

1 We thank the reviewers for their valuable comments. We reply to each outstanding point below.

2 **REVIEWER 1. R.** *Extension to meta-reinforcement learning.* **A.** One starting point in this direction would be to  
3 consider the simplified setting of contextual bandits with linear reward functions. However, even in this case the  
4 extension of the proposed method does not seem straightforward and it requires further investigation. **R.** *...scaling up*  
5 *to...rich observation spaces?* **A.** The scaling properties of our method depend on the specific setting considered. For  
6 instance, for the bias setting, our method is based on the combination of Alg. 5 & Alg. 6 and it scales linearly with  
7 the dimension of the input space. We thus expect that in this setting the method will be appropriate also for datasets  
8 in more rich observation spaces. In the feature map setting, our method in Alg. 7 & Alg. 8 requires to compute the  
9 eigenvalue decomposition of a rank one perturbation of the current matrix. This can be performed using methods such  
10 as [On the efficient update of the singular value decomposition, Stange, 2008], which essentially scale quadratically  
11 with respect to the input dimensionality. An interesting alternative here may be to use Frank-Wolfe as meta-algorithm,  
12 which requires to compute only the maximum eigenvalue. However the better scaling property of this method comes at  
13 the price of a slower learning/convergence rate. **R.** *Related work in supplementary.* **A.** We agree, we will move the  
14 discussion on previous work to the main body. Thanks also for the additional reference, which we will add to the paper.

15 **REVIEWER 2. R.** *I expected more meta-learning insights. It seems that the proposed approach is more a theoretically*  
16 *guaranteed solver ...* **A.** As described in the paper, the proposed method takes inspiration from multitask learning  
17 (MTL) and, as such, it allows to translate key MTL insights to meta-learning. In the paper, we particularly stress  
18 this aspect in Sec. 6, when we specialize our analysis to the settings of the bias and the feature map. In such cases,  
19 we provide an in depth interpretation of the bounds from the meta-learning perspective. We will highlight more the  
20 novelty of the proposed method from the meta-learning point of view also in the rest of the paper. **R.** *Related work in*  
21 *supplementary.* **A.** We agree, we will discuss the most related work in the main body. **R.** *...how the current theoretical*  
22 *framework helps practical algorithm design.* **A.** The proposed framework provides us with a flexible meta-learning  
23 scheme, which is able to cover various kinds of tasks' relatedness and a wide family of inner/meta algorithms, by  
24 choosing in an appropriate way the complexity terms ( $f$  and  $F$ ) and the aggressiveness of the updates. For instance, the  
25 feature learning setting can be immediately extended to other structured sparsity frameworks, by considering alternative  
26 choices for the set  $\Theta$ , different from the matricial simplex. Further examples in our framework are e.g. online Gradient  
27 Descent, Matrix Exponentiated,  $p$ -norm and Follow-The-Regularized-Leader. We will discuss this in the paper.

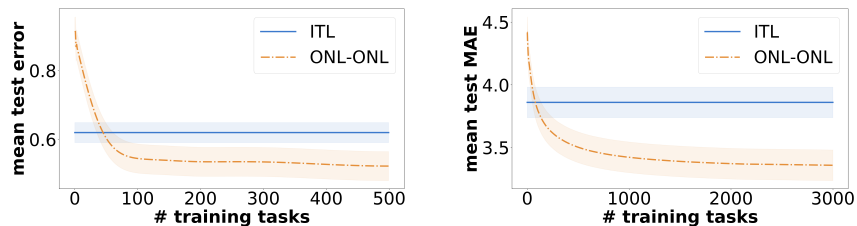


Figure 1: (Left) Mini-Wiki Dataset. (Right) Jester1 Dataset.

28 **REVIEWER 3. R.** *... the proposed error bound seems novel although is related to previous work. I'm not quite sure*  
29 *about the significance of this submission.* **A.** We care to point out that the framework proposed in this work offers key  
30 improvements with respect to previous work in the meta-learning literature. First, differently from recent related work  
31 (see references [9,14] listed in the paper), the online-within-online method we propose can be adapted to a wide class of  
32 algorithms. Second, our analysis is innovative in that it allows us to provide guarantees for our method in the adversarial  
33 setting under very standard assumptions, by leveraging the delicate interplay between the within-task problem and  
34 the outer-task problem. On the contrary, other papers, such as [14,18] require to introduce stronger assumptions like  
35 growth conditions for the loss function. Third, note that when we apply our analysis to the statistical setting, we provide  
36 guarantees for the average of the estimators returned by our method, while in other papers [1,18] the authors provides  
37 guarantees for one estimator sampled from the pool, which requires the memorization of all the estimators and adds  
38 randomness to the process. **R.** *The authors should present more experiments to validate the proposed method.* **A.** We  
39 further validated the proposed method on two additional experimental settings satisfying all the assumptions of our  
40 theory. Specifically, we considered 1) a multi-class classification problem over the Mini-Wiki dataset from [18] and  
41 2) a matrix-completion (reformulated as regression) problem on the Jester1 dataset in [Eigentaste: A constant time  
42 collaborative filtering algorithm, Goldberg et al., 2001], containing user ratings of jokes. In the Mini-Wiki Dataset we  
43 have 813 tasks, 128 available points per task and 50 dimensions. In the Jester1 dataset we have 24983 tasks (users), 100  
44 points (jokes) for each task and 100 dimensions. In fig. 1 we report the averaged test performance of our method in the  
45 feature map framework on these datasets. For the Mini-Wiki dataset we used the multi-class hinge loss, for the Jester1  
46 dataset the absolute loss. Notice that in both cases, coherently to what already observed in the paper, our meta-learning  
47 approach (ONL-ONL) reveals to be an effective approach to transfer the knowledge among the tasks in comparison to  
48 solving each task independently (ITL).