We thank the reviewers for their thoughtful observations and comments. Their input will be used to improve our paper. Overall, we are encouraged by the reviewers' reaction to our work. We don't see the need to go over the high level summary of our contributions at this time, given that all reviewers seem to be clear on this point. Next we address some outstanding comments, hoping that our responses clear any remaining concerns.

Response to Reviewer 1

Regarding a roadmap, equation numbering, result labels

These are all good ideas, thank you.

At this level of review, the paper generally appears correct. The use of base 2 for purposes of illustration (subsection "Weight Decomposition" in Section 5) does not actually work following the proof of Lemma 9 (cf. lines 294-295); it would be better to use base 3/2 from the start and state that the choice is explained via the proof of Lemma 9.

We chose base 2 because it makes it clear that the number can be efficiently decomposed. We expect more readers to be confused by another base. We can explain right after the binary decomposition that a base 3/2 is needed instead later. If the reviewer feels strongly about this, please let us know in the "post-rebuttal" so that we may reconsider.

Response to Reviewer 2

(1) The approach of the paper is significantly more complicated than that of Malach et al.
(2) The hyperbolic sampling distribution does not reflect what is done in practice.

Indeed our analysis is more complicated than that of Malach et al., but this is quite often the case for improvements on first simple theoretical results. We believe this is certainly worth the price. Regarding hyperbolic sampling, we agree that uniform sampling is more traditional and widespread, however our theory also suggests that hyperbolic sampling may be worth investigating in practice.

Response to Reviewer 3

Regarding https://arxiv.org/pdf/2006.07990.pdf and https://www.ics.uci.edu/ lueker/papers/exppart.pdf:

Thank you for these references. As you note, Pensia et al. came out after the submission deadline and we were indeed not aware of Lueker (1998). Both are quite nice pieces of work. Lueker's result is astounding and can indeed be used for uniform sampling, with the caveat that you note. We will certainly include discussions on these two papers in the final version.

(The idea of discarding the samples (which motivates the batch sampling) is a bit irritating imho, since all the weights are choosen iid.)

We are not entirely sure what precisely is irritating, but we can at least say early on that discarding samples is a naive approach if this helps.

Response to Reviewer 4

I find that the exposition of the results is not clear. For example, at Line 66-68, the complexity is described per "target weight", but the results presented in Theorem 3 are not. I find it difficult to digest the results (Remark 4 should be mentioned earlier).

Thank you for spotting some reading hurdles, which we will strive to fix. We will at the very least add a forward reference to Remark 4 in the revision.