Assorter-Based RLAs for Complex Sainte-Laguë Elections: Reducing sample sizes by trusting manually counted ballots

Jurlind Budurushi[2] [0000−0002−6732−4400], Tobias Hilt[1] [0000−0001−9267−5109], Christian Mack[1] [0009−0002−9298−0589], Vanessa Teague[3] [0000−0003−2648−2565], and Melanie Volkamer[1] [0000−0003−2674−4043]

[1] Karlsruhe Institute of Technology, Karlsruhe, Germany
[2] Baden-Wuerttemberg Cooperative State University Karlsruhe, Karlsruhe, Germany
[3] Thinking Cybersecurity Pty. Ltd., Melbourne, Australia

Abstract. In several states of Germany, elections for the city council are very complex. They allow voters to cast their vote for one party (so called unchanged ballots) or to distribute their votes to individual candidates from one or several parties (so called changed ballots). While the unchanged ballots are manually tallied by the poll workers, the changed ones are entered into tallying software. We study how to use assorter-based Risk Limiting Audits (RLAs) for such elections. We examine whether we can reduce sample sizes for a given risk limit by focusing the audit on the changed ballots, i.e., not trusting the tallying software.

Keywords: Risk Limiting Audits · Complex Election Rules

1 Introduction

In Germany, votes are cast on paper, but tallying software is widely employed to record and tally ballots that are too complex for manual processing. Among these elections, there are the city council elections in Baden-Württemberg. For such elections, voters can cast (1) so called *unchanged ballots*, which assign one vote to each candidate on a party list or (2) so called *changed ballots*, which may allocate votes to candidates from one or multiple parties and/or up to three votes to an individual candidate. The first type are straightforward to manually tally and to audit by a full recount. Therefore, we consider the tally of the *unchanged ballots* as trusted. The second type are recorded and tallied by software. Currently, this tallying software must be trusted for the tallying of the *changed ballots* (which is about 60% of all votes). Therefore, we consider the tally of the *changed ballots* as untrusted.

This paper presents an audit process for local elections in Baden-Württemberg that is, *risk-limiting* under the assumption that the tallies of the *unchanged ballots* are correct. Risk limiting means that a limit $\alpha > 0$ can be set in advance and the audit process guarantees that if the announced result is wrong, the audit will not pass, except with probability at most α. The objective, as with any

Risk-limiting Audit (RLA), is to check the result—in this case, the assignment of seats to parties—without necessarily checking that the tallies are exactly right.

The contribution of this paper is to examine whether trusting part of the tally can reduce sample sizes to reasonable levels. Building on the approach of Blom et al. [1], we adapt the SHANGRLA framework [9] to the context of *trusted* and *untrusted* tallies. Thereby, we derive the corresponding set of assertions and formulate their associated assorters. Our simulations show a modest improvement (decrease) in estimated sample sizes, which is most significant when the margin is very small. As most information about the local elections of Baden-Württemberg is in German, we first provide a summary for non-German speakers and also explain why we consider the tally of the *unchanged ballots* as trusted.

2 Related Work

Research on RLAs [7,8] has simplified their practical application and extended them to support various social choice functions, including plurality, majority and ranked-choice. The SHANGRLA [9] framework generalized the conduct of RLAs, in particular for social choice functions that are *scoring rules*.

Stark et al. [10] proposed RLAs for proportional representation in the context of European Parliamentary elections by expressing outcome correctness as the satisfaction of a finite set of linear inequalities, thus paving the way for the construction of assertions and assorters. Subsequent research extended assertion-based RLAs to a range of social choice functions, including Hamiltonian elections [3], Instant-Runoff Voting elections [6], Condorcet elections [5], and Single Transferable Vote elections [2,4]. Most importantly, Blom et al.[1] showed how linear inequalities can be expressed in SHANGRLA's canonical form, thus providing a general method for deriving assorters from such inequalities. They demonstrated their approach on party-list proportional elections, including Hamiltonian free-list elections and elections using the D'Hondt method, such as the local elections in Hesse, Germany.

3 Local Elections in Baden-Württemberg

In some German states, local elections are notably complex, such as the city council elections in Baden-Württemberg. Voters can assign as many votes as seats in the city council (which can be up to 60 votes). Political parties can place the same number of candidates on their list. The ballot is a booklet consisting of multiple pages, with each page listing the candidates of one party. To cast a vote, one or more pages can be detached from the booklet and placed into a ballot envelope before being cast into the ballot box. Voters can either (i) allocate all their votes to a single political party, one vote to each of the candidates, which we refer to as *unchanged ballots*. *Unchanged ballots* can be cast by ripping out the page of the political party, placing it into the envelope and casting it into the ballot box, either without any markings or by marking every candidate of said party once, or (ii) voters can apply cumulating and/or panachaging, i.e., assigning up to three votes for selected candidates (cumulating) and/or distributing votes throughout candidates in different political parties (panachaging). We refer to cumulated and/or panachaged ballots as *changed ballots*.

After closing of the polling places on election day, the number of ballots is counted. This number is put into a protocol signed by all poll workers. All ballots, changed and unchanged, as well as the protocol are stored in a *vote suitcase* which is sealed and delivered to the electoral office.

On the day after the election the total tally is carried out at a central place – enabling voters to observe the process. In a first step, the electoral workers (in groups of three people) check the seals of each *vote suitcase* and separate *unchanged* and *changed ballots* for each suitcase. For each of these suitcases, the tallying of the *unchanged ballots*, which is done manually, is performed first. Note that each ballot is counted at least twice by two of the electoral workers. If these come to the same tally, it is entered into the software. If the numbers are different, the manual tallying is repeated until two electoral officials reach the same result. Note that like all the tallying processes, this process can be observed by any voter or election observers. Therefore, we consider the tally of the *unchanged ballots* to be **trusted**.

Afterwards, *changed ballots* are tallied by entering them into the software one by one. For entering the *changed ballots*, one of the three electoral workers reads out the votes of the ballot, another one enters the votes into the software, and the third manages the paper ballots (unpacking, passing along and storing). Electoral workers write an ID on each changed paper ballot they enter into the software in order to clearly assign them later to the corresponding entries from the software.

After all ballots are entered into the software (from all polling places) the total tally for the city is calculated by the software. Based on this tally the seat distribution is also calculated by the software and candidates are assigned to the seats granted to their party. For the distribution of seats *Sainte-Laguë* is used, in which the total number of votes of each party is divided by odd (uneven) numbers (/1, /3, /5, /7, ...) and the next seat is allocated to the party with the next biggest number. No minimum vote threshold is required for a party or candidate to be considered in this process.

The calculation of the seat distribution is also performed manually by the electoral workers based on the tally by the software to verify the accuracy of the distribution announced by the software. Afterwards, the preliminary results are published. Thus, while the computation of the distribution is audited, full trust is necessary in the software regarding the announced tally. This tally is the sum of the trusted tally of *unchanged ballots* and the tally of the *changed ballots*. Consequently, the tally of the *changed ballots* is considered **untrusted**.

4 General Idea to Audit

We assume that the audit software is independent from the tallying software, i.e., it also runs on a different, independent device. Furthermore, the process described here could be repeated using several different audit software applications from independent entities. In addition, we assume that the tally of the *unchanged ballots* is trusted. Therefore, we only audit the tally of the *changed ballots* provided by the software. Note that while the software outputs the overall election tally, the tally of the *changed ballots* can be computed by subtraction.

The input and the different steps of the audit process are shown in Fig. 1: In step (1), the number of votes for each party is computed (the tally resulting from the *changed ballots*). Afterwards, in step (2), the audit software would add the individual *changed ballots* and verify that the sum was properly computed. Step (3) outlines the computation of actual ballots to be audited, given the number of votes for *changed* and *unchanged ballots*, as well as the seat distribution, which is manually computed as part of the current tallying process. Finally, in step (4), *changed ballots* are randomly selected to compare the physical ones with their corresponding electronic record. Note that paper ballots and their corresponding entries in the software can be traced back due to the IDs written on the paper ballots. If errors are found they may be caused by software, hardware, or human error.

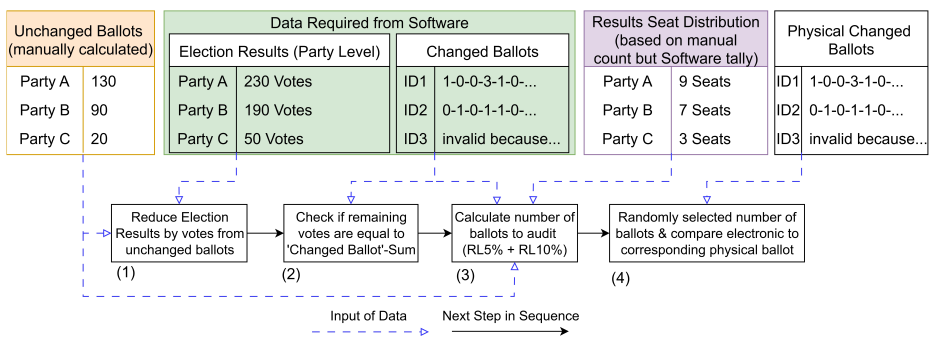

Fig. 1. General idea of the audit process

5 Assertion-based RLA

The SHANGRLA [9] framework verifies a reported election outcome by reducing the correctness of the underlying social choice function to a set of assertions, typically expressed as linear inequalities over the set of validly cast ballots. For use within SHANGRLA, these assertions must be represented in the form of assorters. Following the approach of Blom et al. [1], we first introduce the necessary notation, remaining consistent with their definitions. We then define a partition of the ballot tally into *trusted* and *untrusted* components. Afterwards, we derive the assorters required to conduct risk-limiting audits for the Sainte-Laguë city council elections in Baden-Württemberg.

5.1 Preliminaries

Notation (following [1]) Let S be the total number of seats. Let T_e be the tally of entity e (usually a party or coalition). A highest average method is parameterised by a set of *divisors* $d(1), \ldots, d(S)$. For Sainte-Laguë, the divisors are $d(i) = 2i-1$, for $i = 1, 2, \ldots S$. Let m be the maximum number of votes a voter may assign to a single party.[4]

The S seats are awarded to the S highest *quotients*. That is, e gets s seats if $T_e/d(s)$ is among the S highest such values. The *Winning Set* $\mathcal{W}$ is

$$\mathcal{W} = \{(e, s) : T_e/d(s) \text{ is one of the } S \text{ largest}\}.$$

Let W_e be the number of seats won and L_e the number of the first seat lost by entity e. That is:

$$W_e = \max\{s : (e, s) \in \mathcal{W}\}; \perp \text{ if } e \text{ has no winners.}$$
$$L_e = \min\{s : (e, s) \notin \mathcal{W}\}; \perp \text{ if all } e\text{'s candidates won seats.}$$

The election outcome is the assignment of seats to parties. It is correct if and only if for all parties A and B the following holds:

$$\forall A \ s.t. \ W_A \neq \perp, \forall B \neq A \ s.t. \ L_B \neq \perp, \quad T_A/d(W_A) - T_B/d(L_B) > 0. \tag{1}$$

The audit will test all of these comparisons, which are updated whenever a ballot is sampled.

The rest of the paper considers how to generate one assertion to compare two parties, but of course a real audit would generate all of them, for all necessary comparisons.

5.2 Trusted–Untrusted Tally Decomposition

In analogy with the classification of ballots as *changed* and *unchanged* in Baden-Württemberg city council elections, we introduce a corresponding classification of ballots into *trusted* and *untrusted* categories. Accordingly, we decompose the overall tally into a *trusted* (reliable, denoted by R) and *untrusted* (unreliable, denoted by U) component. Hence, each tally T_e is partitioned into a reliable and an unreliable component, so that

$$T_e = R_e + U_e,$$

where R_e is *trusted* and U_e is *untrusted*. We rewrite Equation 1 as

$$(R_A + U_A)/d(W_A) - (R_B + U_B)/d(L_B) > 0, \tag{2}$$

where, crucially, only U_A and U_B are subject to audit.

[4] This is typically the number of available seats. In Baden-Württemberg, $m = S = 48$.

5.3 Assorters for Trusting Partial Tallies

Let $\mathcal{U}$ be the set of untrusted ballots, and $|\mathcal{U}|$ be its size, i.e. the total number of untrusted ballots. We can now follow a similar derivation to [1], Section 3.1, to derive a SHANGRLA assorter.

Let

$$\Delta = (R_A d(L_B)/d(W_A) - R_B)/|\mathcal{U}|.$$

Then the assorter is:

$$h^u{}_{A,B}(b) = \frac{\frac{b_A d(L_B)}{d(W_A)} - b_B + \Delta}{2(m - \Delta)} + \frac{1}{2}. \quad (3)$$

A derivation and validity checks are in Appendix A.1. This assorter is appropriate for ballot-polling audits, but these can be quite inefficient. Fortunately, for the *untrusted ballots*, we have cast vote records against which we can compare the individual ballot. This allows us to conduct a ballot-level comparison audit.

5.4 Assorter for Ballot-level Comparison

We now derive a ballot-level comparison assorter, which we refer to as h^{comp}. h^{comp} will compare the contents of the actual sampled ballot with the cast vote record. If they are the same, or if the error understates the apparent winner's margin, then the assorter should have a value $> 1/2$. Conversely, if the cast vote record overstates the apparent winner's margin (for example, if the cast vote record contains 48 votes for the winning list while the actual ballot contains 48 votes for the loser), then the assorter should have a lower value (in this example, zero). We should design the assorter so that it has a mean above 1/2 if and only if the assertion is true.

When c is a cast vote record and v is the corresponding (manually audited) vote, we can calculate the discrepancy by applying the ballot-polling assorter to both values. An appropriate ballot-level comparison assorter for comparing cast vote record c with (manually audited) vote v is then:

$$h^{\text{comp}}_{A,B}(c, v) = \frac{u_{A,B} - (h^u{}_{A,B}(c) - h^u_{A,B}(v))}{2u_{A,B} - \nu_u(CVRs)}, \quad (4)$$

where $u_{A,B}$ is the upper bound of $h^u_{A,B}$ and $\nu_u(CVRs)$ is the apparent assorter mean. Some validity checks and other calculations, including $u_{A,B}$ and $\nu_u(CVRs)$, are given in Appendix A.2.

6 Sample Size Estimates for Baden-Württemberg

We estimate the sample size by finding the estimated sample sizes of all the assertions (that is, for all pairs of parties) and reporting the largest. This is a fair (but not perfect) approximation of the overall estimated sample size. The estimates use real election data from Karlsruhe (2019 and 2024), along with an artificial contest with a very close margin. For these initial estimates, we assume that all parties have the same rate of *changed* vs. *unchanged ballots*. However, in practice they can vary substantially, which may affect sample sizes.

Table 1 shows estimated sample sizes for a risk limit of 5% assuming zero errors, for Karlsruhe local council elections 2019 and 2024. Table 2 shows the same comparison, assuming a 0.1% error rate. All estimates are for ballot-level comparison audits using the assorter in Equation 4. The code and sample data are available on GitHub.

Trusting the manually counted subset does reduce the sample size somewhat, but sample size estimates for local elections in Baden-Württemberg remain high even with the efficiency gains proposed in this paper.

Election	**Max-sample comparison**		**Sample size estimate**	
	Winner	**Loser**	**100% unreliable**	**60% unreliable**
Karlsruhe 2024	CDU	Die PARTEI	12072	7377
Karlsruhe 2019	FDP	AfD	9425	5736

Table 1. Comparison of sample size estimates, assuming a zero error rate, comparing the case when all are unreliable vs. the case when 0.6 are unreliable, which is average for Baden-Württemberg.

Election	**Max-sample comparison**		**Sample size estimate**	
	Winner	**Loser**	**100% unreliable**	**60% unreliable**
Karlsruhe 2024	CDU	GRÜNE	135496	21609
Karlsruhe 2019	FDP	AfD	134735	33405

Table 2. Comparison of sample size estimates, assuming a 0.1% error rate, comparing the case when all are unreliable vs. the case when 0.6 are unreliable, which is average for Baden-Württemberg.

7 Conclusion and Future Work

This work extends the SHANGRLA framework to Baden-Würtemberg city council elections, auditing the allocation of seats to parties. It introduces a trust-based decomoposition of the tally motivated by the treatment of *changed* and *unchanged ballots*. Empirical estimation based on real data indicates a significant reduction in the number of samples required to audit the election outcome, though sample sizes remain high.

References

1. Blom, M., Budurushi, J., Rivest, R.L., Stark, P.B., Stuckey, P.J., Teague, V., Vukcevic, D.: Assertion-based approaches to auditing complex elections, with application to party-list proportional elections. In: Electronic Voting (2021)
2. Blom, M., Ek, A., Stuckey, P.J., Teague, V., Vukcevic, D.: 3+ seat risk-limiting audits for single transferable vote elections. In: FC International Workshops (2025)
3. Blom, M., Stark, P.B., Stuckey, P.J., Teague, V., Vukcevic, D.: Auditing hamiltonian elections. In: FC International Workshops (2021)
4. Blom, M., Stuckey, P.J., Teague, V., Vukcevic, D.: A first approach to risk-limiting audits for single transferable vote elections. In: FC International Workshops (2022)
5. Blom, M., Stuckey, P.J., Teague, V., Vukcevic, D.: Risk-limiting audits for condorcet elections. In: FC International Workshops (2023)
6. Blom, M., Stuckey, P.J., Teague, V.J.: Ballot-polling risk limiting audits for irv elections. In: Electronic Voting (2018)
7. Lindeman, M., Stark, P.B.: A gentle introduction to risk-limiting audits. IEEE Security & Privacy **10**(5), 42–49 (2012)
8. Stark, P.B.: Risk-limiting vote-tabulation audits: The importance of cluster size. Chance **23**(3), 9–12 (2010)
9. Stark, P.B.: Sets of half-average nulls generate risk-limiting audits: Shangrla. In: FC International Workshops (2020)
10. Stark, P.B., Teague, V.: Verifiable European elections: Risk-limiting audits for D'Hondt and its relatives. USENIX (JETS) **1**, 18–39 (2014)

A Appendix: Assorter derivations and validity checks

A.1 Deriving the ballot-polling assorter

Derivation Following [1], we rearrange Equation 2 to derive the proto-assorter:

$$g(b) = b_A/d(W_A) - b_B/d(L_B) + [R_A/d(W_A) - R_B/d(L_B)]/|\mathcal{U}|,$$

where b_e is the number of votes assigned to entity e on ballot b.

To see why this proto-assorter is correct, consider what happens when we sum it over all untrusted ballots.

$$\begin{aligned}\sum_{b\in\mathcal{U}} g(b) &= \sum_{b\in\mathcal{U}} b_A/d(W_A) - \sum_{b\in\mathcal{U}} b_B/d(L_B) + |\mathcal{U}|[R_A/d(W_A) - R_B/d(L_B)]/|\mathcal{U}| \\ &= U_A/d(W_A) - U_B/d(L_B) + R_A/d(W_A) - R_B/d(L_B)\end{aligned}$$

which is exactly the left side of Equation 2.

The proto-assorter g has a lower bound of $\frac{\Delta - m}{d(L_B)}$, reached when m votes are given to B (and none to A). Substituting this into Equation 2 of [1] gives us the assorter h^u in Equation 3.

Validity checks The reader can easily check that when b contains m votes for the loser (B), ${h^u}_{A,B}(b) = 0$. A tie occurs when $(R_A + U_A)/d(W_A) = (R_B + U_B)/d(L_B)$. Then, the average of h^u over all untrusted ballots is $\frac{U_A d(L_B)/d(W_A) - U_B + \Delta|\mathcal{U}|}{2|\mathcal{U}|(m-\Delta)} + \frac{1}{2} = \frac{1}{2}$.

When there are no trusted ballots, $\Delta = 0$ and so h^u is the same as the assorter derived in Blom et al [1], Section 5.3.

Other calculations useful for sample size estimation The upper bound of ${h^u}_{A,B}$ is obtained when the ballot contains the maximum votes for A and none for B, that is when $b_A = m$ and $b_B = 0$. Then

$$u_{A,B} = \frac{\frac{md(L_B)}{d(W_A)} + \Delta}{2(m-\Delta)} + \frac{1}{2} = \frac{m(d(L_B)/d(W_A) + 1)}{2(m-\Delta)} \tag{5}$$

It helps to compute the mean of ${h^u}_{A,B}$ over all untrusted ballots. The *apparent mean* is a function of the CVRs.

$$\begin{aligned}
\overline{{h^u}_{A,B}}(CVRs) &= \frac{1}{|\mathcal{U}|}\sum_{b\in\mathcal{U}} h^u_{A,B}(b) && (6)\\
&= \frac{1}{|\mathcal{U}|}\sum_{b\in\mathcal{U}}\left(\frac{\frac{b_A d(L_B)}{d(W_A)} - b_B}{2(m-\Delta)} + \frac{\Delta}{2(m-\Delta)} + \frac{1}{2}\right) && (7)\\
&= \frac{\frac{U_A d(L_B)}{d(W_A)} - U_B}{2|\mathcal{U}|(m-\Delta)} + \frac{\Delta}{2(m-\Delta)} + \frac{1}{2} && (8)
\end{aligned}$$

This then allows us to compute the apparent *ballot-polling assorter margin*:

$$\nu_u(CVRs) = 2\overline{{h^u}_{A,B}(CVRs)} - 1 \tag{9}$$

A.2 Other calculations for the ballot-level comparison assorter

The ballot-level comparison assorter in Equation 4 clearly has a lower bound of zero (when c causes h^u to reach its upper bound and v makes it zero), and an upper bound of $2u_{A,B}/(2u_{A,B} - \nu_u)$ (in the opposite case).

Validity checks To see that the mean is $1/2$ when the election is tied, let V_e be the true sub-tally of entity e over the untrusted (real paper) ballots. We are assuming that R_e is the same for both cast vote records (CVRs) and real ballots. A tie occurs when $(V_A + R_A)/d(W_A) = (V_B + R_B)/d(L_B)$, that is, when $V_A d(L_B)/d(W_A) - V_B + \Delta|\mathcal{U}| = 0$. Now compute the mean of h^{comp}, over (manually inspected) ballot papers in $\mathcal{U}$, for the fixed set of CVRs we used to derive the assorter. Let $CVR(b)$ be the CVR corresponding to ballot b. Then the mean is

$$
\begin{aligned}
\overline{h^{\text{comp}}_{A,B}} &= \frac{1}{|\mathcal{U}|}\sum_{b\in\mathcal{U}} \frac{u_{A,B} - (h^u{}_{A,B}(CVR(b)) - h^u_{A,B}(b))}{2u_{A,B} - \nu_u(CVRs)} \\
&= \frac{u_{A,B} - \overline{h^u_{A,B}(CVRs)} + \frac{1}{|\mathcal{U}|}\sum_{b\in\mathcal{U}} h^u_{A,B}(b)}{2u_{A,B} - \nu_u(CVRs)} \text{ (by Equation 6)} \\
&= \frac{u_{A,B} - (\nu_u(CVRs)+1)/2 + \frac{1}{2(m-\Delta)}(\frac{V_A d(L_B)/d(W_A) - V_B + \Delta}{|\mathcal{U}|}) + 1/2}{2u_{A,B} - \nu_u(CVRs)} \text{ (by Equation 9 and Equation 3)} \\
&= \frac{u_{A,B} - (\nu_u(CVRs)+1)/2 + 1/2}{2u_{A,B} - \nu_u(CVRs)} \text{ (by definition of tie)} \\
&= 1/2.
\end{aligned}
$$

Overstatements When c has m votes for A and v is blank,

$$
h^{\text{comp}}_{A,B}(c,v) = \frac{u_{A,B} - ((\frac{md(L_B)/d(W_A) - 0 + \Delta}{2(m-\Delta)} + 1/2) - (\frac{\Delta}{2(m-\Delta)} + 1/2))}{2u_{A,B} - \nu_u(\text{CVRs})} \tag{10}
$$

$$
= \frac{m}{2(m-\Delta)(2u_{A,B} - \nu_u(\text{CVRs}))} \text{ (by Equation 5)} \tag{11}
$$

This is used for sample-size estimation in the presence of overstatement errors.