Election Security with Risk-Limiting Audits

Oliver Broadrick[1] Poorvi L. Vora[1] Filip Zagórski[23]

[1]Department of Computer Science, The George Washington University [2]Wroclaw University of Science and Technology [3]Votifica

Original Contributions

Risk-Limiting Audits (RLAs) are rigorous election audits performed in rounds. RLAs EoR BRAVO and SO BRAVO [?] rely on the Sequential Probability Ratio Test (SPRT), using a ratio of sample probabilities conditioned on alternative and null hypotheses. More recent RLA MINERVA [?] uses a ratio of the tails of the same distributions, and requires half as many ballots as EoR BRAVO in a first round with probability of stopping 0.9. Understanding audit behavior for later rounds has no easy analytical approach. MINERVA was used in a pilot RLA in Montgomery County, Ohio in 2020, and is recommended by the largest US voting organizations: Verified Voting, the Brennan Center, and Common Cause.

- We provide simulations [?] of EoR BRAVO, SO BRAVO, and MINERVA that help us understand their behavior across multiple rounds and with lower first round stopping probability; this information can be used to advise election officials.

Unlike EoR and SO BRAVO, MINERVA is proven to be risk-limiting only when round sizes are predetermined, meaning round sizes cannot be chosen based on previous samples, flexibility that may have meaningful impact on workload.

- We introduce PROVIDENCE, an RLA which has the efficient tail ratio approach of MINERVA but is resistant to an adversary who can pick future round sizes with knowledge of previous samples.
- PROVIDENCE was named after the Rhode Island city where it was used for a public pilot audit in February 2022.

Results

- BRAVO audits are more conservative than MINERVA, which stops with fewer ballots, for both first round stopping probabilities
- Advantage of using MINERVA decreases considerably for smaller first round stopping probability
- Proofs that PROVIDENCE is risk-limiting and resistant to an adversary who can pick future round sizes with knowledge of previous samples

Risk-Limiting Audits

- A risk-limiting audit (RLA) draws paper ballots in rounds, stopping if a rigorous statistical criterion is satisfied, or proceeding to a full hand count. If the announced outcome of the election is erroneous, an RLA will detect the error with high, predetermined minimum probability.
- An audit $\mathcal{A}$ takes a sample of ballots X as input and gives as output either (1) $Correct$: the audit is complete, or (2) $Uncertain$: continue the audit.
- The maximum risk R of audit $\mathcal{A}$ with sample $X \in \{0,1\}^*$ drawn from the true underlying distribution of ballots is

$$R(\mathcal{A}) = \Pr[\mathcal{A}(X) = Correct \mid H_0],$$

where H_0 is the null hypothesis: the true underlying election is a tie.
- An audit $\mathcal{A}$ is a Risk-Limiting Audit with risk limit α iff

$$R(\mathcal{A}) \leq \alpha.$$

Simulations

- Software: R2B2 library [?] which has Python implementations of several RLAs as well as a simulator
- Risk limit: 10%
- Margins: 2020 US Presidential election statewide results, limiting ourselves to pairwise margins for the two main candidates of 0.05 or larger
- Trials per state: $10,000 = 10^4$ audits assuming the underlying election is as announced, and $10,000 = 10^4$ audits assuming the underlying election is a tie
- Round schedules: EoR BRAVO and SO BRAVO round sizes are chosen to achieve the same probability of stopping in each round given preceding samples. MINERVA first round sizes are chosen to achieve some probability of stopping and subsequent round sizes are given by multiplying the previous round size by 1.5 or by 1.0.

Risk

- All audits have estimated risks below the risk limit
- EoR BRAVO falls an order of magnitude less than the others, unnecessarily conservative

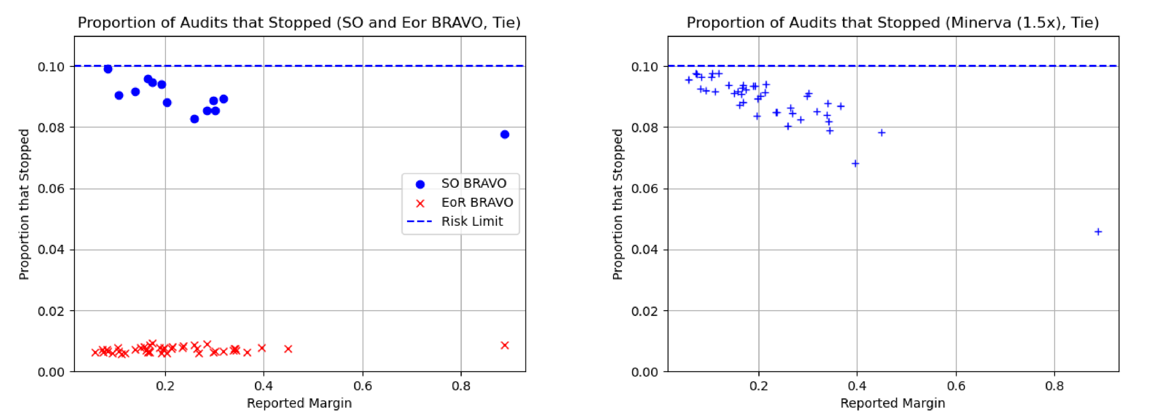

Figure 1. These plots show for $\chi_1 = 0.9$ the fraction of audits (EoR BRAVO (left), SO BRAVO (left), and MINERVA with multiplier 1.5 (right)) that stopped in any of the 5 rounds when the underlying election was a tie. (For SO BRAVO audits we show the 13 states for which simulations are complete.)

Stopping Probability

- Conditional stopping probability, χ_j, is the probability that an audit will stop in round j given that it did not stop earlier, and in simulations χ_j is estimated by the proportion of audits that stop in a round to those that "entered" that round i.e. those that did not stop before round j.

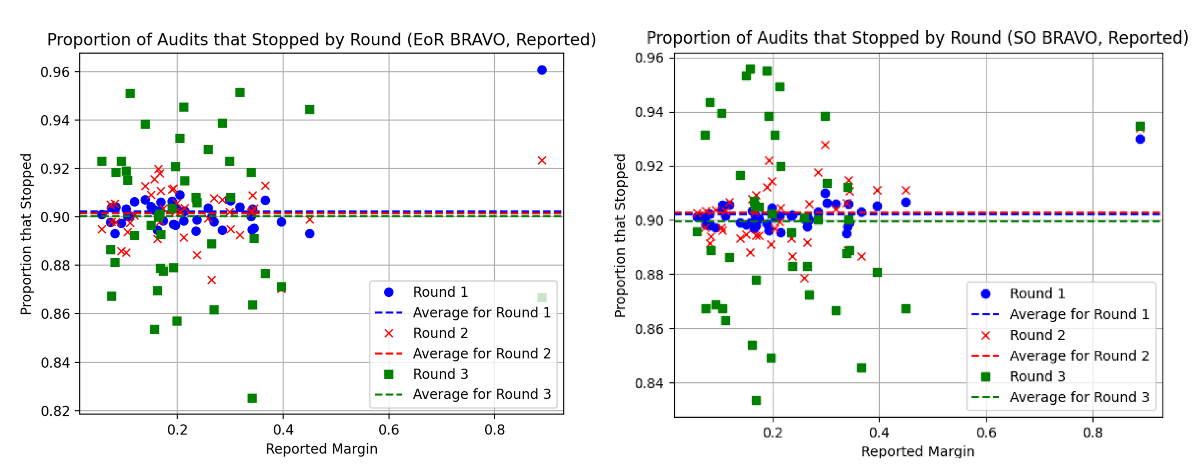

Figure 2. These plots show, for each state margin, when the underlying election is as announced, the number of audits that stopped in the j^{th} round, as a fraction of all audits which had not yet stopped before the j^{th} round for $j = 1, 2, 3$ and $\chi_j = 0.9$. The left shows EoR BRAVO and the right shows SO BRAVO.

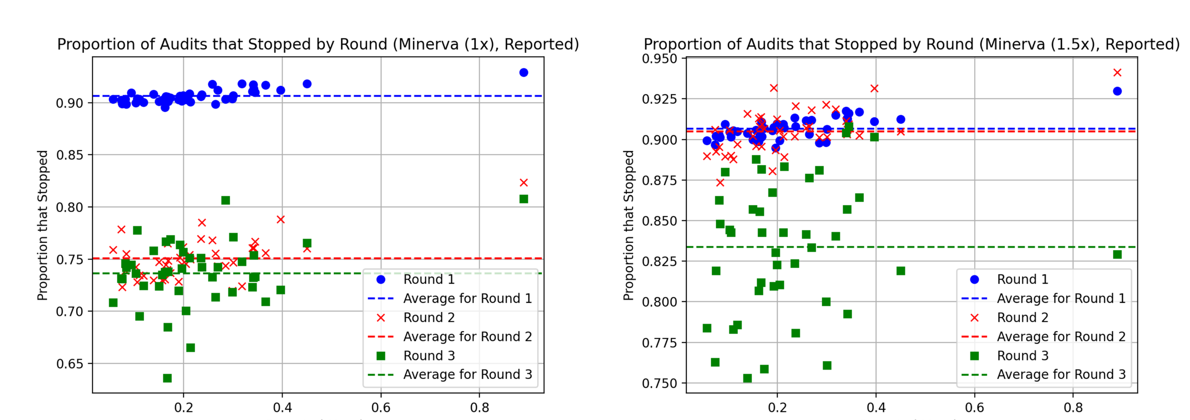

Figure 3. These plots show, for each state margin, when the underlying election is as announced, the number of MINERVA audits that stopped in the j^{th} round, as a fraction of all MINERVA audits which had not yet stopped before the j^{th} round for $j = 1, 2, 3$. The left shows MINERVA with the round schedule obtained by $\chi_1 = 0.9$ and round size multiple 1.5.

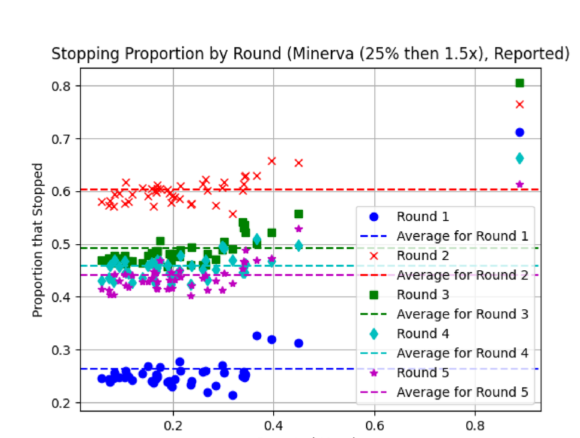

Figure 4. This plot shows, for each state margin, when the underlying election is as announced, the number of MINERVA audits that stopped in the j^{th} round, as a fraction of all MINERVA audits which had not yet stopped before the j^{th} round for $j = 1, 2, 3$, round size multiple of 1.5 and $\chi_1 = 0.25$.

Number of Ballots

- The number of ballots sampled is one crude measure of the workload of an audit. To keep the costs of RLAs low, audits should be designed to stop with as few ballots as possible.
- The following plots show the probability of stopping as a function of the average number of ballots sampled by round in our simulations.

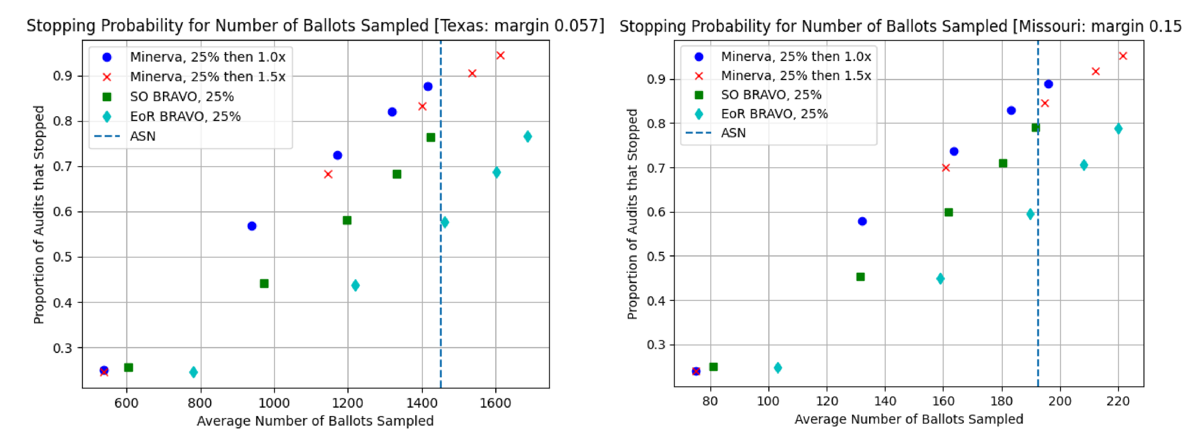

Figure 5. These plots show the cumulative fraction of audits that stopped as a function of average number of sampled ballots for all four audits we studied, for the states of Texas with margin $.057$ (left) and Missouri with margin 0.157 (right), both with first round stopping probability $\chi_1 = 0.25$.

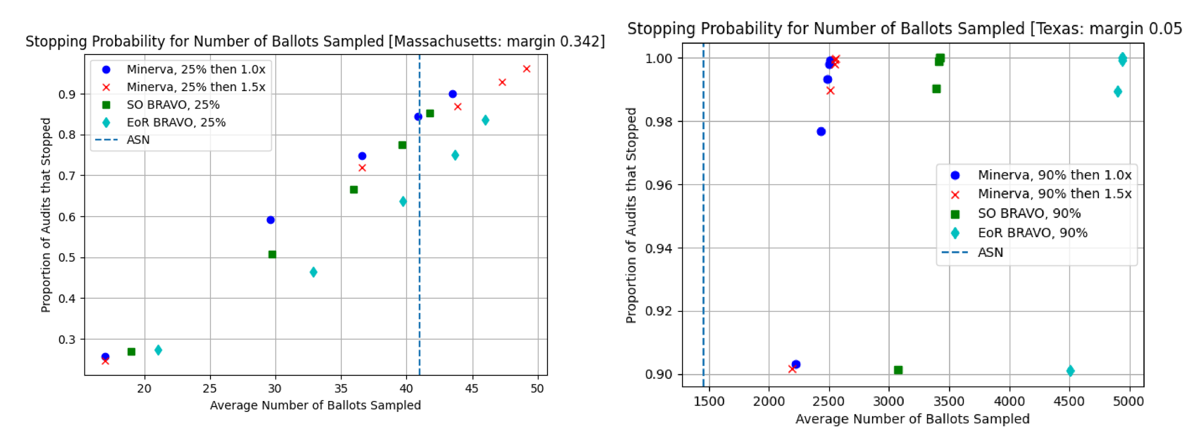

Figure 6. These plots show the cumulative fraction of audits that stopped as a function of average number of sampled ballots for all four audits we studied, for the states Massachusetts (left) with margin 0.342 and $\chi_1 = 0.25$ and Texas (right) with margin 0.057 and $\chi_1 = 0.9$.

For $\chi_1 = 0.25$, the ratio of first round size of EoR BRAVO to MINERVA is 1.45, 1.37, 1.23 for states Texas, Missouri and Massachusetts, and margins 0.057, 0.157 and 0.342 respectively. This may be compared to 2.03, 1.99 and 1.8 respectively for $\chi_1 = 0.9$. Similarly, for $\chi_1 = 0.25$, the ratio of first round size of SO BRAVO to MINERVA is 1.13, 1.08, 1.12 for states Texas, Missouri and Massachusetts, and margins 0.057, 0.157 and 0.342 respectively. This may be compared to 1.38, 1.38 and 1.30 respectively for $\chi_1 = 0.9$.

Providence

- The efficiency of MINERVA is great, but it lacks the flexibility of BRAVO in choosing round sizes based on previous samples.
- PROVIDENCE is our novel RLA which has the efficiency of MINERVA and the flexibility of BRAVO.
- For alternative hypothesis H_a that the election is truly as announced and null hypothesis H_0 that the true election is a tie, BRAVO has the stopping condition that for k cumulative ballots for the winner and n cumulative sampled ballots,

$$\sigma(k, n, p_a, p_0) \triangleq \frac{\Pr[K = k \mid H_a, n]}{\Pr[K = k \mid H_0, n]} \geq \frac{1}{\alpha}.$$

- MINERVA has the stopping condition that in round j with cumulative winner ballots k_j and round sizes $\bar{n}_j = n_1, n_2, \ldots, n_j$

$$\tau_j(k_j, \bar{n}_j, p_a, p_0) \triangleq \frac{\Pr[K_j \geq k_j \wedge \mathcal{A}_{i<j}(X) \neq Correct \mid H_a, \bar{n}_j]}{\Pr[K_j \geq k_j \wedge \mathcal{A}_{i<j}(X) \neq Correct \mid H_0, \bar{n}_j]} \geq \frac{1}{\alpha}.$$

Testing this stopping condition requires computationaly expensive convolutions.
- The PROVIDENCE stopping condition uses ideas from both BRAVO and MINERVA and requires no convolution to test:

$$\omega_j(k_{j-1}, k_j, n_{j-1}, n_j, p_a, p_0) \triangleq \sigma(k_{j-1}, n_{j-1}, p_a, p_0) \cdot \tau_1(k_j, n_j, p_a, p_0) \geq \frac{1}{\alpha}.$$

Providence

Resistance against an adversary

If adversary $\mathcal{A}$ has knowledge of previous samples and chooses round size n_j in round j, PROVIDENCE is still risk-limiting. Proof of this and the RLA property of PROVIDENCE are currently being revised for submission.

Providence pilot

In February 2022, The Rhode Island Board of Elections hosted a public pilot PROVIDENCE RLA which passed in the first round.

Figure 7. Professor Vora tosses her random-seed-generating 10-sided die, in a roll she dedicated to election officials everywhere. The standing people from left to right are (1) Miguel Nunez, Rhode Island Deputy Director of Elections, (2) Professor Vora, and (3) Mark Lindeman, Verified Voting Director.

Simulations

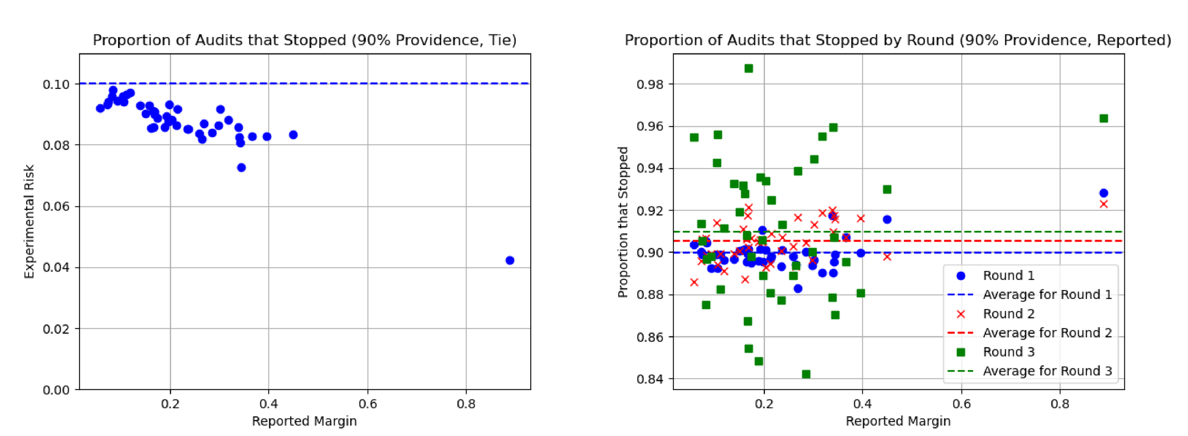

Figure 8. The left plot shows the fraction of PROVIDENCE audits that stopped in any of the 5 rounds when the underlying election was a tie where each round size is selected to achieve $\chi_j = 0.9$. The right plot shows the fraction of PROVIDENCE audits that stopped in any of the 5 rounds when the underlying election was a tie for rounds $j = 1, 2, 3$ and $\chi_j = 0.9$.

References

[1] Oliver Broadrick, Sarah Morin, Grant McClearn, Neal McBurnett, Poorvi L. Vora, and Filip Zagórski. Simulations of ballot polling risk-limiting audits. In *Seventh Workshop on Advances in Secure Electronic Voting, in Association with Financial Crypto*, 2022.

[2] Mark Lindeman, Philip B Stark, and Vincent S Yates. BRAVO: Ballot-polling risk-limiting audits to verify outcomes. In *EVT/WOTE*, 2012.

[3] Sarah Morin and Grant McClearn. The R2B2 (Round-by-Round, Ballot-by-Ballot) library, https://github.com/gwexploratoryaudits/r2b2.

[4] Filip Zagórski, Grant McClearn, Sarah Morin, Neal McBurnett, and Poorvi L. Vora. Minerva– an efficient risk-limiting ballot polling audit. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 3059–3076. USENIX Association, August 2021.

odbroadrick@gmail.com