Relations between Privacy, Verifiability, Accountability and Coercion-Resistance in Voting Protocols

Alisa Pankova[1] and Jan Willemson[1][0000−0002−6290−2099]
{alisa.pankova,jan.willemson}@cyber.ee

Cybernetica AS, Narva mnt 20, 51009 Tartu, Estonia

Abstract. This paper studies quantitative relationships between privacy, verifiability, accountability, and coercion-resistance of voting protocols. We adapt existing definitions to make them better comparable with each other and determine which bounds a certain requirement on one property poses on some other property. It turns out that, in terms of proposed definitions, verifiability and accountability do not necessarily put constraints on privacy and coercion-resistance. However, the relations between these notions become more interesting in the context of particular attacks. Depending on the assumptions and the attacker's goal, voter coercion may benefit from a too weak as well as too strong verifiability.

Keywords: Security and privacy metrics, privacy, anonymity, verifiability, voting

1 Introduction

Voting is a complex process subject to a number of requirements such as eligibility, generality, uniformity, freedom of choice, tally integrity, accessibility, etc [4, 10, 19, 23]. In order to implement these requirements, a number of measures can be applied. For example, in order to express one's preference freely and withstand coercion, voting privately is often required. Tally integrity, on the other hand, can be achieved via various verification procedures.

Even though both privacy and verifiability of voting are well-motivated, they are at least partially contradictory. Intuitively, when targeting full public verifiability without any trust assumptions, it seems necessary to also open all the personalised votes, but this causes privacy loss and potential coercion issues. Of course, this intuition is very informal and the situation becomes more complicated when we consider particular definitions for privacy and verifiability.

In order to study the connections between the two notions, the corresponding definitions must be given in comparable terms. However, it is far from being clear which terms are the best suited for this comparison. Working towards definitions that can be quantitatively compared to each other, and coming up with some comparison results, are the main aims of the current paper.

2 Related Work

There are many definitions of privacy in the context of voting, and an extensive survey discussing their advantages and drawbacks can be found in [3]. Relations between privacy and coercion-resistance for certain formal definitions of these notions have been shown in [9]. In this work, we are using definitions of privacy and coercion-resistance that originate from [16]. The benefit of these definitions is that they allow to measure the corresponding properties quantitatively. We instantiate our definitions of verifiability and accountability in the KTV framework [7, 15, 17]. This framework provides generic definitions for verifiability and accountability, and many other, more specific definitions of verifiability can be instantiated in this framework. Among other results, [16] shows the relation between privacy and coercion resistance, and [15] shows the relation between verifiability and accountability. In all these works, the agents (i.e. the voters and the authorities) of a voting protocol are modeled as some processes, typically specified in pi-calculus.

KTV framework relies on the notion of end-to-end (E2E, global) verifiability, where voters and external observers are able to check whether the final result corresponds to the actual choices of honest voters. An alternative is to consider *universal* and *individual* verifiability as separate properties [22]. Previous research has established the following:

- There can be no unconditional privacy if there *is* universal verifiability [5].
- There can be no privacy if there *is no* individual verifiability [8].

In addition, [5] proves that universal verifiability and receipt-freeness cannot be achieved simultaneously unless private channels are available. A receipt is a witness which allows verifying in an unambiguous way the vote of a certain voter. Intuitively, the existence of a receipt may lead to voter coercion. Different types of realistic coercion methods, both legal and illegal, are discussed in [11].

It has been noted in [14] that universal and individual verifiability are not *sufficient* for E2E verifiability [12]. Indeed, by definition, universal verifiability only checks that the final result corresponds to the submitted votes, but it does not require that the votes are well-formed (e.g. that there are no negative votes). Also by [14], universal verifiability is not *necessary* for E2E verifiability.

In [8], it is shown how manipulation of even one vote may break privacy by observing the change that it caused in the tally. It is important that the attacker knows whose vote he is trying to change, so privacy requires *individual* verifiability. The proposed attack breaks a particular privacy definition, which says that the attacker should not be able to distinguish two protocol transcripts where some honest voters Alice and Bob have decided to swap their votes. The attacker may drop the vote of Alice in both transcripts, and observe the difference in the tally of the two transcripts to determine what the vote of Alice actually was. Such a privacy definition is very strong, and in practice, the attacker does not actually have access to two alternative voting transcripts. If there are many voters, dropping a single vote does not help much in actually guessing some other votes. Nevertheless, if the attacker has a strong prior knowledge of the other voters' choices, such an attack may allow learning the vote of the victim. In this work, we consider similar attacks w.r.t. the privacy definition of [16], which allows assessing the severity of the attack quantitatively.

An interesting approach to estimate voting systems in terms of *distributional differential privacy* has been proposed in [18]. While differential privacy is often achieved by adding noise to the system, which is unacceptable for voting, DDP is achieved by considering the distribution of votes as a source of randomness.

3 Preliminaries

3.1 Protocols

In this section, we present a generic framework for the definitions considered in this paper. The framework originates from [7, 15, 16] and is provided with some simplifications, excluding details that are not relevant for this paper.

First of all, we need the notion of a process that can perform internal computation and can communicate with other processes by sending messages via (external) input/output channels.

Definition 1 (Process). *A process is a set of probabilistic polynomial-time interactive Turing machines (also named programs) that are connected via named tapes (also called channels). We denote by $\mathsf{\Pi}(I, O)$ the set of all processes with external input channels I and external output channels O. A process defines a family of probabilistic distributions over runs, indexed by the security parameter η. The concurrent composition of processes π and π' is denoted by $\pi \| \pi'$.*

A protocol is not a process by itself, but rather a collection of building blocks that will be used to define a process. As noted in [16], since the quantitative level of privacy, coercion-resistance, and verifiability of a voting protocol depends on several parameters such as the number of voters and the number of choices, we consider a protocol *instantiation* for which these parameters are fixed.

Definition 2 (Protocol instantiation). *A* protocol instantiation *is a tuple $P = (\Sigma, \mathsf{Ch}, \mathsf{In}, \mathsf{Out}, \{\mathsf{\Pi}_a\}_{a\in\Sigma})$ where*

- *Σ is a set of protocol* agents.
- *Ch is a set of protocol* channels.
- *In and Out are functions from Σ to 2^{Ch} (i.e assignments of input and output channels for each protocol agent) such that $\mathsf{In}(a) \cap \mathsf{In}(b) = \emptyset$ and $\mathsf{Out}(a) \cap \mathsf{Out}(b) = \emptyset$ for all $a, b \in \Sigma$, $a \neq b$.*
- *$\mathsf{\Pi}_a \subseteq \mathsf{\Pi}(\mathsf{In}(a), \mathsf{Out}(a))$ for $a \in \Sigma$ is the set of honest programs that can be run by the agent a.*

The randomness of agent behaviour, such as probabilistic distribution of choices of an honest voter, is covered by $\mathsf{\Pi}_a$. Particular probability distributions are not relevant for the results of this paper.

A protocol *instance* is the process that will actually be executed.

Definition 3 (Protocol instance, run). *Let* $P = (\Sigma, \mathsf{Ch}, \mathsf{In}, \mathsf{Out}, \{\Pi_a\}_{a\in\Sigma})$ *be a protocol instantiation.*

- *An* instance *of* P *is a process* $\pi_P = \pi_{a_1} \| \dots \| \pi_{a_{|\Sigma|}}$ *for* $\pi_{a_i} \in \Pi_{a_i}$*,*
- *A* run *of* P *is a run of some instance of* P*.*

Similarly to [7, 16], we have not included processes of dishonest parties into the definitions of P and π_P. Instead, the dishonest parties are subsumed by a special *adversary* process.

Definition 4 (Adversary). *A protocol instance* π_P *is typically run in parallel with an adversary process* π_A *as a process* $\pi := \pi_P \| \pi_A$*.*

There is a bidirectional channel between the adversary A *and each protocol agent* $a \in \Sigma$*. The adversary can corrupt an agent* $a \in \Sigma$ *by sending a special message* corrupt*. Upon receiving such a message,* a *reveals its internal state to* A *and from then on is controlled by* A*, i.e. runs a dummy process* dum *which simply forwards all messages between* A *and the interface of* a *in* π_P*. Some agents (honest users and incorruptible authorities) ignore* corrupt *messages. Public information (such as the election result) is output to* A *even without corruption.*

At the end of a run, π_A *produces some output* y*. We use the notation* $\pi \overset{A}{\mapsto} y$ *to say that the output of* π_A *in a run of* π *is* y*.*

We say that an agent $a \in \Sigma$ is *honest* in a run of $\pi := \pi_A \| \pi_P$ if a has not been corrupted in this run, i.e has not accepted the message corrupt. We use notation $\pi \vDash \mathsf{dis}(a)$ to denote an event (viewing π as a probabilistic distribution over runs) that the agent a has been corrupted.

The condition $\mathsf{dis}(a)$ can be viewed as a certain *property* of a protocol P. A property is a function that takes as input a run of a process π and returns a boolean value, telling whether that property is satisfied. For a fixed protocol instantiation P, a property can be viewed as a subset of runs of P.

Definition 5 (Protocol property). *A property* γ *of* P *defines a subset of the set of all runs of* P*. By* $\neg\gamma$ *we denote the complement of* γ*, i.e. the set of runs that do not satisfy* γ*.*

In order to reason about probability distributions of protocol runs taking into account the privacy parameter η, we will need the following definition.

Definition 6 (negligible, overwhelming, δ-bounded [7, 15–17]). *A function* $f : \mathbb{N} \to [0, 1]$ *is* negligible *if, for every* $c > 0$*, there exists* η_0 *such that* $f(\eta) \leq \frac{1}{\eta^c}$ *for all* $\eta > \eta_0$*. The function* f *is* overwhelming *if the function* $1 - f$ *is negligible. A function* f *is* δ*-bounded if, for every* $c > 0$*, there exists* η_0 *such that* $f(\eta) \leq \delta + \frac{1}{\eta^c}$ *for all* $\eta > \eta_0$*.*

The summary of process-related notation used in this paper is given in Table 1.

Table 1. Table of notations. For events, π is viewed as a distribution of runs.

Notation	Type	Meaning
$\pi^{(\eta)}$	process	a process π where all programs use the security parameter η
$\pi_1 \| \pi_2$	process	concurrent composition of processes π_1 and π_2
$\pi(\vec{x})$	process	a process π running with inputs $\vec{x}$
$\pi_{P \setminus \Sigma'}$	process	concurrent composition of all subprocesses of π_P excluding subprocesses π_a of agents $a \in \Sigma' \subseteq \Sigma$.
$\pi_{P \setminus \vec{i}}$	process	same as $\pi_{P \setminus \Sigma'}$ for $\Sigma' = \{v_{i_1}, \ldots, v_{i_k}\}$, where $\vec{i} \subseteq \{1, \ldots, \|V\|\}$
$\pi \mapsto (a : y)$	event	the final output of the agent $a \in \Sigma$ in the run of π is y
$\pi \overset{A}{\mapsto} y$	event	the final output of the adversary π_A in the run of π is y
$\pi \models \gamma$	event	a run of π satisfies a property γ
$\mathsf{dis}(a)$	property	the agent $a \in \Sigma$ has been corrupted
$\mathsf{voted}(i, c)$	property	the voter $v_i \in V$ cast a vote c
$\mathcal{F}_{dis}$	set	the set of boolean formulae over literals $\mathsf{dis}(a)$ for $a \in \Sigma$

3.2 Notation Related to Voting Protocols

We will use V to denote the set of voters, C the set of possible choices made by the voters (a choice does not necessarily represent a single candidate), and R the set of possible election results. Let $V = V_H \cup V_D$ for $V_H \cap V_D = \emptyset$, where V_H are honest voters, and V_D are dishonest voters (controlled by the adversary). Let $|V| = n = n_h + n_d$ be the total number of voters, where $n_h = |V_H|$ and $n_d = |V_D|$. We assume that the voters are somehow ordered, and the voter with index $i \in \{1, \ldots, n\}$ is denoted by v_i. The votes are combined using a result function $\rho : C^n \to R$ whose exact definition depends on the used voting rule.

3.3 Verifiability and Accountability

We start from a generic definition of verifiability from [7]. First of all, we need to state what exactly we are verifying. We assume a certain property γ (Definition 5) that we want to achieve in each protocol run, e.g. that each voter votes at most once, or that all ballots are well-formed. If γ is achieved, then everything is fine. If γ is not achieved, then we at least want to detect such a case.

The definitions of verifiability and accountability used in this paper will be based on the particular γ for quantitative verifiability proposed in [7]. First, let us define the protocol runs covered by γ. The idea of the following definition is that the final tally (i.e. the multiset of ballots before applying ρ) of a voting protocol may differ from the true tally in at most k votes.

Definition 7 (k-correctness of the protocol run [7]). *A protocol run r, where $c_1, \ldots, c_{n_h}$ are the choices of honest voters, is called k-correct if there exist valid choices $c'_1, \ldots, c'_{n_d}$ (representing possible choices of dishonest voters) and $\tilde{c}_1, \ldots, \tilde{c}_n$, such that:*

- *an election result is published in r and it is equal to $\rho(\tilde{c}_1, \ldots, \tilde{c}_n)$;*
- *$d((c_1, \ldots, c_{n_h}, c'_1, \ldots, c'_{n_d}), (\tilde{c}_1, \ldots, \tilde{c}_n)) \leq k$;*

where d is defined as $d(\vec{c}, \vec{c}') = \sum_{c \in C} |f_{count}(\vec{c})[c] - f_{count}(\vec{c}')[c]|$, where C is the set of possible choices, and $f_{count} : C^n \to \mathbb{N}^C$ counts how many times each choice occurs in a vector.

The set of all k-correct runs of a protocol is denoted by γ_k.

In [7], verifiability w.r.t. a property γ is quantified by an upper bound on the probability that:

1. γ is not achieved; and
2. this fact remains undetected by a certain designated party J called the Judge.

The particular definition of γ can be very different, and various choices of γ provide different flavours of verifiability. In this paper, we instantiate the generic verifiability property of [7] on γ_k. This leads to the following definition.

Definition 8 ((k, δ)-verifiability). *Let π_P be an instance of a voting protocol P with the set of agents Σ. Let $\delta \in [0, 1]$ be the tolerance, $J \in \Sigma$ be the Judge, and γ_k be the set of runs of P such that, for all runs $r \in \gamma_k$, r is k-correct according to Definition 7. We say that π_P is (k, δ)-verifiable w.r.t. J if for all adversaries π_A and $\pi = \pi_P \| \pi_A$, the probability*

$$\Pr[(\pi^{(\eta)} \models \neg\gamma_k) \wedge (\pi^{(\eta)} \mapsto (J : accept))]$$

is δ-bounded as a function of η, and

$$\Pr[\pi^{(\eta)} \mapsto (J : reject)] = 0$$

if $\pi \not\models \mathsf{dis}(a)$ for all $a \in \Sigma$.

We do not want that the attacker would be able to abort the elections, so we need to specify what actually happens after the Judge rejects. As proposed in [15], in general verifiability is not enough, and in practice, we want *accountability*. This property assumes that, if the Judge rejects, he needs to come up with a certain *verdict*, which states which parties have potentially misbehaved. A verdict is a boolean formula over statements $\mathsf{dis}(a)$ for $a \in \Sigma$. Let $\mathcal{F}_{dis}$ be the set of all boolean formulae of such a form. It is possible that a verdict has a form of disjunction, e.g. $\mathsf{dis}(v_i) \vee \mathsf{dis}(a)$, for a voter v_i and a voting authority $a \in \Sigma$, which could mean that it is not clear whether a has dropped the message of the voter v_i, or the voter v_i has not sent a valid message. An *accountability constraint* of a protocol P consists of a property α that we want to be satisfied, and a set of possible verdicts $\phi_1, \ldots, \phi_\ell$ the Judge J must come out in the case when α is not satisfied.

Definition 9 (Accountability constraint [15]). *An accountability constraint of a protocol P is a tuple $(\alpha, \phi_1, \ldots, \phi_\ell)$ where α is a property of P (i.e. a subset of runs of P) and $\phi_1, \ldots, \phi_\ell \in \mathcal{F}_{dis}$.*

In this paper, we will be working with the property $\alpha := \gamma_k$ as in Definition 8. This means that we require accountability if the tally error is at least k, and we agree to accept smaller errors in the tally.

Definition 10 (**(k,δ)-accountability**)**.** *Let π_P be an instance of a voting protocol P with the set of agents Σ, and let $J \in \Sigma$ be the Judge. Let $\Phi = (\gamma_k, \phi_1, \ldots, \phi_\ell)$ be an accountability constraint where γ_k is set of runs of P such that, for all runs $r \in \gamma_k$, r is k-correct according to Definition 7.*

We say that π_P is (k,δ)-accountable w.r.t. Φ and J if for all adversaries π_A and $\pi = \pi_P \| \pi_A$, the probability

$$\Pr[(\pi^{(\eta)} \models \neg\gamma_k) \wedge \neg\exists i(\pi^{(\eta)} \mapsto (J : \phi_i))]$$

is δ-bounded as a function of η, and, for all $i \in \{1, \ldots, n\}$,

$$\Pr[\pi^{(\eta)} \mapsto (J : \phi_i)] = 0$$

if $\pi \not\models \phi_i$.

Ideally, we would like to have *individual accountability* where every verdict blames a particular agent. However, as shown in [15], individual accountability is typically not achieved by voting protocols, and in [2] it was shown that resolving a dispute between two agents requires certain assumptions such as undeniable channels or trusted authorities. The problem is the communication between the voter and the voting system, where a voter may always say that "the system does not respond", and the system may always argue that "the voter has not attempted to communicate". In this work, we will consider general accountability.

3.4 Privacy and Coercion-Resistance

We take the definition of voter privacy from [16], defined as the inability to distinguish whether the voter $v \in V$ under observation made the choice $c \in C$ or $c' \in C$. The parameter k quantifies the number of voters under observation.

Definition 11 (**(k,δ)-privacy**)**.** *Let π_P be an instance of a voting protocol P with n voters. Let $\delta \in [0,1]$ be the tolerance. For all $i \in \{1, \ldots, n\}$, let π_{v_i} be the honest process of the voter v_i. Let $\vec{i} = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ be the indices of honest voters under observation, and let $\vec{c}, \vec{c'} \in C^k$ be two assignments of choices to the voters $\vec{i}$. Denote $\pi_{\vec{i},\vec{c}} := \pi_A \| \pi_{v_{i_1}}(c_1) \| \ldots \| \pi_{v_{i_k}}(c_k) \| \pi_{P \setminus \vec{i}}$ for an adversary process π_A. We say that π_P is (k,δ)-private if the difference of probabilities*

$$\left| \Pr[\pi^{(\eta)}_{\vec{i},\vec{c}} \stackrel{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i},\vec{c'}} \stackrel{A}{\mapsto} 1] \right|$$

is δ-bounded as a function of the security parameter η for all $\vec{i}$, $\vec{c}, \vec{c'}$ and for all adversaries π_A.

In contrary to Definition 8, larger k means stronger privacy guarantees, and for $k = 0$, the adversary would need to distinguish two identical distributions.

Proposition 1. *Let $\ell \leq k$. If an instance π_P of a voting protocol P is (k,δ)-private, then it is also (ℓ,δ)-private .*

Intuitively, it can only be easier for the adversary to notice the difference between two distributions if a larger number of voters' votes is fixed in advance. A formal proof of Proposition 1 can be found in App. A.1.

Let us now consider the definition of coercion-resistance from [16]. A protocol is called coercion-resistant if the coerced voter, instead of running the dummy strategy dum (which simply lets all messages be chosen by the coercer), can run some counter-strategy $\pi_{\tilde{v}}$ such that:

1. by running this counter-strategy, the coerced voter achieves their own goal, e.g., votes for a specific candidate; and
2. the coercer is not able to distinguish whether the coerced voter followed coercer's instructions or tried to achieve their own goal (by running $\pi_{\tilde{v}}$).

Similar to the privacy definition, we extend the coercion-resistance of [16] to k voters, where we allow that up to k voters can be coerced simultaneously. Here the coerced voters may share a common goal γ. For example, if the goal of k coerced voters is to give at least $\ell < k$ votes to Alice, then it does not matter who exactly gave a vote to Alice, and only the total multiset of votes in the group matters.

Definition 12 (**(k, δ)-coercion-resistance**). *Let π_P be an instance of a voting protocol P with n voters. Let $\delta \in [0, 1]$ be the tolerance. Let $\vec{i} = \{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$ be the indices of honest voters under observation. Let γ be the joint goal of the voters $\vec{i}$. We say that π_P is (k, δ)-coercion-resistant w.r.t. γ, if the exists a joint strategy $\pi_{\tilde{v}}$ of coerced voters such that the following conditions are satisfied for any adversary π_A connected to $v_{i_1}, \dots, v_{i_k}$ via the interface of* dum*:*

- $\Pr[(\pi_A \| \pi_{\tilde{v}} \| \pi_{P \setminus \vec{i}})^{(\eta)} \models \gamma]$ *is overwhelming as a function of η.*
- $\Pr[(\pi_A \| \mathsf{dum} \| \pi_{P \setminus \vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] - \Pr[(\pi_A \| \pi_{\tilde{v}} \| \pi_{P \setminus \vec{i}})^{(\eta)} \overset{A}{\mapsto} 1]$ *is δ-bounded as a function of η.*

Note that the counter-strategy does not necessarily belong to the set of honest voter processes, and e.g. in order to give k votes to Alice, it is allowed that one of the coerced voters submits a malformed ballot with k votes, while the other $k - 1$ coerced voters abstain from voting.

4 Relations Between Definitions

In this paper, we study relations between the definitions of Sec. 3.3 and Sec. 3.4. A summary of relations considered in this paper is depicted in Figure 1. We note that it does not cover *all* possible relationships between definitions. While Theorem 1 and Theorem 2 are based on related work and merely adapted to our definitions, we still reproduce them for the sake of completeness. On the contrary, Theorems 3, 4, and 5 that establish a bridge between the verifiability-related definitions of Sec. 3.3 and the privacy-related definitions of Sec. 3.4 comprise a new contribution. In this section, we formally state the corresponding theorems and provide proof sketches. The full proofs can be found in App. A.

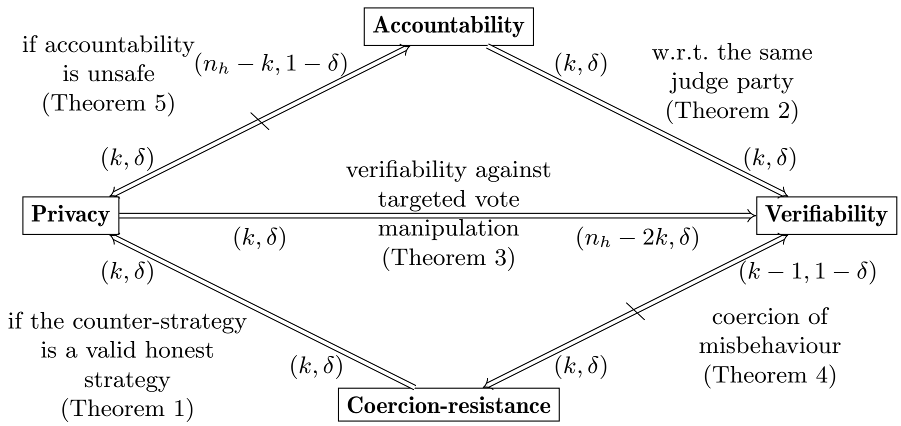

Fig. 1. Summary of the results of this paper (informal, simplified). Here n_h is the total number of honest voters, and k and δ are parameters. The graph depicts relations between these parameters for different properties of a voting protocol. A unidirectional arrow $\Rightarrow$ denotes implication, and a negated bidirectional arrow $\not\Leftrightarrow$ denotes properties that cannot be achieved simultaneously. The arrows can be composed, but one must be careful that the assumptions of corresponding theorems are all taken into account.

4.1 Coercion-Resistance and Privacy

Relationships of coercion-resistance and privacy have been studied in [9, 16]. An interesting outcome of [16] is that, while intuitively coercion-resistance is a stronger notion than privacy, for some protocols it is possible that the level of privacy is *lower* than the level of coercion resistance. The reason is that the counter-strategy of a voter in Definition 12 does not necessarily belong to the set of valid strategies of honest voters, and may protect the vote in a better way than following the protocol honestly. However, coercion-resistance is nevertheless stronger than privacy if we assume that the counter-strategy does not *outperform* an honest strategy, defined as follows.

Definition 13 (non-outperforming counter-strategy [16]). *Let π_P be an instance of a voting protocol P. Let $\vec{i} = \{i_1, \ldots, i_k\}$ be the indices of honest voters under observation. Let $\pi_A^{\vec{c'}}$ be a process that is only connected to the agents $v_{i_1}, \ldots, v_{i_k}$ using the interface of* dum*, and acts on their behalf according to an honest strategy $\pi_v(\vec{c'}) := \pi_{v_{i_1}}(c'_1)\|\ldots\|\pi_{v_{i_k}}(c'_k)$. Let $\pi_v(\vec{c}) := \pi_{v_1}(c_1)\|\ldots\|\pi_{v_k}(c_k)$. Let $\pi_{\tilde{v}}(\vec{c})$ be a joint counter-strategy of the honest voters $\vec{i}$ whose goal is to make choices $\vec{c} = \{c_1, \ldots, c_k\}$. We say that the counter-strategy $\pi_{\tilde{v}}$ does not outperform the honest voting strategy of π_P if, for any adversary process π_A that is not connected to $\pi_A^{\vec{c'}}$, and any choices $\vec{c}$ and $\vec{c'}$,*

$$\Pr[(\pi_A\|\pi_A^{\vec{c'}}\|\pi_{\tilde{v}}(\vec{c})\|\pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] - \Pr[(\pi_A\|\pi_v(\vec{c})\|\pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1]$$

is negligible as a function in the security parameter η.

We adapt a theorem of [16] to our definitions.

Theorem 1. *Let an instance π_P of a voting protocol P be (k,δ)-coercion-resistant. Assume that, for any subset of k coerced voters, the coercion counter-strategy $\pi_{\tilde{v}}$ does not outperform the honest voting strategy of π_P (Definition 13). Then, π_P is (k,δ)-private.*

Proof (Sketch). Suppose that π_P is not (k,δ)-private. There exist k voters $\vec{i}$, choices $\vec{c}$ and $\vec{c}'$, and an adversary process π_A such that

$$\left|\Pr[\pi^{(\eta)}_{\vec{i},\vec{c}} \overset{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i},\vec{c}'} \overset{A}{\mapsto} 1]\right|$$

is not δ-bounded as a function of η, where $\pi_{\vec{i},\vec{c}}$ is defined as in Definition 11.

Let us now describe a coercer that breaks coercion-resistance. Consider a particular setting where the true goals of the voters $\vec{i}$ is to make the choice $\vec{c}$. Let $\pi_A^{\vec{c}'}$ be a coercer that selects for the voters the input $\vec{c}'$, and otherwise acts as an honest voter would. By construction of $\pi_A^{\vec{c}'}$,

$$\Pr[(\pi_A \| \pi_A^{\vec{c}'} \| \mathsf{dum} \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] = \Pr[\pi^{(\eta)}_{\vec{i},\vec{c}'} \overset{A}{\mapsto} 1] \ .$$

Let $\pi_v = \pi_{v_{i_1}} \| \ldots \| \pi_{v_{i_k}}$. By definition of $\pi_{\vec{i},\vec{c}}$,

$$\Pr[(\pi_A \| \pi_v(\vec{c}) \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] = \Pr[\pi^{(\eta)}_{\vec{i},\vec{c}} \overset{A}{\mapsto} 1] \ .$$

Since $\pi_{\tilde{v}}$ does not outperform $\pi_v = \pi_{v_{i_1}} \| \ldots \| \pi_{v_{i_k}}$, and there are no direct connections between π_A and $\pi_A^{\vec{c}}$,

$$\Pr[(\pi_A \| \pi_v(\vec{c}) \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] - \Pr[(\pi_A \| \pi_A^{\vec{c}'} \| \pi_{\tilde{v}}(\vec{c}) \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1]$$

is negligible as a function of η.We get that

$$\Pr[(\pi_A \| \pi_A^{\vec{c}'} \| \mathsf{dum} \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] - \Pr[(\pi_A \| \pi_A^{\vec{c}'} \| \pi_{\tilde{v}}(\vec{c}) \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A'}{\mapsto} 1]$$

is not δ-bounded as a function of η. Let $\pi_{A'} := \pi_A \| \pi_A^{\vec{c}'}$ be an adversary that outputs the final output of π_A. Such $\pi_{A'}$ breaks (k,δ)-coercion-resistance. Since π_A does not interact with $\vec{i}$ (as they are honest), and $\pi_{A'}^{\vec{c}'}$ interacts only with $\vec{i}$ using interface of dum, $\pi_{A'}$ satisfies Definition 12. $\square$

4.2 Accountability and Verifiability

It has been proven in [15] that verifiability can be treated as a special case of accountability. We adapt a theorem of [15] to our definitions.

Theorem 2. *Let an instance π_P of a voting protocol P be (k,δ)-accountable w.r.t. a Jugde J and a property $\Phi = (\gamma_k, \phi_1, \ldots, \phi_\ell)$ where $\forall i: \ \phi_i \in \mathcal{F}_{dis}$. Then, π_P is (k,δ)-verifiable w.r.t. a Judge J' who rejects iff J outputs a verdict ϕ_i.*

Proof (Sketch). Let $\pi := \pi_A \| \pi_P$. Suppose that π_P is not (k, δ)-verifiable w.r.t. J. The verifiability may fail due to one of the following reasons:

1. There is a run where J' outputs *reject*, but all parties are honest. Then, there is a run where J outputs a verdict ϕ_i while all parties are honest. This violates accountability requirement that $\Pr[\pi^{(\eta)} \mapsto (J : \phi_i)] = 0$ if $\pi \not\models \phi_i$.
2. Suppose that there exists an adversary process π_A such that

$$\Pr[(\pi^{(\eta)} \models \neg\gamma_k) \wedge (\pi^{(\eta)} \mapsto (J' : \mathit{accept}))]$$

is not δ-bounded as a function of η.
Let us show that π_A breaks accountability as well. By assumption, J' outputs *reject* iff J outputs a verdict ϕ_i. Hence, the event $\pi^{(\eta)} \mapsto (J' : \mathit{accept})$ is as likely as the event $\neg\exists i(\pi^{(\eta)} \mapsto (J : \phi_i))$, hence,

$$\Pr[(\pi^{(\eta)} \models \neg\gamma_k) \wedge \neg\exists i(\pi^{(\eta)} \mapsto (J : \phi_i))]$$

is also not δ-bounded as a function of η. □

4.3 Privacy and Verifiability

Without additional assumptions, verifiability of Definition 8 is neither essential for the privacy of Definition 11, nor contradicts it. It is not *essential* since e.g. if the adversary violates the property γ_k by directly interacting with the final tally, when the ballots are not linked to the identities of voters anymore, it will not help in breaking privacy. It does not *contradict* privacy e.g. if the Judge's verdict only depends on inputs of dishonest parties.

Considered Attacks. The importance of verifiability for privacy has been demonstrated in [8]. The necessity of avoiding duplicate ballots in order to preserve privacy is mentioned in [3]. While our results and definitions are formally different, the considered actual attacks are of similar nature, and are related to manipulating the ballots which the attacker can link to identities of particular voters. We consider verifiability against particular types of attacks that could be applied to violate the goal γ_k. Let us list the possible cases and briefly summarize our results.

- *Add ballots:* suppose that the attacker is capable of ballot stuffing.
  - If the added ballots *do* depend on the votes of honest voters (e.g. some ballot of an honest voter is replayed), then the attack reduces the privacy of voters whose ballots are replayed.
  - If the added ballots *do not* depend on the votes of honest voters (e.g. are chosen by the attacker or are sampled randomly), then the attack does not directly help in breaking privacy.
- *Drop ballots:* suppose that the attacker is capable of ballot dropping.
  - If the attacker drops ballots of some *honest* voters, it reduces the privacy of the remaining voters who are still counted.

    - If the attacker drops ballots of some *dishonest* voters, it does not directly help in breaking privacy.
- *Substitute ballots:* This attack can be viewed as a combination of ballot adding and dropping. The privacy can be reduced in the following two cases:
    - The inserted ballot *does* depend on the votes of honest voters.
    - The replaced ballot *does not* depend on the votes of honest voters.

It is important that the attacker knows whether ballot manipulation has succeeded or not. For example, if the attacker wanted to replay the ballot of an honest voter k times, but occasionally has replayed the ballot of some dishonest voter, this can lead to a completely wrong decision. We need the notion of a *detectable* protocol property.

Definition 14 (detectable property). *Let $\pi := \pi_A \| \pi_P$ be a voting protocol instance π_P running in parallel with an adversary π_A. Let γ be a property of π. We say that γ is* detectable *in π if*

$$\Pr[(\pi_O \| \pi)^{(\eta)} \overset{O}{\mapsto} 1 \mid \gamma] - \Pr[(\pi_O \| \pi)^{(\eta)} \overset{O}{\mapsto} 1 \mid \neg\gamma] = 1$$

for a passive observer process π_O who has access to the internal state of π_A, but does not directly interact with π_P.

We could quantify the probability in Definition 14 as δ, introducing an extra parameter into relations between privacy and verifiability.

Considered Voting Rules. Many voting systems reveal not just the voting result, but also the full tally, which shows the exact number of votes per candidate. Revealing such information can lead to high privacy leakage. For that reason, some voting systems like Ordinos [13] ensure that only the final result is revealed, e.g. the identity of the winner, and it has been shown in [13] that doing this may reduce privacy leakage significantly. In this work, we want to quantify attacks on privacy that are possible even if only the final result is revealed.

The main idea is that, even if we do not know the particular distribution of votes and cannot compute privacy parameter δ precisely, we can apply the attack on verifiability to change the number of votes that are "known in advance" to the attacker and thus switch between (k, δ_k) and $(k', \delta_{k'})$-privacy. This can be useful for certain kinds of voting rules, satisfying the following definition.

Definition 15 (majority-determined voting rule). *Let n be the total number of voters. A voting rule is called* majority-determined *if it is sufficient to cast $n' = \lfloor \frac{n}{2} \rfloor + 1$ identical votes to determine the election outcome.*

While Definition 15 is trivially satisfied in the case where the election result is a counting histogram of votes, let us show that it holds for a variety of widely used voting rules. The following descriptions of voting rules are taken from [6].

- **Plurality rule.** Each voter votes for one favorite candidate, and the winner is the candidate with the most votes.
  Attack. A candidate j who receives a majority of the votes wins the elections.

- **Borda rule.** Each voter orders candidates by preference and each candidate j gets $m - i$ points in each vote, where i is the rank of j in the vote, and m is the number of candidates; the winner is the candidate with the highest total points.
  Attack. If some candidate j has the first rank in a majority of votes, this ensures the maximum number of points for j.
- **Single transferable vote (STV):** This rule proceeds through a series of rounds. Similar to the Borda rule, each voter orders the candidates by preference. In each round, the candidate that gets the fewest votes ranking it first among the remaining candidates is eliminated, and each of the votes for the eliminated candidate transfers to the next preferred candidate in that vote. The winner is the last remaining candidate.
  Attack. A majority of voters can ensure that a certain candidate wins even if some other candidate receives all the other votes.
- **Maximin.** For any two candidates j and j', let $N(j, j')$ be the number of votes that prefer j to j'. The score of j is $\min_{j'} N(j, j')$.
  Attack. Suppose that the attacker controls a majority n' of the n votes. If the attacker lets j be the most preferable candidate in his controlled votes, then $\min_{j'} N(j, j') \geq n'$. On the other hand, for any other candidate j', we have $N(j', j) \leq n - n'$, hence, $\min_{j''} N(j', j'') \leq n - n' < n'$.
- **Copeland.** For any two candidates j and j', let $C(j, j') = 1$ if $N(j, j') > N(j', j)$, $C(j, j') = 1/2$ if $N(j, j') = N(j', j)$, and $C(j, j') = 0$ if $N(j, j') < N(j', j)$. The Copeland score of candidate j is $s(j) = \sum_{j'} C(j, j')$.
  Attack. If the attacker lets j be the most preferable candidate in his controlled votes, then $N(j, j') > N(j', j)$ for all j', so j gets score $m - 1$, where m is the number of candidates. On the other hand, any other candidate j' will miss the score $C(j', j)$, so $s(j) > s(j')$ for all $j' \leq j$.
- **Bucklin.** For any candidate j and integer l, let $B(j, l)$ be the number of votes that rank candidate j among the top l candidates. The winner is $\arg\min_j(\min_l B(j, l) > n/2)$.
  Attack. If the attacker lets j be the most preferable candidate in his controlled votes, then j wins already for $l = 1$.

While these voting rules guarantee success for an attacker who controls a majority of votes, in practice it is unlikely that all honest voters prefer the same candidate, and the attacker may be successful even controlling way less than half of the votes. This is closely related to the notion of *manipulability* of voting. The authors of [20] have estimated asymptotic bounds for the fraction of voters that are being manipulated to make switching the election outcome hard in the average case. It would be interesting to consider such bounds in future research.

There are some standard voting rules for which Definition 15 does not hold. E.g. in a *veto rule*, each voter gives a score of 0 to one least favorite candidate, and 1 to every other candidate, and the winner is the candidate with the most votes. Here is possible that all voters that are not controlled by the attacker will veto the particular candidate chosen by the attacker, but the attacker does not have enough votes to veto each of the other candidates.

Results. We now show how privacy implies certain types of *targeted* attacks on votes, i.e. where the attacker is able to link manipulated ballots to the identities of corresponding voters who cast these ballots. We will also assume that the attacker *knows* whether the attack has succeeded or not. The main idea is that, for majority-determined voting rules, if $k > n_h/2$, the attacker can always win in the distinguishability game of Definition 11 by taking choices $\vec{c}$ and $\vec{c'}$ that produce different election outcomes. We cannot get a better result without taking into account a particular vote distribution, since it is possible that there is a candidate whom the remaining $n_h - k$ voters will choose with overwhelming probability, resulting in a constant election result r that does not say anything about the victim's choice.

Proposition 2. *Let π_P be an instance of a voting protocol P that uses a majority-determined voting rule, with n_h honest voters V_H. If π_P is (k,δ)-private w.r.t. a subset of voters $V_{pr} \subseteq V_H$ of size k, then π_P is $(n_h - 2k, \delta)$-verifiable against an attacker π_A who has access to $\mathsf{Out}(J)$ who is only able to drop votes of $V_H \setminus V_{pr}$ from the tally, whose success does not depend on the choices of V_H, and the property γ_{n_h-2k} is detectable in $\pi_A \| \pi_P$.*

Proof (Sketch). Regardless of the prior distribution of votes, if a protocol uses a majority-determined voting rule, if $k > n_h/2$, the attacker may always choose votes $c_1, \ldots, c_k$ and $c'_1, \ldots, c'_k$ that determine some election results $r \neq r'$. If $k \leq n_h/2$, the attacker can use the attack on verifiability to drop some of the $n_h - k$ ballots of voters that are not under observation, until a majority of ballots belongs to voters under observation. Suppose that the attacker has managed to drop ℓ ballots. He will control k out of $n - \ell$ ballots. In order to control a majority, he needs $k > (n_h - \ell)/2$, which means $\ell > n_h - 2k$ dropped ballots. If dropping ℓ ballots has failed, the attacker will detect it and output a constant bit, which will be the same regardless of the choices of V_{pr}. Since the protocol is by assumption (k,δ)-private, the attacker should not be able to drop these ℓ ballots with probability larger than δ. $\square$

Proposition 3. *Let π_P be an instance of a voting protocol P that uses a majority-determined voting rule, with n_h honest voters V_H. If π_P is (k,δ)-private w.r.t. a subset of voters $V_{pr} \subseteq V_H$, then P is $(n_h - 2k, \delta)$-verifiable against an attacker π_A who has access to $\mathsf{Out}(J)$, who is only able to duplicate votes of V_{pr} in the tally, whose success does not depend on the choices of V_H, and the property γ_{n_h-2k} is detectable in $\pi_A \| \pi_P$.*

Proof (Sketch). Regardless of the prior distribution of votes, if a protocol uses a majority-determined voting rule, if $k > n_h/2$, the attacker may always choose votes $c_1, \ldots, c_k$ and $c'_1, \ldots, c'_k$ that determine some election results $r \neq r'$. If $k \leq n_h/2$, the attacker can use the attack on verifiability to duplicate some of the k ballots of voters under observation, until a majority of ballots belongs to voters under observation. Suppose that the attacker has managed to produce ℓ duplicates. He will control $k + \ell$ out of $n_h + \ell$ ballots. In order to control a majority, he needs $k + \ell > (n_h + \ell)/2$, which is $\ell > n_h - 2k$ additional ballots.

Since the protocol is by assumption (k, δ)-private, the attacker should not be able to get these additional ℓ ballots with probability larger than δ. $\square$

Propositions 2 and 3 put the same constraint on verifiability, which does not depend on whether the attacker adds or drops the votes. This leads to the following theorem, which is an immediate consequence of the propositions above.

Theorem 3. *Let π_P be an instance of a voting protocol P that uses a majority-determined voting rule, with n_h honest voters V_H. If π_P is (k, δ)-private w.r.t. a subset of voters $V_{pr} \subseteq V_H$, then π_P is $(n_h - 2k, \delta)$-verifiable against an attacker π_A capable of duplicating votes of V_{pr} and dropping votes of $V_H \setminus V_{pr}$, assuming that success of the attack does not depend on the particular choices of the voters V_H, and the property γ_{n_h-2k} is detectable in $\pi_A \| \pi_P$.*

The attacks of Theorem 3 are mostly oriented to small-scale elections with few voters. Suppose that the attacker is interested in a vote of a particular single voter, i.e. $k = 1$. Let there be n_h honest voters for an even n_h. The attacker attempts to drop $\frac{n_h}{2}$ ballots belonging to the remaining $n_h - 1$ voters, and introduces $\frac{n_h}{2}$ copies of the ballot of the vote under observation instead. There are still n_h votes in the final tally, but $\frac{n_h}{2} + 1$ of them are copies of the ballot under observation, so the winner of the election is the main preference of the victim. It is interesting that when the attacker combines vote adding and dropping, in the end, the protocol run may still satisfy γ_{n_h-2k} if the dropped votes occasionally turn out to be the same as the added votes. Such an attack is formally treated as unsuccessful, and in practice, we may get tighter bounds if we measure "success of substituting k votes" instead of "violating γ_{k-1}".

Such types of attack are more interesting in terms of coercion. Suppose that the attacker already controls n_d dishonest voters, and in addition, is able to manipulate ℓ ballots with a high probability of success. If $n_d + \ell < \frac{n}{2}$, then it is not enough to switch the election result and make a certain candidate j the winner. The attacker tries to convince $k = (n_h - \ell)/2$ voters to vote for j. If in the end, j is not the winner, the attacker learns that at least some voters of the coerced group have not obeyed, and may punish them.

4.4 Verifiability and coercion-resistance

Suppose that the attacker is trying to convince a subset of k voters to misbehave. It can be viewed as a variant of coercing abstention from voting (since bad votes are not supposed to be counted), or even an attempt to halt the elections, in the case when Judge's rejection does not allow proceeding with publishing the result. Such kind of attacks, called *fault attacks*, have been considered in [9], and the attacker can apply them to test the loyalty of a voter (or a subset of voters) in a probabilistic way. The following definition allows the attacker to break k-correctness by taking control of a certain number of dishonest voters.

Definition 16 (ballot-corruptible protocol). *An instance π_P of a voting protocol P is called ballot-corruptible if, for all $k \in \mathbb{N}$, there exists a subset of voters*

$V' := \{v_{i1}, \ldots, v_{i\ell}\}$ of size $\ell \leq k+1$, and a joint strategy bad *for these ℓ voters, such that*

$$\Pr[(\pi_{P\setminus V'} \| \mathsf{bad})^{(\eta)} \models \neg\gamma_k] = 1$$

where γ_k is defined as in Definition 7.

We could quantify the probability in Definition 16 as δ, introducing an extra parameter into relations between coercion-resistance and verifiability.

Definition 16 allows the attacker to interact with the protocol in such a way that γ_k will actually be violated and the judging procedure triggered. In practice, the bad voting strategy may correspond to submitting corrupted paper ballots, or malformed digital ballots that e.g. encode several votes in a single ballot. In practice, $\ell \leq k+1$ voters can be sufficient to break γ_k-correctness, e.g. by submitting multiple votes in a single corrupted ballot.

The following theorem estimates the relation between verifiability and coercion-resistance for ballot-corruptible protocols. The idea is that, even if the corrupted final result is not published, the fact that the cheating was detected may already leak something. Since the Judge's decision cannot leak more than a single bit, the attacker needs to encode information into that bit in such a way that it tells whether the inputs of the victim voter(s) are $\vec{c}$ or $\vec{c'}$.

Theorem 4. *Let π_P be an instance of ballot-corruptible voting protocol P with n_h honest voters. Then the following statements cannot be true at once:*

- *π_P is (k,δ)-coercion-resistant (Definition 12) against an attacker who has access to* $\mathsf{Out}(J)$*;*
- *The instance $\pi_{P'}$ of P with $n_h - k$ honest voters is $(k-1, 1-\delta)$-verifiable (Definition 8).*

Proof (Sketch). Let V' be the k voters of π_P to be coerced. Consider the protocol instance $\pi_{P'}$ where V' are treated as corrupted. Let $\pi_{A'}$ be an adversary who sends $\mathsf{corrupt}$ message to V' and follows the strategy bad on their behalf, but does not corrupt any other agents. Let π_A be an adversary that behaves similarly to $\pi_{A'}$, except that it does not send $\mathsf{corrupt}$ message to V', but is just connected to them via the interface of dum. Such π_A satisfies Definition 12. The processes $\pi_{A'} \| \pi_{P'}$ and $\pi_A \| \mathsf{dum} \| \pi_{P\setminus V'}$ differ only in the interface between the protocol and the adversary, but the output of J is the same in these processes.

- If the voters V' obey the attacker in $\pi_A \| \pi_P$, they follow the strategy dum, and since P is ballot-corruptible, the goal γ_{k-1} will be violated. Since $\pi_{P'}$ is $(k-1, 1-\delta)$-verifiable, the Judge will *accept* with probability at most $1-\delta$ in $\pi_{A'} \| \pi_{P'}$, and hence also in $\pi_A \| \mathsf{dum} \| \pi_{P\setminus V'}$.
- While the definition of coercion-resistance does not prohibit that the counter-strategy may violate γ_{k-1}, it is reasonable to assume that the goal of the coerced voters is that the elections end up successfully and the Judge will *accept*. Hence, if the voters V' do not obey the attacker, the Judge will *accept* with a probability 1.

The difference between the probabilities of Judge accepting is at least δ. The attacker outputs 1 iff the Judge accepts, breaking (k, δ)-coercion-resistance. $\square$

In practice, Theorem 4 could be applied by an attacker who coerces k voters to put corrupted ballots into the ballot box. The attacker then looks into the ballot box and sees whether it contains at least k corrupted ballots. In the real world, however, it is not excluded that the "bad" vote can occasionally be cast as well by voters who are not controlled by the attacker, even though it is not intended behaviour. Such voters add certain randomness to the experiment.

If the voting protocol is accountable, the coerced voters might not want that the Judge would accuse them of misbehaviour, so they might not agree to follow the strategy bad unless the attacker threatens them by a more severe punishment than the Judge. However, accountability may in turn provide other means of coercion, as discussed in the following section.

4.5 Privacy and accountability

If the Judge's verdict is independent of the choices of honest participants, it will not harm the privacy of an honest voter in any way. However, as shown in [2], if we want to get a stronger kind of accountability (the *individual accountability*) that allows pinpointing the cheater directly, we may need stronger assumptions. In order to resolve all possible disputes between a voter v_i and a non-voter agent a (such as a voting machine), we need to either assume a semi-trusted a (who processes all received ballots honestly), or the existence of reliable and/or undeniable channels between the voter and the machine, such as voting authorities who actually saw that the voter indeed has interacted with the machine. While an undeniable channel does not leak the exact choice of a voter, it would still at least leak the fact that a voter has voted. Let us formally define an accountability property Φ that does not threaten the privacy of honest voters.

Definition 17 (safe-evidence accountability property). *Let P be a voting protocol instantiation. Let $\delta \in [0,1]$ be the tolerance. Let $\pi_{\vec{i},\vec{c}}$ and $\pi_{\vec{i},\vec{c}'}$ be defined as in Definition 11. We say that the accountability property $\Phi = (\alpha, \phi_1, \ldots, \phi_\ell)$ of P w.r.t. a Judge $J \in \Sigma$ is* (k,δ)-safe-evidence *if*

$$\left|\Pr[\pi_{\vec{i},\vec{c}}^{(\eta)} \overset{A}{\mapsto} 1 \mid \exists j:\ \pi \mapsto (J : \phi_j)] - \Pr[\pi_{\vec{i},\vec{c}'}^{(\eta)} \overset{A}{\mapsto} 1 \mid \exists j:\ \pi \mapsto (J : \phi_j)]\right|$$

is δ-bounded as a function of the security parameter η for all indices of honest voters $\vec{i}$, choices $\vec{c}, \vec{c'}$ and for all adversary processes π_A that have access to the channels $\mathsf{In}(J)$.

Definition 17 says that the evidence for a verdict, based on all inputs that J has received through the channels $\mathsf{In}(J)$, does not depend (much) on the choices of honest voters. The condition $\exists i:\ \pi \mapsto (J : \phi_i)$ ensures that we only consider protocol runs where the Judge has actually made a verdict, which excludes possible attacks that come due to failure of accountability, e.g. leakage via the final result. The definition allows an arbitrary property α.

In order to break privacy, the attacker should first of all be able to violate the condition α, so that the judging procedure would be triggered. Then, in order that the Judge would learn anything interesting, the evidence should depend on the vote of an honest voter under observation, at least telling whether the voter has voted or abstained from voting. The following definition characterizes protocols for which accountability has a direct impact on privacy.

Definition 18 (unsafe accountability property). *Let π_P be an instance of a voting protocol P, Σ the agents of P, $\Phi = (\gamma_k, \phi_1, \ldots, \phi_\ell)$ an accountability property, and $J \in \Sigma$ the Judge. The property Φ is called unsafe in π_P w.r.t J if there exists an adversary π_A such that:*

1. $\Pr[(\pi_P \| \pi_A)^{(\eta)} \vDash \neg\gamma_k] = 1$.
2. *There is a choice $c \in C$ such that, in every run r of π satisfying $\exists i : (J : \phi_i)$, there is a subset $\vec{i_r}$ of $k+1$ honest voters (which can be different in each run) such that $(\pi_P \| \pi_A)^{(\eta)}$ outputs a boolean value $\mathsf{voted}(i, c)$ for all $i \in \vec{i_r}$ to $\mathsf{In}(J)$.*

Intuitively, the second point of Definition 18 says that, whenever the Judge makes a verdict, he learns something about a subset of voters somehow involved in a dispute. The parameter k could be e.g. the minimal number of complaints required to start the dispute resolution procedure. A particular example of an unsafe accountability property would be individual accountability that relies on undeniable channels, assuming that the Judge makes the verdict based on access to these channels. In that case, c would be an abstention vote. Let us show how Definition 18. is related to Definition 17.

Proposition 4. *Let π_P be an instance of a voting protocol P with n_h honest voters. Let Σ be the agents of P, $\Phi = (\gamma_{k'}, \phi_1, \ldots, \phi_\ell)$ an accountability property, and $J \in \Sigma$ the Judge. Let Φ be unsafe in π_P w.r.t J. Then, Φ is not (k, δ)-safe-evidence w.r.t. J and π_A for any $\delta < 1 - \prod_{j=0}^{k'} \left(1 - \frac{k}{n_h - j}\right)$ and any η.*

Proof (Sketch). Let π_A be an adversary that satisfies Definition 18. Consider the runs of $(\pi_P \| \pi_A)^{(\eta)}$ that satisfy $\exists i : (J : \phi_i)$. In each such run r, there is a subset $\vec{i_r}$ of k' voters such that messages $\mathsf{voted}(i, c)$ are sent to a channel of $\mathsf{In}(J)$ for all $i \in \vec{i_r}$. The idea is that the same attacker π_A chooses $\vec{c} = (c, \ldots, c)$ and $\vec{c'} = (c', \ldots, c')$ for $c \neq c'$ to break the safe-evidence property. However, the problems is that $\vec{i_r}$ can be different in each run, but we need a single $\vec{i}$ for all runs. The simplest solution would be to take $k' = n_h - k$, when any subset of size $k' + 1$ always covers at least one victim. However, we can do better since the adversary may choose the $\vec{i}$ itself. In the worst case (from attacker perspective), no subset of voters is preferable, and all voters are equally likely to be exposed to $\mathsf{In}(J)$. The probability that all $k' + 1$ leaked votes are "not interesting" is $\binom{n_h - k}{k'+1} / \binom{n_h}{k'+1}$, which equals $\prod_{j=0}^{k'} \frac{n_h - k - j}{n_h - j} = \prod_{j=0}^{k'} \left(1 - \frac{k}{n_h - j}\right)$. $\square$

The following theorem estimates the relation between privacy and accountability for an unsafe accountability property.

Theorem 5. *Let π_P be an instance of a voting protocol P with n_h honest voters. Let Σ be the agents of P. Let $\Phi = (\gamma_k, \phi_1, \ldots, \phi_\ell)$ and $J \in \Sigma$ be such that Φ is unsafe in π_P w.r.t. J. Then the following statements cannot be true at once:*

- *π_P is (k, δ)-private (Definition 11);*
- *π_P is $(k', 1-\delta/\left(1 - \prod_{j=0}^{k'}\left(1 - \frac{k}{n_h - j}\right)\right))$-accountable w.r.t. Φ (Definition 10).*

Proof (Sketch). Assume that π_P is (k, δ_{acc})-accountable. The condition $\exists i:\ \pi \mapsto (J : \phi_i) \vee \pi \models \gamma_{k'}$ is satisfied with probability at least $1 - \delta_{acc}$. Since Φ is by assumption unsafe in π_P w.r.t. J, there exists an adversary π_A such that $\Pr[(\pi_A \| \pi_P)^{(\eta)} \models \neg\gamma_{k'}] = 1$, so $\exists i:\ \pi \mapsto (J : \phi_i)$ is satisfied with probability at least $1-\delta_{acc}$. Assume that Φ is (k, δ_{ev})-safe-evidence w.r.t. J and π_A. The success of π_A in distinguishing whether the voters $\vec{i}$ have voted or not equals $\delta_{ev}\cdot(1-\delta_{acc})$. Assuming that the protocol is (k, δ_{pr})-private, we have $\delta_{ev} \cdot (1 - \delta_{acc}) < \delta_{pr}$, so $\delta_{ev} < \delta_{pr}/(1-\delta_{acc})$. Now, since Φ is unsafe w.r.t. J, by Proposition 4, it can only be (k, δ_{ev})-safe-evidence w.r.t. J for $\delta_{ev} \geq 1 - \prod_{j=0}^{k'}\left(1 - \frac{k}{n_h - j}\right)$, which gives us $\delta_{acc} > 1 - \delta_{pr}/\left(1 - \prod_{j=0}^{k'}\left(1 - \frac{k}{n_h - j}\right)\right)$, and any smaller δ_{acc} is not suitable. $\square$

In practice, Theorem 5 could be applied by an attacker who takes control over a voting machine that issues receipts for later verification, such as Wombat [1], ThreeBallot, and VAV [21]. The idea is that the corrupted machine will nicely output to all voters appropriate receipts. However, it excludes at least k ballots when displaying information on the bulletin board. With probability at most δ_{acc}, the attack will not be detected, and the Judge does not do anything. Otherwise, there are several outcomes possible.

- The cheating is detected directly by auditors.
- Sufficiently many voters complain after looking at the bulletin board.

In the first case, the Judge does not learn anything interesting from the evidence. In the second case, a subset of voters whose ballots have been dropped come to complain, and the attacker who has corrupted the voting machine can now match the complainer's identity with an affected ballot. If the ballots are not encrypted, the attacker will not only detect that the voter has voted, but also match the corrupted ballot to the complainer's identity and learn the vote.

5 Conclusions and Future Work

In this paper, we have proposed a selection of quantitative definitions of privacy, verifiability, coercion-resistance, and accountability, which are adapted versions of the definitions of the KTV framework. We have shown how these metrics are related to each other, exploring some generic relations that do not depend on the actual distribution of voters' votes.

In practice, the quantitative degree of privacy and coercion-resistance of voting protocols strongly depends on the way in which the voters make their choices.

As the next step, it will be natural to analyse particular vote distributions to get more interesting and tight bounds.

Assuming that the votes are independent, the privacy definition that we have considered in this paper can be viewed as a variant of distributional differential privacy (DDP), albeit DDP estimates the ratio of probabilities instead of the difference. Related work [18] has estimated DDP bounds for various voting rules, and we could study how their definitions of privacy can be combined with verifiability and accountability of the KTV framework.

Acknowledgements The authors are grateful to the anonymous reviewers for their valuable comments. The paper has been supported by the Estonian Research Council under the grant number PRG920.

References

1. Wombat voting system (2011), `http://www.wombat-voting.com/`
2. Basin, D.A., Radomirovic, S., Schmid, L.: Dispute resolution in voting. In: 33rd IEEE Computer Security Foundations Symposium, CSF 2020, Boston, MA, USA, June 22-26, 2020. pp. 1–16. IEEE (2020). https://doi.org/10.1109/CSF49147.2020.00009
3. Bernhard, D., Cortier, V., Galindo, D., Pereira, O., Warinschi, B.: Sok: A comprehensive analysis of game-based ballot privacy definitions. In: 2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015. pp. 499–516. IEEE Computer Society (2015). https://doi.org/10.1109/SP.2015.37
4. Cetinkaya, O.: Analysis of Security Requirements for Cryptographic Voting Protocols (Extended Abstract). In: Proceedings ARES 2008. pp. 1451–1456. IEEE Computer Society (2008)
5. Chevallier-Mames, B., Fouque, P., Pointcheval, D., Stern, J., Traoré, J.: On some incompatible properties of voting schemes. In: Towards Trustworthy Elections, New Directions in Electronic Voting. LNCS, vol. 6000, pp. 191–199. Springer (2010)
6. Conitzer, V., Sandholm, T.: Nonexistence of voting rules that are usually hard to manipulate. In: Proceedings, The Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, July 16-20, 2006, Boston, Massachusetts, USA. pp. 627–634. AAAI Press (2006), `http://www.aaai.org/Library/AAAI/2006/aaai06-100.php`
7. Cortier, V., Galindo, D., Küsters, R., Müller, J., Truderung, T.: Sok: Verifiability notions for e-voting protocols. In: Proceedings of IEEE S&P 2016. pp. 779–798. IEEE Computer Society (2016)
8. Cortier, V., Lallemand, J.: Voting: You can't have privacy without individual verifiability. In: Proceedings of ACM CCS 2018. pp. 53–66 (2018)
9. Delaune, S., Kremer, S., Ryan, M.: Coercion-resistance and receipt-freeness in electronic voting. In: 19th IEEE Computer Security Foundations Workshop, (CSFW-19 2006), 5-7 July 2006, Venice, Italy. pp. 28–42. IEEE Computer Society (2006). https://doi.org/10.1109/CSFW.2006.8
10. Heiberg, S., Willemson, J.: Modeling threats of a voting method. In: Design, Development, and Use of Secure Electronic Voting Systems, pp. 128–148. IGI Global (2014)

11. Jonker, H., Pieters, W.: Anonymity in voting revisited. In: Towards Trustworthy Elections, New Directions in Electronic Voting. pp. 216–230 (2010)
12. Kiayias, A., Zacharias, T., Zhang, B.: End-to-end verifiable elections in the standard model. In: Proceedings of EUROCRYPT 2015, Part II. pp. 468–498 (2015)
13. Küsters, R., Liedtke, J., Müller, J., Rausch, D., Vogt, A.: Ordinos: A verifiable tally-hiding e-voting system. In: IEEE European Symposium on Security and Privacy, EuroS&P 2020, Genoa, Italy, September 7-11, 2020. pp. 216–235. IEEE (2020). https://doi.org/10.1109/EuroSP48549.2020.00022
14. Küsters, R., Müller, J.: Cryptographic security analysis of e-voting systems: Achievements, misconceptions, and limitations. In: Proceedings of E-Vote-ID 2017. pp. 21–41 (2017)
15. Küsters, R., Truderung, T., Vogt, A.: Accountability: definition and relationship to verifiability. In: Proceedings of ACM CCS 2010. pp. 526–535. ACM (2010)
16. Küsters, R., Truderung, T., Vogt, A.: Verifiability, privacy, and coercion-resistance: New insights from a case study. In: Proceedings of IEEE S&P 2011. pp. 538–553. IEEE Computer Society (2011)
17. Küsters, R., Truderung, T., Vogt, A.: Clash attacks on the verifiability of e-voting systems. In: IEEE Symposium on Security and Privacy, SP 2012, 21-23 May 2012, San Francisco, California, USA. pp. 395–409. IEEE Computer Society (2012). https://doi.org/10.1109/SP.2012.32
18. Liu, A., Lu, Y., Xia, L., Zikas, V.: How private are commonly-used voting rules? Cryptology ePrint Archive, Report 2021/392 (2021), `https://eprint.iacr.org/2021/392`
19. Mitrou, L., Gritzalis, D., Katsikas, S.K.: Revisiting Legal and Regulatory Requirements for Secure E-Voting. In: Proceedings of IFIP TC11 17$^{\text{th}}$ International Conference on Information Security (SEC2002. IFIP Conference Proceedings, vol. 214, pp. 469–480. Kluwer (2002)
20. Procaccia, A.D., Rosenschein, J.S.: Average-case tractability of manipulation in voting via the fraction of manipulators. In: Durfee, E.H., Yokoo, M., Huhns, M.N., Shehory, O. (eds.) 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007), Honolulu, Hawaii, USA, May 14-18, 2007. p. 105. IFAAMAS (2007). https://doi.org/10.1145/1329125.1329255
21. RonaldL.Rivest, Smith, W.D.: Three voting protocols: Threeballot, vav, and twin. In: Martinez, R., Wagner, D.A. (eds.) 2007 USENIX/ACCURATE Electronic Voting Technology Workshop, EVT'07, Boston, MA, USA, August 6, 2007. USENIX Association (2007), `https://www.usenix.org/conference/evt-07/three-voting-protocols-threeballot-vav-and-twin`
22. Sako, K., Kilian, J.: Receipt-free mix-type voting scheme - A practical solution to the implementation of a voting booth. In: Proceedings of EUROCRYPT 1995. LNCS, vol. 921, pp. 393–403. Springer (1995)
23. Schryen, G.: Security Aspects of Internet Voting. In: Proceedings of HICSS-37. IEEE Computer Society (2004)

A Full Proofs

A.1 Proof of Proposition 1

Let $\vec{i}_\ell = \{i_1, \ldots, i_\ell\} \subseteq \{1, \ldots, n\}$, and $\vec{c}_\ell, \vec{c'}_\ell \in C^\ell$. We want to estimate the quantity

$$\left|\Pr[\pi^{(\eta)}_{\vec{i}_\ell,\vec{c}_\ell} \overset{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i}_\ell,\vec{c'}_\ell} \overset{A}{\mapsto} 1]\right| \ ,$$

where $\pi_{\vec{i},\vec{c}}$ is defined as in Definition 11. We have

$$\Pr[\pi^{(\eta)}_{\vec{i}_\ell,\vec{c}_\ell} \overset{A}{\mapsto} 1] = \sum_{\vec{c}_{k,\ell} \in C^{k-\ell}} \Pr[\pi^{(\eta)}_{\vec{i}_\ell \| \vec{i}_{k,\ell}, \vec{c}_\ell \| \vec{c}_{k,\ell}} \overset{A}{\mapsto} 1] \cdot \Pr[\vec{c}_{k,\ell}] \ ,$$

where $\Pr[\vec{c}_{k,\ell}]$ is the probability that certain $k-\ell$ voters $\vec{i}_{k,\ell}$ (other than $\vec{i}_\ell$) will vote for $\vec{c}_{k,\ell}$. We have

$$\begin{aligned}
&\left|\Pr[\pi^{(\eta)}_{\vec{i}_\ell,\vec{c}_\ell} \overset{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i}_\ell,\vec{c'}_\ell} \overset{A}{\mapsto} 1]\right| \\
&\quad = \sum_{\vec{c}_{k,\ell} \in C^{k-\ell}} \left|\Pr[\pi^{(\eta)}_{\vec{i}_\ell \| \vec{i}_{k,\ell}, \vec{c}_\ell \| \vec{c}_{k,\ell}} \overset{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i}_\ell \| \vec{i}_{k,\ell}, \vec{c'}_\ell \| \vec{c}_{k,\ell}} \overset{A}{\mapsto} 1]\right| \cdot \Pr[\vec{c}_{k,\ell}] \ .
\end{aligned}$$

For a fixed ℓ, let $\vec{i}_k := \vec{i}_\ell \| \vec{i}_{k,\ell}$, $\vec{c}_k := \vec{c}_\ell \| \vec{c}_{k,\ell}$, and $\vec{c'}_k := \vec{c'}_\ell \| \vec{c'}_{k,\ell}$. Since π_P is by assumption (k,δ)-private,

$$\left|\Pr[\pi^{(\eta)}_{\vec{i}_k,\vec{c}_k} \overset{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i}_k,\vec{c'}_k} \overset{A}{\mapsto} 1]\right| \leq \delta(\eta) \ ,$$

where $\delta(\eta)$ is δ-bounded as a function of η. Hence,

$$\begin{aligned}
\left|\Pr[\pi^{(\eta)}_{\vec{i}_\ell,\vec{c}_\ell} \overset{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i}_\ell,\vec{c'}_\ell} \overset{A}{\mapsto} 1]\right| &\leq \sum_{\vec{c}_{k,\ell} \in C^{k-\ell}} \delta(\eta) \cdot \Pr[\vec{c}_{k,\ell}] \\
&= \delta(\eta) \cdot \sum_{\vec{c}_{k,\ell} \in C^{k-\ell}} \Pr[\vec{c}_{k,\ell}] = \delta(\eta)
\end{aligned}$$

is also δ-bounded as a function of η. □

A.2 Proof of Theorem 1

Let π_P be (k,δ)-coercion-resistant. That is, for any set $\vec{i} = \{i_1, \ldots, i_k\}$ of k coerced voters, any attacker process $\pi_{A'}$ that is connected to $\vec{i}$ using interface of dum, there exists a counter-strategy $\pi_{\tilde{v}}$ such that, for some $\eta_0 > 0$, and all $\eta > \eta_0$, $c > 0$:

$$\Pr[(\pi_{A'} \| \mathsf{dum} \| \pi_{P \backslash \vec{i}})^{(\eta)} \overset{A'}{\mapsto} 1] - \Pr[(\pi_{A'} \| \pi_{\tilde{v}} \| \pi_{P \backslash \vec{i}})^{(\eta)} \overset{A'}{\mapsto} 1] < \delta + \frac{1}{\eta^c} \ . \quad (1)$$

Let $\vec{c}$ and $\vec{c'}$ be vote assignments of the voters $\vec{i}$. Let π_A be an arbitrary attacker process. We want to estimate the quantity

$$\left|\Pr[\pi^{(\eta)}_{\vec{i},\vec{c}} \overset{A}{\mapsto} 1] - \Pr[\pi^{(\eta)}_{\vec{i},\vec{c'}} \overset{A}{\mapsto} 1]\right| = \delta(\eta) \ .$$

Let us now describe a coercer that breaks coercion-resistance. Consider a particular setting where the true goals of the voters $\vec{i}$ is to make the choices $\vec{c}$. Let

$\pi^{\vec{c'}}_{A'}$ be a coercer that selects for the voters the choices $\vec{c'}$, and otherwise acts as an honest voter would. By construction of $\pi^{\vec{c'}}_{A'}$,

$$\Pr[(\pi_A \| \pi^{\vec{c'}}_{A'} \| \mathsf{dum} \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] = \Pr[\pi^{(\eta)}_{\vec{i},\vec{c'}} \overset{A}{\mapsto} 1] \ .$$

Since $\vec{i}$ do not interact with π_A, and $\pi^{\vec{c'}}_{A'}$ is connected to $\vec{i}$ using the interface of dum, the adversary $\pi_{A'} := \pi_A \| \pi^{\vec{c'}}_{A'}$ satisfies Definition 12.

Let $\pi_v = \pi_{v_{i_1}} \| \ldots \| \pi_{v_{i_k}}$. By definition of $\pi_{\vec{i},\vec{c}}$,

$$\Pr[(\pi_A \| \pi_v(\vec{c}) \| \pi_{P\setminus\vec{i}})^{(\eta)} \overset{A}{\mapsto} 1] = \Pr[\pi^{(\eta)}_{\vec{i},\vec{c}} \overset{A}{\mapsto} 1] \ .$$

Since $\pi_{\tilde{v}}$ does not outperform π_v, there exists η_1 such that, for all $\eta > \eta_1$ and all $c > 0$, we have

$$\Pr[(\pi_A \| \pi_v(\vec{c}) \| \pi_{P\setminus\vec{i}})^{(\eta)} \ \overset{A}{\mapsto}\ 1] \ - \ \Pr[(\pi_A \| \pi^{\vec{c'}}_{A'} \| \pi_{\tilde{v}} \| \pi_{P\setminus\vec{i}})^{(\eta)} \ \overset{A}{\mapsto}\ 1] \ < \ \frac{1}{\eta^c} \ .$$

Applying (1), we get that, for all $\eta > \max(\eta_0, \eta_1)$ and all $c > 0$,

$$\delta(\eta) < \delta + \frac{1}{\eta^c} + \frac{1}{\eta^c} \ ,$$

which is δ-bounded as a function of η. Hence, π_P is (k, δ)-private. □

A.3 Proof of Theorem 2

Let π_P be (k, δ)-accountable. That is, there exists η_0 such that, for all $\eta > \eta_0$, all $c > 0$, and all adversaries π_A, for $\pi := (\pi_A \| \pi_P)$, we have

$$\Pr[(\pi^{(\eta)} \models \neg\gamma_k) \wedge \neg\exists i :\ (\pi^{(\eta)} \mapsto (J : \phi_i))] < \delta + \frac{1}{\eta^c} \ ; \tag{2}$$

$$\Pr[\pi^{(\eta)} \mapsto (J : \phi_i)] = 0 \text{ if } \pi \not\models \phi_i \ . \tag{3}$$

Consider the probability

$$\Pr[(\pi^{(\eta)} \models \neg\gamma_k) \wedge (\pi^{(\eta)} \mapsto (J' : \mathit{accept}))] = \delta(\eta) \ .$$

We want to find an upper bound on $\delta(\eta)$. By assumption, J' accepts iff J outputs a verdict ϕ_i. Hence, by (2),

$$\delta(\eta) = \Pr[(\pi^{(\eta)} \models \neg\gamma_k) \wedge \neg\exists i :\ (\pi^{(\eta)} \mapsto (J : \phi_i))] < \delta + \frac{1}{\eta^c} \ .$$

Consider the probability

$$\Pr[\pi^{(\eta)} \mapsto (J' : \mathit{reject})]$$

on the condition that $\pi^{(\eta)} \not\models \mathsf{dis}(a)$ for all $a \in \Sigma$. Since $\phi_i \subseteq \mathcal{F}_{dis}$ for all i, by (3), under this condition, $\Pr[\pi^{(\eta)} \mapsto (J : \phi_i)] = 0$ for any verdict ϕ_i. Hence, $\Pr[\pi^{(\eta)} \mapsto (J : \mathit{reject})] = \Pr[\pi^{(\eta)} \mapsto (J : \phi_i)] = 0$. □

A.4 Proof of Theorem 3

When considering a process π describing a particular instance of a voting protocol, we assume that there is a special channel that publishes the final result $\rho(c_1, \ldots, c_n)$. Let $\pi \overset{R}{\mapsto} y$ denote that the election result of π published through this channel is y.

Theorem 3 follows directly from Propositions 2 and 3. Let us give their full proofs. First of all, we state and prove two helpful lemmata.

Lemma 1. *Let π_P be an instantiation of a voting protocol P with n honest voters V_H. Let $V_{ver}, V_{pr} \subseteq V_H$ be such that $V_{ver} \cap V_{pr} = \emptyset$, $|V_{ver}| = n - n'$, $|V_{pr}| = k$ for some $n' < n$ and $k < n'$. Let $\pi_{P'}$ be an instantiation of P with honest voters $V_H \setminus V_{ver}$. Suppose that:*

- *π_P is (k, δ)-private w.r.t. voters V_{pr};*
- *$\pi_{P'}$ is not (k, δ')-private w.r.t. voters V_{pr} against a passive observer who only observes the final result, for any privacy parameter η;*

Then, π_P is $(n - n' - 1, \delta/\delta')$-verifiable against an attacker $\pi_{A_{ver}}$ such that:

1. *$\pi_{A_{ver}}$ has access to $\mathsf{Out}(J)$;*
2. *$\pi_{A_{ver}}$ is only capable of dropping votes of V_{ver} from the tally and*
3. *the property γ_k is detectable in the process $\pi_{A_{ver}} \| \pi_P$ (Definition 14)*

regardless of the inputs of V_H.

Proof. Suppose that $\pi_{P'}$ is not (k, δ')-private against a passive observer w.r.t. a set of honest voters V_{pr}. That is, there are choices $\vec{c}$ and $\vec{c'}$ for V_{pr} such that

$$\left|\Pr[(\pi_O \| \pi_{P'(\vec{c})})^{(\eta)} \overset{O}{\mapsto} 1] - \Pr[(\pi_O \| \pi_{P'(\vec{c'})})^{(\eta)} \overset{O}{\mapsto} 1]\right| > \delta' \ , \tag{4}$$

for all $\eta > 0$, where $\pi_{P'(\vec{c})} := \pi_{V_{pr}(\vec{c})} \| \pi_{P' \setminus V_{pr}}$ and $\pi_{P'(\vec{c'})} := \pi_{V_{pr}(\vec{c'})} \| \pi_{P' \setminus V_{pr}}$, and π_O is a passive observer process who only observes the final result.

Let $\pi_{A_{ver}}$ be any attacker process that is only capable of vote dropping. Let δ_{ver} be a function of η such that

$$\Pr[((\pi_{A_{ver}} \| \pi_P)^{(\eta)} \vDash \neg\gamma_k) \wedge ((\pi_{A_{ver}} \| \pi_P)^{(\eta)} \mapsto (J : \mathit{accept}))] = \delta_{ver}(\eta) \ .$$

Our goal is to find an upper bound on $\delta_{ver}(\eta)$. Let an event $B(\vec{c})$ be defined as

$$B(\vec{c}) := ((\pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \vDash \neg\gamma_k) \wedge ((\pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \mapsto (J : \mathit{accept})) \ .$$

Since by assumption the attacker's success does not depend on the particular inputs of honest users, for all $\vec{c}$, we have

$$\Pr[B(\vec{c})] = \Pr[B(\vec{c'})] = \delta_{ver}(\eta) \ .$$

Let $\vec{c_h}$ be the choices of the honest voters $V' = V_H \setminus (V_{ver} \cup V_{pr})$. By assumption, $\pi_{A_{ver}}$ is only capable of vote dropping, and there is no other way to violate

γ_k other than by the acts of $\pi_{A_{ver}}$. Hence, every run of $\neg\gamma_k$ ends up in $\rho(\vec{c_h} \uplus \vec{c})$, and the result gets released if J accepts. Hence,

$$\Pr[(\pi_{A_{ver}} \| \pi_{P(\vec{c_h}, \vec{c})})^{(\eta)} \overset{R}{\mapsto} \rho(\vec{c_h} \uplus \vec{c}) \mid B(\vec{c})] = 1 \ . \tag{5}$$

On the other hand, by correctness of a voting protocol,

$$\Pr[(\pi_{P'(\vec{c_h}, \vec{c})})^{(\eta)} \overset{R}{\mapsto} \rho(\vec{c_h} \uplus \vec{c})] = 1 \ . \tag{6}$$

Since (5) and (6) hold for any possible $\vec{c_h}$, we get

$$\Pr[(\pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \overset{R}{\mapsto} r \ | B(\vec{c})] = \Pr[(\pi_{P'(\vec{c})})^{(\eta)} \overset{R}{\mapsto} r]$$

for any result r, and since the decision of π_O only depends on the final result,

$$\Pr[(\pi_O \| \pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \overset{O}{\mapsto} 1 \ | B(\vec{c})] = \Pr[(\pi_O \| \pi_{P'(\vec{c})})^{(\eta)} \overset{O}{\mapsto} 1] \ .$$

Using similar reasoning for choices $\vec{c'}$ of voters V_{pr}, we get

$$\Pr[(\pi_O \| \pi_{A_{ver}} \| \pi_{P(\vec{c'})})^{(\eta)} \overset{O}{\mapsto} 1 \ | B(\vec{c'})] = \Pr[(\pi_O \| \pi_{P'(\vec{c'})})^{(\eta)} \overset{O}{\mapsto} 1] \ .$$

Let us now consider the case $\neg B(\vec{c})$. Since the property γ_k is detectable in $\pi_{A_{ver}} \| \pi_P$, there exists a passive observer $\pi_{O'}$ such that

$$\left| \Pr[(\pi_{O'} \| \pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \mid \gamma_k] - \Pr[(\pi_{O'} \| \pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \mid \neg\gamma_k] \right| = 1 \ . \tag{7}$$

A similar statement holds for $\vec{c'}$. We can now define an attacker π_A who runs $\pi_O \| \pi_{O'} \| \pi_{A_{ver}}$ and acts as follows:

- Outputs a constant bit (either 1 or 0) if J rejects.
- Outputs a constant bit (either 1 or 0) if J accepts and γ_k is detected by $\pi_{O'}$.
- Outputs the output of π_O if J accepts and $\neg\gamma_k$ is detected by $\pi_{O'}$.

From (7), we get that π_A always outputs a constant bit whenever $\neg B(\vec{c})$ or $\neg B(\vec{c'})$ is true. Hence,

$$\left| \Pr[(\pi_A \| \pi_{P(\vec{c})})^{(\eta)} \overset{A}{\mapsto} 1 \mid \neg B(\vec{c})] - \Pr[(\pi_A \| \pi_{P(\vec{c'})})^{(\eta)} \overset{A}{\mapsto} 1 \mid \neg B(\vec{c'}) \right| = 0 \ .$$

Since the observers π_O and $\pi_{O'}$ do not introduce additional corruptions, π_A is a valid attacker on privacy of $V_{pr} \subseteq V_H$. Since π_P is (k, δ)-private, there exists η_0 such that, for all $\eta > \eta_0$ and all $c' > 0$:

$$\left| \Pr[(\pi_A \| \pi_{P(\vec{c})})^{(\eta)} \overset{O}{\mapsto} 1] - \Pr[(\pi_A \| \pi_{P(\vec{c'})})^{(\eta)} \overset{O}{\mapsto} 1] \right| \leq \delta + \frac{1}{\eta^{c'}} \ ,$$

where $\pi_{P(\vec{c})} := \pi_{V_{pr}(\vec{c})} \| \pi_{P \setminus V_{pr}}$ and $\pi_{P(\vec{c'})} := \pi_{V_{pr}(\vec{c'})} \| \pi_{P \setminus V_{pr}}$. We can rewrite this inequality using conditioning over $B(\vec{c})$ and $B(\vec{c'})$ and the reversed triangle inequality as

$$
\begin{aligned}
&|\left|\Pr[(\pi_A \| \pi_{P(\vec{c})})^{(\eta)} \overset{A}{\mapsto} 1 \mid B(\vec{c})] - \Pr[(\pi_A \| \pi_{P(\vec{c'})})^{(\eta)} \overset{A}{\mapsto} 1 \mid B(\vec{c'})]\right| \cdot \Pr[B(\vec{c'})] \\
&- \left|\Pr[(\pi_A \| \pi_{P(\vec{c})})^{(\eta)} \overset{A}{\mapsto} 1 \mid \neg B(\vec{c})] - \Pr[(\pi_A \| \pi_{P(\vec{c'})})^{(\eta)} \overset{A}{\mapsto} 1 \mid \neg B(\vec{c'})\right| \cdot \Pr[\neg B(\vec{c'})]| \\
&\leq \delta + \frac{1}{\eta^{c'}} \ .
\end{aligned}
$$

We can rewrite it as

$$
\left|\Pr[(\pi_A \| \pi_{P'(\vec{c})})^{(\eta)} \overset{A}{\mapsto} 1] - \Pr[(\pi_A \| \pi_{P'(\vec{c'})})^{(\eta)} \overset{A}{\mapsto} 1]\right| \cdot \delta_{ver}(\eta) < \delta + \frac{1}{\eta^{c'}} \ .
$$

Applying (4), we get

$$
\delta' \cdot \delta_{ver}(\eta) < \delta + \frac{1}{\eta^{c}} \ ,
$$

i.e. $\delta_{ver}(\eta) < \frac{\delta + \frac{1}{\eta^c}}{\delta'} = \frac{\delta}{\delta'} + \frac{1}{\delta' \cdot \eta^c}$. We want to show that this quantity is bounded by $\frac{\delta}{\delta'} + \frac{1}{\eta^c}$ for all c. Since the inequality holds for all c', let us take $c' := 2 \cdot c$. We have $\frac{2}{\delta' \cdot \eta^{2c}} \leq \frac{1}{\eta^c}$ for $\eta > \sqrt[c]{\frac{1}{\delta'}}$. Hence, the desired inequality holds for all $\eta > \max(\sqrt[c]{\frac{1}{\delta'}}, \eta_0)$, so the protocol instance π_P is $(n - n' - 1, \delta/\delta')$-verifiable. $\square$

Lemma 2. *Let π_P be an instantiation of a voting protocol P with n honest voters V_H. Let $V_{pr} \subseteq V_H$ and V_{ver} be such that, if $\vec{c}$ are choices made by V_{pr}, then $\tau(\vec{c})$ for a function $\tau : C^{|V_{pr}|} \mapsto C^{|V_{ver}|}$ are choices made by V_{ver}. Let $|V_{ver}| = n' - n$, $|V_{pr}| = k$ for some $n < n'$ and $k < n$. Let $\pi_{P'}$ be an instantiation of P with honest voters $V_H \cup V_{ver}$. Suppose that:*

- *π_P is (k, δ)-private w.r.t. voters V_{pr};*
- *$\pi_{P'}$ is not $(n' - n + k, \delta')$-private w.r.t. voters $V_{pr} \cup V_{ver}$ against a passive observer, for any privacy parameter η;*

Then, π_P is $(n' - n - 1, \delta/\delta')$-verifiable against an attacker $\pi_{A_{ver}}$ such that:

1. *$\pi_{A_{ver}}$ has access to $\mathsf{Out}(J)$;*
2. *$\pi_{A_{ver}}$ is only capable of adding to the tally votes that are result of applying τ to the voters of V_{pr};*
3. *the property γ_k is detectable in the process $\pi_{A_{ver}} \| \pi_P$ (Definition 14)*

regardless of the inputs of V_H.

Proof. Suppose that $\pi_{P'}$ is not (k, δ')-private against a passive observer w.r.t. a set of honest voters $V_{pr} \cup V_{ver}$. That is, there are choices $\vec{c}$ and $\vec{c'}$ of V_{pr} such that

$$
\left|\Pr[(\pi_O \| \pi_{P'(\vec{c}, \tau(\vec{c}))})^{(\eta)} \overset{O}{\mapsto} 1] - \Pr[(\pi_O \| \pi_{P'(\vec{c'}, \tau(\vec{c'}))})^{(\eta)} \overset{O}{\mapsto} 1]\right| \geq \delta' \ , \qquad (8)
$$

for all $\eta > 0$, where

$$\pi_{P'(\vec{c},\tau(\vec{c}))} := \pi_{V_{pr}(\vec{c})} \| \pi_{V_{ver}(\tau(\vec{c}))} \| \pi_{P' \setminus (V_{pr} \cup V_{ver})} \ ;$$
$$\pi_{P'(\vec{c'},\tau(\vec{c'}))} := \pi_{V_{pr}(\vec{c'})} \| \pi_{V_{ver}(\tau(\vec{c'}))} \| \pi_{P' \setminus (V_{pr} \cup V_{ver})} \ ;$$

and π_O is a passive observer process that only sees the final result.

Let $\pi_{A_{ver}}$ be any attacker that is only capable of adding votes $\tau(\vec{c})$. Let δ_{ver} be a function of η such that

$$\Pr[((\pi_{A_{ver}} \| \pi_P)^{(\eta)} \models \neg\gamma_k) \wedge ((\pi_{A_{ver}} \| \pi_P)^{(\eta)} \mapsto (J : \mathit{accept}))] = \delta_{ver}(\eta) \ .$$

Our goal is to find an upper bound on $\delta_{ver}(\eta)$. Let an event $B(\vec{c})$ be defined as

$$B(\vec{c}) := ((\pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \models \neg\gamma_k) \wedge ((\pi_{A_{ver}} \| \pi_{P(\vec{c})})^{(\eta)} \mapsto (J : \mathit{accept})) \ .$$

Since by assumption the attacker's success does not depend on the particular inputs of honest users, for all $\vec{c}$, we have

$$\Pr[B(\vec{c})] = \Pr[B(\vec{c'})] = \delta_{ver}(\eta) \ .$$

Let $\vec{c_h}$ be the choices of the honest voters $V' = V_H \setminus (V_{ver} \cup V_{pr})$. By assumption, $\pi_{A_{ver}}$ is only capable of adding votes $\tau(\vec{c})$, and there is no other way to violate γ_k other than by the acts of $\pi_{A_{ver}}$. Hence, every run of $\neg\gamma_k$ ends up in $\rho(\vec{c_h} \uplus \vec{c} \uplus \tau(\vec{c}))$, and the result gets released if J accepts. Hence,

$$\Pr[(\pi_{A_{ver}} \| \pi_{P(\vec{c_h},\vec{c})})^{(\eta)} \overset{R}{\mapsto} \rho(\vec{c_h} \uplus \vec{c} \uplus \tau(\vec{c})) \mid B(\vec{c})] = 1 \ . \tag{9}$$

On the other hand, by correctness of a voting protocol,

$$\Pr[(\pi_{P'(\vec{c_h},\vec{c},\tau(\vec{c}))})^{(\eta)} \overset{R}{\mapsto} \rho(\vec{c_h} \uplus \vec{c} \uplus \tau(\vec{c}))] = 1 \ . \tag{10}$$

Since (9) and (10) hold for any possible $\vec{c_h}$, we get

$$\Pr[(\pi_{A_{ver}} \| \pi_{P(\vec{c},\tau(\vec{c}))})^{(\eta)} \overset{R}{\mapsto} r \ | B(\vec{c})] = \Pr[(\pi_{P'(\vec{c})})^{(\eta)} \overset{R}{\mapsto} r]$$

for any result r, and since the decision of π_O only depends on the final result,

$$\Pr[(\pi_O \| \pi_{A_{ver}} \| \pi_{P(\vec{c},\tau(\vec{c}))})^{(\eta)} \overset{O}{\mapsto} 1 \ | B(\vec{c})] = \Pr[(\pi_O \| \pi_{P'(\vec{c})})^{(\eta)} \overset{O}{\mapsto} 1] \ .$$

Using similar reasoning for choices $\vec{c'}$ of voters V_{pr}, we get

$$\Pr[(\pi_O \| \pi_{A_{ver}} \| \pi_{P(\vec{c'},\tau(\vec{c'}))})^{(\eta)} \overset{O}{\mapsto} 1 \ | B(\vec{c'})] = \Pr[(\pi_O \| \pi_{P'(\vec{c'})})^{(\eta)} \overset{O}{\mapsto} 1] \ .$$

The rest of the proof is analogous to the proof of Lemma 1. $\square$

Proof of Proposition 2. Let π_P be an instance of a protocol that uses a majority-determined voting rule. That is, for $k > n/2$, the attacker may come up with votes $c_1, \dots, c_k$ and $c'_1, \dots, c'_k$ that determine some election results r and r' such that $r \neq r'$.

Let $n' < n$ be odd. In this case, any instantiation $\pi_{P'}$ of P with n' honest voters can only be $((n'-1)/2+1, 1)$-private.

Fix any subset V_{pr} of size $(n'-1)/2+1$. Fix any subset $V_{ver} \subseteq V_H \setminus V_{pr}$ of size $n-n'$ and instantiate $\pi_{P'}$ on $V_H \setminus V_{ver}$. Consider an attacker A who is capable of dropping votes of V_{ver} from the tally. By Lemma 1, if π_P is $((n'-1)/2+1, n, \delta)$-private w.r.t. V_{pr}, since $\pi_{P'}$ can only be $((n'-1)/2+1, 1)$-private regardless of the choice of V_{ver}, we have that π_P is $(n-n'-1, \delta)$-verifiable against A. This holds for any choice of V_{pr} and V_{ver}.

Putting $k = (n'-1)/2+1$, we get that, for all $k < (n+1)/2$, if π_P is (k, δ)-private, then π_P is $(n-2k, \delta)$-verifiable against vote dropping. □

Proof of Proposition 3. Let π_P be an instance of a protocol that uses a majority-determined voting rule. That is, for $k > n/2$, the attacker may choose votes $c = c_1 = \dots = c_k$ and $c' = c'_1 = \dots = c'_k$ that determine some election results r and r' such that $r \neq r'$.

Let $n' > n$ be odd. Since the election result (r or r') tells whether a majority voted for c or for c', any instantiation $\pi_{P'}$ of P with n' honest voters can only be $((n'-1)/2+1, 1)$-private. In order to match the second condition of Lemma 2, we need k such that $n'-n+k = (n'-1)/2+1$, i.e. $k = n-(n'+1)/2+1$.

Fix any subset V_{pr} of size $(n'-1)/2+1$. Let V_{ver} make the choices $\tau(\vec{c})$ where $\vec{c}$ are choices of V_{pr}, and τ introduces $n'-n$ copies of any vote of V_{pr}. Instantiate $\pi_{P'}$ on $V_H \cup V_{ver}$. Consider an attacker A who is capable of adding votes of V_{ver} to the tally. By Lemma 2, if π_P is $(n-(n'+1)/2+1, \delta)$-private w.r.t. V_{pr}, since $\pi_{P'}$ can only be $((n'-1)/2+1, 1)$-private regardless of the choice of V_{ver}, we have that π_P is $(n'-n-1, \delta)$-verifiable against A. This holds for any choice of V_{pr} and V_{ver}.

Putting $k = n-(n'+1)/2+1$, we get that, for all $k < n-(n+1)/2+1 = (n-1)/2+1$, if π_P is (k, δ)-private, then π_P is $(n-2k, \delta)$-verifiable against vote duplication. □

A.5 Proof of Theorem 4

Let π_P be (k, δ)-coercion-resistant. Let V_H and V_D be the sets of honest and dishonest voters of π_P respectively. Let $V_{\vec{i}} = v_{i1}, \dots, v_{ik} \subseteq V_H$ be the set of coerced voters, and π_A a coercer algorithm which chooses the strategy bad for the coerced voters. Note that π_A does not send $\mathsf{corrupt}$ messages to $V_{\vec{i}}$, but is just connected to them via interface of dum. Let $\pi_{P(1)}$ be the protocol runs where $V_{\vec{i}}$ decided to obey the attacker, and let $\pi_{P(0)}$ be the protocol runs where they decided to follow their own strategy. We assume that the goal of $V_{\vec{i}}$ is to make choices $\vec{c}$, and that they want that J would accept the protocol run. Since P is by assumption ballot-corruptible,

$$\Pr[(\pi_A\|\pi_{P(1)})^{(\eta)} \models \neg\gamma_k] = 1$$

i.e

$$\Pr[(\pi_A\|\pi_{P(1)})^{(\eta)} \models \gamma_k] = 1 - \Pr[(\pi_A\|\pi_{P(1)})^{(\eta)} \models \neg\gamma_k] = 0 \ . \tag{11}$$

Suppose that π_A orders all other agents under his control (other than $V_{\vec{i}}$) to behave honestly. That is,

$$\forall a \in \Sigma \setminus V_{\vec{i}}, \forall b \in \{0,1\} : \Pr[\pi_A\|\pi_{P(b)} \not\models \mathsf{dis}(a)] = 1 \ .$$

Let $\pi_{P'}$ be the protocol instance run with honest voters $V_H \setminus V_{\vec{i}}$ and dishonest voters $V_D \cup V_{\vec{i}}$. The difference between π_P and $\pi_{P'}$ is that $\pi_{P(b)}$ receives b from the environment (voters' own inputs), while $\pi_{P'}$ receives instructions from the adversary. Formally, we have $\pi_A\|\pi_{P(1)} = \pi_{A'}\|\pi_{P'}$, where $\pi_{A'}$ corrupts the voters $V_{\vec{i}}$, runs the strategy bad on their behalf, and in parallel runs π_A, where, instead of linking π_A directly to $V_{\vec{i}}$, $\pi_{A'}$ forwards the messages between the dum interface of $V_{\vec{i}}$ and π_A. Let us estimate the verifiability of $\pi_{P'}$. Suppose that

$$\Pr[((\pi_A\|\pi_{P(1)})^{(\eta)} \models \neg\gamma_k) \wedge ((\pi_A\|\pi_{P(1)})^{(\eta)} \mapsto (J : \mathit{accept}))] < \delta'(\eta) \tag{12}$$

for some function δ'. Applying (11) and (12),

$$\begin{aligned}
&\Pr[((\pi_A\|\pi_{P(1)})^{(\eta)} \mapsto (J : \mathit{accept}))] = \\
&\qquad \Pr[((\pi_A\|\pi_{P(1)})^{(\eta)} \mapsto (J : \mathit{accept})) \wedge (\pi_A\|\pi_{P(1)})^{(\eta)} \models \neg\gamma_k] \\
&\qquad + \Pr[((\pi_A\|\pi_{P(1)})^{(\eta)} \mapsto (J : \mathit{accept})) \wedge (\pi_A\|\pi_{P(1)})^{(\eta)} \models \gamma_k] < \delta'(\eta) \ .
\end{aligned}$$

Since π_P is (k,δ)-coercion-resistant, there exists $\eta_0 > 0$ such that, for all $\eta > \eta_0$, and all $c > 0$:

$$\left|\Pr[(\pi_A\|\pi_{P(0)})^{(\eta)} \overset{C}{\mapsto} 1] - \Pr[(\pi_A\|\pi_{P(1)})^{(\eta)} \overset{C}{\mapsto} 1]\right| < \delta + \frac{1}{\eta^c} \tag{13}$$

We have assumed that the goal of voters $V_{\vec{i}}$ is that J would accept the run. Hence,

$$\Pr[((\pi_A\|\pi_{P(0)})^{(\eta)} \mapsto (J : \mathit{reject})] = 0 \ .$$

By assumption, π_A has access to $\mathsf{Out}(J)$. Let π_A output a bit 1 iff J outputs *accept*. For all $\eta > \eta_0$ and all $c > 0$, we get

$$\begin{aligned}
\delta + \frac{1}{\eta^c} &\geq \left|\Pr[(\pi_A\|\pi_{P(0)})^{(\eta)} \overset{C}{\mapsto} 1] - \Pr[(\pi_A\|\pi_{P(1)})^{(\eta)} \overset{C}{\mapsto} 1]\right| \\
&= \left|\Pr[(\pi_A\|\pi_{P(0)})^{(\eta)} \mapsto (J : \mathit{accept})] - \Pr[(\pi_A\|\pi_{P(1)})^{(\eta)} \overset{C}{\mapsto} \mathit{accept}]\right| \\
&= \left|1 - \Pr[(\pi_A\|\pi_{P(1)})^{(\eta)} \mapsto (J : \mathit{accept})]\right| \\
&= \left|1 - \Pr[(\pi_{A'}\|\pi_{P'})^{(\eta)} \mapsto (J : \mathit{accept})]\right| \\
&> 1 - \delta'(\eta) \ ,
\end{aligned}$$

which gives us

$$\delta'(\eta) > 1 - \delta - \frac{1}{\eta^c} \ .$$

Suppose by contrary that $\pi_{P'}$ is $(k, 1-\delta-\epsilon)$ verifiable for some $\epsilon > 0$. There exists η_1 such that, for all c and all $\eta > \max(\eta_0, \eta_1)$ it should be that $1 - \delta - \epsilon + \frac{1}{\eta^c} > 1 - \delta - \frac{1}{\eta^c}$. This inequality fails for $\epsilon \geq \frac{2}{\eta^c}$, i.e. $\eta \geq \sqrt[c]{\frac{2}{\epsilon}}$, and the only case when it is true for any η is $\epsilon = 0$. Hence, $\pi_{P'}$ can at most be $1 - \delta$-verifiable. Since $V_{\vec{i}}$ is a freely chosen set of k voters, the property holds for any instance $\pi_{P'}$ with $n_h - k$ honest voters. $\square$

A.6 Proof of Theorem 5

For $\vec{i} = \{i_1, \ldots, i_k\}$ and $\vec{c} = (c_1, \ldots, c_k)$, let $\mathsf{voted}(\vec{i}, \vec{c})$ denote the set of protocol runs in which the choice of v_{i_j} is c_j for all $j \in \{1, \ldots, k\}$.

Proof of Proposition 4. Let π_A be an adversary that satisfies Definition 18. Consider the runs of $(\pi_P \| \pi_A)^{(\eta)}$ that satisfy $\exists i : \ (J : \phi_i)$. In each such run r, there is a subset $\vec{i}_r$ of k' voters such that messages $\mathsf{voted}(i, c)$ are sent to a channel of $\mathsf{In}(J)$ for all $i \in \vec{i}_r$.

To break the safe-evidence property, we construct an attacker $\pi_{A'} := \pi_J \| \pi_A$, where π_J is a component that has access to the channels $\mathsf{In}(J)$.

Let $\vec{c} = (c, \ldots, c)$ where c is the same choice that has been used in messages $\mathsf{voted}(i, c)$, and let $\vec{c'} = (c', \ldots, c')$ for $c \neq c'$. in order to break safe-evidence property, we need to distinguish between $\vec{c}$ and $\vec{c'}$ for a certain subset of voters $\vec{i}$. However, the component π_J may receive messages $\mathsf{voted}(i, c)$ for different $\vec{i}_r$, and it is possible that $\vec{i} \cap \vec{i}_r = \emptyset$ for some runs r. Let $\vec{i}$ be chosen in such a way that the probability of getting $\vec{i} \cap \vec{i}_r = \emptyset$ is minimized. In the worst case, no subset $\vec{i}$ is preferable, and all voters are equally likely to be exposed. The probability that all $k' + 1$ leaked votes are "not interesting" is

$$\begin{aligned}
\binom{n_h - k}{k' + 1} \Big/ \binom{n_h}{k' + 1} &= \frac{(n_h - k)!(n_h - k' - 1)!}{n_h!(n_h - k - k' - 1)!} \\
&= \frac{(n_h - k) \cdots (n_h - k - k')}{n_h \cdots (n_h - k')} \\
&= \prod_{j=0}^{k'} \frac{n_h - k - j}{n_h - j} = \prod_{j=0}^{k'} \left(1 - \frac{k}{n_h - j}\right) \ .
\end{aligned}$$

Let $\pi_{A'}$ return 1 iff π_J receives $\mathsf{voted}(i, c)$ for $i \in \vec{i}$. Let $\pi_{A'}$ return 0 in the other cases, i.e. when π_J receives $\mathsf{voted}(i, c')$ for $i \in \vec{i}$, or does not receive any messages for $i \in \vec{i}$ at all. We get

$$|\Pr[(\pi_{A'} \| \pi_P)^{(\eta)} \overset{A'}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c}) \wedge \exists j:\ \pi \mapsto (J : \phi_j)] \\ - \Pr[(\pi_{A'} \| \pi_P)^{(\eta)} \overset{A'}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c'}) \wedge \exists j:\ \pi \mapsto (J : \phi_j)]| \\ \geq 1 - \prod_{j=0}^{k'} \left(1 - \frac{k}{n_h - j}\right) \quad ,$$

so π_P can only be (k, δ_{ev})-safe-evidence for $\delta_{ev} > 1 - \prod_{j=0}^{k'} \left(1 - \frac{k}{n_h - j}\right)$. □

Proof of the theorem. Let π_A be an attacker that satisfies Definition 18. Let $\vec{i}$, $\vec{c}$, $\vec{c'}$ be such that the safe-evidence property w.r.t. π_A is violated. Denote $\pi := (\pi_A \| \pi_P)^{(\eta)}$ and $\pi(\vec{c}) := (\pi_A \| \pi_{P(\vec{c})})^{(\eta)}$. Since π_P is (k, δ_{pr})-private, there exists η_0 such that, for all $\eta > \eta_0$ and all $c' > 0$, we have

$$\begin{aligned} \delta_{pr} + \frac{1}{\eta^{c'}} &> \left|\Pr[(\pi_A \| \pi_{P(\vec{c})})^{(\eta)} \overset{A}{\mapsto} 1] - \Pr[(\pi_A \| \pi_{P(\vec{c'})})^{(\eta)} \overset{A}{\mapsto} 1]\right| \\ &= \left|\Pr[\pi(\vec{c}) \overset{A}{\mapsto} 1] - \Pr[\pi(\vec{c'}) \overset{A}{\mapsto} 1]\right| \\ &= \left|\Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c})] - \Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c'})]\right| \quad . \end{aligned}$$

Using the Bayesian formula, we get

$$\begin{aligned} &\Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c})] \\ &\quad = \frac{\Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)] \cdot \Pr[\exists i:\ \pi \mapsto (J : \phi_i)]}{\Pr[\exists i:\ \pi \mapsto (J : \phi_i) \mid \pi \overset{A}{\mapsto} 1 \wedge \mathsf{voted}(\vec{i}, \vec{c})]} \\ &\quad \geq \Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)] \cdot \Pr[\exists i:\ \pi \mapsto (J : \phi_i)] \end{aligned}$$

Since π_P is (k, δ_{acc}) accountable w.r.t. J and the accountability constraint Φ, there exists η_1 such that, for all $\eta > \eta_1$ and all $c' > 0$, we have

$$\Pr[(\pi \vDash \neg\gamma_k) \wedge \neg\exists i:\ \pi \mapsto (J : \phi_i)] \leq \delta_{acc} + \frac{1}{\eta^{c'}} \quad ,$$

i.e.

$$\Pr[(\pi \vDash \gamma_k) \vee \exists i:\ \pi \mapsto (J : \phi_i)] > 1 - \delta_{acc} - \frac{1}{\eta^{c'}} \quad .$$

By assumption, there exists η_2 such that, for $\eta > \eta_2$, we have $\Pr[\pi^{(\eta)} \vDash \gamma_k] \leq \frac{1}{\eta^{c'}}$. For $\eta > \max(\eta_1, \eta_2)$, we get

$$\begin{aligned} \Pr[\exists i:\ \pi \mapsto (J : \phi_i)] &= \Pr[\pi \vDash \gamma_k] + \Pr[\exists i:\ \pi \mapsto (J : \phi_i)] - \Pr[\pi \vDash \gamma_k] \\ &> \Pr[(\pi \vDash \gamma_k) \vee \exists i:\ \pi \mapsto (J : \phi_i)] - \frac{1}{\eta^{c'}} \\ &> 1 - \delta_{acc} - \frac{2}{\eta^{c'}} \quad . \end{aligned}$$

For $\eta > \max(\eta_0, \eta_1, \eta_2)$, we get

$$\begin{aligned}
\delta_{pr} + \frac{1}{\eta^{c'}} > & \,|\Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)] \\
& - \Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c'}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)]| \cdot \Pr[\exists i:\ \pi \mapsto (J : \phi_i)] \\
> & \,|\Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)] \\
& - \Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c'}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)]| \cdot (1 - \delta_{acc} - \frac{2}{\eta^{c'}})
\end{aligned}$$

which gives us

$$\begin{aligned}
& |\Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)] \\
& \qquad - \Pr[\pi \overset{A}{\mapsto} 1 \mid \mathsf{voted}(\vec{i}, \vec{c'}) \wedge \exists i:\ \pi \mapsto (J : \phi_i)]| \\
& \qquad\qquad < \frac{\delta_{pr} + \frac{1}{\eta^{c'}}}{1 - \delta_{acc} - \frac{2}{\eta^{c'}}} \ . \qquad (14)
\end{aligned}$$

On the other hand, by Proposition 4, this probability is at least as large as $\delta_{ev} = 1 - \prod_{j=0}^{k'} \left(1 - \frac{k}{n_h - j}\right)$. This gives us

$$\delta_{ev} < \frac{\delta_{pr} + \frac{1}{\eta^{c'}}}{1 - \delta_{acc} - \frac{2}{\eta^{c'}}} \ ,$$

i.e.

$$\delta_{acc} > 1 - \frac{\delta_{pr} + \frac{1}{\eta^{c'}}}{\delta_{ev}} - \frac{2}{\eta^{c'}} \ .$$

Since the inequality must hold for an arbitrary large η, we can only have $\delta_{acc} > 1 - \delta_{pr}/\delta_{ev}$.

□