We genuinely appreciate all three reviewers' (#1,#2,#3,#4) valuable suggestions to strengthen our paper. We have addressed all raised questions below: conducting new experiments to compare with hand-designed optimizers (#1) and on more complex optimizee architectures (#1,#3); clarifying the experiment's observations (#2,#4); improving the presentation (#1,#4) and fixing notations (#2). We will also fix all typos (#1) in our final version, and confirm our promise to release our source codes and all trained models publicly, upon NeurIPS 2020 acceptance.

▷ Reviewer #1. **Q1. Comparison against hand-designed optimizers.** *Reply:* We follow your suggestions to conduct comparison against hand-designed optimizers on MLP-orig and Conv-MNIST optimizees with MNIST dataset, as presented in Figure 1 [a] and [b]. Plugging our proposed enhanced training techniques immediately boosts L2O to outperform analytical optimizers (i.e., Adam, RMSProp, and SGD) whose hyperparameters have been optimally tuned via a grid search, while the vanilla L2O-DM quickly collapses and diverges in the same setting. **Q2. Complex optimizees (e.g. searched architecture from NAS).** *Reply:* We evaluate our enhanced L2O on a challenging optimizee, NAS-CIFAR, from the popular NAS-Bench-201's search space [1]. NAS-CIFAR[1] consists of multiple skip connection, convolution, and average pooling operations, which is significantly different with simple MLPs and CNNs. As shown in Figure 1 [c], our proposed techniques enable L2O trained on single-layer MLP to generalize robustly to the much more sophisticated NAS-CIFAR, where vanilla L2O fails. We further validate our proposed training techniques can scale up to LeNet. More details are referred to Reviewer #3's Q2. **Q3. Readability of figure 4 and line 108.** *Reply:* We sincerely appreciate your suggestion and will revise the caption in figure 4 for better readability. The analogy in line 108 is to unify our enhanced training framework under an RL view, which facilitates us to introduce core concepts like the exploration-exploitation balance and imitation learning. We will make sure to picture this more clearly in our final version. **Q4. Typos and Organization.** *Reply:* We will follow your suggestion to use the continue/break statement, fix all typos, and remove the confusing comma in line 272.

▷ Reviewer #2. **Q1. IL generalizes better than CL from MLP to CNN.** *Reply:* It is a great observation. The curriculum learning (CL) technique mainly helps alleviate the notorious L2O truncation bias, while the imitation learning (IL) approach mainly helps L2O refer to the generally applicable optimization rules and hence avoid overfitting specific optimizee structures. Thus, IL appears to be a main contributor for L2O generalizing across different optimizees. In addition, we note that L2O equipped with both IL and CL achieves the best performance, as shown in figures 2 and 5. **Q2. Notation in line 98.** *Reply:* Thank you. We will change the formulation $f(\boldsymbol{\theta}_t)$ to $f(\boldsymbol{\theta}_t, \boldsymbol{\phi})$.

▷ Reviewer #3. **Q1. Connections with RL.** *Reply:* We acknowledge Reviewer #3's comments and agree that our proposed techniques his rooted in RL literature, but is uniquely motivated by and suitable for L2O. **Q2. Apply L2O on LeNet.** *Reply:* We verify that our techniques can scale up the full LeNet: see Figure 1 [d].

▷ Reviewer #4. **Q1. Why imitation learning is superior to self-improving.** *Reply:* Great question. We believe that the key point lies in the *long-term coherency* of optimization trajectories. The imitation learning technique allows L2O to learn from the entire trajectories, with hand-crafted optimizers serving as end-to-end guidance. In comparison, self-improving breaks each optimization trajectory into mixed local "pieces" of applying either analytical or learned optimizers, therefore restricting learned optimizers to capturing only the dependency within local segments (e.g., a few iterations). Experimental comparison details are referred to section 5.3. **Q2. Ablation of proposed strategies.** *Reply:* In fact, we already compared the performance across L2O-DM-CL-IL, L2O-DM-CL and L2O-DM-IL, as presented in figures 2 and 5 (we will add the curve of L2O-DM-CL-IL to figure 5 for better visibility). We observe: i) On MLP optimizees, curriculum learning (CL) improves more than imitation learning (IL); On CNN optimizees, IL contributes more to the performance gain (please also check our explanation in answering Reviewer #2's Q1). ii) Combining CL and IL techniques always enjoys extra performance boost, compared to using either alone. **Q3. Readability of the method section.** *Reply:* We will revise our method's presentation to draw a tighter connection with our design philosophy of proposed techniques. For example, the above answer to Q1 will be integrated when we introduce IL.

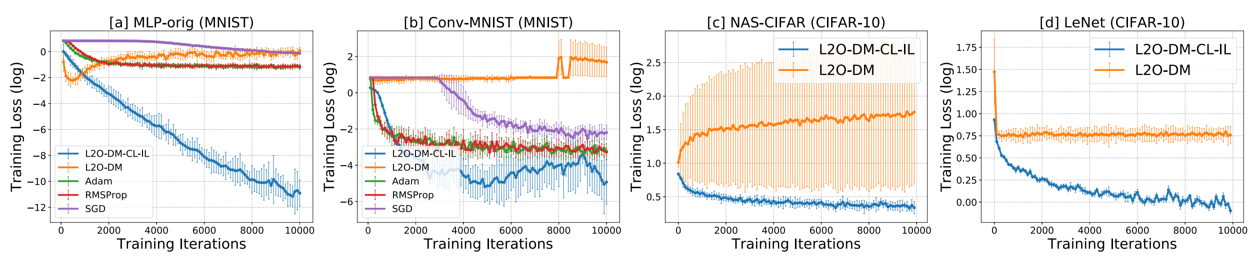

Figure 1: **Ten runs.** [a]/[b] are comparisons against analytical optimizers (@R#1). [c]/[d] are complex optimizees (@R#1,@R#3).

[1] Xuanyi Dong and Yi Yang. Nas-bench-201: Extending the scope of reproducible neural architecture search. In ICLR, 2020.

[1] Detailed Architecture: |nor_conv_3x3∼0|+|nor_conv_3x3∼0|avg_pool_3x3∼1|+|skip_connect∼0|nor_conv_3x3∼1|skip_connect∼2|, the demo architecture in [1]'s GitHub repository `https://github.com/D-X-Y/NAS-Bench-201`.