Differentially Private E-Values

Daniel Csillag
FGV EMAp

Diego Mesquita
FGV EMAp

Abstract

E-values have gained prominence as flexible tools for statistical inference and risk control, enabling anytime- and post-hoc-valid procedures under minimal assumptions. However, many real-world applications fundamentally rely on sensitive data, which can be leaked through e-values. To ensure their safe release, we propose a general framework to transform non-private e-values into differentially private ones. Towards this end, we develop a novel biased multiplicative noise mechanism that ensures our e-values remain statistically valid. We show that our differentially private e-values attain strong statistical power, and are asymptotically as powerful as their non-private counterparts. Experiments across online risk monitoring, private healthcare, and conformal e-prediction demonstrate our approach's effectiveness and illustrate its broad applicability.

1 INTRODUCTION

E-values have emerged as versatile tools for statistical inference and risk control, offering anytime- and post-hoc-valid guarantees under minimal assumptions. They are the backbone of a growing body of methods for continuous risk monitoring [Podkopaev and Ramdas, 2021], change-point detection [Shin et al., 2022, Shekhar and Ramdas, 2023], test-time adaptation [Bar et al., 2024], uncertainty quantification [Balinsky and Balinsky, 2024, Gauthier et al., 2025b], and interpretability [Teneggi and Sulam, 2024], among other inferential tasks. However, many of the domains where these methods are most impactful involve sensitive data. Applying existing e-value-based procedures directly to such data can compromise individual privacy, since their guarantees ensure statistical validity but not protection against information leakage.

Differential privacy provides a principled framework for addressing such challenges. However, standard differential privacy mechanisms are insufficient for our setting: though privacy is attained, the resulting quantities are generally not statistically valid e-values. In this paper, we resolve this by introducing novel biased multiplicative noise mechanisms, which ensure privacy while retaining statistical validity. In this way, we are able to to convert any non-private e-value into a differentially private one. To the best of our knowledge, this is the first instance of valid e-values that satisfy differential privacy.

Beyond establishing validity, we also exactly quantify the statistical power of our differentially private e-values, showing that it differs from that of the non-private e-value only by a factor that decreases with the number of samples, related to the mechanism's bias. Crucially, as the number of samples goes towards infinity, we recover the power of the non-private test. We also show that our private e-values inherit many of the usual properties of e-values, with no loss of privacy.

Our framework is general: the resulting differentially private e-values retain the compositional and optional continuation properties of standard e-values while ensuring privacy, and can be applied not only to hypothesis testing but also to confidence intervals and general e-value-based inference procedures. We demonstrate the effectiveness of our approach through experiments in three real-world tasks: e-value based confidence intervals for the prevalence of diabetes, continuous monitoring of the risk of a deployed model, and e-conformal prediction for detection of phishing attacks.

Our contributions

- We propose the first general framework to privatize e-values, simultaneously satisfying differential privacy and statistical validity. Our framework works by introducing novel biased multiplicative noise mechanisms, which we show to be necessary

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300. Copyright 2026 by the author(s).

to ensure validity of the resulting e-values.

- We derive an exact characterization of the statistical power of our differentially private e-values, in terms of their growth rates. As the number of samples grows, so does the growth rates, matching that of the non-private e-value at the limit.
- We prove that beyond basic validity, our differentially private e-values satisfy many compositional properties of e-values while preserving privacy, including optional continuation, e-to-p conversion and averaging (the latter under some restrictions); this enables the seamless use of many existing e-value-based procedures.
- We demonstrate our framework on three real-world settings spanning private healthcare, online risk monitoring atop private data, and private predictive modelling. For each instance we derive effective procedures that are readily applicable by practitioners.

Related work. The broader idea of statistical inference with differential privacy has already been the target of much attention. For example, [Rogers and Kifer, 2017, Swanberg et al., 2019, Ding et al., 2018, Gaboardi et al., 2018, Campbell et al., 2017, Couch et al., 2019] propose differentially private versions of classical tests, and [Karwa and Vadhan, 2017, Du et al., 2020, Ferrando et al., 2020] propose private confidence intervals. More recently, [Kazan et al., 2023, Pena and Barrientos, 2022] proposed general frameworks from which differentially private hypothesis tests can be obtained. Also of note are the prior works of [Dwork et al., 2018, Zhang et al., 2020], which also derive differentially private statistical inference procedures. However, these are all either inapplicable to e-values or discard essential properties of the e-values — e.g., post-hoc validity, optional continuation, and advantages for multiple testing. Our approach, in contrast, naturally benefits from all the usual properties of e-values, while also being significantly more data efficient (cf. Appendix C.1).

2 BACKGROUND

2.1 Differential privacy

Differential privacy (DP) is a framework for controlling the privacy loss incurred when releasing information about a dataset. Introduced by [Dwork et al., 2006], differential privacy works by incorporating controlled noise to the processing of the dataset, so as to guarantee that the inclusion or exclusion of any single individual's data cannot be inferred from the outputs. The first fully formalized notion of DP was ϵ-differential privacy, which was later generalized to (ϵ, δ)-differential privacy.

Definition 2.1 ((ϵ, δ)-differential privacy)**.** *A randomized algorithm $\mathcal{A}(\cdot)$ satisfies (ϵ, δ)-differential privacy if, for all (fixed) neighboring datasets D and D' (i.e., differing by one record), for any set A,*

$$\mathbb{P}[\mathcal{A}(D) \in A] \leq e^{\epsilon}\mathbb{P}[\mathcal{A}(D') \in A] + \delta.$$

If $\delta = 0$, algorithm $\mathcal{A}(\cdot)$ satisfies ϵ-differential privacy.

Though influential, (ϵ, δ)-differential privacy is known to struggle with the composition of many differentially private procedures. To improve on this, [Mironov, 2017] introduced (α, ϵ)-Rényi differential privacy, which generalizes ϵ-differential privacy:

Definition 2.2 ((α, ϵ)-Rényi differential privacy)**.** *A randomized algorithm $\mathcal{A}(\cdot)$ satisfies (α, ϵ)-Rényi differential privacy for $\alpha > 1$ if, for all (fixed) neighboring datasets D and D' (i.e., differing by one record),*

$$D_\alpha\big(\mathcal{A}(D) \,\|\, \mathcal{A}(D')\big) \coloneqq \frac{1}{\alpha - 1} \log \mathbb{E}_{z\sim\mathcal{A}(D')}\left[\left(\frac{\mathrm{d}\mathcal{A}(D)}{\mathrm{d}\mathcal{A}(D')}(z)\right)^{\alpha}\right] \leq \epsilon,$$

where $D_\alpha(P \,\|\, Q)$ is the α-Rényi divergence between distributions P and Q.

It is worth mentioning that (α, ϵ)-Rényi differential privacy is closely related to classical (ϵ, δ)-differential privacy. When $\alpha \to \infty$ we recover ϵ-differential privacy [Mironov, 2017], and satisfying (α, ϵ)-Rényi differential privacy also implies (ϵ', δ)-differential privacy with $\epsilon' = \epsilon + \log(1/\delta)/(\alpha - 1)$, for any $\delta > 0$.

Crucially, differential privacy benefits from two key properties: composition and post-processing. Composition states that differentially private algorithms can be combined into new differentially private algorithms, whereas post-processing states that any post-processing of a differentially private algorithm retains its privacy.

Proposition 2.3 (Composition of (α, ϵ)-Rényi differential privacy)**.** *For any $\alpha > 1$, let algorithms $\mathcal{A}_1(\cdot)$ and $\mathcal{A}_2(\cdot)$ be (α, ϵ_1)- and (α, ϵ_2)-Rényi differentially private. Then the algorithm $\mathcal{A}(\cdot) = (\mathcal{A}_1(\cdot), \mathcal{A}_2(\cdot))$ is $(\alpha, \epsilon_1 + \epsilon_2)$-Rényi differentially private.*

Proposition 2.4 (Post-processing of (α, ϵ)-Rényi differential privacy)**.** *For any $\alpha > 1$, let $\mathcal{A}(\cdot)$ be a (α, ϵ)-Rényi differentially private algorithm. Then, for any post-processing operation f, the algorithm $\mathcal{A}_f(\cdot) = f(\mathcal{A}(\cdot))$ is also (α, ϵ)-Rényi differentially private.*

Throughout the paper, we will refer primarily to (α, ϵ)-Rényi differential privacy. Nevertheless, our framework is just as applicable to standard (ϵ, δ)-differential privacy; see Appendix B.1.

2.2 E-values

Consider the problem of testing a null hypothesis H_0 with data $[Y_1, Y_2, \ldots, Y_n] =: D$. To this end, e-values serve as alternatives to the classic p-values as measures of evidence against the null. Formally, an e-value $E(D)$ for the null hypothesis H_0 is a nonnegative random variable whose expectation is at most one under the null.

Definition 2.5 (E-value)**.** *A nonnegative real random variable $E(D)$ is an e-value for a null hypothesis H_0 if $\mathbb{E}[E(D)] \leq 1$ under H_0.*

A random variable is commonly said to be a "valid" e-value when it satisfies Definition 2.5.

Any e-value can be converted to a p-value by simply taking its reciprocal, and any p-value can be converted to an e-value by a process termed calibration [Vovk and Wang, 2019], albeit at a slight loss of power. In a sense, e-values can be seen as p-values with richer structure [Koning, 2023], in particular satisfying post-hoc validity, optional continuation, and merging through averaging.

Proposition 2.6 (E-to-p conversion; post-hoc validity)**.** *If $E(D)$ is an e-value for a null hypothesis H_0, then $1/E(D)$ is a p-value for H_0.*[1] *Moreover, it is a* post-hoc valid *p-value [Koning, 2023], i.e., a p-value that allows for an arbitrarily data dependent significance level α, and all post-hoc valid p-values can be written as the reciprocal of some e-value.*

Proposition 2.7 (Optional continuation)**.** *If $E_1(D_1)$ and $E_2(D_2)$ are e-values for a null hypothesis H_0 over independent*[2] *datasets D_1 and D_2, then $E_1(D_1) \cdot E_2(D_2)$ is also an e-value for H_0.*

Proposition 2.8 (Averaging)**.** *If $E_1(D_1)$ and $E_2(D_2)$ are e-values for a null hypothesis H_0, then all convex combinations $\eta E_1(D_1) + (1-\eta) E_2(D_2)$ for $\eta \in [0,1]$ are e-values for H_0, regardless of any dependence between D_1 and D_2.*

Defined here for hypothesis testing, e-values then serve as the building blocks of many higher-level procedures, from parameter inference [Ramdas et al., 2022, Csillag et al., 2025, Kilian et al., 2025], risk monitoring [Podkopaev and Ramdas, 2021], change-point detection [Shin et al., 2022, Shekhar and Ramdas, 2023], test-time adaptation [Bar et al., 2024], uncertainty quantification [Balinsky and Balinsky, 2024, Gauthier et al., 2025b, Gauthier et al., 2025a], interpretability [Teneggi and Sulam, 2024], and more.

[1] Considering $1/0/+\infty$.

[2] This can be refined to require only a sequential structure of the data, rather than full independence. We keep to independence for simplicity.

3 DIFFERENTIALLY PRIVATE E-VALUES

In this section we present our framework for differentially private e-values, which guarantees privacy while retaining the validity of the resulting e-values. We will first construct our differentially private e-values in the context of hypothesis testing, and then show how these can be used for general e-value-based procdures and confidence intervals.

3.1 Hypothesis testing

Suppose we have a (non-private) e-value $E(D)$ for a null hypothesis H_0, i.e., a nonnegative real random variable whose expectation over the data D is at most 1 under the null hypothesis. Our goal is to leverage $E(D)$ to construct a new e-value $E^{\mathrm{DP}}(D)$ satisfying (α, ϵ)-Rényi differential privacy. We do this by introducing a measured amount of independent noise to the e-value, as such:

$$E^{\mathrm{DP}}(D) \coloneqq E(D) \cdot e^{-\xi}, \tag{1}$$

with random noise ξ independent from $E(D)$. By passing the noise variable ξ through the exponential and incorporating it multiplicatively we ensure that $E^{\mathrm{DP}}(D)$ remains nonnegative.

With well-designed choices for the distribution of ξ, we can simultaneously ensure differential privacy and validity. First, note that since ξ is independent from $E(D)$,

$$\mathbb{E}[E^{\mathrm{DP}}(D)] = \mathbb{E}[E(D) \cdot e^{-\xi}] = \mathbb{E}[E(D)] \cdot \mathbb{E}[e^{-\xi}];$$

and so, as long as the moment-generating function $t \mapsto \mathbb{E}[e^{t\xi}]$ exists and is at most 1 at $t = -1$, we have that $E^{\mathrm{DP}}(D)$ will be a valid e-value: under the null,

$$\mathbb{E}[E^{\mathrm{DP}}(D)] = \mathbb{E}[E(D)] \cdot \mathbb{E}[e^{-\xi}] \leq \mathbb{E}[E(D)] \leq 1,$$

where the last step follows from the fact that $E(D)$ is an e-value.

To obtain differential privacy, we then appeal to the post-processing theorem: $E^{\mathrm{DP}}(D)$ is (α, ϵ)-Rényi differentially private iff $\log E^{\mathrm{DP}}(D) = \log E(D) - \xi$ is (α, ϵ)-Rényi differentially private. This allows us to leverage existing additive noise mechanisms such as Gaussian and Laplace mechanisms to attain differential privacy, as a function of the *log-sensitivity* of the e-value, defined as

$$\Delta_{\log}(E) \coloneqq \sup_{|D \Delta D'| \leq 1} |\log E(D) - \log E(D')|, \tag{2}$$

with $|D \Delta D'| \leq 1$ denoting that D and D' differ by a single record.

To ensure validity of the resulting e-values, we bias ξ so as to ensure its moment-generating function satisfies the required bound. This leads to the following biased Gaussian and Laplace mechanisms:

Theorem 3.1 (Biased Gaussian mechanism)**.** *For any $\alpha > 1$ and $\epsilon > 0$, let*

$$E^{\mathrm{DP}}(D) = E(D) \cdot e^{-\xi},$$
$$\xi \sim \mathcal{N}\left(\frac{\alpha[\Delta_{\log}(E)]^2}{4\epsilon}, \frac{\alpha[\Delta_{\log}(E)]^2}{2\epsilon}\right).$$

Then $E^{\mathrm{DP}}(D)$ is a valid e-value satisfying (α, ϵ)-Rényi differential privacy.

Theorem 3.2 (Biased Laplace mechanism)**.** *For any $\alpha > 1$ and $\epsilon > 0$, let*

$$E^{\mathrm{DP}}(D) = E(D) \cdot e^{-\xi},$$
$$\xi \sim \mathrm{Laplace}\big(-\log(1 - b_{\alpha,\epsilon}^2), b_{\alpha,\epsilon}\big),$$

where

$$b_{\alpha,\epsilon} \coloneqq 1 \,/\, \mathbf{h}_\alpha^{-1}\left((2\alpha - 1)e^{(\alpha-1)\epsilon}\right),$$
$$\mathbf{h}_\alpha(t) \coloneqq \alpha e^{(\alpha-1)\Delta_{\log}(E)t} + (\alpha - 1)e^{-\alpha\Delta_{\log}(E)t} \quad \text{for } t \geq 0.$$

Then, as long as $b_{\alpha,\epsilon} < 1$, $E^{\mathrm{DP}}(D)$ is a valid e-value satisfying (α, ϵ)-Rényi differential privacy.

Note that, in practice, $\mathbf{h}_\alpha^{-1}$ can be computed through bracketing and bisection.

Interestingly, the biased Laplace mechanism is only viable when the sensitivity is not too high, due to the $b_{\alpha,\epsilon} < 1$ requirement. This is in contrast to the more usual (non-biased) Laplace mechanism for differential privacy, which is always applicable (but does not ensure validity of the e-values). The biased Gaussian mechanism, however, is always available.

We can also exactly characterize the stastical power of our differentially private e-values, in terms of their expected growth rates [Kelly, 1956]:

Proposition 3.3. *Let $E^{\mathrm{DP}}(D)$ be as in Equation 1. Then*

$$\mathbb{E}\left[\frac{1}{n}\log E^{\mathrm{DP}}(D)\right] = \mathbb{E}\left[\frac{1}{n}\log E(D)\right] - \frac{\mathbb{E}[\xi]}{n}.$$

Crucially, since the sensitivity of the log of e-values is usually independent of n (cf. Section 4), the penalty $\mathbb{E}[\xi]/n$ decays with a fast rate of $O(1/n)$. At the limit, our differentially private e-values are as powerful as their non-private counterparts.

Our differentially private e-values also behave in many of the usual ways, while preserving privacy.

Proposition 3.4 (Optional continuation)**.** *If $E_1^{\mathrm{DP}}(D_1)$ and $E_2^{\mathrm{DP}}(D_2)$ are (α, ϵ)-Rényi differentially private e-values for a null hypothesis H_0, with data D_1 independent from D_2, then $E_1^{\mathrm{DP}}(D_1) \cdot E_2^{\mathrm{DP}}(D_2)$ is also an (α, ϵ)-Rényi differentially private e-value for H_0. Moreover, the release of both $E_1^{\mathrm{DP}}(D_1)$ and $E_1^{\mathrm{DP}}(D_1) \cdot E_2^{\mathrm{DP}}(D_2)$ is also (α, ϵ)-Rényi differentially private.*

Proposition 3.5 (E-to-p conversion)**.** *If $E^{\mathrm{DP}}(D)$ is an (α, ϵ)-Rényi differentially private e-value for a null hypothesis H_0, then $1/E^{\mathrm{DP}}(D)$ is an (α, ϵ)-Rényi differentially private post-hoc valid p-value for H_0.*

A notable exception, however, is averaging. Though the average of e-evalues is always an e-value, the privacy guarantee may degrade under certain conditions.

Proposition 3.6 (Independent Averaging)**.** *If $E_1^{\mathrm{DP}}(D_1)$ and $E_2^{\mathrm{DP}}(D_2)$ are (α, ϵ)-Rényi differentially private e-values for a null hypothesis H_0, then for any $\eta \in [0, 1]$, $\eta E_1^{\mathrm{DP}}(D_1) + (1 - \eta)E_2^{\mathrm{DP}}(D_2)$ is also an (α, ϵ)-differentially private e-value for H_0.*

Proposition 3.7 (Dependent Averaging)**.** *More generally, if $E_1^{\mathrm{DP}}(D)$ and $E_2^{\mathrm{DP}}(D)$ are (α, ϵ)-Rényi differentially private e-values for a null hypothesis H_0, then for any $\eta \in [0, 1]$, $\eta E_1^{\mathrm{DP}}(D) + (1 - \eta)E_2^{\mathrm{DP}}(D)$ is an $(\alpha, 2\epsilon)$-differentially private e-value for H_0.*

3.2 Algorithms atop e-values and confidence intervals

Beyond hypothesis testing, e-values are also commonly used as fundamental building blocks in larger algorithms. As long as the number of e-values used is finite, the standard composition theorems of differential privacy apply, and so for appropriately chosen values of (α, ϵ) we can simply replace the procedure's e-values with our differentially private ones and attain validity with privacy.

Formally, we have an algorithm $\mathcal{A}(E_1, \ldots, E_k)$, receiving as input k e-values for k respective null hypotheses $H_0^{(k)}$. We need some notion of validity of the overall algorithm, which we assume holds whenever the input e-values are all valid.

Assumption 3.8. *If $E_1, \ldots, E_k$ are valid e-values for the nulls $H^{(1)}, \ldots, H^{(k)}$, then $\mathcal{A}(E_1, \ldots, E_k)$ is valid.*

Then the next result follows by the standard composition theorem of (Rényi) differential privacy.

Theorem 3.9. *Under Assumption 3.8, let $\alpha > 1$ and $\epsilon > 0$. For each $j = 1, \ldots, k$, let E_j^{DP} be an $(\alpha, \epsilon/k)$-Rényi differentially private e-value for the null $H_0^{(j)}$ (e.g., obtained through Theorems 3.1 and 3.2). Then $\mathcal{A}(E_1^{\mathrm{DP}}, \ldots, E_k^{\mathrm{DP}})$ is valid and (α, ϵ)-Rényi differentially private.*

This works out of the box for many algorithms, including multiple testing and online monitoring procedures. However, it is not enough for algorithms that (formally) depend on an infinite set of e-values. A particularly notable example of this are confidence intervals; an e-value-based confidence interval for some parameter $\theta^\star \in \Theta$ is typically defined by taking a family of e-values $(E_\theta(D))_{\theta\in\Theta}$ for nulls $H_0^{(\theta)} : \theta^\star = \theta$, and inverting the test as $C_\alpha(D) \coloneqq \{\theta \in \Theta : E_\theta(D) < 1/\alpha\}$.

To resolve this, we define a procedure that leverages a finite amount of our differentially private e-values to provide a provably valid confidence interval, under the assumption that the log of the (non-private) e-value, $\log E_\theta(D)$, is locally Lipschitz in the parameter θ. For simplicity we present here the scalar case where $\Theta \subset \mathbb{R}$, but the same technique can be applied in $\mathbb{R}^d$: Let $0 = a_0 < \cdots < a_k = 1$ be a partition of $[0, 1]$, and define the corresponding midpoints $\theta_j \coloneqq (a_{j-1} + a_j)/2$. Because $\log E_\theta(D)$ is locally Lipschitz, it is L_j-Lipschitz within $[a_{j-1}, a_j]$, for each $j = 1, \ldots, k$; and thus, for all $\theta' \in [a_{j-1}, a_j]$,

$$\begin{aligned} &\log E_{\theta'}(D) \geq \log E_{\theta_j}(D) - L_j|\theta' - \theta_j| \\ \implies &E_{\theta'}(D) \geq E_{\theta_j}(D) \cdot e^{-L_j|\theta'-\theta_j|} \\ &\qquad\qquad \geq E_{\theta_j}(D) \cdot e^{-L_j(a_j-a_{j-1})} =: \widetilde{E}_{\theta_j}(D), \end{aligned} \tag{3}$$

and the last right-hand-side is independent of θ'. So $\widetilde{E}_{\theta_j}(D)$ is simultaneously an e-value for all nulls $H_0 : \theta^\star = \theta'$, $\theta' \in [a_{j-1}, a_j]$, as

$$\mathbb{E}[E_{\theta_j}(D) \cdot e^{-L_j(a_j-a_{j-1})}] \leq \mathbb{E}[E_{\theta'}(D)] \leq 1.$$

Now that we have a finite amount of e-values covering all our nulls, we can straightforwardly define a procedure atop them to generate an actual CI from the finite e-values:

$$\mathrm{CI}_\alpha\Big(\widetilde{E}_{\theta_1}(D), \ldots, \widetilde{E}_{\theta_k}(D)\Big) \coloneqq \bigcup_{\substack{j=1 \\ \widetilde{E}_{\theta_j}(D)\leq 1/\alpha}}^{k} [a_{j-1}, a_j].$$

Differential privacy is now easily attained by creating differentially private versions of $\widetilde{E}_{\theta_1}(D), \ldots, \widetilde{E}_{\theta_k}(D)$ (e.g., by Theorems 3.2 and 3.1) and applying Theorem 3.9.

Corollary 3.10. *Suppose $\log E_\theta(D)$ is locally Lipschitz in θ. Let $\alpha > 1$ and $\epsilon > 0$, and let $\widetilde{E}^{\mathrm{DP}}_{\theta_1}, \ldots, \widetilde{E}^{\mathrm{DP}}_{\theta_k}$ be $(\alpha, \epsilon/k)$-Rényi differentially private versions of $\widetilde{E}_{\theta_1}, \ldots, \widetilde{E}_{\theta_k}$ as defined in Equation 3. Then $\mathrm{CI}_\alpha\big(\widetilde{E}^{\mathrm{DP}}_{\theta_1}(D), \ldots, \widetilde{E}^{\mathrm{DP}}_{\theta_k}(D)\big)$ is an (α, ϵ)-Rényi differentially private confidence interval for $\theta^\star$, i.e., it satisfies (α, ϵ)-Rényi differential privacy, and*

$$\mathbb{P}\left[\theta^\star \in \mathrm{CI}_\alpha\Big(\widetilde{E}^{\mathrm{DP}}_{\theta_1}(D), \ldots, \widetilde{E}^{\mathrm{DP}}_{\theta_k}(D)\Big)\right] \geq 1 - \alpha.$$

4 EXPERIMENTS AND APPLICATIONS

In this section we empirically evaluate our method in three settings: (i) confidence intervals for the prevalence of diabetes with private patient data (Section 4.1); (ii) private anytime-valid hypothesis testing for online risk monitoring (Section 4.2); and (iii) e-conformal prediction for the predictive modelling of online phishing attacks (Section 4.3). Experiment details can be found in the supplementary material.

Code for all experiments can be found on GitHub.[3] All experiments were run on an AMD Ryzen 9 5950X CPU, with 64GB of RAM. That said, they are lightweight and should easily run on weaker hardware.

4.1 Private e-confidence intervals

We start by the problem of producing a confidence interval for the mean of a bounded random variable, following the work of [Waudby-Smith and Ramdas, 2020]. To do so, we leverage the e-value for the mean of [Waudby-Smith and Ramdas, 2020] with betting following Cover's universal portfolios [Cover, 1996, Orabona and Jun, 2021]:

$$E_\theta(D) = \prod_{i=1}^{n}\big(1 + \lambda_i(Y_i - \theta)\big) \tag{4}$$
$$\text{with } \lambda_i = \frac{\mathbb{E}_{\lambda\sim F}\left[\lambda \cdot \prod_{j=1}^{i}\big(1 + \lambda(Y_j - \theta)\big)\right]}{\mathbb{E}_{\lambda\sim F}\left[\prod_{j=1}^{i}\big(1 + \lambda(Y_j - \theta)\big)\right]},$$

where F is a distribution with support in $[\lambda_{\inf}, \lambda_{\sup}] \subset (-1/(1-\theta), 1/\theta)$.

Besides great performance, this choice of betting strategy $(\lambda_i)_{i=1}^n$ has the key property that the e-values $E_\theta(D)$ are invariant to permutations of the data $Y_1, \ldots, Y_n$ [Orabona and Jun, 2021, Ville, 1939]. Straightfoward computation then yields the following bound on the log–sensitivity of $E_\theta(D)$:

Proposition 4.1. *Let F have support in $[\lambda_{\inf}, \lambda_{\sup}] \subset (-1/(1-\theta), 1/\theta)$. The log-sensitivity of $E_\theta(D)$, $\Delta_{\log}(E_\theta)$, is upper bounded as*

$$\begin{aligned}\Delta_{\log}(E_\theta) \leq \max\big\{&\log(1 + \max\{\lambda_{\sup}(1-\theta), -\lambda_{\inf}\theta\}), \\ &-\log(1 + \min\{\lambda_{\inf}(1-\theta), -\lambda_{\sup}\theta\})\big\}.\end{aligned}$$

Moreover, to apply our Proposition 3.10 for constructing valid confidence intervals, we further bound the Lipschitz constant of the log of the e-value over $\theta \in [\theta_{\inf}, \theta_{\sup}]$:

[3] https://github.com/dccsillag/experiments-evalue-dp.

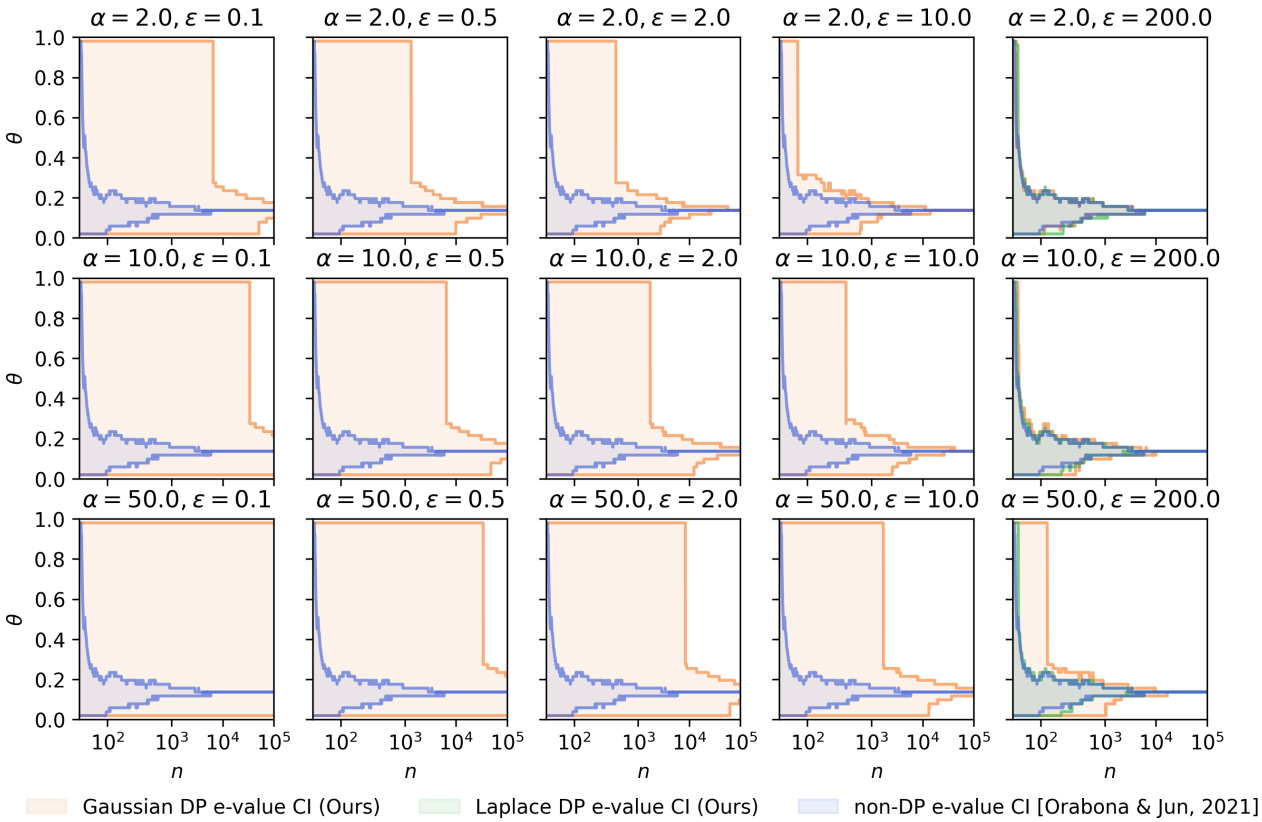

Figure 1: **Private e-confidence intervals.** We use our differentially private e-values to produce private confidence intervals for the prevalence of diabetes over a population. We illustrate the confidence intervals varying the number of samples n (x axis) and Rényi privacy budgets (α, ϵ). We display the non-private confidence intervals (blue) along with those obtained through our biased Gaussian mechanism (orange), and when applicable also the as the ones obtained via our biased Laplace mechanism (green). Note that our biased Gaussian mechanism always converges to the non-private confidence intervals, as does the biased Laplace mechanism when available.

Proposition 4.2. *For distributions F with support contained in $[\lambda_{\text{inf}}, \lambda_{\text{sup}}]$, the Lipschitz constant of $\log E_\theta(D)$ over $\theta \in [\theta_{\text{inf}}, \theta_{\text{sup}}]$ is upper bounded by*

$$\|\theta \mapsto \log E_\theta(D)\|_{\text{Lip}} \leq \max\left\{\left|\frac{\lambda_{\text{sup}}}{1-\lambda_{\text{sup}}\theta_{\text{sup}}}\right|, \left|\frac{\lambda_{\text{inf}}}{1+\lambda_{\text{inf}}(1-\theta_{\text{inf}})}\right|\right\}.$$

Altogether, we can now directly use the biased Gaussian and Laplace mechanisms defined in Section 3.1 along with Proposition 3.10 to construct differentially private e-value-based confidence intervals for our data.

For our experiment, we use the e-values from Equation 4 atop the dataset of [CDC, 2015] for estimating the prevalence of diabetes, and vary the Rényi privacy budget over $\alpha \in \{2, 10, 50\}$ and $\epsilon \in \{0.1, 0.5, 2, 10, 200\}$. The results can be seen in Figure 1; we note that our method with biased Gaussian noise always converges to the non-private intervals as $n \to \infty$, and very closely matches the non-private confidence intervals over all n with $\epsilon = 10$. For smaller values of ϵ (e.g. 0.1), our method requires only mild increases in data size, past a certain point. The biased Laplace mechanism, however, is not defined for most combinations of (α, ϵ), due to the $b_{\alpha,\epsilon} < 1$ requirement (cf. Theorem 3.2). But when it is defined, it is quite accurate.

4.2 Anytime-valid hypothesis testing for online risk monitoring

We now turn our attention to the problem of online risk monitoring with private data. In this setting we have a pre-trained predictive model $\widehat{\mu}$, and want to continuously track its test loss so as to ensure it does not go over some predetermined safety threshold. Inspired by the work of [Podkopaev and Ramdas, 2021] and [Csillag et al., 2025], we frame this as an anytime-valid hypothesis test for the null hypothesis $H_0 : \mathbb{E}[\text{Loss}(\widehat{\mu}(X_i), Y_i)] \leq \text{SafetyThreshold}$ for all i. This corresponds to a one-sided test for the mean, which we can perform using the same e-value from Equation 4, but with F now being the uniform distri-

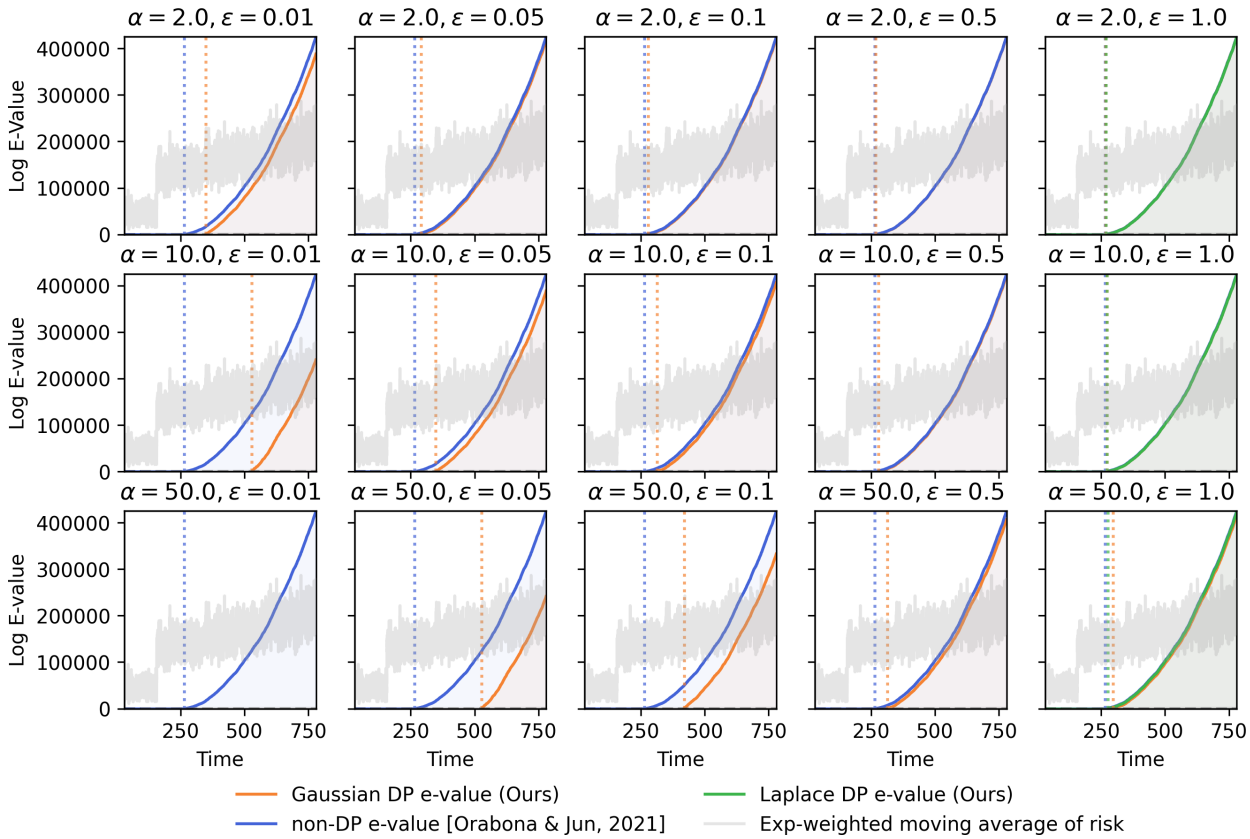

Figure 2: **Anytime-valid private hypothesis testing.** We apply our differentially private e-values for private continuous risk monitoring by the means of anytime-valid hypothesis testing. Our goal is to detect significant increases in the risk of a deployed predictive model, whose moving average can be seen in the background. The curves indicate the log of the e-values for the null hypothesis that the risk is under a certain safety threshold. Our private e-values closely match the non-private ones, quickly rejecting the null hypothesis after the change-point even for very small values of ϵ.

bution with support in $[0, c/\theta)$ for some $0 < c < 1$.

Crucially, in to attain anytime-validity with differential privacy we conduct our inference in batches, and appeal to the private optional continuation property (Proposition 3.4). In each batch we compute the privatized form of the e-value of Equation 4 with the one-sided distribution F, and combine the e-values across batches through multiplication.

For this experiment we use the data of [Blackard, 1998] with a simulated change-point. Figure 2 shows this procedure in action, with the Rényi privacy budget varying over $\alpha \in \{2, 10, 50\}$ and $\epsilon \in \{0.01, 0.05, 0.1, 0.5, 1\}$. We note the remarkable accuracy of our biased Gaussian mechanism in this setting, even for small values of ϵ, which often rejects the null hypothesis (and thus triggers an alarm of a change of distribution) very close to the non-private procedure, and thus also close to the actual change-point in the data. However, similar to our previous example in Section 4.1, the biased Laplace mechanism is only defined for most higher values of ϵ.

4.3 Private e-conformal prediction

Finally, we apply our method to privatize e-conformal prediction [Gauthier et al., 2025b, Balinsky and Balinsky, 2024]. E-conformal prediction provides predictive uncertainty quantification for machine learning models, achieving strong guarantees (e.g. post-hoc validity) by leveraging e-values. Given calibration samples $D = (X_i, Y_i)_{i=1}^n \subset \mathcal{X} \times \mathcal{Y}$ and a a score function $s : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_{>0}$ (which typically incorporates a predictive model), e-conformal prediction produces predictive sets as

$$C_\alpha(x) \coloneqq \left\{ y \in \mathcal{Y} : E^{\text{exch}}(D; s(x, y)) < 1/\alpha \right\},$$

where $E^{\text{exch}}(D; s(x, y))$ is an e-value for the null that $s(X_1, Y_1), \ldots, s(X_n, Y_n), s(x, y)$ are exchangeable random variables. A common choice is

$$E^{\text{exch}}(D; S^{\text{test}}) = \frac{(n+1)S^{\text{test}}}{\sum_{i=1}^n s(X_i, Y_i) + S^{\text{test}}}.$$

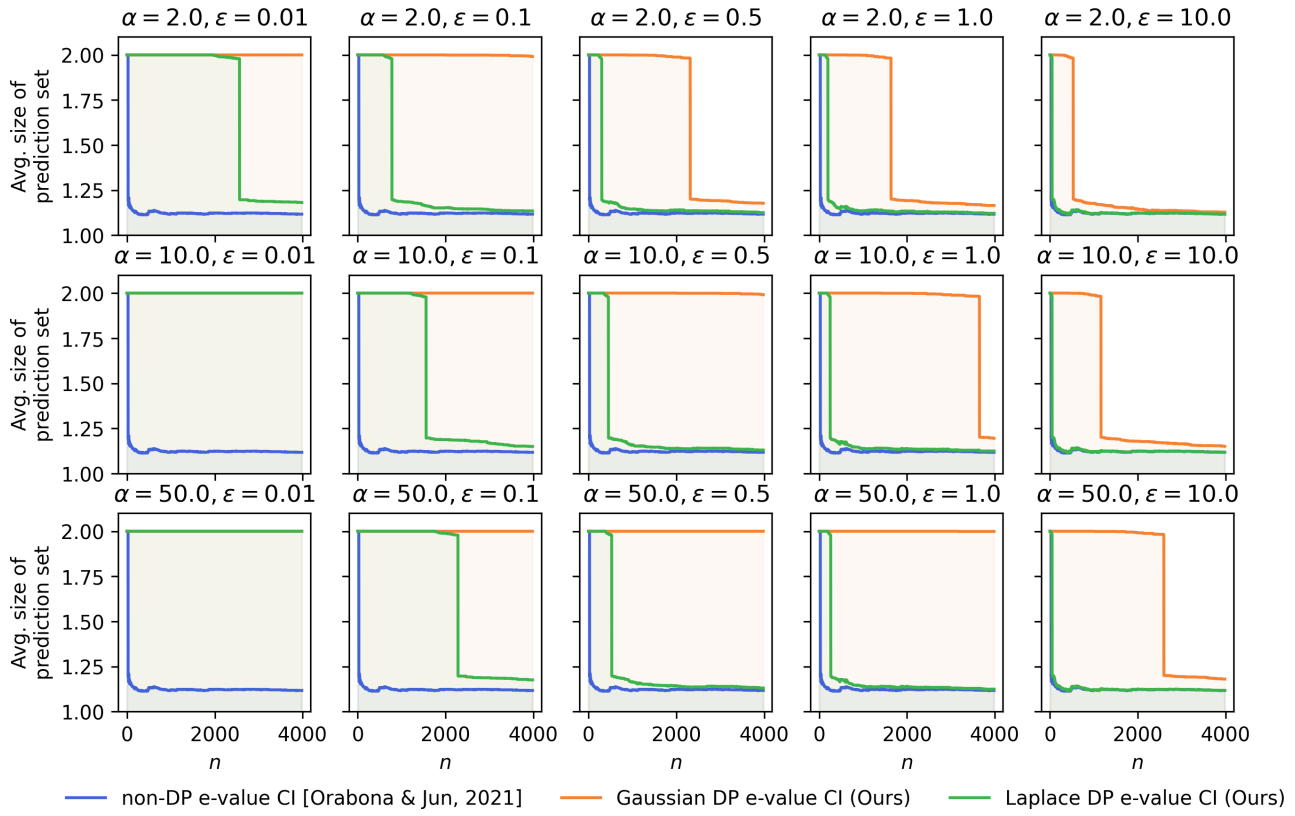

Figure 3: **Private e-conformal prediction.** We also apply our e-values to produce private predictive sets with post-hoc guarantees through e-conformal prediction. In this setting the biased Laplace mechanism is generally applicable, and consistently outperforms the biased Gaussian mechanism, closely matching the average size of non-private prediction sets.

We furthermore consider that the conformity score s has its image contained within a finite set $s_1, \ldots, s_K \subset [a, b]$. This makes it so that $C_\alpha(\cdot)$ is a function of a finite amount of e-values (one for each s_i, $i = 1, \ldots, K$) and affords us the following bound on the log-sensitivity:

Proposition 4.3. *Suppose the score function and S^{test} are all contained in $[a, b]$. Then the log-sensitivity of $E^{\text{exch}}(D; S^{\text{test}})$, $\Delta_{\log}(E^{\text{exch}})$, is upper bounded as*

$$\Delta_{\log}(E^{\text{exch}}) \leq 2 \cdot \frac{b/a}{n+1}.$$

Then, by privatizing these e-values for exchangeability and using Theorem 3.9, we obtain private e-conformal predictive sets.

Figure 3 shows the predictive set sizes for conformal prediction atop the phishing classification dataset of [Mohammad and McCluskey, 2012], for Rényi privacy budgets with $\alpha \in \{2, 10, 50\}$ and $\epsilon \in \{0.01, 0.1, 0.5, 1, 10\}$. For our conformity score, we use a truncated version of the score proposed by [Gauthier et al., 2025b], quantized onto 500 bins in $[1, 100]$, which affords us reasonably low sensitivities. Remarkably, we find that in this instance the biased Laplace is (i) always defined, and (ii) consistently outperforms the biased Gaussian mechanism, quickly approaching the non-private average predictive set size even for small values of ϵ.

5 CONCLUSION

In this work we introduced the first general framework for constructing differentially private e-values that simultaneously guarantee privacy protection and statistical validity. Through novel biased multiplicative noise mechanisms, we showed that our private e-values retain strong statistical power – asymptotically matching their non-private counterparts – while preserving their key compositional properties. Our experiments on online risk monitoring, private healthcare analytics, and conformal prediction demonstrate the practical effectiveness of our approach across diverse privacy-sensitive applications. By bridging differential privacy and e-value-based inference, this work enables the principled deployment of flexible e-value-based statistical procedures on sensitive data, broadening the

applicability of modern inference methods to domains with rigorous privacy requirements.

Acknowledgements

We acknowledge the support by the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (SEI-260003/020348/2025, SEI-260003/020694/2025) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (404336/2023-0, 305692/2025-9).

References

[Balinsky and Balinsky, 2024] Balinsky, A. A. and Balinsky, A. D. (2024). Enhancing conformal prediction using e-test statistics. *ArXiv*, abs/2403.19082.

[Bar et al., 2024] Bar, Y., Shaer, S., and Romano, Y. (2024). Protected test-time adaptation via online entropy matching: A betting approach. *ArXiv*, abs/2408.07511.

[Blackard, 1998] Blackard, J. (1998). Covertype. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C50K5N.

[Campbell et al., 2017] Campbell, Z., Bray, A., Ritz, A. M., and Groce, A. (2017). Differentially private anova testing. *2018 1st International Conference on Data Intelligence and Security (ICDIS)*, pages 281–285.

[CDC, 2015] CDC (2015). Cdc – 2014 brfss survey data and documentation. Last accessed 30 January 2025.

[Couch et al., 2019] Couch, S., Kazan, Z., Shi, K., Bray, A., and Groce, A. (2019). Differentially private nonparametric hypothesis testing. *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*.

[Cover, 1996] Cover, T. M. (1996). Universal portfolios.

[Csillag et al., 2025] Csillag, D., Struchiner, C. J., and Goedert, G. T. (2025). Prediction-powered e-values. *ArXiv*, abs/2502.04294.

[Ding et al., 2018] Ding, B., Nori, H., Li, P. L., and Allen, J. (2018). Comparing population means under local differential privacy: with significance and power. In *AAAI Conference on Artificial Intelligence*.

[Du et al., 2020] Du, W., Foot, C., Moniot, M., Bray, A., and Groce, A. (2020). Differentially private confidence intervals. *ArXiv*, abs/2001.02285.

[Dwork et al., 2006] Dwork, C., McSherry, F., Nissim, K., and Smith, A. D. (2006). Calibrating noise to sensitivity in private data analysis. *J. Priv. Confidentiality*, 7:17–51.

[Dwork and Roth, 2014] Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9:211–407.

[Dwork et al., 2018] Dwork, C., Su, W., and Zhang, L. (2018). Differentially private false discovery rate control. *J. Priv. Confidentiality*, 11.

[Ferrando et al., 2020] Ferrando, C., Wang, S., and Sheldon, D. (2020). Parametric bootstrap for differentially private confidence intervals. In *International Conference on Artificial Intelligence and Statistics*.

[Gaboardi et al., 2018] Gaboardi, M., Rogers, R. M., and Sheffet, O. (2018). Locally private mean estimation: Z-test and tight confidence intervals. *ArXiv*, abs/1810.08054.

[Gauthier et al., 2025a] Gauthier, E., Bach, F., and Jordan, M. I. (2025a). Backward conformal prediction. *ArXiv*, abs/2505.13732.

[Gauthier et al., 2025b] Gauthier, E., Bach, F., and Jordan, M. I. (2025b). E-values expand the scope of conformal prediction. *ArXiv*, abs/2503.13050.

[Karwa and Vadhan, 2017] Karwa, V. and Vadhan, S. P. (2017). Finite sample differentially private confidence intervals. In *Information Technology Convergence and Services*.

[Kazan et al., 2023] Kazan, Z., Shi, K., Groce, A., and Bray, A. (2023). The test of tests: A framework for differentially private hypothesis testing. *ArXiv*, abs/2302.04260.

[Kelly, 1956] Kelly, J. L. (1956). A new interpretation of information rate. *IRE Trans. Inf. Theory*, 2:185–189.

[Kilian et al., 2025] Kilian, V., Cortinovis, S., and Caron, F. (2025). Anytime-valid, bayes-assisted,prediction-powered inference. *ArXiv*, abs/2505.18000.

[Koning, 2023] Koning, N. W. (2023). Post-hoc α hypothesis testing and the post-hoc p-value.

[Mironov, 2017] Mironov, I. (2017). Rényi differential privacy. *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275.

[Mohammad and McCluskey, 2012] Mohammad, R. and McCluskey, L. (2012). Phishing Websites. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C51W2X.

[Orabona and Jun, 2021] Orabona, F. and Jun, K.-S. (2021). Tight concentrations and confidence sequences from the regret of universal portfolio. *IEEE Transactions on Information Theory*, 70:436–455.

[Pena and Barrientos, 2022] Pena, V. and Barrientos, A. F. (2022). Differentially private hypothesis testing with the subsampled and aggregated randomized response mechanism. *Statistica Sinica*.

[Podkopaev and Ramdas, 2021] Podkopaev, A. and Ramdas, A. (2021). Tracking the risk of a deployed model and detecting harmful distribution shifts. *ArXiv*, abs/2110.06177.

[Ramdas et al., 2022] Ramdas, A., Grünwald, P. D., Vovk, V., and Shafer, G. (2022). Game-theoretic statistics and safe anytime-valid inference. *ArXiv*, abs/2210.01948.

[Rogers and Kifer, 2017] Rogers, R. M. and Kifer, D. (2017). A new class of private chi-square hypothesis tests. In *International Conference on Artificial Intelligence and Statistics*.

[Shekhar and Ramdas, 2023] Shekhar, S. and Ramdas, A. (2023). Sequential change detection via backward confidence sequences. *ArXiv*, abs/2302.02544.

[Shin et al., 2022] Shin, J., Ramdas, A., and Rinaldo, A. (2022). E-detectors: A nonparametric framework for sequential change detection. *The New England Journal of Statistics in Data Science*.

[Swanberg et al., 2019] Swanberg, M., Globus-Harris, I., Griffith, I., Ritz, A. M., Groce, A., and Bray, A. (2019). Improved differentially private analysis of variance. *Proceedings on Privacy Enhancing Technologies*, 2019:310 – 330.

[Teneggi and Sulam, 2024] Teneggi, J. and Sulam, J. (2024). I bet you did not mean that: Testing semantic importance via betting. *ArXiv*, abs/2405.19146.

[Ville, 1939] Ville, J.-L. (1939). Étude critique de la notion de collectif.

[Vovk and Wang, 2019] Vovk, V. and Wang, R. (2019). E-values: Calibration, combination, and applications. *Political Methods: Quantitative Methods eJournal*.

[Waudby-Smith and Ramdas, 2020] Waudby-Smith, I. and Ramdas, A. (2020). Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society Series B: Statistical Methodology*.

[Zhang et al., 2020] Zhang, W., Kamath, G., and Cummings, R. (2020). Paprika: Private online false discovery rate control. *ArXiv*, abs/2002.12321.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes] (e.g. statistical power; time and memory complexities are trivial.)
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Yes]

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]
   (b) Complete proofs of all theoretical results. [Yes]
   (c) Clear explanations of any assumptions. [Yes]

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Yes]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Yes]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Yes]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Yes]

(b) The license information of the assets, if applicable. [Not Applicable]

(c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]

(d) Information about consent from data providers/curators. [Not Applicable]

(e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:

   (a) The full text of instructions given to participants and screenshots. [Not Applicable]

   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]

   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Differentially Private E-Values: Supplementary Materials

A Proofs from the main text

For proofs of Propositions 2.3 and 2.4, see [Mironov, 2017]. The proof of Proposition 2.6 can be found in [Koning, 2023]. For Propositions 2.7 and 2.8, see [Vovk and Wang, 2019].

Theorem A.1 (Biased Gaussian mechanism for a single e-value; Theorem 3.1 in the main text)**.** *For any $\alpha > 1$ and $\epsilon > 0$, let*

$$E^{\mathrm{DP}}(D) = E(D) \cdot e^{-\xi}, \qquad \xi \sim \mathcal{N}\left(\frac{\alpha[\Delta_{\log}(E)]^2}{4\epsilon}, \frac{\alpha[\Delta_{\log}(E)]^2}{2\epsilon}\right).$$

Then $E^{\mathrm{DP}}(D)$ is a valid e-value satisfying (α, ϵ)-Rényi differential privacy.

Proof. We are considering $\xi \sim \mathcal{N}(\mu, \sigma^2)$ for some μ and σ^2. By the post-processing theorem, $E^{\mathrm{DP}}(D) = E(D) \cdot e^{-\xi}$ is (α, ϵ)-RDP iff $\log E^{\mathrm{DP}}(D) = \log\left(E(D) \cdot e^{-\xi}\right) = \log E(D) - \xi$ is (α, ϵ)-RDP; note that $(\log E(D) - xi) \sim \mathcal{N}(\log E(D) - \mu, \sigma^2)$. It then follows, using Proposition 7 of [Mironov, 2017], that

$$\begin{aligned}
D_\alpha(\log E^{\mathrm{DP}}(D) \,\|\, \log E^{\mathrm{DP}}(D')) &= D_\alpha(\mathcal{N}(\log E(D) - \mu, \sigma^2) \,\|\, \mathcal{N}(\log E(D') - \mu, \sigma^2)) \\
&= D_\alpha(\mathcal{N}(0, \sigma^2) \,\|\, \mathcal{N}((\log E(D') - \mu) - (\log E(D) - \mu), \sigma^2)) \\
&= D_\alpha(\mathcal{N}(0, \sigma^2) \,\|\, \mathcal{N}(\log E(D') - \log E(D), \sigma^2)) \\
&= \frac{\alpha(\log E(D') - \log E(D))^2}{2\sigma^2}.
\end{aligned}$$

For this to be at most ϵ, we require:

$$\begin{aligned}
& \frac{\alpha(\log E(D') - \log E(D))^2}{2\sigma^2} \leq \epsilon \\
\iff & \frac{\alpha(\log E(D') - \log E(D))^2}{2\epsilon} \leq \sigma^2 \\
\Longleftarrow & \frac{\alpha[\Delta_{\log}(E)]^2}{2\epsilon} \leq \sigma^2;
\end{aligned}$$

So, taking $\sigma^2 = \alpha[\Delta_{\log}(E)]^2/2\epsilon$ ensures (α, ϵ)-Renyi differential privacy.

Thus, all that remains is to find the smallest bias μ that ensures validity of $E^{\mathrm{DP}}(D)$. As argued in Section 3.1, all we need is that the MGF of ξ is at most one at -1. So it follows, using the form of the MGF of an univariate normal:

$$\begin{aligned}
& \mathbb{E}[e^{-\xi}] = \exp\left(-\mu + \frac{\alpha[\Delta_{\log}(E)]^2/2\epsilon}{2}\right) \leq 1 \\
\iff & -\mu + \frac{\alpha[\Delta_{\log}(E)]^2/2\epsilon}{2} \leq 0 \\
\iff & \mu \geq \frac{\alpha[\Delta_{\log}(E)]^2/2\epsilon}{2} = \frac{\alpha[\Delta_{\log}(E)]^2}{4\epsilon}.
\end{aligned}$$

□

In order to derive the biased Laplace mechanism, we first introduce the following Lemma, which is a proper generalization of Proposition 6 of [Mironov, 2017].

Lemma A.2. *Let $\alpha > 1$. Then the α-Rényi divergence between* $\mathrm{Laplace}(0, b)$ *and* $\mathrm{Laplace}(\mu, b)$ *for any $\mu \in \mathbb{R}$ and $b > 0$ is given by*

$$D_\alpha\big(\mathrm{Laplace}(0,b) \,\|\, \mathrm{Laplace}(\mu,b)\big) = \frac{1}{\alpha-1}\log\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)|\mu|/b} + \frac{\alpha-1}{2\alpha-1}e^{-\alpha|\mu|/b}\right).$$

Additionally, $D_\alpha\big(\mathrm{Laplace}(0,b) \,\|\, \mathrm{Laplace}(\mu,b)\big)$ is increasing in $|\mu|$.

Proof. By definition of D_α,

$$\begin{aligned}
D_\alpha\big(\mathrm{Laplace}(0,b) \,\|\, \mathrm{Laplace}(\mu,b)\big) &= \frac{1}{\alpha-1}\log\int_{-\infty}^{\infty}\left(\frac{1}{2b}e^{-|x|/b}\right)^{\alpha}\left(\frac{1}{2b}e^{-|x-\mu|/b}\right)^{1-\alpha}dx \\
&= \frac{1}{\alpha-1}\log\int_{-\infty}^{\infty}\frac{1}{2b}e^{-\alpha|x|/b-(1-\alpha)|x-\mu|/b}.
\end{aligned}$$

We need to evaluate

$$I = \int_{-\infty}^{\infty}\frac{1}{2b}e^{-\alpha|x|/b-(1-\alpha)|x-\mu|/b}\,dx.$$

First, let us assume that $\mu \geq 0$. We then split the integral into three regions:

Region 1: $x < 0$

$$\int_{-\infty}^{0}\frac{1}{2b}e^{\alpha x/b-(1-\alpha)(\mu-x)/b}\,dx = \frac{1}{2b}e^{-(1-\alpha)\mu/b}\int_{-\infty}^{0}e^{x/b}\,dx = \frac{1}{2}e^{-(1-\alpha)\mu/b}$$

Region 2: $0 \leq x < \mu$

$$\begin{aligned}
\int_{0}^{\mu}\frac{1}{2b}e^{-\alpha x/b-(1-\alpha)(\mu-x)/b}\,dx &= \frac{1}{2b}e^{-(1-\alpha)\mu/b}\int_{0}^{\mu}e^{-(2\alpha-1)x/b}\,dx \\
&= \frac{1}{2}e^{-(1-\alpha)\mu/b}\frac{1-e^{-(2\alpha-1)\mu/b}}{2\alpha-1}
\end{aligned}$$

Region 3: $x \geq \mu$

$$\int_{\mu}^{\infty}\frac{1}{2b}e^{-\alpha x/b-(1-\alpha)(x-\mu)/b}\,dx = \frac{1}{2b}e^{(1-\alpha)\mu/b}\int_{\mu}^{\infty}e^{-x/b}\,dx = \frac{1}{2}e^{-\alpha\mu/b}$$

We thus have

$$\begin{aligned}
I &= \frac{1}{2}e^{-(1-\alpha)\mu/b}\left[1+\frac{1-e^{-(2\alpha-1)\mu/b}}{2\alpha-1}\right] + \frac{1}{2}e^{-\alpha\mu/b} \\
&= \frac{1}{2}e^{(\alpha-1)\mu/b}\left[1+\frac{1-e^{-(2\alpha-1)\mu/b}}{2\alpha-1}\right] + \frac{1}{2}e^{-\alpha\mu/b} \\
&= \frac{1}{2}e^{(\alpha-1)\mu/b}\left[\frac{2\alpha-1+1-e^{-(2\alpha-1)\mu/b}}{2\alpha-1}\right] + \frac{1}{2}e^{-\alpha\mu/b} \\
&= \frac{1}{2}e^{(\alpha-1)\mu/b}\left[\frac{2\alpha-e^{-(2\alpha-1)\mu/b}}{2\alpha-1}\right] + \frac{1}{2}e^{-\alpha\mu/b} \\
&= \frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b} - \frac{1}{2}e^{(\alpha-1)\mu/b}\frac{e^{-(2\alpha-1)\mu/b}}{2\alpha-1} + \frac{1}{2}e^{-\alpha\mu/b} \\
&= \frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b} - \frac{1}{2}\frac{e^{-\alpha\mu/b}}{2\alpha-1} + \frac{1}{2}e^{-\alpha\mu/b} \\
&= \frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b} + \frac{\alpha-1}{2\alpha-1}e^{-\alpha\mu/b}.
\end{aligned}$$

For $\mu < 0$, we have

$$\begin{aligned} D_\alpha\big(\text{Laplace}(0,b) \,\|\, \text{Laplace}(\mu,b)\big) &= D_\alpha\big(-\text{Laplace}(0,b) \,\|\, -\text{Laplace}(\mu,b)\big) \\ &= D_\alpha\big(\text{Laplace}(0,b) \,\|\, \text{Laplace}(-\mu,b)\big) \\ &= D_\alpha\big(\text{Laplace}(0,b) \,\|\, \text{Laplace}(|\mu|,b)\big), \end{aligned}$$

from which we conclude the computation.

Finally, to see that this is increasing in $|\mu|$, we take the derivative:

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\mu}\frac{1}{\alpha-1}\log\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha\mu/b}\right) \\ &= \frac{1}{\alpha-1}\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha\mu/b}\right)^{-1}\frac{\mathrm{d}}{\mathrm{d}\mu}\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha\mu/b}\right) > 0 \\ &\iff \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha\mu/b}\right) > 0 \\ &\iff \frac{\alpha}{2\alpha-1}\frac{\mathrm{d}}{\mathrm{d}\mu}e^{(\alpha-1)\mu/b}+\frac{\alpha-1}{2\alpha-1}\frac{\mathrm{d}}{\mathrm{d}\mu}e^{-\alpha\mu/b} > 0 \\ &\iff \frac{\alpha}{2\alpha-1}e^{(\alpha-1)\mu/b}\frac{\mathrm{d}}{\mathrm{d}\mu}((\alpha-1)\mu/b)+\frac{\alpha-1}{2\alpha-1}e^{-\alpha\mu/b}\frac{\mathrm{d}}{\mathrm{d}(-\alpha\mu/b)\mu} > 0 \\ &\iff \alpha e^{(\alpha-1)\mu/b}\frac{\mathrm{d}}{\mathrm{d}\mu}((\alpha-1)\mu/b)+(\alpha-1)e^{-\alpha\mu/b}\frac{\mathrm{d}}{\mathrm{d}\mu}(-\alpha\mu/b) > 0 \\ &\iff \alpha e^{(\alpha-1)\mu/b}\frac{\alpha-1}{b}-(\alpha-1)e^{-\alpha\mu/b}\frac{\alpha}{b} > 0 \iff \frac{\alpha(\alpha-1)}{b}e^{(\alpha-1)\mu/b}-\frac{\alpha(\alpha-1)}{b}e^{-\alpha\mu/b} > 0 \\ &\iff e^{(\alpha-1)\mu/b}-e^{-\alpha\mu/b} > 0 \iff (\alpha-1)\mu/b > -\alpha\mu/b \iff \alpha-1 > -\alpha \iff \alpha > \frac{1}{2}, \end{aligned}$$

which is always true. □

Theorem A.3 (Biased Laplace mechanism for a single e-value; Theorem 3.2 in the main text)**.** *For any $\alpha > 1$ and $\epsilon > 0$, let*

$$E^{\mathrm{DP}}(D) = E(D)\cdot e^{-\xi}, \qquad \xi \sim \text{Laplace}\big(-\log(1-b_{\alpha,\epsilon}^2), b_{\alpha,\epsilon}\big),$$

where

$$\begin{aligned} b_{\alpha,\epsilon} &:= 1 \,/\, \mathbf{h}_\alpha^{-1}\left((2\alpha-1)e^{(\alpha-1)\epsilon}\right), \\ \mathbf{h}_\alpha(t) &:= \alpha e^{(\alpha-1)\Delta_{\log}(E)t} + (\alpha-1)e^{-\alpha\Delta_{\log}(E)t} \qquad \textit{for } t \geq 0. \end{aligned}$$

Then, as long as $b_{\alpha,\epsilon} < 1$, $E^{\mathrm{DP}}(D)$ is a valid e-value satisfying (α,ϵ)-Rényi differential privacy.

Proof. We are considering $\xi \sim \text{Laplace}(\mu, b)$ for some μ and b. By the post-processing theorem, $E^{\mathrm{DP}}(D) = E(D)\cdot e^{-\xi}$ is (α,ϵ)-RDP iff $\log E^{\mathrm{DP}}(D) = \log\big(E(D)\cdot e^{-\xi}\big) = \log E(D) - \xi$ is (α,ϵ)-RDP; note that $(\log E(D) - xi) \sim \text{Laplace}(\log E(D) - \mu, b)$. It then follows, using using Lemma A.2, that

$$\begin{aligned} &D_\alpha(\log E^{\mathrm{DP}}(D) \,\|\, \log E^{\mathrm{DP}}(D')) \\ &= D_\alpha(\text{Laplace}(\log E(D)-\mu, b) \,\|\, \text{Laplace}(\log E(D')-\mu, b)) \\ &= D_\alpha(\text{Laplace}(0,b) \,\|\, \text{Laplace}((\log E(D')-\mu)-(\log E(D)-\mu), b)) \\ &= D_\alpha(\text{Laplace}(0,b) \,\|\, \text{Laplace}(\log E(D')-\log E(D), b)) \\ &= \frac{1}{\alpha-1}\log\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)|\log E(D')-\log E(D)|/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha|\log E(D')-\log E(D)|/b}\right). \end{aligned}$$

For this to be at most ϵ, we require:

$$\frac{1}{\alpha-1}\log\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)|\log E(D')-\log E(D)|/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha|\log E(D')-\log E(D)|/b}\right) \leq \epsilon$$

$$
\begin{aligned}
&\Longleftarrow \frac{1}{\alpha-1}\log\left(\frac{\alpha}{2\alpha-1}e^{(\alpha-1)\Delta_{\log}(E)/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha\Delta_{\log}(E)/b}\right)\leq\epsilon\\
&\Longleftrightarrow \frac{\alpha}{2\alpha-1}e^{(\alpha-1)\Delta_{\log}(E)/b}+\frac{\alpha-1}{2\alpha-1}e^{-\alpha\Delta_{\log}(E)/b}\leq e^{(\alpha-1)\epsilon}\\
&\Longleftrightarrow \alpha e^{(\alpha-1)\Delta_{\log}(E)/b}+(\alpha-1)e^{-\alpha\Delta_{\log}(E)/b}\leq(2\alpha-1)e^{(\alpha-1)\epsilon}.
\end{aligned}
$$

Let $\mathbf{h}_\alpha(t)=\alpha e^{(\alpha-1)\Delta_{\log}(E)t}+(\alpha-1)e^{-\alpha\Delta_{\log}(E)t}$, for $t\geq 0$. Note that this is strictly increasing in t, as

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{h}_\alpha(t)&=\frac{\mathrm{d}}{\mathrm{d}t}\alpha e^{(\alpha-1)ct}+\frac{\mathrm{d}}{\mathrm{d}t}(\alpha-1)e^{-\alpha ct}\\
&=\alpha e^{(\alpha-1)ct}(\alpha-1)c+(\alpha-1)e^{-\alpha ct}(-\alpha c)\\
&=\alpha(\alpha-1)c\left(e^{(\alpha-1)ct}-e^{-\alpha ct}\right)>0\\
&\Longleftrightarrow e^{(\alpha-1)ct}>e^{-\alpha ct}\iff(\alpha-1)ct>-\alpha ct\\
&\Longleftrightarrow \alpha-1>-\alpha\iff\alpha>\frac{1}{2},
\end{aligned}
$$

which is always true. Hence, the inverse $\mathbf{h}_\alpha^{-1}(y)$ exists. This allows us to write

$$
\begin{aligned}
&\alpha e^{(\alpha-1)\Delta_{\log}(E)/b}+(\alpha-1)e^{-\alpha\Delta_{\log}(E)/b}\leq(2\alpha-1)e^{(\alpha-1)\epsilon}\\
\Longleftrightarrow\ &\mathbf{h}_\alpha(1/b)\leq(2\alpha-1)e^{(\alpha-1)\epsilon}\\
\Longleftrightarrow\ &1/b\leq\mathbf{h}_\alpha^{-1}\left((2\alpha-1)e^{(\alpha-1)\epsilon}\right)\\
\Longleftrightarrow\ &b\geq 1/\mathbf{h}_\alpha^{-1}\left((2\alpha-1)e^{(\alpha-1)\epsilon}\right).
\end{aligned}
$$

So, taking $b=1/\mathbf{h}_\alpha^{-1}\left((2\alpha-1)e^{(\alpha-1)\epsilon}\right)$ ensures (α,ϵ)-Renyi differential privacy.

Thus, all that remains is to find the smallest bias μ that ensures validity of $E^{\mathrm{DP}}(D)$. As argued in Section 3.1, all we need is that the MGF of ξ is at most one at -1. The MGF of a Laplace distribution with mean μ and scale parameter b is given by $t\mapsto e^{t\mu}/(1-b^2t^2)$, defined at $|t|<1/b$. Thus we need that $|-1|=1<1/b$, i.e., that $b<1$. As long as this is satisfied, it then follows:

$$
\begin{aligned}
&\mathbb{E}[e^{-\xi}]=e^{-\mu}/\left(1-b^2\right)\leq 1\\
&\Longleftrightarrow e^{-\mu}\leq 1-b^2\\
&\Longleftrightarrow \mu\geq-\log(1-b^2).
\end{aligned}
$$

□

Proposition A.4 (Proposition 3.3 in the main text)**.** *Let $E^{\mathrm{DP}}(D)$ be as in Equation 1 in the main text. Then*

$$
\mathbb{E}\left[\frac{1}{n}\log E^{\mathrm{DP}}(D)\right]=\mathbb{E}\left[\frac{1}{n}\log E(D)\right]-\frac{\mathbb{E}[\xi]}{n}.
$$

Proof.

$$
\begin{aligned}
\mathbb{E}\left[\frac{1}{n}\log E^{\mathrm{DP}}(D)\right]=\mathbb{E}\left[\frac{1}{n}\log\big(E(D)\cdot e^{-\xi}\big)\right]&=\mathbb{E}\left[\frac{1}{n}\log E(D)-\frac{1}{n}\xi\right]\\
&=\mathbb{E}\left[\frac{1}{n}\log E(D)\right]-\frac{\mathbb{E}[\xi]}{n}.
\end{aligned}
$$

□

Proposition A.5 (Optional continuation; Proposition 3.4 in the main text)**.** *If $E_1^{\mathrm{DP}}(D_1)$ and $E_2^{\mathrm{DP}}(D_2)$ are (α,ϵ)-Rényi differentially private e-values for a null hypothesis H_0, with data D_1 independent from D_2, then $E_1^{\mathrm{DP}}(D_1)\cdot E_2^{\mathrm{DP}}(D_2)$ is also an (α,ϵ)-Rényi differentially private e-value for H_0. Moreover, the release of both $E_1^{\mathrm{DP}}(D_1)$ and $E_1^{\mathrm{DP}}(D_1)\cdot E_2^{\mathrm{DP}}(D_2)$ is also (α,ϵ)-Rényi differentially private.*

Proof. Let (D_1',D_2') be a neighboring dataset to (D_1,D_2), i.e., either (i) D_1' and D_1 differ by a single element and $D_2'=D_2$; (ii) D_2' and D_2 differ by a single element and $D_1'=D_1$; or (iii) $D_1'=D_1$ and $D_2'=D_2$. And since $E_1^{\mathrm{DP}}(\cdot)$ and $E_2^{\mathrm{DP}}(\cdot)$ are (α,ϵ)-Rényi differentially private, we have that

$$
D_\alpha\big(E_1^{\mathrm{DP}}(D_1')\,\|\,E_1^{\mathrm{DP}}(D_1)\big)\leq\epsilon\quad\text{and}\quad D_\alpha\big(E_1^{\mathrm{DP}}(D_1')\,\|\,E_1^{\mathrm{DP}}(D_1)\big)\leq\epsilon.
$$

Then, using the fact that the randomness in $E_1^{\mathrm{DP}}(\cdot)$ and $E_2^{\mathrm{DP}}(\cdot)$ (due to the noise) are independent,

$$\begin{aligned}&D_\alpha\left((E_1^{\mathrm{DP}}(D_1), E_1^{\mathrm{DP}}(D_2)) \,\|\, (E_1^{\mathrm{DP}}(D_1'), E_1^{\mathrm{DP}}(D_2'))\right)\\&= D_\alpha\left(E_1^{\mathrm{DP}}(D_1') \,\|\, E_1^{\mathrm{DP}}(D_1')\right) + D_\alpha\left(E_1^{\mathrm{DP}}(D_2) \,\|\, E_1^{\mathrm{DP}}(D_2')\right);\end{aligned}$$

and since we must be in one of the three cases delineated above, this must be at most ϵ. Hence, $(E_1^{\mathrm{DP}}(\cdot), E_2^{\mathrm{DP}}(\cdot))$ is (α, ϵ)-Rényi differentially private.

Now considering randomness over the data, since the two datasets are assumed to be independent it is immediate to see that the product is an e-value. Under the null,

$$\mathbb{E}[E_1^{\mathrm{DP}}(D_1) \cdot E_2^{\mathrm{DP}}(D_2)] = \mathbb{E}[E_1^{\mathrm{DP}}(D_1)] \cdot \mathbb{E}[E_2^{\mathrm{DP}}(D_2)] \leq 1 \cdot 1 = 1.$$

And, by the post-processing theorem, $(E_1^{\mathrm{DP}}(\cdot), E_1^{\mathrm{DP}}(\cdot) \cdot E_2^{\mathrm{DP}}(\cdot))$ is also (α, ϵ)-Rényi DP. □

Proposition A.6 (E-to-p conversion; Proposition 3.5 in the main text)**.** *If $E^{\mathrm{DP}}(D)$ is an (α, ϵ)-Rényi differentially private e-value for a null hypothesis H_0, then $1/E^{\mathrm{DP}}(D)$ is an (α, ϵ)-Rényi differentially private post-hoc valid p-value for H_0.*

Proof. By Theorem 2 of [Koning, 2023], $1/E^{\mathrm{DP}}(D)$ is a post-hoc valid p-value. All that remains is to show that it satisfies (ϵ, δ)-Rényi differential privacy, which follows directly by the post-processing theorem for $t \mapsto 1/t$. □

Proposition A.7 (Independent Averaging; Proposition 3.6 in the main text)**.** *If $E_1^{\mathrm{DP}}(D_1)$ and $E_2^{\mathrm{DP}}(D_2)$ are (α, ϵ)-Rényi differentially private e-values for a null hypothesis H_0, then for any $\eta \in [0, 1]$, $\eta E_1^{\mathrm{DP}}(D_1) + (1 - \eta)E_2^{\mathrm{DP}}(D_2)$ is also an (α, ϵ)-differentially private e-value for H_0.*

Proof. Let (D_1', D_2') be a neighboring dataset to (D_1, D_2), i.e., either (i) D_1' and D_1 differ by a single element and $D_2' = D_2$; (ii) D_2' and D_2 differ by a single element and $D_1' = D_1$; or (iii) $D_1' = D_1$ and $D_2' = D_2$. And since $E_1^{\mathrm{DP}}(\cdot)$ and $E_2^{\mathrm{DP}}(\cdot)$ are (α, ϵ)-Rényi differentially private, we have that

$$D_\alpha\big(E_1^{\mathrm{DP}}(D_1') \,\|\, E_1^{\mathrm{DP}}(D_1)\big) \leq \epsilon \quad \text{and} \quad D_\alpha\big(E_1^{\mathrm{DP}}(D_1') \,\|\, E_1^{\mathrm{DP}}(D_1)\big) \leq \epsilon.$$

Then, using the fact that the randomness in $E_1^{\mathrm{DP}}(\cdot)$ and $E_2^{\mathrm{DP}}(\cdot)$ (due to the noise) are independent,

$$\begin{aligned}&D_\alpha\left((E_1^{\mathrm{DP}}(D_1), E_1^{\mathrm{DP}}(D_2)) \,\|\, (E_1^{\mathrm{DP}}(D_1'), E_1^{\mathrm{DP}}(D_2'))\right)\\&= D_\alpha\left(E_1^{\mathrm{DP}}(D_1') \,\|\, E_1^{\mathrm{DP}}(D_1')\right) + D_\alpha\left(E_1^{\mathrm{DP}}(D_2) \,\|\, E_1^{\mathrm{DP}}(D_2')\right);\end{aligned}$$

and since we must be in one of the three cases delineated above, this must be at most ϵ. Hence, $(E_1^{\mathrm{DP}}(\cdot), E_2^{\mathrm{DP}}(\cdot))$ is (α, ϵ)-Rényi differentially private.

Now considering randomness over the data, it is immediate to see that the convex combination is an e-value: under the null,

$$\mathbb{E}[\eta E_1^{\mathrm{DP}}(D_1) + (1 - \eta)E_2^{\mathrm{DP}}(D_2)] = \eta\mathbb{E}[E_1^{\mathrm{DP}}(D_1)] + (1 - \eta)\mathbb{E}[E_2^{\mathrm{DP}}(D_2)] \leq \eta 1 + (1 - \eta)1 = 1.$$

And, by the post-processing theorem, $(E_1^{\mathrm{DP}}(\cdot), \eta E_1^{\mathrm{DP}}(\cdot) + (1 - \eta)E_2^{\mathrm{DP}}(\cdot))$ is also (α, ϵ)-Rényi DP. □

Proposition A.8 (Dependent Averaging; Proposition 3.7 in the main text)**.** *If $E_1^{\mathrm{DP}}(D)$ and $E_2^{\mathrm{DP}}(D)$ are (α, ϵ)-Rényi differentially private e-values for a null hypothesis H_0, then for any $\eta \in [0, 1]$, $\eta E_1^{\mathrm{DP}}(D) + (1 - \eta)E_2^{\mathrm{DP}}(D)$ is an $(\alpha, 2\epsilon)$-differentially private e-value for H_0.*

Proof. By the composition theorem, the release of $(E_1^{\mathrm{DP}}(D), E_2^{\mathrm{DP}}(D))$ is $(\alpha, 2\epsilon)$-RDP. The post-processing then ensures that the release of $\eta E_1^{\mathrm{DP}}(D) + (1 - \eta)E_2^{\mathrm{DP}}(D)$ is also $(\alpha, 2\epsilon)$-RDP.

The validity of the e-value is due to linearity: under the null,

$$\mathbb{E}[\eta E_1^{\mathrm{DP}}(D_1) + (1 - \eta)E_2^{\mathrm{DP}}(D_2)] = \eta\mathbb{E}[E_1^{\mathrm{DP}}(D_1)] + (1 - \eta)\mathbb{E}[E_2^{\mathrm{DP}}(D_2)] \leq \eta 1 + (1 - \eta)1 = 1.$$

□

Theorem A.9 (Theorem 3.9 in the main text)**.** *Under Assumption 3.8, let $\alpha > 1$ and $\epsilon > 0$. For each $j = 1, \dots, k$, let E_j^{DP} be an $(\alpha, \epsilon/k)$-Rényi differentially private e-value for the null $H_0^{(j)}$ (e.g., obtained through Theorems 3.1 and 3.2). Then $\mathcal{A}(E_1^{\mathrm{DP}}, \dots, E_k^{\mathrm{DP}})$ is valid and (α, ϵ)-Rényi differentially private.*

Proof. That $\mathcal{A}(E_1^{\mathrm{DP}}, \dots, E_k^{\mathrm{DP}})$ is valid is immediate: by Assumption 3.8 it suffices to have $E_1^{\mathrm{DP}}, \dots, E_k^{\mathrm{DP}}$ be valid, which we assume them to be.

As for privacy, this follows as an immediate use of the composition and post-processing theorems (Propositions 2.3 and 2.4): since $E_1^{\mathrm{DP}}, \dots, E_k^{\mathrm{DP}}$ are each $(\alpha, \epsilon/k)$-Rényi differentially private, the tuple $(E_1^{\mathrm{DP}}, \dots, E_k^{\mathrm{DP}})$ is (α, ϵ)-Rényi differentially private. Thus, by post-processing, $\mathcal{A}(E_1^{\mathrm{DP}}, \dots, E_k^{\mathrm{DP}})$ is also (α, ϵ)-Rényi differentially private. □

Corollary A.10 (Corollary 3.10 in the main text)**.** *Suppose $\log E_\theta(D)$ is locally Lipschitz in θ. Let $\alpha > 1$ and $\epsilon > 0$, and let $\widetilde{E}_{\theta_1}^{\mathrm{DP}}, \dots, \widetilde{E}_{\theta_k}^{\mathrm{DP}}$ be $(\alpha, \epsilon/k)$-Rényi differentially private versions of $\widetilde{E}_{\theta_1}, \dots, \widetilde{E}_{\theta_k}$ as defined in Equation 3 in the main text. Then $\mathrm{CI}_\alpha\big(\widetilde{E}_{\theta_1}^{\mathrm{DP}}(D), \dots, \widetilde{E}_{\theta_k}^{\mathrm{DP}}(D)\big)$ is an (α, ϵ)-Rényi differentially private confidence interval for $\theta^\star$, i.e., it satisfies (α, ϵ)-Rényi differential privacy, and*

$$\mathbb{P}\left[\theta^\star \in \mathrm{CI}_\alpha\Big(\widetilde{E}_{\theta_1}^{\mathrm{DP}}(D), \dots, \widetilde{E}_{\theta_k}^{\mathrm{DP}}(D)\Big)\right] \geq 1 - \alpha.$$

Proof. As noted in the main text, each $\widetilde{E}_{\theta_j}(D)$ is an simultaneously an e-value for all nulls $H_0 : \theta^\star = \theta'$, $\theta' \in [a_{j-1}, a_j]$. Thus, the $\widetilde{CI}_\alpha(\cdot, \dots, \cdot)$ algorithm constructs valid confidence intervals:

$$\begin{aligned}
\mathbb{P}\left[\theta^\star \in \mathrm{CI}_\alpha\Big(\widetilde{E}_{\theta_1}(D), \dots, \widetilde{E}_{\theta_k}(D)\Big)\right] &= \mathbb{P}\left[\theta^\star \in \bigcup_{\substack{j=1 \\ \widetilde{E}_{\theta_j}(D) \leq 1/\alpha}}^{k} [a_{j-1}, a_j]\right] \\
&\geq \mathbb{P}\left[\widetilde{E}_{\theta_{j'}}(D) \leq 1/\alpha\right], \quad \text{for } j' \text{ s.t. } \theta^\star \in [a_{j'-1}, a_{j'}] \\
&= 1 - \mathbb{P}\left[\widetilde{E}_{\theta_{j'}}(D) > 1/\alpha\right] \\
&\geq 1 - \frac{\mathbb{E}[\widetilde{E}_{\theta_{j'}}(D)]}{1/\alpha} \geq 1 - \frac{1}{1/\alpha} = 1 - \alpha.
\end{aligned}$$

By applying Theorem 3.9, we conclude. □

Proposition A.11 (Proposition 4.1 in the main text)**.** *Let F have support in $[\lambda_{\inf}, \lambda_{\sup}] \subset (-1/(1-\theta), 1/\theta)$. The log-sensitivity of $E_\theta(D)$, $\Delta_{\log}(E_\theta)$, is upper bounded as*

$$\begin{aligned}
\Delta_{\log}(E_\theta) \leq \max\big\{&\log(1 + \max\{\lambda_{\sup}(1-\theta), -\lambda_{\inf}\theta\}), \\
&- \log(1 + \min\{\lambda_{\inf}(1-\theta), -\lambda_{\sup}\theta\})\big\}.
\end{aligned}$$

Proof. Let D and D' be neighboring datasets. I.e., either D' is D with an additional sample Y_{n+1}, or it is D with a sample Y_i removed. Since the e-value $E_{D,\theta}$ is invariant to permutations of the data, we can, without loss of generality, consider that Y_n was the element removed.

When adding one new element, it follows:

$$\begin{aligned}
|\log E_{D,\theta} - \log E_{D',\theta}| &= \left|\sum_{i=1}^{n} \log\big(1 + \lambda_i(Y_i - \theta)\big) - \sum_{i=1}^{n+1} \log\big(1 + \lambda_i(Y_i - \theta)\big)\right| \\
&= \left|\log\big(1 + \lambda_{n+1}(Y_{n+1} - \theta)\big)\right|;
\end{aligned}$$

and, when removing an element – wlog. Y_n, we have

$$|\log E_{D,\theta} - \log E_{D',\theta}| = \left|\sum_{i=1}^{n} \log\big(1 + \lambda_i(Y_i - \theta)\big) - \sum_{i=1}^{n-1} \log\big(1 + \lambda_i(Y_i - \theta)\big)\right|$$

$$= \left|\log\big(1 + \lambda_n(Y_n - \theta)\big)\right|.$$

Hence, all that remains is to upper bound $\left|\log\big(1 + \lambda(Y - \theta)\big)\right|$, for any $\lambda \in (-c/(1-\theta), c/\theta)$ and $Y \in [0, 1]$. We split into a few cases:

1. When $Y \geq \theta$ and $\lambda \geq 0$, it follows:

$$\log\big(1 + \lambda(Y - \theta)\big) \leq \log\big(1 + \lambda(1 - \theta)\big) \leq \log\big(1 + \lambda_{\sup}(1 - \theta)\big)$$
$$\log\big(1 + \lambda(Y - \theta)\big) \geq \log\big(1 + \lambda 0\big) = \log 1 = 0;$$

2. When $Y \geq \theta$ and $\lambda \leq 0$, it follows:

$$\log\big(1 + \lambda(Y - \theta)\big) \leq \log\big(1 + \lambda 0\big) = \log 1 = 0$$
$$\log\big(1 + \lambda(Y - \theta)\big) \geq \log\big(1 + \lambda(1 - \theta)\big) \geq \log\big(1 + \lambda_{\inf}(1 - \theta)\big)$$

3. When $Y \leq \theta$ and $\lambda \geq 0$, it follows:

$$\log\big(1 + \lambda(Y - \theta)\big) \leq \log\big(1 + \lambda 0\big) = \log 1 = 0$$
$$\log\big(1 + \lambda(Y - \theta)\big) \geq \log\big(1 + \lambda(0 - \theta)\big) \geq \log\big(1 + \lambda_{\sup}(0 - \theta)\big) = \log\big(1 - \lambda_{\sup}\theta\big)$$

4. When $Y \leq \theta$ and $\lambda \leq 0$, it follows:

$$\log\big(1 + \lambda(Y - \theta)\big) \leq \log\big(1 + \lambda(0 - \theta)\big) \leq \log\big(1 + \lambda_{\inf}(0 - \theta)\big) \leq \log\big(1 - \lambda_{\inf}\theta\big)$$
$$\log\big(1 + \lambda(Y - \theta)\big) \geq \log\big(1 + \lambda 0\big) = \log 1 = 0;$$

Thus, we conclude that

$$\begin{aligned}
\left|\log\big(1 + \lambda(Y - \theta)\big)\right| &\leq \max\{0, |\log(1 + \lambda_{\sup}(1 - \theta))|, |\log(1 + \lambda_{\inf}(1 - \theta))|, \\
&\qquad |\log(1 - \lambda_{\sup}\theta)|, |\log(1 - \lambda_{\inf}\theta)|\} \\
&= \max\{|\log(1 + \lambda_{\sup}(1 - \theta))|, |\log(1 + \lambda_{\inf}(1 - \theta))|, \\
&\qquad |\log(1 - \lambda_{\sup}\theta)|, |\log(1 - \lambda_{\inf}\theta)|\}. \\
&= \max\big\{\log(1 + \max\{\lambda_{\sup}(1 - \theta), \lambda_{\inf}(1 - \theta), -\lambda_{\sup}\theta, -\lambda_{\inf}\theta\}), \\
&\qquad - \log(1 + \min\{\lambda_{\sup}(1 - \theta), \lambda_{\inf}(1 - \theta), -\lambda_{\sup}\theta, -\lambda_{\inf}\theta\})\big\}. \\
&\leq \max\big\{\log(1 + \max\{\lambda_{\sup}(1 - \theta), -\lambda_{\inf}\theta\}), \\
&\qquad - \log(1 + \min\{\lambda_{\inf}(1 - \theta), -\lambda_{\sup}\theta\})\big\}.
\end{aligned}$$

□

Proposition A.12 (Proposition 4.2 in the main text)**.** *For distributions F with support contained in $[\lambda_{\inf}, \lambda_{\sup}]$, the Lipschitz constant of $\log E_\theta(D)$ over $\theta \in [\theta_{\inf}, \theta_{\sup}]$ is upper bounded by*

$$\|\theta \mapsto \log E_\theta(D)\|_{\text{Lip}} \leq \max\left\{\left|\frac{\lambda_{\sup}}{1 - \lambda_{\sup}\theta_{\sup}}\right|, \left|\frac{\lambda_{\inf}}{1 + \lambda_{\inf}(1 - \theta_{\inf})}\right|\right\}.$$

Proof. Bounding the Lipschitz constant is equivalent to bounding the absolute value of the derivative. So it follows:

$$\begin{aligned}
\left|\frac{\mathrm{d}}{\mathrm{d}\theta}\log E_\theta(D)\right| &= \left|\frac{\frac{\mathrm{d}}{\mathrm{d}\theta}E_\theta(D)}{E_\theta(D)}\right| = \left|\frac{\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbb{E}_{\lambda\sim F}[\prod_{i=1}^n(1 + \lambda(Y_i - \theta))]}{E_\theta(D)}\right| \\
&= \left|\frac{\mathbb{E}_{\lambda\sim F}[\frac{\mathrm{d}}{\mathrm{d}\theta}\prod_{i=1}^n(1 + \lambda(Y_i - \theta))]}{E_\theta(D)}\right|
\end{aligned}$$

$$= \left| \frac{\mathbb{E}_{\lambda\sim F}[\frac{\mathrm{d}}{\mathrm{d}\theta}\exp(\sum_{i=1}^n \log(1+\lambda(Y_i-\theta)))]}{E_\theta(D)} \right|$$
$$= \left| \frac{\mathbb{E}_{\lambda\sim F}[\exp(\sum_{i=1}^n \log(1+\lambda(Y_i-\theta)))\frac{\mathrm{d}}{\mathrm{d}\theta}\sum_{i=1}^n \log(1+\lambda(Y_i-\theta))]}{E_\theta(D)} \right|$$

Write $E_\theta^{(\lambda)}(D) = \prod_{i=1}^n (1+\lambda(Y_i-\theta))$. Then:

$$\left| \frac{\mathbb{E}_{\lambda\sim F}[\exp(\sum_{i=1}^n \log(1+\lambda(Y_i-\theta)))\frac{\mathrm{d}}{\mathrm{d}\theta}\sum_{i=1}^n \log(1+\lambda(Y_i-\theta))]}{E_\theta(D)} \right|$$
$$= \left| \frac{\mathbb{E}_{\lambda\sim F}[E_\theta^{(\lambda)}(D)\frac{\mathrm{d}}{\mathrm{d}\theta}\sum_{i=1}^n \log(1+\lambda(Y_i-\theta))]}{E_\theta(D)} \right|$$
$$= \left| \frac{\mathbb{E}_{\lambda\sim F}[E_\theta^{(\lambda)}(D)\frac{\mathrm{d}}{\mathrm{d}\theta}\sum_{i=1}^n \log(1+\lambda(Y_i-\theta))]}{\mathbb{E}_{\lambda\sim F}[E_\theta^{(\lambda)}(D)]} \right|;$$

Now let $F_{n,\theta}$ be the distribution such that $\mathrm{d}F_{n,\theta}/\mathrm{d}F \propto E_\theta^{(\lambda)}(D)$. This allows us to rewrite the ratio as

$$\left| \frac{\mathbb{E}_{\lambda\sim F}[E_\theta^{(\lambda)}(D)\frac{\mathrm{d}}{\mathrm{d}\theta}\sum_{i=1}^n \log(1+\lambda(Y_i-\theta))]}{\mathbb{E}_{\lambda\sim F}[E_\theta^{(\lambda)}(D)]} \right|$$
$$= \left| \mathbb{E}_{\lambda\sim F_{n,\theta}}\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\sum_{i=1}^n \log(1+\lambda(Y_i-\theta))\right] \right|$$
$$= \left| \mathbb{E}_{\lambda\sim F_{n,\theta}}\left[\sum_{i=1}^n \frac{-\lambda}{1+\lambda(Y_i-\theta)}\right] \right|.$$

So all that remains is to bound this. Now, note that $-\lambda/(1+\lambda(Y_i-\theta))$ is (i) increasing in $Y-\theta$ and (ii) decreasing in λ. Thus we have

$$-\frac{\lambda_{\sup}}{1-\lambda_{\sup}\theta_{\sup}} \leq -\lambda/(1+\lambda(Y_i-\theta)) \leq -\frac{\lambda_{\inf}}{1+\lambda_{\inf}(1-\theta_{\inf})},$$

and

$$\left| \mathbb{E}_{\lambda\sim F_{n,\theta}}\left[\sum_{i=1}^n \frac{-\lambda}{1+\lambda(Y_i-\theta)}\right] \right| \leq \max\left\{ \left| -\frac{\lambda_{\sup}}{1-\lambda_{\sup}\theta_{\sup}} \right|, \left| -\frac{\lambda_{\inf}}{1+\lambda_{\inf}(1-\theta_{\inf})} \right| \right\}$$
$$= \max\left\{ \left| \frac{\lambda_{\sup}}{1-\lambda_{\sup}\theta_{\sup}} \right|, \left| \frac{\lambda_{\inf}}{1+\lambda_{\inf}(1-\theta_{\inf})} \right| \right\}.$$

□

Proposition A.13 (Proposition 4.3 in the main text)**.** *Suppose the scores are all contained in $[a,b]$ with $a>0$. Then the log-sensitivity of $E^{\mathrm{exch}}(D;S^{\mathrm{test}})$, $\Delta_{\log}(E^{\mathrm{exch}})$, is upper bounded as*

$$\Delta_{\log}(E^{\mathrm{exch}}) \leq 2\cdot\frac{b/a}{n+1}.$$

Proof. We want to compute $\Delta_{\log}(E^{\mathrm{exch}}) = \sup_{|D\Delta D'|\leq 1}|\log E^{\mathrm{exch}}(D;S^{\mathrm{test}}) - \log E^{\mathrm{exch}}(D';S^{\mathrm{test}})|$. For simplicity, we will rewrite the e-value as

$$E^{\mathrm{exch}}(D;S^{\mathrm{test}}) = \frac{S^{\mathrm{test}}}{\frac{1}{n+1}\left(\sum_{i=1}^n S_i + S^{\mathrm{test}}\right)}$$

There are three cases:

(i) $D=D'$, in which case $|\log E^{\mathrm{exch}}(D;S^{\mathrm{test}}) - \log E^{\mathrm{exch}}(D';S^{\mathrm{test}})| = 0$.

(ii) D' differs from D by one new element, Y_{n+1}; in this case we have

$$
\begin{aligned}
&|\log E^{\text{exch}}(D; S^{\text{test}}) - \log E^{\text{exch}}(D'; S^{\text{test}})| \\
&= \left| \left(\log S^{\text{test}} - \log \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} \right) - \left(\log S^{\text{test}} - \log \frac{\sum_{i=1}^{n+1} S_i + S^{\text{test}}}{n+2} \right) \right| \\
&= \left| \log \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} - \log \frac{\sum_{i=1}^{n+1} S_i + S^{\text{test}}}{n+2} \right| \\
&\leq \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} - \frac{\sum_{i=1}^{n+1} S_i + S^{\text{test}}}{n+2} \right| \Big/ \left(\frac{\sum_{i=1}^{n} a + a}{n+2} \right) \\
&= \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} - \frac{\sum_{i=1}^{n+1} S_i + S^{\text{test}}}{n+2} \right| \Big/ \left(\frac{n+1}{n+2} a \right) \\
&= \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} \cdot \frac{n+2}{n+1} - \frac{\sum_{i=1}^{n+1} S_i + S^{\text{test}}}{n+2} \cdot \frac{n+2}{n+1} \right| / a \\
&= \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} \cdot \frac{n+2}{n+1} - \frac{\sum_{i=1}^{n+1} S_i + S^{\text{test}}}{n+1} \right| / a \\
&= \left| \sum_{i=1}^{n} \left(\frac{1}{n+1} \cdot \frac{n+2}{n+1} - \frac{1}{n+1} \right) S_i + \left(\frac{1}{n+1} \cdot \frac{n+2}{n+1} - \frac{1}{n+1} \right) S^{\text{test}} - \frac{S_{n+1}}{n+1} \right| / a \\
&= \left| \sum_{i=1}^{n} \left(\frac{n+2}{n+1} - 1 \right) S_i + \left(\frac{n+2}{n+1} - 1 \right) S^{\text{test}} - S_{n+1} \right| / a(n+1) \\
&\leq \left(\left(\frac{n+2}{n+1} - 1 \right) \left| \sum_{i=1}^{n} S_i + S^{\text{test}} \right| + |S_{n+1}| \right) / a(n+1) \\
&= \left(\frac{1}{n+1} \left| \sum_{i=1}^{n} S_i + S^{\text{test}} \right| + |S_{n+1}| \right) / a(n+1) \\
&\leq \left(\frac{1}{n+1} \left| \sum_{i=1}^{n} b + b \right| + b \right) / a(n+1) \\
&= (b+b) / a(n+1) = 2b/a(n+1) = 2 \cdot \frac{b/a}{n+1}.
\end{aligned}
$$

(iii) D' differs from D by one less element; in this case we have

$$
\begin{aligned}
&|\log E^{\text{exch}}(D; S^{\text{test}}) - \log E^{\text{exch}}(D'; S^{\text{test}})| \\
&= \left| \left(\log S^{\text{test}} - \log \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} \right) - \left(\log S^{\text{test}} - \log \frac{\sum_{i=1}^{n-1} S_i + S^{\text{test}}}{n} \right) \right| \\
&= \left| \log \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} - \log \frac{\sum_{i=1}^{n-1} S_i + S^{\text{test}}}{n} \right| \\
&\leq \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} - \frac{\sum_{i=1}^{n-1} S_i + S^{\text{test}}}{n} \right| \Big/ \left(\frac{\sum_{i=1}^{n-1} a + a}{n+1} \right) \\
&= \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} - \frac{\sum_{i=1}^{n-1} S_i + S^{\text{test}}}{n} \right| \Big/ \left(\frac{n}{n+1} a \right) \\
&= \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} \cdot \frac{n}{n+1} - \frac{\sum_{i=1}^{n-1} S_i + S^{\text{test}}}{n} \cdot \frac{n}{n+1} \right| / a \\
&= \left| \frac{\sum_{i=1}^{n} S_i + S^{\text{test}}}{n+1} \cdot \frac{n}{n+1} - \frac{\sum_{i=1}^{n-1} S_i + S^{\text{test}}}{n+1} \right| / a
\end{aligned}
$$

$$
\begin{aligned}
&= \left| \sum_{i=1}^{n-1} \left(\frac{1}{n+1} \cdot \frac{n}{n+1} - \frac{1}{n+1} \right) S_i + \left(\frac{1}{n+1} \cdot \frac{n}{n+1} - \frac{1}{n+1} \right) S^{\text{test}} + \frac{S_n}{n+1} \right| / a \\
&= \left| \sum_{i=1}^{n-1} \left(\frac{n}{n+1} - 1 \right) S_i + \left(\frac{n}{n+1} - 1 \right) S^{\text{test}} + S_n \right| / a(n+1) \\
&\leq \left(\left(1 - \frac{n}{n+1} \right) \left| \sum_{i=1}^{n-1} S_i + S^{\text{test}} \right| + |S_n| \right) / a(n+1) \\
&= \left(\frac{1}{n+1} \left| \sum_{i=1}^{n-1} S_i + S^{\text{test}} \right| + |S_n| \right) / a(n+1) \\
&\leq \left(\frac{1}{n+1} \left| \sum_{i=1}^{n-1} b + b \right| + b \right) / a(n+1) \\
&= \left(\frac{n}{n+1} b + b \right) / a(n+1) \\
&\leq 2b/a(n+1) = 2 \cdot \frac{b/a}{n+1}.
\end{aligned}
$$

□

B Additional theoretical results

B.1 (ϵ, δ)-differential privacy

Theorem B.1 ((ϵ, δ)-DP Biased Gaussian mechanism for a single e-value)**.** *For any $\epsilon, \delta > 0$, let $E^{\text{DP}}(D)$ be as in Equation 1 in the main text with $\xi \sim \mathcal{N}(c^2[\Delta_{\log}(E)]^2/2\epsilon^2, c^2[\Delta_{\log}(E)]^2/\epsilon^2)$, for $c^2 = 2\log 1.25/\delta$. Then $E^{\text{DP}}(D)$ is a valid e-value satisfying (ϵ, δ)-differential privacy.*

Proof. First note that for any bias μ, $E^{\text{DP}}(D)$ with $\xi \sim \mathcal{N}(\mu, c^2[\Delta_{\log}(E)]^2/\epsilon^2)$ is (ϵ, δ)-differentially private: by the post-processing theorem, $E^{\text{DP}}(D) = E(D) \cdot e^{-\xi}$ is (ϵ, δ)-DP iff $\log E^{\text{DP}}(D) = \log \left(E(D) \cdot e^{-\xi} \right) = \log E(D) - \xi$ is (ϵ, δ)-DP iff $\log E(D) - (\xi - \mu)$ is (ϵ, δ)-DP; and, by Theorem A.1 of [Dwork and Roth, 2014], the latter is (ϵ, δ)-DP.

Thus, all that remains is to find the smallest bias μ that ensures validity of $E^{\text{DP}}(D)$. As argued in Section 3.1, all we need is that the MGF of ξ is at most one at -1. So it follows, using the form of the MGF of an univariate normal:

$$
\begin{aligned}
&\mathbb{E}[e^{-\xi}] = \exp\left(-\mu + \frac{c^2[\Delta_{\log}(E)]^2/\epsilon^2}{2} \right) \leq 1 \\
&\iff -\mu + \frac{c^2[\Delta_{\log}(E)]^2/\epsilon^2}{2} \leq 0 \\
&\iff \mu \geq \frac{c^2[\Delta_{\log}(E)]^2/\epsilon^2}{2} = \frac{c^2[\Delta_{\log}(E)]^2}{2\epsilon^2}.
\end{aligned}
$$

□

Theorem B.2 ($(\epsilon, 0)$-DP Biased Laplace mechanism for a single e-value)**.** *For any $\epsilon > 0$ satisfying $\epsilon > \Delta_{\log}(E)$, let $E^{\text{DP}}(D)$ be as in Equation 1 in the main text with $\xi \sim \text{Laplace}(-\log(1 - [\Delta_{\log}(E)]^2/\epsilon^2), \Delta_{\log}(E)/\epsilon)$. Then $E^{\text{DP}}(D)$ is a valid e-value satisfying $(\epsilon, 0)$-differential privacy.*

Proof. First note that for any bias μ, $E^{\text{DP}}(D)$ with $\xi \sim \text{Laplace}(\mu, \Delta_{\log}(E)/\epsilon)$ is $(\epsilon, 0)$-differentially private: by the post-processing theorem, $E^{\text{DP}}(D) = E(D) \cdot e^{-\xi}$ is $(\epsilon, 0)$-DP iff $\log E^{\text{DP}}(D) = \log \left(E(D) \cdot e^{-\xi} \right) = \log E(D) - \xi$ is $(\epsilon, 0)$-DP iff $\log E(D) - (\xi - \mu)$ is $(\epsilon, 0)$-DP; and, by Theorem 3.6 of [Dwork and Roth, 2014], the latter is $(\epsilon, 0)$-DP.

Thus, all that remains is to find the smallest bias μ that ensures validity of $E^{\text{DP}}(D)$. As argued in Section 3.1, all we need is that the MGF of ξ is at most one at -1. The MGF of a Laplace distribution with mean μ and

scale parameter b is given by $t \mapsto e^{t\mu}/(1-b^2t^2)$, defined at $|t| < 1/b$. Thus we need that

$$|-1| = 1 < \frac{1}{\Delta_{\log}(E)/\epsilon} = \frac{\epsilon}{\Delta_{\log}(E)} \iff \Delta_{\log}(E) < \epsilon.$$

As long as this is satisfied, it then follows:

$$\begin{aligned}
&\mathbb{E}[e^{-\xi}] = e^{-\mu}/(1-[\Delta_{\log}(E)]^2/\epsilon^2) \leq 1 \\
&\iff e^{-\mu} \leq 1-[\Delta_{\log}(E)]^2/\epsilon^2 \\
&\iff -\mu \leq \log(1-[\Delta_{\log}(E)]^2/\epsilon^2) \\
&\iff \mu \geq -\log(1-[\Delta_{\log}(E)]^2/\epsilon^2).
\end{aligned}$$

□

B.2 Bias is necessary

The following is a simple result showing that a biased mechanism is necessary for differentially-private e-values. We consider here that our non-private e-value has expectation exactly equal to 1 and that $E(D)$ is not already differentially private itself.

Proposition B.3 (Bias is necessary)**.** *Let $\mathbb{E}[E(D)] = 1$ and assume that $E(D)$ is not already differentially private (either (ϵ, δ)-DP or (α, ϵ)-DP). Then for $E^{\mathrm{DP}}(D)$ (as per Equation 1 in the main text) to be a differentially private e-value, we must take some ξ with $\mathbb{E}[\xi] > 0$.*

Proof. Since $E(D)$ is not already differentially private, for $E^{\mathrm{DP}}(D)$ to be differentially private ξ must not be constant. Then, for $E^{\mathrm{DP}}(D)$ to be a valid e-value, we require that, under the null,

$$\mathbb{E}[E^{\mathrm{DP}}(D)] = \mathbb{E}[E(D) \cdot e^{-\xi}] = \mathbb{E}[E(D)] \cdot \mathbb{E}[e^{-\xi}] = \mathbb{E}[e^{-\xi}] \leq 1;$$

and, since $s \mapsto e^{-s}$ is strictly convex, using the strict variant of Jensen's inequality (which holds for strictly convex functions and non-constant random variables),

$$1 \geq \mathbb{E}[e^{-\xi}] > e^{-\mathbb{E}[\xi]}.$$

Taking the log on both sides, we conclude that

$$\mathbb{E}[\xi] > 0.$$

□

C Additional experimental results

C.1 Comparison with the test-of-tests

In this section we compare our framework to the private hypothesis testing method of [Kazan et al., 2023] (the "test-of-tests"). Their work takes as input a (non-private) hypothesis test and privatizes it, guaranteeing $(\epsilon, 0)$-differential privacy.

We consider here hypothesis tests for nulls $H_0^{(p)} : p^\star = p$, for data $Y_1, \ldots, Y_n \sim \mathrm{Bern}(p^\star)$ with $p^\star = 0.5$. For this, we have two base hypothesis tests: (i) the e-value test for the mean [Waudby-Smith and Ramdas, 2020], from which we can get a p-value by taking the reciprocal; and (ii) the exact p-value given by a binomial test. We compute both of these non-private p-values. As for private procedures, we compute:

- The test-of-tests [Kazan et al., 2023] atop the reciprocal of the e-value for the mean;
- The test-of-tests [Kazan et al., 2023] atop the exact p-value;
- Our method for $(\epsilon, 0)$-differential privacy with Laplace noise (Theorem B.2); and
- Our method for (ϵ, δ)-differential privacy with Gaussian noise (Theorem B.1), for a fixed $\delta = 0.01$.

The results can be seen in Tables 1-4, for varying values of ϵ and of c (the latter being a hyperparameter in the e-value for the mean). Throughout, we note the better performance of our method, achieving much lower p-values even when compared to the test-of-tests atop the exact test. However, it should be noted that for our biased Laplace mechanism to be defined (which is required in order to obtain the $(\epsilon, 0)$-differential privacy analogous to the test-of-tests), we need a sufficiently low c.

Table 1: $\epsilon = 0.1$, $c = 0.2$

p	1/e-value (non-private)	Exact p-value (non-private)	Ours, Laplace	Ours, Gauss. ($\delta = 0.01$)	ToT atop 1/e-value	ToT atop exact p-value
0.05	$4.52 \cdot 10^{-68}$	$5.81 \cdot 10^{-209}$	N/A	$\mathbf{6.89 \cdot 10^{-60}}$	$1.70 \cdot 10^{-02}$	$2.22 \cdot 10^{-02}$
0.10	$2.47 \cdot 10^{-60}$	$1.34 \cdot 10^{-161}$	N/A	$\mathbf{3.93 \cdot 10^{-47}}$	$1.74 \cdot 10^{-01}$	$3.83 \cdot 10^{-01}$
0.15	$1.92 \cdot 10^{-49}$	$1.90 \cdot 10^{-110}$	N/A	$\mathbf{5.44 \cdot 10^{-36}}$	$1.50 \cdot 10^{-01}$	$2.37 \cdot 10^{-02}$
0.20	$7.00 \cdot 10^{-45}$	$6.20 \cdot 10^{-93}$	N/A	$\mathbf{4.68 \cdot 10^{-33}}$	$6.46 \cdot 10^{-01}$	$6.66 \cdot 10^{-01}$
0.25	$3.45 \cdot 10^{-30}$	$1.23 \cdot 10^{-48}$	N/A	$\mathbf{1.17 \cdot 10^{-22}}$	$6.72 \cdot 10^{-01}$	$9.06 \cdot 10^{-01}$
0.30	$1.56 \cdot 10^{-26}$	$4.10 \cdot 10^{-40}$	N/A	$\mathbf{2.72 \cdot 10^{-21}}$	$1.14 \cdot 10^{-02}$	$5.56 \cdot 10^{-01}$
0.35	$4.46 \cdot 10^{-18}$	$9.59 \cdot 10^{-24}$	N/A	$\mathbf{2.99 \cdot 10^{-09}}$	$9.74 \cdot 10^{-01}$	$2.76 \cdot 10^{-01}$
0.40	$2.17 \cdot 10^{-06}$	$5.95 \cdot 10^{-08}$	N/A	$1.00 \cdot 10^{+00}$	$\mathbf{2.38 \cdot 10^{-01}}$	$6.05 \cdot 10^{-01}$
0.45	$2.80 \cdot 10^{-02}$	$1.39 \cdot 10^{-03}$	N/A	$1.00 \cdot 10^{+00}$	$7.63 \cdot 10^{-01}$	$\mathbf{7.61 \cdot 10^{-01}}$
0.50	$1.00 \cdot 10^{+00}$	$4.67 \cdot 10^{-01}$	N/A	$1.00 \cdot 10^{+00}$	$\mathbf{4.24 \cdot 10^{-01}}$	$8.59 \cdot 10^{-01}$
0.55	$1.00 \cdot 10^{+00}$	$9.37 \cdot 10^{-02}$	N/A	$1.00 \cdot 10^{+00}$	$9.33 \cdot 10^{-01}$	$\mathbf{7.30 \cdot 10^{-01}}$
0.60	$8.65 \cdot 10^{-09}$	$1.18 \cdot 10^{-10}$	N/A	$1.00 \cdot 10^{+00}$	$5.91 \cdot 10^{-01}$	$\mathbf{2.96 \cdot 10^{-02}}$
0.65	$1.08 \cdot 10^{-15}$	$6.29 \cdot 10^{-20}$	N/A	$\mathbf{7.70 \cdot 10^{-05}}$	$7.09 \cdot 10^{-01}$	$2.24 \cdot 10^{-01}$
0.70	$3.46 \cdot 10^{-26}$	$2.38 \cdot 10^{-39}$	N/A	$\mathbf{1.75 \cdot 10^{-10}}$	$3.93 \cdot 10^{-01}$	$2.90 \cdot 10^{-01}$
0.75	$5.00 \cdot 10^{-34}$	$1.35 \cdot 10^{-58}$	N/A	$\mathbf{1.69 \cdot 10^{-23}}$	$6.35 \cdot 10^{-01}$	$1.79 \cdot 10^{-01}$
0.80	$6.68 \cdot 10^{-42}$	$1.56 \cdot 10^{-82}$	N/A	$\mathbf{2.52 \cdot 10^{-33}}$	$5.25 \cdot 10^{-01}$	$2.61 \cdot 10^{-01}$
0.85	$2.88 \cdot 10^{-49}$	$9.78 \cdot 10^{-110}$	N/A	$\mathbf{4.77 \cdot 10^{-34}}$	$8.10 \cdot 10^{-02}$	$1.57 \cdot 10^{-01}$
0.90	$4.31 \cdot 10^{-62}$	$2.24 \cdot 10^{-171}$	N/A	$\mathbf{5.29 \cdot 10^{-45}}$	$4.16 \cdot 10^{-01}$	$8.74 \cdot 10^{-01}$
0.95	$2.01 \cdot 10^{-68}$	$2.00 \cdot 10^{-211}$	N/A	$\mathbf{2.31 \cdot 10^{-53}}$	$1.29 \cdot 10^{-01}$	$9.41 \cdot 10^{-02}$

Table 2: $\epsilon = 0.5$, $c = 0.2$

p	1/e-value (non-private)	Exact p-value (non-private)	Ours, Laplace	Ours, Gauss. ($\delta = 0.01$)	ToT atop 1/e-value	ToT atop exact p-value
0.05	$5.97 \cdot 10^{-69}$	$3.48 \cdot 10^{-215}$	$\mathbf{4.87 \cdot 10^{-69}}$	$9.60 \cdot 10^{-68}$	$5.57 \cdot 10^{-03}$	$1.15 \cdot 10^{-02}$
0.10	$4.88 \cdot 10^{-61}$	$1.89 \cdot 10^{-165}$	$3.14 \cdot 10^{-60}$	$\mathbf{3.24 \cdot 10^{-61}}$	$3.04 \cdot 10^{-03}$	$1.33 \cdot 10^{-02}$
0.15	$1.64 \cdot 10^{-47}$	$8.64 \cdot 10^{-103}$	$\mathbf{8.51 \cdot 10^{-48}}$	$3.78 \cdot 10^{-47}$	$1.47 \cdot 10^{-02}$	$5.43 \cdot 10^{-02}$
0.20	$2.35 \cdot 10^{-44}$	$4.85 \cdot 10^{-91}$	$\mathbf{3.46 \cdot 10^{-44}}$	$2.13 \cdot 10^{-43}$	$6.67 \cdot 10^{-03}$	$3.11 \cdot 10^{-03}$
0.25	$3.11 \cdot 10^{-31}$	$3.03 \cdot 10^{-51}$	$8.49 \cdot 10^{-30}$	$\mathbf{4.23 \cdot 10^{-31}}$	$5.30 \cdot 10^{-03}$	$3.91 \cdot 10^{-03}$
0.30	$1.45 \cdot 10^{-17}$	$6.78 \cdot 10^{-23}$	$1.05 \cdot 10^{-17}$	$\mathbf{5.91 \cdot 10^{-18}}$	$5.59 \cdot 10^{-03}$	$2.11 \cdot 10^{-02}$
0.35	$1.28 \cdot 10^{-19}$	$2.14 \cdot 10^{-26}$	$1.66 \cdot 10^{-19}$	$\mathbf{1.24 \cdot 10^{-19}}$	$3.05 \cdot 10^{-02}$	$1.11 \cdot 10^{-03}$
0.40	$2.36 \cdot 10^{-11}$	$8.29 \cdot 10^{-14}$	$3.01 \cdot 10^{-11}$	$\mathbf{2.75 \cdot 10^{-11}}$	$9.12 \cdot 10^{-01}$	$8.38 \cdot 10^{-02}$
0.45	$1.20 \cdot 10^{-02}$	$5.61 \cdot 10^{-04}$	$\mathbf{1.69 \cdot 10^{-02}}$	$1.99 \cdot 10^{-01}$	$6.30 \cdot 10^{-02}$	$5.19 \cdot 10^{-01}$
0.50	$1.00 \cdot 10^{+00}$	$6.35 \cdot 10^{-01}$	$1.00 \cdot 10^{+00}$	$1.00 \cdot 10^{+00}$	$7.69 \cdot 10^{-01}$	$\mathbf{7.52 \cdot 10^{-01}}$
0.55	$4.18 \cdot 10^{-02}$	$2.14 \cdot 10^{-03}$	$6.37 \cdot 10^{-02}$	$\mathbf{1.21 \cdot 10^{-02}}$	$3.91 \cdot 10^{-01}$	$9.35 \cdot 10^{-01}$
0.60	$2.03 \cdot 10^{-09}$	$2.13 \cdot 10^{-11}$	$\mathbf{1.49 \cdot 10^{-09}}$	$6.53 \cdot 10^{-09}$	$9.07 \cdot 10^{-02}$	$3.05 \cdot 10^{-01}$
0.65	$7.99 \cdot 10^{-21}$	$1.45 \cdot 10^{-28}$	$9.04 \cdot 10^{-21}$	$\mathbf{4.62 \cdot 10^{-21}}$	$1.41 \cdot 10^{-01}$	$7.15 \cdot 10^{-02}$
0.70	$3.14 \cdot 10^{-27}$	$1.15 \cdot 10^{-41}$	$\mathbf{5.02 \cdot 10^{-27}}$	$1.10 \cdot 10^{-25}$	$8.66 \cdot 10^{-03}$	$1.28 \cdot 10^{-02}$
0.75	$4.00 \cdot 10^{-36}$	$1.70 \cdot 10^{-64}$	$\mathbf{3.28 \cdot 10^{-36}}$	$1.53 \cdot 10^{-35}$	$2.27 \cdot 10^{-03}$	$8.96 \cdot 10^{-03}$
0.80	$3.36 \cdot 10^{-41}$	$3.37 \cdot 10^{-80}$	$\mathbf{1.27 \cdot 10^{-41}}$	$1.23 \cdot 10^{-40}$	$4.51 \cdot 10^{-02}$	$3.13 \cdot 10^{-02}$
0.85	$3.37 \cdot 10^{-51}$	$9.60 \cdot 10^{-118}$	$5.67 \cdot 10^{-51}$	$\mathbf{5.27 \cdot 10^{-51}}$	$1.14 \cdot 10^{-02}$	$2.01 \cdot 10^{-02}$
0.90	$1.60 \cdot 10^{-57}$	$6.11 \cdot 10^{-147}$	$1.95 \cdot 10^{-57}$	$\mathbf{3.18 \cdot 10^{-58}}$	$2.69 \cdot 10^{-03}$	$7.39 \cdot 10^{-03}$
0.95	$3.98 \cdot 10^{-69}$	$1.86 \cdot 10^{-216}$	$\mathbf{4.23 \cdot 10^{-69}}$	$2.99 \cdot 10^{-68}$	$4.41 \cdot 10^{-03}$	$3.18 \cdot 10^{-03}$

D Experiment details

D.1 Section 4.1 in the main text

We use the e-value for the mean with distribution $F \sim \text{Unif}(-1, 1)$ independent of θ. We use the first 100 000 samples of our dataset.

For the plot and confidence interval algorithm, we compute e-values for the mean over $k = 50$ values of θ, evenly spaced within $(0, 1)$. For the confidence intervals we use a significance level of $\alpha = 0.05$.

Table 3: $\epsilon = 0.1$, $c = 0.05$

p	1/e-value (non-private)	Exact p-value (non-private)	Ours, Laplace	Ours, Gauss. ($\delta = 0.01$)	ToT atop 1/e-value	ToT atop exact p-value
0.05	$2.08 \cdot 10^{-18}$	$1.08 \cdot 10^{-229}$	$\mathbf{3.26 \cdot 10^{-18}}$	$4.24 \cdot 10^{-17}$	$9.00 \cdot 10^{-01}$	$2.58 \cdot 10^{-01}$
0.10	$2.74 \cdot 10^{-15}$	$1.33 \cdot 10^{-149}$	$\mathbf{2.63 \cdot 10^{-15}}$	$3.67 \cdot 10^{-14}$	$7.91 \cdot 10^{-01}$	$3.76 \cdot 10^{-02}$
0.15	$6.63 \cdot 10^{-13}$	$1.57 \cdot 10^{-105}$	$\mathbf{2.52 \cdot 10^{-13}}$	$5.72 \cdot 10^{-12}$	$2.38 \cdot 10^{-01}$	$1.41 \cdot 10^{-02}$
0.20	$5.68 \cdot 10^{-12}$	$4.85 \cdot 10^{-91}$	$6.95 \cdot 10^{-12}$	$\mathbf{2.26 \cdot 10^{-13}}$	$6.64 \cdot 10^{-04}$	$3.65 \cdot 10^{-02}$
0.25	$2.78 \cdot 10^{-10}$	$4.26 \cdot 10^{-68}$	$\mathbf{4.06 \cdot 10^{-10}}$	$1.84 \cdot 10^{-09}$	$8.57 \cdot 10^{-01}$	$9.85 \cdot 10^{-02}$
0.30	$8.87 \cdot 10^{-07}$	$3.58 \cdot 10^{-32}$	$1.71 \cdot 10^{-06}$	$\mathbf{8.86 \cdot 10^{-08}}$	$1.82 \cdot 10^{-02}$	$1.81 \cdot 10^{-02}$
0.35	$7.30 \cdot 10^{-05}$	$2.10 \cdot 10^{-18}$	$\mathbf{5.06 \cdot 10^{-05}}$	$1.03 \cdot 10^{-04}$	$9.86 \cdot 10^{-01}$	$5.53 \cdot 10^{-02}$
0.40	$3.91 \cdot 10^{-03}$	$3.02 \cdot 10^{-09}$	$6.99 \cdot 10^{-03}$	$\mathbf{1.46 \cdot 10^{-03}}$	$7.44 \cdot 10^{-01}$	$2.58 \cdot 10^{-01}$
0.45	$2.41 \cdot 10^{-01}$	$7.16 \cdot 10^{-03}$	$7.69 \cdot 10^{-01}$	$1.00 \cdot 10^{+00}$	$\mathbf{2.98 \cdot 10^{-01}}$	$5.13 \cdot 10^{-01}$
0.50	$1.00 \cdot 10^{+00}$	$9.75 \cdot 10^{-01}$	$1.00 \cdot 10^{+00}$	$1.00 \cdot 10^{+00}$	$\mathbf{5.67 \cdot 10^{-01}}$	$6.03 \cdot 10^{-01}$
0.55	$3.38 \cdot 10^{-01}$	$1.77 \cdot 10^{-02}$	$5.55 \cdot 10^{-01}$	$5.71 \cdot 10^{-01}$	$\mathbf{4.23 \cdot 10^{-02}}$	$4.66 \cdot 10^{-02}$
0.60	$3.91 \cdot 10^{-03}$	$3.02 \cdot 10^{-09}$	$5.88 \cdot 10^{-03}$	$\mathbf{5.58 \cdot 10^{-04}}$	$7.84 \cdot 10^{-01}$	$4.34 \cdot 10^{-01}$
0.65	$2.63 \cdot 10^{-05}$	$3.01 \cdot 10^{-21}$	$4.49 \cdot 10^{-05}$	$\mathbf{1.66 \cdot 10^{-05}}$	$4.90 \cdot 10^{-01}$	$2.57 \cdot 10^{-01}$
0.70	$9.91 \cdot 10^{-08}$	$1.69 \cdot 10^{-40}$	$3.86 \cdot 10^{-07}$	$\mathbf{3.39 \cdot 10^{-07}}$	$4.10 \cdot 10^{-02}$	$1.57 \cdot 10^{-01}$
0.75	$3.71 \cdot 10^{-10}$	$1.54 \cdot 10^{-66}$	$\mathbf{1.45 \cdot 10^{-09}}$	$1.63 \cdot 10^{-08}$	$7.98 \cdot 10^{-01}$	$9.63 \cdot 10^{-02}$
0.80	$1.83 \cdot 10^{-11}$	$1.02 \cdot 10^{-83}$	$\mathbf{9.61 \cdot 10^{-12}}$	$9.10 \cdot 10^{-11}$	$7.43 \cdot 10^{-01}$	$5.66 \cdot 10^{-02}$
0.85	$7.71 \cdot 10^{-14}$	$1.62 \cdot 10^{-121}$	$\mathbf{1.19 \cdot 10^{-13}}$	$5.74 \cdot 10^{-13}$	$8.99 \cdot 10^{-01}$	$9.04 \cdot 10^{-02}$
0.90	$1.13 \cdot 10^{-15}$	$7.96 \cdot 10^{-158}$	$\mathbf{1.66 \cdot 10^{-15}}$	$4.25 \cdot 10^{-14}$	$9.31 \cdot 10^{-01}$	$3.42 \cdot 10^{-01}$
0.95	$3.75 \cdot 10^{-18}$	$1.25 \cdot 10^{-221}$	$\mathbf{2.82 \cdot 10^{-18}}$	$4.03 \cdot 10^{-18}$	$7.65 \cdot 10^{-01}$	$1.62 \cdot 10^{-01}$

Table 4: $\epsilon = 0.05$, $c = 0.05$

p	1/e-value (non-private)	Exact p-value (non-private)	Ours, Laplace	Ours, Gauss. ($\delta = 0.01$)	ToT atop 1/e-value	ToT atop exact p-value
0.05	$5.04 \cdot 10^{-18}$	$9.79 \cdot 10^{-218}$	$\mathbf{5.71 \cdot 10^{-18}}$	$5.84 \cdot 10^{-18}$	$3.57 \cdot 10^{-01}$	$3.56 \cdot 10^{-02}$
0.10	$9.31 \cdot 10^{-16}$	$1.06 \cdot 10^{-159}$	$\mathbf{1.02 \cdot 10^{-15}}$	$1.61 \cdot 10^{-15}$	$1.43 \cdot 10^{-01}$	$5.31 \cdot 10^{-03}$
0.15	$2.75 \cdot 10^{-13}$	$7.00 \cdot 10^{-112}$	$\mathbf{2.94 \cdot 10^{-13}}$	$3.06 \cdot 10^{-13}$	$8.85 \cdot 10^{-01}$	$4.55 \cdot 10^{-03}$
0.20	$6.90 \cdot 10^{-12}$	$8.60 \cdot 10^{-90}$	$\mathbf{6.71 \cdot 10^{-12}}$	$6.86 \cdot 10^{-12}$	$9.91 \cdot 10^{-01}$	$5.43 \cdot 10^{-03}$
0.25	$1.55 \cdot 10^{-10}$	$2.81 \cdot 10^{-71}$	$\mathbf{1.49 \cdot 10^{-10}}$	$1.92 \cdot 10^{-10}$	$4.83 \cdot 10^{-01}$	$5.28 \cdot 10^{-03}$
0.30	$5.87 \cdot 10^{-06}$	$8.57 \cdot 10^{-26}$	$5.99 \cdot 10^{-06}$	$\mathbf{3.75 \cdot 10^{-06}}$	$8.28 \cdot 10^{-01}$	$1.52 \cdot 10^{-03}$
0.35	$8.00 \cdot 10^{-05}$	$3.71 \cdot 10^{-18}$	$\mathbf{7.42 \cdot 10^{-05}}$	$1.05 \cdot 10^{-04}$	$5.95 \cdot 10^{-01}$	$3.63 \cdot 10^{-02}$
0.40	$3.00 \cdot 10^{-03}$	$9.25 \cdot 10^{-10}$	$\mathbf{3.02 \cdot 10^{-03}}$	$5.10 \cdot 10^{-03}$	$1.60 \cdot 10^{-01}$	$8.82 \cdot 10^{-01}$
0.45	$4.62 \cdot 10^{-01}$	$3.98 \cdot 10^{-02}$	$6.66 \cdot 10^{-01}$	$\mathbf{5.54 \cdot 10^{-01}}$	$9.64 \cdot 10^{-01}$	$7.51 \cdot 10^{-01}$
0.50	$1.00 \cdot 10^{+00}$	$5.48 \cdot 10^{-01}$	$1.00 \cdot 10^{+00}$	$8.77 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$	$\mathbf{6.41 \cdot 10^{-02}}$
0.55	$9.27 \cdot 10^{-02}$	$4.42 \cdot 10^{-04}$	$1.10 \cdot 10^{-01}$	$\mathbf{5.35 \cdot 10^{-02}}$	$7.74 \cdot 10^{-01}$	$7.36 \cdot 10^{-01}$
0.60	$5.54 \cdot 10^{-03}$	$1.38 \cdot 10^{-08}$	$5.83 \cdot 10^{-03}$	$\mathbf{3.82 \cdot 10^{-03}}$	$5.24 \cdot 10^{-01}$	$1.55 \cdot 10^{-01}$
0.65	$1.67 \cdot 10^{-04}$	$3.03 \cdot 10^{-16}$	$\mathbf{1.71 \cdot 10^{-04}}$	$1.84 \cdot 10^{-04}$	$5.56 \cdot 10^{-01}$	$6.10 \cdot 10^{-02}$
0.70	$7.78 \cdot 10^{-06}$	$6.65 \cdot 10^{-25}$	$8.23 \cdot 10^{-06}$	$\mathbf{6.01 \cdot 10^{-06}}$	$3.64 \cdot 10^{-01}$	$6.52 \cdot 10^{-04}$
0.75	$4.59 \cdot 10^{-09}$	$6.20 \cdot 10^{-54}$	$4.55 \cdot 10^{-09}$	$\mathbf{3.25 \cdot 10^{-09}}$	$1.06 \cdot 10^{-01}$	$6.70 \cdot 10^{-03}$
0.80	$1.83 \cdot 10^{-11}$	$1.02 \cdot 10^{-83}$	$1.85 \cdot 10^{-11}$	$\mathbf{1.65 \cdot 10^{-11}}$	$5.40 \cdot 10^{-01}$	$2.58 \cdot 10^{-03}$
0.85	$1.95 \cdot 10^{-14}$	$1.51 \cdot 10^{-132}$	$2.00 \cdot 10^{-14}$	$\mathbf{1.73 \cdot 10^{-14}}$	$5.20 \cdot 10^{-01}$	$2.43 \cdot 10^{-03}$
0.90	$9.31 \cdot 10^{-16}$	$1.06 \cdot 10^{-159}$	$\mathbf{1.03 \cdot 10^{-15}}$	$1.03 \cdot 10^{-15}$	$4.89 \cdot 10^{-01}$	$2.93 \cdot 10^{-02}$
0.95	$2.08 \cdot 10^{-18}$	$1.08 \cdot 10^{-229}$	$2.55 \cdot 10^{-18}$	$\mathbf{2.42 \cdot 10^{-18}}$	$5.72 \cdot 10^{-01}$	$1.88 \cdot 10^{-03}$

D.2 Section 4.2 in the main text

We start by splitting the first 100 000 samples of the dataset into three splits: training (45%), validation (15%) and test (40%). We train a random forest classifier on the training data, and compute a validation loss on the validation set. With this we define the desired safety threshold to be the validation 0-1 loss plus a tolerance of 0.05.

We use the e-value for the mean with distribution $F \sim \text{Unif}(0, c/\theta)$, where θ is the determined safety threshold and $c = 0.2$. For the anytime-validity, we use a batch size of 128 samples per batch. The null is rejected at a significance level of $\alpha = 0.05$, i.e., when the e-value crosses $1/0.05 = 20$.

D.3 Section 4.3 in the main text

We split the data into three splits: training (54%), calibration (36%) and testing (10%). On the training data we fit a XGBoost classifier. Then, on the calibration data we run our procedure, with conformity score $(x, y) \mapsto 1/\mathrm{clamp}_{[0.01,1]}(\widehat{p}(y|x))^{1/4}$. We then linearly transform this score to have support in $[1, 100]$, and quantize it into 500 uniform bins over $[1, 100]$. Atop this, we run our method. Finally, we assess the resulting predictive interval sizes in the test split. Empty prediction sets (if any) are converted into singleton predictive sets.