留下还是更换？在线候选人到达下的度量失真

Kaito Fujii[1],Yeeseok Oh[2], 和 Shinsaku Sakaue[3]

[1]京都大学[2]东京大学[3]
CyberAgent、NII、RIKEN
AIP

arXiv:2611.30683v1 [cs.GT] 23 Nov 2026

摘要

我们研究在线候选人到达的序贯选举中的度量失真。在每一轮，新到达的挑战者与现任者竞争，投票规则必须根据选民的成对偏好从两者中选出一人。我们的目标是确保每一轮的度量失真都较低，相对于迄今为止所有已到达候选人中的最佳候选人而言。对于确定性规则，我们证明了第 t轮的渐近最优界为 $\Theta(\sqrt{t})$，该界由偏向现任者的阈值规则实现。相比之下，随机化可实现常数失真。我们证明，持久随机独裁在面对观察此前获胜者的自适应对手时，可实现最优失真 5；在面对不观察此前获胜者的无知对手时，可实现最优失真 3。最后，我们从可用信息和行动空间两个维度推广该模型，并为每种情形建立失真保证。

目录

1 引言 **2**
1.1 我们的贡献 .21.2 相关工作 .3

2 模型 **4**
2.1 预备知识 .42.2 在线候选人到达情况下的度量失真 .5

3 部分信息下的在任者—挑战者模型 **6**
3.1 确定性投票规则 .73.2 随机化投票规则 .10

4 入驻　完全信息下的现任者—挑战者模型 1 **6**
4.1 确定性投票规则 .164.2 随机化投票规则 .20

5 结论 **23**

A 附录：部分信息下的完整候选人模型 27

1. 引言

回到罗马共和国时期，公民会投票决定是否接受一项拟议法律。他们可以选择按提议执行 （“照你的提议执行”）或保留旧法 （“我保留旧法”）。对于单一政策或议题，替换旧法的提案和投票会在多年间接连举行。例如，罗马关于公共土地的立法在公元前133年至111年间经历了连续改革。这种集体决策方式在各种政治体系中一直很常见，包括美国在法律变更方面的做法。提出新法案后，国会会投票决定是否采纳该法案，或保留现状。

除立法决策外，在现代选举中，在任者—挑战者选举的结构也是一种非常普遍的现象。一个近期例子来自2026年美国中期选举：参议员苏珊·柯林斯在缅因州自1996年以来连续赢得五次参议院选举后，最终在竞选连任中落败。这类例子表明，现任者可能在接连不断的挑战中持续胜出，直到最终被取代。受这种常见的序贯投票结构启发，我们研究在线候选人到达下的集体决策，其中只有当前现任者和新到达的挑战者竞争当选。此类现任者与挑战者选举能否保证在每一轮都选出一位优秀的代表?

1.1. 我们的贡献

我们开启了对在线候选人到达下度量失真的研究。[1] 在每一轮中，一位新候选人到达，我们必须为该轮选出一位获胜者。

我们从行动空间和信息结构两个维度考察限制条件。首先，我们考虑行动空间。正如引言中所述，我们引入了在任者—挑战者 （IC）模型，在该模型中，每一轮我们只能选出现任者或挑战者。相比之下，在完整候选人空间中，我们可以选出在先前轮次中落败的候选人。

接下来，我们考察信息结构。对于完全信息设置，我们假设所有选民对迄今为止已到达的候选人的排名都已公开。相比之下，在部分信息设置中，我们只能观察选民对现任者和挑战者的偏好。这些设置在第 2 节中有完整说明。

基于这两个维度，我们的结果总结于表1中。我们考察确定性规则和随机化规则。对于随机化规则，根据对手能否即时观察之前的获胜者，我们考察两种不同的对手模型：自适应对手和无知对手。

我们主要关注部分信息下的现任者—挑战者模型。我们认为，这一模型刻画了一类自然的重复二元选举。第 3 节推导了该情形下的界，对应于表1的最底行。对于确定性规则，我们展示了一种偏好现任者的规则，其失真达到 $O(\sqrt{T})$，并证明这是渐近最优的。转向随机化规则，我们证明，对于自适应对手和无知对手，持久随机独裁分别实现常数度量失真 5 和 3。我们进一步证明这些界是最优的，从而证明持久随机独裁在度量失真方面的最优性。这给出了右下角单元格中的界。此外，我们还证明，允许将完整候选者集合用作行动空间并不能改善这些界，从而得到左下角单元格。

在第4节中，我们转向完全信息的情形，对应于表1的顶行。在完整候选集设置下，这种情况与静态度量失真一致，因此其界可由经典结果得出（Anshelevich 等人，2018年；Gkatzelis 等人，2020年；Charikar 和 Ramakrishnan，2022年；Cai 等人，2026年）。这给出了左上单元格中的界。对于

[1] 一项独立的同期工作研究了类似的模型；参见相关工作部分。

信息 \ 行动	完整候选者		在任者—挑战者	
full	(D)	$\Delta^* = 3$	(D)	$\Delta^* = 2+\sqrt{5} \approx 4.2361$
	(自适应随机)	$2.1126 \leq \Delta^* \leq 2.1248$	(自适应随机)	$3 \leq \Delta^* \leq 2+\sqrt{5}$
	(无知随机)	$2.1126 \leq \Delta^* \leq 2.1248$	(无知随机)	$\Delta^* = 3$
部分	(D)	$\Delta^* = \Theta(\sqrt{T})$	(D)	$\Delta^* = \Theta(\sqrt{T})$
	(自适应随机)	$\Delta^* = 5$	（自适应随机）	$\Delta^* = 5$
	(无知随机)	$\Delta^* = 3$	（无知随机）	$\Delta^* = 3$

表1：各种模型下度量失真界的结果汇总。 Δ^* 表示最优最坏情况失真保证。（D）表示确定性投票规则，（自适应随机）表示面对自适应对手的随机化投票规则，（无知随机）表示面对无知对手的随机化投票规则。

IC 模型中，与部分信息设置不同，确定性规则可以实现 $2+\sqrt{5}$ 的常数度量失真。这意味着，对于自适应对手，持久随机独裁不再是最优的。然而，对于无知对手，持久随机独裁再次成为最优方案。这些结果给出了右上单元格中的界。

1.2. 相关工作

与我们的工作独立且同时，Ghasemi等人（2026年研究了候选人依次到达的在线投票。他们的模型对应于我们的完全信息现任者—挑战者模型。Ghasemi等人还推导出 $2+\sqrt{5}$ 确定性规则的上界，以及随机化规则的上界3。他们证明了确定性下界为 $\sqrt{13}$，对于三名候选人而言这是最优的。对于随机化算法，他们沿用了静态下界。这些结果使两个最优界仍未确定。此外，随机化保证是在无知对手下建立的。在我们的工作中，我们明确考虑了无知对手和自适应对手，并分别推导出针对两类对手的不同界。此外，我们弥合了他们针对完全信息现任者—挑战者模型这一特定模型所得结果中的差距，证明 $2+\sqrt{5}$ 是确定性规则的最优值，而3是无知随机化规则的最优值。我们还刻画了不同模型下的最优失真。他们的工作还研究了功利主义失真，而不仅仅是度量失真。

社会选择中的失真。失真这一概念最早由Procaccia 和 Rosenschein (2006年) 提出，用于衡量因观察到序数偏好而非基数效用所造成的社会福利损失。后续研究探讨了确定性和随机投票规则能够达到的失真（Caragiannis 和 Procaccia，2011年；Boutilier 等人，2015；Ebadian 等人，2024年）；综述请参见Anshelevich 等人(2021)。

将失真与空间投票模型联系起来，Anshelevich 等人 (2018) 开启了对度量失真的研究，证明了确定性下界为 3、上界为 5。Munagala 和 Wang (2019年) 将上界改进为 $2+\sqrt{5}$，所使用的是加权未覆盖集。随后，Gkatzelis 等人 (2020) 确立了最优确定性度量失真为 3。Kizilkaya 和 Kempe (2022年) 提出了一种简单的多数否决替代规则，该规则同样达到了 3。Anshelevich 和 Postl (2017) 转而研究随机化规则，并表明随机独裁实现了 3 的度量失真。随后一系列工作缩小了目前已知随机化下界与上界之间的差距（Charikar 和

Ramakrishnan，2022年；Charikar 等人，2024年；Cai 等人，2026年）。

不同选票下的度量失真。关于度量失真的文献考察了不同类型的选票，其中许多选票的信息受到限制。一系列研究探讨了加权锦标赛规则，即 C2规则，在这类规则中，只能观察到两两投票计数。Munagala 和 Wang(2019证明了可以达到 $2+\sqrt{5}$ 的度量失真，后续研究进一步改进了界（Charikar 等人，2025；Lin 和 Wang，2026）。研究者也考察了基于认可的偏好下的失真（Pierczy´nski 和 Skowron，2019）。Anshelevich 等人(2025)考察了额外的阈值认可信息，而Frank 和 Peters(2026)则考察了相对多数投票中的认可。我们的部分信息设置同样限制了选票，但方式有所不同：比较信息会随着候选人的到来而逐步揭示，因此先前的选择可能限制后续轮次中可获得的信息和可采取的行动。

现任者与挑战者结构。经济学和政治学文献对采用现任者与挑战者结构的序贯选举进行了长期研究。现任者与挑战者结构被作为修正案程序进行研究，其历史可追溯至Black (1948)，此后又有大量研究（Riker，1958；Farquharson，1969；Miller，1977；McKelvey 和 Niemi，1978；Banks，1985）。这些研究传统上从议程操纵、策略性投票和均衡分析的角度研究这一模型。Bredereck 等人(2017)研究了这些问题的计算复杂性。

将在任者—挑战者结构视为在线过程，并通过竞争分析或遗憾分析来研究，似乎仍鲜有人涉足。我们考虑一种基于度量失真的竞争比方法来研究该模型。

在线社会选择。 Oren 和 Lucier (2014年) 研究了一种在线设置：选民以在线方式到达，决策者必须决定是否将某位候选人永久加入委员会。 Dong 和 Peters(2025年) 研究了候选人到达和离开情况下的多赢家投票。在序数、认可投票和度量设置中，他们提出了在候选人集合变化过程中维持比例性的算法。他们还证明了不可能性结果，并通过少量替换的算法处理相关情形。

2. 模型

2.1. 预备知识

对于正整数 t，令 $[t]=\{1,\dots,t\}$。选举实例 $E=(N,\ C,\ \succ)$ 由有限的选民集合 $N=[n]$、有限的候选人集合 C，以及偏好剖面 $()$组成，其中每个 都是候选人 上的严格排序。投票规则是一个函数，它以选举实例 $(,\ C,\ \succ)$ 为输入，并输出获胜者 $f(\succ)\in C$。

空间投票模型。在空间投票模型中，集合 N 和 C 位于同一个度量空间中。换言之，存在一个（伪）度量[2] d 定义在 $N\cup C$ 上，满足$d(x,x)=0$, $d(x,y)=d(y,x)$，且 $d(x,y)\le d(x,z)+d(z,y)$，对所有 $x,y,z\in N\cup C$均成立。选民对候选人的偏好由其到各候选人的距离决定，他们更偏好距离较近的候选人。给定定义在 $N\cup C$上的（伪）度量 d，如果对每个 $i\in N$, $x\succ_i y$，$d(i,x)\le d(i,y)$，则称偏好剖面$\succ=(\succ_i)_{i\in N}$由 d 诱导。在这种情况下，我们也称 d 与 $\succ$ 一致，并记作 $d\triangleright\succ$。

[2]在全文中，我们将伪度量和度量这两个术语互换使用。

度量失真。给定定义在 $N \cup C$上的度量 d，候选人 $c \in C$ 的社会成本 为 $\mathrm{sc}_d(c) = \sum_{i \in N} d(i, c)$。当 d 的含义明确时，我们省略下标 d。于是，我们可以考虑选择一名使社会成本最小的候选人这一问题。可以将投票程序视为旨在最小化社会成本。然而，由于选举实例输入的是序数选票，而非底层度量信息，我们无法优化社会成本。度量失真 衡量仅凭序数信息所能达到的相对于最优社会成本的近似比。

形式化地，给定一个选举实例 $\succ$ 和一位候选人 $x \in C$，x 在 $\succ$上的度量失真定义为

$$\max_{d \rhd \succ} \frac{\sum_{i \in N} d(i, x)}{\min_{y \in C} \sum_{i \in N} d(i, y)} = \max_{d \rhd \succ} \frac{\mathrm{sc}(x)}{\min_{y \in C} \mathrm{sc}(y)},$$

其中，对于 $a > 0$，$0/0 = 1$ 和 $a/0 = +\infty$。对于投票规则 f，该规则 f的度量失真定义为

$$\Delta(f) = \sup_{(N, C, \succ)} \max_{d \rhd \succ} \frac{\mathrm{sc}(f(\succ))}{\min_{y \in C} \mathrm{sc}(y)},$$

也就是说，所产生社会成本的最坏情况近似比。对于 $f(\succ)$ 是 C上的概率分布的随机化投票规则，度量失真定义为

$$\Delta(f) = \sup_{(N, C, \succ)} \max_{d \rhd \succ} \frac{\mathbb{E}[\mathrm{sc}(f(\succ))]}{\min_{y \in C} \mathrm{sc}(y)}.$$

我们可以通过考虑以下对手模型来理解这一数值：

1. 对手选择 N、C、定义在 $N \cup C$上的度量 d 以及偏好剖面 $\succ$，使得$d \rhd \succ$。
2. 投票规则 f 选择获胜者 $f(\succ) = x \in C$。
3. 社会的成本为 $\sum_{i \in N} d(i, f(\succ))$。

我们将这一视角发展为一种在线模型。

2.2. 在线候选人到达下的度量失真

选举按轮次反复举行，轮次为 $t = 1$、2、.....、T。每一轮 t，都会有一位新候选人 c_t到达。我们称 c_t第 t 轮的挑战者 。用 $C_t = \{c_1, \ldots, c_t\}$ 表示截至第 t 轮已到达的候选人。用 x_t 表示第 t 轮的获胜者。我们称 x_{t-1} 第 t 轮的现任者 。注意，$x_1 = c_1$ 是固定的。

我们考虑在第 $t \geq 2$轮获得的信息或选票的两种模型。在完全信息 设定下，每位选民 $i \in N$都会报告对 $C_t = \{c_1, \ldots, c_t\}$的排名。在部分信息 设定下，每位选民 $i \in N$ 都会报告对 $\{x_{t-1}, c_t\}$的排名。类似地，我们考虑两种基于候选人集合的行动模型：X_t 是我们在第 t 轮可以从中选出获胜者的集合。在完整候选者 模型中，$X_t = C_t = \{c_1, \ldots, c_t\}$。在在任者—挑战者 （IC）模型中，$X_t = \{x_{t-1}, c_t\}$。

这些组合总共构成四种可能的选举结构。对于每种组合，我们考虑以下问题。

1. 对手选择 N、c_1、$N \cup C_1$ 上的度量 d，以及轮数 T .
2. 对于 $t = 2, \ldots, T$:

   a）对手放置 c_t，将 d 扩展为定义在 $N \cup C_t$ 上的内容，并将之前的偏好剖面 $\succ_{t-1}$ 扩展为偏好剖面 $\succ_t$，使得 $d \rhd \succ_t$。

b）投票规则 f 根据截至目前获得的信息选出获胜者 $x_t \in X_t$。c）社会承担的社会成本为 $\mathrm{sc}(x_t) = \sum_{i \in N} d(i, x_t)$。

请注意，选民之间的距离不能随轮次变化，任何已放置点之间的距离也不能改变。对于每一轮 t，我们考虑相对于 C_t（截至目前的候选人）的度量失真：

$$\frac{\mathrm{sc}(x_t)}{\min_{y \in C_t} \mathrm{sc}(y)}.$$

我们的目标是保证每一轮的度量失真都较低。因此，对于一个输出 x_1、……、x_T 的（确定性）规则 f，f 的度量失真定义为所有轮次 $t = 1$、……、T 中的最坏情况失真：

$$\Delta(f) = \sup_{(N,C,\succ),T} \max_{t \in [T]} \max_{d \rhd \succ} \frac{\mathrm{sc}(x_t)}{\min_{y \in C_t} \mathrm{sc}(y)}.$$

一个直接的观察是，在完全信息和完整候选人空间下，度量失真界与经典度量失真一致。这是因为我们只需在每一轮运行静态投票规则，即可达到上界，[3] 而静态实例所给出的同一下界，则可通过以任意顺序引入候选人推导得出。这给出了 表1 的左上单元格。

对手。由于问题具有序贯性质，我们可以考虑两类对手。在第 t 轮选择 c_t 后，自适应对手会根据此前选出的获胜者 x_1、……、x_{t-1} 的信息选择 c_t。无知对手则在不观察此前获胜者的情况下选择 c_t。

对于确定性规则，这两种对手并无区别。无知对手可以直接模拟投票规则 f，以预测获胜者 x_1、……、x_{t-1}。然而，对于随机化规则，度量失真可能因对手模型而异。对于无知对手，底层度量在规则实现之前预先选定，因此度量失真为

$$\Delta(f) = \sup_{(N,C,\succ),T} \max_{t \in [T]} \max_{d \rhd \succ} \frac{\mathbb{E}[\mathrm{sc}(x_t)]}{\min_{y \in C_t} \mathrm{sc}(y)}.$$

然而，对于自适应对手，c_t 的位置可以取决于 x_1、……、x_{t-1}先前的实现，这意味着第 t 轮的最优候选人 $y_t^* \in \mathrm{argmin}_{y \in C_t} \mathrm{sc}_d(y)$ 可能会因 x_1、……、x_{t-1} 的实现而变化。也就是说，y_t^*本身是一个随机变量。因此，我们将度量失真定义为

$$\Delta(f) = \sup_{\mathcal{A}} \max_{t \in [T]} \frac{\mathbb{E}[\mathrm{sc}(x_t)]}{\mathbb{E}[\mathrm{sc}(y_t^*)]},$$

其中，上确界是针对自适应对手 $\mathcal{A}$取的。

3. 部分信息下的现任者—挑战者模型

我们考虑在部分信息条件下的在任者—挑战者行动模型。我们认为，这是现实生活中最常见的情形。

[3] 对于 ra 针对自适应对手的随机静态规则，每轮明确使用新的随机硬币 .

3.1. 确定性投票规则

一种自然的投票规则是在每一轮都采用多数规则。对于只有一轮且有两位候选人的情形，多数规则可实现 3 的失真，且这是最优的。然而，对于我们的序贯模型，失真会随轮数线性增长。我们首先介绍一个引理，它是以下内容的直接推论：Kempe (2020)。

引理3.1 设 a_1、…、a_m、a' 为候选人，使得对于每个 $j=2$、…、m，至少有一半的选民更偏好候选人 a_{j-1} 而非 a_j。设 $\alpha=\frac{1}{n}|\{i\in N:\neg\, a'\prec_i a_m\}|>0$。则，

$$\mathrm{sc}(a_1)\le\left(2m+\frac{2}{\alpha}-3\right)\mathrm{sc}(a').$$

特别地，如果 $\alpha\ge 1/2$，那么 $\mathrm{sc}(a_1)\le(2m+1)\mathrm{sc}(a')$。

证明。对于 $m=1$，根据推论5.3（Kempe，2020），该情形显然成立。对于 $m\ge 2$，令 p_j($2\le j\le m$) 表示更偏好 a_{j-1} 而非 a_j的选民比例。使用推论5.1（Kempe，2020）中的记号，我们有数 $\lambda_1=1$、$\lambda_2=\frac{2}{p_2}-1\le 3$、$\lambda_j=\frac{2}{p_j}\le 4$。于是 Λ，在（λ_1、…、λ_m、$2/\alpha$）中不含两个连续数的数之和的最大值满足 $\Lambda\le 2m+\frac{2}{\alpha}-3$，因此结论得证。 □

命题 3.2 在部分信息 *IC* 模型中，逐轮多数规则实现的度量失真至多为 $2T-1$，且这一界是紧的。

证明。在每一轮 t，从多数获胜者指向失败者，在 c_t 和 x_{t-1} 之间添加一条有向边。于是得到的底层无向图是一棵树，因为它是连通的，并且有$T-1$ 条边和 T 个顶点。注意，t 轮的现任者 x_t 到所有截至 t轮已经到达的候选人 $\{c_1,\dots,c_t\}\setminus\{x_t\}$ 都有一条有向路径。对此可以对 t进行如下归纳证明。如果 $x_t=x_{t-1}$，则存在一条有向边 $x_{t-1}c_t$，而其他候选人根据归纳假设是可达的。如果 $x_t=c_t$，则存在一条有向边 c_tx_{t-1}，而其他候选人可通过 x_{t-1}到达。因此，对于每个 $i\in[t]$，都存在一条从 x_t 到 c_i 的有向路径，其长度至多为 $t-1$。根据引理3.1，对于所有 $i\in[t]$和所有 $t\in[T]$，我们有 $\mathrm{sc}(x_t)\le(2t-1)\,\mathrm{sc}(c_i)$。

现在我们给出一个匹配的下界示例。考虑以下实例。首先，将 $2M$名选民放置在一条直线上：M名选民位于 0，M名选民位于 1。将候选人 $c_1,\dots,c_T$放置在 0、2、−2、4、−4、…，也就是说，若 t为偶数，则将 c_t放置在 t；若 t为奇数，则将 $1-t$放置在那里。最后，放置一名额外选民 w，使其到直线上每个选民或候选人的距离均为 T。注意，所得度量是一个有效的（伪）度量：对于任意直线上的点 y，z，$d(y,z)\le 2T=d(w,y)+d(w,z)$ 以及$d(w,y)=T\le d(w,z)+d(z,y)$。

对于每位选民，若距离相同，则优先选择到达较晚的候选人。现在注意，在每一轮 t，挑战者 c_t 都会获胜：位于 0 的 M 名选民（若 t 为奇数），或位于 1 的 M 名选民（若 t 为偶数），再加上选民 w，总是更偏好 c_t 而非 x_{t-1}。然而，我们有 $\mathrm{sc}(c_T)=M(2T-1)+T$，且 $\mathrm{sc}(c_1)=M+T$。因此，由于 $M\to\infty$，

$$\frac{\mathrm{sc}(c_T)}{\mathrm{sc}(c_1)}\to 2T-1.$$ □

我们能否获得更低的失真？事实证明，通过以有利于现任者的方式调整挑战者获胜所需的票数门槛，我们可以将失真降至 $O(\sqrt{T})$。此外，这在 T 中是渐近最优的。

定理 3.3 设 $(\alpha_t)_{t\ge 2}$ 为一个非递增序列，满足 $0<\alpha_t\le 1/4$。在 算法 *1*下，对于每个 $t\ge 2$ 和 $j\le t$，

$$\frac{\mathrm{sc}(x_t)}{\mathrm{sc}(c_j)}\le\frac{2}{\alpha_t}+8\sum_{s=2}^{t}\alpha_s+1.$$

Alg算法 1 α-阈值 ma　　jorit

y输入：一个包含 T 轮次的在线投票实例，以及一个满足$0 < \alpha_t \le 1/4$的非递增序列（α_t）$_{t\ge 2}$ **输出**：一个获胜者序列 x_1，……，x_T，满足：对于每个 $t \ge 2x_1 \leftarrow c_1$，**对于** $t = 2$，…… ，T**执行**$q_t \leftarrow \frac{1}{n}|\{i \in N:\ x_{t-1} \succ_i c_t\}|$**如果** $q_t \le \alpha_t$ **则**$x_t \leftarrow c_t$**否则**$x_t \leftarrow x_{t-1}$**返回** $(x_1$，……，$x_T)$

证明。 如下构造加权有向图。在第 t 轮，从所选候选人 x_t 向失败者添加一条有向边。若 $x_t = c_t$，则称该边为切换边，并赋予权重$1 - q_t$。若 $x_t = x_{t-1}$，则称该边为拒绝边，并赋予权重 q_t。也就是说，我们将权重设为偏好 x_t 而非失败者的选民比例。

与 命题 3.2中一样，该图是一棵有向树，且存在一条从x_t 到每个候选人 c_j的有向路径，$j \in [t]$。此外，从 x_t 到 $c_j \ne x_t$ 的有向路径由连续的切换边构成，最后一条边也可能是拒绝边。这是因为拒绝边的终点会成为叶节点。

固定一轮 $t \ge 2$ 和 $j \le t$，并考虑有向 x_t-c_j 路径。设x_t-c_j路径上的切换边按路径顺序为 e_1、……、e_ℓ。我们按如下方式将边划分为连续边块：对于切换边 e_p，令其权重为 $1 - \alpha'_p$。贪心地尽可能延长每个区块，但要求区块中所有边 p 上的 α'_p 之和不超过 $1/2$。令 R 为区块数。（若不存在切换边，则令 $R = 0$。）

注意，对于权重为 ℓ 的连续切换边 $1 - \alpha'_1$、$1 - \alpha'_2$、……、$1 - \alpha'_\ell$，根据并集界，至少有 $1 - \sum_{p=1}^{\ell} \alpha'_p$ 比例的选民支持所有边。因此，如果$\sum_{p=1}^{\ell} \alpha'_p \le 1/2$，根据传递性，至少半数选民更偏好源候选人而非汇候选人。因此，我们可以将每个区块替换为权重为 $\ge 1/2$ 的多数边。这样便得到一条包含 x_t-c_j 条路径，其中有 R 条权重为 $\ge 1/2$ 的边，并且可能有一条权重为$q_j \ge \alpha_j$ 的最终拒绝边。由于 $j \le t$，因此 $\alpha_j \ge \alpha_t$。因此，根据 引理3.1，我们有

$$\mathrm{sc}(x_t) \le \left(2R + \frac{2}{\alpha_t} - 1\right) \mathrm{sc}(c_j).$$

最后，我们对 R 作界定。除最后一个区块外，每个区块都是极大的：若加入下一条边的 α，α之和就会超过 $1/2$。由于每个 α 至多为 $1/4$，每个非末尾区块的 α 之和至少为 $1/4$。因此，原始 x_t-c_j路径中的切换边的 α 之和至少为 $(R-1)\cdot\frac{1}{4}$。由于 α 之和有上界 $\sum_{s=2}^{t} \alpha_s$，我们有 $R \le 4\sum_{s=2}^{t} \alpha_s + 1$。因此，我们达到了该界。 □

通过对所有 t 设置 $\alpha_t = 1/(4\sqrt{T})$，我们可以得到 $O(\sqrt{T})$ 的度量失真。事实上，即使不知道时域 T，我们也能在第 t 轮次保证 $O(\sqrt{t})$ 的失真。

推论 3.4 算法 *1* 在 $\alpha_s = 1/(4\sqrt{s})$ 时，在第 $O(\sqrt{t})$ 轮次达到 t 的度量失真，适用于部分信息 *IC* 模型。

证明。令 $\alpha_s = 1/(4\sqrt{s})$。于是，

$$\begin{aligned}\frac{2}{\alpha_t} + 8\sum_{s=2}^{t}\alpha_s + 1 &= 8\sqrt{t} + 2\sum_{s=2}^{t}\frac{1}{\sqrt{s}} + 1 \\ &\le 8\sqrt{t} + 2\int_1^t \frac{1}{\sqrt{x}}dx + 1 \\ &= 12\sqrt{t} - 3.\end{aligned}$$

□

此外，该界在渐近意义下是紧的；确定性投票规则无法避免失真以轮数平方根的速度累积。

命题3.5 在部分信息 *IC* 模型中，每个确定性投票规则的度量失真均为$\Omega(\sqrt{T})$。

证明。固定一个确定性规则。对于整数 $k \ge 2$，令 $T = k(k-1) + 1$。考虑按 $c_1, \dots, c_T$ 顺序到达的 k 个选民和 T 个候选人。在第 t 轮，只有选民 $i_t = 1 + ((t-2) \bmod k)$ 更偏好 x_{t-1} 而非 c_t，而所有其他 $k-1$ 个选民都更偏好 c_t。我们根据该规则是否曾选择过现任者，分情况讨论。

情况 1：该规则总是选择挑战者。 定义选民—候选人距离为$d(i, c_t) = 1 + 2\left\lfloor \frac{t-1+k-i}{k} \right\rfloor$。若出现平局，则优先选择较大的 t。那么，对于每个$2 \le t \le T$，

$$d(i, c_t) - d(i, c_{t-1}) = \begin{cases} 2, & i = i_t, \\ 0, & i \ne i_t. \end{cases}$$

因此，选民 i_t 更偏好 c_{t-1} 而非 c_t，而按照指定的平局处理规则，其他每位选民都更偏好 c_t。使用最短路径距离定义其他距离。

让我们验证所得度量满足度量公理。由于我们已在选民与候选人构成的完全二部图上定义了距离，因此只需检查三路径三角不等式：$d(i,c) \le d(i,c') + d(i',c') + d(i',c)$，其中任意 $i, i' \in N$和 $c, c' \in C$均成立。对于每个 t 以及任意一对选民 i, i'，我们都有 $d(i, c_t) \ge 1$ 和 $|d(i, c_t) - d(i', c_t)| \le 2$。因此，$d(i,c) - d(i',c) \le 2 \le d(i,c') + d(i',c')$，三路径三角不等式成立。

现在让我们计算失真。由于 $T - 1 = k(k-1)$，对于每位选民 i，我们有 $d(i, c_1) = 1$且 $d(i, c_T) = 2k - 1$。由于 $x_T = c_T$，我们有

$$\frac{\mathrm{sc}(x_T)}{\mathrm{sc}(c_1)} = \frac{k(2k-1)}{k} = 2k - 1 = \Theta(\sqrt{T}).$$

情形 2：该规则至少有一次选择现任者。 令 t 为该规则首次选择现任者 x_{t-1}的轮次。注意，$x_{t-1} = c_{t-1}$。接下来，考虑以下选民和候选人位于直线上的实例。将候选人 c_t 放在 2 处，并将其他所有候选人$c_s \ne c_t$ 放在 0 处。将选民 i_t 放在 1 处，并将其他所有选民 $i \ne i_t$ 放在 2 处。在以有利方式打破平局、令除 i_s 之外的所有选民都偏好在第 s轮较晚到达的候选人的情况下，会诱导出上述指定的偏好剖面。

由于除 c_t 之外的所有候选人都位于 0 处，且 c_t 从未成为获胜者，因此 x_T 位于 0 处。于是，

$$\frac{\mathrm{sc}(x_T)}{\mathrm{sc}(c_t)} = \frac{2(k-1)+1}{1} = 2k - 1 = \Theta(\sqrt{T}).$$

请注意，对于任意 T，令 $k = \lfloor \sqrt{T} \rfloor$，并将上述论证应用于 $T_0 = k(k-1) + 1$，再补入虚拟候选项，即可得到相同的渐近界。 □

算法 2 持久随机独裁**输入**：一个包含 T 轮次的在线投票实例**输出**：获胜者序列 $x_1, \ldots, x_T$，满足：对于每个 $t \geq 2$，$x_t \in \{x_{t-1}, c_t\}$ 随机均匀抽取一名选民 $i^\star \in N$ **对于** $t = 2, \ldots, T$，**若** $c_t \succ_{i^\star} x_{t-1}$ **则**$x_t \leftarrow c_t$**否则**$x_t \leftarrow x_{t-1}$**返回** $(x_1, \ldots, x_T)$



注释 3.6 将候选人空间扩大到 $C_t = \{c_1, \ldots, c_t\}$ 并无帮助；只需在规则选择旧候选人而非挑战者的首个轮次t 评估失真，上述下界实例对这个更大的行动空间仍然成立。由于适用于 IC 模型的投票规则也适用于完整候选人空间，因此 $\Theta(\sqrt{T})$ 的失真在这两种部分信息模型下都是最优的。

3.2. 随机化投票规则

允许随机化能使静态度量失真出现分离。那么，在在线模型中还会如此吗?

我们将随机独裁推广到这一设定，使其在静态设定下的度量失真为 3（Anshelevich 和 Postl，2017）。一种方法是为每个轮次随机选择一位独裁者 t。然而，即使允许采用完整候选人空间 C_t，我们也不一定能找出独裁者在目前已到达的候选人中最偏好的候选人。这是因为我们只能观察现任者 x_{t-1} 与挑战者c_t之间的偏好。另一种方法是选择独裁者在 $\{x_{t-1}, c_t\}$中更偏好的候选人。然而，这种方法的失真会随轮数线性增长 T。

例 3.7（逐轮随机独裁下的线性失真）考虑将 n 名选民放置在一条直线上，其中 $n-1$ 名选民位于 0，另有一名选民 i^* 位于 1。将 c_1 放在 0，并将 c_t 放在 2，适用于 $t \geq 2$。打破平局时，让 i^* 偏好 c_2，c_3，.....而非 c_1。当现任者为 c_1时，若选中 i^* 作为独裁者，则获胜者将转为挑战者；若独裁者不是 i^*，则现任者保持不变。因此，$\Pr(x_T \neq c_1) = 1 - (1 - 1/n)^{T-1}$。于是，第 T轮的失真为

$$\frac{\mathbb{E}[\mathrm{sc}(x_T)]}{\mathrm{sc}(c_1)} = 1 + 2(n-1)\left(1 - \left(1 - \frac{1}{n}\right)^{T-1}\right).$$

固定 T并取 $n \to \infty$ 时，失真可达到 $2T - 1$。

相反，我们考虑让独裁者在每一轮中保持不变： 算法 2。独裁者固定后，我们可以通过传递性推断出独裁者的首选。直接观察可知，在无知对手的情况下，我们获得与静态情形相同的保证。

命题 3.8 在无知对手的情况下，*算法 2* 实现的度量失真至多为 3在部分信息 *IC* 模型中。

证明。 算法 2 在 $i^\star$ 每一轮为随机独裁者选择最佳候选人，其度量失真至多为 $3 - \frac{2}{n}$ （Anshelevich 和 Postl，2017）。 – □

有趣的情况是在自适应对手下。这样的对手可以通过多轮博弈获知持久独裁者，从而造成更大的失真。

例 3.9（学习持久独裁者）考虑对手将 n 名选民置于直线上，其中一名选民 i 位于 i处。在第 t轮，将 c_t 置于 t处。设 i^* 为持久独裁者。那么，获胜者为（c_1、c_2、……、c_{i^*}、c_{i^*}、……），也就是说，独裁者总是在第 i^* 轮之前选择挑战者，之后则持续选择现任者。因此，至多经过 n 轮，我们就能确定谁是独裁者。

在得知独裁者是谁之后，对手可以引入使失真最大化的候选人。然而，令人惊讶的是，算法 2 仍能达到 5 的常数度量失真。从对手的角度来看，这类似于强化学习中的探索—利用权衡。对手可以尝试探索独裁者是谁，但这样做的代价是，利用阶段受到限制，因而无法造成严重的失真。

定理 3.10 在自适应对手下，*算法 2* 在部分信息 *IC* 模型中实现的度量失真至多为5 。

证明。 对于 $n \le 2$，情形显然。假设 $n \ge 3$。固定一个轮次 t 和一个自适应对手。对于独裁者被选为 $i \in N$的情形，令 x_i 为在轮次 t选出的候选人，即 x_i 在 i 中被 C_t排在首位。令 d_i 为独裁者被选为 i时，对手在轮次 t 选择的度量；令 y_i^* 为相对于 d_i，在轮次 t使社会成本最小的最优候选人。

于是，轮次 t 的度量失真为

$$\frac{\mathbb{E}\left[\mathrm{sc}(x_t)\right]}{\mathbb{E}\left[\mathrm{sc}(y_t^*)\right]} = \frac{\frac{1}{n}\sum_{i\in N}\mathrm{sc}_{d_i}(x_i)}{\frac{1}{n}\sum_{i\in N}\mathrm{sc}_{d_i}(y_i^*)}.$$

首先注意到，我们有

$$\sum_{i\in N}\mathrm{sc}_{d_i}(y_i^*) = \sum_{i,j\in N} d_i(j,y_i^*) = \sum_{i\in N} d_i(i,y_i^*) + \sum_{i\neq j} d_i(j,y_i^*).$$

另一方面，由于 $d_i(j,\ x_i) \le d(j,\ i) + d_i(i,\ x_i) \le d(j,\ i) + d_i(i,\ y_i^*)$，我们有

$$\sum_{i\in N}\mathrm{sc}_{d_i}(x_i) = \sum_{i,j\in N} d_i(j,x_i) \le \sum_{i\neq j} d(j,i) + n\cdot\sum_{i\in N} d_i(i,y_i^*).$$

注意，由于随机独裁者不同实现下的选民之间的距离相同，因此我们可以直接写作 $d(j,\ i)$，而无需添加下标。因此，我们有

$$\sum_{i\neq j} d(j,i) = 2\sum_{i<j} d(j,i).$$

固定 $i \neq j$。对于另一位选民 k，由于 $d(j,\ i) = d_k(j,\ i) \le d_k(j,\ y_k^*) + d_k(i,\ y_k^*)$，对所有 $k \neq i,\ j$求和，可得

$$(n-2)d(j,i) \le \sum_{k\neq i,j}\left(d_k(j,y_k^*) + d_k(i,y_k^*)\right).$$

对所有满足 $i < j$ 的 $i,\ j$ 求和，我们得到

$$\begin{aligned}(n-2)\sum_{i<j} d(i,j) &\le \sum_{i<j}\sum_{k\neq i,j}\left(d_k(j,y_k^*) + d_k(i,y_k^*)\right)\\ &= \sum_{k\in N}\sum_{i<j,\,i,j\neq k}\left(d_k(j,y_k^*) + d_k(i,y_k^*)\right)\\ &= (n-2)\sum_{k\in N}\sum_{\ell\neq k} d_k(\ell,y_k^*),\end{aligned}$$

最后一个等式成立，是因为对于每个固定的 $\ell \neq k$， $d_k(\ell,\ y_k^*)$ 对 $N \setminus \{k,\ell\}$ 中的每位其他选民恰好出现一次。由于这样的选民有 $n-2$ 位，因此该等式成立。将两边除以 $n-2$，并按 $i,\ j$ 对 $k,\ \ell$ 重新索引，我们得到

$$\sum_{i<j} d(j,i) \leq \sum_{i \neq j} d_i(j, y_i^*).$$

因此，

$$\frac{\sum_{i\in N} \mathrm{sc}_{d_i}(x_i)}{\sum_{i\in N} \mathrm{sc}_{d_i}(y_i^*)} \leq \frac{nS_0 + 2S_1}{S_0 + S_1},$$

其中 $S_0 = \sum_{i\in N} d_i(i,\ y_i^*)$ 和 $S_1 = \sum_{i\neq j} d_i(j,\ y_i^*)$。

注意，$d_i(i,\ y_i^*) \leq d(i,\ j) + d_i(j,\ y_i^*)$。对所有满足 $i \neq j$ 的 $i,\ j$ 求和，得到

$$(n-1)S_0 \leq \sum_{i\neq j} d(i,j) + S_1 \leq 3S_1.$$

因此，$S_0 \leq \frac{3}{n-1} S_1$ 成立，并且在此约束下，上述右侧在 $S_0 = \frac{3}{n-1} S_1$ 时取最大值：

$$\frac{\sum_{i\in N} \mathrm{sc}_{d_i}(x_i)}{\sum_{i\in N} \mathrm{sc}_{d_i}(y_i^*)} \leq \frac{5n-2}{n+2} < 5. \qquad \square$$

现在，我们用显著的下界补充上界。结果表明，在两种对抗模型中，持久随机独裁都能实现最优失真。我们证明，对于无知对手和自适应对手，下界分别为 3 和 5。

我们首先考虑自适应对手。该下界由两类度量实例构造而成，算法无法根据观测到的信息区分它们。在一个实例中，第一个候选人 c_1 是最优候选人，而接受挑战者会使社会成本不断累积。另一个实例中，有一个最优挑战者，不接受它会导致很高的社会成本。这两个实例之间存在权衡：接受许多挑战者代价高昂，因为这会使第一个实例中的成本不断累积；但接受得少，则会增加错过最优候选人的概率。

定理 3.11 在自适应对手下，在部分信息 *IC* 模型中，任何随机投票规则都无法保证低于5 的常数度量失真。

证明。 固定 $n \geq 3$ 名选民和时域 T。我们首先指定可能到达的潜在候选人。对于每个子集 $S \subseteq \{2,\ldots,T\}$，创建一个潜在候选人 v_S。对于潜在候选人 v_S，令 $h(v_S) = |S|$ 表示其深度。

设定 $c_1 = v_\emptyset$。在轮次 $t \geq 2$，如果现任者是 $x_{t-1} = v_S$，则引入 $c_t = v_{S\cup\{t\}}$。

对于整数 h，令 $i_h = 1 + (h \bmod n)$。我们将构造两个不同的度量，它们都会向算法提供以下完全相同的输入：在任意轮次中，深度为 h 的现任者仅受到选民 i_h 的偏好而胜过挑战者，而其他每位选民 $i \neq i_h$ 都更偏好挑战者。

第一种度量。 令 $u^0 = (1,\ \ldots,\ 1) \in \mathbb{R}^n$，并递归定义 $u^{h+1} = u^h + 2e_{i_h}$，其中 e_i 是仅在坐标 i 上为 1、其他位置均为 0 的向量。对于深度为$h(v) = h$的潜在候选人 v，将选民—候选人距离定义为 $d(i,\ v) = u_i^h$。注意，如果轮次 t 的现任者 x_{t-1}深度为 h，则挑战者 c_t 的深度为 $h+1$。因此，我们有

$$d(i_h, c_t) = u_{i_h}^{h+1} = u_{i_h}^h + 2 = d(i_h, x_{t-1}) + 2 > d(i_h, x_{t-1}),$$

并且选民 i_h 更偏好现任者而非挑战者。对于其他选民，他们到 x_{t-1} 和 c_t 的距离相等，我们规定平局时优先选择挑战者。因此，该度量会产生所描述的输入。

注意，这可以扩展为一个完整度量。固定一个完整度量，例如将所有选民间距离设为 2。同样，只需检查三路径三角不等式： $d(i,c) \le d(i,c') + d(i',c') + d(i',c)$，对所有 i、$i' \in N$ 和 c、$c' \in C$ 均成立。根据定义，对于每个 t 以及每一对选民 i、i'，我们有$d(i,c_t) \ge 1$，且 $|d(i,c_t) - d(i',c_t)| \le 2$。因此， $d(i,c) - d(i',c) \le 2 \le d(i,c') + d(i',c')$，三路径三角不等式成立。

在此度量下， v_S的社会成本为 $n+2|S|$，因此初始候选人 c_1始终是最优的，其成本为 n。因此，第 T 轮的度量失真变为

$$\frac{\mathbb{E}[\mathrm{sc}(x_T)]}{n} = 1 + \frac{2}{n}\mathbb{E}[h(x_T)].$$

第二个度量。接下来，我们构造一个不同的度量，它诱导出相同的偏好剖面。固定一个数 $k \in \{2,\dots,T\}$。如果满足 $k \in S$，则称潜在候选人 v_S为优，否则称其为劣。因此，在第 k轮之前，所有候选人都是劣的；在第 k轮，挑战者是优的；在第 k轮之后，现任者和挑战者要么都是优的，要么都是劣的。

对于选民 $i \in N$，定义 n向量 $b^i = (5,\dots,5) - 2e_i$和 $g^i = (1,\dots,1) + 2e_i$。然后定义 v_S 在 $h(v_S) = h$ 下的选民—候选人距离为

$$d(i,v_S) = \begin{cases} b_i^{i_h}, & k \notin S, \\ g_i^{i_{h-1}}, & k \in S. \end{cases}$$

要验证它能扩展为完整度量，再次注意，根据定义，对于每个 t 以及每一对选民 i，i'，我们都有 $d(i,\ c_t) \ge 1$ 和$|d(i,\ c_t) - d(i',\ c_t)| \le 2$。因此， $d(i,\ c) - d(i',\ c) \le 2 \le d(i,\ c') + d(i',\ c')$，且三路径三角不等式成立。

平局时优先选择劣质候选人而非优质候选人，然后优先选择深度较大的候选人。我们声称，这个度量也会产生所描述的输入。对于深度为 h的现任者，挑战者的深度为 $h+1$。如果现任者是劣质候选人，它到选民的距离向量为 b^{i_h}。那么，如果挑战者处于第 k 轮，其距离向量为 g^{i_h}；否则为 $b^{i_{h+1}}$。如果现任者是优质候选人，其距离向量为 $g^{i_{h-1}}$，而挑战者的距离向量为 g^{i_h}。在这三种情况下，选民 i_h到现任者的距离分别为3、3、1，到挑战者的距离分别为5、3、3。因此，i_h总是更倾向于现任者。选民 i_{h-1}到现任者的距离分别为5、5、3，到挑战者的距离分别为5、1、1。因此， i_{h-1} 总是更倾向于挑战者。选民 i_{h+1}到现任者的距离分别为5、5、1，到挑战者的距离分别为3、1、1。因此， i_{h+1}总是更倾向于挑战者。其他每位选民 $i \notin \{i_{h-1}, i_h, i_{h+1}\}$到现任者的距离分别为5、5、1，到挑战者的距离分别为5、1、1。因此， i 总是更倾向于挑战者，由此产生了所描述的输入。

在此度量下，劣质候选人的社会成本为 $5n-2$，而优质候选人的社会成本为$n+2$。优质候选人总会到来，因此始终是最优选择，其社会成本为 $n+2$。让我们计算最后一轮的获胜者 x_T的期望社会成本。

由于算法无法观察到隐藏的 k， x_T 与 k的选择相互独立。对每个 $k \in \{2,\dots,T\}$，令$p_k = \Pr(k \in S_T)$，其中 $x_T = v_{S_T}$。由于当且仅当 $k \in S_T$时 x_T为优质候选人， x_T的期望社会成本变为 $p_k(n+2) + (1-p_k)(5n-2)$。因此，第 T 轮的度量失真变为

$$\frac{\mathbb{E}[\mathrm{sc}(x_T)]}{n+2} = \frac{5n-2}{n+2} - \frac{4n-4}{n+2}p_k.$$

现在我们证明，可以选择 k 使其 p_k较小。由于 x_T 当且仅当 $k \in S_T$时是好的，我们有

$$\sum_{k=2}^{T} \mathbf{1}(k \in S_T) = |S_T| = h(x_T).$$

取期望后，我们有 $\sum_{k=2}^{T} p_k = \mathbb{E}[h(x_T)]$。因此，存在 k 使得 $p_k \le \mathbb{E}[h(x_T)]/(T-1)$，并且对于这个 k，x_T 的度量失真至少为

$$\frac{5n-2}{n+2} - \frac{4n-4}{(T-1)(n+2)}\mathbb{E}[h(x_T)] \ge \frac{5n-2}{n+2} - \frac{4}{T-1}\mathbb{E}[h(x_T)].$$

请注意，仅凭算法的概率分布即可推断这一选择，而不必知道实际实现的随机选择。

选择度量。 现在取 $T = n^2+1$。利用算法的概率分布，根据所述输入计算 $\mathbb{E}[h(x_T)]$。如果 $\mathbb{E}[h(x_T)] \ge 2n$，就选择第一个度量来模拟该输入，从而得到至少为 $1 + \frac{2}{n}\cdot 2n = 5$ 的失真。如果 $\mathbb{E}[h(x_T)] < 2n$，就选择第二个度量，从而得到至少为

$$\frac{5n-2}{n+2} - \frac{4}{n^2}2n \ge 5 - \frac{20}{n}.$$

对于任何小于 5 的常数，当 n 足够大时，下界都会超过该常数。 □

备注 3.12 上述下界实例仍适用于更大的完整候选人空间C_t。对手会在每次选定现任者 $x_{t-1} = v_S$ 后（包括返回先前的候选人时）引入 $c_t = v_{S\cup\{t\}}$。因此，每个挑战者始终都是实际现任者的子节点，而且所有构造出的度量都会继续产生相同的现任者–挑战者比较结果。由于失真界是根据最终获胜者的深度推导出的，因此无论算法在中间阶段如何选择，该结论都仍然成立。

最后，我们证明，在无知对手面前，持久随机独裁同样是最优的。我们再次考虑算法无法区分的两个实例。

定理 3.13 *在无知对手面前，在部分信息 IC 模型中，没有任何随机投票规则能够保证度量失真低于常数 3。*

证明。 固定 $n \ge 2$ 和 $D < 3 - \frac{2}{n}$。我们证明，没有任何随机投票规则能够保证失真至多为 D。固定一个随机投票规则、一轮 $t \ge 2$，以及关于前 $t-1$ 个候选人 c_1、.....、c_{t-1} 的完整偏好剖面。假设该算法保证失真为 D。对于每位选民 $i \in N$，令 X_i 为指示随机变量，用以表示当恰好选民 i 更偏好 x_{t-1} 而非 c_t 时，算法在第 t 轮是否保留现任者 x_{t-1}。X_i 的随机性来自算法（以及对先前实际获胜者的观测）。我们首先对 X_i 的期望进行界定。

考虑直线上的一个实例：将前 $t-1$ 个候选人置于 0 处，将选民 i 置于 1 处，并将挑战者 c_t 和所有其他选民置于 2 处。以一致的方式打破平局，使得对于每个 $s < t$，选民 i 都更偏好 c_s 而非 c_t。前 $t-1$ 个候选人的社会成本为 $2n-1$，而 c_t 的社会成本为 1。于是，在第 t 轮，算法的失真为

$$\Pr(X_i = 0)\cdot 1 + \Pr(X_i = 1)\cdot(2n-1) = 1 + 2(n-1)\mathbb{E}[X_i] \le D,$$

或者等价地，$\mathbb{E}[X_i] \le \frac{D-1}{2n-1}$。

现在我们构造另一个实例。对于每个候选人 c，将其与一个集合 $B(c) \subseteq [n]$ 关联起来。定义选民—候选人距离如下：

$$d(i,c) = \begin{cases} 1, & i \in B(c), \\ 3, & i \notin B(c). \end{cases}$$

将选民之间的距离和候选人之间的距离设为 2，可得到一个有效度量。平局时优先选择后到者。因此，在第 t 轮，现任者的支持者恰好是$B(x_{t-1}) \setminus B(c_t)$。

注意，候选人 c 的社会成本为 $3n-2|B(c)|$。设置 $B(c_1)=[n]$，使得 c_1 始终具有最优社会成本。由于该算法具有 D的失真保证，因此对于每一轮 t，都有 $\mathbb{E}[3n-2|B(x_t)|]/n \le D$，或者等价地，

$$\mathbb{E}[|B(x_t)|] \ge \frac{n(3-D)}{2} > 1.$$

我们声称 我们可以用一种特定方式选择集合 $B(c_t)$，使得 $t \ge 2$ 导致矛盾 .

我们使用概率方法证明此类集合的存在。对于参数 $\lambda > 1$，根据以下分布独立地对每个 $B(c)$ 进行抽样:

$$\Pr(B(c)=S) = \frac{\lambda^{-|S|}}{Z},$$

其中 $S \subseteq [n]$ 且 $Z=(1+\lambda^{-1})^n$。这些集合对应于实例的随机性。

对于非负整数 k，令 $\phi(k) = \sum_{\lambda^{-1}}^{Z} \lambda^k$。（令$\phi(-1)=0$。）令 $b_t = |B(x_t)|$。我们将跟踪势能 $\phi(b_t)$ 的期望变化。在第 $t \ge 2$ 轮，如果 $x_t = x_{t-1}$，则有$\phi(b_t)=\phi(b_{t-1})$。如果 $x_t = c_t$，增量 $\phi(b_t)-\phi(b_{t-1})$ 有界于 $\phi(b_t)=\phi(|B(c_t)|)$。由于$c_t \ne c_1$，我们有 $\mathbb{E}[\phi(|B(c_t)|)] = 2^n/(\lambda-1) =: \eta > 0$。

也存在 $\phi(b_t)$ 递减的情况。固定一轮 t。令 $B_{<t} = (B(c_s))_{2 \le s \le t-1}$ 表示截至第 $t-1$ 轮实例的随机性。我们将以 $B_{<t}$ 为条件。

对于每个 $i \in N$，考虑事件 $E_i = \{i \in B(x_{t-1}) \ \ B(c_t) = B(x_{t-1}) \setminus \{i\}\}$，并令 $E = \cup_{i \in N} E_i$。事件 E_i两两不相交，因为对于 $i,\ j \in N$, $i,\ j \in B(x_{t-1})$ 和 $B(x_{t-1}) \setminus \{i\} = B(x_{t-1}) \setminus \{j\}$ 意味着 $i=j$。在事件 E_i 上，由于 $B(x_{t-1}) \setminus B(c_t) = \{i\}$，也就是说，恰好只有选民 i偏好现任者，因此当且仅当 $X_i = 1$ 时，算法才会选出现任者。注意，在以 $B_{<t}$ 为条件时，$B(x_{t-1})$ 和 b_{t-1} 仍可能是随机的。由于 $B(c_t)$ 与 $B_{<t}$及算法随机性相互独立，我们有

$$\begin{aligned}
&\mathbb{E}[(\phi(b_t)-\phi(b_{t-1}))\mathbf{1}_{E_i}|B_{<t}] \\
&= \mathbb{E}[(\phi(b_{t-1}-1)-\phi(b_{t-1}))\mathbf{1}(\{i \in B(x_{t-1})\})\mathbf{1}(B(c_t)=B(x_{t-1})\setminus\{i\})(1-X_i)|B_{<t}] \\
&= \mathbb{E}\left[(\phi(b_{t-1}-1)-\phi(b_{t-1}))\mathbf{1}(\{i \in B(x_{t-1})\})\frac{\lambda^{-(b_{t-1}-1)}}{Z}(1-X_i)\middle| B_{<t}\right] \\
&= -\mathbb{E}\left[\mathbf{1}(\{i \in B(x_{t-1})\})(1-X_i)|B_{<t}\right],
\end{aligned}$$

其中，最后一个等式由 $\Pr(B(c)=S)(\phi(|S|+1)-\phi(|S|))=1$ 得出。因此，

$$\begin{aligned}
\mathbb{E}[(\phi(b_t)-\phi(b_{t-1}))\mathbf{1}_E|B_{<t}] &= \sum_{i \in N}\mathbb{E}[(\phi(b_t)-\phi(b_{t-1}))\mathbf{1}_{E_i}|B_{<t}] \\
&= -\mathbb{E}[b_{t-1}|B_{<t}] + \mathbb{E}\left[\sum_{i \in B(x_{t-1})} X_i \middle| B_{<t}\right].
\end{aligned}$$

此外，根据前面的分析，$\mathbb{E}[(\phi(b_t)-\phi(b_{t-1}))\mathbf{1}_{E^c}|B_{<t}] \le \mathbb{E}[\phi(|B(c_t)|)] = \eta$。综合以上结果可得

$$\mathbb{E}[\phi(b_t)-\phi(b_{t-1})|B_{<t}] \le -\mathbb{E}[b_{t-1}|B_{<t}] + \mathbb{E}\left[\sum_{i \in B(x_{t-1})} X_i \middle| B_{<t}\right] + \eta.$$

由于 $X_i \ge 0$，我们有

$$\mathbb{E}\left[\sum_{i \in B(x_{t-1})} X_i \middle| B_{<t}\right] \le \sum_{i=1}^{n}\mathbb{E}[X_i|B_{<t}] \le \frac{n(D-1)}{2(n-1)} =: \beta < 1,$$

其中第二个不等式由上述论证得出。现在对 $B_{<t}$ 取期望，根据塔式性质，我们有

$$\mathbb{E}[\phi(b_t) - \phi(b_{t-1})] \leq -\mathbb{E}[b_{t-1}] + \beta + \eta.$$

取足够大的 λ，使得 $\delta = 1-\beta-\eta > 0$。由于 $\mathbb{E}[b_{t-1}] > 1$，我们有 $\mathbb{E}[\phi(b_t)-\phi(b_{t-1})] \leq -\delta < 0$。由于 $b_1 = n$，对各轮次求和可得 $\mathbb{E}[\phi(b_T)] \leq \phi(n) - (T-1)\delta$。取足够大的 T，得到 $\mathbb{E}[\phi(b_T)] < 0$，这与 ϕ 非负矛盾。 □

注 3.14 我们可以再次通过推广证明，将结果扩展到完整候选人空间。细节留在附录 A中。

4. 完全信息下的现任者—挑战者模型

现在我们考虑完全信息下的 IC 模型。在第 t 轮，我们观察到选民对 $C_t = \{c_1, \ldots, c_t\}$ 的完整排名。

4.1. 确定性投票规则

确定性投票规则能否在此模型下实现常数度量失真？可以注意到，即使在完全信息下，多数否决（Kizilkaya 和 Kempe，2022年）也无法直接应用，因为它可能会选出此前已被淘汰的候选人。

例 4.1 考虑以下情形：$T = 3$，3 位候选人 a、b、c 按此顺序到达。共有 9 位选民，其离线排名如下：

选民人数	排名
2	$a \succ b \succ c$
4	$b \succ c \succ a$
3	$c \succ a \succ b$

每轮运行多数否决。在第 $t = 2$ 轮，a 获胜，因为 5 位选民将 $a \succ b$ 排在首位，而 b 被淘汰。然而，在第 $t = 3$ 轮，多数得分为 $s(a) = 2$、$s(b) = 4$、$s(c) = 3$，无论选民的顺序如何，a 都会首先被否决。接着，由于有 6 位选民将 b 排在 c 之前，c 随后被否决。因此，多数否决的获胜者是 b，而在现任者—挑战者模型中无法选出该候选人。

尽管如此，我们证明常数失真是可能的。Copeland 的度量失真为 5（Anshelevich 等人，2018），但它再次与在任者—挑战者模型不兼容，因为它可能会选择先前已淘汰的候选人。然而，得出度量失真界的关键分析基于这样一个事实：Copeland 会选出竞赛图中未被覆盖的候选人。Munagala 和 Wang (2019年) 将这一观察推广到加权竞赛图，将度量失真界改进为 $2+\sqrt{5}$。我们证明，即使是这一推广版本，也可以与 IC 模型兼容，并且每轮的度量失真至多为 $2+\sqrt{5}$。

我们首先引入一些记号。固定 $\lambda \in [\frac{1}{2}, 1]$。给定一个选举实例，候选人 $c \in C$属于 λ-加权未覆盖集 (Munagala 和 Wang，2019年)，当且仅当对于每个候选人 $c' \neq c$，

- 要么 $|\{i \in N : c \succ_i c'\}| \geq (1-\lambda)|N|$，
- 或者存在 $c'' \notin \{c, c'\}$，使得 $|\{i \in N : c \succ_i c''\}| \geq (1-\lambda)|N|$ 且 $|\{i \in N : c'' \succ_i c'\}| \geq \lambda|N|$。

算法 3 在线加权未覆盖集规则**输入**：一个包含 T 轮次的在线投票实例**输出**：一个获胜者序列 x_1，……，x_T，满足对于每个 $t \geq 2$**令** $\varphi \leftarrow \frac{\sqrt{5}-1}{2}$ $x_1 \leftarrow c_1$**对于** $t = 2$，……，T，**执行以下操作：若** x_{t-1} 属于 $E_t = (N,\ C_t,\ \succ|_{C_t \times C_t})$ 的 φ 加权未覆盖集，**则** $x_t \leftarrow x_{t-1}$**否则**$x_t \leftarrow c_t$**返回** $(x_1,\ \ldots\ldots,\ x_T)$

引理 4.2 (Munagala 和 Wang，2019年) 设 $\varphi = \frac{\sqrt{5}-1}{2} \approx 0.618$。在 φ 加权未被支配集合中始终存在一位候选人，其度量失真至多为 $2+\sqrt{5}$。

这使我们能够设计一种具有常数度量失真的确定性规则。

定理 4.3 算法 3 在完全信息 *IC* 模型中达到的度量失真至多为 $2+\sqrt{5}$。

证明。根据引理 4.2，只需证明第 t 轮的获胜者 x_t 属于 $E_t = (N, C_t, \succ|_{C_t \times C_t})$ 的 φ 加权未被支配集合。我们对 t 使用归纳法证明。当$t = 1$ 或 $t = 2$ 时，结论显然成立，因为 φ 加权未被支配集合非空。假设 x_{t-1} 属于 E_{t-1} 的 φ 加权未被支配集合。若 $x_t = x_{t-1}$，根据定义，x_t 属于 E_t 的 φ 加权未被支配集合。另一方面，若 $x_t = c_t$，则 x_{t-1} 不属于 E_t 的 φ 加权未被支配集合。由于添加 c_t 不会改变 C_{t-1} 中候选人之间的两两排名，因此必定是c_t 阻止了 x_{t-1} 进入 φ 加权未被支配集合，即，

(1) $|\{i \in N:\ x_{t-1} \succ_i c_t\}| < (1-\varphi)|N|$ 且

(2) 对每个 $c' \in C_t \setminus \{x_{t-1}, c_t\}$，$|\{i \in N:\ x_{t-1} \succ_i c'\}| < (1-\varphi)|N|$ 或 $|\{i \in N:\ c' \succ_i c_t\}| < \varphi|N|$。

现在我们证明 c_t 属于 E_t的 φ-加权未覆盖集。对于 $x_{t-1} \in C_t$，由于$|\{i \in N:\ c_t \succ_i x_{t-1}\}| > \varphi|N| \geq (1-\varphi)|N|$，它不会阻挡 c_t。对于 $c' \in C_t \setminus \{x_{t-1}, c_t\}$，由于$x_{t-1}$ 属于 E_{t-1}的 φ-加权未覆盖集，

- 要么 $|\{i \in N:\ x_{t-1} \succ_i c'\}| \geq (1-\varphi)|N|$，
- 或者存在 $c'' \in C_{t-1} \setminus \{x_{t-1}, c'\}$，使得 $|\{i \in N:\ x_{t-1} \succ_i c''\}| \geq (1-\varphi)|N|$ 且$|\{i \in N:\ c'' \succ_i c'\}| \geq \varphi|N|$。

对于前一种情况，根据 (2)，我们有 $|\{i \in N:\ c_t \succ_i c'\}| \geq (1-\varphi)|N|$。对于后一种情况，由于$|\{i \in N:\ x_{t-1} \succ_i c''\}| \geq (1-\varphi)|N|$ 对某个 c''成立，再次根据 (2)，我们有 $|\{i \in N:\ c_t \succ_i c''\}| \geq (1-\varphi)|N|$。由于后一种情况下 $|\{i \in N:\ c'' \succ_i c'\}| \geq \varphi|N|$，$c'$不会阻挡 c_t。因此，x_t属于 E_t的 φ-加权未覆盖集。□

因此，在整个过程中都可以通过确定性投票规则实现常数失真，这与部分信息情形不同。事实上，常数 $2+\sqrt{5}$ 是最优的。

定理 4.4 在完全信息 *IC* 模型中，任何确定性投票规则都无法保证常数度量失真低于$2+\sqrt{5}$。

证明再次针对每个 $\theta \in (0,1)$，考虑一个与多个度量相一致的序数偏好剖面。确定性投票规则要么在每一轮都接受挑战者，要么在某一轮拒绝某个挑战者。一种度量会使前一种情况的失真为 $\frac{1+\theta}{1-\theta}$，另一种度量则会使失真为 $1+\frac{2}{\theta}$。因此，对 $\theta \in (0,1)$ 上的 $\left\{1+\frac{2}{\theta}, \frac{1+\theta}{1-\theta}\right\}$ 取最大最小值，可得失真为 $2+\sqrt{5}$，由 $\theta = (\sqrt{5}-1)/2$ 给出。

我们首先证明一个引理，该引理允许我们根据某些偏好剖面构造一致度量。

引理 4.5 给定一个偏好剖面 $\succ$、候选人的非负有理高度$(h_c)_{c\in C}$ 以及非负有理数$(\delta_i)_{i\in N}$。对于选民 $i \in N$，令 $F_i(c) = \min\{h_{c'} : c' \succsim_i c\}$表示在偏好程度至多为 c 的候选人中最低的高度。如果

(1) $\sum_{i\in N} \delta_i > 0$ 并且对于每个选民 $i \in N$，$c \succ_i c'$ 推出 $h_c - h_{c'} \le \delta_i$，且

(2) 存在一个候选人 y 满足 $h_y = 0$,

那么存在一个一致伪度量 d，使得

$$\frac{\mathrm{sc}(x)}{\mathrm{sc}(y)} = 1 + 2\frac{\sum_{i\in N} F_i(x)}{\sum_{i\in N} \delta_i},$$

对于每个候选人 $x \in C$。

证明。我们构造所需的伪度量 d。定义选民—候选人距离为 $d(i,c) = \delta_i/2 + F_i(c)$。对于不同的候选人 c、c'，令 $d(c,c') = \max\{h_c, h_{c'}\}$。对于不同的选民 i、j，通过最短路径距离定义$d(i,j)$：$d(i,j) = \min_{c\in C}\{d(i,c) + d(j,c)\}$。为证明这是一个有效度量，注意对于不同的候选人 a、b，$c \in C$ $d(a,c) = \max\{h_a, h_c\} \le \max\{h_a, h_b, h_c\} \le d(a,b) + d(b,c)$。因此，只需证明对于每个$i \in N$ 以及 c、$c' \in C$，都有 $|d(i,c) - d(i,c')| \le d(c,c') \le d(i,c) + d(i,c')$。这是因为选民之间的距离是通过最短路径距离定义的。

注意，根据假设，$\max\{0, h_c - \delta_i\} \le F_i(c) \le h_c$。因此，对于任意选民 i 和互不相同的候选人 c, c',

$$|d(i,c) - d(i,c')| = |F_i(c) - F_i(c')| \le d(c,c').$$

此外，假设 $h_c \ge h_{c'}$，我们有

$$d(i,c) + d(i,c') = \delta_i + F_i(c) + F_i(c') \ge \delta_i + F_i(c) \ge h_c = d(c,c').$$

因此，该伪度量是有效的。

根据定义，$c \succ_i c'$ 意味着 $F_i(c) \le F_i(c')$，因此得到的伪度量与$\succ$一致。最后，由于 $h_y = 0$，对于每位选民 i，我们有 $F_i(y) = 0$。因此，对于每位候选人 x，

$$\frac{\mathrm{sc}(x)}{\mathrm{sc}(y)} = \frac{\sum_{i\in N}(\delta_i/2 + F_i(x))}{\sum_{i\in N}\delta_i/2} = 1 + 2\frac{\sum_{i\in N} F_i(x)}{\sum_{i\in N}\delta_i}.$$ □

现在我们来证明下界。

定理 *4.4* 的证明定理 *4.4*。固定一个确定性投票规则、一个有理数 $\theta \in (0,1)$，以及一个整数$L \ge 2$。令 $T = L(L-1)+1$，并考虑 T按以下顺序到达的候选人：c_1，c_2，…，c_T。

我们构造两类选民。比例为 $1-\theta$ 的选民（记为类型 A）更偏好较晚到达的候选人：$c_T \succ c_{T-1} \succ \cdots \succ c_1$。剩余比例 θ 平均分为 L 个组 B_1，……，B_L。对于每个 $\ell \in [L]$，B_ℓ中的选民按照以下方式将候选人划分为区块：

$$b_\ell(t) = \left\lfloor \frac{t+\ell-2}{L} \right\rfloor.$$

也就是说，候选人 c_t 位于区块 b_ℓ（t）中。选民更偏好较早区块中的候选人，而非较晚区块中的候选人；但在同一区块内，他们更偏好较晚到达的候选人。注意，由于 θ 是有理数，所有选民比例都可以由一个有限选民集 N 实现。固定一个这样的实现，并令 n 表示选民人数。

注意，对于每对连续候选人 c_{t-1}，c_t，恰有一个群组 B_ℓ 偏好c_{t-1} 而非 c_t，而所有其他选民（包括类型 A）都偏好 c_t。因此，

$$\frac{1}{n}|\{i \in N : c_t \succ_i c_{t-1}\}| = 1 - \frac{\theta}{L}. \tag{1}$$

另一方面，群组 B_ℓ 中的所有选民都偏好 c_1 而非 c_T。

我们考 虑两种情况，具体取决于投票规则在此偏好剖面上的表现。

情况 1：该规则拒绝了某个挑战者。 令 t 为该规则首次选出现任者的轮次。注意，$x_{t-1} = c_{t-1}$。考虑限制在 C_t上的偏好剖面。令 $s < t$的高度分别为 $h_{c_t} = 0$ 和 $h_{c_s} = 1$，并令 $i \in A$ 的高度为 $\delta_i = 0$，其他情况下则令 $\delta_i = 1$。

这些取值满足引理 4.5的假设。我们有 $\sum_{i\in N} \delta_i = n\theta > 0$。此外，由于一名选民 $i \in A$ 将 c_t 排在可选候选人中的首位，$c \succ_i c'$ 意味着 $h_c \le h_{c'}$。对于其他类型的选民，该不等式显然成立。

此外，由于 $F_i(c_{t-1}) = 1$ 当且仅当 $c_t \succ_i c_{t-1}$，根据 (1)，我们有 $\sum_{i\in N} F_i(c_{t-1}) = n(1 - \theta/L)$。应用 引理 4.5，令 $y = c_t$，可得第 t 轮的失真至少为

$$\frac{\mathrm{sc}(c_{t-1})}{\mathrm{sc}(c_t)} = 1 + 2\frac{n(1-\theta/L)}{n\theta} = 1 + \frac{2}{\theta} - \frac{2}{L}. \tag{2}$$

对于剩余轮次，由于我们已经得到了必要的界，可以任意安排挑战者。剩余候选人的偏好不必与上述实例相符。

情况 2：该规则接受每一位挑战者。 在这种情况下，最终轮获胜者是 $x_T = c_T$。指定高度 $h_{c_t} = (t-1)/(T-1)$，并令 $\delta_i = 1$ 对 $i \in A$ 成立，否则令 $\delta_i = 1/L$。

这些值同样满足 引理 4.5。事实上，对于 A 中的选民，所有高度差都至多为 1。对于 B_ℓ 中的选民，沿其排名递减的高度仅发生在区块内，且递减至多为 $(L-1)/(T-1) = 1/L$。此外，$\sum_{i\in N} \delta_i = n\left(1 - \theta + \frac{\theta}{L}\right) > 0$。

现在我们计算第 T 轮的失真。对于类型为 A 的选民 i，有 $F_i(c_T) = \min_t h_{c_t} = 0$。对于类型为 B_ℓ 的选民 i，最后一个区块从 $c_{T-\ell+1}$ 开始，因此 $F_i(c_T) = (T-\ell)/(T-1)$。因此，

$$\sum_i F_i(c_T) = \frac{n\theta}{L}\sum_{\ell=1}^{L}\frac{T-\ell}{T-1} = n\theta\left(1 - \frac{1}{2L}\right).$$

应用引理 4.5 并结合 $y = c_1$，我们得到

$$\begin{aligned}\frac{\mathrm{sc}(c_T)}{\mathrm{sc}(c_1)} &= 1 + 2\frac{\theta(1-1/(2L))}{1-\theta+\theta/L} \\ &= \frac{1+\theta}{1-\theta+\theta/L}.\end{aligned} \tag{3}$$

综合各个情形。 由（2）和（3，该规则的最坏情况失真至少为 $\min\left\{1 + \frac{2}{\theta} - \frac{2}{L},\ \frac{1+\theta}{1-\theta+\theta/L}\right\}$。取 $L \to \infty$，该值趋近于 $\min\left\{1 + \frac{2}{\theta},\ \frac{1+\theta}{1-\theta}\right\}$。因此，在 $\theta \in (0, 1)$ 上对上式取最大值，可得 $2 + \sqrt{5}$ 由 $\theta^* = (\sqrt{5}-1)/2$。

因此，对于低于 D 的常数 $2 + \sqrt{5}$，我们可以取足够大的 L，并取一个足够接近 θ^* 的有理数 θ，使得失真高于 D。由于该规则是确定性的，

对手可以预先确定该规则是否以及何时首次拒绝某个挑战者。因此，它可以在执行前固定相应的完整实例和度量，实际上使失真超过 D。

□

注 4.6 在部分信息设置下，我们无法实现 算法 3 ，因为我们无法判断 x_{t-1} 是否位于 φ-加权未覆盖集合中，除非掌握 c_t与 $C_{t-1} \setminus \{x_{t-1}\}$之间的信息。

4.2. 随机化投票规则

算法 2 显然适用于完全信息 IC 模型，并且能够保证对无知对手和自适应对手的失真界分别为 3 和 5。持久随机独裁是否仍然是最优的?

我们先考虑无知对手。事实证明，算法 2在这种情况下仍然是最优的。

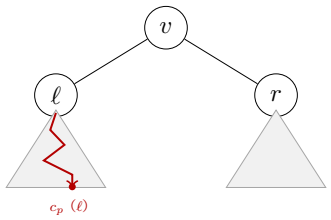
定理 4.7 在无知对手下，完全信息 *IC* 模型中不存在任何随机投票规则能够保证低于3 的常数度量失真。

证明。 固定一个随机投票规则和 $\varepsilon > 0$。我们构造一个无知实例，使得其在最后一轮的期望失真至少为 $3 - \varepsilon$。

实例。 考虑一棵深度为 d 、根节点为 v_0 的完美二叉树，其叶节点从左到右分别代表候选人。因此，总共有 $T = 2^d$ 轮次。我们使用概率方法证明所需实例的存在性。对于每个内部节点 v，考虑一个子节点指针 $p_v \in \{L, R\}$ ，它要么指向左子节点，要么指向右子节点。给定一个指针配置 $p = (p_v)_v$，令 $c_p(v)$ 表示从 v出发沿着这些指针得到的叶节点。对于每个 $p =(p_v)_v$ ，我们将构造一个实例，使得 $c_p^* = c_p(v_0)$ 在最后一轮中成为最优候选人。我们从 $\{L, R\}$ 中独立且均匀地采样每个指针 p_v ，并证明某个实现会产生所需的实例。

现在，我们为每个 $p =(p_v)_v$构造实例。将树的内部节点编号为 v_1、.....、 v_m，其中 $m = 2^d - 1$。固定一个整数 $q \geq 3$，使得 $\alpha = 1/q < \varepsilon/4$。这些参数对所有配置 p均相同。构造一个包含 $n = q^m$ 名选民的集合，并将其编号为 $i = (i_1$、.....、 $i_m) \in [q]^m$。对于每名选民 i 和每个内部节点 v_j，如果 $i_j = 1$，则定义 $X_{i,\ v_j} = 1$ ；否则定义 $X_{i,\ v_j} = 0$ 。因此，对于从 N中均匀随机抽取的选民 i ，各值 $(X_{i,\ v})_v$ 是成功概率为 α 的独立同分布伯努利随机变量。

现在，我们递归地构造每位选民 i 的偏好排序。固定一个指针配置 p。回忆一下，候选人就是叶节点。对于节点 v，记 T_v 为以 v 为根的子树，记 C_v 为 T_v的叶节点。对于叶节点，其排名仅包含该候选人。对于一个内部节点 v ，其左右子节点为 ℓ、 r，假设两个子树叶节点内部的排名已经构造完成。如果$X_{i,\ v} = 1$，则将所有 C_r 排在所有 C_ℓ之前，同时保持各自的排名不变。否则，将 C_r 的整个排名插入 C_ℓ 的排名中，紧接在 $c_p\ (\ell)$ 之后，同样



$$X_{i,v} = 1 \implies R \succ L^+ \succ c_p(\ell)\ \succ L^-$$
$$X_{i,v} = 0 \implies L^+ \succ c_p(\ell)\ \succ R \succ L^-$$

保留两个子树内部的排名。根节点处的排名定义了 $\succ_i$ 选民 i的完整偏好排序。

注意到，将排名限制在 C_v 以及所有在 C_v 之前到达的候选人范围内后，排名不依赖于 p_v。因此，直到 C_v 中的最后一位候选人到达的轮次， p_v 仍独立于算法的观测和决策。

偏好比例界。如前所述，我们稍后将构造一个度量，使得 $c_p^* = c_p(v_0)$ 达到最优社会成本。为了界定算法的失真，令 $S(x,y) = \frac{1}{n}|\{i \in N : x \succsim_i y\}|$ 表示将 x 排在不低于 y 的位置的选民比例。等价地，$S(x,y) = \Pr_{i\sim\mathrm{Unif}(N)}(x \succsim_i y)$。我们将证明，$\mathbb{E}[S(x_T, c_p^*)]$ 可以任意小；结合稍后考虑的特定度量，这将得到所需的失真。注意，对于 $S(x,y)$，根据实例的构造，以下结论成立。对于任意叶节点 x，$S(x,x) = 1$。对于任意内部顶点 v，若其左右子节点为 ℓ, r，且 $x \in C_v$ 是一个叶节点，

$$S(x, c_p(v)) = \begin{cases} S(x, c_p(\ell)), & \text{if } x \in C_\ell \text{ and } p_v = L, \\ S(x, c_p(r)), & \text{if } x \in C_r \text{ and } p_v = R, \\ (1-\alpha)S(x, c_p(\ell)), & \text{if } x \in C_\ell \text{ and } p_v = R, \\ \alpha, & \text{if } x \in C_r \text{ and } p_v = L. \end{cases}$$

前两种情况根据定义是显然的。若 $x \in C_\ell$ 且 $p_v = R$，那么由于 $c_p(v) = c_p(r)$，对于选民 i 而言，要将 x 排在 $c_p(r) \in C_r$之上，等价于 $X_{i,v} = 0$ 且 x 的排名至少与 $c_p(\ell)$一样高。因此，在固定指针配置 p 的同时，

$$\begin{aligned} S(x, c_p(v)) &= \Pr_{i\sim\mathrm{Unif}(N)}(X_{i,v} = 0, x \succsim_i c_p(\ell)) \\ &= \Pr_{i\sim\mathrm{Unif}(N)}(X_{i,v} = 0) \Pr_{i\sim\mathrm{Unif}(N)}(x \succsim_i c_p(\ell)) \\ &= (1-\alpha)S(x, c_p(\ell)). \end{aligned}$$

同样地，若 $x \in C_r$ 且 $p_v = L$，当且仅当 $X_{i,v} = 1$时，选民 i 将 x 排在 $c_p(v) = c_p(\ell)$之上。因此，$S(x, c_p(v)) = \Pr_{i\sim\mathrm{Unif}(N)}(X_{i,v} = 1) = \alpha$。

因此，如果从根 v_0 到叶节点 x 的路径中出现了与指针 p方向相反的移动情况——如果路径从父节点移动到左子节点，而指针指向右子节点，则会累积一个 $(1-\alpha)$ 因子。将这种情况称为软错误。如果路径移动到右子节点，而指针指向左侧，则 $S(x, c_p^*) \le \alpha$。将这种情况称为致命错误。如果没有发生致命错误，并且存在 r 个软错误，则 $S(x, c_p^*) = (1-\alpha)^r$。

存活概率。将每个 p_v 视为在$\{L, R\}$上均匀独立的随机变量。对于顶点 v，令 x_v 表示算法处理完 T_v中最后一个候选人的轮次后的获胜者。对于非负整数 k，令 $E_{v,k}$ 表示这样一个事件：$x_v \in C_v$，且从 v 到 x_v 的路径中不包含致命错误，并且至多包含 k 个软错误。注意，$E_{v,k}$的随机性来自算法随机性（在整棵树上运行）以及 p的随机性。

对于非负整数 j, k，令 $M_{j,k}$ 表示在所有随机化算法以及满足深度$(T_v) = j$的顶点 v 上，$\Pr(E_{v,k})$ 的上确界。我们声称，对于固定的 k和 $M_{j,k} \to 0$，当 $j \to \infty$时成立。

令 $M_{0,k} = 1$，其中 $k \ge 0$ 且$M_{j,-1} = 0$。我们首先证明如下递推关系：

$$M_{j,k} \le M_{j-1,k} - \frac{1}{2}M_{j-1,k}^2 + \frac{1}{2}M_{j-1,k-1}.$$

为说明这一点，考虑任意投票规则以及一个满足深度$(T_v) = j$的顶点 v。令 ℓ 和 r分别为 v的左右子节点。

我们将概率 $\Pr(E_{v,k})$ 分解为：

$$\begin{aligned} \Pr(E_{v,k}) = \Pr(E_{v,k}|E_{\ell,k-1}) \Pr(E_{\ell,k-1}) \\ + \Pr(E_{v,k}|E_{\ell,k} \setminus E_{\ell,k-1}) \Pr(E_{\ell,k} \setminus E_{\ell,k-1}) + \Pr(E_{v,k}|E_{\ell,k}^c) \Pr(E_{\ell,k}^c). \end{aligned}$$

注意，$\Pr(E_{\ell,k}) \le M_{j-1,k}$ 且 $\Pr(E_{\ell,k-1}) \le M_{j-1,k-1}$。下面分别对每个条件概率进行界定。在 $E_{\ell,k-1}$ 中，我们只需用 $\Pr(E_{v,k}|E_{\ell,k-1}) \le 1$ 作为上界。

在 $E_{\ell,k} \setminus E_{\ell,k-1}$ 中，可能的获胜者 x_v 有两种情况。首先，若 $x_v \in C_\ell$，由于之后所有挑战者都属于 C_r，因此必有 $x_v = x_\ell$。在这种情况下，若 $p_v = L$，则不会再出现软错误，且 $E_{v,k}$ 成立。但若 $p_v = R$，该路径会出现 $k+1$ 个软错误，因此 $E_{v,k}$ 不成立。其次，若 $x_v \in C_r$，且若 $p_v = L$，则会发生致命错误，因此 $E_{v,k}$ 不成立。因此，在这两种情况下，给定 $E_{\ell,k} \setminus E_{\ell,k-1}$ 时，$E_{v,k}$ 发生的概率至多为 $1/2$。

在 $E^c_{\ell,k}$ 中，必有 $x_v \neq x_\ell$，使得对于 $E_{v,k}$，$x_v \neq x_\ell$ 成立。因此，必有 $x_v \in C_r$，随后还必须满足 $p_v = R$，以避免致命错误。对于 $p_v = R$，当且仅当 $E_{r,k}$ 成立时，$E_{v,k}$ 才会发生。因此，

$$\Pr(E_{v,k}|E^c_{\ell,k}) = \Pr(E_{r,k}, p_v = R|E^c_{\ell,k}) = \frac{1}{2}\Pr(E_{r,k}|E^c_{\ell,k}) \leq \frac{1}{2}M_{j-1,k}.$$

这里第二个等式成立，是因为 p_v 与 $E_{\ell,k}$ 以及 $E_{r,k}$ 联合独立；而最后一个不等式成立，是因为 T_r 中的指针与先前历史相互独立。

因此，

$$\begin{aligned}
\Pr(E_{v,k}) &\leq \Pr(E_{\ell,k-1}) + \frac{1}{2}\left(\Pr(E_{\ell,k}) - \Pr(E_{\ell,k-1})\right) + \frac{1}{2}M_{j-1,k}(1 - \Pr(E_{\ell,k})) \\
&= \frac{1}{2}\Pr(E_{\ell,k-1}) + \frac{1}{2}M_{j-1,k} + \frac{1}{2}\Pr(E_{\ell,k})(1 - M_{j-1,k}) \\
&\leq \frac{1}{2}M_{j-1,k-1} + \frac{1}{2}M_{j-1,k} + \frac{1}{2}M_{j-1,k}(1 - M_{j-1,k}) \\
&= M_{j-1,k} - \frac{1}{2}M^2_{j-1,k} + \frac{1}{2}M_{j-1,k-1},
\end{aligned}$$

whe因为 $\Pr(E_{\ell,k-1}) \leq M_{j-1,k-1}$ 且 $\Pr(E_{\ell,k}) \leq M_{j-1,k}$，所以第二个不等式成立。

现在我们证明这一结论。我们对 k 使用归纳法，从 $M_{j,-1} = 0$ 开始。假设 $M_{j,k-1} \to 0$ 满足 $j \to \infty$，令 $M =$ 为 $\limsup_{j\to\infty} M_{j,k}$。由于函数 $t \mapsto t - \frac{1}{2}t^2$ 在 $[0,1]$, 上单调不减，对递推不等式两边取 $\limsup_{j\to\infty}$ 可得 $M \leq M - \frac{1}{2}M^2$，从而推出 $M = 0$。因此，$\lim_{j\to\infty} M_{j,k} = 0$。 –

失真。回想一下，我们选择了 q，使得 $\alpha < \varepsilon/4$。选择足够大的 k，使得 $(1-\alpha)^{k+1} \leq \alpha$。对于最终获胜者 x_T 和 $c^*_p = c_p(v_0)$，令 $E = E_{v_0,k}$ 表示从根节点 v_0 到 x_T 的路径上不包含致命错误且至多包含 k 个软错误这一事件。那么在 E^c中，我们有 $S(x_T, c^*_p) \leq \alpha$。因此，

$$\mathbb{E}[S(x_T, c^*_p)] \leq \Pr(E) + \alpha\Pr(E^c) \leq M_{d,k} + \alpha.$$

根据前面的结论，选择足够大的 d 可得 $\mathbb{E}[S(x_T, c^*_p)] < \varepsilon/2$。最后，对于指针 $p = (p_v)_v$的每一种实现，按如下方式定义选民—候选人距离：$d(i,c) = 1$，若 $c \gtrsim_i c^*_p$ 且$d(i,c) = 3$，否则。将选民之间的距离和候选人之间的距离设为 2，即得到一个有效度量。

在此度量下，每位候选人的社会成本至少为 n，而 c^*_p 的社会成本恰好为 n。对于候选人 x，$\mathrm{sc}(x) = n(3 - 2S(x, c^*_p))$ 成立，因此，

$$\mathbb{E}_p\left[\frac{\mathbb{E}[\mathrm{sc}(x_T)|p]}{\mathrm{sc}(c^*_p)}\right] = 3 - 2\mathbb{E}[S(x_T, c^*_p)] > 3 - \varepsilon,$$

其中，分子的内期望是相对于算法随机性取的。

因此，存在一种指针实现 p^*，使得 $\mathbb{E}[\mathrm{sc}(x_T)|p = p^*]/\mathrm{sc}(c^*_{p^*}) > 3-\varepsilon$。由于指针 p 是独立于算法随机性采样的，p^*，因此可以通过无知对手获得这个固定实例。

□

接下来，我们考虑自适应对手。由于算法 3保证失真为 $2+\sqrt{5}\approx 4.2361<5$，因此，要么持久随机独裁不是最优的，要么完全信息 IC 模型中的失真 5 分析并不紧。事实证明，即使在完全信息下，因子 5 也是紧的，这证明持久随机独裁在自适应对手下并非最优。

命题4.8 在自适应对手下，算法 *2* 无法保证低于5 的常数度量失真，在完全信息 *IC* 模型中。

证明。固定一个较小的 $\varepsilon>0$。考虑如下针对 n 个选民的实例，其中对于所有选民， $d(i,j)=2$ 。第一位候选人 c 到来时，对于每位选民， $d(i,c)=4$。然后对于 $i=1,\dots,n$，候选人 c'_i 在第 $i+1$ 轮到来，对于每位选民，其距离为 $d(i,c'_i)=3-\varepsilon$ 以及 $d(j,c'_i)=5-\varepsilon$。对于每个$j\neq i$，由于 $3-\varepsilon<4<5-\varepsilon$，每位选民 $i\in N$ 将 $c'_i\succ_i c\succ_i c'_j$ 排在所有 $j\in[n]\setminus\{i\}$之前。于是，在第 $n+1$ 轮之后，对手可以获知独裁者是谁： $i^\star\in N$，其 $x_{n+1}=c'_{i^\star}$。然后在第 $n+2$ 轮，对手引入候选人 $c^\star$，对于所有 $j\neq i^\star$，其距离为 $d(i^\star,c^\star)=3$ 以及 $d(j,c^\star)=1$。其余距离由最短路径距离给出。

在第 $n+2$ 轮，由于 $3-\varepsilon<3$，我们有 $c'_{i^\star}\succ_{i^\star}c^\star$ 和 $x_{n+2}=c'_{i^\star}$。社会成本为

$$\mathrm{sc}(c'_{i^\star})=(3-\varepsilon)+(n-1)(5-\varepsilon)=5n-2-n\varepsilon,$$

而

$$\mathrm{sc}(c^\star)=3+(n-1)\cdot 1=n+2.$$

因此，度量失真为

$$\frac{5n-2-n\varepsilon}{n+2}\to\frac{5n-2}{n+2},$$

因为 $\varepsilon\to 0$。令 $n\to\infty$。 □

这使得自适应对手的最优失真界位于区间 $3\leq\Delta^*\leq 2+\sqrt{5}$ 内。下界来自无知对手给出的下界，而上界来自一个确定性算法。我们将缩小这一差距留作一个开放问题。由于下界没有利用对手的自适应性，而上界没有利用随机性，我们猜想最优失真严格位于两个数 $\Delta^*\in(3,2+\sqrt{5})$ 之间。

5. 结论

我们研究了在线候选人到达情形下的度量失真，为度量失真理论提供了在线算法视角。基于信息空间和行动空间的两个维度，我们刻画了在确定性和随机化设置下能够实现最优失真的规则；唯一的例外是完全信息 IC 模型中对抗自适应对手的随机化投票规则。这一差距仍有待弥合。未来，可以研究这两个维度下、面对两种对手时的功利主义失真。

致谢

Shinsaku Sakaue 获得了日本 JST BOOST计划的资助，资助编号为 JP- MJBY24D1。

参考文献

Elliot Anshelevich、Onkar Bhardwaj、Edith Elkind、John Postl 和 Piotr Skowron。2018 年。《度量偏好下社会选择的近似》。*Artificial Intelligence* 264（2018），27–51。doi:10.1016/j.artint.2018.07.006 [→页。 2, 3, 16]

Elliot Anshelevich、Aris Filos-Ratsikas、Christopher Jerrett 和 Alexandros A. Voudouris。2025年。通过阈值批准改进度量失真。*Artificial Intelligence* 341 (2025), 104295。doi:10.1016/j.artint.2025。104295 [→ 页。 4]

Elliot Anshelevich、Aris Filos-Ratsikas、Nisarg Shah 和 Alexandros A. Voudouris。2021年。社会选择问题中的失真：前十五年及以后。载于第三十届国际人工智能联合会议（*IJCAI*）论文集。4294–4301。arXiv:2103.00911doi:10.24963/ijcai.2021/589 [→页。 3]

Elliot Anshelevich 和 John Postl。2017年。度量偏好下的随机化社会选择函数。人工智能研究杂志58 (2017), 797–827。doi:10.1613/jair.5340[→ 页。 3, 10]

Jeffrey S. Banks。1985 年。《复杂投票结果与议程控制》。*Social Choice andWelfare* 1, 4（1985），295–306。doi:10.1007/BF00649265 [→页。 4]

Duncan Black。1948 年。《群体决策的理论依据》。*Journal of Political Economy*56, 1（1948），23–34。doi:10.1086/256633 [→ 页。 4]

Craig Boutilier、Ioannis Caragiannis、Simi Haber、Tyler Lu、Ariel D. Procaccia 和 Or Sheffet。2015 年。《最优社会选择函数：功利主义视角》。*Artificial Intelligence* 227（2015），190–213。doi:10.1016/j.artint.2015.06.003 [→ 页。 3]

Robert Bredereck、Jiehua Chen、Rolf Niedermeier 和 Toby Walsh。2017 年。《议会投票程序：议程控制、操纵与不确定性》。*Journal of Artificial IntelligenceResearch* 59（2017），133–173。doi:10.1613/jair.5407 [→页。 4]

Ziyi Cai、Fabian Frank、JabariHastings、Prasanna Ramakrishnan、Kangning Wang 和 Qilin Ye。2026年。接近最优度量失真的稳定投票规则。arXiv:2609。08259 [cs.GT] https://arxiv.org/abs/2609.08259v2 2026年11月3日，版本 2。 [→第 2, 4]

Ioannis Caragiannis 和 Ariel D. Procaccia。2011年。尽管通信受限，投票几乎最大化社会福利。*Artificial Intelligence* 175, 9–10（2011年），1655–1671。doi:10.1016/j.artint.2011.03.005 [→ p. 3]

Moses Charikar 和 Prasanna Ramakrishnan。2022年。《随机化社会选择的度量失真界》。载于 *2022*年 *ACM-SIAM* 离散算法年度研讨会论文集（*SODA*）。SIAM，2986–3004。arXiv:2111。03694 [cs.GT] doi:10.1137/1.9781611977073.116[→ 第 2, 3]

Moses Charikar、Prasanna Ramakrishnan、Zihan Tan 和 Kangning Wang。2025年。《锦标赛投票及其他领域的度量失真》。载于 第*26*届 *ACM*经济学与计算会议（*EC*）论文集。ACM，790–818。arXiv:2505.13630doi:10.1145/3736252。3742629 [→ 页 4]

Moses Charikar、Prasanna Ramakrishnan、Kangning Wang 和 Hongxun Wu。2024年。突破度量投票失真障碍。*J. ACM* 71, 6（2024年），42:1–42:33。arXiv:2306。17838 [cs.GT]doi:10.1145/3689625 [→ p. 4]

Chris Dong 和 Jannik Peters。2025年。维护动态候选人集合中的比例代表委员会。载于 第*42*届国际机器学习会议论文集（*ICML*）（机器学习研究论文集，第*267*卷）。PMLR，14152–14168。https://proceedings.mlr.press/v267/dong25j.html [→第 4]

Soroush Ebadian、Anson Kahng、Dominik Peters 和 Nisarg Shah。2024年。投票中的优化失真与比例公平。 *ACM*经济学与计算汇刊 12, 1 (2024年), 3:1–3:39. doi:10.1145/3640760 [→页 3]

Robin Farquharson。1969年。投票理论。Blackwell。 [→页 4]

Fabian Frank 和 Jannik Peters。2026年。基于多数票民调的认可投票常数度量失真协议。arXiv:2608.28340 [cs.GT] https://arxiv.org/abs/2608.28340 [→页 4]

MohammadHossein Ghasemi、Masoud Seddighin 和 Saeed Seddighin。2026年。在线投票中的失真。arXiv:2611.25709 [cs.GT] https://arxiv.org/abs/2611.25709 [→ 页 3]

Vasilis Gkatzelis、Daniel Halpern 和 Nisarg Shah。2020年。解决最优度量失真猜想。发表于第*61*届*IEEE*计算机科学基础年会（*FOCS*）论文集。IEEE，1427–1438。arXiv:2004。07447 [cs.GT] doi:10.1109/FOCS46700.2020。00134 [→ 页。 2, 3]

David Kempe。2020年。基于线性规划对偶性的度量投票分析框架。发表于*AAAI*人工智能会议论文集，第34卷。2079–2086。doi:10.1609/ aaai.v34i02。5581 [→ 页。 7]

Fatih Erdem Kizilkaya 和 David Kempe。2022年。多数否决：实现最优度量失真的简单投票规则。发表于第三十一届国际人工智能联合会议（*IJCAI*）论文集。349–355。arXiv:2206。07098 [cs.GT] doi:10.24963/ijcai.2022/50[→ 页。 3, 16]

Jianan Lin 和 Chenhao Wang。2026年。确定性加权锦标赛投票规则的改进度量失真界。arXiv:2608。15247 [cs.GT] https://arxiv.org/abs/2608。15247v2 版本 2，2026年11月11日。 [→第 4]

Richard D. McKelvey 和 Richard G. Niemi。1978年。二元程序中复杂投票的多阶段博弈表示。《经济理论杂志》18，1（1978年），1–22。doi:10.1016/0022-0531(78)90039-X [→第 4]

Nicholas R. Miller。1977年。投票理论的图论方法。《美国政治科学杂志》21，4（1977年），769–803。 doi:10.2307/2110736 [→第 4]

Kamesh Munagala 和 Kangning Wang。2019年。《确定性社会选择规则的改进度量失真》。载于*2019*年 *ACM*经济学与计算会议（*EC*）论文集。ACM，245–262。arXiv:1905.01401 [cs.GT] doi:10.1145/3328526.3329550 [→ 第3、4、16页，17]

Joel Oren 和 Brendan Lucier。2014。《在线（预算约束）社会选择》。载于第二十八届 *AAAI*人工智能会议（*AAAI*）论文集。AAAI Press，1456–1462。doi:10.1609/aaai.v28i1.8891 [→第 4]

Grzegorz Pierczy´nski 和 Piotr Skowron。2019年。基于批准的选举与投票规则失真。载于第二十八届国际人工智能联合会议（*IJCAI*）论文集。543–549。arXiv:1901.06709doi:10.24963/ijcai.2019/77 [→页 4]

Ariel D. Procaccia 和 Jeffrey S. Rosenschein。2006年。投票中基数偏好的失真。载于第十届合作信息代理国际研讨会（*CIA*）论文集（计算机科学讲义，第*4149*卷）。Springer，317–331。doi:10.1007/11839354 23 [→ 页 3]

William H. Riker。1958年。《投票悖论与国会修正案投票规则》。《美国政治学评论》52, 2（1958年），349–366。doi:10.2307/1952321 [→ 第4]

AI 使用披露

我们以以下方式使用了 AI 模型（ChatGPT 6.2 Astra、6.1 Sol、6 Astra、5.6 Sol、5.5 和 5.4）。首先，我们用它们搜索该主题的相关研究。其次，我们用它们润色作者草稿的措辞和表达。第三，我们将它们用作额外工具，检查作者撰写的证明。最后，我们用它们为模型推导界（主要是下界）。从作者提供的较弱界出发，这些模型帮助我们将部分界加强为紧界结果。尽管部分界由 AI 推导得出，我们仍独立核查了所有最终论证，并对证明进行了大量改写，使其便于人类读者理解。作者对此承担责任。在线候选人到达模型包含两个维度的信息与行动空间，以及两种不同的对手，该模型由作者提出。

A. 附录：部分信息下的全候选人模型

现在我们考虑部分信息完整候选人模型。如第 3 节所述，自适应对手下确定性算法和随机化算法的最优界可自然推广，且无需更改证明。

为了将针对无知对手的随机化算法下界 3 推广到全候选人模型，我们首先证明以下引理。

引理 A.1 假设某个随机投票规则具有常数失真。若所有选民都更偏好 c_t而非 x_{t-1}，那么在部分信息完整候选人模型中，对每个 $s \geq t$, $x_s \in \{c_t, c_{t+1}, \ldots, c_T\}$ 几乎必然成立。

证明。 反设存在某个实例、轮次 t和 $s \geq t$，使得所有选民都更偏好 c_t而非 x_{t-1}，但 $x_s \notin \{c_t, c_{t+1}, \ldots, c_T\}$ 以正概率 $p > 0$。令 s为第一个满足这一条件的 $s \geq t$。由于截至轮次 s之前可能的观测和决策历史数量有限，因此存在一个以概率 $q > 0$ 发生且满足上述条件的历史 h 。

设 $M > 1$。考虑直线上的另一个实例，其中所有选民都位于 0，候选人 c_1、……、c_{t-1}位于 M，而所有其他候选人 c_t、c_{t+1}、……、c_T位于 1。于是，所有选民都更偏好 c_t而非 x_{t-1}。在 $C_{t-1} = \{c_1, \ldots, c_{t-1}\}$ 内部打破平局，使其与原实例的偏好剖面一致。因此，直到第 t 轮，算法都无法区分这两个实例。若 $s > t$，则在历史h中，c_t 被选中。因此，对于第 $t+1$、……、s轮，我们只比较 $\{c_t, c_{t+1}, \ldots, c_s\}$中的候选人。因此，我们可以在 $\{c_t, c_{t+1}, \ldots, c_s\}$ 内部打破平局，使其与原实例的偏好剖面一致。因此，当算法以正概率 q模拟历史h 时，它无法区分这两个实例。

因此在第 s 轮，$\{c_1, \ldots, c_{t-1}\}$ 中被选中的旧候选人的社会成本为 nM，而最优候选人为 n。因此，期望失真至少为 qM；当$M \to \infty$时，该值趋于无穷大，矛盾。

□

这使我们能够转移 3 的下界。

推论 A.2 在无知对手下，在部分信息全候选人模型中，没有任何随机投票规则能够保证低于 3 的常数度量失真。

证明。 我们修改 定理 3.13 的证明中的构造。此前，对于每个候选人$c \neq c_1$，我们都有一个关联集合 $B(c)$。现在，我们复制每个候选人，并让两个连续候选人共享同一个关联集合。

固定 $n \geq 2$和 $D < 3 - \frac{2}{n}$，并假设某个随机投票规则保证失真至多为D。考虑与前一个证明中相同的度量和抽样分布。令 $B(c_1) = N$。

对于每个 $r \geq 1$，根据以下规则独立抽取 $S_r \subseteq N$

$$\Pr(S_r = S) = \frac{\lambda^{-|S|}}{Z},$$

其中 $S \subseteq N$ 和 $Z = (1+\lambda^{-1})^n$。设 $B(c_{2r}) = B(c_{2r+1}) = S_r$。所有集合均预先抽样。回顾势能函数 $\phi(k) = \frac{Z}{\lambda - 1}\lambda^k$ 和 $\eta = \mathbb{E}[\phi(|S_r|)] = \frac{2^n}{\lambda - 1} > 0$。

我们修改势能函数，以考虑返回先前候选人的可能性；在完整候选人空间下运行的算法中，这种情况可能发生。每当算法选择 c_{2r} 时，所有选民都更偏好 c_{2r+1} 而非 c_{2r}，因为它们的关联集合相同，均为 S_r，且平局时优先选择较晚到达的候选人。根据前面的引理，所有先于 c_{2r+1} 到达的候选人都将永久不可用。我们称此事件为重置。

设 $M_0 = n$，并对每个 $r \geq 1$，定义

$$M_r = \begin{cases} |S_r|, & \text{if } x_{2r} = c_{2r}, \\ \max\{M_{r-1}, |S_r|\}, & \text{otherwise.} \end{cases}$$

量 M_r 界定了在已到达的候选人中，算法仍可能选择的所有候选人的关联集合大小。事实上，重置后，在迄今为止已看到的候选人中，只有 c_{2r+1} 仍然可用。否则，之前的界仍然成立，而两个新候选人的关联集合大小均为 $|S_r|$。我们跟踪势能 $\phi(M_r)$。（与之前跟踪 $\phi(|B(x_t)|)$ 的方式相比。）

固定 $r \geq 1$，并令 $b = |B(x_{2r-1})|$。根据前述观察，$b \leq M_{r-1}$。对于每个 $i \in N$，令 X_i 表示当且仅有选民 i 在第 $2r$ 轮更偏好现任者 x_{2r-1} 时，算法选择了挑战者 c_{2r} 以外的候选人。与前一个证明中一样，我们于是有

$$\mathbb{E}[X_i] \leq \frac{D-1}{2(n-1)}.$$

我们分析 $\phi(M_r)$ 的期望变化。对先前抽样得到的集合和算法的随机性进行条件化，使得 $B(x_{2r-1})$、b、M_{r-1}和 X_i 均已确定。对于每个$i \in B(x_{2r-1})$，考虑事件 $E_i = \{S_r = B(x_{2r-1}) \setminus \{i\}\}$，并令 $E = \bigcup_{i \in B(x_{2r-1})} E_i$。这些事件互不相交。

在 E_i 上，恰好选民 i 更偏好 x_{2r-1} 而非 c_{2r}。如果 $X_i = 0$，算法会选择 c_{2r}，从而导致重置，因此 $M_r = b - 1$。否则，$X_i = 1$，且 $M_r = M_{r-1}$，因为 $b - 1 \leq M_{r-1}$。因此，在 E_i 上，

$$\phi(M_r) - \phi(M_{r-1}) \leq (1 - X_i)(\phi(b-1) - \phi(b)).$$

由于 S_r 是独立于先前集合以及算法随机性进行采样的，在该条件化下，我们有$\Pr(E_i) = \frac{\lambda^{-(b-1)}}{Z}$。由此可得

$$\Pr(E_i)(\phi(b) - \phi(b-1)) = \frac{\lambda^{-(b-1)}}{Z}(\phi(b) - \phi(b-1)) = 1,$$

因此，E_i 的条件期望贡献至多为 $-(1 - X_i)$。

此外，根据 M_r 的定义，我们始终有 $M_r \leq \max\{M_{r-1}, |S_r|\}$，因此，由于 ϕ 非负且递增，

$$\phi(M_r) - \phi(M_{r-1}) \leq \phi(|S_r|).$$

因此，其余事件在期望中至多贡献 η。结合这些界并消除条件化，我们得到

$$\begin{aligned}\mathbb{E}[\phi(M_r) - \phi(M_{r-1})] &\le -\mathbb{E}[b] + \mathbb{E}\left[\sum_{i \in B(x_{2r-1})} X_i\right] + \eta \\ &\le -\mathbb{E}[b] + \sum_{i \in N} \mathbb{E}\left[X_i\right] + \eta \\ &\le -\mathbb{E}[b] + \beta + \eta,\end{aligned}$$

其中 $\beta = {}^{n}\!/_{(D-1)\ 2(n-1)} < 1$，与前面的证明相同。

其余论证也类似。由于 c_1 仍然是最优的，且社会成本为 n，我们有 $\mathbb{E}[b] \ge {}^{n}\!/_{(3-D)\ 2} > 1$。取足够大的 λ，使得 $\delta = 1 - \beta - \eta > 0$，可得$\mathbb{E}[\phi(M_r) - \phi(M_{r-1})] \le -\delta$。对 $r = 1, \ldots, R$ 求和，得到 $\mathbb{E}[\phi(M_R)] \le \phi(n) - R\delta$；再取足够大的 R，使右侧为负数，这与 ϕ 的非负性矛盾。 □