We would like to sincerely thank all the reviewers for their very valuable feedback. Below we address reviewers' comments in more detail.

Overview: We are happy that the reviewers appreciate the importance of adaptive methods in ES and point out lots of applications and possible extensions of the presented results. We will discuss them as well as some of our own extension ideas in the final version.

Additional Experiments: In the final version we will include additional suggested experiments for linear policies and on RL tasks with no termination condition, that as noted by reviewers, might be particularly amenable to adaptive methods. Additionally, we agree that it would be very interesting to test ASEBO on even higher dimensional tasks such as Humanoid. This is well aligned with very recent research on intrinsic dimensionality for RL objective landscapes suggesting that tasks such as Humanoid exhibit much lower intrinsic dimensionality [1]. It would be useful to see how ASEBO can take advantage of this and thus we will include the results of these experiments in the final version.

Theory Clarification: We will also add a paragraph summarizing our theoretical findings. In particular, we will emphasize the adaptivity aspects (e.g. non-sensitivity to fixed hyperparameters). We will comment on Theorem 3.3 explaining that it shows the presented algorithm automatically finds the optimal (in terms of variance reduction) strategy of sampling from ES-active subspaces without any pre-tuned hyperparameters. In Theorem 3.3 this is expressed via convergence result through loss function l. We will clarify it.

Open Sourcing: Given the potential impact of adaptive methods to scale ES algorithms, we will release an open-source version of the algorithm with the final version of the paper (that will also contain an updated set of benchmark tests). We hope this will help other researchers in applying our algorithms for their RL problems, as well as developing them further.

Additional Clarifications:

- The curve for $n = 100$ is indeed hidden behind $n = 212$ in Fig.3 (as both failed to learn). Thank you for pointing this out. We will make this clearer in the final version.
- In the final version we will also simplify algorithmic block 2, as suggested and give an additional explanation in a separate paragraph.
- We will also add extra definitions and explanations of the used concepts that will be presented at the beginning of the paper, as recommended, to improve clarity.
- We will replace one of the plots in Fig.1 with the plot visualizing how the gradient direction changes as training progresses.
- We will also fix all listed typos regarding notations.

References

[1] C. Li, H. Farkhoor, R. Liu, and J. Yosinski. Measuring the intrinsic dimension of objective landscapes. In *International Conference on Learning Representations*, 2018.