We thank the reviewers for their positive feedback and will revise the paper accordingly. The following are answers to the questions posed by each reviewer:

Reviewer 1

The CLT assumption does not affect the privacy of the method. If the assumption is heavily violated during an MCMC run, the method still guarantees privacy but the result might be far from the true posterior. We do not treat the CLT in more depth in the paper since this is one of the central themes of Seita et al. 2017, and we do not claim any contribution to this theory. This is clearly one of the most interesting questions for further research.

Reviewer 2

We will include a list of the different factors to improve clarity:

1. Common parameters
   (a) α parameter for RDP.
   (b) T number of iterations.
   (c) N full data size.
2. DP MCMC with full data
   (a) $C \in (0, \pi^2/3)$ freely chosen constant for noise variance: higher C improves privacy but also increases decomposition error.
   (b) B assumed bound for the log-likelihood ratios (llr) w.r.t. data OR the parameters. This holds for Lipschitz llr, or it can be enforced by using small enough proposal variance (for suitable models), or by clipping (for any arbitrary model). Smaller B improves privacy, while smaller proposal variance means the chain moves slower, and tighter clipping means less accurate posterior estimation.
3. DP MCMC with subsampling
   (a) b batch size s.t. $\alpha < b/5$: smaller b means better privacy amplification but generally also more error in the CLT approximation.
   (b) Set noise variance $C = 2$ (analysis could be done for other values as well).
   (c) Assume llr w.r.t. the parameters $\leq \sqrt{b}/N$
   OR
   optionally with tempering: choose β in the coarsened posterior s.t. $N_0 = N\beta/(\beta + N)$ and assume llr w.r.t. the parameters $\leq \sqrt{b}/N_0$.
   As above, either condition holds for Lipschitz llr or can be enforced by tuning the proposal variance or by clipping.

We agree that a good comparison of different approaches to DP samplers would be interesting. However, we feel that this is out of scope for the current submission, since the various methods have different assumptions.

Reviewer 3

We will clarify our contribution as opposed to the Seita et al. 2017 paper more clearly starting already from the Introduction.

The decomposition idea is discussed more extensively by Seita et al. 2017 (for achieving subsampling without any privacy notion). In our paper, most of the discussion regarding the decomposition is in the Supplement, while in the main paper we focus more on the privacy.

The privacy setting we consider is the standard centralised setting for DP: a single trusted party (data curator) has access to all the data and runs the MCMC algorithm with the aim of releasing a trained DP model and/or the samples from the chain while protecting the privacy of the training data (with one sample corresponding to one individual). The adversary has access to the trained model/samples from the chain and (almost) arbitrary side information as per DP definition. We will clarify this in the final version.