We thank the reviewers for their useful feedback and their time. We are happy to see that all the reviews are positive and we appreciate that the reviewers found our article clear and of high quality.

We have corrected all the minor comments, as suggested. We now provide specific answers to each reviewer below.

Reviewer #1:

We thank the reviewer for their positive evaluation of our work and their comments.

"All the theoretical contributions seem to me a bit marginal" Since the Sliced-Wasserstein distance is an average of one-dimensional Wasserstein distances, the proofs for the existence, measurability and consistency of Sliced-Wasserstein estimators and the Central Limit Theorems (CLTs) are, indeed, inevitably similar to the proofs done in Bernton et al. [1]. However, the adaptation of these techniques was made possible by the derivation of novel properties regarding the topology induced by the Sliced-Wasserstein distance, which have never been investigated before and whose proofs differ from the ones in [1]: in Theorem 1, we show that the convergence in Sliced-Wasserstein implies weak convergence using non-standard techniques, and the same observation holds true for Lemma S6 (in the supplementary document), which establishes the lower semi-continuity of Sliced-Wasserstein. We believe that these two results are important contributions on their own and explain why the adaptation of [1] is not straightforward in the first place. Specifically, without Theorem 1, the formulation and use of Assumption A2 in our proofs would not make sense: see lines 174 and 221 in the supplementary document. We will explain these observations more explicitly to clarify our contributions.

On the other hand, we appreciate that the reviewer describes our Central Limit Theorems as *"a very nice observation"*. We think that it is another significant result compared to [1] since it applies to the multidimensional case, as the reviewer noticed. In their conclusion, [1] conjectures that the rate of the minimum Wasserstein estimators would "depend negatively on the dimension of the observation space rather than that of the parameter space", which suggests that the rate would suffer from the curse of dimensionality. Our result shows that it is not the case for the Sliced-Wasserstein distance. This curse of dimensionality has created a pessimism in the machine learning community about the use of Wasserstein-based methods in large dimensional settings (e.g. this fact has been popularly used for motivating regularized optimal transport). We believe that our findings provide an important counter-example to this conception and that the derived CLTs are thus another key contribution to the field. We will underline this observation in the paper.

"Question for the authors: Does the obtained distributional limit here gives any intuition about what should be the limit in the multidimensional case for the vanilla wasserstein distance?" This is a very interesting question. We believe it is definitively worth investigating whether our findings would help in the derivation of a CLT for the Wasserstein distance in the multidimensional case; however, we leave it out of the scope of this study.

Reviewer #2:

We thank the reviewer for the positive evaluation. We appreciate that they pointed out the importance of the problem and of our analysis given the rise in popularity of computational optimal transport.

"The analysis is quite compactly presented in 10 pages and it could be useful to consider a longer version where the key findings are better elaborated in more details." As suggested, we will add a new section to the supplementary document to further elaborate our theoretical findings.

Reviewer #3:

We thank the reviewer for giving a positive evaluation and a useful feedback. In particular, the reviewer noticed two minor problems:

"Line 104, the second $\mathcal{P}_p$ should be $\mathcal{P}$" We agree. We fixed this typo in the manuscript.

"In thoerems 5 and 6, since the limit distribution is degenerated, why not claim the convergence as "in probability"?" The limit distribution is in fact not degenerated since $G_\star$ is a random element: see Assumption A8. Therefore, we can not claim that the convergence in distribution derived in these theorems implies the convergence in probability. We will add further clarification in the manuscript to reflect this comment and avoid any confusion.

References

[1] E. Bernton, P. E. Jacob, M. Gerber, and C. P. Robert. On parameter estimation with the Wasserstein distance. *Information and Inference: A Journal of the IMA*, Jan 2019.