in the first round, hence A_3 ends up winning; under the proposed manipulation, A_1 is eliminated in the first round, and A_2 ends up winning. Hence, the manipulation is indeed beneficial.

By Lemma 2.2(b), we know each of the two pivotal scenarios happens with probability $\sim (1/2)\sqrt{1/\pi(1/2^{M-1}+\epsilon)N}$. Therefore, the total probability that the manipulator is pivotal is twice this quantity. We have thus shown

$$\sigma_N^{STV} \gtrsim \sqrt{\frac{1}{\pi\left(\frac{1}{2^{M-1}}+\epsilon\right)N}}.$$

Taking $\epsilon \to 0$ gives the result. □

F Analysis of the pair-or-plurality voting system

Proof of Lemma 3.7: For readability, we will refer to the candidates as A, B, C, with the understanding that this does not necessarily represent the tie-breaking order.

Pareto efficiency is immediate: if all voters rank A above B, then B is not viable and so cannot win. Hence we focus on monotonicity.

Consider a profile P at which some voter reports preference ABC. We need only consider what happens when this voter changes his preference by transposing the winner $f(P)$ with the candidate ranked immediately above her.

If $f(P) = C$, and the voter changes his preference to ACB, this cannot change the set of viable candidates, nor can it cause C to lose in a majority vote against another candidate given that C previously won this pairwise contest. This leaves only the case in which all three candidates are viable; in this case, the change can only increase C's score and decrease B's (while leaving A's unchanged), so that C remains the winner.

It remains to consider the case in which $f(P) = B$, and the voter changes his preference to BAC. Let P' be the resulting profile. The change cannot affect the set of viable candidates except by making A inviable. If this happens, A gets exactly K first-place votes at P. Suppose that $f(P') = C$ (otherwise $f(P') = B$ and we are done). Then, A, B, C are all viable at P, while only B and C are viable at P'.

We claim that B and C both have at least L first-place votes at P. If B has less than L first-place votes, then C has at least $N + 1 - K - L > (N + 1)/2$ first-place votes and so gets more than half the total points, giving $f(P) = C$, a contradiction. If C has less than L first-place votes, then B likewise has more than $(N+1)/2$ first-place votes and so $f(P') = B$.

Hence, at P, all the points from voters ranking B first go to B, and all the points from voters ranking C first go to C. Since there are K voters ranking A first, their points go to B and C in the same quantities as rank B or C second, respectively. So the outcome is effectively determined by a pairwise vote between B and C — exactly the same as at P'. Thus $f(P) = f(P')$, a contradiction.

Thus we can assume that the same set of candidates is viable at P as at P'. Since the change from ABC to BAC can only improve B's standing in a pairwise majority vote, we only need to concern ourselves with the case where all three candidates are viable at both P and P'.

Let us consider then the effect of changing ABC to BAC on each candidate's score. Given a profile where all three candidates are viable, write $s_A(A), s_A(B), s_A(C)$ for the number of points awarded to A, B, C, respectively, from the voters ranking A first. Let us consider the effect of removing an ABC vote on the vector $s_A = (s_A(A), s_A(B), s_A(C))$. If this leaves at least L total voters with A as their first-place vote, the net change in s_A is $(-1, 0, 0)$. Otherwise, $s_A(A)$ is changed by $-L/(L - K)$ and $s_A(C)$ is changed by $\leq L/2(L - K)$, so $s_A(B)$ is changed by $\geq -1 - (-L/(L - K)) - (L/2(L - K)) = (2K - L)/2(L - K)$. In short, the net change in s_A is of the form

$$\Delta s_A \quad = \quad (-1, 0, 0) \quad \text{or} \quad \left(-\frac{L}{L-K}, \ \geq \frac{2K-L}{2(L-K)}, \ \leq \frac{L}{2(L-K)}\right). \tag{F.1}$$

Now consider the effect of adding a BAC vote on the corresponding vector $s_B = (s_B(A), s_B(B), s_B(C))$ of points from the voters ranking B first. If there are initially at least L such voters, the net change is $(0,1,0)$; otherwise, $s_B(B)$ changes by $L/(L-K)$, $s_B(A)$ changes by at most $(L-2K)/2(L-K)$, and $s_B(C)$ changes by at most 0. So the net change in s_B is

$$\Delta s_B \quad = \quad (0,1,0) \quad \text{or} \quad \left(\leq \frac{L-2K}{2(L-K)}, \ \frac{L}{L-K}, \ \leq 0\right). \tag{F.2}$$

Finally, when one voter's preference changes from ABC to BAC, the net effect on the scores of the three candidates is given by the vector sum $\Delta s = \Delta s_A + \Delta s_B$. From (F.1), $\Delta s_A(B) \geq \Delta s_A(A)$, and from (F.2), $\Delta s_B(B) \geq \Delta s_B(A)$; thus $\Delta s(B) \geq \Delta s(A)$. From (F.1), $\Delta s_A(B) \geq \Delta s_A(C) - 1$, and from (F.2), $\Delta s_B(B) \geq \Delta s_B(C) + 1$; thus $\Delta s(B) \geq \Delta s(C)$. We conclude that the net change in B's score from P to P' is at least as large as the net change in A's score or C's score. Since B was the winner at the original profile P, then, B again wins at P'. So we have $f(P') = f(P)$ in this case as well, as required.

□

The proof of Proposition 3.8 will make use of the following two lemmas.

Lemma F.1 *Let x_N, y_N be sequences of positive integers with $(y_N - x_N)/N > \epsilon$, where $\epsilon > 0$ is some constant; $y_N \leq N - x_N$; and $x_N \to \infty$ as $N \to \infty$. Also let b be a fixed positive integer, and let a_N be any sequence of integers. Then*

$$\max_{\alpha_N \in [0,1]} \left[b \times \sum_{\substack{x_N < x < y_N \\ x \equiv a_N \pmod{b}}} \mathbf{P}\left(\begin{array}{c} x \\ N-x \end{array} \middle| \, N; \begin{array}{c} \alpha_N \\ 1-\alpha_N \end{array} \right) \right] \to 1. \tag{F.3}$$

(Here, for each N separately, we are maximizing over α_N.)

Proof: The sum of $\mathbf{P}(x, N-x \mid N; \alpha_N, 1-\alpha_N)$ over *all* x congruent to a_N mod b, without the restriction $x_N < x < y_N$, equals the sum of the coefficients of the corresponding terms z^x in the polynomial $(\alpha_N z + (1-\alpha_N))^N$, and so is computed by the formula

$$\frac{1}{b} \sum_{i=0}^{b-1} \zeta^{-a_N i} (\alpha_N \zeta^i + (1-\alpha_N))^N \tag{F.4}$$

where ζ is a primitive complex bth root of unity.

Let λ be a positive number such that $|\alpha\zeta^i + (1-\alpha)| < e^{-\lambda\alpha}$ for each $i = 1, \ldots, b-1$

and each $\alpha \in (0, 1/2)$; this holds if λ is sufficiently small. Also consider any $t > 0$, held fixed as N grows. For any N, if $t/N < \alpha_N < 1 - t/N$, then the term in the sum (F.4) corresponding to $i = 0$ always equals 1, and the other terms are all bounded above in absolute value by $e^{-\lambda t}$. On the other hand, if $\alpha_N \leq t/N$, then a variable following the multinomial distribution $\mathbf{M}(N; \alpha_N, 1 - \alpha_N)$ has mean $\leq t$ and variance $\leq t$, so its probability of exceeding x_N is at most $t/(x_N - t)^2$ by Chebyshev's inequality. This quantity goes to zero as $N \to \infty$ (since $x_N \to \infty$), so the sum in (F.3) does as well. A similar argument applies when $\alpha_N \geq 1 - t/N$. So in all three cases, the maximand in (F.3) is $\leq 1 + (b-1)e^{-\lambda t}$, once N is sufficiently large. By choosing t arbitrarily large, we see that the left-hand side of (F.3) is $\lesssim 1$.

To see that it is $\gtrsim 1$, simply take $\alpha_N = (x_N + y_N)/2N$ and apply (F.4) to obtain $\sum_{x \equiv a_N \pmod{b}} \mathbf{P}(x, N - x \mid N; \alpha_N, 1 - \alpha_N) \to 1/b$, and note that the probability of realizing a value $x < x_N$ or $x > y_N$ tends to 0, again by a Chebyshev argument (such a realization would require a multinomial $\mathbf{M}(N; \alpha_N, 1 - \alpha_N)$ to deviate from its mean $\alpha_N N$ by at least $N \cdot \epsilon/2$, which has probability $\leq N/(N \cdot \epsilon/2)^2 = 4/\epsilon^2 N \to 0$).

□

Lemma F.2 *Let S be a set of r-vectors of integers $(x_1, \ldots, x_r)$ each of which has sum N and satisfies $x_1, x_2 \geq K$, for some integer K. Then for any distribution α,*

$$\sum_{(x_1,\ldots,x_r) \in S} \left[\mathbf{P}\left(\begin{array}{c|c} x_1 & \\ x_2 & \\ x_3 & N;\ \alpha \\ \vdots & \\ x_r & \end{array} \right) - \mathbf{P}\left(\begin{array}{c|c} x_1 + 1 & \\ x_2 - 1 & \\ x_3 & N;\ \alpha \\ \vdots & \\ x_r & \end{array} \right) \right] \leq \frac{e^{1/12}}{\sqrt{\pi K}}.$$

Proof: We first prove the result for $r = 2$. In this case, taking N and α as fixed, the expression $h(x) = \mathbf{P}(x, N - x \mid N; \alpha_1, \alpha_2)$ is unimodal as a function of x. Letting x^* denote its maximum on the range $K \leq x \leq N - K$, the specified difference is negative for $x < x^*$ and nonnegative for $x \geq x^*$, so the sum in the lemma statement is maximized when S is the set of pairs whose x_1-values are $x^*, x^* + 1, \ldots, N - K$. In this case, the sum of differences telescopes and the sum is simply $h(x^*) - h(N - K + 1) \leq h(x^*)$. It follows from Lemma C.7 that $h(x^*) \leq e^{1/12}/\sqrt{\pi K}$. This completes the proof in the case $r = 2$.

For the general case, let S' be the set of all values of the $(r-1)$-tuple $x' = (x_1 + x_2, x_3, \ldots, x_r)$ for $(x_1, \ldots, x_r) \in S$. For each such $x' \in S'$, let $S_{x'}$ be the set of pairs (x_1, x_2) such that $(x_1, x_2, \ldots, x_r) \in S$.

By Lemma C.2, we can rewrite the sum in the lemma statement as as

$$\sum_{(x_1,\dots,x_r)\in S} \mathbf{P}\left(\begin{array}{c} x_{1+2} \\ x_3 \\ \vdots \\ x_r \end{array} \middle| N; \begin{array}{c} \alpha_{1+2} \\ \alpha_3 \\ \vdots \\ \alpha_r \end{array}\right) \times$$

$$\left[\mathbf{P}\left(\begin{array}{c} x_1 \\ x_2 \end{array} \middle| x_{1+2}; \begin{array}{c} \beta_1 \\ \beta_2 \end{array}\right) - \mathbf{P}\left(\begin{array}{c} x_1+1 \\ x_2-1 \end{array} \middle| x_{1+2}; \begin{array}{c} \beta_1 \\ \beta_2 \end{array}\right)\right]$$

(with $\beta_1 = \alpha_1/\alpha_{1+2}$ and β_2 similarly)

$$= \sum_{x'=(x_{1+2},x_3,\dots,x_r)\in S'} \mathbf{P}\left(x' \middle| N; \begin{array}{c} \alpha_1+\alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_r \end{array}\right) \times$$

$$\left(\sum_{(x_1,x_2)\in S_{x'}} \left[\mathbf{P}\left(\begin{array}{c} x_1 \\ x_2 \end{array} \middle| x_{1+2}; \begin{array}{c} \beta_1 \\ \beta_2 \end{array}\right) - \mathbf{P}\left(\begin{array}{c} x_1+1 \\ x_2-1 \end{array} \middle| x_{1+2}; \begin{array}{c} \beta_1 \\ \beta_2 \end{array}\right)\right]\right).$$

By the $r = 2$ case, the expression in square brackets is at most $e^{1/12}/\sqrt{\pi K}$, so the whole sum is

$$\leq \frac{e^{1/12}}{\sqrt{\pi K}} \sum_{x'\in S'} \mathbf{P}\left(x' \middle| N; \begin{array}{c} \alpha_1+\alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_r \end{array}\right).$$

Since the sum of probabilities is at most 1, we are done.

□

Proof of Proposition 3.8: We will prove the following claim: If λ is an integer such that $L/K > \lambda$ for each (sufficiently large) N, and $K \to \infty$ as $N \to \infty$, then the pair-or-plurality voting rule satisfies

$$\sigma_N^{POP} \lesssim \left(\frac{1}{2} + \frac{2}{\lambda-1}\right)\sqrt{\frac{3}{\pi N\left(1-\frac{1}{16\lambda^2}\right)}}. \tag{F.5}$$

The desired bound will then follow by taking $\lambda \to \infty$.

Our proof will make frequent use of the following observation: At any profile where all three candidates are viable, if a candidate A_i wins, then A_i must have score at least $(N+1)/3$. For each $A_j \neq A_i$, the voters ranking A_j first can contribute at most K points to A_i, so there must be at least $(N+1)/3 - 2K > L$ voters ranking A_i first. In particular, all the points from these voters are awarded to A_i; and even if we change one of their votes, the other such voters still award all their points to A_i.

Henceforth, as in the proof of Proposition 3.3, we will notate the manipulator's true preference as ABC for readability; this does not necessarily correspond to the tie-breaking order.

We first narrow down the manipulations we need to consider. With reference to (2.3), we have either $\mathcal{C}^+ = \{A\}$ or $\mathcal{C}^+ = \{A, B\}$. In the first case, the manipulator wants to maximize the probability of A winning. By monotonicity (Lemma 3.7) an optimal manipulation ranks A first, so we need only consider the manipulation ACB. However, we can show this manipulation cannot improve the outcome from B or C to A. To see this, first notice that it cannot change the set of viable candidates. It also cannot change the outcome of a pairwise vote between A and B, or between A and C, and so cannot have any effect if only one or two candidates are viable. This leaves the case when all three candidates are viable (at both the true profile, namely (ABC, P) for some opponent-profile P, and the manipulated profile, (ACB, P)). But if A wins at the manipulated profile, then our initial observation implies all voters ranking A first assign all their points to A, at both profiles. So the manipulation from ABC to ACB actually has no effect on any candidate's score, and thus no effect on the outcome.

This leaves us with the case $\mathcal{C}^+ = \{A, B\}$, so that the manipulator wishes to minimize the probability of C winning. By monotonicity again, an optimal manipulation ranks C last, so we can focus on the manipulation BAC. This manipulation cannot improve the outcome from C to A, again by monotonicity, so we need only consider the possibility that it improves from C to B.

So, let S denote the set of all opponent-profile realizations P such that $f(ABC, P) = C$ and $f(BAC, P) = B$. Now consider any such P. By our initial observation, P must include at least L first-place votes for B and L first-place votes for C. If the manipulation does not change the set of viable candidates, and only one or two candidates are viable, then manipulation cannot improve the outcome from C to B, by the same arguments as in the $\mathcal{C}^+ = \{A\}$ case. If manipulation does change the set of viable candidates, then it cannot make B become viable (since B has more than L votes), so it can only make A inviable: all three candidates are viable at (ABC, P), but only B, C are viable at (BAC, P). Then P contains exactly $K - 1$ first-place votes for A. But then the outcome at both (ABC, P) and (BAC, P) is determined by a a pairwise vote between B and C, which means the winner is the same at both profiles, a contradiction. Hence, for any $P \in S$, all three candidates are viable at both (ABC, P) and (BAC, P), and at both these profiles, all voters ranking B or C first award all their points to their first-choice candidate.

Now we can get quantitative. Our goal is to show that $\max_\phi \mathbf{P}(S \mid N; \phi) \lesssim G(N)$, where $G(N)$ is the right-hand side of (F.5). This will imply the proposition.

We will assume that $\phi \in \Delta(\mathcal{L})$ is the distribution attaining the max. (ϕ depends implicitly on N.) Also, in a generic opponent-profile P, write w, x, y, z, respectively, for the number of ABC votes, the number of ACB votes, the number of first-place votes for B, and the number of first-place votes for C. (We can regard BAC and BCA votes as equivalent, and CAB and CBA votes as equivalent.)

Let S_A be the set of pairs (w, x) such that there exist y and z with $(w, x, y, z) \in S$. For any fixed (w, x), there exists at most one such pair (y, z). Indeed: start with a P corresponding to some $(w, x, y, z) \in S$, and let P' be the $(N-1)$-profile obtained by removing a B-first vote. Then $f(ABC, BAC, P') = C$. So $f(BAC, CAB, P') = f(CAB, BAC, P') = C$ by monotonicity. Thus, if we change one B-first vote in P to a C-first vote, we get a profile P'' such that $f(BAC, P'') = C$, hence $P'' \notin S$. By monotonicity this remains true if we change further B-first votes to C-first votes.

Thus, we can regard y, z as functions of $(w, x) \in S_A$. Our desired probability $\mathbf{P}(S \mid N; \phi)$ can then be written by Lemma C.2 as

$$\sum_{(w,x,y,z)\in S} \mathbf{P}\left(\begin{array}{c} w \\ x \\ y+z \end{array} \,\middle|\, N; \begin{array}{c} \phi_w \\ \phi_x \\ \phi_{y+z} \end{array} \right) \mathbf{P}\left(\begin{array}{c} y \\ z \end{array} \,\middle|\, y+z; \begin{array}{c} \phi_y/\phi_{y+z} \\ \phi_z/\phi_{y+z} \end{array} \right)$$

$$= \sum_{(w,x)\in S_A} \mathbf{P}\left(\begin{array}{c} w \\ x \\ N-(w+x) \end{array} \,\middle|\, N; \begin{array}{c} \phi_w \\ \phi_x \\ 1-\phi_{w+x} \end{array} \right) \times \mathbf{P}\left(\begin{array}{c} y(w,x) \\ z(w,x) \end{array} \,\middle|\, N-(w+x); \begin{array}{c} \chi \\ 1-\chi \end{array} \right) \tag{F.6}$$

with $\chi = \phi_y/\phi_{y+z}$.

We will now proceed to bound each factor separately: we will show that

$$\sum_{(w,x)\in S_A} \mathbf{P}\left(\begin{matrix} w \\ x \\ N-(w+x) \end{matrix} \;\middle|\; N; \;\; \begin{matrix} \phi_w \\ \phi_x \\ 1-\phi_{w+x} \end{matrix} \right) \lesssim \frac{1}{2} + \frac{2}{\lambda - 1} \tag{F.7}$$

and

$$\max_{\substack{(w,x,y,z)\in S \\ \chi\in[0,1]}} \mathbf{P}\left(\begin{matrix} y \\ z \end{matrix} \;\middle|\; y+z; \;\; \begin{matrix} \chi \\ 1-\chi \end{matrix} \right) \lesssim \sqrt{\frac{3}{\pi N\left(1-\frac{1}{16\lambda^2}\right)}}. \tag{F.8}$$

The expression in (F.6) is bounded above by the product of the left-hand sides of (F.7) and (F.8). So if we can prove (F.7) and (F.8), then combining will give (F.5), which will in particular imply our desired result.

We first prove (F.8), which is the easier of the two. First note that the absolute difference between points awarded to B and points awarded to C from the voters who rank A first is always at most K; so for any $P \in S$, where either B or C can be made to win by adding just one vote, we must have $|y-z| \leq K+1$. Moreover, we claim that any profile in S satisfies $y+z \geq (2N-1)/3$. Indeed, if this is false, then $w+x > N/3 > L$, so at both of the profiles (ABC, P) and (BAC, P), the voters ranking A first award all their points to A, and the winner is simply determined by plurality vote. Then, for C and B to win at these two profiles respectively, we must have $z \geq (N+1)/3$ and $y+1 \geq (N+1)/3$, implying $y+z \geq (2N-1)/3$ after all.

Combining $|y-z| \leq K+1$ and $y+z \geq (2N-1)/3$ gives $y/(y+z), z/(y+z) \leq (2N+3K+2)/(4N-2)$. More simply, using $K \leq N/6\lambda$, we have for any $\epsilon > 0$ that

$$y, z \leq \kappa(y+z) \qquad \text{with} \qquad \kappa = \frac{1}{2} + \frac{1}{8\lambda} + \epsilon \qquad \text{for large } N.$$

Thus, the left-hand side of (F.8) is at most

$$\max_{\substack{\chi; y,z \\ y+z\geq(2N-1)/3 \\ y,z\leq\kappa(y+z)}} \mathbf{P}\left(\begin{matrix} y \\ z \end{matrix} \;\middle|\; y+z; \;\; \begin{matrix} \chi \\ 1-\chi \end{matrix} \right).$$

By Lemmas 2.3 and 2.4, for any given value of $s = y+z$, this maximum is attained by taking $y = \lfloor \kappa s \rfloor$ and $\chi = y/s$, and by Lemma 2.1, the value is asymptotically (in s)

bounded above by $\sqrt{1/2\pi s\kappa(1-\kappa)}$. Using $s \gtrsim 2N/3$ and the value of κ, we get

$$\max_{\substack{(w,x,y,z)\in S\\ \chi\in[0,1]}} \mathbf{P}\left(\begin{matrix} y \\ z \end{matrix} \;\middle|\; y+z;\; \begin{matrix} \chi \\ 1-\chi \end{matrix} \right) \lesssim \sqrt{\frac{3}{\pi N\left(1-\left(\frac{1}{4\lambda}+2\epsilon\right)^2\right)}}.$$

Letting $\epsilon \to 0$ gives (F.8).

It remains to prove (F.7).

To begin this, divide the set of pairs of nonnegative integers (w, x), with $w + x \geq K$, into four regions:

- R_1: $w + x < L$ and $w \geq x(1 - 2K/L) + K$.
- R_2: $w + x < L$ and $x \geq w(1 - 2K/L) + K$.
- R_3: $w + x < L$, $w < x(1 - 2K/L) + K$ and $x < w(1 - 2K/L) + K$.
- R_4: $w + x \geq L$.

When an opponent-profile $P = (w, x, y, z)$ is drawn from $\mathbf{M}(N;\phi)$, the probability that (w, x) lies in one region and $(w + 1, x)$ lies in a different region converges to zero (uniformly over ϕ) as $N \to \infty$. Indeed, we have the following exhaustive list of possible subcases:

- $(w, x) \in R_1$, R_2 or R_3 but $(w+1, x) \in R_4$: Can happen only if $w+x = L-1$, which occurs with probability $\leq e^{1/12}/\sqrt{\pi(L-1)}$ by Lemmas C.2 and C.7.
- $(w, x) \in R_2$ but $(w + 1, x) \in R_1$ or R_3: This requires $w(1 - 2K/L) + K \leq x < (w+1)(1-2K/L)+K$. For any given value of w, there is at most one integer value of x in this range, and it must be at least K. So conditional on the realization of w, this x occurs with probability at most $e^{1/12}/\sqrt{\pi K}$ (again by Lemmas C.2 and C.7), and hence this same bound applies to the unconditional probability of this subcase.
- $(w, x) \in R_3$ but $(w + 1, x) \in R_1$: Requires $w = x(1 - 2K/L) + K - 1$, so similarly to the previous subcase, this occurs with probability at most $e^{1/12}/\sqrt{\pi(K-1)}$.

Since $K, L \to \infty$ as $N \to \infty$, we see that the total probability of any of these subcases goes to zero.

Thus, since our goal is to prove (F.7), we can modify S_A by removing all pairs such that (w, x) and $(w + 1, x)$ are not in the same region; this introduces negligible error. So from here on we work with the modified S_A.

Given now that (w, x) and $(w+1, x)$ are assumed to lie in the same region, we compute the scores of B and C associated with the true profile (ABC, P) $(= (w+1, x, y, z))$ and the manipulated profile (BAC, P) $(= (w, x, y+1, z))$, for each of the four regions:

- R_1: The associated scores are

$$(ABC, P): \qquad \frac{K(L-w-x-1)}{L-K} + y \text{ for } B, \qquad z \text{ for } C$$

$$(BAC, P): \qquad \frac{K(L-w-x)}{L-K} + y + 1 \text{ for } B, \qquad z \text{ for } C.$$

  Thus, in this region, the manipulation increases B's score by $L/(L-K)$ and leaves C's score unaffected.

- R_2:

$$(ABC, P): \qquad y \text{ for } B, \qquad \frac{K(L-w-x-1)}{L-K} + z \text{ for } C$$

$$(BAC, P): \qquad y+1 \text{ for } B, \qquad \frac{K(L-w-x)}{L-K} + z \text{ for } C.$$

  Thus, manipulation increases B's score by 1 and C's score by $K/(L-K)$.

- R_3:

$$(ABC, P): \quad w+1-\frac{(w+x+1-K)L}{2(L-K)}+y \text{ for } B, \quad x-\frac{(w+x+1-K)L}{2(L-K)}+z \text{ for } C$$

$$(BAC, P): \qquad w-\frac{(w+x-K)L}{2(L-K)}+y+1 \text{ for } B, \qquad x-\frac{(w+x-K)L}{2(L-K)}+z \text{ for } C.$$

  Thus, in this region, manipulation has no effect on the difference between B's and C's scores, so it cannot change the winner from C to B: $R_3 \cap S_A = \emptyset$.

- R_4:

$$(ABC, P): \qquad y \text{ for } B, \qquad z \text{ for } C$$

$$(BAC, P): \qquad y+1 \text{ for } B, \qquad z \text{ for } C.$$

  Thus, in this region, manipulation increases B's score by 1 and leaves C's score unaffected.

Henceforth, we assume that tie-breaking favors B over C. (If the reverse is the case, all the same arguments will go through with only minor adjustments.)

Let T be the set of all pairs (w, x) such that $K \leq w + x \leq N - K$ and $N - w - x$ is odd.

By Lemma F.1, we know that the maximum probability (over all ϕ) of drawing $(w, x) \in T$ is $\sim 1/2$. Our strategy will be to show that the probability of $(w, x) \in S_A$ cannot be much larger than the probability of $(w, x) \in T$.

We begin by comparing the probability of $(w, x) \in R_1 \cap S_A$ and the probability of $(w, x) \in R_1 \cap T$.

First consider the possibility that

$$(w, x) \in R_1 \cap S_A \qquad \text{and} \qquad (w + 1, x) \in R_1 \cap S_A. \tag{F.9}$$

We observe that if

$$\left\lfloor \frac{K(L - w - x - 2)}{L - K} \right\rfloor = \left\lfloor \frac{K(L - w - x - 1)}{L - K} \right\rfloor = \left\lfloor \frac{K(L - w - x)}{L - K} \right\rfloor, \tag{F.10}$$

then (F.9) cannot occur. Indeed, $(w, x) \in R_1 \cap S_A$ means that, for some suitable choice of (y, z),

$$\frac{K(L - w - x - 1)}{L - K} + y < z \qquad \text{but} \qquad \frac{K(L - w - x)}{L - K} + y + 1 \geq z$$

which, under (F.10), implies that $\lfloor K(L - w - x - 1)/(L - K) \rfloor + y = z - 1$. Then $\lfloor K(L-w-x-1)/(L-K) \rfloor + (N-w-x)$ must be odd (since $N-w-x = y+z$). Likewise, $(w+1, x) \in R_1 \cap S_A$ together with (F.10) requires $\lfloor K(L-w-x-1)/(L-K) \rfloor + (N-w-1-x)$ to be odd. But these expressions cannot both be odd.

Let

$$V = \left\{ v \;\middle|\; K \leq v \leq N - K \text{ and } \left\lfloor \frac{K(L - v - 1)}{L - K} \right\rfloor < \left\lfloor \frac{K(L - v)}{L - K} \right\rfloor \right\},$$

$$V' = \left\{ v \;\middle|\; K \leq v \leq N - K \text{ and } \left\lfloor \frac{K(L - v - 2)}{L - K} \right\rfloor < \left\lfloor \frac{K(L - v - 1)}{L - K} \right\rfloor \right\}.$$

Thus, the probability that (F.9) arises is at most the probability of $w + x \in V \cup V'$. ($w + x \leq N - K$ holds since B and C need to be viable.)

Since $(L - K)/K > \lambda - 1$, it follows that of any $\lambda - 1$ consecutive integers, at most one can be in V.

Now, we claim that for any set $V \subseteq \{K, K+1, \ldots, N-K\}$ with this property, there exists a such that

$$\sum_{v\in V} \mathbf{P}\left(\begin{array}{c} v \\ N-v \end{array} \middle| N; \begin{array}{c} \phi_{w+x} \\ 1-\phi_{w+x} \end{array} \right) \leq \sum_{\substack{K\leq v\leq N-K \\ v\equiv a \bmod \lambda-1}} \mathbf{P}\left(\begin{array}{c} v \\ N-v \end{array} \middle| N; \begin{array}{c} \phi_{w+x} \\ 1-\phi_{w+x} \end{array} \right). \quad \text{(F.11)}$$

Indeed, choose a to be the value of $v \in V$ for which $\mathbf{P}(v, N-v \mid N; \phi_{w+x}, 1-\phi_{w+x})$ is maximized. Since this latter expression is unimodal in v, for any two successive elements of V that differ by more than $\lambda - 1$, either the lower element can be increased by 1 or the higher element can be decreased by 1 in such a way that the expression on the left side of (F.11) is increased. We thus replace V by a new set for which the left-hand side of (F.11) is higher than before. This operation cannot be repeated forever; when it terminates, it must be that every two consecutive elements of the current set differ by $\lambda - 1$. The resulting set clearly satisfies (F.11), and so the original set V did as well.

By Lemma F.1, the right-hand side of (F.11) is $\lesssim 1/(\lambda - 1)$. Hence, the same holds for the probability of $w + x \in V$. The same argument applies to $w + x \in V'$ as well. We conclude that the probability of (F.9) is $\lesssim 2/(\lambda - 1)$.

Now we are ready to compare the probability of $(w, x) \in R_1 \cap S_A$ with that of $(w, x) \in R_1 \cap T$. To economize on notation, henceforth, we write simply $\mathbf{P}(w, x)$ for the probability of realizing (w, x) under $\mathbf{M}(N; \phi)$.

Notice that

$$\begin{aligned}
\sum_{(w,x)\in R_1\cap S_A} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} \leq \sum_{(w,x)\in R_1\cap T} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} &+ \\
&\sum_{\substack{(w,x)\in R_1\cap S_A\setminus T \\ (w+1,x)\in R_1\cap T\setminus S_A}} \left[\mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} - \mathbf{P}\begin{pmatrix} w+1 \\ x \end{pmatrix} \right] + \\
&\sum_{\substack{(w,x)\in R_1\cap S_A\setminus T \\ (w+1,x)\in R_1\cap S_A}} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} + \\
&\sum_{\substack{(w,x)\in R_1\cap S_A\setminus T \\ (w+1,x)\in R_1\setminus T}} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix}.
\end{aligned}$$

The second sum on the right-hand side is at most $e^{1/12}/\sqrt{2\pi K}$, by Lemma F.2 (notice that $(w,x) \in R_1$ ensures $w \geq K$). The third sum consists of pairs (w,x) satisfying (F.9), which have a total probability $\lesssim 2/(\lambda-1)$, by the preceding argument. And the fourth sum is empty.

The bound $e^{1/12}/\sqrt{\pi K}$ goes to 0 as $N \to \infty$, and thus we get

$$\sum_{(w,x)\in R_1\cap S_A} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} - \sum_{(w,x)\in R_1\cap T} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} \lesssim \frac{2}{\lambda-1}. \tag{F.12}$$

This takes care of $R_1 \cap S_A$ for now. Next let us perform a similar analysis for pairs $(w,x) \in R_2 \cap S_A$.

Consider the possibility that

$$(w,x) \in R_2 \cap S_A \qquad \text{and} \qquad (w+1,x) \in R_2 \cap S_A. \tag{F.13}$$

If $(w,x) \in R_2 \cap S_A$, then for suitable choices of (y,z),

$$y < \frac{K(L-w-x-1)}{L-K} + z \qquad \text{and} \qquad y+1 \geq \frac{K(L-w-x)}{L-K} + z.$$

This means that

$$\left\lceil \frac{K(L-w-x-1)}{L-K} \right\rceil = \left\lceil \frac{K(L-w-x)}{L-K} \right\rceil = y - z + 1.$$

Hence, $\lceil K(L-w-x-1)/(L-K)\rceil + (N-w-x)$ must be odd; and if $(w+1,x) \in R_2 \cap S_A$, then $\lceil K(L-w-x-1)/(L-K)\rceil + (N-w-x-1)$ must be odd. These quantities cannot both be odd, however. So we see that (F.13) can never occur.

Thus, we can perform an analysis for the probability of $(w,x) \in R_2 \cap S_A$ that entirely parallels what we did for $R_1 \cap S_A$, but our life is now simplified by the fact that (F.13) has probability zero (unlike its counterpart (F.9)). The result is

$$\sum_{(w,x)\in R_2\cap S_A} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} - \sum_{(w,x)\in R_2\cap T} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} \leq \frac{e^{1/12}}{\sqrt{\pi K}} \to 0. \tag{F.14}$$

Next we turn to R_3. Since $R_3 \cap S_A = \emptyset$, we simply have

$$\sum_{(w,x)\in R_3\cap S_A} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} = 0 \leq \sum_{(w,x)\in R_3\cap T} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix}. \tag{F.15}$$

Finally, we claim that $R_4 \cap S_A \subseteq T$. Check: if $(w,x) \in R_4 \cap S_A$, then $y < z$ but $y+1 \geq z$, so $y = z-1$, and hence $N-w-x = y+z$ is odd; also $w+x \leq N-K$ by the viability of B and C. So it is immediate that

$$\sum_{(w,x)\in R_4\cap S_A} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} \leq \sum_{(w,x)\in R_4\cap T} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix}. \tag{F.16}$$

Finally, using the fact that every $(w,x) \in S_A$ must lie in one of the regions R_1, R_2, R_3, R_4, we can combine (F.12), (F.14), (F.15), (F.16):

$$\begin{aligned}
\sum_{(w,x)\in S_A} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} &= \sum_{i=1}^{4} \sum_{(w,x)\in R_i\cap S_A} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} \\
&\lesssim \sum_{i=1}^{4} \sum_{(w,x)\in R_i\cap T} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} + \frac{2}{\lambda-1} \\
&\leq \sum_{(w,x)\in T} \mathbf{P}\begin{pmatrix} w \\ x \end{pmatrix} + \frac{2}{\lambda-1} \\
&\lesssim \frac{1}{2} + \frac{2}{\lambda-1}.
\end{aligned}$$

Thus, we have proven (F.7). Combining with (F.8), as previously mentioned, gives the result.

□

G Proofs of lower bounds

The proofs of the results from Section 4 are in this appendix (except for results that are proven in the main text, and Theorem 4.3 which is in the next appendix). We present the proofs in the same order that they are sketched in the text.

Proof of Theorem 4.5: For any two candidates A, B, let $K^*(A;B)$ be the maximum number K such that $f(K\ A, N+1-K\ B) \neq A$. By unanimity, $K^*(A;B) < N+1$.

Let us call a triple (A,B,C) of distinct candidates *unobtrusive* if

$$f(1\ A, K\ B, N-K\ C) \neq A \qquad \text{for all} \qquad K.$$

Note that this also implies $f(K\ B, N+1-K\ C) \neq A$ for all K (otherwise, change one of the B or C votes to A, and we get a violation of monotonicity).

Fix any triple (A, B, C). Write K^* for $K^*(B, C)$ defined above. Also write $\widetilde{K}^*$ for the maximum value of K such that $f(1\ A, K\ B, N-K\ C) \neq B$ (or $\widetilde{K}^* = -1$ if no such K exists). Notice that $\widetilde{K}^* \geq K^* - 1$, since otherwise $f(1\ A, \widetilde{K}^*+1\ B, N-\widetilde{K}^*-1\ C) = B$ and $f(K^*\ B, N+1-K^*\ C) \neq B$ would violate monotonicity.

We will show the inequality $\sigma \geq \sigma_N^*$ in each of the following cases:

(i) $\widetilde{K}^* = K^* - 1$;

(ii) $\widetilde{K}^* = K^*$ and (A, B, C) is unobtrusive;

(iii) $\widetilde{K}^* > K^*$.

First note that

$$f(K\ B, N+1-K\ C) = B \qquad \text{for all} \qquad K > K^* \tag{G.1}$$

by definition, and

$$f(K\ B, N+1-K\ C) \neq B \qquad \text{for all} \qquad K \leq K^* \tag{G.2}$$

since otherwise monotonicity would imply $f(K^*\ B, N+1-K^*\ C) = B$, a contradiction. By similar arguments,

$$f(1\ A, K\ B, N-K\ C) = B \qquad \text{for all} \qquad K > \widetilde{K}^*; \tag{G.3}$$

$$f(1\ A, K\ B, N-K\ C) \neq B \qquad \text{for all} \qquad K \leq \widetilde{K}^*. \tag{G.4}$$

Now for the case analysis:

- Case (i): Let the manipulator's true preference be any ordering with A ranked first and B last; let the proposed manipulation be a vote for C; and let the manipulator's belief be $\phi = (\phi_B\ B, (1-\phi_B)\ C)$ with $\phi_B = K^*/N$. Since the other voters all vote for B or C, (G.1)–(G.4) imply that the manipulator cannot affect whether B wins, unless the realized opponent-profile is $(K^*\ B, N-K^*\ C)$, in which case a vote for A leads to B winning and a vote for C leads to B losing. So considering the definition (2.3) of susceptibility with $\mathcal{C}^+ = \mathcal{C} \setminus \{B\}$, we have

$$\sigma \geq \mathbf{P}\left(\begin{array}{cc} K^* & B \\ N-K^* & C \end{array} \middle| N;\ \phi \right),$$

which is $\geq \sigma_N^*$ by Lemma 2.4.

- Case (ii): Let the manipulator's true preference be any ordering with A ranked first and B second; let the proposed manipulation be a vote for B; and again let the manipulator's belief be $\phi = (\phi_B\ B, (1-\phi_B)\ C)$ with $\phi_B = K^*/N$. By unobtrusiveness, no matter whether the manipulator votes for A or B, A cannot win. Again, (G.1)-(G.4) imply that the manipulator cannot affect whether or not B wins, unless the realized opponent-profile is $(K^*\ B, N-K^*\ C)$ in which case a vote for A leads to B losing and a vote for B leads to B winning. So considering (2.3) with $\mathcal{C}^+ = \{A, B\}$, we again have

$$\sigma \geq \mathbf{P}\left(\begin{array}{cc} K^* & B \\ N-K^* & C \end{array} \middle| N;\ \phi \right) \geq \sigma_N^*.$$

- Case (iii): Suppose $\widetilde{K}^* > K^*$. Let the manipulator's true preference be any ordering with C ranked first and B last; let the proposed manipulation be a vote for A; and let the belief be $\phi = (\phi_B\ B, (1-\phi_B)\ C)$ with $\phi_B = (K^*+1)/N$. Once again, the manipulator cannot affect whether or not B wins, unless the opponent-profile is $(K\ B, N-K\ C)$ for some K with $K^* < K \leq \widetilde{K}^*$, in which case a vote for C leads to B winning and a vote for A leads to B losing. Considering (2.3) with $\mathcal{C}^+ = \mathcal{C} \setminus \{B\}$, we again have

$$\sigma \geq \sum_{K=K^*+1}^{\widetilde{K}^*} \mathbf{P}\left(\begin{array}{cc} K & B \\ N-K & C \end{array} \middle| N;\ \phi \right) \geq \mathbf{P}\left(\begin{array}{cc} K^*+1 & B \\ N-K^*-1 & C \end{array} \middle| N;\ \phi \right) \geq \sigma_N^*.$$

Now, if any triple of candidates (A, B, C) is unobtrusive, then since we already observed $\widetilde{K}^* \geq K^* - 1$, one of the cases (i)–(iii) must hold. So to prove the inequality, it remains only to consider the case that no unobtrusive triple exists.

In this case, choose A, B, C so that $f(1\ A, K\ B, N-K\ C) = A$ for K as large as possible. By assumption, there also exists K' such that $f(1\ C, K'\ B, N-K'\ A) = C$, and by maximality $K' \leq K$. If $K < N$, then monotonicity implies $f(N-K'\ A, K'\ B, 1\ C) = A$, a contradiction. Therefore $K = N$, so that $f(1\ A, N\ B) = A$. Again by assumption, there exists K'' such that $f(1\ B, K''\ C, N-K''\ A) = B$. If $K'' < N$ then monotonicity implies $f(1\ A, N\ B) = B$, a contradiction. So $K'' = N$, or $f(1\ B, N\ C) = B$. By monotonicity again, $f(N\ B, 1\ C) = B$.

Suppose the manipulator's true preference ranks C first and B last; let the proposed manipulation be a vote for A, and let the belief be that everyone else votes for B with probability 1. Then a truthful vote for C leads to B winning, while manipulation leads to A winning, hence (taking $\mathcal{C}^+ = \mathcal{C} \setminus \{B\}$) we have susceptibility $\sigma = 1$.

This proves that the inequality $\sigma \geq \sigma_N^*$ always holds.

It remains to study the equality case. This proof roughly follows the above case analysis but requires further splitting into subcases. We prove the contrapositive: suppose that f is not a majority rule; we will show that $\sigma > \sigma_N^*$ strictly. So there is a profile at which strictly more than half the voters vote for some candidate — say C — but some other candidate wins — say B. We may assume B and C are chosen so as to maximize the number of voters voting for C with B winning.

By monotonicity, B still wins when all the non-C votes are replaced by B's, and it follows that $K^* = K^*(B;C) \leq (N-2)/2$. The extremal choice of B and C implies that whenever at least $N + 1 - K^*$ voters vote for C, then C wins.

Let A be an arbtirary candidate distinct from B and C. Define $\widetilde{K}^*$ as before. We have $\widetilde{K}^* \geq K^* - 1$ again. We now review the cases from the previous analysis, making amendments as needed; but now we also add the case where $\widetilde{K}^* = K^*$ and (A, B, C) is not unobtrusive.

- In case (i), the same argument as before applies. Since $K^* < (N-1)/2$, Lemma 2.4 implies that the inequality at the end of case (i) holds strictly.
- In case (ii) the analysis goes through as before and again the final inequality holds strictly.
- Case (ii'), where $\widetilde{K}^* = K^*$ but (A, B, C) is not unobtrusive. Then we have $f(1\ A, K\ B, N - K\ C) = B$ whenever $K > K^*$. Since at least $N + 1 - K^*$ votes for C make C win, then, obtrusiveness can only happen for $K = K^*$: $f(1\ A, K^*\ B, N - K^*\ C) = A$. By monotonicity, we then have

  $$f(J\ A, K\ B, N+1-J-K\ C) = A \qquad \text{for all} \qquad K \leq K^*,\ J+K-1 \geq K^*. \quad \text{(G.5)}$$

  And the extremal property of B and C implies that

  $$f(J\ A, K\ B, N+1-J-K\ C) = C \qquad \text{whenever} \qquad J+K-1 < K^*. \quad \text{(G.6)}$$

  If $K^* \geq 1$ then (G.5) and (G.6) imply that we can use the triple (B, A, C) instead of (A, B, C): this triple has the same value of K^*, but falls into case (i), from which the proof is complete.

Finally suppose $K^* = 0$. Then we have $f(1\ A, 1\ B, N-1\ C) = B$; $f(1\ A, N\ C) = A$ (and by monotonicity $f(K\ A, N+1-K\ C) = A$ for all $K \geq 1$); $f(1\ B, N\ C) = B$; and $f(N+1\ C) = C$. Let the manipulator have true preference ranking C first, B second, and A last; let the proposed manipulation be a vote for B, and let the belief ϕ be $1/N\ A, (N-1)/N\ C$. If the realized opponent profile is that all others vote for C, then truthful voting leads to C winning, while manipulating leads to B winning. For any other possible opponent-profile, telling the truth leads to A winning, and at least when the opponent-profile is $(1\ A, N-1\ C)$, manipulation leads to B winning instead. It follows by taking $\mathcal{C}^+ = \mathcal{C} \setminus \{A\}$ that

$$\sigma \geq \mathbf{P}\left(\begin{array}{cc} 1 & A \\ N-1 & C \end{array} \middle| \ N;\ \phi \right) > \sigma_N^*.$$

- In case (iii), if $K^* \leq (N-4)/2$, then the final inequality in case (iii) becomes strict, again by Lemma 2.4. So we may assume $K^* > (N-4)/2 \geq 0$.

  If $f(1\ A, K^*\ B, N-K^*\ C) = A$, we again have (G.5) and (G.6), so that just as in case (ii) above, we can replace the triple (A, B, C) by (B, A, C), and end up in case (i), for which the proof has been completed. (Note that this uses our assumption $K^* > 0$.)

  Finally, suppose $f(1\ A, K^*\ B, N-K^*\ C) \neq A$. We also have $f(1\ A, K^*\ B, N-K^*\ C) \neq B$ by the assumption of case (iii).

  As before, the extremal property of B and C implies $f(1\ A, K\ B, N-K\ C) = C$ for $K < K^*$. In this case, consider the same preferences, belief, and proposed manipulation as in the original analysis for case (ii). If the realized opponent-profile is $(K\ B, N-K\ C)$ for $K < K^*$, then C wins regardless of whether the manipulator votes for A or B. Otherwise, a vote for B will ensure that B wins, while a vote for A will fail to ensure an outcome in $\mathcal{C}^+ = \{A, B\}$ if the realized opponent-profile is $(K^*\ B, N-K^*\ C)$. Hence (2.3) with $\mathcal{C}^+ = \{A, B\}$ gives

$$\sigma \geq \mathbf{P}\left(\begin{array}{cc} K^* & B \\ N-K^* & C \end{array} \middle| \ N;\ \phi \right) > \sigma_N^*.$$

This shows that $\sigma > \sigma_N^*$ in every possible case.

□

Next we proceed to the proof of Theorem 4.4, for monotone and Pareto-efficient voting rules. This proof makes reference to proof techniques from Theorem 4.7, which was given in the main text.

Proof of Lemma 4.10: For each $K = 0, \ldots, \overline{K}$, let $J(K)$ be the highest value such that $f(P_{J,K}) = A_j$, or $J(K) = \underline{J} - 1$ if no such value exists. By (i) and (iii), $f(P_{J,K}) = A_j$ for $J \leq J(K)$ and $= A_i$ for $J > J(K)$. Also (iii) ensures that $J(K-1) \leq J(K) + 1$ (whenever these quantities are defined); hence $J(K) + K$ is weakly increasing in K. We further note for later reference that

$$J(K) + K < \widetilde{N} \qquad \text{for each} \qquad K, \tag{G.7}$$

since otherwise $f(P_{J(K),K}) = A_j$, together with $f(P_{\widetilde{N}-\overline{K},\overline{K}}) = A_i$ from (v), would contradict (iii).

Choose integer values $0 = K_0 < K_1 < K_2 < \cdots < K_r = \overline{K}$, where any two successive K_i differ by at most $40\sqrt{\widetilde{N}}/\kappa$ and with $r \leq \sqrt{\widetilde{N}}\kappa/20$. Certainly this can be done, as long as N is sufficiently large.

Now, by (iv), $J(0) = \overline{J}$, while by (v), $J(\overline{K}) = \underline{J} - 1$. Therefore

$$J(0) - J(\overline{K}) > \overline{J} - \underline{J} > \kappa\widetilde{N}.$$

Therefore, there exists some $i \in \{1, \ldots, r\}$ such that

$$J(K_{i-1}) - J(K_i) > \frac{\kappa\widetilde{N}}{r} \geq 20\sqrt{\widetilde{N}}.$$

Put

$$\gamma = \frac{J(K_{i-1}) + J(K_i)}{2\widetilde{N}},$$

$$\delta_1 = \frac{K_{i-1} + \sqrt{2\widetilde{N}}}{\widetilde{N}}, \qquad \phi_1 = \begin{pmatrix} \gamma & \succ \\ \delta_1 & \succ' \\ 1-\gamma-\delta_1 & \succ'' \end{pmatrix},$$

$$\delta_2 = \min\left\{\frac{K_i - \sqrt{2\widetilde{N}}}{\widetilde{N}}, 1-\gamma\right\}, \qquad \phi_2 = \begin{pmatrix} \gamma & \succ \\ \delta_2 & \succ' \\ 1-\gamma-\delta_2 & \succ'' \end{pmatrix}.$$

It is straightforward to check that ϕ_1 and ϕ_2 are legitimate probability distributions (that is, all entries are nonnegative); the only nontrivial part is $K_i - \sqrt{2\widetilde{N}} \geq 0$ which follows from $K_i - K_{i-1} \geq J(K_{i-1}) - J(K_i) \geq 20\sqrt{\widetilde{N}}$.

We will show that

$$\overline{f}_{A_j}(\phi_1) > 3/4 \tag{G.8}$$

and

$$\overline{f}_{A_i}(\phi_2) > 3/4. \tag{G.9}$$

Suppose that the $\widetilde{N}$-profile $P = (x \ \succ, y \ \succ', z \ \succ'')$ is drawn according to $IID(\phi_1)$. If the inequalities

$$x \geq \underline{J} \tag{G.10}$$

$$y \geq K_{i-1} \tag{G.11}$$

$$x + y \leq K_{i-1} + J(K_{i-1}) \tag{G.12}$$

are satisfied, then we must have $P \in R$ (since $x \leq J(K_{i-1}) \leq \overline{J}$, and $x + y \leq \overline{K} + J(\overline{K})$ implying $y < \overline{K}$). Moreover in this case $f(P) = A_j$, on account of $f(P_{J(K_{i-1}),J_{i-1}}) = A_j$ and the monotonicity relation (iii). Notice also that if

$$x \leq (3J(K_{i-1}) + J(K_i))/4 \tag{G.13}$$

$$y \leq K_{i-1} + 4\sqrt{\widetilde{N}} \tag{G.14}$$

are satisfied, then (G.12) will automatically hold.

Now we apply the same Chebyshev argument as in the proof of Lemma 4.2. We have $(x, \widetilde{N} - x) \sim \mathbf{M}(\widetilde{N}; \gamma, 1 - \gamma)$, so (G.10) and (G.13) are satisfied unless $|x - E[x]| \geq (J(K_{i-1}) - J(K_i))/4$, which happens with probability

$$\Pr\left(|x - E[x]| \geq \frac{J(K_{i-1}) - J(K_i)}{4}\right) \leq \frac{Var(x)}{\left(\frac{J(K_{i-1})-J(K_i)}{4}\right)^2} \leq \frac{\widetilde{N}/4}{(5\sqrt{\widetilde{N}})^2} = \frac{1}{100}.$$

Likewise, $(y, \widetilde{N} - y) \sim \mathbf{M}(\widetilde{N}; \delta_1, 1 - \delta_1)$, so (G.11) and (G.14) are satisfied unless $|y - E[y]| \geq \sqrt{2\widetilde{N}}$, which happens with probability

$$\Pr(|y - E[y]| \geq \sqrt{2\widetilde{N}}) \leq \frac{Var(y)}{\left(\sqrt{2\widetilde{N}}\right)^2} \leq \frac{\widetilde{N}/4}{2\widetilde{N}} = \frac{1}{8}.$$

We conclude that (G.10), (G.11), (G.13), (G.14) are all satisfied — and hence $f(P) = A_j$ — with probability at least $1 - 1/100 - 1/8 > 3/4$. This gives (G.8).

Similarly, suppose that the $\widetilde{N}$-profile $P = (x \ \succ, y \ \succ', z \ \succ'')$ is drawn according to $IID(\phi_2)$. If the inequalities

$$x \leq \overline{J} \tag{G.15}$$

$$y \leq K_i \tag{G.16}$$

$$x + y > K_i + J(K_i) \tag{G.17}$$

are satisfied, then we again must have $P \in R$ (since $x > J(K_i) \geq \underline{J}$), and then $f(P) = A_i$, in consequence of $f(P_{J(K_i)+1,K_i}) = A_i$ and the monotonicity condition (iii).

Notice also that if

$$x \geq (J(K_{i-1}) + 3J(K_i))/4 \tag{G.18}$$

$$y > K_i - 4\sqrt{\widetilde{N}} \tag{G.19}$$

are satisfied, then (G.17) will automatically hold.

If $\delta_2 = (K_i - \sqrt{2\widetilde{N}})/\widetilde{N}$, then exactly the same Chebyshev arguments as before give that (G.15), (G.16), (G.18), (G.19) are all satisfied — and hence $f(P) = A_i$ — with probability greater than 3/4. Otherwise, we necessarily have $x + y = \widetilde{N}$ so that (G.17) is always satisfied (recall (G.7)), and then the same arguments show that (G.15), (G.16) are both satisfied with probability greater than 3/4. In either case, then, we get (G.9).

Now that (G.8) and (G.9) are proven, we use Lemma 4.9 to complete the argument. Notice that $\phi_2 - \phi_1 = \Delta(\succ' - \succ'')$, where

$$0 \leq \Delta < \frac{K_i}{\widetilde{N}} - \frac{K_{i-1}}{\widetilde{N}} < \frac{40}{\kappa\sqrt{\widetilde{N}}}.$$

By (ii), preferences $\succ'$ and $\succ''$ rank A_i and A_j in the same way. If they both rank A_i above A_j, then let $\mathcal{C}^+$ be the set of candidates weakly preferred to A_i under $\succ''$. Lemma 4.9(a) gives

$$\sum_{A \in \mathcal{C}^+} \overline{f}_A(\phi_2) - \sum_{A \in \mathcal{C}^+} \overline{f}_A(\phi_1) \leq c_0 \widetilde{N} \Delta \sigma, \tag{G.20}$$

where c_0 is the constant promised by that lemma. By (G.8) and (G.9), the left-hand side of (G.20) is at least $3/4 - (1 - 3/4) = 1/2$, so

$$\frac{1}{2} \leq c_0 \widetilde{N} \Delta \sigma \leq c_0 \sqrt{\widetilde{N}} \frac{40}{\kappa} \sigma.$$

If $\succ'$ and $\succ''$ both rank A_j above A_i, then let $\mathcal{C}^+$ be the set of candidates weakly preferred to A_j under $\succ'$. Lemma 4.9(a) gives

$$\sum_{A\in\mathcal{C}^+} \overline{f}_A(\phi_1) - \sum_{A\in\mathcal{C}^+} \overline{f}_A(\phi_2) \le c_0 \widetilde{N}\Delta\sigma$$

and we again arrive at $1/2 \le c_0\sqrt{\widetilde{N}}(40/\kappa)\sigma$. Thus in either case we have

$$\sigma \ge \frac{\kappa}{80c_0}\widetilde{N}^{-1/2}$$

which is the promised result. $\square$

Proof of Theorem 4.4: We first suppose there are just three candidates, $\mathcal{C} = \{A, B, C\}$. For every K, we have $f(K\ ABC, \widetilde{N} - K\ BCA) \in \{A, B\}$ by Pareto efficiency (and this value is B when $K = 0$ and A when $K = \widetilde{N}$). Moreover, by monotonicity, if this expression equals A for some K then it also equals A for all higher K. So, writing

$$K_{AB} = \max\left\{K \,\middle|\, f\begin{pmatrix} K & ABC \\ \widetilde{N}-K & BCA\end{pmatrix} = B\right\},$$

we have $f(K\ ABC, \widetilde{N} - K\ BCA) = B$ if $K \le K_{AB}$ and A if $K > K_{AB}$. Likewise define

$$K_{BC} = \max\left\{K \,\middle|\, f\begin{pmatrix} K & BCA \\ \widetilde{N}-K & CAB\end{pmatrix} = C\right\},$$

$$K_{CA} = \max\left\{K \,\middle|\, f\begin{pmatrix} K & CAB \\ \widetilde{N}-K & ABC\end{pmatrix} = A\right\},$$

$$K_{CB} = \max\left\{K \mid \,\middle|\, f\begin{pmatrix} K & CBA \\ \widetilde{N}-K & BAC\end{pmatrix} = B\right\},$$

$$K_{BA} = \max\left\{K \,\middle|\, f\begin{pmatrix} K & BAC \\ \widetilde{N}-K & ACB\end{pmatrix} = A\right\},$$

$$K_{AC} = \max\left\{K \,\middle|\, f\begin{pmatrix} K & ACB \\ \widetilde{N}-K & CBA\end{pmatrix} = C\right\}.$$

We now have two cases.

(i) $K_{AB} + K_{BC} + K_{CA} + K_{CB} + K_{BA} + K_{AC} > 7\widetilde{N}/2$.

In this case one of the three quantities $K_{AB} + K_{BA}, K_{BC} + K_{CB}, K_{CA} + K_{AC}$ is greater than $7\widetilde{N}/6$. Without loss of generality we will assume $K_{CA}+K_{AC} > 7\widetilde{N}/6$, which is the case shown in Figure 4.4.

Let
$$K^* = \max\left\{ K \;\middle|\; f\begin{pmatrix} K & CAB \\ \widetilde{N}-K & ACB \end{pmatrix} = A \right\}.$$

Similarly to before, $f(K\ CAB, \widetilde{N}-K\ ACB) = A$ for $K \le K^*$ and $= C$ for $K > K^*$. Now one of the following two inequalities must hold:

$$K_{CA} - K^* > \frac{\widetilde{N}}{12}; \qquad K_{AC} - (N - K^*) > \frac{\widetilde{N}}{12}.$$

We assume henceforth that the first inequality holds (otherwise, the argument is the same with A and C reversed).

Now we apply Lemma 4.10 with

$$\succ = CAB, \qquad \succ' = ACB, \qquad \succ'' = ABC,$$

$$\underline{J} = K^*+1, \qquad \overline{J} = K_{CA}, \qquad \kappa = \frac{1}{13}, \qquad \overline{K} = \widetilde{N} - (K^*+1),$$
$$A_i = C, \qquad A_j = A.$$

The condition $\overline{J} - \underline{J} > \kappa\widetilde{N}$ is evidently satsified (as long as N is large), so we need to verify conditions (i)-(v) of the lemma. (i) follows from Pareto efficiency. (ii) is immediate. (iii) follows from monotonicity. (iv) comes from the definition of K_{CA} (and our monotonicity observation earlier). (v) comes from the definition of K^*. Hence, the lemma applies, and σ is bounded below by a constant times $N^{-1/2}$. This takes care of case (i).

(ii) $K_{AB} + K_{BC} + K_{CA} + K_{CB} + K_{BA} + K_{AC} \le 7\widetilde{N}/2$.

In this case one of the quantities $K_{AB}+K_{BC}+K_{CA}$, $K_{CB}+K_{BA}+K_{AC}$ is at most $7\widetilde{N}/4$. Without loss of generality we will assume

$$K_{AB} + K_{BC} + K_{CA} \le \frac{7\widetilde{N}}{4}. \tag{G.21}$$

We can now focus our attention on the $ABC - BCA - CAB$ simplex.

We will also assume for now the inequalities

$$K_{AB} + K_{BC}, K_{BC} + K_{CA}, K_{CA} + K_{AB} \geq \frac{89}{90}\widetilde{N}. \tag{G.22}$$

Afterwards we will come back to address the (easier) case where one of these inequalities is violated.

In terms of Figure 4.4, we will show that one of the boundaries between regions within the $ABC - BCA - CAB$ simplex has a portion sufficiently sloped so that we can apply Lemma 4.10. However, to locate such a portion of a boundary we will need some more detailed case analysis. An outline of the argument is illustrated in Figure G.1. In the top-left panel, the dots marked on the edges of the simplex are the profiles $(K_{AB}\ ABC, \widetilde{N} - K_{AB}\ BCA)$, $(K_{BC}\ BCA, \widetilde{N} - K_{BC}\ CAB)$, and $(K_{CA}\ CAB, \widetilde{N} - K_{CA}\ ABC)$. The assumption $K_{AB} + K_{BC} + K_{CA} \leq 7\widetilde{N}/4$ ensures that the downward-pointing triangle in the middle of the simplex has side length at least $\widetilde{N}/4$. Consider the profile at the center of the triangle, and without loss of generality assume that the winner there is A. Then consider the smaller downward-pointing triangle (shown in the bottom two panels). Using monotonicity we can show that at each profile in the smaller triangle, f must choose either A or B. If f chooses A at the center of the smaller triangle, then consider the shaded trapezoid in the bottom-left panel of Figure G.1. By monotonicity arguments, f chooses either A or C at each profile in the trapezoid, and chooses A near the left edge and C at the right edge. Then, this trapezoid gives a region where the $A - C$ boundary is nontrivially sloped, and we can apply Lemma 4.10. If instead f chooses B at the center of the smaller triangle, then we consider the parallelogram shown in the bottom-right panel, and similarly apply Lemma 4.10.

Now we begin the proof properly. Let P_0 be a profile with

$$P_0 = \begin{pmatrix} x_0\ ABC \\ y_0\ BCA \\ z_0\ CAB \end{pmatrix} \approx \begin{pmatrix} (\widetilde{N} + K_{AB} + K_{BC} - 2K_{CA})/3\ ABC \\ (\widetilde{N} - 2K_{AB} + K_{BC} + K_{CA})/3\ BCA \\ (\widetilde{N} + K_{AB} - 2K_{BC} + K_{CA})/3\ CAB \end{pmatrix},$$

where the approximation means that we add or subtract at most 1 to each component to ensure x_0, y_0, z_0 are integers. Inequality (G.21), together with (G.22), ensure that x_0, y_0, z_0 are all positive. We have $f(P_0) = A$, B, or C. Without loss of generality, suppose henceforth that $f(P_0) = A$.

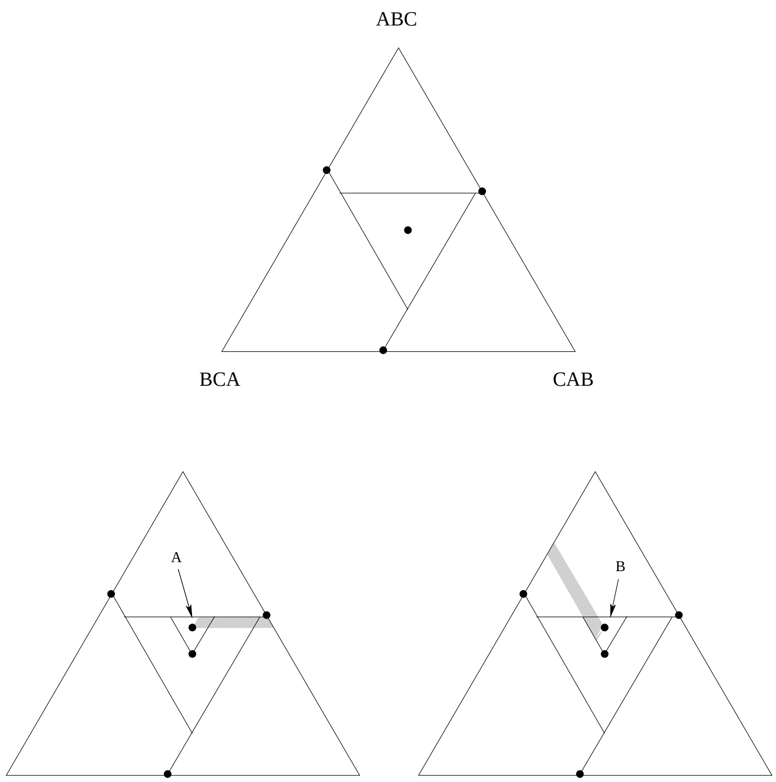

Figure G.1: Proof of Theorem 4.4 (case (ii))

Now take

$$T = \left\{ \begin{pmatrix} \widetilde{N} - s - t & ABC \\ s & BCA \\ t & CAB \end{pmatrix} \middle| \, s \le y_0, t \le z_0, s + t \ge K_{CA} \right\}.$$

If P is any profile in T with $t = z_0$, then P can be obtained from P_0 by changing some BCA votes to ABC, so by monotonicity $f(P) = A$ for each such P. Consequently, we cannot have $f(P) = C$ for any $P \in T$: if $f(\widetilde{N} - s - t \; ABC, s \; BCA, t \; CAB) = C$, then by monotonicity $f(\widetilde{N} - s - z_0 \; ABC, s \; BCA, z_0 \; CAB) = C$, but this profile is also in T and we just saw that A must win there, a contradiction. Hence, $f(P) \in \{A, B\}$ for all $P \in T$.

Let P_1 be a profile with

$$P_1 = \begin{pmatrix} x_1 \; ABC \\ y_1 \; BCA \\ z_1 \; CAB \end{pmatrix} \approx \begin{pmatrix} (7\widetilde{N} + K_{AB} + K_{BC} - 8K_{CA})/9 \; ABC \\ (\widetilde{N} - 5K_{AB} + 4K_{BC} + 4K_{CA})/9 \; BCA \\ (\widetilde{N} + 4K_{AB} - 5K_{BC} + 4K_{CA})/9 \; CAB \end{pmatrix}.$$

This profile is the "center of the smaller triangle" in Figure G.1. Again, one can verify that all components are positive. Moreover, $P_1 \in T$: all of the relevant inequalities reduce (up to rounding error which is bounded by a constant) to $K_{AB} + K_{BC} + K_{CA} \le 2\widetilde{N}$, which is true by (G.21). Therefore, $f(P_1) \in \{A, B\}$. We have two subcases.

– If $f(P_1) = A$, then we will apply Lemma 4.10 with

$$\succ = ABC, \qquad \succ' = BCA, \qquad \succ'' = CAB,$$

$$\underline{J} = x_1, \qquad \overline{J} = \widetilde{N} - K_{CA} - 4, \qquad \kappa = \frac{1}{40}, \qquad \overline{K} = y_1 - (\widetilde{N} - K_{CA} - 4 - x_1),$$

$$A_i = A, \qquad A_j = C.$$

The required inequality $\overline{J} - \underline{J} > \kappa\widetilde{N}$ follows directly from (G.21). We proceed to verify conditions (i)-(v) of the lemma.

To verify condition (i), suppose for contradiction that $f(J \; ABC, K \; BCA, \widetilde{N} - J - K \; CAB) = B$ for some $\underline{J} \le J \le \overline{J}$ and $0 \le K \le \overline{K}$. By monotonicity, $f(J \; ABC, \widetilde{N} - J - z_1 \; BCA, z_1 \; CAB) = B$ also. (Note that since $\widetilde{N} - \overline{J} - z_1 = \overline{K}$, then indeed $\widetilde{N} - J - z_1 \ge K \ge 0$.) But since $J \ge x_1$, $f(P_1) = A$ and monotonicity imply $f(J\ ABC, \widetilde{N}-J-z_1\ BCA, z_1\ CAB) = A$, a contradiction. Thus condition (i) of Lemma 4.10 holds.

Condition (ii) is immediate. (iii) follows from monotonicity given (i): if $f(P_{J,K}) = A$ then $f(P_{J+1,K-1}) = A$ by monotonicity, and $f(P_{J+1,K})$ cannot equal C because then monotonicity would require $f(P_{J,K}) = C$, so $f(P_{J+1,K}) = A$ instead. (iv) follows from the definition of K_{CA}. And (v) holds because each of the relevant profiles $P_{J,\overline{K}}$ lies in the set T (checking the relevant linear inequalities is tedious but straightforward), hence $f(P_{J,\overline{K}}) = A$ or B; since we have already ruled out B with condition (i), we must have $f(P_{J,\overline{K}}) = A$ for each J, and condition (v) is satisfied. This checks all the conditions to apply Lemma 4.10, and we conclude that σ is bounded below by a constant times $N^{-1/2}$.

– If on the other hand $f(P_1) = B$, then we will apply Lemma 4.10 with

$$\succ = BCA, \qquad \succ' = CAB, \qquad \succ'' = ABC,$$

$$\underline{J} = y_1, \qquad \overline{J} = y_0 - 4, \qquad \kappa = \frac{1}{40}, \qquad \overline{K} = z_1,$$

$$A_i = B, \qquad A_j = A.$$

Again, the requirement $\overline{J} - \underline{J} > \kappa\widetilde{N}$ follows from (G.21), so we proceed to verify conditions (i)-(v) of the lemma.

If $f(J\ BCA,\ K\ CAB,\ \widetilde{N} - J - K\ ABC) = C$ for some (J, K), then by monotonicity we also have $f(J\ BCA, z_1\ CAB, \widetilde{N} - J - z_1\ ABC) = C$. ((G.21) and (G.22) ensure this is a valid profile.) But this profile lies in T, so we should have $f(J\ BCA, z_1\ CAB, \widetilde{N} - J - z_1\ ABC) \in \{A, B\}$, a contradiction. This shows that condition (i) is satisfied. Condition (ii) is immediate. (iii) follows from monotonicity given (i): if $f(P_{J,K}) = B$, then $f(P_{J+1,K-1}) = B$ by monotonicity, and we cannot have $f(P_{J+1,K}) = A$ since then monotonicity would require $f(P_{J,K}) = A$ as well, so we must have $f(P_{J+1,K}) = B$. (iv) follows from $\widetilde{N} - \overline{J} > K_{AB}$ (which in turn follows from (G.21)). Finally, $f(P_{\underline{J},\overline{K}}) = f(P_1) = B$, so $f(P_{J,\overline{K}}) = B$ for all J (by condition (iii)), verifying (v). So we have checked all the conditions, and Lemma 4.10 applies. We again conclude that σ is bounded below by a constant times $N^{-1/2}$.

This completes the proof of case (ii) of the theorem as long as we have the maintained assumption (G.22). It remains to address what happens when (G.22) is violated.

Without loss of generality, we assume that

$$K_{CA} + K_{AB} < \frac{89}{90}\widetilde{N}.$$

Then we can apply Lemma 4.10 with

$$\succ = ABC, \qquad \succ' = BCA, \qquad \succ'' = CAB,$$

$$\underline{J} = K_{AB} + 1, \qquad \overline{J} = \widetilde{N} - K_{CA} - 1, \qquad \kappa = \frac{1}{100}, \qquad \overline{K} = \widetilde{N} - K_{AB} - 1,$$
$$A_i = A, \qquad A_j = C.$$

It is clear that $\overline{J} - \underline{J} > \kappa\widetilde{N}$ as long as N is large, so we check (i)-(v) of the lemma. For (i), suppose $f(J\ ABC, K\ BCA, \widetilde{N} - J - K\ CAB) = B$ for some J, K with $J \geq K_{AB} + 1$. By monotonicity, $f(J\ ABC, \widetilde{N} - J\ BCA) = B$ also. This contradicts $J > K_{AB}$. Then (i) follows. (ii) is immediate. (iii) holds by monotonicity: if $f(P_{J,K}) = A$, then $f(P_{J+1,K-1}) = A$ by monotonicity directly; and $f(P_{J+1,K}) = C$ would imply $f(P_{J,K}) = C$ by monotonicity, a contradiction, so from (i) we must have $f(P_{J+1,K}) = A$ instead. (iv) follows from the definition of K_{CA}, and (v) follows from the definition of K_{AB}. Thus all the conditions hold and once again Lemma 4.10 assures us that σ is bounded below by a constant times $N^{-1/2}$.

This completes the analysis of cases (i) and (ii). We have had to apply Lemma 4.10 with only finitely many values of κ, so if we simply let c be the smallest of the corresponding values of $c(\kappa)$, then we have $\sigma \geq cN^{-1/2}$ in every subcase (as always, assuming N is sufficiently large).

Finally, the foregoing analysis assumed that $\mathcal{C}$ consisted of just three candidates. If there are more than three candidates, then let A, B, C be any three of them, and restrict attention to profiles (and beliefs) at which each voter ranks A, B, C higher than any other candidate, with the remaining candidates all ranked according to some fixed order. By Pareto efficiency, only A, B, or C can win at any such profile. Then, all of the preceding analysis carries through directly, with the preferences ABC replaced by $ABC\ldots$, BCA replaced by $BCA\ldots$, and so forth.

□

Next, we round out Subsection 4.5 by supplying the proof of Theorem 4.2, for any unanimous, tops-only voting rule.

Proof of Theorem 4.2: Let c_1 be the constant given by Lemma C.8. Take A, B, C to be any three different candidates. We consider two possibilities.

(i) Suppose there exists some K such that $f(K\ B, N+1-K\ C) \notin \{B, C\}$. Let $S \subseteq \{1, \ldots, N\}$ be the set of all such values. Let α be the value given by Lemma C.8 for the set S. Put $\phi = (\alpha\ B, 1-\alpha\ C)$. The conclusion of the lemma can be written as

$$\Pr_{IID(\phi)}(f(C,P) \notin \{B,C\}) - \Pr_{IID(\phi)}(f(B,P) \notin \{B,C\}) \geq \frac{c_1}{N},$$

where the probabilities are over opponent-profiles P drawn according to $IID(\phi)$; or equivalently,

$$\Pr_{IID(\phi)}(f(B,P) \in \{B,C\}) - \Pr_{IID(\phi)}(f(C,P) \in \{B,C\}) \geq \frac{c_1}{N}. \tag{G.23}$$

If the manipulator's true preference ranks C first and B second, then consider the manipulation to reporting C, with the set of preferred candidates $\mathcal{C}^+ = \{B, C\}$. The left side of (G.23) is $\leq \sigma$, by (2.3). So we get $\sigma \geq c_1/N$ in this case.

(ii) Suppose that $f(K\ B, N+1-K\ C) \in \{B, C\}$ for all K. Assume that $\sigma < 1/(N+1)$ (otherwise we are done). We will first show the following

Claim. There exists exactly one value of K such that $f(1\ A, K\ B, N-K\ C) \notin \{B, C\}$.

For any $\alpha \in [0,1]$, consider (2.3) for a manipulator with true preference $B \ldots C$, considering a manipulation to A, with belief $\phi = (\alpha\ B, 1-\alpha\ C)$ and $\mathcal{C}^+ = \mathcal{C} \setminus \{C\}$. We get

$$\sum_{K=0}^{N} \mathbf{P}\left(\begin{array}{c} K \\ N-K \end{array} \middle| N; \begin{array}{c} \alpha \\ 1-\alpha \end{array} \right) \times \left[\mathbf{I}\left(f\left(\begin{array}{cc} 1 & A \\ K & B \\ N-K & C \end{array} \right) \in \mathcal{C}^+ \right) - \mathbf{I}\left(f\left(\begin{array}{cc} K+1 & B \\ N-K & C \end{array} \right) \in \mathcal{C}^+ \right) \right]$$

$$\leq \sigma < \frac{1}{N+1}.$$

Now integrate over α from 0 to 1, using the well-known identity $\int_0^1 \binom{N}{K}\alpha^K(1-\alpha)^{N-K}\,d\alpha = 1/(N+1)$.[1] This gives

$$\sum_{K=0}^{N} \frac{1}{N+1}\left[\mathbf{I}\left(f\begin{pmatrix} 1 & A \\ K & B \\ N-K & C \end{pmatrix} \in \mathcal{C}^+\right) - \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} \in \mathcal{C}^+\right)\right] < \frac{1}{N+1}.$$

Applying (2.3) for a manipulator with true preference $A \dots C$, considering a manipulation to B, gives the same inequality with the left-hand side negated. Hence, after multiplying through by $N+1$, we get

$$\left|\sum_{K=0}^{N} \mathbf{I}\left(f\begin{pmatrix} 1 & A \\ K & B \\ N-K & C \end{pmatrix} \neq C\right) - \sum_{K=0}^{N} \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} \neq C\right)\right| < 1.$$

The left side is an integer, so it must be zero.

After subtracting both of the sums from $N+1$, we get the simpler equation

$$\sum_{K=0}^{N} \mathbf{I}\left(f\begin{pmatrix} 1 & A \\ K & B \\ N-K & C \end{pmatrix} = C\right) = \sum_{K=0}^{N} \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} = C\right).$$

Moreover, the upper bound of summation on the right side can be replaced by $N-1$, since we know $f(N+1\ C) = C$ by unanimity.

Now, we can repeat the same argument with B and C reversed, giving

$$\sum_{K=0}^{N} \mathbf{I}\left(f\begin{pmatrix} 1 & A \\ K & B \\ N-K & C \end{pmatrix} = B\right) = \sum_{K=0}^{N-1} \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} = B\right).$$

Adding these two equations gives

$$\sum_{K=0}^{N} \mathbf{I}\left(f\begin{pmatrix} 1 & A \\ K & B \\ N-K & C \end{pmatrix} \in \{B, C\}\right) = \sum_{K=0}^{N-1} \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} \in \{B, C\}\right).$$

[1] The identity can be proven by showing that the integral is equal at two successive values of K, since the difference between the integrals at K and $K+1$ is $\frac{1}{N+1}\binom{N+1}{K+1}\alpha^{K+1}(1-\alpha)^{N-K}|_0^1 = 0$.

Since the right side is N by assumption, we see there are exactly N profiles of the form $(1\ A, K\ B, N-K\ C)$ at which either B or C wins. This proves the claim.

Let K^* be the unique value for which $f(1\ A, K^*\ B, N-K^*\ C) \notin \{B, C\}$.

Now let K_B be the minimum value such that $f(K_B\ B, N+1-K_B\ C) = B$. Let K_C be the maximum value such that $f(K_C\ B, N+1-K_C\ C) = C$. Our next step is to show that K_B and K_C are both close to K^*, and therefore close to each other. From there, we will be able to repeat the argument from the proof of Theorem 4.7.

Let $S = \{K \mid f(K\ B, N+1-K\ C) = C\}$. Let $\alpha_C \geq K_C/N$ be the value given by Lemma C.8 for this set. Assume that $\sigma < c_1/3N$ (otherwise we are done).

Consider a manipulator with belief $\phi_C = (\alpha_C\ B, 1-\alpha_C\ C)$, preference $A \ldots B$, manipulating to C. We will write $\mathbf{P}[K]$ rather than $\mathbf{P}(K, N-K \mid N; \phi_C)$ to save on notation. The manipulation cannot decrease the probability of B by more than σ, hence

$$\sum_{K=0}^{N} \mathbf{P}[K] \left[\mathbf{I}\left(f\begin{pmatrix} 1 & A \\ K & B \\ N-K & C \end{pmatrix} = B \right) - \mathbf{I}\left(f\begin{pmatrix} K & B \\ N+1-K & C \end{pmatrix} = B \right) \right] \leq \frac{c_1}{3N}. \tag{G.24}$$

Similarly, a manipulator with the same belief and preference $A \ldots C$, manipulating to B, cannot decrease the probability of C by more than σ, hence

$$\sum_{K=0}^{N} \mathbf{P}[K] \left[\mathbf{I}\left(f\begin{pmatrix} 1 & A \\ K & B \\ N-K & C \end{pmatrix} = C \right) - \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} \right) = C \right] \leq \frac{c_1}{3N}. \tag{G.25}$$

Now add (G.24) and (G.25). Notice that the $f(1\ A, K\ B, N-K\ C)$ terms add up to cover each possible value of K exactly once, except for $K = K^*$. Thus we get

$$\begin{aligned} &1 - \mathbf{P}[K^*] - \\ &\sum_{K=0}^{N} \mathbf{P}[K] \times \left[\mathbf{I}\left(f\begin{pmatrix} K & B \\ N+1-K & C \end{pmatrix} = B \right) + \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} = C \right) \right] \\ &\leq \frac{2c_1}{3N}. \end{aligned} \tag{G.26}$$

But the summation on the left side comes under control because

$$\begin{aligned}
\sum_{K=0}^{N} \mathbf{P}[K] & \left[\mathbf{I}\left(f\begin{pmatrix} K & B \\ N+1-K & C \end{pmatrix} = B \right) + \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} = C \right) \right] \\
&= \sum_{K=0}^{N} \mathbf{P}[K] \left[1 - \mathbf{I}\left(f\begin{pmatrix} K & B \\ N+1-K & C \end{pmatrix} = C \right) + \right. \\
&\qquad\qquad \left. \mathbf{I}\left(f\begin{pmatrix} K+1 & B \\ N-K & C \end{pmatrix} = C \right) \right] \\
&= 1 + \sum_{K=0}^{N} \left[\mathbf{P}[K-1] - \mathbf{P}[K] \right] \mathbf{I}\left(f\begin{pmatrix} K & B \\ N+1-K & C \end{pmatrix} = C \right) \\
&\leq 1 - \frac{c_1}{N}
\end{aligned}$$

where the second equality comes from reindexing the sum, and the final inequality comes from Lemma C.8.

Combining with (G.26) gives

$$1 - \mathbf{P}[K^*] - \left[1 - \frac{c_1}{N}\right] \leq \frac{2c_1}{3N}$$

or, finally,

$$\mathbf{P}\left(\begin{matrix} K^* \\ N-K^* \end{matrix} \;\middle|\; N;\; \begin{matrix} \alpha_C \\ 1-\alpha_C \end{matrix} \right) \geq \frac{c_1}{3N}. \tag{G.27}$$

Now combining (G.27) with Lemma C.4 gives

$$\frac{c_1}{3N} \leq e^{-N\frac{(\alpha_C - K^*/N)^2}{2}}$$

from which

$$\left| \alpha_C - \frac{K^*}{N} \right| \leq \sqrt{\frac{2(\ln N - \ln(c_1/3))}{N}}.$$

As long as N is sufficiently large, the right-hand side is $\leq N^{-1/3}$. So we can conclude

$$K_C \leq \alpha_C N \leq K^* + N^{2/3}.$$

Now, exactly the same argument with the roles of B and C reversed leads to the conclusion that

$$K_B \geq K^* - N^{2/3}.$$

Therefore, we have

$$K_C - K_B \leq 2N^{2/3}. \tag{G.28}$$

This is the assertion that K_B and K_C are close to each other, as promised. (Notice also from the definitions that $K_B \leq K_C + 1$.)

From here, we will continue to assume that f has susceptibility $\sigma < 1/\widetilde{N}$ and obtain a contradiction, following the same steps as for Theorem 4.7. As long as N is large enough, we may assume that $K_B \leq 2\widetilde{N}/3$ (otherwise $K_C \geq \widetilde{N}/3$, so just switch B and C). Let

$$\phi_1 = (\alpha_1\ B, 1 - \alpha_1\ C) \qquad \text{with} \qquad \alpha_1 = \min\left\{\frac{K_C + \sqrt{2\widetilde{N}}}{\widetilde{N}}, 1\right\}.$$

Whenever more than K_C voters vote for B and the rest vote for C, B wins; so the same Chebyshev argument as in the proof of Theorem 4.7 gives $\overline{f}(\phi_1) = (\gamma_1\ B, 1 - \gamma_1\ C)$ where $\gamma_1 \geq 7/8$. Let

$$\phi_2 = (\alpha_2\ B, 1 - \alpha_2\ C) \qquad \text{with} \qquad \alpha_2 = \max\left\{\frac{K_B - \sqrt{2\widetilde{N}}}{\widetilde{N}}, 0\right\},$$

and obtain $\overline{f}(\phi_2) = (\gamma_2\ B, 1 - \gamma_2\ C)$ where $\gamma_2 \leq 1/8$.

Write $\phi_1 - \phi_2 = \Delta(B - C)$, with $\Delta \geq 0$. On account of (G.28), we have

$$\Delta = \alpha_1 - \alpha_2 \leq 2\sqrt{2}\widetilde{N}^{-1/2} + 2\widetilde{N}^{-1/3} \leq 3\widetilde{N}^{-1/3}$$

as long as N is large. Again taking c_0 to be the constant from Lemma 4.9, we have

$$c_0\widetilde{N}\Delta\sigma < 3c_0\widetilde{N}^{-1/3} < \frac{1}{8}$$

as long as N is large. Exactly as for Theorem 4.7, we now define $\phi_3 = \phi_1 + \Delta(A - B)$ and $\phi_4 = \phi_1 + \Delta(A - C)$. We check that these are valid probability distributions as long as $1 - \alpha_1 > \Delta$; and indeed, we have

$$\alpha_1 + \Delta \leq \frac{K_C + \sqrt{2\widetilde{N}}}{\widetilde{N}} + 3\widetilde{N}^{-1/3} \leq \frac{K_B + 2N^{2/3} + \sqrt{2\widetilde{N}}}{\widetilde{N}} + 3\widetilde{N}^{-1/3} < 1$$

as long as N is sufficiently large (using (G.28) and $K_B \leq 2\widetilde{N}/3$). We can then apply Lemma 4.9 to each of the pairs connected by thick lines in Figure 4.5, obtaining constraints on the values of $\overline{f}(\phi_3)$ and $\overline{f}(\phi_4)$ until we reach a contradiction.

□

Finally, we give proofs of the ingredients for Theorem 4.1, covering any weakly unanimous voting rule. We begin with Lemma 4.11.

Proof of Lemma 4.11: Let v denote the four-way difference on the left-hand side of (4.7).

Put

$$w_1 = \overline{f}\begin{pmatrix} \alpha & \succ_1 \\ \beta & \succ_3 \\ \gamma & \phi \end{pmatrix} - \overline{f}\begin{pmatrix} \alpha & \succ_2 \\ \beta & \succ_3 \\ \gamma & \phi \end{pmatrix},$$

$$w_2 = \overline{f}\begin{pmatrix} \alpha & \succ_1 \\ \beta & \succ_4 \\ \gamma & \phi \end{pmatrix} - \overline{f}\begin{pmatrix} \alpha & \succ_2 \\ \beta & \succ_4 \\ \gamma & \phi \end{pmatrix}.$$

Apply Lemma 4.9(b) twice to the difference represented by w_1: once letting $\mathcal{C}'$ be the set of candidates $A \neq A_i, A_j$ such that $(w_1)_A \geq 0$, and once letting $\mathcal{C}'$ be the set of candidates $A \neq A_i, A_j$ such that $(w_1)_A < 0$. We obtain

$$\sum_{A \neq A_i, A_j} |(w_1)_A| \leq 2c_0\widetilde{N}\alpha\sigma \leq 2c_0\widetilde{N}\sigma.$$

Likewise,

$$\sum_{A \neq A_i, A_j} |(w_2)_A| \leq 2c_0\widetilde{N}\sigma.$$

Then, since $v = w_1 - w_2$, we get

$$\sum_{A \neq A_i, A_j} |v_A| \leq 4c_0\widetilde{N}\sigma. \tag{G.29}$$

Now put

$$w_3 = \overline{f}\begin{pmatrix} \alpha & \succ_1 \\ \beta & \succ_3 \\ \gamma & \phi \end{pmatrix} - \overline{f}\begin{pmatrix} \alpha & \succ_1 \\ \beta & \succ_4 \\ \gamma & \phi \end{pmatrix},$$

$$w_4 = \overline{f}\begin{pmatrix} \alpha & \succ_2 \\ \beta & \succ_3 \\ \gamma & \phi \end{pmatrix} - \overline{f}\begin{pmatrix} \alpha & \succ_2 \\ \beta & \succ_4 \\ \gamma & \phi \end{pmatrix}.$$

Using $v = w_3 - w_4$, analogous computations give

$$\sum_{A \neq A_k, A_l} |v_A| \leq 4c_0\widetilde{N}\sigma. \tag{G.30}$$

Now if $\{A_i, A_j\}$ is disjoint from $\{A_k, A_l\}$, then (G.29) and (G.30) immediately lead us to $\sum_{A\in\mathcal{C}} |v_A| \leq 8c_0\widetilde{N}\sigma$ which is stronger than (4.7). Otherwise, $\{A_i, A_j\}$ and $\{A_k, A_l\}$ have one element in common — say A_i — in which case (G.29) and (G.30) give $\sum_{A\neq A_i} |v_A| \leq 8c_0\widetilde{N}\sigma$. Since the sum of the components of v is zero, we also have $|v_{A_i}| \leq 8c_0\widetilde{N}\sigma$, and (4.7) follows. □

We now prove the three main lemmas that combine to give the theorem.

Proof of Lemma 4.12: Suppose the conclusion does not hold. Then the same reasoning as in case (i) of Theorem 4.2 — applying Lemma C.8 to the set of all K such that $f(K\ CAB, \widetilde{N} - K\ CBA) \neq C$ — gives a distribution ϕ such that

$$\Pr_{IID(\phi)}(f(CAB, P) = C) - \Pr_{IID(\phi)}(f(CBA, P) = C) \geq \frac{c_1}{N}.$$

If we consider a manipulator with true preference CBA, manipulating to CAB, with the set of preferred candidates $\mathcal{C}^+ = \{C\}$, then this gives us $\sigma \geq c_1/N$, contradicting the given. □

Proof of Lemma 4.13: Define the following vectors in $\mathbb{R}^M$:

$$v_1 \quad = \quad \overline{f}\begin{pmatrix} x & ABC\ldots \\ y & BAC\ldots \\ z & BAC\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x & ABC\ldots \\ y & ABC\ldots \\ z & BAC\ldots \end{pmatrix},$$

(we write e.g. $(x\ ABC\ldots, y\ BAC\ldots, z\ BAC\ldots)$ rather than $(x\ ABC\ldots, y{+}z\ BAC\ldots)$ to aid readability; no confusion should result)

$$v_2 \quad = \quad \overline{f}\begin{pmatrix} x & ACB\ldots \\ y & BAC\ldots \\ z & BAC\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x & ACB\ldots \\ y & ABC\ldots \\ z & BAC\ldots \end{pmatrix},$$

$$v_3 \quad = \quad \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & BAC\ldots \\ z & BAC\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & ABC\ldots \\ z & BAC\ldots \end{pmatrix},$$

$$v_4 \quad = \quad \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & BAC\ldots \\ z & BCA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & ABC\ldots \\ z & BCA\ldots \end{pmatrix},$$

$$v_5 \quad = \quad \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & BAC\ldots \\ z & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & ABC\ldots \\ z & CBA\ldots \end{pmatrix}$$

By applying Lemma 4.11 repeatedly, we get

$$|v_1 - v_2| \leq 16c_0\widetilde{N}\sigma; \qquad |v_2 - v_3| \leq 16c_0\widetilde{N}\sigma; \qquad |v_3 - v_4| \leq 16c_0\widetilde{N}\sigma; \qquad |v_4 - v_5| \leq 16c_0\widetilde{N}\sigma.$$

Adding these and using the triangle inequality gives

$$|v_1 - v_5| \leq 64c_0\widetilde{N}\sigma.$$

Next, define

$$v_1' \quad = \quad \overline{f}\begin{pmatrix} x' & ABC\ldots \\ y & BAC\ldots \\ z' & BAC\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & ABC\ldots \\ y & ABC\ldots \\ z' & BAC\ldots \end{pmatrix},$$

$$v_5' \quad = \quad \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & BAC\ldots \\ z' & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & ABC\ldots \\ z' & CBA\ldots \end{pmatrix}.$$

Then the above reasoning also gives

$$|v_1' - v_5'| \leq 64c_0\widetilde{N}\sigma,$$

and hence we obtain

$$|(v_1 - v_1') - (v_5 - v_5')| \leq 128c_0\widetilde{N}\sigma. \tag{G.31}$$

Now define

$$w_1 = \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & BAC\ldots \\ z & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & BAC\ldots \\ z' & CBA\ldots \end{pmatrix},$$

$$w_2 = \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & BCA\ldots \\ z & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & BCA\ldots \\ z' & CBA\ldots \end{pmatrix},$$

$$w_3 = \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & CBA\ldots \\ z & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & CBA\ldots \\ z' & CBA\ldots \end{pmatrix}.$$

Then Lemma 4.11 gives

$$|w_1 - w_2| \leq 16c_0\widetilde{N}\sigma; \qquad |w_2 - w_3| \leq 16c_0\widetilde{N}\sigma,$$

so by the triangle inequality,

$$|w_1 - w_3| \leq 32c_0\widetilde{N}\sigma.$$

However, $w_3 = 0$, because our assumption (4.8) implies that both $\overline{f}(\cdots)$ values in the definition of w_3 are just C with probability 1. Thus we actually have

$$|w_1| \leq 32c_0\widetilde{N}\sigma. \tag{G.32}$$

Similarly define

$$w_4 = \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & ABC\ldots \\ z & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & ABC\ldots \\ z' & CBA\ldots \end{pmatrix},$$

$$w_5 = \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & ACB\ldots \\ z & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & ACB\ldots \\ z' & CBA\ldots \end{pmatrix},$$

$$w_6 = \overline{f}\begin{pmatrix} x & CAB\ldots \\ y & CAB\ldots \\ z & CBA\ldots \end{pmatrix} - \overline{f}\begin{pmatrix} x' & CAB\ldots \\ y & CAB\ldots \\ z' & CBA\ldots \end{pmatrix}.$$

Lemma 4.11 gives

$$|w_4 - w_5| \leq 16c_0\widetilde{N}\sigma; \qquad |w_5 - w_6| \leq 16c_0\widetilde{N}\sigma,$$

and as before we actually have $w_6 = 0$, so we conclude

$$|w_4| \leq 32c_0\widetilde{N}\sigma. \tag{G.33}$$

Notice now that $v_5 - v_5' = w_1 - w_4$, so (G.32) and (G.33) give us

$$|v_5 - v_5'| \leq 64c_0\widetilde{N}\sigma,$$

and combining this with (G.31) we obtain

$$|v_1 - v_1'| \leq 192c_0\widetilde{N}\sigma. \tag{G.34}$$

This is exactly what we sought to prove. □

Proof of Lemma 4.14: We proceed by considering the behavior of f near the endpoints of the $ABC\ldots - BAC\ldots$ edge of the vote simplex, showing that $\overline{f}$ cannot be very close to linearity.

Given f, let Γ denote the supremum of the left-hand side of (4.10), over all choices of x, y, z, x', z'. Also define v_1 and v_1' as in the proof of Lemma 4.13. Now, we consider two cases.

(i) There is some $K \leq \sqrt{\widetilde{N}}/2$ such that $f(K\ BAC\ldots, \widetilde{N} - K\ ABC\ldots) \neq A$. In this case, as long as N is sufficiently large, we have

$$\begin{aligned} \overline{f}_A\begin{pmatrix} 1-\frac{K}{\widetilde{N}} & ABC\ldots \\ \frac{K}{\widetilde{N}} & BAC\ldots \end{pmatrix} &\leq 1 - \mathbf{P}\left(\begin{array}{c} \widetilde{N}-K \\ K \end{array} \middle| \widetilde{N};\ \begin{array}{c} 1-\frac{K}{\widetilde{N}} \\ \frac{K}{\widetilde{N}} \end{array}\right) \\ &\leq 1 - \sigma_N^* \\ &< 1 - \frac{1}{\sqrt{2\widetilde{N}}} \end{aligned}$$

by Lemma 2.4 and the asymptotic behavior of σ_N^*. Therefore by taking $x = 1-K/\widetilde{N}$, $y = K/\widetilde{N}$, $z = 0$, and noting $\overline{f}(x+y\ ABC\ldots, z\ BAC\ldots) = f(\widetilde{N}\ ABC) = A$ by weak unanimity, we get

$$(v_1)_A \leq -\frac{1}{\sqrt{2\widetilde{N}}}.$$

It follows that for any choices of $x', z' \geq 0$ with $x' + z' = 1 - K/\widetilde{N}$,

$$(v_1')_A \leq -\frac{1}{\sqrt{2\widetilde{N}}} + \Gamma.$$

In particular, for any positive integer $r \leq \lfloor 2\sqrt{\widetilde{N}} \rfloor$, we may take $x' = 1 - rK/\widetilde{N}, z = (r-1)K/\widetilde{N}$ to obtain

$$\overline{f}_A \begin{pmatrix} 1 - r\frac{K}{\widetilde{N}} & ABC\ldots \\ r\frac{K}{\widetilde{N}} & BAC\ldots \end{pmatrix} - \overline{f}_A \begin{pmatrix} 1 - (r-1)\frac{K}{\widetilde{N}} & ABC\ldots \\ (r-1)\frac{K}{\widetilde{N}} & BAC\ldots \end{pmatrix} \leq -\frac{1}{\sqrt{2\widetilde{N}}} + \Gamma.$$

Put $\overline{r} = 2\sqrt{\widetilde{N}}\rfloor$, and apply the above inequality for each $r = 1, 2, \ldots, \overline{r}$ and telescope. This gives

$$\overline{f}_A \begin{pmatrix} 1 - \overline{r}\frac{K}{\widetilde{N}} & ABC\ldots \\ \overline{r}\frac{K}{\widetilde{N}} & BAC\ldots \end{pmatrix} - \overline{f}_A \begin{pmatrix} 1 & ABC\ldots \\ 0 & BAC\ldots \end{pmatrix} \leq \overline{r}\left(-\frac{1}{\sqrt{2\widetilde{N}}} + \Gamma\right).$$

The left side cannot be lower than $0 - 1 = -1$, so

$$-1 \leq \overline{r}\left(-\frac{1}{\sqrt{2\widetilde{N}}} + \Gamma\right)$$

which leads to

$$\Gamma \geq \frac{1}{\sqrt{2\widetilde{N}}} - \frac{1}{\overline{r}} \sim \left(\frac{\sqrt{2}-1}{2}\right) \cdot \frac{1}{\sqrt{\widetilde{N}}}.$$

(ii) For all $K \leq \sqrt{\widetilde{N}}/2$, $f(K\ BAC\ldots, \widetilde{N} - K\ ABC\ldots) = A$. Then apply Lemma C.9 with $c = 1/6$ to conclude that if an $\widetilde{N}$-profile P is drawn $IID(\alpha\ BAC\ldots, 1 - \alpha\ ABC\ldots)$ for any $\alpha \leq 1/6\sqrt{\widetilde{N}}$, then the probability that $f(P) \neq A$ is at most $1/\widetilde{N}$, as long as N is sufficiently large.

Let s be an integer with $6\sqrt{\widetilde{N}} < s < 7\sqrt{\widetilde{N}}$. Then taking $x = 1 - 1/s$, $y = 1/s$, $z = 0$, and again using $f(x + y\ ABC\ldots, z\ BAC\ldots) = A$ by weak unanimity, we get

$$(v_1)_A \geq -\frac{1}{\widetilde{N}}.$$

So for any choices of $x', z' \geq 0$ with $x' + z' = 1 - 1/s$, we have

$$(v_1')_A \geq -\frac{1}{\widetilde{\widetilde{N}}} - \Gamma.$$

In particular, for any $r = 0, \ldots, s-1$, we can take $x' = (s-1-r)/s$ and $z' = r/s$ to obtain

$$\overline{f}_A \begin{pmatrix} 1 - \frac{r+1}{s} & ABC\ldots \\ \frac{r+1}{s} & BAC\ldots \end{pmatrix} - \overline{f}_A \begin{pmatrix} 1 - \frac{r}{s} & ABC\ldots \\ \frac{r}{s} & BAC\ldots \end{pmatrix} \geq -\frac{1}{\widetilde{\widetilde{N}}} - \Gamma.$$

Summing for $r = 0, \ldots, s-1$ and telescoping gives

$$\overline{f}_A \begin{pmatrix} 0 & ABC\ldots \\ 1 & BAC\ldots \end{pmatrix} - \overline{f}_A \begin{pmatrix} 1 & ABC\ldots \\ 0 & BAC\ldots \end{pmatrix} \geq s\left(-\frac{1}{\widetilde{\widetilde{N}}} - \Gamma\right).$$

Using weak unanimity, the left side equals $0 - 1 = -1$, so

$$-1 \geq s\left(-\frac{1}{\widetilde{\widetilde{N}}} - \Gamma\right)$$

from which

$$\Gamma \geq \frac{1}{s} - \frac{1}{\widetilde{\widetilde{N}}} \gtrsim \frac{1}{8\sqrt{\widetilde{\widetilde{N}}}}.$$

In both cases (i) and (ii), we showed that Γ was bounded below (asymptotically) by a constant times $1/\sqrt{\widetilde{\widetilde{N}}}$, which is exactly what the lemma claims.

□

This wraps up Theorem 4.1. Now we give the proof of Theorem 4.6, for simple and weakly unanimous voting rules. Essentially, we just need to replace Lemma 4.14 with a corresponding statement giving a sharper bound when the rule is simple:

Lemma G.1 *There exists some absolute constant c_3, independent of N, with the following property: As long as N is large enough, for any f that is weakly unanimous and simple over A and B, there exist some nonnegative x, y, z, x', z' with*

$$\left| \left(\overline{f} \begin{pmatrix} x & ABC\ldots \\ y+z & BAC\ldots \end{pmatrix} - \overline{f} \begin{pmatrix} x+y & ABC\ldots \\ z & BAC\ldots \end{pmatrix} \right) - \left(\overline{f} \begin{pmatrix} x' & ABC\ldots \\ y+z' & BAC\ldots \end{pmatrix} - \overline{f} \begin{pmatrix} x'+y & ABC\ldots \\ z' & BAC\ldots \end{pmatrix} \right) \right| \geq c_3. \quad \text{(G.35)}$$

Proof: Let K^* be the threshold such that $f(K\ ABC\ldots, \widetilde{N}-K\ BAC\ldots) = A$ iff $K \geq K^*$. Just as in the proof of Theorem 4.7, assume that $K^* \leq \widetilde{N}/2$ (otherwise switch A and B), and put

$$\phi_1 = (\alpha_1\ ABC\ldots, 1-\alpha_1\ BAC\ldots) \qquad \text{with} \qquad \alpha_1 = \frac{K^* + \sqrt{2\widetilde{N}}}{\widetilde{N}},$$

$$\phi_2 = (\alpha_2\ ABC\ldots, 1-\alpha_2\ BAC\ldots) \qquad \text{with} \qquad \alpha_2 = \max\left\{\frac{K^* - \sqrt{2\widetilde{N}}}{\widetilde{N}}, 0\right\}.$$

By simplicity, $\overline{f}(\phi_1), \overline{f}(\phi_2)$ both put positive weight only on A and B, and by the same Chebyshev argument as in Theorem 4.2, $\overline{f}(\phi_1)$ puts probability at least 7/8 on A, while $\overline{f}(\phi_2)$ puts probability at most 1/8 on A.

Next put

$$\phi_3 = (\alpha_3\ ABC\ldots, 1-\alpha_3\ BAC\ldots) \qquad \text{with} \qquad \alpha_3 = 2\alpha_1 - \alpha_2.$$

The weight on $ABC\ldots$ is $\leq (K^* + \sqrt{2\widetilde{N}})/\widetilde{N} + 2\sqrt{2\widetilde{N}}/\widetilde{N} < 1$ for large N, so this is a valid distribution. Since $\alpha_3 > \alpha_1$, the same Chebyshev argument gives that $\overline{f}(\phi_3)$ puts probability at least 7/8 on A (and the remaining probability on B). We now have

$$|\overline{f}(\phi_2) - \overline{f}(\phi_1)| \geq 3/2,$$

$$|\overline{f}(\phi_1) - \overline{f}(\phi_3)| \leq 1/4.$$

Now take

$$x = \alpha_2, \qquad y = \alpha_1 - \alpha_2, \qquad z = 1 - \alpha_1,$$

$$x' = \alpha_1, \qquad z' = 1 - \alpha_3.$$

The expression on the left side of (G.35) reduces to

$$\left|\left(\overline{f}(\phi_2) - \overline{f}(\phi_1)\right) - \left(\overline{f}(\phi_1) - \overline{f}(\phi_3)\right)\right| \geq \frac{3}{2} - \frac{1}{4} = \frac{5}{4}$$

which proves the lemma. □

Proof of Theorem 4.6: As usual, it suffices to assume N is large enough so that Lemma G.1 applies. Assume A, B, C are chosen so that f is simple over A and B. Let c_1, c_0, c_3 be as in Lemmas 4.12, 4.13, G.1. Either $\sigma \geq c_1/N$, and we are done; or else

Lemma 4.12 applies, in which case (4.9) and (by simplicity) (G.35) apply; combining these gives $\sigma \geq c_3/192c_0\widetilde{N}$. □

H Construction for quickly-decaying susceptibility

We provide here the construction of a tops-only voting rule whose susceptibility shrinks in N at rate $N^{-\kappa}$ with $\kappa > 1/2$, as required by Theorem 4.3. The actual construction is more elaborate than the approximate random dictatorship sketched in the main paper, so we first give a more detailed overview.

The main idea behind the construction is to subdivide the simplex of vote profiles into *blocks*, as illustrated in Figure H.1. Within each block, we then assign winners $A_1, \ldots, A_M$ to the various profiles, in proportions that correspond to the position of the block in the vote simplex.

More specifically, in order to avoid creating especially large opportunities for manipulation near the edge of the vote simplex, we need to focus on *viable* candidates at each vote profile, much as in the construction of the pair-or-plurality system in Subsection 3.3. Roughly speaking, each candidate needs to get more than some threshold number of votes to be considered viable; the threshold will be taken to be (asymptotically) some constant λ times N. Then, for each set $\mathcal{C}'$ of candidates, we consider the set of all vote profiles in which the viable candidates are precisely the members of $\mathcal{C}'$, and carve up this set of profiles into blocks, depending on how many votes each viable candidate receives. All blocks have equal size S along each of the dimensions corresponding to a viable candidate.

For any given block, we define a weight for each viable candidate by subtracting λN from her vote total. We then define the voting rule within the block by assigning a winner at each vote profile, in such a way that the fraction of profiles assigned to any (viable) candidate is approximately proportional to her weight. This principle tells us how many profiles each candidate gets within the block; to decide exactly *which* profiles she gets, so as to keep the difference between her relative probability of winning and her weight tightly controlled, we use Lemma C.11.

Consider now the susceptibility of a voting rule defined in this way, with blocks of size S. When the manipulator changes his vote, this affects the distribution over realized vote profiles in two ways: it changes the distribution over blocks, and it changes the distribution over profiles within each block. Because the distribution within each block is close to the weights in that block, we are assured that the distribution across blocks approximately pins down the distribution of winners — more concretely, the error in this approximation is of order $S^d N^{-\frac{d-2}{2}}$ (ignoring constant factors). Here d is the value used in applying Lemma C.11. We also show that the change across blocks affects the distribution over candidates on the order of $S^{-\frac{1}{d}} N^{-\frac{1}{2}}$. Hence, our construction gives an upper bound for susceptibility that is approximately on the same order as $\max\{S^d N^{-\frac{d-2}{2}}, S^{-\frac{1}{d}} N^{-\frac{1}{2}}\}$. In order to achieve the fastest possible rate of decline in susceptibility as $N \to \infty$, we choose $d = 6$ and $S \approx N^{\frac{9}{37}}$, with the resulting rate of decline $N^{-\frac{20}{37}}$. We will henceforth use these numbers for concreteness.[2]

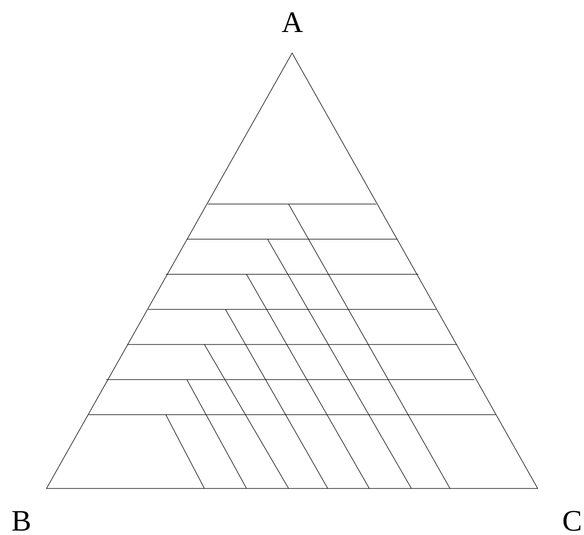

Figure H.1: Sketch of the construction for Theorem 4.3: each region of the vote simplex bounded by the lines shown represents a block

Proof of Theorem 4.3: We first give the exact construction of the voting system. Fix constants λ, μ with $0 < \lambda < 1/M$ and $0 < \mu < 1 - M\lambda$. Also fix $\underline{\alpha}$ with $0 < \underline{\alpha} < \min\{\lambda/3, \mu/3, (1 - M\lambda - \mu)/(M + 1)\}$.

For each value of N, choose integers S_N, L_N, R_N such that

- $S_N = 2^{6h}$ for some integer h and $N^{\frac{9}{37}} \lesssim S_N \lesssim 2^6 \cdot N^{\frac{9}{37}}$;
- $L_N \sim \lambda N$;
- $R_N S_N \sim \mu N$.

We will henceforth refer to these as S, L, R, with the dependence on N implicit. The roles played by these values will be as follows: L is the number of votes needed for a

[2]It is possible to achieve faster rates of convergence through minor improvements on the construction, but we do not include the details here since they do not seem to yield substantial new insights.

candidate to be viable; S is the size of a block, and R measures the number of blocks (in each dimension).

A *block label* is a sequence consisting of $M-1$ or fewer (possibly zero) nonnegative integers, whose sum is at most R.

Given a profile $P = (x_1\ A_1, \ldots, x_M\ A_M)$ of first-place votes, we compute a corresponding block label $BL(P)$ by the following algorithm:

1. For each $i = 1, \ldots, M$, if $x_i < L$, put $\Lambda_i = 0$. Otherwise, put $\Lambda_i = \lfloor (x_i - L)/S \rfloor + 1$.

2. Let t be the smallest index such that $\Lambda_1 + \cdots + \Lambda_t > R$. Notice that t must exist (for large enough N), since

$$\frac{\Lambda_1 + \cdots + \Lambda_M}{R} \geq \frac{x_1 + \cdots + x_M - ML}{RS} \sim \frac{(1 - M\lambda)N}{RS} \sim \frac{1 - M\lambda}{\mu} > 1.$$

   Then define $BL(P)$ to be the $(t-1)$-term sequence $\Lambda = (\Lambda_1, \ldots, \Lambda_{t-1})$.

Given a block label Λ, we define the corresponding *block* as $BL^{-1}(\Lambda)$, the set of profiles that can be obtained by inverting the above procedure. For each candidate A_i we construct a lower bound $\underline{x}_i$ and an upper bound $\overline{x}_i$ on the number of votes: If $\Lambda = (\Lambda_1, \ldots, \Lambda_{t-1})$, then

- if $i \leq t-1$ and $\Lambda_i = 0$, put $\underline{x}_i = 0$ and $\overline{x}_i = L - 1$;
- if $i \leq t-1$ and $\Lambda_i > 0$, put $\underline{x}_i = S(\Lambda_i - 1) + L$ and $\overline{x}_i = S\Lambda_i + L - 1$;
- for $i = t$, put $\underline{x}_i = S(R - \sum_j \Lambda_j) + L$ and $\overline{x}_i = N + 1$;
- for $i > t$, put $\underline{x}_i = 0$ and $\overline{x}_i = N + 1$.

Then one readily checks that $BL^{-1}(\Lambda)$ is the set of all $(N+1)$-profiles of votes $(x_1, \ldots, x_M)$ such that $\underline{x}_i \leq x_i \leq \overline{x}_i$ for all i.

We also define *weights* $W_i(\Lambda)$, for each block label Λ and each candidate A_i. If $\Lambda = (\Lambda_1, \ldots, \Lambda_{t-1})$ then the weights are defined by

- for $i \leq t-1$, $W_i(\Lambda) = \Lambda_i/(R+1)$;
- for $i = t$, $W_i(\Lambda) = 1 - \sum_j \Lambda_j/(R+1)$;
- for $i > t$, $W_i(\Lambda) = 0$.

Thus we always have $\sum_i W_i(\Lambda) = 1$.

We further modify these weights by rounding down to integer multiples of $1/2^h$:

- for $i \leq t-1$, $\widetilde{W}_i(\Lambda) = \lfloor 2^h W_i(\Lambda) \rfloor / 2^h$;
- for $i = t$, $\widetilde{W}_i(\Lambda) = 1 - \sum_{j<t} \widetilde{W}_j(\Lambda)$;
- for $i > t$, $\widetilde{W}_i(t) = 0$.

Let $S = 2^{6h}$, and let $Z = \{0, 1, \ldots, 2^{6h}-1\}$ be partitioned into 2^h subsets $Z_0, \ldots, Z_{2^h-1}$ according to Lemma C.11. For each block label Λ, we let $g_\Lambda : \{0, 1, \ldots, 2^h - 1\} \to \mathcal{C}$ be any function such that $|g_\Lambda^{-1}(A_i)| = 2^h \widetilde{W}_i(\Lambda)$ for each candidate A_i. Thus, the proportion of values of y on which g_Λ takes the value A_i equals the rounded weight of A_i.

Finally, we are ready to define the voting rule f. Given a profile of votes, $P = (x_1\ A_1, \ldots, x_M\ A_M)$, we define $f(P)$ as follows:

- Let $(\Lambda_1, \ldots, \Lambda_{t-1}) = \Lambda = BL(P)$ be the block label.
- If every term Λ_i is zero, then let $f(P) = A_t$.
- Otherwise, consider the smallest i such that $\Lambda_i > 0$. Let

  $$\widehat{x}_i = (x_i - L) - S \left\lfloor \frac{x_i - L}{S} \right\rfloor.$$

  Then $\widehat{x}_i$ is an element of Z. So $\widehat{x}_i \in Z_y$ for exactly one y. Put $f(P) = g_\Lambda(y)$.

This defines the voting rule. The statement of Theorem 4.3 promised that it would be Pareto efficient and tops-only. Tops-onliness is clear from the construction, so we should check Pareto efficiency. Evidently we must check that $f(P)$ is always a candidate who gets at least one vote in profile P. If every term Λ_i of the block label $BL(P)$ is zero, then $\Lambda_t > R > 1$ so that $x_t > 0$. Otherwise, notice that whenever A_i is a candidate with $\Lambda_i = 0$, then $\widetilde{W}_i(\Lambda) = 0$, and so $g_\Lambda(y) \neq A_i$ for all y. Consequently we cannot have $f(P) = A_i$ for any such i. Thus, $f(P)$ must be a candidate A_i for whom $\Lambda_i > 0$, implying $x_i > 0$.

Our remaining task is to prove the susceptibility bound. The proof of the bound is based on two claims. Let ϵ be an arbitrary small positive constant, and $\mathcal{BL}$ the set of all block labels (which depends on N).

Claim I. There is a constant c_I such that the following holds, for every value of N: For all distributions $\phi \in \Delta(\mathcal{C})$, all candidates A_i, A_j,

$$\left| \Pr_\phi(f(A_j, P) = A_i) - \sum_{\Lambda \in \mathcal{BL}} \Pr_\phi((A_j, P) \in BL^{-1}(\Lambda)) \widetilde{W}_i(\Lambda) \right| < c_I N^{-\left(\frac{20}{37} - \epsilon\right)}.$$

Here the $\Pr_\phi(\cdots)$ expressions refer to probabilities concerning the profile (A_j, P), given that P is formed by having each of the N other votes drawn independently from ϕ.

Claim II. There is a constant c_{II} such that the following holds, for every value of N: For all distributions $\phi \in \Delta(\mathcal{C})$, and all candidates A_i, A_j, A_k,

$$\left| \sum_{\Lambda \in \mathcal{BL}} \Pr_\phi((A_j, P) \in BL^{-1}(\Lambda)) \widetilde{W}_i(\Lambda) - \sum_{\Lambda \in \mathcal{BL}} \Pr_\phi((A_k, P) \in BL^{-1}(\Lambda)) \widetilde{W}_i(\Lambda) \right| < c_{II} N^{-\left(\frac{20}{37} - \epsilon\right)}. \quad \text{(H.1)}$$

We shall prove these two claims, then show how this quickly completes the proof of the theorem.

Proof of Claim I. We rewrite the expression inside the absolute value as

$$\sum_{\Lambda \in \mathcal{BL}} \left[\Pr_\phi((A_j, P) \in BL^{-1}(\Lambda) \text{ and } f(A_j, P) = A_i) - \Pr_\phi((A_j, P) \in BL^{-1}(\Lambda)) \widetilde{W}_i(\Lambda) \right].$$

For each block label Λ consisting of zeroes, the relevant difference is zero. (If $t-1$ is the length of Λ, then $\widetilde{W}_t(\Lambda) = 1$, while $\widetilde{W}_i(\Lambda) = 0$ for $i \neq t$; and f takes the value A_t throughout $BL^{-1}(\Lambda)$.) So we can restrict to the sum over Λ having a nonzero component.

For each pair (k, t) with $1 \leq k < t \leq M$, let $\mathcal{BL}_{k,t}$ be the set of all block labels Λ with length $t-1$ such that $\Lambda_l = 0$ for all $l < k$, but $\Lambda_k > 0$. It suffices to show that there is a constant c', independent of ϕ (or N), such that

$$\left| \sum_{\Lambda \in \mathcal{BL}_{k,t}} \left[\Pr_\phi((A_j, P) \in BL^{-1}(\Lambda) \cap f^{-1}(A_i)) - \Pr_\phi((A_j, P) \in BL^{-1}(\Lambda)) \widetilde{W}_i(\Lambda) \right] \right| < c' N^{-\left(\frac{20}{37} - \epsilon\right)}. \quad \text{(H.2)}$$

Accordingly, we take k, t as fixed from here on.

First consider any distribution ϕ such that $\phi_t < \underline{\alpha}$. If $P \sim IID(\phi)$, then the number of votes received by candidate A_t in P has expectation $\phi_t N \leq \underline{\alpha} N$ and variance $\phi_t(1-\phi_t)N \leq$

$\underline{\alpha}N$, so by Chebyshev, the probability that A_t's vote count is at least $2\underline{\alpha}N$ is $\leq 1/\underline{\alpha}N$. Consequently, the probability that (A_j, P) gives A_t at least $2\underline{\alpha}N + 1$ votes is $\leq 1/\underline{\alpha}N$. Notice that at every profile in any block $\Lambda \in \mathcal{BL}_{k,t}$, we must have $\Lambda_t \geq 1$ from which $x_t \geq L > 2\underline{\alpha}N + 1$. Thus

$$\sum_{\Lambda \in \mathcal{BL}_{k,t}} \Pr_{\phi}((A_j, P) \in BL^{-1}(\Lambda)) \leq 1/\underline{\alpha}N. \tag{H.3}$$

But both probabilities on the left side of (H.2) are bounded above by the sum in (H.3), hence (H.2) holds in this case (with the appropriate choice of c').

Similarly, consider any distribution ϕ such that $\phi_k < \underline{\alpha}$. Because every block $\Lambda \in \mathcal{BL}_{k,t}$ must have $\Lambda_k \geq 1$, from which any profile in such a block must have $x_k \geq L > 2\underline{\alpha}N + 1$, we can follow the same argument to show that (H.2) is satisfied again.

This means we can now restrict to distributions ϕ such that

$$\phi_t \geq \underline{\alpha} \qquad \text{and} \qquad \phi_k \geq \underline{\alpha}.$$

Take ϕ as given, and let l be the highest index such that $\phi_l \geq \underline{\alpha}$; thus $l \geq t$.

Consider any block $\Lambda = (\Lambda_1, \ldots, \Lambda_{t-1}) \in \mathcal{BL}_{k,t}$. For each $s = 1, \ldots, M$, define bounds $\underline{x}_s, \overline{x}_s$ as in the computation of $BL^{-1}(\Lambda)$ above. Consider any given values x_s, with $\underline{x}_s \leq x_s \leq \overline{x}_s$, for each $s \neq k, l$; write x_{-kl} for the vector of such values. Define $[x_{-kl}]$ to be the set of all profiles having the specified number of votes for each candidate A_s, $s \neq k, l$.

We further break down the left-hand side of (H.2) by summing over different values of x_{-kl}. Define notations

$$\begin{aligned}
\Pi_1(\Lambda, x_{-kl}) &= \Pr_{\phi}\left((A_j, P) \in BL^{-1}(\Lambda) \cap [x_{-kl}] \cap f^{-1}(A_i)\right), \\
\Pi_2(\Lambda, x_{-kl}) &= \Pr_{\phi}\left((A_j, P) \in BL^{-1}(\Lambda) \cap [x_{-kl}]\right) \cdot \widetilde{W}_i(\Lambda).
\end{aligned}$$

Then in the left-hand side of (H.2), the first probability is $\sum_{x_{-kl}} \Pi_1(\Lambda, x_{-kl})$ (where the sum is over all vectors x_{-kl} of $M - 2$ nonnegative integers), and the second probability is $\sum_{x_{-kl}} \Pi_2(\Lambda, x_{-kl})$. Thus, (H.2) is equivalent to

$$\left| \sum_{\substack{\Lambda \in \mathcal{BL}_{k,t} \\ x_{-kl}}} [\Pi_1(\Lambda, x_{-kl}) - \Pi_2(\Lambda, x_{-kl})] \right| < c' N^{-\left(\frac{20}{37} - \epsilon\right)}. \tag{H.4}$$

We prove this by breaking into three cases depending on the choice of Λ and x_{-kl}. We first deal with cases that have low probability, so that their contribution to the sums of Π_1, Π_2 in (H.4) is small; and then we can deal with the substantive case where we actually make use of the elaborate construction behind f within each block.

(i) First, for each $s \neq k, l$, consider choices of x_{-kl} that have $|x_s - \phi_s N| > \underline{\alpha} N/M$. The probability that (A_j, P) gives candidate A_s such a number of votes is at most the probability that P gives A_s a number of votes more than $\underline{\alpha} N/2M$ away from $\phi_s N$. By the usual Chebyshev argument, this probability is $\leq 4M^2/\underline{\alpha}^2 N$.

Since $\Pi_1(\Lambda, x_{-kl}) \leq \mathrm{Pr}_\phi((A_j, P) \in BL^{-1}(\Lambda) \cap [x_{-kl}])$, the sum of $\Pi_1(\Lambda, x_{-kl})$ over all Λ and all x_{-kl} with $|x_s - \phi_s N| > \underline{\alpha} N/M$ is at most $4M^2/\underline{\alpha}^2 N$. Similarly, the same holds for Π_2. Thus, all the pairs (Λ, x_{-kl}) for which $|x_s - \phi_s N| > \underline{\alpha} N/M$ make a total contribution to the left side of (H.4) that is bounded above by a constant times N^{-1}.

(ii) Next, consider choices of $\Lambda \in \mathcal{BL}_{k,t}$ that have $|(S(\Lambda_k - 1) + L) - \phi_k N| > \underline{\alpha} N/M$. If (A_j, P) is in such a block $BL^{-1}(\Lambda)$, then the number of votes for candidate A_k is between $\underline{x}_k = S(\Lambda_k - 1) + L$ and $\overline{x}_k = S\Lambda_k + L - 1$. For N sufficiently large, this means that the number of votes for A_k in P is more than $\underline{\alpha} N/2M$ away from $\phi_k N$. Again, this occurs with probability $\leq 4M^2/\underline{\alpha} N$.

Thus, the pairs (Λ, x_{-kl}) for which $|(S(\Lambda_k - 1) + L) - \phi_k N| > \underline{\alpha} N/M$ make a total contribution to the left side of (H.4) that is bounded above by a constant times N^{-1}.

From this and the previous bullet point, we see that in proving (H.4) it suffices to restrict attention to pairs (Λ, x_{-kl}) for which

$$|x_s - \phi_s N| \leq \underline{\alpha} N/M \qquad \text{for all } s \neq k, l; \tag{H.5}$$

$$|(S(\Lambda_k - 1) + L) - \phi_k N| \leq \underline{\alpha} N/M. \tag{H.6}$$

That is, the contribution of all other pairs to the sum in (H.4) is negligible.

(iii) We will show that (H.5) and (H.6) imply

$$|\Pi_1(\Lambda, x_{-kl}) - \Pi_2(\Lambda, x_{-kl})| \leq c'' N^{-\left(\frac{48}{37} - \epsilon\right)} \Pr_\phi((A_j, P) \in [x_{-kl}]) \tag{H.7}$$

where c'' is a constant not depending on ϕ, N, or x_{-kl}.

For each candidate A_s, define bounds $\underline{x}_s, \overline{x}_s$ as in the calculation of $BL^{-1}(\Lambda)$. We will first show that $BL^{-1}(\Lambda) \cap [x_{-kl}]$ contains exactly $\overline{x}_k - \underline{x}_k + 1$ profiles. That is, for every choice of x_k with $\underline{x}_k \leq x_k \leq \overline{x}_k$, there is exactly one choice of x_l such that the profile (x_k, x_l, x_{-kl}) is in $BL^{-1}(\Lambda) \cap [x_{-kl}]$. The relevant choice of x_l would of course be $x_l = x_{k+l} - x_k$, where $x_{k+l} = N + 1 - \sum_{s \neq k,l} x_s$, so we just need to check that this value of x_l always lies between the bounds $\underline{x}_l$ and $\overline{x}_l$.

There are two cases for the lower bound:

- If $l > t$, then $\underline{x}_l = 0$. We have

$$\begin{aligned}
x_l &= N + 1 - \sum_{s \neq k,l} x_s - x_k \\
&\geq N + 1 - \sum_{s \neq k,l} (\phi_s N + \underline{\alpha} N/M) - \overline{x}_k \\
&\geq (\phi_k + \phi_l)N - (M-2)\underline{\alpha} N/M - S\Lambda_k - L \\
&\geq (\phi_k + \phi_l)N - (M-2)\underline{\alpha} N/M - (S - L + \phi_k N + \underline{\alpha} N/M) - L \\
&= \phi_l N - (M-1)\underline{\alpha} N/M - S \\
&\geq \underline{\alpha} N/M - S \\
&\geq 0
\end{aligned}$$

  as long as N is sufficiently large.

- If $l = t$, then $\underline{x}_l = S(R - \sum_{s<t} \Lambda_s) + L$. We also know, by definition of l, that $\phi_s < \underline{\alpha}$ for each $s > t$, so (H.5) implies $x_s \leq \underline{\alpha}(1 + 1/M)N$ for each such s.

Let ν be the constant $1 - \underline{\alpha}(M+1) - (M\lambda + \mu) > 0$. We have

$$\begin{aligned}
x_l &= N + 1 - \sum_{s \neq t} x_s \\
&\geq N + 1 - \sum_{s<t} \overline{x}_s - \sum_{s>t} \underline{\alpha}(1 + 1/M)N \\
&\geq N + 1 - \sum_{s<t} (S\Lambda_s + L - 1) - \underline{\alpha}(M+1)N \\
&\geq N(1 - \underline{\alpha}(M+1)) - S \sum_{s<t} \Lambda_s - (M-1)L \\
&= (M\lambda + \mu + \nu)N - S \sum_{s<t} \Lambda_s - ML + L \\
&\geq ML + RS - S \sum_{s<t} \Lambda_s - ML + L \\
&\qquad\qquad \text{(when } N \text{ is sufficiently large)} \\
&= S\left(R - \sum_{s<t} \Lambda_s\right) + L.
\end{aligned}$$

Thus the lower bound is satisfied when $l = t$ as well.

As for the upper bound, in both cases, $\overline{x}_l = N + 1$. Then

$$x_l = N + 1 - \sum_{s \neq k,l} x_s - x_k \leq N + 1$$

so the upper bound is always satisfied.

Thus $BL^{-1}(\Lambda) \cap [x_{-kl}]$ contains exactly $\overline{x}_k - \underline{x}_k + 1 = S$ profiles.

For each profile (x_k, x_l, x_{-kl}), we will explicitly write out the probability of achieving this profile as the realized value of (A_j, P). Specifically, let $(x_k^-, x_l^-, x_{-kl}^-)$ be identical to (x_k, x_l, x_{-kl}) except that the j-component has been decreased by 1. (There is no more succinct way to write this without breaking into cases depending whether $j = k$, $j = l$, or neither.) Likewise put $x_{k+l}^- = x_k^- + x_l^-$, and as usual write $\phi_{k+l} = \phi_k + \phi_l$ and ϕ_{-kl} for the vector of other components of ϕ. Then the probability of achieving (x_k, x_l, x_{-kl}) is

$$\mathbf{P}\left(\begin{array}{c} x_k^- \\ x_l^- \\ x_{kl}^- \end{array} \middle| \; N; \; \phi \right) = \mathbf{P}\left(\begin{array}{c} x_{k+l}^- \\ x_{-kl}^- \end{array} \middle| \; N; \; \begin{array}{c} \phi_{k+l} \\ \phi_{-kl}^- \end{array} \right) \mathbf{P}\left(\begin{array}{c} x_k^- \\ x_l^- \end{array} \middle| \; x_{k+l}^-; \; \begin{array}{c} \phi_k/\phi_{k+l} \\ \phi_l/\phi_{k+l} \end{array} \right)$$

by Lemma C.2.

For succinctness, write

$$\beta = \mathbf{P}\left(\begin{array}{c} x_{k+l}^{-} \\ x_{-kl}^{-} \end{array} \middle| \, N; \begin{array}{c} \phi_{k+l} \\ \phi_{-kl}^{-} \end{array} \right)$$

for the first factor, which is independent of x_k, and

$$\widehat{\mathbf{P}}(x_k^-) = \mathbf{P}\left(\begin{array}{c} x_k^- \\ x_{k+l}^- - x_k^- \end{array} \middle| \, x_{k+l}^-; \begin{array}{c} \phi_k/\phi_{k+l} \\ \phi_l/\phi_{k+l} \end{array} \right)$$

for the second factor.

Let $\underline{x}_k^- = \underline{x}_k$ or $\underline{x}_k - 1$ (depending whether $j \neq k$ or $j = k$). Then the possible values of x_k^- corresponding to profiles $(x_k, x_l, x_{-kl}) \in BL^{-1}(\Lambda) \cap [x_{-kl}]$ are exactly the numbers $\underline{x}_k^- + z$, for $z \in Z$. (Recall we defined $Z = \{0, 1, \ldots, S-1\}$.)

Now, we have $\phi_k/\phi_{k+l} \geq \phi_k \geq \underline{\alpha}$, and likewise $\phi_l/\phi_{k+l} \geq \underline{\alpha}$. We also have $x_{k+l}^- \geq \phi_{k+l}N - \underline{\alpha}N - 1$ (using (H.5)) $\geq \underline{\alpha}N - 1$, a lower bound that grows linearly in N.

Consequently, we can apply Lemma C.11, with $d = 6$. As long as N is greater than some absolute threshold N_0, we have the inequality for any two values $y, y' \in \{0, 1, \ldots, 2^h - 1\}$:

$$\left| \sum_{z \in Z_y} \widehat{\mathbf{P}}(\underline{x}_k^- + z) - \sum_{z \in Z_{y'}} \widehat{\mathbf{P}}(\underline{x}_k^- + z) \right| \leq 2^{41h} h \left(x_{k+l}^-\right)^{-6\left(\frac{1}{2} - \frac{\epsilon}{18}\right)}. \tag{H.8}$$

This inequality is the key step in the proof of Claim I; it was for this reason that we needed to use the sets Z_y in designing f.

We have the bounds $h \leq \ln(N) \leq N^{\frac{\epsilon}{3}}$ for large N; $x_{k+l}^- \geq \underline{\alpha}N - 1$; and

$$2^{41h} = S^{\frac{41}{6}} \leq (constant) \cdot N^{\frac{123}{74}}.$$

Applying these to simplify the right side of (H.8) gives

$$\left| \sum_{z \in Z_y} \widehat{\mathbf{P}}(\underline{x}_k^- + z) - \sum_{z \in Z_{y'}} \widehat{\mathbf{P}}(\underline{x}_k^- + z) \right| \leq (constant) \cdot N^{-\left(\frac{99}{74} - \epsilon\right)}.$$

Next, sum over all choices of $y' \in \{0, 1, \ldots, 2^h - 1\}$ and use the triangle inequality.

Since $2^h = S^{1/6} \lesssim (constant) \cdot N^{3/74}$, we obtain

$$\left| 2^h \sum_{z \in Z_y} \widehat{\mathbf{P}}(\underline{x}_k^- + z) - \sum_{z=0}^{2^{6h}-1} \widehat{\mathbf{P}}(\underline{x}_k^- + z) \right| \leq (constant) \cdot N^{-\left(\frac{96}{74} - \epsilon\right)}.$$

Sum again over all y with $g_\Lambda(y) = A_i$, and then also divide by 2^h. The right-hand side has been multiplied by $\widetilde{W}_i(\Lambda) \leq 1$, and so we get

$$\left| \sum_{\substack{y \in g_\Lambda^{-1}(A_i) \\ z \in Z_y}} \widehat{\mathbf{P}}(\underline{x}_k^- + z) - \widetilde{W}_i(\Lambda) \sum_{z=0}^{2^{6h}-1} \widehat{\mathbf{P}}(\underline{x}_k^- + z) \right| \leq (constant) \cdot N^{-\left(\frac{48}{37} - \epsilon\right)} \tag{H.9}$$

(after simplifying the exponent on the right side).

Now we return to the definitions of Π_1 and Π_2. By definition, $\Pi_1(\Lambda, x_{-kl})$ is the sum of the probabilities of all realizations of P such that $(A_j, P) \in BL^{-1}(\Lambda) \cap [x_{-kl}]$ and $f(A_j, P) = A_i$. Continuing to write $(x_k, x_l, x_{-kl}) = (A_j, P)$ for such a P, and writing $\widehat{f}(x_k) = f(x_k, x_{k+l} - x_k, x_{-kl})$ for each possible value of x_k, we have

$$\Pi_1(\Lambda, x_{-kl}) \quad = \sum_{z : \widehat{f}(\underline{x}_k + z) = A_i} \beta \widehat{\mathbf{P}}(\underline{x}_k^- + z).$$

Moreover, by assumption $\Lambda_k > 0$, while $\Lambda_s = 0$ for all $s < k$. Therefore, the construction of f on the block $BL^{-1}(\Lambda)$ implies that $\widehat{f}(\underline{x}_k + z) = A_i$ if and only if $z \in Z_y$ for some y such that $g_\Lambda(y) = A_i$. That is,

$$\Pi_1(\Lambda, x_{-kl}) \quad = \sum_{\substack{y \in g_\Lambda^{-1}(A_i) \\ z \in Z_y}} \beta \widehat{\mathbf{P}}(\underline{x}_k^- + z).$$

Meanwhile, $\Pi_2(\Lambda, x_{-kl})$ is the sum of the probabilities of all profiles in $BL^{-1}(\Lambda) \cap [x_{-kl}]$, regardless of the corresponding values of f, multiplied by $\widetilde{W}_i(\Lambda)$. This can be written as

$$\Pi_2(\Lambda, x_{-kl}) = \left(\sum_{z=0}^{2^{6h}-1} \beta \widehat{\mathbf{P}}(\underline{x}_k^- + z) \right) \cdot \widetilde{W}_i(\Lambda).$$

Now we see that multiplying (H.9) by β gives

$$|\Pi_1(\Lambda, x_{-kl}) - \Pi_2(\Lambda, x_{-kl})| \leq (constant) \cdot \beta \cdot N^{-\left(\frac{48}{37}-\epsilon\right)}.$$

Since $\beta = \Pr_\phi((A_j, P) \in [x_{-kl}])$, we see that this is exactly (H.7), as promised.

This completes the main goal of item (iii). Before leaving this case, however, let us consider what happens when we hold fixed x_{-kl} and sum over Λ. If $BL^{-1}(\Lambda) \cap [x_{kl}] = \emptyset$, then $\Pi_1(\Lambda, x_{-kl}) = \Pi_2(\Lambda, x_{-kl}) = 0$, so these choices of Λ will contribute nothing to the sum on the left-hand side of (H.4). How many block labels Λ make a nonzero contribution, i.e. satisfy $BL^{-1}(\Lambda) \cap [x_{kl}] \neq \emptyset$? Suppose Λ is such a block label, with length $t-1$. For each $s \leq t-1$ except for $s = k$, the value of Λ_s is uniquely determined by the constraint $\underline{x}_s \leq x_s \leq \overline{x}_s$. (Recall that $l \geq t$.) This determines every component of Λ except for Λ_k, and so we get at most $R+1$ such block labels.

Now we are ready to complete the proof of (H.2). Consider the sum

$$\sum_{\substack{\Lambda \in \mathcal{BL}_{k,t} \\ x_{-kl}}} [\Pi_1(\Lambda, x_{-kl}) - \Pi_2(\Lambda, x_{-kl})]$$

on the left side of (H.2). Each term of the sum is indexed by a pair (Λ, x_{-kl}). Again, we can consider only terms with $BL^{-1}(\Lambda) \cap [x_{-kl}] \neq \emptyset$, because the other terms are all zero.

All the terms for which x_{-kl} violates (H.5) have a total sum whose absolute value is bounded by a constant times N^{-1} (this was case (i)). All the terms for which Λ violates (H.6) have a sum that is again bounded by a constant times N^{-1} (this was case (ii)). For the remaining terms, we apply case (iii). Consider any x_{-kl} satisfying (H.5). Sum over all Λ that satisfy (H.6). Using (H.7), and our previous observation that at most $R+1$ choices of Λ make a nonzero contribution to the left-hand side, we get

$$\begin{aligned}\left| \sum_{\substack{\Lambda \in \mathcal{BL}_{k,t} \\ \Lambda \text{ satisfies (H.6)}}} [\Pi_1(\Lambda, x_{-kl}) - \Pi_2(\Lambda, x_{-kl})] \right| & \\ \leq (constant) \cdot N^{-\left(\frac{48}{37}-\epsilon\right)} & \Pr_\phi((A_j, P) \in [x_{-kl}]) \cdot (R+1) \\ \leq (constant) \cdot N^{-\left(\frac{20}{37}-\epsilon\right)} & \Pr_\phi((A_j, P) \in [x_{-kl}])\end{aligned}$$

since $R+1 \leq (constant) \cdot N^{28/37}$. Summing over all choices of x_{-kl}, and using the obvious fact that

$$\sum_{x_{-kl} \text{ satisfies (H.5)}} \Pr_{\phi}((A_j, P) \in [x_{-kl}]) \leq 1,$$

we obtain

$$\left| \sum_{\substack{\Lambda \in \mathcal{BL}_{k,t} \text{ satisfying (H.6)} \\ x_{-kl} \text{ satisfying (H.5)}}} [\Pi_1(\Lambda, x_{-kl}) - \Pi_2(\Lambda, x_{-kl})] \right| \leq (constant) \cdot N^{-\left(\frac{20}{37} - \epsilon\right)}.$$

These three cases (i)–(iii) together cover every possible pair (Λ, x_{-kl}). So, adding them together, we obtain (H.4). We already saw that (H.4) was equivalent to (H.2), so we have proven (H.2) and the proof of Claim I is complete.

Proof of Claim II. Rewrite the asserted bound as a sum over all N-profiles P:

$$\left| \sum_P \Pr_{\phi}(P) \widetilde{W}_i(BL(A_j, P)) - \sum_P \Pr_{\phi}(P) \widetilde{W}_i(BL(A_k, P)) \right| < c_{II} N^{-\left(\frac{20}{37} - \epsilon\right)}$$

or equivalently

$$\left| \sum_P \mathbf{P}(P \mid N; \phi) \left[\widetilde{W}_i(BL(A_j, P)) - \widetilde{W}_i(BL(A_k, P)) \right] \right| < c_{II} N^{-\left(\frac{20}{37} - \epsilon\right)}. \tag{H.10}$$

Notice that the P term on the left side can only be nonzero if (A_j, P) and (A_k, P) are in different blocks. In fact, it is necessary not only that these two terms be in different blocks but that these blocks have different rounded weights for A_i. We will bound the left side of (H.10) by bounding both the probability of drawing a P for which $BL(A_j, P)$ and $BL(A_k, P)$ have different rounded weights for A_i, and the amount by which these rounded weights can differ.

Specifically, we will show

$$\Pr_{\phi} \left(\widetilde{W}_i(BL(A_j, P)) \neq \widetilde{W}_i(BL(A_k, P)) \right) < (constant) \cdot N^{-\left(\frac{1}{2} - \epsilon\right)} \tag{H.11}$$

and

$$\left| \widetilde{W}_i(BL(A_j, P)) - \widetilde{W}_i(BL(A_k, P)) \right| < (constant) \cdot N^{-\frac{3}{74}} \qquad \text{for each } P. \tag{H.12}$$

First, we prove (H.11).

Define $\Lambda_1, \ldots, \Lambda_M$ from the profile (A_j, P) following the block label algorithm, and put $\Lambda = (\Lambda_1, \ldots, \Lambda_{t-1}) = BL(A_j, P)$. Similarly define $\Lambda'_1, \ldots, \Lambda'_M$ from (A_k, P), and put $\Lambda' = (\Lambda'_1, \ldots, \Lambda'_{t'-1}) = BL(A_k, P)$. Notice that $\Lambda_s = \Lambda'_s$ for each s, except possibly if $s = j$ or $s = k$, in which case we may have $\Lambda'_j = \Lambda_j - 1$ or $\Lambda'_k = \Lambda_k + 1$, respectively.

We consider all the cases in which $\widetilde{W}_i(\Lambda) \neq \widetilde{W}_i(\Lambda')$. There are several possibilities, depending whether the lengths t, t' are different or equal.

(a) It may be that $t < t'$.

(b) It may be that $t > t'$.

If $t = t'$, then we must have $\widetilde{W}_s(\Lambda) \neq \widetilde{W}_s(\Lambda')$ for some $s < t$, which in turn can only happen if $W_s(\Lambda) \neq W_s(\Lambda')$. Since this can occur only for $s = j$ or k, we have two remaining possibilities:

(c) $j < t$ and $\widetilde{W}_j(\Lambda) \neq \widetilde{W}_j(\Lambda')$.

(d) $k < t$ and $\widetilde{W}_k(\Lambda) \neq \widetilde{W}_k(\Lambda')$.

We will deal with each of these cases in turn, and show that the probability of each one is bounded above by a constant times $N^{-(1/2-\epsilon)}$.

(a) If $t < t'$, then $\Lambda_1 + \cdots + \Lambda_t > R$ but $\Lambda'_1 + \cdots + \Lambda'_t \leq R$. This can only happen if $j \leq t$, $\Lambda'_j = \Lambda_j - 1$ and $\Lambda_1 + \cdots + \Lambda_t = R + 1$. We will estimate the probability of these latter two equalities jointly occurring, for any *fixed* value of $t \geq j$.

Write $(A_j, P) = (x_1\ A_1, \ldots, x_M\ A_M)$ as usual. To have $\Lambda'_j = \Lambda_j - 1$ we must have $x_j = L + (\Lambda_j - 1)S$ exactly. We claim that we need only worry about realizations for which $x_j - 1$ is within $2N^{\frac{1}{2}}\sqrt{\ln N}$ of $\phi_j N$. Indeed, using Lemma C.4, the probability of realizing any given value of x_j outside this range is at most

$$e^{-N \cdot \frac{(2N^{-\frac{1}{2}}\sqrt{\ln N})^2}{2}} = e^{-2\ln N} = N^{-2},$$

so the total probablity of realizing all such x_j is at most N^{-1}. Certainly, then, it is sufficient to focus on values of x_j that are within $3N^{\frac{1}{2}}\sqrt{\ln N} > 2N^{\frac{1}{2}}\sqrt{\ln N} + 1$ of $\phi_j N$.

We may also assume that $\underline{\alpha} \leq \phi_j \leq 1 - \underline{\alpha}$. For if $\phi_j < \underline{\alpha}$ and $x_j \leq \phi_j N + 3N^{\frac{1}{2}}\sqrt{\ln N}$, then $x_j < L$ (as long as N is large enough); and if $\phi_j > 1 - \underline{\alpha}$ and $x_j \geq \phi_j N - 3N^{\frac{1}{2}}\sqrt{\ln N}$, then $x_j > L + RS \geq L + (\Lambda_j - 1)S$ (again for large N).

The number of possible values of $x_j = L + (\Lambda_j - 1)S$ that are within $3N^{\frac{1}{2}}\sqrt{\ln N}$ of $\phi_j N$ is at most a constant times $N^{\frac{1}{2}}\sqrt{\ln N}/S < N^{\frac{1}{2}+\epsilon}/S$. Moreover, for each such value, the probability of realizing it is at most a constant times $N^{-\frac{1}{2}}$, by Lemma C.7.

Therefore,

$$\Pr_{\phi}(x_j = L + (\Lambda_j - 1)S) \le (\mathit{constant}) \cdot N^{\epsilon}/S.$$

Now, *conditional* on the value of $x_j = L + (\Lambda_j - 1)S$, the remaining terms x_{-j} are distributed multinomially (by Lemma C.2). What is the probability that $\Lambda_1 + \cdots + \Lambda_t = R + 1$?

Since we are holding fixed the values of x_j and Λ_j, let us denote them by x_j^* and Λ_j^* respectively, while the other x_s and Λ_s follow their corresponding conditional distributions. As long as $\Lambda_j^* < R + 1$, what we are looking is for the probability, under the specified multinomial distribution, that

$$\sum_{\substack{1 \le s \le t \\ s \ne j}} \Lambda_s = R + 1 - \Lambda_j^*.$$

Consider any realization of the profile for which this occurs. If we let Γ be the set of indices s ($1 \le s \le t$, $s \ne j$) such that $\Lambda_s > 0$, then we also have $\sum_{s\in\Gamma} \Lambda_s = R + 1 - \Lambda_j^* > 0$.

Consider *any* possible choice of the nonempty set Γ not containing j, and estimate the probability that $\sum_{s\in\Gamma} \Lambda_s = R + 1 - \Lambda_j^*$ with each $\Lambda_s > 0$, conditional on the value of $x_j = x_j^* = L + (\Lambda_j^* - 1)S$. Since $0 \le x_s - (L + S(\Lambda_s - 1)) < S$ for each $s \in \Gamma$, the desired event can happen only if

$$0 \le \sum_{s\in\Gamma} x_s - (|\Gamma|L + S(R + 1 - \Lambda_j^* - |\Gamma|)) < |\Gamma| \cdot S.$$

This requires that the sum $\sum_{s\in\Gamma} x_s$ — which is binomially distributed — should lie between the lower bound

$$|\Gamma|L + S(R + 1 - \Lambda_j^* - |\Gamma|)$$

and the strict upper bound

$$|\Gamma|L + S(R + 1 - \Lambda_j^* - |\Gamma|) + |\Gamma|S.$$

The lower bound is at least

$$|\Gamma|L + S(1 - M) \geq \frac{\lambda}{2}N$$

when N is large, and the upper bound is at most

$$|\Gamma|L + S(R + 1) \leq ML + (R + 1)S \leq \frac{1 + (M\lambda + \mu)}{2}N$$

when N is large. Therefore, each such realization of $\sum_{s\in\Gamma} x_s$ has probability bounded by a constant times $N^{-1/2}$ by Lemma C.7, and so their total probability is at most

$$|\Gamma| \cdot S \cdot (constant) \cdot N^{-1/2}.$$

Summing over all possible sets Γ (there are certainly at most 2^{M-1} possibilities), we see that

$$\Pr_{\phi}\Bigg(\text{there exists some set } \Gamma \text{ with } \sum_{s\in\Gamma} \Lambda_s = R + 1 - \Lambda_j^*,\ \Lambda_s > 0 \text{ for all } s \in \Gamma, \text{ and } j \notin \Gamma \mid x_j = L + (\Lambda_j^* - 1)S\Bigg) \quad \leq \quad (constant) \cdot S \cdot N^{-1/2} \tag{H.13}$$

for each fixed choice of $\Lambda_j^* < R + 1$.

Therefore,

$$\Pr_{\phi}(\Lambda_1 + \cdots + \Lambda_t = R + 1 \mid x_j = L + (\Lambda_j^* - 1)S) \leq (constant) \cdot S \cdot N^{-1/2}$$

for each fixed choice of $\Lambda_j^* < R + 1$.

Finally,

$$
\begin{aligned}
&\Pr_{\phi}(\Lambda'_j = \Lambda_j - 1 \text{ and } \Lambda_1 + \cdots + \Lambda_t = R+1) \\
&\qquad \le \sum_{\Lambda^*_j} \Pr_{\phi}(x_j = L + (\Lambda^*_j - 1)S \text{ and } \Lambda_1 + \cdots + \Lambda_t = R+1) \\
&\qquad \le \Bigg(\sum_{\Lambda^*_j < R+1} \Pr_{\phi}(x_j = L + (\Lambda^*_j - 1)S) \times \\
&\qquad\qquad \Pr_{\phi}(\Lambda_1 + \cdots + \Lambda_t = R+1 \mid x_j = L + (\Lambda^*_j - 1)S) \Bigg) \\
&\qquad\qquad + \Pr_{\phi}(x_j = L + RS) \\
&\qquad \le \left(\sum_{\Lambda^*_j < R+1} \Pr_{\phi}(x_j = L + (\Lambda^*_j - 1)S) \cdot (\mathit{constant}) \cdot N^{-\frac{1}{2}} S \right) \\
&\qquad\qquad + (\mathit{constant}) \cdot N^{-\frac{1}{2}} \\
&\qquad \le \left(\sum_{\Lambda^*_j} \Pr_{\phi}(x_j = L + (\Lambda^*_j - 1)S) \right) \cdot (\mathit{constant}) \cdot N^{-\frac{1}{2}} S \\
&\qquad\qquad + (\mathit{constant}) \cdot N^{-\frac{1}{2}} \\
&\qquad \le (\mathit{constant}) \cdot (N^{\epsilon}/S) \cdot N^{-\frac{1}{2}} S + (\mathit{constant}) \cdot N^{-\frac{1}{2}} \\
&\qquad \le (\mathit{constant}) \cdot N^{-\left(\frac{1}{2} - \epsilon\right)}.
\end{aligned}
$$

This shows that the total probability of case (a) is at most a constant times $N^{-\left(\frac{1}{2}-\epsilon\right)}$.

(b) If $t > t'$, then $\Lambda_1 + \cdots + \Lambda_{t'} \le R$ but $\Lambda'_1 + \cdots + \Lambda'_{t'} > R$. This can only happen if $k < t$, $\Lambda'_k = \Lambda_k + 1$ and $\Lambda'_1 + \cdots + \Lambda'_{t'} = R + 1$. From here we proceed exactly as in case (a), with Λ and Λ' interchanged, and with the role of j played instead by k. We thus see that the probability of case (b) is also at most a constant times $N^{-\left(\frac{1}{2}-\epsilon\right)}$.

(c) Suppose $j < t$. If $\widetilde{W}_j(\Lambda) \ne \widetilde{W}_j(\Lambda')$, it must certainly happen that $W_j(\Lambda) \ne W_j(\Lambda')$, which requires $\Lambda_j \ne \Lambda'_j$. As in (a), this requires $\Lambda'_j = \Lambda_j - 1$ and $x_j = L + (\Lambda_j - 1)S$ exactly. Also as in (a), we need only worry about values of x_j that are within $3N^{1/2}\sqrt{\ln N}$ of $\phi_j N$, because the total probability of all other values of x_j is at most N^{-1}. Note that

$$|x_j - \phi_j N| \le 3N^{\frac{1}{2}}\sqrt{\ln N}$$

is equivalent to

$$\left|\Lambda_j - \frac{\phi_j N - L + S}{S}\right| \leq \frac{3N^{\frac{1}{2}}\sqrt{\ln N}}{S}. \tag{H.14}$$

However, $\Lambda'_j = \Lambda_j - 1$ implies $W_j(\Lambda') = W_j(\Lambda) - 1/(R+1)$, and therefore

$$2^h W_j(\Lambda') = 2^h W_j(\Lambda) - \frac{2^h}{R+1}.$$

For the corresponding rounded weights to differ, that is to say equivalently

$$\lfloor 2^h W_j(\Lambda) \rfloor \neq \lfloor 2^h W_j(\Lambda) \rfloor,$$

it must be that

$$K \leq 2^h W_j(\Lambda) < K + \frac{2^h}{R+1}$$

for some integer K. Writing this in terms of Λ_j, we have

$$K \leq \frac{2^h}{R+1}\Lambda_j < K + \frac{2^h}{R+1}.$$

Now, for each integer K, we get exactly one choice of Λ_j that satisfies this. Moreover, the difference between any two such possible values of Λ_j (corresponding to different K's) is at least

$$\left\lfloor \frac{1 - 2^h/(R+1)}{2^h/(R+1)} \right\rfloor \geq \left\lfloor (constant) \cdot \frac{N/S - 1}{S^{1/6}} \right\rfloor \geq \left\lfloor (constant) \cdot N^{\frac{53}{74}} \right\rfloor.$$

For N sufficiently large, this is bigger than the width of the window in (H.14), since the latter is

$$3\frac{2N^{1/2}\sqrt{\ln N}}{S} \lesssim N^{1/2}.$$

Therefore, for N sufficiently large, no matter what ϕ is, there is at most one possible value of Λ_j — call it Λ^*_j — that falls in the window (H.14) and allows $\widetilde{W}_j(\Lambda) \neq \widetilde{W}_j(\Lambda')$.

We also know that a realization of this case requires $\Lambda_j > 0$ (since $\Lambda'_j = \Lambda_j - 1$), and $\Lambda_j \leq R$ (since $j < t$). Thus, using the same arguments as in case (a), $(L + (\Lambda^*_j - 1)S)/N$ is bounded strictly between 0 and 1, and so the probability of realizing $x_j = L + (\Lambda^*_j - 1)S$ is bounded by a constant times $N^{-\frac{1}{2}}$.

In summary: for case (c), to happen, either (H.14) must be violated, which happens with probability at most N^{-1}; or we must have $x_j = L + (\Lambda_j^* - 1)S$ for a specific value $0 < \Lambda_j^* \leq R$ (although this value may depend on ϕ), which happens with probability at most a constant times $N^{-\frac{1}{2}}$. This shows that the total probability of case (c) is at most a constant times $N^{-\frac{1}{2}}$.

(d) For this case to happen, we must have $\Lambda_k = \Lambda'_k - 1$. From here we proceed exactly as in (c), with the roles of Λ and Λ' interchanged, and the role of j played by k.

This covers all four cases (a)-(d), completing the proof of (H.11).

Next we prove (H.12). We retain the notation Λ, Λ', and so forth from the proof of (H.11). We regard j, k, P as fixed, and prove that (H.12) holds for every possible choice of $i = 1, \dots, M$.

First we will show the analogue for the unrounded weights:

$$|W_i(\Lambda) - W_i(\Lambda')| < (constant) \cdot N^{-\frac{3}{74}}. \tag{H.15}$$

To show this, suppose first that $i < \min\{t, t'\}$. Since Λ_i and Λ'_i can differ by at most 1 (with a difference possible only when $i = j$ or k), then

$$|W_i(\Lambda) - W_i(\Lambda')| = \left| \frac{\Lambda_i - \Lambda'_i}{R+1} \right| \leq \frac{1}{R+1} < (constant) \cdot N^{-\frac{28}{37}}$$

which is a much stronger bound than $N^{-\frac{3}{74}}$. And if $i > \max\{t, t'\}$, then $W_i(\Lambda) = W_i(\Lambda') = 0$.

Thus, if $t = t'$, we have proven (H.15) for all i except possibly $i = t$. But then since

$$\sum_{i=1}^{M} (W_i(\Lambda) - W_i(\Lambda')) = 1 - 1 = 0, \tag{H.16}$$

the fact that (H.15) holds for all $i \neq t$ means it holds for $i = t$ as well.

Now suppose $t < t'$. As in case (a) of the proof of (H.11), this implies $j \leq t$, $\Lambda'_j = \Lambda_j - 1$ and $\Lambda_1 + \dots + \Lambda_t = R + 1$, whereas $\Lambda'_s = \Lambda_s$ for every $s \leq t$, $s \neq j$. Hence $\Lambda'_1 + \dots + \Lambda'_t = R$, and then $\Lambda'_s = 0$ for all $t < s < t'$ (because otherwise we would have $\Lambda'_1 + \dots + \Lambda'_s > R$ contradicting the minimality of t').

Since $\Lambda_1 + \dots + \Lambda_t = R + 1$, we have

$$W_t(\Lambda) = 1 - \frac{\sum_{s<t} \Lambda_s}{R+1} = \frac{\Lambda_t}{R+1}$$

while also $W_t(\Lambda) = \Lambda'_t/(R+1)$, so the same argument used to verify (H.15) above for $i < t$ also holds for $i = t$. And if $t < i < t'$ then $W_i(\Lambda) = 0 = W_i(\Lambda')$. Thus, we have now shown (H.15) for all $i \neq t'$. By (H.16), it holds for $i = t'$ as well.

This proves (H.15) for the case $t < t'$. The case $t > t'$ is identical, with the roles of Λ and Λ' interchanged and k in place of j. Thus (H.15) is proven in all cases.

Moreover, for all Λ and all i, we have

$$\left| W_i(\Lambda) - \widetilde{W}_i(\Lambda) \right| < (constant) \cdot N^{-\frac{3}{74}}. \tag{H.17}$$

Indeed, if $i < t$, then the definition of $\widetilde{W}_i(\Lambda)$ implies

$$0 \leq W_i(\Lambda) - \widetilde{W}_i(\Lambda) < \frac{1}{2^h} = \frac{1}{S^{\frac{1}{6}}} < (constant) \cdot N^{-\frac{3}{74}}.$$

If $i > t$ then $\widetilde{W}_i(\Lambda) = W_i(\Lambda) = 0$. And now that we have (H.17) for all $i \neq t$, the identity

$$\sum_{i=1}^{M} \Big(W_i(\Lambda) - \widetilde{W}_i(\Lambda) \Big) = 1 - 1 = 0$$

implies that it holds for $i = t$ as well.

Combining (H.15), (H.17), and another application of (H.17) with Λ' in place of Λ, we get (H.12), as claimed.

Now we can prove (H.10). Let Ω be the set of all N-profiles P for which $\widetilde{W}_i(BL(A_j, P)) \neq \widetilde{W}_i(BL(A_k, P))$. We have

$$\begin{aligned}
&\left| \sum_{P} \mathbf{P}(P \mid N; \phi) \left[\widetilde{W}_i(BL(A_j, P)) - \widetilde{W}_i(BL(A_k, P)) \right] \right| \\
&\qquad = \left| \sum_{P \in \Omega} \mathbf{P}(P \mid N; \phi) \left[\widetilde{W}_i(BL(A_j, P)) - \widetilde{W}_i(BL(A_k, P)) \right] \right| \\
&\qquad \leq \Pr_{\phi}(P \in \Omega) \cdot (constant) \cdot N^{-\frac{3}{74}} \\
&\qquad\qquad \text{by (H.12)} \\
&\qquad \leq (constant) \cdot N^{-\left(\frac{1}{2} - \epsilon\right) - \frac{3}{74}} \\
&\qquad\qquad \text{by (H.11).}
\end{aligned}$$

This gives (H.10), and so Claim II is proven.

Completion of Proof of Theorem 4.3. Suppose the manipulator has belief ϕ and considers a change in his vote from A_j to A_k. We show that this manipulation can change the probability of any candidate A_i winning by (asymptotically) no more than a constant times $N^{-\left(\frac{20}{37}-\epsilon\right)}$. We have

$$\left|\Pr_\phi(f(A_j,P)=A_i)-\sum_\Lambda \Pr_\phi((A_j,P)\in BL^{-1}(\Lambda))\widetilde{W}_i(\Lambda)\right| \lesssim c_I N^{-\left(\frac{20}{37}-\epsilon\right)}$$

by Claim I;

$$\begin{aligned}\left|\sum_\Lambda \Pr_\phi((A_j,P)\in BL^{-1}(\Lambda))\widetilde{W}_i(\Lambda)-\sum_\Lambda \Pr_\phi((A_k,P)\in BL^{-1}(\Lambda))\widetilde{W}_i(\Lambda)\right|& \\ \lesssim \quad c_{II} N^{-\left(\frac{20}{37}-\epsilon\right)}&\end{aligned}$$

by Claim II;

$$\left|\sum_\Lambda \Pr_\phi((A_k,P)\in BL^{-1}(\Lambda))\widetilde{W}_i(\Lambda)-\Pr_\phi(f(A_k,P)=A_i)\right| \lesssim c_I N^{-\left(\frac{20}{37}-\epsilon\right)}$$

by Claim I again. The triangle inequality then gives

$$\left|\Pr_\phi(f(A_j,P)=A_i)-\Pr_\phi(f(A_k,P)=A_i)\right| \lesssim (2c_I+c_{II}) N^{-\left(\frac{20}{37}-\epsilon\right)}.$$

The theorem follows, with (say) $\kappa = 20/37 - 2\epsilon$.

□

References

[1] Milton Abramowitz and Irene A. Stegun (1972), *Handbook of Mathematical Functions* (Washington: U.S. Government Printing Office), tenth printing.

[2] Fuad Aleskerov and Eldeniz Kurbanov (1999), "Degree of Manipulability of Social Choice Procedures," in Ahmet Alkan, Charalambos D. Aliprantis, and Nicholas C. Yannelis, eds., *Current Trends in Economics: Theory and Applications* (Berlin, Heidelberg: Springer-Verlag).

[3] Wassily Hoeffding (1963), "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association* 58 (301), 13-30.