Firstly, we would like to thank all the reviewers for their careful reading and valuable suggestions. We will next address some of the main concerns raised in the reviews, as well as answer some of the questions that were posed.

Although the analysis of the secretary problem augmented with predictions (from a technical perspective) is mostly there to provide a clean and illustrative example, we give, in our opinion, a non-trivial analysis of the augmented online bipartite matching and the graphic matroid secretary problem. For these problems, we design novel algorithms whose analysis requires non-trivial probabilistic arguments, and crucially relies on the specific order of the procedures carried out in the Phases II and III. Albeit, due to space limitation, we defer them to the supplementary material:

i) for the online bipartite matching problem – see pages 14-15;

ii) for the graphic matroid secretary problem – see pages 19-20 and Appendix E.

This will be emphasized more in the paper.

We, indeed, do not provide any lower bounds for our algorithms, apart from the trivial lower bound induced by the original secretary problem. We agree that finding lower bounds is an important question (and based on personal experience, a difficult one) that would be very interesting to be addressed in future work. Further, finding a more general meta-property under which our results hold, would be very intriguing.

We also want to mention that λ and c in the analysis of the secretary problem are independent parameters that provide the most general description of the competitive ratio. Here λ is our confidence of the predictions and c describes how much we are willing to lose in the worst case. Although these parameters can be set independently, some combinations of them are indeed not very sensible, as one might not get an improved performance guarantee, even when the prediction error is small (for instance, if $c = 1$, i.e., we are not willing to lose anything in the worst case, then it is not helpful to consider the prediction at all). In particular, as one does not know the prediction error, there is no way in choosing these parameters optimally, since different prediction errors require different settings of λ and c.

For the graphic matroid secretary problem, the predictions model the optimal spanning tree in the case when all edge-weights are pairwise distinct. Otherwise, there can be many (offline) optimal spanning trees, and thus the predictions do not encode a unique optimal spanning tree. We intentionally chose not to use predictions regarding what edges are part of an optimal solution, as in our opinion, such an assumption would be too strong.

With respect to the Lambert-W function, it appears naturally in the analysis of the classical secretary problem. In the classical analysis, 1/e is exactly the optimum solution where the two branches of the Lambert-W function "meet". We discuss this in more detail in the Preliminaries section and the proof sketch of Theorem 1.2.

There is indeed a missing expression "for any $c \geq 1$" in the statement of Theorem 1.2. Thank you for pointing this out.