A. Algorithmic Subroutines

In this section, we present three key subroutines in our work. Algorithm 2 is the SPARSEVECTOR algorithm, a classic primitive in differential privacy. Algorithm 3 is SAFFRON and algorithm 4 is LORD++, both are recent methods for online control of false discovery rate.

Algorithm 2 Sparse Vector: SPARSEVECTOR$(D, \Delta, \{f_1, f_2, \ldots\}, T, c, \varepsilon)$

Input: database D, stream of queries $\{f_1, f_2, \ldots\}$ each with sensitivity Δ, threshold T, a cutoff point c, privacy parameter ε
Let $\hat{T}_0 = T + \text{Lap}(\frac{2\Delta c}{\varepsilon})$
Let count $= 0$
for each query i **do**
  Let $Z_i \sim \text{Lap}(\frac{4\Delta c}{\varepsilon})$
  if $f_i(X) + Z_i > \hat{T}$ **then**
    Output $a_i = \top$
    Let count $=$ count $+1$
    Let $\hat{T}_{\text{count}} = T + \text{Lap}(\frac{2\Delta c}{\varepsilon})$
  else
    Output $a_i = \perp$
  end if
  if count $\geq c$ **then**
    Halt.
  end if
end for

Algorithm 3 SAFFRON$(\alpha, W_0, \{\gamma_j\}_{j=0}^{\infty})$

Input: stream of p-values $\{p_1, p_2, \ldots\}$, target FDR level α, initial wealth $W_0 < \alpha$, positive non-increasing sequence $\{\gamma_j\}_{j=0}^{\infty}$ of summing to one.
Set rejection number $i = 0$
for each p-value p_t **do**
  Set $\lambda_t = g_t(R_{1:t-1}, C_{1:t-1})$
  Set the indicator for candidacy $C_t = I(p_t < \lambda_t)$. Set the candidates after the j-th rejection as $C_{j+} = \sum_{i=\tau_j+1}^{t-1} C_i$
  if $t = 1$ **then**
    Set $\alpha_1 = (1 - \lambda_1)\gamma_1 W_0$
  else
    Compute $\alpha_t = (1 - \lambda_t)(W_0 \gamma_{t-C_{0+}} + (\alpha - W_0)\gamma_{t-\tau_1-C_{1+}} + \sum_{j\geq 2} \alpha\gamma_{t-\tau_j-C_{j+}})$
  end if
  Output $R_t = I(p_t \leq \alpha_t)$
  if $R_t = 1$ **then**
    Update rejection number $i = i + 1$. Set the i-th rejection time as $\tau_i = t$
  end if
end for

Algorithm 4 LORD++$(\alpha, W_0, \{\gamma_j\}_{j=0}^{\infty})$

Input: stream of p-values $\{p_1, p_2, \ldots\}$, target FDR level α, initial wealth $W_0 < \alpha$, positive non-increasing sequence $\{\gamma_j\}_{j=0}^{\infty}$ of summing to one.
Set rejection number $i = 0$
for each p-value p_t **do**
  Compute $\alpha_t = W_0\gamma_t + (\alpha - W_0)\gamma_{t-\tau_1} + \sum_{j\geq 2} \alpha\gamma_{t-\tau_j}$
  Output $R_t = I(p_t \leq \alpha_t)$
  if $R_t = 1$ **then**
    Update rejection number $i = i + 1$. Set the i-th rejection time as $\tau_i = t$
  end if
end for

We conclude this section by formally describing the conditional super-uniformity condition condition required by SAFFRON's guarantees, as stated in Theorem 2. It requires that the input sequence of p-values are not too correlated under the null hypothesis. This condition is formalized through a *filtration* on the sequence of candidacy and rejection decisions. Intuitively, this means that the sequence of hypotheses cannot be too adaptively chosen, otherwise the p-values may become overly correlated and violate this condition. Denote by $R_j := I(p_j \leq \alpha_j)$ the indicator for rejection, and let $C_j := I(p_j \leq \lambda_j)$ be the indicator for candidacy. Define the filtration formed by the sequences of σ-fields $\mathcal{F}^t := \sigma(R_1, \ldots, R_t, C_1, \ldots, C_t)$, and let $\alpha_t := f_t(R_1, \ldots, R_{t-1}, C_1, \ldots, C_{t-1})$, where f_t is an arbitrary function of the first $t-1$ indicators for rejections and candidacy. The null p-values are said to be *conditionally super-uniformly distributed* with respect to the filtration $\mathcal{F}$ if:

$$\text{If null hypothesis } H_i \text{ is true, then } \Pr(p_t \leq \alpha_t | \mathcal{F}^{t-1}) \leq \alpha_t. \tag{2}$$

We note that independent p-values is a special case of the conditional super-uniformity condition of (2). When p-values are independent, they satisfy the following condition:

$$\text{If the null hypothesis } H_i \text{ is true, then } \Pr(p_t \leq u) \leq u \text{ for all } u \in [0, 1].$$

B. PrivLORD and PrivLORD2

In this section, we present two private versions of LORD++: PrivLORD in Algorithm 5 and PrivLORD2 in Algorithm 6. The former combines SPARSEVECTOR and LORD++, with the same threshold shifting as in PAPRIKA. The latter adds the candidacy checking step with constant λ on top of PrivLORD. The privacy of PrivLORD follows immediately from SPARSEVECTOR , and the privacy proof for PAPRIKA also applies to PrivLORD2 with a different choice of α_t.

Algorithm 5 PrivLORD$(\alpha, W_0, \gamma, c, \varepsilon, A)$

Input: stream of p-values $\{p_1, p_2, \ldots\}$ with mutiplicative sensitivity (η,μ), target FDR level α, initial wealth $W_0 < \alpha$, positive non-increasing sequence $\{\gamma_j\}_{j=0}^{\infty}$ of summing to one, expected number of rejections c, privacy parameters ε, threshold shift A.
Let $Z_\alpha^0 \sim \text{Lap}(2\eta c/\varepsilon)$, count $= 0$
for each p-value p_t **do**
  if count $\geq c$ **then** Output $R_t = 0$
  else
  Sample $Z_t \sim \text{Lap}(4\eta c/\varepsilon)$.
  if $t = 1$
  then Set $\alpha_1 = \gamma_1 W_0$
  else
  Compute $\alpha_t = W_0\gamma_t + (\alpha - W_0)\gamma_{t-\tau_1} + \sum_{j\geq 2} \alpha\gamma_{t-\tau_j}$
  if $\log p_t + Z_t \leq \log \alpha_t - A + Z_\alpha^{\text{count}}$
  then Output $R_t = 1$. Set count = count $+1$ and sample $Z_\alpha^{\text{count}} \sim \text{Lap}(2\eta c/\varepsilon)$
  else Output $R_t = 0$
end for

Algorithm 6 PrivLORD2$(\alpha, \lambda, W_0, \gamma, c, \varepsilon, \delta, A)$

Input: stream of p-values $\{p_1, p_2, \ldots\}$ with mutiplicative sensitivity (η,μ), target FDR level α, candidacy threshold λ, initial wealth $W_0 < \alpha$, positive non-increasing sequence $\{\gamma_j\}_{j=0}^{\infty}$ of summing to one, expected number of rejections c, privacy parameters ε, δ, threshold shift A.
Let $Z_\alpha^0 \sim \text{Lap}(2\eta c/\varepsilon)$, count $= 0$
for each p-value p_t **do**
  if count $\geq c$ **then** Output $R_t = 0$
  else
  Sample $Z_t \sim \text{Lap}(4\eta c/\varepsilon)$. Set the indicator for candidacy $C_t = I(\log p_t < \log \lambda)$.
  if $t = 1$
  then Set $\alpha_1 = \gamma_1 W_0$
  else
  Compute $\alpha_t = W_0\gamma_t + (\alpha - W_0)\gamma_{t-\tau_1} + \sum_{j\geq 2} \alpha\gamma_{t-\tau_j}$
  if $C_t = 1$ and $\log p_t + Z_t \leq \log \alpha_t - A + Z_\alpha^{\text{count}}$
  then Output $R_t = 1$. Set count = count $+1$ and sample $Z_\alpha^{\text{count}} \sim \text{Lap}(2\eta c/\varepsilon)$
  else Output $R_t = 0$
end for

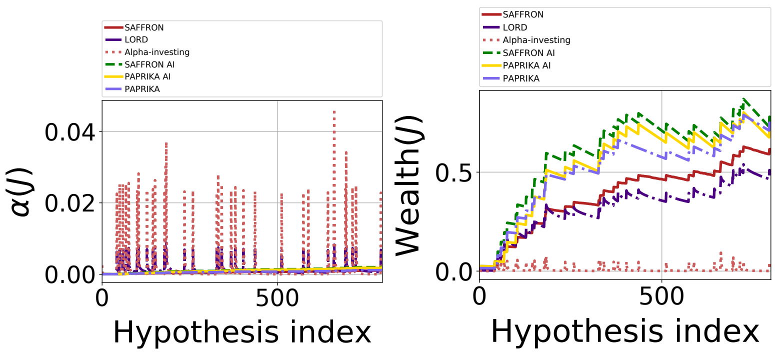
C. Additional Experimental Results

We provide a further illustration of our experiments on truncated exponentials in Figure 4. In particular, we plot the rejection threshold α_t and wealth versus the hypothesis index. Each "jump" of the wealth corresponds to a rejection. We observe that the rejections of our private algorithms are consistent with the rejections of the non-private algorithms, another perspective which empirically confirms their accuracy.

One hypothesis for the good performance observed in Figure 3 is that the signal between the null and alternative hypotheses as parameterized by θ_i is very strong, meaning the algorithms can easily discriminate between the true null and true non-null hypotheses based on the observed p-values. To measure this, we also varied the value of θ_i in the alternative hypotheses. These results are shown in Figure 5, which plots FDR and power of PAPRIKA and PAPRIKA AI with when the alternative hypotheses have parameter $\theta_i = 1.90, 1.95, 2.00$. As expected, the performance gets better as we increase the signal, and we observe that when the signal is too weak ($\theta_i = 1.90$), performance begins to decline.

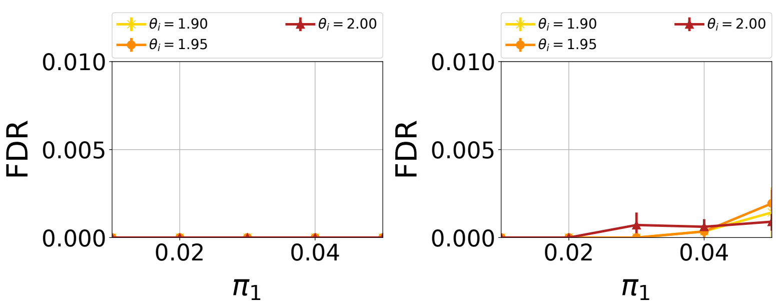
We compare PAPRIKA and PAPRIKA AI against PrivLORD and PrivLORD2 with truncated exponential observations in Figure 6. All methods except for PrivLORD perform well, suggesting that the candidacy checking step is critical for private

FDR control algorithms.

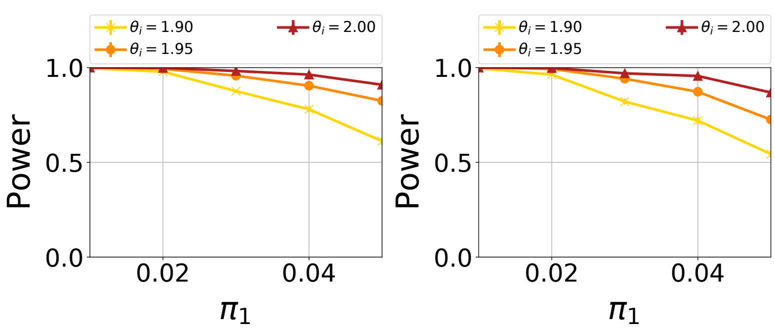


(a) $\varepsilon = 5$ (b) $\varepsilon = 5$

Figure 4. Wealth and rejection threshold α_t versus hypothesis index with privacy parameter $\varepsilon = 5$ when the database consists of truncated exponential observations. PAPRIKA AI and SAFFRON AI used $\lambda_j = \alpha_j$, PAPRIKA used $\lambda_j = 0.2$, and SAFFRON used $\lambda_j = 0.5$.

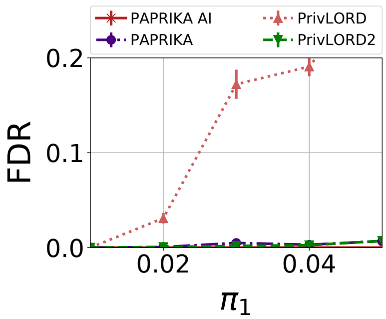


(a) PAPRIKA AI (b) PAPRIKA AI

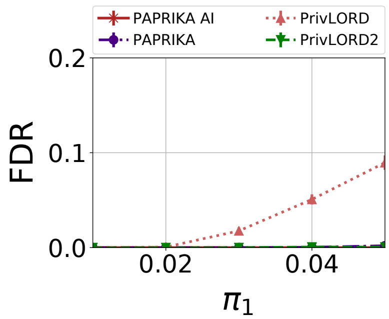


(c) PAPRIKA (d) PAPRIKA

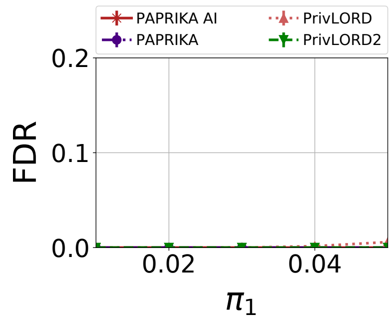
Figure 5. FDR and statistical power versus expected fraction of non-null hypotheses π_1 under various choices of signal $\theta_i = 1.90, 1.95, 2.00$ for alternative hypothesis parameters. The privacy parameter is $\varepsilon = 5$, and the database consists of truncated exponential observations. The first row shows performance of PAPRIKA AI where $\lambda_j = \alpha_j$, and the second row shows performance of PAPRIKA where $\lambda_j = 0.2$.



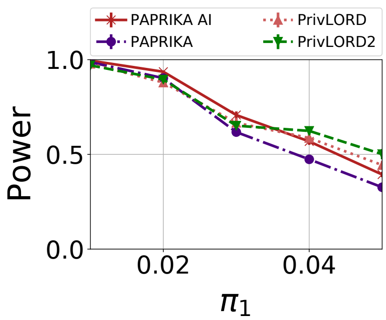
(a) $\varepsilon = 3$



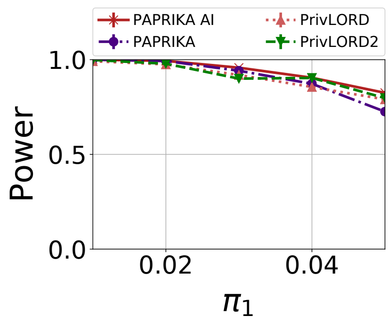
(b) $\varepsilon = 5$



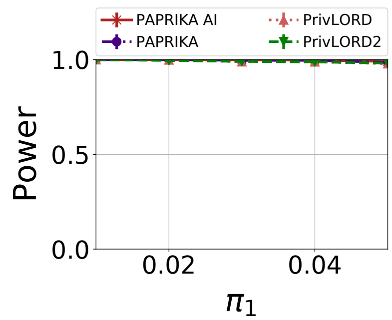
(c) $\varepsilon = 10$



(d) $\varepsilon = 3$



(e) $\varepsilon = 5$



(f) $\varepsilon = 10$

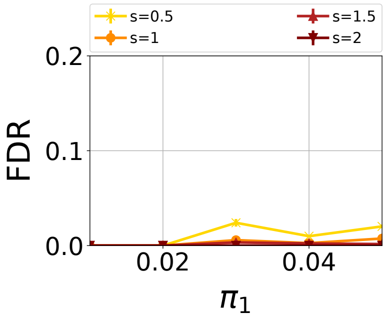
Figure 6. FDR and statistical power versus fraction of non-nulls π_1 for PAPRIKA (with $\lambda_j = 0.2$), PAPRIKA AI (with $\lambda_j = \alpha_j$), and PrivLORD and PrivLORD2 when the database consists of truncated exponential observations.

C.1. Choice of shift A

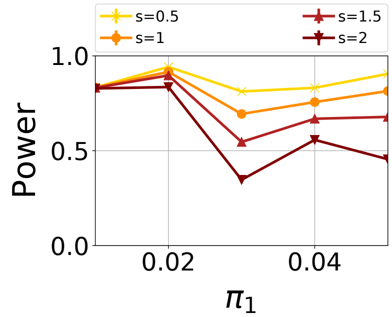
We now discuss how to choose the shift parameter A. Theorem 3 gives a theoretical lower bound for A in terms of the privacy parameter δ, but this bound may be overly conservative. Since the shift A is closely related to the performance of FDR and statistical power, we wish to pick a value of A that yields good performance in practice. In Theorem 4, we show that FDR(t) is less than our desired bound α plus the privacy parameter δt, which naturally requires that the privacy loss parameter δ be small. For a more detailed explanation, we bound Inequality (22) in the proof of Theorem 4 using Inequality (14) from the proof of Theorem 3, and therefore, the empirical δ is naturally tied to the empirical FDR. As long as we can guarantee the empirical FDR to be bounded by the target FDR level, our privacy loss is bounded by the nominal δ.

We use the Bernoulli example in Section 4.1 to investigate the performance under different choices of the shift A with privacy parameter $\varepsilon = 5$. The results are summarized in Figure 7, which plots the FDR and power versus the expected fraction of non-nulls when we vary the shift size with $s = 0.5, 1, 1.5, 2$.

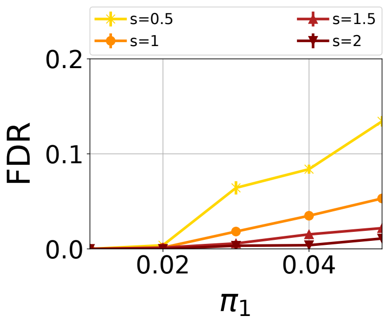
Larger shifts (corresponding to larger values of s) will lower the rejection threshold, which causes fewer hypotheses to be rejected. This improves FDR of the algorithm, but harms Power, as the threshold may be too low to reject true nulls. Figure 7 shows that the shift size parameter s should be chosen by the analyst to balance the tradeoff between FDR and Power, as demanded by the application.



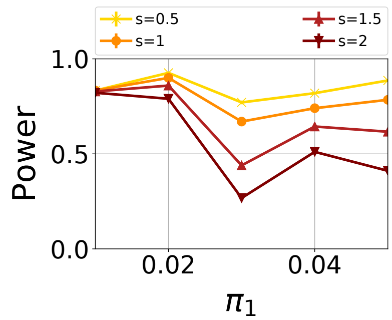
(a) PAPRIKA AI



(b) PAPRIKA AI



(c) PAPRIKA



(d) PAPRIKA

Figure 7. FDR and statistical power versus expected fraction of non-null hypotheses π_1 under various choices of shift magnitude s. The privacy parameter is $\varepsilon = 5$, and the database consists of Bernoulli observations. The first row shows performance of PAPRIKA AI where $\lambda_j = \alpha_j$, and the second row shows performance of PAPRIKA where $\lambda_j = 0.2$.

C.2. Additional Tables

Tables 1 and 2 report the numerical values for our experiments on Bernoulli and truncated exponential data, respectively. This information is also presented visually in Figures 1 and 3.

D. Proof of Theorem 3

Before proving Theorem 3, we will state and prove the following lemma, which will be useful in the proofs of Theorem 3 and Theorem 4.

Lemma 2. *If $Z_1 \sim Lap(2b)$, $Z_2 \sim Lap(b)$ and $C > 0$ is a constant, we have* $\Pr(Z_1 \geq Z_2 - C) = 1 - \frac{2}{3}\exp(-\frac{C}{2b}) + \frac{1}{6}\exp(-C/b)$.

π	ε	PAPRIKA AI		PAPRIKA		SAFFRON AI		SAFFRON		LORD		Alpha-investing		LapSAFFRON	
		FDR	power	FDR	power	FDR	power	FDR	power	FDR	power	FDR	power	FDR	power
0.01	3	0	.825	0	.817	0	.833	0	.833	0	.833	0	.833	.990	.485
	5	0	.833	0	.833										
	10	0	.833	0	.833										
0.02	3	0	.844	.017	.810	0	.938	0	.938	0	.938	0	.875	.973	.509
	5	0	.916	.001	.900										
	10	0	.941	0	.938										
0.03	3	.008	.457	.103	.389	.077	.923	0	.846	0	.846	0	.692	.977	.509
	5	.006	.694	.018	.670										
	10	.015	.849	.007	.808										
0.04	3	.003	.604	.120	.580	.030	.970	0	.879	0	.940	0	.848	.943	.512
	5	.003	.756	.035	.740										
	10	.060	.860	.008	.836										
0.05	3	.009	.560	.168	.514	.056	.971	.056	.971	.105	.971	.056	.971	.940	.505
	5	.007	.815	.053	.785										
	10	.017	.938	.012	.922										

Table 1. Numerical values of FDR and power for Bernoulli observations experiments. LapSAFFRON corresponds to running SAFFRON on the naïve Laplace privatization of the p-values.

Proof.

$$\begin{aligned}
\Pr(Z_1 \geq Z_2 - C) &= \int_{-\infty}^{\infty}\int_{x-C}^{\infty} \frac{1}{4b}\exp(-\frac{|y|}{2b})\frac{1}{2b}\exp(-\frac{|x|}{b})dydx \\
&= \int_{\infty}^{C} (1-\frac{1}{2}\exp(-\frac{|x-C|}{2b}))\frac{1}{2b}\exp(-\frac{|x|}{b})dx + \int_{C}^{\infty}\frac{1}{2}\exp(-\frac{|x-C|}{2b})\frac{1}{2b}\exp(-\frac{|x|}{b})dx \\
&= \int_{-\infty}^{C}\frac{1}{2b}\exp(-\frac{|x|}{b})dx - \int_{-\infty}^{0}\frac{1}{4b}\exp(-\frac{|3x-C|}{2b})dx \\
&- \int_{0}^{C}\frac{1}{4b}\exp(-\frac{C+x}{2b})dx + \int_{C}^{\infty}\frac{1}{4b}\exp(-\frac{|3x-C|}{2b})dx \\
&= 1 - \frac{1}{2}\exp(-\frac{C}{b}) - \frac{1}{6}\exp(-\frac{C}{2b}) - \frac{1}{2}\exp(-\frac{C}{2b}) + \frac{1}{2}\exp(-\frac{C}{b}) + \frac{1}{6}\exp(-\frac{C}{b}) \\
&= 1 - \frac{2}{3}\exp(-\frac{C}{2b}) + \frac{1}{6}\exp(-\frac{C}{b})
\end{aligned}$$

□

Theorem 3. *For any stream of p-values* $\{p_1, p_2, \ldots\}$, PAPRIKA *is* (ε, δ)*-differentially private.*

Proof. Fix any two neighboring databases D and D'. Let R denote the random variable representing the output of PAPRIKA$(D, \alpha, \lambda, W_0, \{\gamma_j\}_{j=0}^{\infty}, c, \varepsilon, \delta, s)$ and let R' denote the random variable representing the output of PAPRIKA$(D', \alpha, \lambda, W_0, \{\gamma_j\}_{j=0}^{\infty}, c, \varepsilon, \delta, s)$. Let k denote the total number of hypotheses. When $\log p_t \geq \log 2\lambda$ and $\log p'_t \geq \log 2\lambda$ for all t, $\Pr(R = \{0, 0, \ldots, 0\}) = 1 = \Pr(R' = \{0, 0, \ldots, 0\})$. When $\log p_t < \log 2\lambda$ and $\log p'_t < \log 2\lambda$ for all t, privacy follows from the privacy of SparseVector with dynamic thresholds. Since the threshold at each time t only depends on the threshold at time $t-1$ and and private rejection $R(t-1)$, by post-processing, the threshold α_t is private. Then by post-processing and the privacy of SparseVector , the rejection $R(t)$ is also private. We give the formal probability argument as follows. For any neighboring D, D' and any sequence of hypotheses, we first consider the output up to the first rejection, which is AboveThresh . Consider any output $r \in \{0, 1\}^l$. Let $r = \{r_1, r_2, \ldots, r_l\}$, with $r_l = 1$ and $r_1 = \ldots = r_{l-1} = 0$. Let

$$\begin{aligned}
f_i(D, z, \alpha_i) &= \Pr(\log p_i(D) + Z_i < \log \alpha_i - A + z) \\
g_i(D, z, \alpha_i) &= \Pr(\log p_i(D) + Z_i \geq \log \alpha_i - A + z),
\end{aligned}$$

π	ε	PAPRIKA AI		PAPRIKA		SAFFRON AI		SAFFRON		LORD		Alpha-investing		LapSAFFRON	
		FDR	power	FDR	power	FDR	power	FDR	power	FDR	power	FDR	power	FDR	power
0.01	3	0	.995	0	.987	0	1.00	0	1.00	0	1.00	0	.638	.989	.543
	5	0	1.00	0	1.00										
	10	0	1.00	0	1.00										
0.02	3	0	.936	0	.903	0	1.00	0	1.00	0	.999	0	.676	.973	.505
	5	0	.994	0	.993										
	10	0	.999	0	1.00										
0.03	3	0	.708	.005	.618	0	1.00	0	1.00	0	1.00	0	.982	.977	.516
	5	0	.958	0	.942										
	10	0	.999	0	.996										
0.04	3	0	.569	.003	.474	0	1.00	0	1.00	0	1.00	0	.999	.944	.503
	5	0	.905	0	.873										
	10	0	.998	0	.996										
0.05	3	0	.394	.007	.327	0	1.00	0	1.00	0	1.00	0	1.00	.940	.505
	5	0	.825	.002	.726										
	10	0	.990	0	.986										

Table 2. Numerical values of FDR and power for truncated exponential observations experiments. LapSAFFRON corresponds to running SAFFRON on the naïve Laplace privatization of the p-values.

where $\alpha_1, \dots, \alpha_t$ is a fixed sequence of thresholds determined by the r. We have

$$\begin{aligned}
\frac{\Pr(R=r|D)}{\Pr(R'=r|D')} &= \frac{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z) \Pr(R_l(D) = r_l | r_{l-1}, \dots, r_1) \Pr(R_2(D) = r_2|r_1) \Pr(R_1(D) = r_1) dz}{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z) \Pr(R_l(D') = r_l | r_{l-1}, \dots, r_1) \Pr(R_2(D') = r_2|r_1) \Pr(R_1(D') = r_1) dz} \\
&= \frac{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z) g_l(D, z, \alpha_l) \prod_{i=1}^{l-1} f_i(D, z, \alpha_i) dz}{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z) g_l(D', z, \alpha_l) \prod_{i=1}^{l-1} f_i(D', z, \alpha_i) dz}, \\
&= \frac{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z-\eta) g_l(D, z-\eta, \alpha_l) \prod_{i=1}^{l-1} f_i(D, z-\eta, \alpha_i) dz}{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z) g_l(D', z, \alpha_l) \prod_{i=1}^{l-1} f_i(D', z, \alpha_i) dz}, && (3) \\
&\le \frac{\int_{-\infty}^{\infty} \exp(\varepsilon/2c) \Pr(Z_\alpha = z) g_l(D, z-\eta, \alpha_l) \prod_{i=1}^{l-1} f_i(D', z, \alpha_i) dz}{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z) g_l(D', z, \alpha_l) \prod_{i=1}^{l-1} f_i(D', z, \alpha_i) dz}, && (4) \\
&\le \frac{\int_{-\infty}^{\infty} \exp(\varepsilon/2c) \Pr(Z_\alpha = z) \exp(\varepsilon/2c) g_l(D', z, \alpha_l) \prod_{i=1}^{l-1} f_i(D', z, \alpha_i) dz}{\int_{-\infty}^{\infty} \Pr(Z_\alpha = z) g_l(D', z, \alpha_l) \prod_{i=1}^{l-1} f_i(D', z, \alpha_i) dz}, && (5) \\
&= \exp(\varepsilon/c). && (6)
\end{aligned}$$

Equation (3) is from change of integration variable z to $z-\eta$. Inequality (4) is because Z_α follows $\text{Lap}(2\eta c/\varepsilon)$ and $\log p_i(D) - \eta \le \log p_i(D')$. Inequality (5) is because

$$\begin{aligned}
g_l(D, z-\eta, \alpha_l) &= \Pr(\log p_l(D) + Z_l \ge \log \alpha_l - A + z - \eta) \\
&\le \Pr(\log p_l(D') + \eta + Z_l \ge \log \alpha_l - A + z - \eta) \\
&\le \Pr(\log p_l(D') + Z_l \ge \log \alpha_l - A + z - 2\eta) \\
&\le \exp(\varepsilon/2c) \Pr(\log p_l(D') + Z_l \ge \log \alpha_l - A + z) \\
&\le \exp(\varepsilon/2c) g_l(D', z, \alpha_l).
\end{aligned}$$

When we restart AboveThresh after the first rejection, the inital threshold is the post-processing of the previous ouputs, which is also private. Then by simple composition, the overall privacy loss is ε.

For other cases, the worst case is that for all t, $\log p_t < \log 2\lambda$ and $\log p'_t \ge \log 2\lambda$. In this setting, we have

$$\Pr(R' = r) = \begin{cases} 1 & \text{if } r = \{0, 0, \dots, 0\} \\ 0 & \text{otherwise.} \end{cases}$$

To satisfy (ε, δ)-differential privacy, we need to bound the probability of outputting r for database D. We first consider $r = \{0, 0 \ldots, 0\}$. We wish to bound $\Pr(R' = \{0, 0 \ldots, 0\}) \leq \exp(\varepsilon) \Pr(R = \{0, 0, \ldots, 0\}) + \delta$ and $\Pr(R = \{0, 0 \ldots, 0\}) \leq \exp(\varepsilon) \Pr(R' = \{0, 0, \ldots, 0\}) + \delta$. The latter is trivial since $\exp(\varepsilon) \Pr(R' = \{0, 0, \ldots, 0\}) + \delta = \exp(\varepsilon) + \delta$, which is greater than 1. It remains to satisfy $\Pr(R' = \{0, 0 \ldots, 0\}) \leq \exp(\varepsilon) \Pr(R = \{0, 0, \ldots, 0\}) + \delta$, which is equivalent to $1 - \delta \leq \exp(\varepsilon) \Pr(R = \{0, 0, \ldots, 0\})$. We have

$$
\begin{aligned}
\Pr(R = \{0, 0 \ldots, 0\}) &= \Pr(R_1 = 0) \Pr(R_2 = 0 | R_1 = 0) \ldots \Pr(R_k = 0 | R_{k-1} = 0) \\
&= \prod_{t=1}^{k} \Pr(\log p_t + Z_t \geq \log \alpha_t - A + Z_\alpha) \\
&> \prod_{t=1}^{k} \Pr(\log 2\lambda - \eta + Z_t \geq \log \alpha_t - A + Z_\alpha) \qquad (7) \\
&= \prod_{t=1}^{k} \Pr(Z_t \geq Z_\alpha + \log \alpha_t - \log 2\lambda + \eta - A) \\
&= \prod_{t=1}^{k} \left(1 - \frac{2}{3} \exp(-\frac{\varepsilon(A + \log(2\lambda/\alpha_t) - \eta)}{4\eta c}) + \frac{1}{6} \exp(-\frac{\varepsilon(A + \log(2\lambda/\alpha_t) - \eta)}{2\eta c}\right) \qquad (8) \\
&\geq \left(1 - \frac{2}{3} \exp(-\frac{\varepsilon(A + \log 2 - \eta)}{4\eta c})\right)^k, \qquad (9)
\end{aligned}
$$

where Inequality (7) is because the worst case happens when p_t is η below the candidacy threshold $\log 2\lambda$, Equation (8) applies Lemma 2, and Inequality (9) follows from the facts that $\alpha_t \leq \lambda$ for all t and that the third term in (8) is positive. Setting (9) to be larger than $(1 - \delta)/\exp(\varepsilon)$, we have,

$$
\frac{2}{3} \exp\left(-\frac{\varepsilon(A + \log 2 - \eta)}{4\eta c}\right) \leq 1 - \left(\frac{1 - \delta}{\exp(\varepsilon)}\right)^{\frac{1}{k}}. \qquad (10)
$$

Next, we consider all other possible outputs r. Define the set $S := \{r \mid \text{there exists a } t \text{ such that } r_t = 1\}$. We wish to bound $\Pr(R \in S) \leq \exp(\varepsilon) \Pr(R' \in S) + \delta$ and $\Pr(R' \in S) \leq \exp(\varepsilon) \Pr(R \in S) + \delta$. The latter is trivial since $\Pr(R' \in S) = 0$. It remains to bound $\Pr(R \in S) \leq \delta$. For any t, we have

$$
\begin{aligned}
\Pr(R \in S) &\leq \Pr(R_t = 1) \\
&= \Pr(\log p_t + Z_t \leq \log \alpha_t - A + Z_\alpha) \\
&\leq \Pr(\log 2\lambda + Z_t \leq \log \alpha_t - A + Z_\alpha) \qquad (11) \\
&= \Pr(Z_t \leq Z_\alpha - (\log(2\lambda/\alpha_t) + A)) \\
&\leq \Pr(Z_t \leq Z_\alpha - (\log 2 + A)) \\
&= \frac{2}{3} \exp\left(-\frac{\varepsilon(A + \log 2)}{4\eta c}\right) - \frac{1}{6} \exp\left(-\frac{\varepsilon(A + \log 2)}{2\eta c}\right) \qquad (12) \\
&\leq \frac{2}{3} \exp\left(-\frac{\varepsilon(A + \log 2)}{4\eta c}\right), \qquad (13)
\end{aligned}
$$

where Inequality (11) is because the worst case occurs when $\log p_t = \log 2\lambda$, Equality (12) applies Lemma 2, and Inequality (13) follows from the facts that $\alpha_t \leq \lambda$ for all t and that the second term in (12) is negative. Setting (13) to be less than δ, we have,

$$
\frac{2}{3} \exp\left(-\frac{\varepsilon(A + \log 2)}{4\eta c}\right) \leq \delta. \qquad (14)
$$

Combining Equations (14) and (10), we have the condition that $\frac{2}{3} \exp\left(-\frac{\varepsilon(A + \log 2 - \eta)}{4\eta c}\right) \leq \min\{\delta, 1 - ((1-\delta)/\exp(\varepsilon))^{1/k}\}$.

Rearranging this inequality for A gives

$$
A \geq \frac{4\eta c}{\varepsilon} \left(\log \frac{2}{3 \min\{\delta, 1 - ((1 - \delta)/\exp(\varepsilon))^{1/k}\}} - \log 2 + \eta\right),
$$

which is how the shift term A is set in PAPRIKA.

$\square$

E. Proof of Theorem 4

Theorem 4. *If the null p-values are conditionally super-uniformly distributed, then we have:*
(a) $\mathbb{E}\left[\sum_{j\leq t, j\in\mathcal{H}^0}\alpha_j \frac{I(p_j>2\lambda_j)}{1-2\lambda_j}\right]+\delta t\geq \mathbb{E}\left[|\mathcal{H}^0\cap\mathcal{R}(t)|\right]$*;*
(b)The condition $\widehat{FDP}_{\text{PAPRIKA}}(t)\leq\alpha$ *for all* $t\in\mathbb{N}$ *implies that mFDR*$(t)\leq\alpha+\delta t$ *for all* $t\in\mathbb{N}$*.*
If the null p-values are independent of each other and of the non-null p-values, and $\{\alpha_t\}$ *and* $\{\lambda_t\}$ *are coordinate-wise non-decreasing functions of the vector* $R_1,\dots,R_{t-1},C_1,\dots,C_{t-1}$*, then*
(c) $\mathbb{E}\left[\widehat{FDP}_{\text{PAPRIKA}}(t)\right]+\delta t\geq\mathbb{E}\left[FDP(t)\right]:=FDR(t)$ *for all* $t\in\mathbb{N}$*;*
(d) The condition $\widehat{FDP}_{\text{PAPRIKA}}(t)\leq\alpha$ *for all* t *implies that* $FDR(t)\leq\alpha+\delta t$ *for all* $t\in\mathbb{N}$*.*

Proof. For any time $t>0$, before the total number of rejections reaches c we bound the number of false rejections as follows:

$$\begin{aligned}\mathbb{E}\left[|\mathcal{H}^0\cap\mathcal{R}(t)|\right] &\leq \sum_{j\leq t, j\in\mathcal{H}^0}\mathbb{E}\left[I(\log p_j+Z_j\leq\log\alpha_j-A+Z_\alpha)\right] &(15)\\ &\leq \sum_{j\leq t, j\in\mathcal{H}^0} Pr(\log p_j\leq\log\alpha_j)+Pr(Z_j\leq Z_\alpha-A)\\ &\leq \sum_{j\leq t, j\in\mathcal{H}^0}\mathbb{E}\left[\alpha_j\right]+Pr(Z_j\leq Z_\alpha-A), &(16)\end{aligned}$$

where Inequality (15) follows from the rejection rule before the total number of rejections reaches c, and the number of false rejections is always 0 afterwards. Inequality (16) follows from the conditional super-uniformity property. We bound each term in (16) separately. Using the law of iterated expectations by conditioning on ${\mathcal{F}'}^{t-1}$, we can bound the first term of (16) as follows:

$$\begin{aligned}\sum_{j\leq t, j\in\mathcal{H}^0}\mathbb{E}\left[\alpha_j\right] \leq&\mathbb{E}\left[\sum_{j\leq t, j\in\mathcal{H}^0}\alpha_j\mathbb{E}\left[\frac{I(p_j>2\lambda_j)}{1-2\lambda_j}|{\mathcal{F}'}^{t-1}\right]\right]\\ =&\mathbb{E}\left[\mathbb{E}\left[\sum_{j\leq t, j\in\mathcal{H}^0}\alpha_j\frac{I(p_j>2\lambda_j)}{1-2\lambda_j}|{\mathcal{F}'}^{t-1}\right]\right]\\ =&\mathbb{E}\left[\sum_{j\leq t, j\in\mathcal{H}^0}\alpha_j\frac{I(p_j>2\lambda_j)}{1-2\lambda_j}\right], &(17)\end{aligned}$$

where Equation (17) applies the conditional super-uniformity. Since $\widehat{\text{FDP}}_{\text{PAPRIKA}}(t)\leq\alpha$, we have,

$$\mathbb{E}\left[\sum_{j\leq t, j\in\mathcal{H}^0}\alpha_j\frac{I(p_j>2\lambda_j)}{1-2\lambda_j}\right]\leq\alpha\mathbb{E}\left[|\mathcal{R}(t)|\right].$$

Next, we bound the second term in (16) as follows:

$$\begin{aligned}\sum_{j\leq t, j\in\mathcal{H}^0}\Pr(Z_j\leq Z_\alpha-A)\leq&\frac{2t}{3}\exp\left(-\frac{A\varepsilon}{4\eta c}\right)-\frac{t}{6}\exp\left(-\frac{A\varepsilon}{2\eta c}\right)\\ \leq& t\min\left\{\delta, 1-\left(\frac{1-\delta}{\exp(\varepsilon)}\right)^{\frac{1}{k}}\right\}.\end{aligned}$$

Combining this inequality with (17), we bound mFDR as

$$\begin{aligned} mFDR :=& \frac{\mathbb{E}\left[|\mathcal{H}^0 \cap \mathcal{R}(t)|\right]}{\mathbb{E}\left[|\mathcal{R}(t)|\right]} \\ \leq& \alpha + \frac{1}{\mathbb{E}\left[|\mathcal{R}(t)|\right]} \sum_{j\leq t, j\in\mathcal{H}^0} \Pr(Z_j \leq Z_\alpha - A) \\ \leq& \alpha + \min\left\{\delta, 1 - \left(\frac{1-\delta}{\exp(\varepsilon)}\right)^{\frac{1}{k}}\right\} t \\ \leq& \alpha + \delta t. \end{aligned}$$

If the null p-values are independent of each other and the non-nullls, and $\{\alpha_t\}$ is a coordinate-wise non-decreasing function of the vector $R_1, \ldots, R_{t-1}$, then we have

$$\begin{aligned} FDR(t) =& \mathbb{E}\left[\frac{|\mathcal{H}^0 \cap \mathcal{R}(t)|}{|\mathcal{R}(t)|}\right] \\ =& \sum_{j\leq t, j\in\mathcal{H}^0} \mathbb{E}\left[\frac{I(\log p_j + Z_j \leq \log \alpha_j - A + Z_\alpha)}{|\mathcal{R}(t)|}\right] \\ \leq& \sum_{j\leq t, j\in\mathcal{H}^0} \mathbb{E}\left[\frac{\min\{\alpha_j \exp(Z_\alpha - Z_j - A), 1\}}{|\mathcal{R}(t)|}\right] && (18) \\ \leq& \sum_{j\leq t, j\in\mathcal{H}^0} \mathbb{E}\left[\frac{\alpha_j}{|\mathcal{R}(t)|}\right] + \Pr(Z_j \leq Z_\alpha - A), && (19) \end{aligned}$$

where Inequality (18) applies the law of iterated expectations by conditioning on $\mathcal{F'}^{t-1}$ and Lemma 1. Inequality (19) follows by a case analysis: if $Z_j > Z_\alpha - A$, then $\exp(Z_\alpha - Z_j - A) < 1$, and thus $\frac{\min\{\alpha_j \exp(Z_\alpha - Z_j - A), 1\}}{|\mathcal{R}(t)|}$ reduces to $\frac{\alpha_j}{|\mathcal{R}(t)|}$. On the other hand, if $Z_j \leq Z_\alpha - A$, then $\frac{\min\{\alpha_j \exp(Z_\alpha - Z_j - A), 1\}}{|\mathcal{R}(t)|} \leq \frac{1}{|\mathcal{R}(t)|} \leq 1$, allowing us to upper bound the expectation by the probability of this event.

We bound the first term in (19) as follows:

$$\begin{aligned} \sum_{j\leq t, j\in\mathcal{H}^0} \mathbb{E}\left[\frac{\alpha_j}{|\mathcal{R}(t)|}\right] \leq& \sum_{j\leq t, j\in\mathcal{H}^0} \mathbb{E}\left[\frac{\alpha_j I(p_j > 2\lambda_j)}{(1-2\lambda_j)|\mathcal{R}(t)|}\right] && (20) \\ \leq& \mathbb{E}\left[\frac{\sum_{j\leq t} \alpha_j I(p_j > 2\lambda_j)}{(1-2\lambda_j)|\mathcal{R}(t)|}\right] \\ =& \mathbb{E}\left[\widehat{\text{FDP}}_{\text{PAPRIKA}}(t)\right] \\ \leq& \alpha, && (21) \end{aligned}$$

where Inequality (20) applies Lemma 1.

It remains to bound the second term in (19), which we do using Lemma 2 as follows:

$$\begin{aligned} \sum_{j\leq t, j\in\mathcal{H}^0} \Pr(Z_j \leq Z_\alpha - A) \leq& \sum_{j\leq t} \Pr(Z_j \leq Z_\alpha - A) \\ =& \frac{2t}{3} \exp(-\frac{A\varepsilon}{4\eta c}) - \frac{t}{6} \exp(-\frac{A\varepsilon}{2\eta c}) \\ \leq& \min\left\{\delta, 1 - \left(\frac{1-\delta}{\exp(\varepsilon)}\right)^{\frac{1}{k}}\right\} t. && (22) \end{aligned}$$

Combining (21) and (22), we reach the conclusion that $\text{FDR}(t) \leq \alpha + \min\{\delta, 1 - ((1-\delta)/\exp(\varepsilon))^{1/k}\} t \leq \alpha + \delta t$.

□

F. Proof of Lemma 1

Lemma 1. *Assume $p_1, p_2, \dots$ are all independent and let $h : \{0,1\}^k \to R$ be any coordinate-wise non-decreasing function. Assume f_t and g_t are coordinate-wise non-decreasing functions and that $\alpha_t = f_t(R_{1:t-1}, C_{1:t-1})$ and $\lambda_t = g_t(R_{1:t-1}, C_{1:t-1})$. Then for any $t \le k$ such that $H_t \in \mathcal{H}^0$, we have $\mathbb{E}\left[\frac{\alpha_t I(p_t > 2\lambda_t)}{(1-2\lambda_t)h(R_{1:k})}|{\mathcal{F}'}^{t-1}\right] \ge \mathbb{E}\left[\frac{\alpha_t}{h(R_{1:k})}|{\mathcal{F}'}^{t-1}\right]$ and $\mathbb{E}\left[\frac{\min\{\alpha_t \exp(Z_\alpha - Z_t - A), 1\}}{h(R_{1:k})}|{\mathcal{F}'}^{t-1}\right] \ge \mathbb{E}\left[\frac{I(\log p_t + Z_t \le \log \alpha_t + Z_\alpha - A)}{h(R_{1:k})}|{\mathcal{F}'}^{t-1}\right]$.*

Proof. The proof is similar to the proof of Lemma 2 in (Ramdas et al., 2018) with the addition of i.i.d. Laplace noise.

In a high level, we hallucinate a vector of p-values that are same as the original vector of p-values, except for the t-th index. This allows us to apply the conditional uniformity property, since now p_t is independent of the hallucinated rejections. We then connect the original rejections and the hallucinated rejections by the monotonicity of the rejections.

We perform our analysis using a hallucinated process: let $\tilde{p}^t_{1:k}$ be a copy of $p_{1:k}$ that is identical everywhere except for the t-th p-value which is set to be 1. That is,

$$\tilde{p}_i = \begin{cases} 1 & \text{if } i = t \\ p_i & \text{otherwise.} \end{cases}$$

Also let the hallucinated Laplace noises $\tilde{Z}^t_{1:k}$ be an identical copy of $Z_{1:k}$, and let $\tilde{Z}_\alpha$ be an identical copy of Z_α. The t-th value of $\tilde{Z}^t_{1:k}$ can be arbitrary since we have ensure the event $\{\tilde{p}_t > 2\lambda_t\}$, so it will fail to become a candidate and the values of $\tilde{Z}_t$ will not be relevant. We denote $\tilde{C}_{1:k}$ and $\tilde{R}_{1:k}$ as the candidates and rejections made using $\tilde{p}^t_{1:k}$, $\tilde{Z}^t_{1:k}$, and $\tilde{Z}_\alpha$.

By construction, we have $\tilde{R}_{1:t-1} = R_{1:t-1}$. On the event $\{p_t > 2\lambda_t\}$, we have $R_t = \tilde{R}_t = 0$ and $C_t = \tilde{C}_t = 0$ because $\tilde{p}_t = 1$, so both will fail to become candidates, and hence we have $\tilde{R}_{1:k} = R_{1:k}$ and the following equation holds:

$$\frac{\alpha_t I(p_t > 2\lambda_t)}{(1-2\lambda_t)h(R_{1:k})} = \frac{\alpha_t I(p_t > 2\lambda_t)}{(1-2\lambda_t)h(\tilde{R}_{1:k})}.$$

We note that when $p_t \le 2\lambda_t$, the above equation still holds since both sides will be zero. Since $\tilde{R}^t_{1:k}$ is independent of p_t, we have

$$\begin{aligned} \mathbb{E}\left[\frac{\alpha_t I(p_t > 2\lambda_t)}{(1-2\lambda_t)h(R_{1:k})}|{\mathcal{F}'}^{t-1}\right] &= \mathbb{E}\left[\frac{\alpha_t I(p_t > 2\lambda_t)}{(1-2\lambda_t)h(\tilde{R}_{1:k})}|{\mathcal{F}'}^{t-1}\right] \\ &\ge \mathbb{E}\left[\frac{\alpha_t}{h(\tilde{R}_{1:k})}|{\mathcal{F}'}^{t-1}\right] \quad (23) \\ &\ge \mathbb{E}\left[\frac{\alpha_t}{h(R_{1:k})}|{\mathcal{F}'}^{t-1}\right] \quad (24) \end{aligned}$$

where Inequality (23) is obtained by taking the expectation only with respect to p_t by invoking the conditional super-uniformity property and independence of p_t and $h(\tilde{R}_{1:k})$, and Inequality (24) follows from the facts that $R_i \ge \tilde{R}_i$ for all i and that the function h is non-decreasing.

For the second inequality in the lemma statement, we hallucinate a vector of p-values $\bar{p}^t_{1:k}$ that equals $p_{1:k}$ everywhere except for the t-th p-value which is set to be 0. That is,

$$\bar{p}_i = \begin{cases} 0 & \text{if } i = t \\ p_i & \text{otherwise.} \end{cases}$$

Also let the hallucinated Laplace noises $\bar{Z}^t_{1:k}$ be an identical copy of $Z_{1:k}$, and let $\bar{Z}_\alpha$ be an identical copy of Z_α. We denote $\bar{C}_{1:k}$ and $\bar{R}_{1:k}$ as the candidates and rejections made using $\bar{p}^t_{1:k}$ and $\bar{Z}^t_{1:k}$. By construction, we have $\bar{R}_i = R_i$ and $\bar{C}_i = C_i$ for all $i < t$. On the event that $\{\log p_t + Z_t \le \log \alpha_t + Z_\alpha - A\}$, since $\bar{p}_t = 0$ and we inject the same Laplace noise, we have $\bar{R}_t = R_t = 1$ and $\bar{C}_t = C_t = 1$, and hence also $\bar{R}_{1:k} = R_{1:k}$. Then the following equation holds:

$$\frac{I(\log p_t + Z_t \le \log \alpha_t + Z_\alpha - A)}{h(R_{1:k})} = \frac{I(\log p_t + Z_t \le \log \alpha_t + Z_\alpha - A)}{h(\bar{R}_{1:k})}.$$

We note that when $\log p_t + Z_t > \log \alpha_t + Z_\alpha - A$, the above equation still holds since both sides will be zero. Since $\bar{R}_{1:k}$ and Z_t, Z_α are independent of p_t, we can take conditional expectations to obtain

$$\begin{aligned}
\mathbb{E}\left[\frac{I(\log p_t + Z_t \leq \log \alpha_t + Z_\alpha - A)}{h(R_{1:k})}|\mathcal{F}'^{t-1}\right] &= \mathbb{E}\left[\frac{I(\log p_t + Z_t \leq \log \alpha_t + Z_\alpha - A)}{h(\bar{R}_{1:k})}|\mathcal{F}'^{t-1}\right] \\
&\leq \mathbb{E}\left[\frac{\min\{\alpha_t \exp(Z_\alpha - Z_t - A), 1\})}{h(\bar{R}_{1:k})}|\mathcal{F}'^{t-1}\right] \quad (25) \\
&\leq \mathbb{E}\left[\frac{\min\{\alpha_t \exp(Z_\alpha - Z_t - A), 1\})}{h(R_{1:k})}|\mathcal{F}'^{t-1}\right], \quad (26)
\end{aligned}$$

where Inequality (25) follows by taking expectation only with respect to p_t by invoking the conditional uniformity property and the fact that the support of p-values is $[0, 1]$, and Inequality (26) follows from the facts that $h(R_{1:k}) \leq h(\bar{R}_{1:k})$ since $R_i \leq \bar{R}_i$ for all i and that the function h is non-decreasing. $\square$