[To Reviewer #1] Thanks for your constructive comments. We answer your main questions as follows.

Question 1. "Is there any hope to avoid the $\log(T)$ blowup in runtime? Is it at least possible to avoid having to aggregate all experts from both algorithms to get the desired best-of-both worlds bound?"
Answer 1. From our current understanding, the meta-expert aggregation is the standard framework for handling the uncertainty of non-stationary online learning. For your second question, yes, we can design an algorithm that avoids aggregating all experts from both algorithms to achieve $\mathcal{O}(\sqrt{(1 + P_T + \min\{\bar{V}_T, F_T\})(1 + P_T)})$ dynamic regret, where $\bar{V}_T = \sum_{t=2}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|^2$. To do so, we require the expert-algorithm to run *optimistic gradient descent* with the learned optimism. We will add a remark in the paper to discuss this point more thoroughly.

Question 2. "Technically, I think in order for Lemma 4 to hold, f needs to be defined on the whole vector space"
Answer 2. Thanks for your comments! You are correct. The issue has also been identified by Reviewer #3. We can resolve it by requiring f_t to be defined on $\mathbb{R}^d$, and we will correct it in the revised version.

[To Reviewer #2] Thanks for your helpful comments, and we address your concerns as follows.

Question 1. about improving paper by spending more efforts in explaining the advantages of universal dynamic regret
Answer 1. Thanks for your suggestion. Universal dynamic regret supports arbitrary comparator sequence, and thus it subsumes the worst-case dynamic regret and static regret as special cases. So it is a more adaptive performance measure for non-stationary online learning. We will improve the paper writing to make this point more clear.

Question 2. "what regret ... if ... only access to 1 gradient query per step, rather than the two used in OEGD."
Answer 2. If the expert-algorithm only has access to 1 gradient per step, $\mathcal{O}(\sqrt{(1 + P_T + \bar{V}_T)(1 + P_T)})$ dynamic regret is attainable, where $\bar{V}_T = \sum_{t=2}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|^2$ is slightly different from the gradient-variation V_T: V_T takes the sup over all $\mathbf{x}$ and thus is fully problem-dependent; while $\bar{V}_T$ will depend on the decision path $\{\mathbf{x}_t\}_{t=1}^{T}$.

[To Reviewer #3] Thanks for your appreciation and insightful review! We address your main questions as follows.

Question 1. "how would the lower-bound of function appear in your bounds if we assume they are not positive"
Answer 1. The variation bound (Theorem 3) does not require this assumption; while it is necessary for small-loss bound (Theorem 5), because the self-bounding property only holds for non-negative smooth functions.

Question 2. "how would the algorithms / results change if $\mathbf{0}$ is not in $\mathcal{X}$?"
Answer 2. There are three places we use this assumption:

- Line 544: we use $\|\mathbf{x}_t\|_2 \le D$ to obtain path-length P_T. Note that actually we can bound this term without assuming $\mathbf{0} \in \mathcal{X}$, by an alternative argument (similar to the analysis of term (ii) in Line 450).
- Line 148: we use $\|\mathbf{x}_{t,i}\|_2 \le D$ to bound adaptive quantity D_∞. We clarify that we can bound the term without $\mathbf{0} \in \mathcal{X}$, by alternatively setting surrogate loss as $\ell_{t,i} = \langle f_t(\mathbf{x}_t), \mathbf{x}_{t,i} - \mathbf{x}_t \rangle$ and optimism as $m_{t,i} = \langle f_{t-1}(\bar{\mathbf{x}}_t), \mathbf{x}_{t,i} - \mathbf{x}_t \rangle$ (note that the term $\mathbf{x}_t$ in optimism will be canceled out in the weight update equation).
- Line 150: we use $\|\mathbf{x}_{t,i}\|_2 \le D$ to bound the term $\|\mathbf{x}_t - \bar{\mathbf{x}}_t\|_2$. It seems necessary to assume $\mathbf{0} \in \mathcal{X}$, and we do not figure out how to eliminate the requirement yet.

About the self-bounding property of smooth functions, you are absolutely correct. We will resolve it according to your suggestion. For other minor issues, we will carefully revise the paper according to your constructive comments.

[To Reviewer #4] Thanks for your review. Below we address your concerns and clarify the misunderstandings.

Question 1. "In my opinion, such a paper must demonstrate the results on several tasks (even on toy problems)."
Answer 1. We performed the empirical studies in the rebuttal period. Due to page limits, we present comparisons of OGD [1], Ader [10] and Sword (our method) only on synthetic non-stationary data in Figure 1, which clearly shows the effectiveness of our algorithms. More results and elaborations will be included in the revised version.

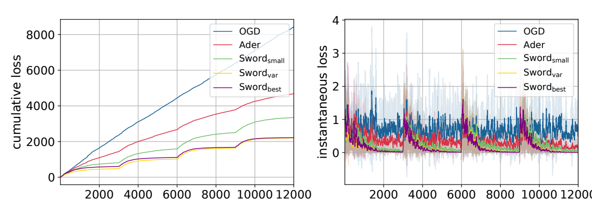

Figure 1: cumulative loss (left); instantaneous loss (right).

Question 2. "The novelty of the paper is limited. The algorithm is a combination of known and well-studied ideas. "
Answer 2. We cannot agree with this claim. We highlight the challenge and technical contributions in Line 62–78 of the paper. Note that a trivial combination of existing meta-algorithms and expert-algorithms cannot obtain desired variation bounds. We resolve the challenge by carefully designing the optimism for OptimisticHedge used as the meta-algorithm (see more discussions in Line 122–127 & 146–151). The lower bound for universal dynamic regret is also novel.

We hope the reviewer could check the comments from other reviewers. For example, Reviewer #1 said "The paper motivates the problem well. The lower bound for regret at least P is nice to have.". Reviewer #2 said "On a technical level, the authors provide an interesting construction of a new online learning algorithm that brings together key ideas from various existing procedures.". Reviewer # 3 said "The results will be interesting to the online learning community, especially given the new interest in dynamic regret and meta-gradient algorithms."

Overall, we believe our work has sufficient novelty and contributions to the online learning community.