Effects of Privacy-Inducing Noise on Welfare and Influence of Referendum Systems

Suat Evren[1]
Massachusetts Institute of Technology (MIT), Cambridge, MA, USA

Praneeth Vepakomma
MIT Institute for Data, Systems and Society (IDSS), Massachusetts Institute of Technology (MIT), Cambridge, MA, USA
Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE

Abstract

Social choice functions help aggregate individual preferences while differentially private mechanisms provide formal privacy guarantees to release answers of queries operating on sensitive data. However, preserving differential privacy requires introducing noise to the system, and therefore may lead to undesired byproducts. Does an increase in the level of privacy for releasing the outputs of social choice functions increase or decrease the level of *influence* and *welfare*, and at what rate? In this paper, we mainly address this question in more precise terms in a referendum setting with two candidates when the celebrated randomized response mechanism is used. We show that the level of privacy is inversely proportional to society's welfare and influence.

2012 ACM Subject Classification Security and privacy → Economics of security and privacy; Applied computing → Economics; Theory of computation → Dynamic programming

Keywords and phrases Welfare, influence, social choice functions, differential privacy, randomized response

Digital Object Identifier 10.4230/LIPIcs.FORC.2024.1

Related Version *Full Version*: `https://arxiv.org/abs/2201.10115`

Funding Research was sponsored by the United States Air Force Research Laboratory and the Department of the Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Department of the Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.
Suat Evren: Partially supported by MIT UROP Direct Funding and MIT Schwarzman SERC Scholarship.
Praneeth Vepakomma: Researcher was partially supported by MBZUAI Startup Grant, ADIA Lab Fellowship, MIT Schwarzman SERC Scholarship and Meta PhD Research Fellowship.

Acknowledgements We would like to thank Alex Pentland, Ramesh Raskar, and Ashwin Sah for helpful comments, support and feedback.

1 Introduction

Differential privacy [6] provides a compelling privacy guarantee to ensure that the outcome of a query over any dataset is substantially not influenced based on the presence or absence of an individual's record. This form of privacy has recently been studied in the context of social choice theory [26, 15, 12]. A predominant strategy to achieve differential privacy in general

[1] Corresponding author.

5th Symposium on Foundations of Responsible Computing (FORC 2024).
Editor: Guy N. Rothblum; Article No. 1; pp. 1:1–1:20
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

even outside the context of social choice theory is to introduce noise or randomization to the system. One of the issues that has been widely studied in this context is the loss of accuracy upon introducing noise and the trade-off between accuracy and increasing levels of privacy preservation. This has been commonly referred to as the *privacy-accuracy* or *privacy-utility trade-off*. Recent work has involved the formalization of other trade-offs such as the trade-off between privacy and fairness [4]. In this work, we analyze two other trade-offs. We show that introducing noise to privatize systems that aggregate the preferences of individuals may affect several other fundamental phenomena such as *influence* and *welfare*.

Does an increase in the level of privacy for releasing the outputs of social choice functions, increase or decrease the level of influence and welfare, and at what rate? In this paper, we mainly address this question in more precise terms and affirmatively answer that this relation is inversely-proportional and shares specific corresponding rates for the popular *ρ-correlated randomized response* mechanism of privatization when used in a referendum setting with two candidates.

The noisy mechanism that we propose and analyze in regards to influence and welfare in this paper is based on a simple coin-flipping perturbation of the input as follows. Let ρ be an exogenous constant in $[0, 1]$ and let each original vote made in the ballot take a value of either 1 or -1. The randomized response records each original vote in the ballot as it is with a probability ρ while with probability $1 - \rho$, it ignores the original vote and instead records it as either a 1 or -1 with a uniformly random pick. The resulting probability space is known as *ρ-correlated distribution* or *noisy distribution* in the field of analysis of Boolean functions, and it is referred to as the *randomized response* mechanism in the field of differential privacy.[2] We show that this mechanism preserves ordinal relations between the influences of voters for *any* social choice function. Therefore, if Alice had more influence before than Bob, she will still continue to have more influence.

In the field of analysis of Boolean functions, the notion of the *influence* of a voter is used to measure the power of an individual on the final result of a social choice function. We extend this definition of influence to our probabilistic setting where noise is introduced for privacy, and term this new notion of influence as *probabilistic influence*. Similarly, we define *welfare* to address the second issue of capturing how *ideal* a voting rule is. First, we define it for deterministic functions and then we extend this definition to any probabilistic mechanism. We then show the effect of our privacy inducing randomized response on the welfare of the system. In particular, we show that it preserves the ordinal relations between the welfare of voting systems. That is, if a social choice function f had a greater welfare than g in the deterministic setting after the randomized response M_ρ is applied based on the exogenous parameter ρ, the welfare of $M_\rho f$ will continue to be greater than that of $M_\rho g$.

In this setting, we share precise statements connecting the noising probabilities ρ used in the mechanism M_ρ, their effect on level of privacy ϵ which in turn results in a specific level of influence and welfare expressed in terms of ρ. We precisely show that as the level of privacy increases, the welfare and influence happen to decrease at correspondingly specific rates. Arguably, having a higher welfare in a voting system is desirable and therefore we shine light on this new trade-off between privacy and welfare. In terms of influence, it is questionable whether a decrease in influence with an increase in privacy is desirable or not. We believe it depends on the context, and therefore in this case, we do not refer to it as a trade-off but instead call it a scaling law. However, as we show in Section 5, welfare of the society is equal to total influence of the society under the monotonicity assumption.

[2] For a survey of the field of analysis of Boolean functions, see [22]. For a survey of the field of differential privacy, see [7].

1.1 Contributions

We contribute towards bridging differential privacy and social choice theory by deriving the following results on the effect of randomized response over influence, welfare, and accuracy.

1. **Privacy-Influence scaling law:** A notion of *influence* is widely used in the analysis of Boolean functions to study social choice functions. We extend the notion of influence to the noisy setting, and call it *probabilistic influence*. We then show a result relating the trade-off between $\rho-$correlated distribution based differential privacy and probabilistic influence. We show that such privatization changes the influence of every single voter by a factor of $\frac{1+\rho^2}{2}$. Thus, the randomized response preserves the ordinal relations between influences of agents while scaling them by a factor depending on ρ while still ensuring their privacy is preserved.
2. **Privacy-Welfare trade-off:** We define *welfare* $W(f)$ of a social choice function f and extend the definition to probabilistic mechanisms. Then, we show that $W(M_\rho f) = \rho{\cdot}W(f)$, i.e. the randomized response scales the welfare by a factor of ρ, whereby preserving the ordinal relations between the welfare of social choice functions.
3. **Accuracy analysis:** We restrict the analysis of *accuracy*[3] of our mechanism to social choice functions, i.e. the functions with range $\{-1, 1\}$. We give the accuracy for Dictatorship, Majority, AND, and OR functions. For dictatorship, AND, and OR functions, we provide a theoretical analysis of accuracy. For the Majority function, we give an asymptotic accuracy when n goes to ∞ based on the existing results in the literature. We also give an exact analysis of accuracy for the Majority function for small n by using a computational method that involves dynamic programming.

1.2 Organization

The rest of the paper is organized as follows. In Section 2, we provide further motivation and background. In Section 3, we formally describe the differentially private randomized response mechanism. In Section 4, we introduce the notion of probabilistic influence, and give one of our main results that influence scales down by the same constant for every individual. In Section 5, we introduce the concept of welfare for general probabilistic mechanisms, and analyze it for randomized response. We shed light into the connection between influence and welfare, and give our second main result that randomized response scales down welfare by the same factor for any given social choice function. In Section 6, we provide an analysis of the accuracy for the randomized response mechanism. In Section 7, we discuss the possible future work and the limitations of this paper, and we conclude. Some preliminaries from social choice theory are provided in Appendix B. All of the proofs are relegated to Appendix A.

2 Motivation

To intuitively expand on the potential relation between privacy and influence, consider an instance where it might be the case that introduction of noise for the sake of obtaining privacy results in undesired shifts of the power held by different individuals in deciding society's final outcome. For example, say that a voter Alice would have had more impact on the outcome than Bob in a case where there is no privatization. It could as well be the case

[3] It is common to refer to *accuracy* with the name *utility* in the differential privacy literature. However, since this term is overloaded also in the economics and social choice theory literature with different meanings, we will opt to call it accuracy throughout our analysis to prevent possible confusion.

that the power balance shifts to Bob having more impact than Alice after a privacy-inducing noise is introduced. We conclusively show that this cannot be the case as the influence scales down for every voter with the increasing level of privacy by the same constant in the case of the popular randomized response privacy mechanism.

Second, regarding the potential relation between privacy and welfare, consider an instance where it may be the case that upon introduction of noise, the chosen social choice function that was originally used to aggregate the individual preferences into a final outcome ends up not being ideal anymore. Hence, it may instead be desirable to switch to another social choice function. For example, suppose that a system uses the majority function to decide which one of the two candidates is elected in the deterministic case. However, the majority function could be severely affected in some instances upon introduction of noise, and another function could end up being a *better* choice. We show that as the privacy increases in the randomized response mechanism, the welfare of each social choice function scales down proportionally under our definition of welfare, which is similar to the notions used in mathematical social choice theory. This implies that if a function is a welfare maximizer before introducing noise, it still is a welfare maximizer after the introduction of the noisy mechanism. These two results are especially useful, as they imply that the designers of the initial deterministic social choice mechanism do not have to be concerned about whether their design is robust to the introduction of noise in terms of influence and welfare.

We now discuss the work that has been done regarding influence and welfare in the context of social choice theory. Influences have long been studied in discrete Fourier analysis and theoretical computer science. The notion of influence was first introduced in [23] and it was first systematically studied in [3]. Some other novel works related to influences in the context of social choice theory include, but are not limited to, the KKL Theorem [14] and the Majority is Stablest Theorem [20]. We extend the notion of influence to the noisy setting and call it *probabilistic influence*, and prove a direct linear relation between deterministic influence and probabilistic influence.

The question of the ideal voting rule has long been a matter of discussion in social choice theory. When there are only two candidates, the answer is relatively simple as the majority function seems to be the most ideal voting rule. It is known that majority is the only social choice function that is anonymous and monotone among all two-candidate voting rules [19]. For more than two candidates, different objectives may result in different voting rules, or even in impossibility results [1, 2, 11, 9, 10]. Various aspects of utilitarian voting is studied in [13]. Finding the best function in computationally efficient ways has been studied in the recent field of computational social choice theory. There is a line of work [16, 17] that aims to maximize welfare given each voter's utility for candidates in a "distortion framework" in which there is a lack of information about voter's utilities. In that framework, a typical approach is to attempt to maximize the worst-case objective.

To the best of our knowledge, a definition of welfare that is closest to ours is the one given by O'Donnell ([22], page 51). Although the author does not explicitly define welfare of a social choice function, there is an affine relation between the expected value of their objective function and the way we define welfare. However, our main conceptual contribution is that our definitions are extended to hold for probabilistic mechanisms and we analyze the effects of privacy on influence and welfare. O'Donnel proves that among all two-candidate voting rules, majority is the unique maximizer of welfare, whose proof is essentially based on [27]. Our main objective is not to find the function that maximizes the welfare; that is rather a simple question. In fact, we show that majority is the unique welfare maximizer as well in an almost identical way to [22]. The primary motivation of the paper is to show that if a voting rule is better in the deterministic setting, it is still better after the privacy-inducing noise is introduced.

3 Model: Randomized Response and Privacy Guarantee

There are three main reasons as to why we chose the randomized response as the privacy-preserving mechanism to focus our attention. First, it is simple, in addition to being one of the earliest, and yet one of the most popularly used privacy-preserving mechanisms to date, be it in the classic form or as a variant of it. As an example, RAPPOR [8] is a recent popular real-world use-case of randomized response, otherwise classically used a few decades ago [28, 18]. Second, the mechanism is based on perturbations of the input which allows it to be applied to *any* social choice function. This enables us to talk about the ordinal relations between the welfare of potential social choice functions before and after the mechanism is applied. Third, ρ-correlated distributions are well studied in mathematical social choice theory [22].

Our randomized mechanism is an input-perturbing mechanism. That is, the mechanism introduces noise to the votes in the ballot so that one can use any social function afterward, yet the same privacy guarantee will continue to hold due to the post-processing property of differential privacy [5]. Randomized response introduces noise by utilizing a simple coin-flip scheme that is based on the following distribution that is widely used in the analysis of Boolean functions.

▶ **Definition 1.** *Let $\rho \in [0,1]$ and $x \in \{-1,1\}^n$ be fixed. y is called* ρ-correlated with x *if for every $i \in [n]$, $y_i = x_i$ with probability ρ and uniformly distributed with probability $1-\rho$, and it is denoted by $y \sim N_\rho x$.*

Note the symmetry in the definition of ρ-correlation. We formalize this symmetry in the following fact, which we will often use in the proofs of our results.

▶ **Observation 2.** *$x \sim \{-1,1\}^n, y \sim N_\rho x$ if and only if $y \sim \{-1,1\}^n, x \sim N_\rho y$. If $x \sim \{-1,1\}^n, y \sim N_\rho x$, we say (x,y) is a* ρ-correlated uniformly random pair*.*

In the literature, ρ-correlated distribution is sometimes referred to as *noisy distribution*. A famous analogy for this definition is as follows. Suppose the votes are recorded by a *noisy* machine. That is, the machine records each ballot correctly with probability ρ, and blurs the ballot with probability $1-\rho$ and instead records it at uniform random. As a result, the vote gets misrecorded with probability $(1-\rho)/2$. In fact, our mechanism corresponds to this noisy machine. Hence, we will call it by the generic name *randomized response*, or *ρ-correlated randomized response* when we need to specify ρ and denote a mechanism that applies it by M_ρ as defined below.[4] It is worth noting that *ρ-correlated randomized response* is in essence just like *randomized response* [28], a classic scheme that inspired several privacy mechanisms.

▶ **Definition 3.** *Let $f : \{-1,1\}^n \to \mathbb{R}$ be any function. For every $x \in \{-1,1\}^n$, the* randomized response *$M_\rho f(x)$ outputs $f(y)$ where $y \sim N_\rho x$.*

Now that we formally defined the randomized response mechanism, we can give the formal definition of differential privacy in our context.

▶ **Definition 4** (ϵ-Differential Privacy [6])**.** *A randomized voting mechanism $\mathcal{A} : \{-1,1\}^n \to \{-1,1\}$ is ϵ-differentially private if for all pair of neighboring voting profiles $\mathbf{x}, \mathbf{x}' \in \{-1,1\}^n$ that differ in exactly one bit and for all $\mathbf{s} \in \{-1,1\}$,*

$$\Pr[\mathcal{A}(\mathbf{x}) = \mathbf{s}] \le e^\epsilon \Pr[\mathcal{A}(\mathbf{x}') = \mathbf{s}]$$

[4] Note the subtle distinction between M_ρ and N_ρ. The former is a randomized query function, i.e. a random variable; whereas the latter denotes a probability distribution.

The above definition of differential privacy is specific to our context. For the general definition of differential privacy and a broad survey of the field, see [7]. The randomized response mechanism preserves ε-differential privacy. The following result holds for any Boolean function f.

▶ **Proposition 5.** *For any $\rho \in [0,1]$, randomized response $M_\rho f$ preserves $\log(\frac{1+\rho}{1-\rho})$-differential privacy regardless of the function $f : \{-1,1\}^n \to \mathbb{R}$. (or, ($\varepsilon$,0)-differential privacy when $\rho \leq 1 - \frac{2}{\exp(\varepsilon)+1}$).*

Proof. Proof is relegated to Appendix A.1. ◀

▶ Remark 6. The equality case is satisfied if f is a dictatorship, which implies that the bound $\log(\frac{1+\rho}{1-\rho})$ is tight. That is, when f is a dictatorship, $M_\rho f$ is not ε-differentially private for any $\varepsilon < \log(\frac{1+\rho}{1-\rho})$. In fact, it can be shown that a social choice function f satisfies the equality case if and only if there is a triple (r,b,i) where $r \in \mathbb{R}, b \in \{-1,1\}, i \in [n]$ such that $\emptyset \neq \{z \in \{-1,1\}^n | f(z) = r\} \subseteq \{z \in \{-1,1\}^n | z_i = b\}$.

The reason our mechanism preserves differential privacy for any Boolean function f is that the mechanism is input-perturbing. In this sense, we could instead present the mechanism as $M_\rho : \{-1,1\}^n \to \{-1,1\}^n$ and write $f \circ M_\rho$ instead of $M_\rho f$. Then we could prove the analogous version of Proposition 5, and by using the post-processing property of differential privacy, we would again obtain Proposition 5. In fact, one can see that in the proof, we also re-prove the post-processing property, seemingly for no reason. However, the reason we choose to give the mechanism altogether after post-processing with f is to make the all equality cases in the above remark apparent. Once post-processing is applied black-box, whether the privacy result is robust is not clear anymore. For example, consider any constant function f, e.g. $f(x) = 1$ for any $x \in \{-1,1\}^n$. In this case, $M_\rho f$ is not only $\log(\frac{1+\rho}{1-\rho})$-differentially private but 0-differentially private. On the other hand, as Remark 6 implies, the privacy guarantee in Proposition 5 is tight, which we would not be able to show without an explicit proof.

4 Probabilistic Influence

Influence of a voter is a notion that is used to measure the power of an individual on a deterministic social choice function. Influences of Boolean functions have long been studied in computer science and the field of analysis of Boolean functions starting with [3]. The *influence* of a voter in a voting system is defined to be the probability of the change in outcome when the voter changes their vote *ceteris paribus*. For example, in the case of a dictatorship, the dictator has influence 1 while every other voter has influence 0. In the majority function with $n = 2k+1$ voters, each voter's influence is the same and equal to $\binom{2k}{k}/2^{2k}$.

We use $x_{i\to 1} = (x_1, \cdots, x_{i-1}, 1, x_{i+1}, \cdots, x_n)$ to denote the case where the i-th voter chooses to vote for 1, and every other voter follows x. Similarly, we denote the alternate case where the i-th voter chooses to vote for -1 and every other voter follows x by $x_{i\to -1} = (x_1, \cdots, x_{i-1}, -1, x_{i+1}, \cdots, x_n)$. Using this notation, influence in the deterministic setting is defined as follows.

▶ **Definition 7.** *For $f : \{-1,1\}^n \to \{-1,1\}$, the influence of elector i is defined as*

$$I_i[f] = \mathbb{P}_{x\sim\{-1,1\}^n}[f(x_{i\to 1}) \neq f(x_{i\to -1})].$$

The total influence of the function f is defined to be

$$I[f] = \sum_{i=1}^{n} I_i[f].$$

A similar notion can be introduced in the probabilistic setting where the randomized response $M_\rho f(x)$ is applied. To do so, we consider the case where everybody casts their votes, following which $M_\rho f(x)$ is applied and the voter i changes their vote. That is, we leave all the noisy versions of the votes cast by everyone as is except for the elector i's vote. For this particular vote, we re-run the randomized response on coordinate i. The probability of result being different is called the *probabilistic influence* of coordinate i. We now introduce the formal definition of the proposed probabilistic influence, which applies not only to social choice functions with range $\{-1,1\}$ but to all Boolean functions with range in $\mathbb{R}$ as follows. In the notation of the following definition, $y_i \sim N_\rho(1)$ refers to the case where voter i chooses to vote for 1 while $z_i \sim N_\rho(-1)$ refers to the case where voter i chooses to vote for -1.

▶ **Definition 8.** *Let $f : \{-1,1\}^n \to \mathbb{R}$ and the probabilistic influence of coordinate i in a mechanism $M_\rho f(x)$ is defined as*

$$I_i[M_\rho f] = \mathbb{E}_{x\sim\{-1,1\}^n, \forall j\neq i\ z_j=y_j=x_j, y_i\sim N_\rho(1), z_i\sim N_\rho(-1)}\left[\left(\frac{f(y)-f(z)}{2}\right)^2\right].$$

The total influence of the mechanism $M_\rho f$ is defined to be

$$I[M_\rho f] = \sum_{i=1}^{n} I_i[M_\rho f].$$

We showed in Proposition 5 that our probabilistic voting mechanism preserves ε-differential privacy. Inducing such privacy requires probabilistic mechanisms as opposed to using deterministic functions. For example, in the majority voting with $2k+1$ voters, if the votes are split k to $k+1$, then changing only one bit in the input may change the outcome of the voting mechanism. Thus, it is not differentially private. Similarly, no deterministic Boolean function can preserve differential privacy unless it is a constant function.

On the other hand, introducing noise may cause several issues in the voting system, one of which is the accuracy of the mechanism, which we will discuss in more detail in Section 6. Another possible issue is that when noise is introduced, we might be altering the voting system in favor of a particular voter. For example, voter A might have more influence relative to voter B in the system now even if that was not the case before. For symmetric social choice functions, it is natural to expect that the randomized response mechanism would have the same effect for any voter since the noise is also symmetric. However, it is not as trivial for arbitrary social choice functions. Yet, we show that each voter's probabilistic influence is proportional to her influence in the deterministic setting, which is one of our main results.

▶ **Theorem 9.** *Let $\rho \in [0,1]$ be any real number and $f : \{-1,1\}^n \to \mathbb{R}$ be any function. For every $i \in [n]$, $I_i[M_\rho f] = \frac{1+\rho^2}{2} I_i[f]$.*

Proof. Proof is relegated to Appendix A.2. ◀

This result shows that the randomized response preserves the ordinal relations between influences of the voters regardless of the original social choice function being used. In other words, if voter A had greater influence than another voter B, she will still have a greater influence on the system after the noise is introduced.

5 Welfare

In this section, we introduce a formal definition of *welfare* of social choice functions. Then we extend this definition to probabilistic mechanisms, and we show that the randomized response preserves the ordinal relations between the welfare of social choice functions.

5.1 Welfare of Deterministic Voting Systems

[24] argues in his *Social Contract* that an ideal voting rule should maximize the number of votes that agree with the outcome. For a more comprehensive discussion on this, see [25]. [22] proves that the majority function is the unique ideal function based on Rousseau's perception of the ideal voting rule without formally introducing welfare. Perhaps, when he proved this result, he had some form of welfare in his mind, especially because he uses the letter w to denote the number of votes that agrees with the outcome. In this section, we will formally define welfare, which will be slightly different than what the w notation of O'Donnell describes. In particular, we define *welfare* of a social choice function $f : \{-1,1\}^n \to \{-1,1\}$ as the average difference between the number of votes that agree with the outcome and the number of votes that do not agree with the outcome under the impartial culture assumption.

▶ **Definition 10.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ and $x \in \{-1,1\}^n$, and let $w_x(f) = |\{i; x_i = f(x)\}| - |\{i; x_i \neq f(x)\}|$. Welfare of the social choice function f is defined to be*

$$W(f) = \mathbb{E}_x[w_x(f)].$$

We can still prove that the majority function is the unique maximizer of welfare when n is odd by using a similar method as in the proof of Theorem 2.33 in [22].

▶ **Proposition 11.** *When n is odd, the unique maximizer of $W(f)$ is the majority function.*

Proof. Proof is relegated to A.3. ◀

Without further assessment, it is not possible to say whether we prefer total influence to be larger or smaller for the welfare of society in a voting system. As we show in the following result, if the social choice function is monotone – that is if a voter changes her vote in favor of a candidate, then this candidate should be weakly better off – then these two notions collide.

▶ **Proposition 12.** *Let f be any monotone social choice function $f : \{-1,1\}^n \to \{-1,1\}$. Then, $W(f) = I[f]$.*

Proof. Proof is relegated to Appendix A.4. ◀

This result has implications beyond being a simple identity, making the case that if we want to achieve a greater social welfare while adhering to monotone social choice functions, we must choose a function with a greater total influence.

5.2 Welfare of Noisy Mechanisms

To capture the same notion for the probabilistic functions as well, we similarly define welfare of a randomized mechanism applied on a social choice function as follows. Note that the following definition is not only for the randomized response M_ρ, but any mechanism defined on social choice functions.

▶ **Definition 13.** *Let $f : \{-1,1\}^n \to \{-1,1\}$, $x \in \{-1,1\}^n$, and M be any mechanism. Let $w_x(Mf) = |\{i; x_i = Mf(x)\}| - |\{i; x_i \neq Mf(x)\}|$. Welfare of the mechanism M with the social choice function f is defined to be*

$$W(Mf) = \mathbb{E}_{x,M}[w_x(Mf)]$$

where the expectation is both over x and the mechanism M.

We showed in Theorem 9 that although introducing ρ-correlated noise in a voting system has negative effects on influences, it does not provide an unfair advantage to any agent. Another possible undesired byproduct of a randomized mechanism could be that the effect of randomization on the welfare of a particular voting system is more severe compared to the other voting systems. For example, we showed in Proposition 11 that the majority function is the unique welfare maximizer. It could be the case that after we introduce noise, it is more likely in the majority function that the outcome will change. Within this context, the following result implies that every voting system is equally affected by the input-perturbing randomized response mechanism. Therefore the randomized response preserves the ordinal relations between the welfare of two-candidate voting systems.

▶ **Theorem 14.** *Let f be any social choice function $f : \{-1,1\}^n \to \{-1,1\}$. Then, $W(M_\rho f) = \rho \cdot W(f)$.*

Proof. Proof is relegated to Appendix A.5 ◀

This result, together with Proposition 11, implies that the majority function is the unique welfare maximizer also after the noise is introduced in applying the randomized response mechanism.

6 Accuracy Analysis

There is one significant drawback of the randomized response privatization mechanism in consideration. It is hard to analyze the accuracy of releasing the output of social choice functions upon privatizing it with the randomized response. Although our main objective in this work is not about the analysis of accuracy, we will dedicate a section to the analysis of accuracy for the sake of completeness. As a first pass, we easily find a *generic* lower-bound on accuracy of the randomized response, but it ends up to be so low that it makes it redundant. Therefore, we restrict our analysis to *specific* social choice functions. We theoretically provide results on accuracy for dictatorship, AND, and OR functions.[5] In addition, we give a tight lower bound as well as an upper bound for the accuracy of majority function. We also give an algorithm to calculate exact accuracy of majority function by using dynamic programming via memoization. The dynamic programming approach avoids the need to make calculations over every entry in the power-set and instead is much more efficient, while still resulting in an exact solution for computing the accuracy. Our definition of accuracy is in-fact the average of accuracy under the impartial culture assumption. That is,

$$Acc(M_\rho f) = \mathbb{P}_{\substack{x \sim \{-1,1\}^n \\ M_\rho}}[M_\rho f(x) = f(x)].$$

Now, we define the *noise operator*, also referred to as the noisy Markov operator, which is a linear operator on the set of Boolean functions. This operator will be useful for accuracy calculations.

▶ **Definition 15.** *For any $\rho \in [0,1]$, the noise operator T_ρ is the linear operator on the set of functions $f : \{-1,1\} \to \mathbb{R}$ defined by*

$$T_\rho f(x) = \mathbb{E}_{y \sim N_\rho x}[f(y)].$$

[5] For formal definitions of these widely known social choice functions, see Appendix B.

Before we start our analysis, let us also give the definition of *noise stability.*

▶ **Definition 16.** *For any* $\rho \in [0,1]$ *and* $f : \{-1,1\}^n \to \mathbb{R}$*,* ρ*-correlated noise stability of* f *is given by*

$$Stab_\rho(f) = \mathbb{E}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho(x)}}[f(x)\cdot f(y)]$$

There is a linear relation between the noise stability of a function and accuracy of the randomized response on this function. Note that $M_\rho f(x)\cdot f(x) = 1$ if $M_\rho f(x) = f(x)$, $M_\rho f(x)\cdot f(x) = -1$ otherwise. Thus,

$$2\cdot Acc(M_\rho f)-1 = 2\cdot\mathbb{P}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho(x)}}[f(y)=f(x)]-1 = \mathbb{E}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho(x)}}[f(y)\cdot f(x)] = Stab_\rho(f). \quad (1)$$

Also, note that

$$Stab_\rho(f) = \mathbb{E}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho(x)}}[f(x)\cdot f(y)] = \mathbb{E}_{x\sim\{-1,1\}^n}[f(x)T_\rho f(x)]. \quad (2)$$

The reason we feel the need to write accuracy in terms of stability is that in the field of Analysis of Boolean functions most results are given in terms of stability for convenience. Yet, we use stability explicitly only when we analyze the accuracy of the majority function.

6.1 Majority

In this section, we will give the asymptotic accuracy for Maj_n function where n is an odd number that goes to infinity.

▶ **Lemma 17** (Proposition 10, [21]). *For any* $\rho \in [0,1)$*,* $Stab_\rho[Maj_n]$ *is a decreasing function of* n *where* n *is an odd number, with*

$$\frac{2}{\pi}\arcsin(\rho) \le Stab_\rho[Maj_n] \le \frac{2}{\pi}\arcsin(\rho) + O(\frac{1}{\sqrt{1-\rho^2}\sqrt{n}}).$$

By using the fact that accuracy is equal to $\frac{1}{2}+\frac{1}{2}Stab_\rho(f)$ due to Equation (1), we get that

$$\frac{1}{2}+\frac{1}{\pi}\arcsin(\rho) \le Acc[M_\rho(Maj_n)] \le \frac{1}{2}+\frac{1}{\pi}\arcsin(\rho) + O(\frac{1}{\sqrt{1-\rho^2}\sqrt{n}}). \quad (3)$$

Despite this fact being quite useful, there is no convenient way to calculate the exact value of accuracy of the randomized response on Majority function. Hence, we compute it using dynamic programming via memoization in the following section.

6.1.1 Algorithm to compute the exact accuracy for small n

We now provide a dynamic programming algorithm with memoization to compute the accuracy of the randomized response. In particular, we give the algorithm to calculate the accuracy of the *threshold functions*, that are of the form

$$f_\theta(x) = \begin{cases} 1 & \text{if } \sum_{i\in[n]} x_i > \theta \\ -1 & \text{if } \sum_{i\in[n]} x_i \le \theta \end{cases}$$

Note that $Maj_n = f_0(\cdot)$ where it takes care of ties by considering them as if -1 is the winner. In general, we work with the odd number of voters when we talk about the majority function. But as a simple trick, we will compute it for any n based on the generic definition of the threshold function we gave above since it makes the algorithm less involved.

We now state the noise operator $T_\rho f_{\theta_0}(x)$ as introduced in Definition 15 when applied to threshold functions as a way to quantify the expected accuracy as

$$T_\rho f_{\theta_0}(x) = \mathbb{E}_{y\sim N_\rho x}\left[\mathbb{1}\left(y_1 + \ldots y_n > \theta_0\right)\right].$$

Let x_{-n} denote x without the last bit. In particular, if $x = (x_1, x_2, \cdots, x_{n-1}, x_n)$, then $x_{-n} = (x_1, x_2, \cdots, x_{n-1})$. Note that $x_{-n} \in \{-1,1\}^{n-1}$ while $x \in \{-1,1\}^n$. Then, the stability can be defined using two calls of recursion as follows

$$T_\rho f_{\theta_0}(x) = \frac{1+\rho}{2} T_\rho f_{\theta_0 - x_n}(x_{-n}) + \frac{1-\rho}{2} T_\rho f_{\theta_0 + x_n}(x_{-n})$$

That is because

$$\begin{aligned}
&\mathbb{E}_{y\sim N_\rho x}\left[\mathbb{1}\left(y_1 + \cdots + y_n > \theta_0\right)\right] \\
&\qquad = \mathbb{E}_{y_n\sim N_\rho x_n}\left[\mathbb{E}_{y_{-n}\sim N_\rho x_{-n}}\left[\mathbb{1}\left(y_1 + \cdots + y_{n-1} > \theta_0 - y_n\right) \mid y_n\right]\right] \\
&\qquad = \frac{1+\rho}{2} \mathop{\mathbb{E}}_{y_{-n}\sim N_\rho(x_{-n})}\left[\mathbb{1}\left(y_1 + \cdots + y_{n-1} > \theta_0 - x_n\right)\right] \\
&\qquad\qquad + \frac{1-\rho}{2} \mathop{\mathbb{E}}_{y_{-n}\sim N_\rho(x_{-n})}\left[\mathbb{1}\left(y_1 + \cdots + y_{n-1} > \theta_0 + x_n\right)\right] \\
&\qquad = \frac{1+\rho}{2} T_\rho f_{\theta_0 - x_n}(x_{-n}) + \frac{1-\rho}{2} T_\rho f_{\theta_0 + x_n}(x_{-n})
\end{aligned}$$

To summarize, this dynamic programming with memoization algorithm is as shown in Algorithm 1 below. In terms of notation we denote a specific dictionary (in terms of popular programming terminology of dictionary data types) as Dictionary: $\{(\rho, n, s, \theta) = T_\rho f_{\theta_0}(x)$ for some x s.t $sum(x) = s\}$.

Our approach is to use this proposed recursive relation with an appropriate initial condition to exactly compute the noise operator $T_\rho f(x)$. Then, by using Equation (2), we calculate the Stability of the function. Finally, by using the linear relation between stability and accuracy from Equation (1), we compute the exact accuracy. This dynamic programming approach avoids having to make 2^n computations, given that $x \sim \{-1,1\}^n$. Note that, $T_p f_{\theta_0}(x) = T_p f_{\theta_0}(z)$ if $sum(x) = sum(z)$. Therefore we iterate over i from 1 to n to represent vectors with i number of $1's$. Then as the rest of entries are -1, and since the length of the array is n, this approach can model the exact sum of all possible vectors. Since the calculation of the stability is one-to-one with respect to sums, we store the intermediate results in a dictionary indexed by this sum. As there are $\binom{n}{i}$ vectors that can be represented this way, we just compute once per each i and multiply it by $\binom{n}{i}$. This enables us to model all possible vectors efficiently but allows us to not have to compute the intermediate results every time via our recursive approach.

In Figure 1, we plot the accuracy curves of the randomized response mechanism with varying values of ρ applied to the majority function as the number of voters increases. Note that as n goes to ∞, the accuracy asymptotically approaches to $\frac{1}{2} + \frac{1}{\pi}\arcsin(\rho)$ as implied by Equation (3).

Algorithm 1 Proposed dynamic programming algorithm with memoization.

```
**Result:** Accuracy
**Initialization**;
Define Dictionary:$\{(\rho, n, s, \theta) = T_\rho f_{\theta_0}(x)$ for some $x$ s.t $\text{sum}(x) = s\}$
Def $T_\rho f_{\theta_0}(x)$ :
$s = \text{sum}(x)$;
**if** $(\rho, n, s, \theta_0)$ *is in dictionary* **then**
    **return** dictionary $[(\rho, n, s, \theta_0)]$;
**else**
    Using 2 recursive calls in summands, compute:
        $\alpha = \frac{1+\rho}{2} T_\rho f_{\theta_0 - x_n}(x_{-n}) + \frac{1-\rho}{2} T_\rho f_{\theta_0 + x_n}(x_{-n})$
    Save $(\rho, n, s, \theta_0) = \alpha$ to dictionary
**end**
Def $\text{Acc}_\rho(f_\theta)$ :
total $= 0$
**for** $i \leftarrow 1$ **to** $n+1$ **do**
        total $+= \binom{n}{i} \cdot f_{\theta_0}(x) \cdot T_\rho f_\theta(x)$ for some $x$ s.t. $x$ has $i$ different $+1$ bits
**end**
**return** $\frac{1}{2} + \frac{\text{total}/2^n}{2}$
```

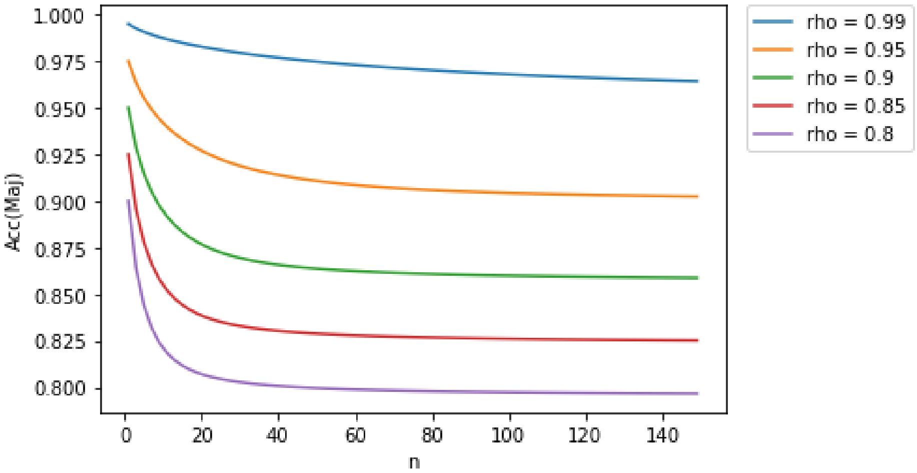

Figure 1 The accuracy curves of the randomized response mechanism with varying values of ρ applied to the majority function as the number of voters increases.

6.2 Dictatorship

Let $f : \{-1,1\}^n \to \{-1,1\}$ be the dictatorship of voter-i, that is $f(x) = 1$ if and only if $x_i = 1$.

Then, for any given $x \in \{-1,1\}^n$,

$$\mathbb{P}[M_\rho f(x) = f(x)] = \mathbb{P}_{y\sim N_\rho(x)}[f(y) = f(x)] = \mathbb{P}_{y_i\sim N_\rho(x_i)}[y_i = x_i] = \frac{1+\rho}{2}.$$

Hence, the average accuracy is also equal to $\frac{1+\rho}{2}$.

6.3 AND_n and OR_n

We will first make the calculations for AND_n and the results will be analogous due to symmetry. We will make use of Fact 2 in the analysis.

First, we start with a generic calculation that holds for any social choice function f. In the calculations in this section, our probability space is $x \sim \{-1,1\}^n, M_\rho f(x) \sim f(y)$ where $y \sim N_\rho x$.

Note that by Fact 2,

$$\mathbb{P}_{x,M_\rho}[M_\rho f(x) = 1] = \mathbb{P}_x[f(x) = 1].$$

$$\mathbb{P}\left[M_\rho f(x) = f(x)\right] = \mathbb{P}\left[M_\rho f(x) = 1 \wedge f(x) = 1\right] + \mathbb{P}\left[M_\rho f(x) = -1 \wedge f(x) = -1\right]$$

and

$$\begin{aligned}\mathbb{P}\left[M_\rho f(x) = -1 \wedge f(x) = -1\right] &= 1 - \mathbb{P}\left[M_\rho f(x) = 1 \vee f(x) = 1\right]\\ &= 1 - \mathbb{P}\left[M_\rho f(x) = 1\right] - \mathbb{P}\left[f(x) = 1\right] + \mathbb{P}\left[M_\rho f(x) = 1 \wedge f(x) = 1\right]\\ &= 1 - 2\cdot\mathbb{P}\left[f(x) = 1\right] + \mathbb{P}\left[M_\rho f(x) = 1 \wedge f(x) = 1\right].\end{aligned}$$

Thus for any social choice function f,

$$\mathbb{P}\left[M_\rho f(x) = f(x)\right] = 1 - 2\cdot\mathbb{P}\left[f(x) = 1\right] + 2\cdot\mathbb{P}\left[M_\rho f(x) = 1 \wedge f(x) = 1\right]$$

For $f = AND_n$,

$$\mathbb{P}[f(x) = 1] = \prod_{i\in[n]} \mathbb{P}[x_i = 1] = 2^{-n},$$

and

$$\mathbb{P}\left[M_\rho f(x) = 1 \wedge f(x) = 1\right] = \mathbb{P}[f(x) = 1]\cdot\mathbb{P}[M_\rho f(x) = 1 | f(x) = 1] = 2^{-n}\cdot(\frac{1+\rho}{2})^{-n}.$$

Hence, the accuracy of M_ρ for AND_n function is equal to $1 - 2^{-n+1}(1 - (\frac{1+\rho}{2})^n)$, whose limit goes to 1 as n goes to ∞. Due to symmetry, accuracy analysis is the same for OR_n function.

7 Conclusion

The main objective in this work is to study the privacy-welfare trade-off and the relation between privacy and probabilistic influence. The proposed definition of welfare happens to hold for any mechanism while on the other hand, the defined probabilistic influence is only specific to the randomized response mechanism. In fact, a more general definition of influence could be coined and a similar property could potentially be observed. We leave out this potential generalization of influence to future work. In terms of welfare, the analysis done in this paper can be replicated in a similar style to other popular privatization schemes such as the Laplace and exponential mechanisms. The privacy-accuracy trade-off of the current mechanism for the majority function may also be further improved. Note that Dictatorship, AND, and OR functions satisfy the equality condition in Proposition 5 as discussed in Remark 6. Thus, the accuracy-privacy analyses for these functions are tight. On the other hand, for a given ρ, the asymptotic accuracy of majority is tight whereas the privacy result is a possibly loose upper bound.

Also, our definitions of influence and welfare assume that the votes are unbiased, that is, they consider everybody to be equally likely to vote for -1 or $+1$. In fact, these definitions can be further generalized to cover the same concept, but for the case of biased voting. For example, one can extend the definitions to be *p-biased* for a given $p \in [-1, 1]$, that is the expected value of each vote is p instead of 0. p-biased distribution is also well-studied in the field of Analysis of Boolean functions.

Finally, our voting model in this paper is a classical referendum model with two candidates. However, in most real-world applications, we generally have multiple candidates and we have to aggregate the rankings. If there is a Condorcet winner in a voting system, then the results regarding two-candidate elections can be directly applied in the multiple-candidate setting. Yet, in many cases, there is no Condorcet winner. Restricting the number of candidates to two has the primary advantage that both the definitions and analyses of welfare and influence naturally follow. We believe that extending the definitions and the tools developed in this paper to multiple-candidate settings would be interesting.

In a broader perspective, we study the effect of using privacy inducing randomized responses in the voting process. We construct a relation between the level of privacy and the resulting level of influence of voters involved in the voting system and the welfare of the chosen social choice function. An insightful takeaway that we can deduce from the derived relationships in this paper is that the ordering of voters' influences and the ordering of welfare amongst the considered social choice functions remain unchanged upon introducing noise via the celebrated randomized response mechanism. Existing works have extensively studied the relationship between privacy and the resulting accuracy in preserving the output of the query that was privatized. At a high level we are the first to shed light on the relationship between privacy and other important phenomena of influence and welfare. We hope that this bridge we have proposed between the two important fields of differential privacy and social choice theory will be further studied and extended as part of future works.

References

1 Kenneth J Arrow. A difficulty in the concept of social welfare. *Journal of political economy*, 58(4):328–346, 1950.

2 Kenneth J Arrow. *Social choice and individual values.* Yale university press, 1951.

3 Michael Ben-Or and Nathan Linial. Collective coin flipping, robust voting schemes and minima of banzhaf values. *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science*, pages 408–416, 1985.

4 Rachel Cummings, Varun Gupta, Dhamma Kimpara, and Jamie Morgenstern. On the compatibility of privacy and fairness. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, pages 309–315, 2019.

5 Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 371–380, 2009.

6 Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Shai Halevi and Tal Rabin, editors, *Theory of Cryptography*, pages 265–284, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

7 Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

8 Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*, pages 1054–1067, 2014.

9 Mark B. Garman and Morton I. Kamien. The paradox of voting: Probability calculations. *Behavioral Science*, 13(4):306–316, 1968.

10 Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

11 Georges-Théodule Guilbaud. Les théories de l'intérêt général et le problème logique de l'agrégation. *Revue économique*, 63(4):659–720, 2012.

12 Michael Hay, Liudmila Elagina, and Gerome Miklau. Differentially private rank aggregation. In *Proceedings of the 2017 SIAM International Conference on Data Mining (SDM)*, pages 669–677, 2017.

13 Claude Hillinger. The case for utilitarian voting. *Homo Oeconomicus*, 22(3):295–321, 2005.

14 J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *[Proceedings 1988] 29th Annual Symposium on Foundations of Computer Science*, pages 68–80, 1988.

15 David Timothy Lee. Efficient, private, and eps-strategyproof elicitation of tournament voting rules. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

16 Debmalya Mandal, Ariel D Procaccia, Nisarg Shah, and David Woodruff. Efficient and thrifty voting by any means necessary. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

17 Debmalya Mandal, Nisarg Shah, and David P. Woodruff. Optimal communication-distortion tradeoff in voting. In *Proceedings of the 21st ACM Conference on Economics and Computation*, EC '20, pages 795–813, New York, NY, USA, 2020. Association for Computing Machinery.

18 Naurang S Mangat. An improved randomized response strategy. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(1):93–95, 1994.

19 Kenneth O. May. A set of independent necessary and sufficient conditions for simple majority decisions. *Econometrica*, 20(4):680–684, 1952.

20 Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, 171(1):295–341, 2010.

21 Ryan O'Donnell. Hardness amplification within np. *Journal of Computer and System Sciences*, 69(1):68–94, 2004. Special Issue on Computational Complexity 2002.

22 Ryan O'Donnell. *Analysis of boolean functions.* Cambridge University Press, 2014.

23 Lionel Penrose. The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, pages 109(1):53–57, 1946.

24 Jean-Jacques Rousseau. Du contrat social. *Marc-Michel Rey*, 1762.

25 Melissa Schwartzberg. Voting the general will: Rousseau on decision rules. *Political Theory*, 36(3):403–423, 2008. URL: `http://www.jstor.org/stable/20452639`.

26 Shang Shang, Tiance Wang, Paul Cuff, and Sanjeev Kulkarni. The application of differential privacy for rank aggregation: Privacy and accuracy, 2014. `arXiv:1409.6831`.

27 Robert Titsworth. Correlation properties of cyclic sequences. *PhD thesis, CalTech*, 1962.

28 Stanley L Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

A Proofs

A.1 Proof of Proposition 5

Proof. Let r be any element in the range of $M_\rho f$. Let $Z = \{z \in \{-1,1\}^n | f(z) = r\}$. Let x and x' differ only at x_i for some $i \in [n]$.

$$\frac{\mathbb{P}[M_\rho f(x) = r]}{\mathbb{P}[M_\rho f(x') = r]} = \frac{\sum_{z\in Z} \mathbb{P}_{y\sim N_\rho x}[y = z]}{\sum_{z\in Z} \mathbb{P}_{y\sim N_\rho x'}[y = z]} = \frac{\sum_{z\in Z} \prod_{j\in[n]} \mathbb{P}_{y_j\sim N_\rho x_j}[y_j = z_j]}{\sum_{z\in Z} \prod_{j\in[n]} \mathbb{P}_{y_j\sim N_\rho x'_j}[y_j = z_j]}.$$

The first equality is upon considering all cases of output of the randomized response resulting in a $z \in Z$. Then by definition that would result in the function f evaluated on this output z to be r. The second equality is due to the independence assumption across the voters choices. Now, for any $z \in Z$,

$$\mathbb{P}_{y_j\sim N_\rho x_j}[y_j = z_j] = \begin{cases} \frac{1+\rho}{2} & \text{if } x_j = z_j \\ \frac{1-\rho}{2} & \text{if } x_j \neq z_j \end{cases} \quad \text{and} \quad \mathbb{P}_{y_j\sim N_\rho x'_j}[y_j = z_j] = \begin{cases} \frac{1+\rho}{2} & \text{if } x'_j = z_j \\ \frac{1-\rho}{2} & \text{if } x'_j \neq z_j \end{cases}$$

This is because $\frac{1-\rho}{2}$ is the probability of a misrecorded vote and $1 - \frac{1-\rho}{2} = \frac{1+\rho}{2}$ is the probability otherwise. More explicitly, with probability $1-\rho$, it chooses to blur the ballot and the blurring is then done by picking uniformly out of the two options of $\{-1, 1\}$ with probability 0.5 each, out of which one pick would result in no change to the vote and the other would result in a misrecorded vote. Also, for any $j \neq i$,

$$\mathbb{P}_{y_j\sim N_\rho x_j}[y_j = z_j] = \mathbb{P}_{y_j\sim N_\rho x'_j}[y_j = z_j].$$

Thus,

$$\frac{1-\rho}{1+\rho} \leq \frac{\sum_{z\in Z} \prod_{j\in[n]} \mathbb{P}_{y_j\sim N_\rho x_j}[y_j = z_j]}{\sum_{z\in Z} \prod_{j\in[n]} \mathbb{P}_{y_j\sim N_\rho x'_j}[y_j = z_j]} \leq \frac{1+\rho}{1-\rho},$$

which completes the proof. ◀

A.2 Proof of Theorem 9

Proof. Using conditional probability, we get that

$$\begin{aligned} I_i[M_\rho f] &= \mathbb{E}_{x\sim\{-1,1\}^n, \forall j\neq i\ z_j=y_j=x_j, y_i\sim N_\rho(1), z_i\sim N_\rho(-1)}\left[\left(\frac{f(y)-f(z)}{2}\right)^2\right] \\ &= \mathbb{P}_{y_i\sim N_\rho(1), z_i\sim N_\rho(-1)}[y_i = 1, z_i = -1] \cdot \mathbb{E}_{x\sim\{-1,1\}^n}\left[\left(\frac{f(x_{i\to 1}) - f(x_{i\to -1})}{2}\right)^2\right] \\ &+ \mathbb{P}_{y_i\sim N_\rho(1), z_i\sim N_\rho(-1)}[y_i = -1, z_i = 1] \cdot \mathbb{E}_{x\sim\{-1,1\}^n}\left[\left(\frac{f(x_{i\to 1}) - f(x_{i\to -1})}{2}\right)^2\right] \end{aligned}$$

Noting that

$$\mathbb{P}_{y_i\sim N_\rho(1), z_i\sim N_\rho(-1)}\left[y_i = 1, z_i = -1\right] = \left(\frac{1+\rho}{2}\right)^2,$$

$$\mathbb{P}_{y_i\sim N_\rho(1), z_i\sim N_\rho(-1)}\left[y_i = -1, z_i = 1\right] = \left(\frac{1-\rho}{2}\right)^2,$$

and that

$$\mathbb{E}_{x\sim\{-1,1\}^n}\left[\left(\frac{f(x_{i\to 1})-f(x_{i\to -1})}{2}\right)^2\right]=I_i[f],$$

we get that

$$I_i[M_\rho f]=\frac{1+\rho^2}{2}I_i[f].$$ ◀

A.3 Proof of Proposition 11

Proof. First, let us fix x. Note that

$$w_x(f)=f(x)\cdot\sum_{i\in[n]}x_i.$$

Since $f(x)\in\{-1,1\}$, $f(x)\cdot\sum_{i\in[n]}x_i$ is maximized when $f(x)=sign(\sum_{i\in[n]}x_i)$. Hence, $W(f)$ is maximized if $\forall x\in\{-1,1\}^n$, $f(x)=sign(\sum_{i\in[n]}x_i)$, which is exactly the definition of the majority function. ◀

▶ Remark 18. Note that we used the condition that n is odd to ensure that *sign* function is well-defined. If n was even, then the maximizers of $W(f)$ are again the majority functions where it does not matter who is elected if it is tied.

A.4 Proof of Proposition 12

In the proof of this result, we use discrete Fourier analysis. It is a well-known result from the field of analysis of Boolean functions, that every function $f:\{-1,1\}^n\to\mathbb{R}$ can be uniquely expressed as a multilinear polynomial,

$$f(x)=\sum_{S\subseteq[n]}\widehat{f}(S)\chi_S(x)$$

where for any $S\in[n]$

$$\chi_S(x)=\prod_{i\in S}x_i.$$

This expression is called the Fourier expansion of f, and the real number $\widehat{f}(S)$ is called the Fourier coefficient of f on S. Collectively, the coefficients are called the Fourier spectrum of f. The following is an essential result from discrete Fourier Analysis.

▶ **Lemma 19** (Plancherel's Theorem). *For any functions $f,g:\{-1,1\}^n\to\mathbb{R}$,*

$$E_{x\sim\{-1,1\}^n}[f(x)g(x)]=\sum_{S\subseteq[n]}\widehat{f}(S)\widehat{g}(S).$$

It is possible to neatly calculate many features of f including the influences in terms of Fourier coefficients.

▶ **Lemma 20** (Proposition 2.21, [22]). *Let $f:\{-1,1\}^n\to\{-1,1\}$ be a monotone function and let the Fourier spectrum of f be $f(x)=\sum_{S\subseteq[n]}\widehat{f}(S)\chi_S(x)$. Then, for any $i\in[n]$,*

$$I_i[f]=\widehat{f}(\{i\}).$$

It is also possible to calculate the welfare in terms of the Fourier coefficients by taking one step further from the proof of Proposition 11.

▶ **Lemma 21.** *Let f be any social choice function $f : \{-1,1\}^n \to \{-1,1\}$. Then, $W(f) = \sum_{i\in[n]} \widehat{f}(\{i\})$.*

Proof. By the definition of welfare,

$$W(f) = \mathbb{E}_x[w_x(f)] = \mathbb{E}_x[f(x) \cdot \sum_{i\in[n]} x_i] = \sum_{i\in[n]} \widehat{f}(\{i\})$$

where the last equation follows from Lemma 19. ◀

We are ready to finish the proof.

Proof of Proposition 12. The proof follows immediately from Lemma 20 and Lemma 21. ◀

A.5 Proof of Theorem 14

Proof. We prove this identity by using a double-counting method and linearity of expectation. Fix f. For any $i \in [n]$, let $1_{i,x,\rho}$ be the indicator random variable defined as follows:

$$1_{i,x,\rho} = \begin{cases} 1 & \text{if } M_\rho f(x) = x_i \\ -1 & \text{if } M_\rho f(x) \neq x_i \end{cases}$$

where the randomization is due to the randomized response. Note then when x is given and $\rho = 1$, there is no randomization because $M_\rho f(x) = f(x)$ with probability 1. Therefore, $1_{i,x,1}$ is a deterministic function. For the sake of simplicity, we will abuse the notation and write $1_{i,x}$ instead of $1_{i,x,1}$ in the deterministic case. Then,

$$w_x(M_\rho f) = \sum_{i\in[n]} 1_{i,x,\rho} \quad \text{and} \quad w_x(f) = \sum_{i\in[n]} 1_{i,x}.$$

Thus,

$$W(M_\rho f) = \mathbb{E}_{M_\rho,x}[w_x(f)] = \mathbb{E}_{x,M_\rho}[\sum_{i\in[n]} 1_{i,x,\rho}] = \sum_{i\in[n]} \mathbb{E}_{x,M_\rho}[1_{i,x,\rho}]$$

and so

$$W(f) = \sum_{i\in[n]} \mathbb{E}_x[1_{i,x}].$$

Now, we will show that for any $i \in [n]$,

$$\mathbb{E}_{x,M_\rho}[1_{i,x,\rho}] = \rho \cdot \mathbb{E}_x[1_{i,x}].$$

First, note that

$$\mathbb{E}_{x,M_\rho}[1_{i,x,\rho}] = \mathbb{P}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho x}}[f(y) = x_i] - \mathbb{P}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho x}}[f(y) \neq x_i].$$

By using

$$\mathbb{P}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho x}}[f(y) = x_i] + \mathbb{P}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho x}}[f(y) \neq x_i] = 1,$$

we get that

$$\mathbb{E}_{x,M_\rho}[1_{i,x,\rho}] = 2 \cdot \mathbb{P}_{\substack{x\sim\{-1,1\}^n \\ y\sim N_\rho x}}[f(y) = x_i] - 1.$$

By Fact 2, we can replace $x \sim \{-1,1\}^n, y \sim N_\rho x$ with $y \sim \{-1,1\}^n, x \sim N_\rho y$. Thus, by using conditional probability,

$$\begin{aligned}
\mathbb{E}_{x,M_\rho}[1_{i,x,\rho}] &= 2 \cdot \mathbb{P}_{\substack{y\sim\{-1,1\}^n \\ x\sim N_\rho y}}[f(y) = x_i] - 1 \\
&= 2(\mathbb{P}_{x\sim N_\rho y}[x_i = y_i] \cdot \mathbb{P}_{y\sim\{-1,1\}^n}[f(y) = y_i] \\
&\quad + \mathbb{P}_{x\sim N_\rho y}[x_i = -y_i] \cdot \mathbb{P}_{y\sim\{-1,1\}^n}[f(y) = -y_i]) - 1 \\
&= (1+\rho) \cdot \mathbb{P}_{y\sim\{-1,1\}^n}[f(y) = y_i] + (1-\rho) \cdot \mathbb{P}_{y\sim\{-1,1\}^n}[f(y) \neq y_i] - 1 \\
&= \rho \cdot (\mathbb{P}_{y\sim\{-1,1\}^n}[f(y) = y_i] - \mathbb{P}_{y\sim\{-1,1\}^n}[f(y) \neq y_i]) \\
&= \rho \cdot \mathbb{E}_x[1_{i,x}]
\end{aligned}$$

which completes the proof. ◀

B Social Choice Functions

In this paper, we exclusively focus on social choice functions with two alternatives. There are many ways to interpret these functions. It can be considered as a two-candidate election or as a referendum in the context of political science. It can also be interpreted as a classifier in the context of Machine Learning. In this paper, we will generally give the interpretations in the context of two-candidate elections.

In general, we work with the Boolean functions defined as $f : \{-1,1\}^n \to \mathbb{R}$, and we denote the bit i of the input x by x_i for any $i \in [n]$. However, we define welfare only for *social choice functions*, that is the Boolean functions whose ranges are $\{-1,1\}$. We analyze accuracy only for the following specific social choice functions.

- **Majority:** Suppose that n is an odd number. The majority function of n agents/voters is denoted by Maj_n and defined as

  $$f(x) = sign(\sum_{i\in[n]} x_i)$$

  for any $x \in \{-1,1\}^n$ where $sign : \mathbb{R} \to \{-1,0,1\}$ is the function such that

  $$sign(a) = \frac{a}{|a|}$$

  for any $a \in \mathbb{R}, a \neq 0$ and sign(0)=0.
- **Dictatorship:** For a given number n and $i \in [n]$, the dictatorship of voter-i is defined as

  $$f(x) = x_i$$

  for any $x \in \{-1,1\}^n$.
- **AND$_n$:** The AND_n function outputs 1 if there is unanimity on 1, outputs -1 otherwise. Namely,

  $$f(x) = \begin{cases} 1 & \text{if } \forall i \in [n], x_i = 1 \\ -1 & \text{otherwise} \end{cases}$$

- **OR$_{\boldsymbol{n}}$:** The OR_n function outputs 1 if at least one voter votes for 1, and outputs -1 otherwise. In other words, it outputs -1 if there is unanimity on -1, outputs 1 otherwise. Namely,

$$f(x) = \begin{cases} -1 & \text{if } \forall i \in [n], x_i = -1 \\ 1 & \text{otherwise} \end{cases}$$

Note that, in this paper, we assume *the impartial culture assumption*, that is the voters are not affected by each other and they vote independently uniform at random between two candidates.