How Private Are Commonly-Used Voting Rules?*

Ao Liu[1], **Yun Lu**[2], **Lirong Xia**[1], **Vassilis Zikas**[2]
[1] Rensselaer Polytechnic Institute, liua6@rpi.edu, xial@cs.rpi.edu
[2] The University of Edinburgh, y.lu-59@sms.ed.ac.uk, vzikas@inf.ed.ac.uk

Abstract

Differential privacy has been widely applied to provide privacy guarantees by adding random noise to the function output. However, it inevitably fails in many high-stakes voting scenarios, where voting rules are required to be deterministic. In this work, we present the first framework for answering the question: *"How private are commonly-used voting rules?"* Our answers are two-fold. First, we show that deterministic voting rules provide sufficient privacy in the sense of *distributional differential privacy (DDP)*. We show that assuming the adversarial observer has uncertainty about individual votes, even publishing the histogram of votes achieves good DDP. Second, we introduce the notion of *exact* privacy to compare the privacy preserved in various commonly-studied voting rules, and obtain dichotomy theorems of exact DDP within a large subset of voting rules called *generalized scoring rules*.

1 INTRODUCTION

Differential privacy (DP) has gained much public attention recently, partly due to its use in the United States 2020 Census. Improving upon ad-hoc privacy techniques that were broken in the previous census [Garfinkel *et al.*, 2018], formal privacy definition like DP are much more suitable for controlling the leakage of sensitive data.

Yet, sensitive data is still published today without necessarily understanding the privacy leakage it incurs. In particular, voting data has been surprisingly accessible. In the US, histograms of votes are revealed per county, and voting and registration tables are released [US Census Bureau, 2019], which include fields like sex, race, age, location, and marital status. This abundance of information has enabled politicians to buy voter profiles from data mining companies to manipulate public opinion [Verini, 2007; Bradshaw and Howard, 2018].

Unfortunately, it is not easy to achieve (differential) privacy for voting. It is insufficient to protect voter registration tables with proven privacy techniques; releasing the election outcome can also be a cause of information leakage. To see how an individual's vote can be inferred by observing the winner of the election, we consider the following example. Suppose Alice cast a vote in an election, and then the winner is announced. Further suppose that an adversary can accurately estimate other votes from questionnaires or by machine learning from the other voters' social media, and it turns out these other votes ended up with a tie among the candidates. In this case, the adversary can distinguish Alice's vote even if he knows nothing about Alice, since Alice must have voted for the winner as the tie-breaker.

The strict definition of differential privacy means the mere *possibility* of the above scenario is a privacy violation. Moreover, ties do occur quite often in real life elections. For example, 9.2% of STV elections on Preflib election data [Mattei and Walsh, 2013] are tied [Wang *et al.*, 2019]. Even if we consider another formal privacy definition that accepts the uncertainly stemming from machine learning methods or low likelihood of ties as helpful in disguising votes, it is unclear how to quantitatively measure the effect of such uncertainty, and how

*VZ: Research supported in part by Sunday Group, Inc., and part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via 2019-1902070008. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein. LX: acknowledges NSF #1453542 and #1716333, and ONR #N00014-17-1-2621 for support. AL: acknowledges an IBM AIHN scholarship for support.

Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI), PMLR volume 124, 2020.

(or whether) privacy differs for different voting rules.

Motivated by the privacy concern in voting, we focus on the following key question in this paper.

How private are commonly-used voting rules?

The importance of answering this question is both practical and theoretical. On the practical side, minimizing the amount of information leakage from voting rules helps protect against censorship, coercion, and vote buying. On the theoretical side, privacy provides a new angle to comparing voting rules and designing new ones.

A first attempt would be to employ *differential privacy (DP)* [Dwork, 2006], measure of privacy widely-accepted and widely-applied in the cryptographic community. Mathematically, a voting rule M for $n \in \mathbb{N}$ voters is a mapping $\mathsf{M} : \mathcal{U}^n \to \mathcal{R}$, where $\mathcal{U}$ is the set of all possible votes; $\mathcal{R}$ is the set of all possible outcomes of voting, e.g. winners or histograms of votes. M is $(\ ;\ )-$*differentially private* if for any pair of *preference profiles* $\mathbf{X} \in \mathcal{U}^n$ and $\mathbf{X}^0 \in \mathcal{U}^n$ that only differ on one vote, and any subset of outcomes $\mathcal{S} \subseteq \mathcal{R}$, the following inequality holds:

$$\Pr\left[\mathsf{M}(\mathbf{X}) \in \mathcal{S}\right] \quad e \Pr\left[\mathsf{M}(\mathbf{X}') \in \mathcal{S}\right] + \ . \qquad (1)$$

Smaller , are desirable as it means the outcome of M is not affected much by one vote, and thus reveals little about an individual voter. Note in general M must be randomized to satisfy Inequality (1); indeed [Shang *et al.*, 2014; Lee, 2015; Hay *et al.*, 2017] achieved DP via randomized voting.

Yet most, if not all, voting rules used in high-stakes political elections are deterministic, since randomized voting rules suffer from difficulties in verifying implementation correctness, e.g. the controversy in the 2016 Democratic primary election in Iowa [Clayworth and Noble, 2016]. Unfortunately, the randomness in Inequality (1) comes from the voting rule itself, so deterministic rules cannot achieve DP except with the trivial parameter of ≥ 1, which always holds (see Example 1 for more details).

1.1 OUR CONTRIBUTIONS

To overcome the critical limitation of DP in high-stakes voting scenarios, we study the privacy of deterministic voting rules using *distributional differential privacy (DDP)* [Bassily *et al.*, 2013], a well-accepted notion of privacy that works for deterministic functions. DDP measures the amount of individual information leakage, assuming the adversary only has uncertain information about voter preferences, for example when using a machine learning algorithm. Our result on the DDP of commonly-used voting rules carries the following encouraging message:

Main Message 1: Many commonly-used voting rules achieve good DDP in natural settings.

More precisely, we focus on a natural DDP setting where the adversary's information is represented by a set of i.i.d. distribution's over preference profiles, denoted by $\subseteq (\mathcal{U})$, where $(\mathcal{U})$ is the set of all probability distributions over $\mathcal{U}$ *with full support*. A voting rule M's DDP is now measured by three parameters $(\ ;\ ;\ )$. A deterministic function is DDP (Definition 2) if it satisfies an inequality similar to Inequality (1), but now the randomness is replenished by the adversary's uncertainty about the profile $\mathbf{X}$, represented by . Like DP, smaller and in DDP are more desirable.

With DDP, we can quantitatively measure the privacy of the histogram rule Hist, which outputs the frequency of each type of vote in the preference profile, in the following Theorem 1. As an immediate consequence, many common voting rules also achieve good privacy.

Theorem 1 (DDP for Hist). *Given any* $\mathcal{U} = \{x_1; \ldots; x_l\}$ *and* $\subseteq (\mathcal{U})$ *with* $|\ | < \infty$, *let* $p_{\min} = \min_{2\Delta; i\ l}(\ (x_i))$. *For any* $n \in \mathbb{N}$ *and any* $\geq 2\ln\ 1 + \frac{1}{p_{\min} n}$, Hist *for* n *voters is* $(\ ;\ ;\ )$*-DDP where* $= \exp(-\ (np_{\min}[\min(2\ln(2);\ )]^2))$.

Theorem 1 states that Hist is private with good parameters, as even a small results in that is considered *negligible* in cryptography literature. The winner of many commonly-used voting rules depends only on the outcome of Hist, and thus contain (often stricly) less information than Hist. Thus, they achieve *at least as good* privacy w.r.t. DDP as simply outputting the histogram.

Next, we highlight that DDP (as well as DP and its variants) parameters only describe loose bounds on privacy—by definition, if a voting rule satisfies $(\ ;\ ;\ )$-DDP, it also satisfies $(\ + 0.1;\ + 0.1;\ )$-DDP. To compare the privacy-preserving capability of voting rules, we introduce the notion of *exact distributional differential privacy (eDDP)*, whose parameters describe tight bounds on and . We focus on the $= 0$ case as a first step to compare various voting rules with their eDDP in the parameter. Our results on the eDDP of commonly-used voting rules carry the following message:

Main Message 2: For many combinations of commonly-used voting rules and , the $(0; p\ ;\ )$-eDDP exhibits a dichotomy between $= (\ \sqrt{1/n})$ and $= \exp(-\ (n))$.

More precisely, we prove the following dichotomy theorem for two candidates $\{a; b\}$ and *-biased majority rules* with $\in (0; 1)$, which chooses a as the winner

iff at least n out of n votes prefer a.

Theorem 2 (Dichotomy in Exact DDP for -Majority Rules over Two Candidates, Informal) *Fix two candidates $\{a; b\}$ and $\subseteq (\{a; b\})$ with $| \ | < \infty$. For any $\in (0; 1)$, the -biased majority rule is $(0; \ ; \ )$-eDDP for all n, where is either $(\sqrt{1=n})$, when contains a distribution with $(a) = \ $, or exponentially small otherwise.*

For more than two candidates, we prove the following dichotomy theorem for a large family of voting rules and $\subseteq (\mathcal{U})$.

Theorem 3 (Dichotomy in Exact DDP of A Large Class of Voting Rules and , Informal) *For any fixed number of candidates, and any voting rule in a large family, the $(0; \ ; \ )$-eDDP is $= (\sqrt{1=n})$, when contains the uniform distribution, or $= \exp(- (n))$, when is finite and does not contain any unstable distributions.*

Intuitively, a distribution is *unstable* under a voting rule M if adding small perturbations can cause a different candidate to win (Definition 7). Instead of conducting case-by-case studies of eDDP for commonly-used voting rules, we prove Theorem 3 for a large family of voting rules called *generalized scoring rules* [Xia and Conitzer, 2008] that further satisfy *monotonicity, local stability*, and *canceling-out*. We show that many commonly-used voting rules satisfy these conditions (Section 5). We also compute and compare the concrete values for small elections (Table 1, Section 6 and Appendix E).

1.2 RELATED WORK

Differential privacy [Dwork, 2006] was recently used to add privacy to voting. Shang *et al.* (2014) applied Gaussian noise to the histogram of linear orders, while Hay *et al.* (2017) used Laplace and Exponential mechanisms applied to specific voting rules. Lee (2015) also developed a method of random selection of votes to achieve differential privacy. One interesting aspect of adding noise to the output that was observed in [Birrell and Pass, 2011; Lee, 2015] is that it enables an approximate strategy-proofness; the idea here is that the added noise dilutes the effect of any individual deviation, thereby making strategies which would slightly perturb the outcome irrelevant. We remark that if one wishes to achieve DP for a large number of voting rules, well-known DP mechanisms (like adding Laplace noise [Dwork *et al.*, 2006]) can be applied to rules in GSR in a straightforward way, by adding noise to each component of the score vector and outputting the winner based on the noised score vector. Our work is different because we focus on exact privacy of deterministic voting rules.

In our work, we compare deterministic functions by their exact privacy. In differential privacy literature where functions must be randomized, their accuracy, or utility, is used to compare them. A number of works have defined utility as a metric which describes the comparative desirability of -DP mechanisms. In [McSherry and Talwar, 2007], utility is an arbitrary user-defined function, used in the exponential mechanism. The works of [Blum *et al.*, 2008; Hardt and Talwar, 2010; Bassily and Smith, 2015] define utility in terms of error, where the closer (by some metric) the output of the function, which uses this mechanism to apply noise, is from the desired (deterministic) query's, the higher the utility; the definition of [Ghosh *et al.*, 2009] in addition allows the user to define as a parameter, the prior distribution on the query output. In contrast, our results imply that in the context of distributional differential privacy, voting rules achieve a well-accepted notion of privacy while preserving perfect accuracy, or utility.

1.3 DISCUSSIONS

While DP has been widely applied to measure privacy and has been applied to voting, as we discussed in the Introduction, it fails for deterministic functions such as voting rules in high-stakes elections. This critical limitation motivates our study. To the best of our knowledge, we are the first to illustrate how to measure privacy in high-stakes voting using (e)DDP in a natural setting. We will see that the problem, though challenging, can be solved by our novel *trails* technique. Below we explicitly discuss our conceptual and technical contributions and closely related works. More comprehensive discussions of related work can be found in Appendix A.

Conceptual contributions. Our first conceptual contribution is the application of DDP to deterministic voting rules. As discussed earlier, while previous works add random noise to achieve DP, to the best of our knowledge, no previous studies were done for deterministic voting rules. We note that the *truncated* histogram result of [Bassily *et al.*, 2013] does not suffice, since in general, votes are not removed in an election. Moreover, we prove our results in a simpler definition than DDP; the equivalence of this definition and DDP is proven in Appendix B.1. Our second conceptual contribution is the introduction of *exact DDP*, addressing the issue that parameters of DDP (and other relaxations of DP [Bassily *et al.*, 2013; Groce, 2014; Kasiviswanathan and Smith, 2008; Hall *et al.*, 2012; Duan, 2009; Bhaskar *et al.*, 2011]) describe only upper bounds on privacy. We are not aware of other works that explicitly propose to characterize tight bounds on the pri-

M	Borda	STV	Maximin	Plurality	2-approval
(n)	$\frac{1}{\sqrt{1.347n + 0.5263}}$	$\frac{1}{\sqrt{1.495n + 0.02669}}$	$\frac{1}{\sqrt{1.553n + 4.433}}$	$\frac{1}{\sqrt{1.717n - 0.09225}}$	$\frac{1}{\sqrt{1.786n + 0.3536}}$

Table 1: values in $(0; ;)$-eDDP for some commonly-used voting rules under the i.i.d. uniform distribution, $m = 3$ and n 50. From left to right, we rank rules from least to most private.

vacy parameters and .

Technical contributions. Our first theorem (Theorem 1) is quite positive, showing the privacy of outputting histograms. Theorem 2 and 3 characterize eDDP in terms of values by fixing $= 0$. We do so for the two reasons: (1) it is the common convention to compute based on a fixed for DP or DDP; (2) $= 0$ is the most informative choice, since Theorem 1 shows that even for small non-zero , any difference we can observe in the of two voting rules is exponentially small—considered negligible in cryptography literature. While our theorems appear similar and related to the dichotomy theorems on the probability of ties in voting [Xia and Conitzer, 2008; Xia, 2015], the definition and mathematical analysis are quite different, and previous techniques do not work for all cases; see more discussions in the proof sketch for Theorem 3. To address the challenge, we developed the *trails* technique, which significantly simplifies calculations.

Generality of our setting. As the first work towards answering our key question, we assume the adversary's beliefs are modeled by a set of i.i.d. distributions over the votes. A special case is the i.i.d. uniform distribution, which is known as the *impartial culture* assumption in social choice [Georges-Théodule, 1952]. Extending to general $(,)$, and non-i.i.d. distributions is an important and challenging future direction. Lastly, though our definitions and results are presented in the context of voting for the sake of presentation, they can easily be extended to general applications.

2 PRELIMINARIES

Let $\mathcal{C} = \{c_1; \cdots; c_m\}$ be a set of $m \geq 2$ candidates, and $\mathcal{L}(\mathcal{C})$ denote the set of all *linear orders* over $\mathcal{C}$, that is, the set of all antisymmetric, transitive, and total binary relations. Let $\mathcal{U}$ denote the set of all possible votes. Given $n \in \mathbb{N}$, we let $\mathbf{X} = (X_1; \cdots; X_n) \in \mathcal{U}^n$ denote a collection of n votes called a *preference profile*. Let $\mathcal{R}$ denote the set of outcomes of voting. A (deterministic) voting rule for n voters is a mapping $\mathsf{M} : \mathcal{U}^n \to \mathcal{R}$.

For example, in the *plurality* rule, $\mathcal{U} = \mathcal{R} = \mathcal{C}$; each voter votes for one favorite candidate, and the winner is the candidate with the most votes. In the *Borda* rule, $\mathcal{U} = \mathcal{L}(\mathcal{C})$ and $\mathcal{R} = \mathcal{C}$; each voter cast a linear order X over $\mathcal{C}$, denoted by $c_{i_1} \succ c_{i_2} \succ \cdots \succ c_{i_m}$, where $a \succ b$ means that a is preferred over b; each candidate c gets $m - i$ points in each vote, where i is the rank of c in the vote; the winner is the candidate with the highest total points. A tie-breaking mechanism is used when there are ties in plurality and Borda.

Definition 1 (The histogram rule)**.** *Let* $\mathcal{U} = \{x_1; \cdots; x_l\}$. *For any* $n \in \mathbb{N}$, *the* histogram function, *denoted by* $\mathsf{Hist} : \mathcal{U}^n \to \mathbb{N}^l$, *takes as input a preference profile* $\mathbf{X} = (X_1; \cdots; X_n) \in \mathcal{U}^n$ *and outputs a* l*-dimensional integer vector whose* i*th component is* $|\{j : X_j = x_i; j \in \{1; \cdots; n\}\}|$.

For example, when applied to the setting of the plurality rule, $l = m$ and Hist outputs the number of votes each candidate receives. When applied to the setting of the Borda rule, $l = m!$ and Hist outputs the number of occurrences of each linear order.

3 DISTRIBUTIONAL DIFFERENTIAL PRIVACY FOR VOTING

As we discussed, DP is not a suitable notion to analyze nontrivial deterministic voting rules as shown in the following example, which motivates our use of distributional differential privacy (DDP) [Bassily *et al.*, 2013].

Example 1 (DP fails for deterministic voting rules)**.** *Consider the plurality rule for two candidates* $\{a; b\}$ *and three voters* ($n = 3$)*. We have* $\mathcal{U} = \mathcal{R} = \{a; b\}$. *In Inequality (1), let* $\mathbf{X} = (a; a; b)$, $\mathbf{X}^0 = (b; a; b)$, *and* $\mathcal{S} = \{a\}$. *Then, (1) becomes* $1 \leq e \times 0 +$, *which means that* ≥ 1.

At a high level, the DDP of a (deterministic or randomized) function is characterized by three parameters $(; ;)$, where and are privacy parameters similar to DP, and is a set describing the adversary's knowledge about the preference profile. We consider adversaries that can be modeled as $\subseteq (\mathcal{U})$, which encodes each of the adversary's possible uncertainties as a distribution where each vote is i.i.d..

Example 2 (Adversary's information)**.** *Suppose* $\mathcal{U} = \mathcal{R} = \mathcal{C} = \{a; b\}$, *and the* n *votes could be i.i.d. generated from either* $_{0.2}$ *or* $_{0.7}$. *Here, for any* $\in [0; 1]$,

$(a) = \quad$. *Then, the adversary's information is represented by* $\quad = \{ \ _{0:2}; \ _{0:7}\}$. *Say we prove that some voting rule is* $(\ = 0.5; \ = 0.1; \ )$*-DDP for the above* $\quad$. *Intuitively, this means that the voting rule has privacy* $\ = 0.5; \ = 0.1$, *given the adversary's knowledge can be modeled by any distribution in* $\quad$. *We remark that this privacy holds* without *the need to add noise to the outcome of the election, contrasting with DP.*

To simplify presentation, below we will introduce the definition of DDP studied in this paper. In our setting of this paper, our simpler definition is equivalent to the original DDP. More details can be found in Appendix B.1.

Definition 2 (DDP studied in this paper)**.** *For any* $\subseteq (\mathcal{U})$, $\ > 0$, *and* $\ > 0$, *a voting rule* $\mathsf{M} : \mathcal{U}^n \to \mathcal{R}$ *is* $(\ ; \ ; \ )$*-DDP if for every* $\in \ , i \leq n, x; x^0 \in \mathcal{U}$, *and* $\mathcal{S} \subseteq \mathcal{R}$, *the following inequality holds.*

$$\begin{aligned} &\Pr_{\vec{X}\sim} (\mathsf{M}(\vec{X}) \ 2 \ S | X_i = x) \\ &e \ \Pr_{\vec{X}\sim} (\mathsf{M}(\vec{X}) \ 2 \ S | X_i = x') + \ ; \end{aligned} \tag{2}$$

where $\vec{X} = (X_1; : : : ; X_n)$ *is a preference profile where each vote is i.i.d. generated from* $\quad$.

For deterministic M, the randomness in Inequality (2) comes from the adversary's incomplete information, captured by $\quad$. We show that Hist satisfies good DDP.

Theorem 1 (DDP of Hist, proof in Appendix B.2)**.** *Given any* $\mathcal{U} = \{x_1; : : : ; x_l\}$ *and* $\quad \subseteq \quad (\mathcal{U})$ *with* $| \quad | < \infty$, *let* $p_{\min} = \min_{\ 2\Delta; i \ \ l}(\ (x_i))$. *For any* $n \in \mathbb{N}$ *and any* $\ \geq 2\ln(1 + \frac{1}{p_{\min} n})$, Hist *for* n *voters is* $(\ ; \ ; \ )$*-DDP where* $\ = \exp(- \ (np_{\min}[\min(2\ln(2); \ )]^2))$.

As corollary, these privacy parameters of Hist automatically apply to all functions that only depend on the output of Hist, i.e. most voting rules, or outputting the histogram in addition to the winner as in US presidential elections. This follows immediately from a property of DDP called *immunity to post processing* (see Lemma 3 in Appendix B.2). We note the result is similar to that of [Bassily *et al.*, 2013], but they assume lower-frequency items in the histogram are truncated (which is not the case in general when election results are posted) and describe a less precise $\quad$.

4 EXACT PRIVACY OF VOTING RULES: TWO-CANDIDATE CASE

In this section, we first present the definition of *exact distributional differential privacy* (exact DDP or eDDP), then characterize $(0; \ ; \ )$-eDDP for two candidates under any $\quad$-biased majority rule. The proof of this theorem will serve as a toy application of our *trails technique*, useful for proving our main result Theorem 3.

Intuitively, a function has *exact privacy* with parameters $\quad$ and $\quad$ if the function cannot satisfy the privacy definition with strictly better parameters. We remark that this definition can easily be altered to define $(\ ; \ )$*-exact differential privacy (eDP)* by omitting $\quad$.

Definition 3 (Exact Distributional Differential Privacy (eDDP))**.** *A voting rule* M *is* $(\ ; \ ; \ )$-Exact Distributional Differential Privacy (eDDP) *if it is* $(\ ; \ ; \ )$*-DDP and there does not exist* $(\ ^0 \leq \ ; \ ^0 < \ )$ *nor* $(\ ^0 < \ ; \ ^0 \leq \ )$ *such that* M *is* $(\ ^0; \ ^0; \ )$*-DDP.*

The $\quad$*-biased majority rule*, denoted by $\mathsf{M} \ \ \

easier to compute (below, $\vec{X} \sim$ is implicit).

$$= \max_{\mathcal{S};x;x^0;i;\ \in} \left[\Pr(f(\mathrm{Hist}(\vec{X})) \ 2\ \mathcal{S} j X_i = x) \right.$$
$$\left. \Pr(f(\mathrm{Hist}(\vec{X})) \ 2\ \mathcal{S} j X_i = x') \right]$$
$$= \max_{\mathcal{S};x;x^0;i;\ \in} \left[\sum_{t:\ f(t) \in \mathcal{S}} \Pr(\mathrm{Hist}(\vec{X}) = t j X_i = x) \right.$$
$$\left. \sum_{t:\ f(t) \in \mathcal{S}} \Pr(\mathrm{Hist}(\vec{X}) = t j X_i = x') \right]; \tag{4}$$

For example, if $\mathcal{S} = \{a\}$, then $\mathsf{T} \equiv \{t : f(t) \in \mathcal{S}\} = \{(5;0);(4;1);(3;2)\}$ is an example of what we call a *trail*. Intuitively, a trail T is a set of histograms *consecutive* in the sense that, starting from some t, we can list exactly the elements of T by iteratively subtracting 1 from and adding 1 to two components of t, respectively. We see that T can be listed in such a way, starting from entry $\mathrm{Enter}(\mathsf{T}) = (5;0)$ and ending at exit $\mathrm{Exit}(\mathsf{T}) = (3;2)$, by interatively subtracting from the first component and adding to the second component of $(5;0)$ (we say the *direction* of T is $(1;2)$). See Figure 1.

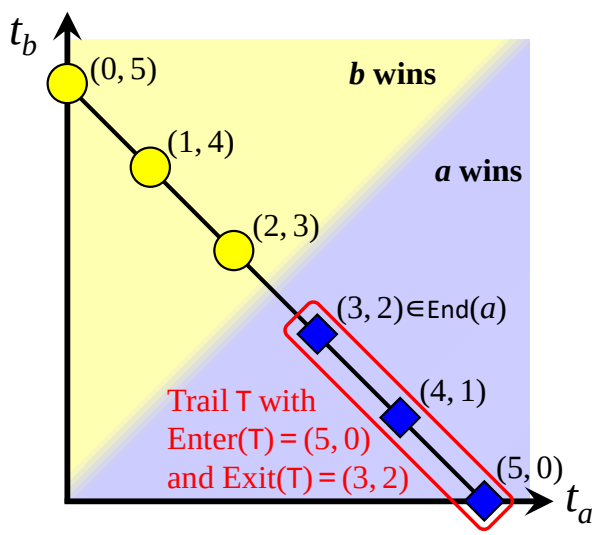

Figure 1: A trail for two candidates. A graph of number of votes for candidate a ($= t_a$) versus votes for candidate b ($= t_b$). Each point in the line is a histogram where the total number of votes is $n = 5$. The set $\{(5;0);(4;1);(3;2)\}$ forms a trail. We denote by $\mathrm{End}(a)$ (used in the proof of Theorem 3) the set of histograms which are exits of trails where a is the winner. In this example $\mathrm{End}(a) = \{(3;2)\}$.

We now give intuition for our key Lemma 1 presented below using this example. Suppose in Equation (4) the maximizing $\mathcal{S}$ is $\{a\}$ (so that $\{t : f(t) \in \mathcal{S}\} = \mathsf{T}$), $x = a$, and $x^0 = b$. Then, for any i, and any $\in$:

$$= \sum_{t \in \{(5;0);(4;1);(3;2)\}} \Pr(\mathrm{Hist}(\vec{X}) = t j X_i = a)$$
$$\sum_{t \in \{(5;0);(4;1);(3;2)\}} \Pr(\mathrm{Hist}(\vec{X}) = t j X_i = b);$$

The core of Lemma 1 is the observation that when votes are independent (e.g. when $\subseteq (\{a;b\})$), then for all $t = (t_a;t_b)$ such that $t_a > 0$, the following holds

$$\Pr(\mathrm{Hist}(\vec{X}) = (t_a;t_b) j X_i = a)$$
$$= \Pr(\mathrm{Hist}(\vec{X}) = (t_a \quad 1; t_b + 1) j X_i = b);$$

In light of this, $\Pr(\mathrm{Hist}(\vec{X}) = (5;0) | X_i = a)$ cancels out with $\Pr(\mathrm{Hist}(\vec{X}) = (4;1) | X_i = b)$, and $\Pr(\mathrm{Hist}(\vec{X}) = (4;1) | X_i = a)$ cancels out with $\Pr(\mathrm{Hist}(\vec{X}) = (3;2) | X_i = b)$. This leaves

$$= \Pr(\mathrm{Hist}(\vec{X}) = (3;2) = \mathrm{Exit}(\mathsf{T}) j X_i = a)$$
$$\Pr(\mathrm{Hist}(\vec{X}) = (5;0) = \mathrm{Enter}(\mathsf{T}) j X_i = b);$$

We note that here $\Pr(\mathrm{Hist}(X) = \mathrm{Enter}(\mathsf{T}) | X_i = b) = 0$, but this does not hold generally for all trails for $m \geq 2$. This calculation can be extended to the more general Lemma 1 below. Before that, let us formally define trails. For any histogram $t = (t_1;\cdots;t_l) \in \mathbb{N}^l$, any $z \in \mathbb{Z}$ and $j \leq l$, we let $(t_1;\cdots;t_l) + z x_j$ denote the histogram $(t_1;\cdots;t_j + z;\cdots t_l)$.

Definition 4 (Trails)**.** *Given a pair of indices $(j;k)$ where $j \neq k$, a histogram t, and a length q, we define the* trail $\mathsf{T}_{t;x_j;x_k;q} = \{t - z x_j + z x_k) : 0 \leq z \leq q\}$, *where $(j;k)$ is called the direction of the trail, t is then the* entry *of this trail, also denoted by Enter$(\mathsf{T}_{t;x_j;x_k;q})$, and $t - q x_j + q x_k$ is called the* exit *of the trail, denoted by Exit$(\mathsf{T}_{t;x_j;x_k;q})$.*

Alternatively, a trail T can be defined by just its entry and exit.

Lemma 1. *Let T be a trail with direction $(j;k)$, and let $\in (\mathcal{U})$. For any i, $x_j;x_k \in \mathcal{U}$, we have:*

$$\Pr_{\vec{X} \sim} (\mathrm{Hist}(\vec{X}) \ 2\ \mathsf{T}\ j\ X_i = x_j)$$
$$\Pr_{\vec{X} \sim} (\mathrm{Hist}(\vec{X}) \ 2\ \mathsf{T}\ j\ X_i = x_k)$$
$$= \Pr_{\vec{X} \sim} (\mathrm{Hist}(\vec{X}) = \mathit{Exit}(\mathsf{T})\ j\ X_i = x_j)$$
$$\Pr_{\vec{X} \sim} (\mathrm{Hist}(\vec{X}) = \mathit{Enter}(\mathsf{T})\ j\ X_i = x_k);$$

Proof. Fix distribution over n votes, where each vote is independently distributed. For $\vec{X} \sim$, denote $\vec{X}_{\ i}$ as the random variable $\vec{X}$ but without the ith vote. The equality in the lemma comes from the simple observation

that when votes are independently distributed, for any histogram $t \in \mathbb{N}^l$ and any $j \in [l]$

$$\Pr_{\vec{X} \sim} (\mathrm{Hist}(\vec{X}) = t | X_i = x_j) = \Pr_{\vec{X} \sim} (\mathrm{Hist}(X_{-i}) = t \quad x_j)$$

(Below, $\vec{X} \sim$ is implicit). Let q be the length of the trail. For any $0 \leq z < q$, let $t_z = \mathrm{Enter}(\mathsf{T}) - z x_j + z x_k$. Then,

$$\begin{aligned}
&\Pr(\mathrm{Hist}(\vec{X}) = t_z | X_i = x_j) \\
&= \Pr(\mathrm{Hist}(X_{-i}) = t_z \quad x_j) \\
&= \Pr(\mathrm{Hist}(\vec{X}) = t_z \quad x_j + x_k | X_i = x_k) \\
&= \Pr(\mathrm{Hist}(\vec{X}) = t_{z+1} | X_i = x_k):
\end{aligned}$$

In other words,

$$\begin{aligned}
&\Pr(\mathrm{Hist}(\vec{X}) \ 2 \ \mathsf{T} | X_i = x_j) \\
&\quad \Pr(\mathrm{Hist}(\vec{X}) \ 2 \ \mathsf{T} | X_i = x_k) \\
&= \Pr(\mathrm{Hist}(\vec{X}) = t_q | X_i = x_j) \\
&\quad \Pr(\mathrm{Hist}(\vec{X}) = t_0 | X_i = x_k) \\
&\quad + \sum_{0 \leq z < q} \Pr(\mathrm{Hist}(\vec{X}) = t_z | X_i = x_j) \\
&\qquad \Pr(\mathrm{Hist}(\vec{X}) = t_{z+1} | X_i = x_k) \\
&= \Pr(\mathrm{Hist}(\vec{X}) = t_q | X_i = x_j) \\
&\qquad \Pr(\mathrm{Hist}(\vec{X}) = t_0 | X_i = x_k)
\end{aligned}$$

(Every term in the summation of differences cancels out.)

$$\begin{aligned}
&= \Pr(\mathrm{Hist}(\vec{X}) = \mathrm{Exit}(\mathsf{T}) | X_i = x_j) \\
&\qquad \Pr(\mathrm{Hist}(\vec{X}) = \mathrm{Enter}(\mathsf{T}) | X_i = x_k)
\end{aligned}$$

□

Remark. In this subsection's example, no matter the $\mathcal{S}$, the set $\{t : \mathsf{f}(t) \in \mathcal{S}\}$ forms one single trail, but this does not hold in general. Instead, to prove our main theorem we will partition this set into multiple trails, and apply Lemma 1 to simplify probabilities over each trail.

4.2 A SIMPLE APPLICATION OF TRAILS TECHNIQUE: PROOF OF THEOREM 2

Proof. [Proof sketch for Theorem 2, see Appendix C.2 for the full proof]. For any $\in$, let $p = (a)$. Let trails $\mathsf{T}_a = \{t : t = (k; n-k); k \geq \ n\}$ and $\mathsf{T}_b = \{t : t = (k; n-k); k < \ n\}$. It follows that any histogram in T_a results in a being the winner, and any in T_b results in b as the winner. Also, Equation (4) implies we should *not* consider $\mathcal{S} = \{a; b\}$ nor $\mathcal{S} = \emptyset$ as otherwise $= 0$ (the default lower bound on). Thus, we only consider $\mathcal{S} = \{a\}$ (when winner is a, corresponding to trail T_a) or $\mathcal{S} = \{b\}$ (trail T_b). Then Equation (4) becomes (we disregard the value of i since votes are i.i.d.):

$$\begin{aligned}
&= \max_{j \in \{a;b\}; x; x^0} \Big[\Pr_{\vec{X} \sim} (\mathrm{Hist}(\vec{X}) \ 2 \ \mathsf{T}_j | X_i = x) \\
&\qquad \Pr_{\vec{X} \sim} (\mathrm{Hist}(\vec{X}) \ 2 \ \mathsf{T}_j | X_i = x') \Big] \quad \text{(Equation (4))} \\
&= \max_{j \in \{a;b\}; x; x^0} \Big[\Pr(\mathrm{Hist}(\vec{X}) = \mathrm{Exit}(\mathsf{T}_j) | X_i = x) \\
&\qquad \Pr(\mathrm{Hist}(\vec{X}) = \mathrm{Enter}(\mathsf{T}_j) | X_i = x') \Big] : \quad \text{(Lemma 1)}
\end{aligned}$$

We first discuss $\mathcal{S} = \{a\}$ whose corresponding trail T_a starts at $\mathrm{Enter}(\mathsf{T}_a) = (n; 0)$ and exits at $\mathrm{Exit}(\mathsf{T}_a) = (\lceil \ n \rceil; \lfloor (1 - \ ) n \rfloor)$. Here, $x = a$ and $x^0 = b$ maximize . Then,

$$\begin{aligned}
&\Pr(\mathrm{Hist}(\vec{X}) = \mathrm{Enter}(\mathsf{T}_a) | X_i = b) \\
&= \Pr(\mathrm{Hist}(\vec{X}) = (n; 0) | X_i = b) = 0;
\end{aligned}$$

and

$$\begin{aligned}
&\Pr(\mathrm{Hist}(\vec{X}) = \mathrm{Exit}(\mathsf{T}_a) | X_i = a) \\
&= \left(\frac{1}{n} \quad \frac{p}{} \quad \frac{1 \quad p}{1} \right)^{1 - \ \#_n} :
\end{aligned}$$

The case for $\mathcal{S} = \{b\}$ is similar and Theorem 2 follows by maximizing over $\in$. □

5 EXACT PRIVACY OF VOTING RULES: GENERAL CASE

The main result of this section, Theorem 3, characterizes $(0; \ ; \ )$-exact DDP of *generalized scoring rules (GSR)* for arbitrary number of candidates, defined below. The main message is that the characterization holds for commonly-used voting rules (Corollary 1). Therefore, to get the main message, a reader can skip the technical descriptions and definitions below to Corollary 1.

Definition 5 (Generalized Scoring Rules (GSR) [Xia and Conitzer, 2008]). *A Generalized Scoring Rule (GSR) is defined by a number $K \in \mathbb{N}$ and two functions $\mathsf{f} : \mathcal{L}(\mathcal{C}) \to \mathbb{R}^K$ and g, which maps weak orders over the set $\{1; \ldots; K\}$ to $\mathcal{C}$. Given a vote $V \in \mathcal{L}(\mathcal{C})$, $\mathsf{f}(V)$ is the* generalized score vector *of V. Given a profile P, we call $\mathsf{f}(P) = \sum_{V \in P} \mathsf{f}(V)$ the* score*. Then, the winner is given by $\mathsf{g}(\mathrm{Ord}(\mathsf{f}(P)))$, where Ord outputs the weak order of the K components in $\mathsf{f}(P)$.*

We say that a rule is a GSR if it can be described by some f, g as above. Most popular voting rules (i.e., Borda, Plurality, k-approval and ranked pairs) are GSRs. See Example 3 and Example 4 for f, g for plurality rule and majority rule. The domain of GSRs can be naturally extended to *weighted* profiles, where each type of vote is weighted by a real number, due to the linearity of f.

Example 3. *The simplest example of a GSR is* plurality. *This is the voting rule where each voter chooses exactly one candidate, and the candidate with the most votes is the winner. Here,* K *is equal to the number of candidates* m*. Suppose* V *is a vote (linear order over candidates) where the top candidate is* x_i*. The function* f *would map* V *to a vector* $\mathsf{f}(V) = (0; \cdots; 0; 1; 0; \cdots; 0)$ *where the 1 is at position* i *in the vector. Then,* $\mathsf{f}(P)$ *is exactly the histogram counting the number of times each candidate is ranked at the top of a vote. Finally, the function* g *chooses the winner.*

We now define a set of properties of GSRs to present our characterization of eDDP in Theorem 3.

Definition 6 (Canceling-out, Monotonicity, and Local stability)**.** *A voting rule* M *satisfies* canceling-out *if for any profile* $\mathcal{X}$*, adding a copy of every ranking does not change the winner. That is,* $\mathsf{M}(\mathcal{X}) = \mathsf{M}(\mathcal{X} \cup \mathcal{L}(C))$.

A voting rule satisfies monotonicity *one cannot prevent a candidate from winning by raising its ranking in a vote while maintaining the order of other candidates.*

A voting rule M *satisfies* local stability *if there exist locally stable profile. A profile* $\mathcal{X}$ *is* locally stable *(to* M*), if there exists a candidate* a*, a ranking* W*, and another ranking* V *that is obtained from* W *by raising the position of* a *without changing the order of other candidates, such that for any* $\mathcal{X}^0$ *in the neighborhood of* $\mathcal{X}$ *in terms of* L_1 *norm, we have (1)* $\mathsf{M}(\mathcal{X}^0) \neq a$*, and (2) the winner is* a *when all* W *votes in* $\mathcal{X}^0$ *becomes* V *votes.*

Definition 7 (Unstable distributions)**.** *Given a GSR* M*, a distribution over* $\mathcal{U}$ *is* unstable*, if for any* > 0*, there exists* $\mathbf{p}$ *and* $\mathbf{q}$ *with* $\|\mathbf{p}\|_2 = \|\mathbf{q}\|_2 <$ *, such that* $\mathsf{M}(+ \mathbf{q}) \neq \mathsf{M}(+ \mathbf{q})$[1]*, where* $\|\cdot\|_2$ *is the* ℓ_2*-norm.*

Theorem 3 (Dichotomy of Exact DDP for GSR, full proof in Appendix D.1)**.** *Fix* $m \geq 2$ *and* $\subseteq (\mathcal{L}(\mathcal{C}))$ *with* $| | < \infty$*. For any* n*, any GSR* M *that satisfies monotonicity, local stability, and canceling-out is* $(0; ;)$*-DDP, where is*

- $(\sqrt{1/n})$*, if contains the uniform distribution over* $\mathcal{L}(\mathcal{C})$*, or*
- $\exp(- (n))$*, if does not contain any unstable distribution.*

Proof sketch for Theorem 3. (See Appendix D.1 for the full proof) We first prove the $= \exp[- (n)]$ case. Recalling the proof of Theorem 2, we know that is closely related to the probability of End(a) for some $a \in \mathcal{C}$. It turns out that this is also the case for any GSR M that also satisfies monotonicity. Applying our trails technique, we have

$$\max_a \sum_{P \in \mathrm{End}(a)} \Pr(P \quad V);$$

where V is a vote s.t. there exists vote W with $\mathsf{M}(P - V + W) \neq a$. Thus, we know is upper bounded by the probability of all profiles $(P - V)$ "close" to a tie of voting rule r. For any unstable distribution , we can prove that the center of is reasonably "far" away from any profile in End(a) (or "far" away from any ties). Then, the exponential upper bound follows after Chernoff bound and union bound. The proof for this part is similar to the analysis of probabilities of tied profiles as in [Xia and Conitzer, 2008].

We now move on to the $= (\sqrt{1/n})$ case. The upper bound $O(\sqrt{1/n})$ also derived from the trails technique's result: $\leq \max_a \sqrt{\sum_{P \, 2 \mathrm{End}(a)} \Pr(P - V)}$. General framework of our proof is similar with the $= \exp[- (n)]$ case. Since adding any vote to a uniform profile results in a new winner, we know the uniform distribution of preferences is always an unstable distribution when requirements in Theorem 3 are met. Thus, we can prove that the center of the profiles' distribution (multinomial distribution in $m!$-dimensional space) is "close" to a tie. Then, we apply Stirling's formula to each trails and give an upper bounds to $\Pr(P - V)$ for profiles $P \in$ End(a).

For the lower bound $(\sqrt{1/n})$, canceling-out and locally stability are used to construct a "good" subset of profiles. At a high level, canceling-out ensures that the constructed subset is large enough, and locally stability ensures the trails constructed from the selected subset is long enough. Our subset is contracted by certain profiles with $O(\sqrt{n})$ distance[2] from the center of profile distribution in the direction of local stable profile. Giving a lower bound to the $\Pr(P - V)$ for any profile P in our selected subset is the most non-trivial part of this proof and is quite different from the proof in [Xia and Conitzer, 2008]. Unlike the profiles P in our selected subset of profiles, $P - V$ do not necessarily concentrated in a specific region in the space of profiles. Here, we use a non-i.i.d. version of Lindeberg-Levy central limit theorem [Greene, 2003] to analyze the multinomial distribution of $m!$ kinds of votes. □

Next, we use a simple example of majority rule to show the results in Theorem 3 matches the 2-candidate results in Section 4. In the following example, we also provide the intuitions on how to describe voting rules in the language of GSR.

[1] We slightly abuse notation—$\mathsf{M}( )$ denotes the output of M when the voters cast fractional votes according to .

[2] we use ℓ_2 distance in the $m!$-dimensional space of profile.