We thank all reviewers for the encouraging and helpful comments. All typos and minor comments will be fixed. Detailed responses are given below.

– – – – – – – – – – – – – – – – – – – –**Reviewer 1**– – – – – – – – – – – – – – – – – – – –

Thanks for your helpful suggestions and comments!

» **“General guideline fo section 3”:** Thanks for the suggestion, we will revise our exposition of the results accordingly.

» **“More discussions on related works”:** Thanks for sharing these related body of work that we missed. Re: [3,4,5], we will discuss these exciting progress in automating the DP proofs and the fact that SVT partially motivated them as in (Lyu et al., 2017). Re: [1,2]. Our technique should be directly applicable to the BetweenThreshold variant as in [1] and to also release the “gap” as in [2], but we need to take a closer look on [2] to ensure.

» **“challenges (and novelty) in the proof”:** Thanks for checking the fine details of our proof. We acknowledge borrowing some ideas from the classical SVT analysis as we stated up front in the paper (Line 41). This particularly refers to the “change-of-variable” trick (which exists since [DNR+09,HR10]). However, the rest of the proof of Theorem 8 is quite different as we need to (1) bound the moments of the density ratio rather than it maximum; (2) formalize the *reduction* to the RDP of subroutines. The new and delicate arguments include:

1. The application of Jensen’s inequality to bivariate joint-convex function $f(x, y) = x^{\alpha} y^{1-\alpha}$ (line 329).
2. The “fictitious query” argument (which is used in stating and proving Lemma 17) is actually important in formalizing the reduction.
3. The stopping time random variable and the use the alternative definition of expectation (line 366-374) are new.
4. Lemma 18 is somewhat cute too, in how it handles the case with ∞.

» **“Is there any specific setting where one shall always prefer Gaussian SVT?”**

Thanks for the thoughtful question. Gaussian SVT is not always preferred over Laplace SVT due to many different dimensions in comparison. But in many cases, it could work better in practice. Our observation is that, when the tail of the noise plays a significant role, e.g. the threshold T is large (Figure b in Experiment 1), Gaussian-SVT is more advantageous due to a more concentrated noise. To further improve Gaussian-SVT, the stage-wise Gaussian-SVT that uses hybrid composition outperforms Laplace-SVT significantly.

» **Question on lemma 7:** Lemma 7 is a combination of Theorem 3.25 from (Dwork and Roth, 2014) and Theorem 2 from (Lyu et al., 2017). For “strong composition"-style result, we were referring to a SVT with cut-off 1 where the dependence on c in the final bound is $O(\sqrt{c})$. To the best of our knowledge, “non-resampling" versions of SVT has not been shown to achieve such a bound, except in restricted settings (from our Theorem 11).

– – – – – – – – – – – – – – – – – – – –**Reviewer 4**– – – – – – – – – – – – – – – – – – – –

Thank you for your helpful comments!

» **“There seems to be a mistake in the parameter settings of ϵ_1 and ϵ_2 in Lemma 7.”**

Thanks for spotting the typo, it shall be $\frac{\triangle}{\epsilon_1} = \frac{\sqrt{32c\log(1/\delta)}}{\epsilon}$.

» **“Is the expectation on the stopping time necessary?”**

This is an excellent observation. We have good reasons to believe that it is necessary unless pure-DP subroutines are used. A lower bound would be nice, but is beyond the scope of this paper. We focused on further bounding that conditional moments involving the stopping time RV by considering a non-negative query class (see Line 156-170).

» **Remark in lines 218-220: doing RDP to DP in each round and then doing optimal DP composition on each one of these?** The reviewer might have missed the motivation of the stage-wise approach. Direct composition of the RDP bound in Theorem 8 with a single conversion to (ϵ, δ)-DP does not allow a $O(\sqrt{c})$ scaling unless δ is very small, see our discussion in line 188-199. Line 218-220 is a comment about the choice of c', k' in Theorem 12 and 13.

» **“Only Laplace or Gaussian noise can be used?”** Any noise-adding procedures that admit an RDP bound can be used, e.g., the optimal stair-case noise from (Geng and Viswanath, 2014).

– – – – – – – – – – – – – – – – – – – –**Reviewer 5**– – – – – – – – – – – – – – – – – – – –

Thanks for your kind comments! We will add experiments related to adaptive data analysis following [DFHPRR-14] in the appendix to demonstrate the merits of Gaussian-SVT. Notice that they added Gausssian noise, rather than the Laplace noise they analyzed, so we are confident that the experimental results will be to our favor.