REVIEWER 2 Thank you for your encouraging comments.

REVIEWER 3 Thank you for your comments. In order to help clarify our contributions and organize them for readers, we provide the following table to summarize the differences between regrets.

Regret	Non-convex Models	Concept Drift	Update Rule
Standard Regret	✗	✗	$x_{t+1} = x_t - \frac{\eta}{\sqrt{t}} \hat{\nabla} f_t(x_t)$
Static Local Regret (Hazan et al.)	✓	✗	$x_{t+1} = x_t - \frac{\eta}{w} \sum_{i=0}^{w-1} \hat{\nabla} f_{t-i}(x_t)$
Dynamic Local Regret (Ours)	✓	✓	$x_{t+1} = x_t - \frac{\eta}{W} \sum_{i=0}^{w-1} \alpha^i \hat{\nabla} f_{t-i}(x_{t-i})$

REVIEWER 4 Thank you for your comments. First we provide a toy example and some additional theoretical motivation for our regret in response to the following comment:

Without some formal notion or even toy scenario for concept drift, it's not clear what theoretical basis there is to prefer this notion of regret to other notions other than some vague heuristics.

Motivation via a Toy Example We demonstrate the motivation of our dynamic regret via a toy example where the static local regret fails. Concept drift occurs when the optimal model at time t may no longer be the optimal model at time $t+1$. Consider an online learning problem with concept drift with $T = 3$ time periods and loss functions: $f_1(x) = (x-1)^2, f_2(x) = (x-2)^2, f_3(x) = (x-3)^2$. Obviously, the best possible sequence of parameters is $x_1 = 1, x_2 = 2, x_3 = 3$. Call this the *oracle policy*. Also consider a suboptimal sequence, where the model does not react quickly enough to concept drift: $x_1 = 1, x_2 = 1.5, x_3 = 2$. Call this the *stale policy*. The values of the *stale policy* were chosen to minimize Static Local Regret. Recall the formulation of static and dynamic local regrets:

$$HR_3(3) = \left\| \frac{\nabla f_3(x_3) + \nabla f_2(x_3) + \nabla f_1(x_3)}{3} \right\|^2 + \left\| \frac{\nabla f_2(x_2) + \nabla f_1(x_2)}{3} \right\|^2 + \left\| \frac{\nabla f_1(x_1)}{3} \right\|^2 \qquad \text{(Hazan's)}$$

$$PR_3(3) = \left\| \frac{\nabla f_3(x_3) + \nabla f_2(x_2) + \nabla f_1(x_1)}{3} \right\|^2 + \left\| \frac{\nabla f_2(x_2) + \nabla f_1(x_1)}{3} \right\|^2 + \left\| \frac{\nabla f_1(x_1)}{3} \right\|^2 \qquad \text{(Ours)}$$

Note that, for the local regrets, we use $w = 3$ and assume $f_t(x) = 0$ for $t \leq 0$. We also set $\alpha = 1$ for our Dynamic Local Regret but other values would not change the results for this example. The formulation of the Standard Regret is $\sum_{t=1}^{T} f_t(x_t) - \min_x \sum_{t=1}^{T} f_t(x)$. Although the *oracle policy* achieves globally minimal loss, Hazan et al.'s Static Local Regret favors the *stale policy*. We can verify this by computing the loss and regret for these policies, as shown in the table below.

Regret	Oracle Policy	Stale Policy	Decision
Cumulative Loss	0	5/4	Oracle policy is better
Standard Regret	-2	-3/8	Oracle policy is better
Static Local Regret (Hazan et al.)	40/9	4/9	Stale policy is better
Dynamic Local Regret (Ours)	0	10/9	Oracle policy is better

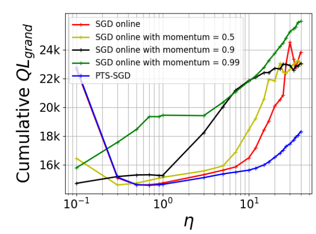

Figure 1: SGD online with momentum

Theoretical motivation via Calibration: A more formal motivation of our regret can be related to the concept of calibration [1]. The comment on line 110 can be rewritten as: *If the updates* $\{x_1, \cdots, x_T\}$ *are* ***well-calibrated****, then perturbing* x_t *by any* u *cannot substantially reduce the cumulative loss.* Hence, it can be said that the sequence $\{x_1, \cdots, x_T\}$ is asymptotically calibrated with respect to $\{f_1, \cdots, f_T\}$ if: $\limsup_{T\to\infty} \sup_{\|u\|=1} \frac{\sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x_t+u)}{T} \leq 0$. Consequently, using the first order Taylor series expansion, we can write the following equation that motivates the left hand side of the equation 3 in the paper: $\limsup_{T\to\infty} \sup_{\|u\|=1} -\frac{1}{T} \langle u, \nabla f_t(x_t) \rangle \leq 0$. Thus our regret ensures asymptotic calibration.

This analysis was dropped for simplicity, but thanks to the reviewer's comments we will put this analysis back into the paper.

Next we provide some additional discussion of momentum to address the following comment:

However, as the authors note, previous work has shown that PTS-SGD coincides with SGD with momentum as the decay factor approaches 1, so it seems like a better empirical comparison might be with SGD with momentum.

We indeed ran experiments using SGD with momentum for various decay parameters and concluded that SGD with momentum is not even as stable as SGD-online (standard SGD without momentum) as shown in Figure 1. Our PTS-SGD is still more robust to the learning rate. On the other hand, we observed that SGD with momentum yields better accuracy for offline learning (results are not shown here). We will add these results to the paper.

References

[1] Dean P Foster and Rakesh V Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998.