# Supplementary Materials for Multi-Objective Online Learning

**Anonymous Author(s)**
Affiliation
Address
`email`

Supplementary materials are organized as follows. Appendix A reviews related work. Appendix B validates the correctness of our definition of PSG. Appendix C discusses the physical meaning of our proposed multi-objective static regret. Appendix D provides more details of DR-OMMD, including a discussion on the lazy version of mirror descent and how to efficiently compute the composition weights for the min-regularized-norm solver. Appendix E proves the tightness of DR-OMMD's bounds and gives a more general dynamic regret bound for arbitrary temporal variability. Appendix F supplements more experimental results, including more results of adaptive regularization and an addition experiment in the online non-convex setting. In Appendix F.4, we discuss how to set hyperparameters when $V_T$ is unknown in the dynamic setting. Appendix G, H, and I provide detailed proofs of all theoretical claims in this work. Note that in Appendix D we fix a typo in the solution of min-regularized-norm (line 277), and in Appendix E we fix a typo in the remark below Theorem 4.4 (line 304); both typos do not affect the following parts.

## A Related Work

In this section, we review previous work in three related fields, i.e., online learning, multi-objective optimization, and multi-objective multi-armed bandits.

### A.1 Online Learning

Online learning arms to make sequential predictions for streaming data. Please refer to the introduction books [13, 23] for more background knowledges.

Most of the previous works on online learning are conducted in the single-objective setting. As far as we are concerned, there are only two lines of work concerning multi-objective learning. The first line of works provides a multi-objective perspective of the prediction-with-expert-advice (PEA) problem [16, 17]. Specifically, they view each individual expert as a multi-objective criterion, and characterize the Pareto optimal trade-offs among different experts. These works have two main distinctions from our proposed MO-OCO. First, they are still built upon the original PEA problem where the payoff of each expert (or decision) is a scalar, while we focus on vectoral payoffs. Second, their framework is restricted to an absolute loss game, whereas our framework is general and can be applied to any coordinate-wise convex loss functions.

The second line of work studies online learning with vectoral payoffs via Blackwell approachability [4, 21, 1]. In their framework, the learner is given a target set $\mathcal{T} \subset \mathbb{R}^m$ and its goal is to generate decisions $\{x_t\}_{t=1}^T$ to minimize the distance between the average loss $\frac{1}{T}\sum_{t=1}^T l_t(x_t)$ and the target set $\mathcal{T}$. There are two major differences between Blackwell approachability and our proposed MO-OCO: previous works on Blackwell approachability are zero-order methods and the target set $\mathcal{T}$ is often known beforehand (also see the discussion in [5]), while in MO-OCO we intend to develop a first-order method to reach the unknown Pareto front.

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

### A.2 Multi-Objective Optimization

Multi-objective optimization aims to optimize multiple objectives concurrently. Most of the previous works on multi-objective optimization are conducted in the offline setting, including the batch optimization setting [10, 19] and the stochastic optimization setting [27, 18, 32, 8, 15]. These methods are based on gradient composition, and have shown very promising results in multi-task learning applications.

Despite the existence of previous works on multi-objective optimization, as the first work of multi-objective optimization in the OCO setting, our work is largely different from them in three aspects. First, we contribute the first formal framework of multi-objective online convex optimization. In particular, our framework is based on a novel equivalent transformation of the PSG metric, which is intrinsically different from previous offline optimization frameworks. Second, we provide a showcase in which a commonly used method in the offline setting, namely min-norm [10, 27], fail to attain sublinear regret in online setting. Our proposed min-regularized-norm is a novel design when tailoring offline methods to the online setting. Third, the regret analysis of multi-objective online learning is intrinsically different from the convergence analysis in the offline setting [32].

### A.3 Multi-Objective Multi-Armed Bandits

Another branch of related works study multi-objective optimization in the multi-armed bandits setting [5, 29, 30, 20, 9]. Among these works, the most relevant one to ours is [30], which introduces the Pareto suboptimality gap (PSG) metric to characterize the multi-objective regret in the bandits setting, and proposes a zero-order zooming algorithm to minimize the regret.

In this work, our regret definition also utilizes the PSG metric [30]. However, as the first study of multi-objective optimization in the OCO setting, our work is intrinsically different from these previous works in the following aspects. First, as PSG is a zero-order metric, we perform a novel equivalent transformation, making it amenable to the OCO setting. Second, our proposed algorithm is a first-order multiple gradient algorithm, whose design principles are completely distinct from zero-order algorithms. For example, the concept of the stability of composite weights does not even exist in the design of previous zero-order methods for multi-objective bandits [30, 20]. Third, the regret analysis of MO-OCO is intrinsically different from that in the bandits setting.

## B More Details of the Definition of PSG

Recall that in Definition 2.3, we formulate the PSG metric as a constrained optimization problem. We note that, since the PSG metric is based on the notion of "non-dominance" [30], its most direct form is actually

$$
\begin{aligned}
\Delta'(x; \mathcal{K}^*, H) = &\inf_{\epsilon \geq 0} \epsilon, \\
\text{s.t.} \quad &\forall x'' \in \mathcal{K}^*, \exists i \in \{1, \ldots, m\}, h^i(x) - \epsilon < h^i(x'') \\
\text{or} \quad &\forall i \in \{1, \ldots, m\}, h^i(x) - \epsilon = h^i(x'').
\end{aligned}
$$

At the first glance, the above definition seems to be quite different from Definition 2.3, since it has an extra condition "$\forall i \in \{1, \ldots, m\}, h^i(x) - \epsilon = h^i(x'')$". In the following, we prove that both definitions actually yield the same value due to the infimum operation on $\epsilon$.

Specifically, for any possible $(x, \mathcal{K}^*, H)$, we denote $\Delta'(x; \mathcal{K}^*, H) = \epsilon_0'$ and $\Delta(x; \mathcal{K}^*, H) = \epsilon_0$. By comparing the constraints of both definitions, it is obvious that $\epsilon_0$ must satisfy the constraint of $\Delta'(x; \mathcal{K}^*, H)$, hence the infimum operation guarantees that $\epsilon_0' \leq \epsilon_0$. It remains to prove that $\epsilon_0' \geq \epsilon_0$. To this end, we only need to show that $\epsilon_0' + \xi$ satisfies the constraint of $\Delta(x; \mathcal{K}^*, H)$ for any $\xi > 0$. Consider an arbitrary $x'' \in \mathcal{K}^*$. From the definition of $\Delta'(x; \mathcal{K}^*, H)$, we know that either $\exists i \in \{1, \ldots, m\}, h^i(x) - \epsilon_0' < h^i(x'')$ or $\forall i \in \{1, \ldots, m\}, h^i(x) - \epsilon_0' = h^i(x'')$. Whichever condition holds, we must have $\exists i \in \{1, \ldots, m\}, h^i(x) - \epsilon_0' - \xi < h^i(x'')$ for any $\xi > 0$. Since it holds for any $x'' \in \mathcal{K}^*$, $\epsilon_0' + \xi$ lies in the feasible region of $\Delta(x; \mathcal{K}^*, H)$, hence we have $\epsilon_0 \leq \epsilon_0' + \xi, \forall \xi > 0$ and thus $\epsilon_0 \leq \epsilon_0'$. In summary, we have $\Delta'(x; \mathcal{K}^*, H) = \Delta(x; \mathcal{K}^*, H)$ for any $(x, \mathcal{K}^*, H)$.

## C More Details of The Multi-Objective Static Regret

Recall that we formally formulate the multi-objective static regret $R_{\text{MOS}}(T)$ by making an extension of the equivalent form of the dynamic variant in Proposition 3.1. In the following, we show that such an alternative form also has a clear physical meaning, i.e., it measures the distance from the cumulative loss $\sum_{t=1}^{T} F_t(x_t)$ to the global Pareto front $\mathcal{P}^* = \mathcal{P}_X(\sum_{t=1}^{T} F_t)$, which is in the same spirit of PSG.

**Proposition C.1.** *Denote the Pareto optimal set of the cumulative loss function $\sum_{t=1}^{T} F_t$ as $\mathcal{X}^*$, and define a metric*

$$\Delta''(\{x_t\}_{t=1}^T; \{F_t\}_{t=1}^T) = \inf_{\epsilon \geq 0} \epsilon,$$

$$\textit{s.t.} \quad \forall x'' \in \mathcal{X}^*, \exists i \in \{1, \dots, m\}, \sum_{t=1}^{T} F^i(x_t) - \epsilon < \sum_{t=1}^{T} F^i(x'').$$

*Then the aforementioned multi-objective static regret $R_{MOS}(T)$ equals to $\Delta''(\{x_t\}_{t=1}^T; \{F_t\}_{t=1}^T)$ if $R_{MOS}(T) \geq 0$.*

The proof is very similar to our derivation of the equivalent form of the dynamic variant, which is omitted here. The above defined metric $\Delta''(\{x_t\}_{t=1}^T; \{F_t\}_{t=1}^T)$ can be viewed as PSG of the decision sequence $\{x_t\}_{t=1}^T$ w.r.t. the loss sequence $\{F_t\}_{t=1}^T$, as it measures the distance between $\sum_{t=1}^{T} F_t(x_t)$ to $\mathcal{P}^*$. From the above proposition, we know that when $R_{\text{MOS}}(T) \geq 0$, it is exactly equivalent to $\Delta''(\{x_t\}_{t=1}^T; \{F_t\}_{t=1}^T)$. Note that in the regret analysis, we are more interested in the case of $R_{\text{MOS}}(T) \geq 0$, since when $R_{\text{MOS}}(T) < 0$, it is naturally bounded by any sublinear regret bound. Hence, bounding $R_{\text{MOS}}(T)$ is essentially bounding $\Delta''(\{x_t\}_{t=1}^T; \{F_t\}_{t=1}^T)$, which naturally drives the cumulative loss $\sum_{t=1}^{T} F_t(x_t)$ to the Pareto front.

***Remark.*** At its first glance, our defined regret $R_{\text{MOS}}(T)$ can be optimized by using the linearization method with fixed weights $\lambda_0 \in \mathcal{S}_m$, or alternatively, optimizing a single objective $i \in \{1, ..., m\}$. But we remark that this is not a problem of our regret definition, but **an intrinsic requirement** of Pareto optimality. Specifically, Pareto optimality is the objective of multi-objective optimization, which characterizes the status where no objective can be improved without hurting the others. Hence merely optimizing a single objective naturally achieves Pareto optimality. Please refer to Proposition 8 in [12] for the rigorous proof. As a general performance metric, our regret should indeed incorporate this special case. Moreover, simple linearization may not solve MOO optimally, and more advanced algorithms need to be motivated by extra insights from other aspects, such as the concept of "common descent" in MGDA [10]. Indeed, in both theory (see the remark below Theorem 4.4) and experiment (see empirical results in Figure 1), DR-OMMD achieves a smaller regret than linearization.

## D More Details of The Algorithm

In this section, we provide more details to help better understand the design of DR-OMMD.

First, we remark that our proposed DR-OMMD is actually based on the agile version of online mirror descent [13], where the updated model directly moves to its projecting point onto the decision set at each round. However, we can easily make an analogy and devise a lazy version of DR-OMMD. Specifically, we only need to use a lazy projection operation in the mirror descent step with the composite gradient instead of the agile projection operation (line 8 in Algorithm 1). Note that, the analysis of the lazy version is very similar to that of the agile version [13].

In the following, we give more details of the calculation of min-regularized-norm, and discuss an adaptive version of DR-OMMD with adaptive learning rates $\eta_t$ and regularization strengths $\alpha_t$. Note that there is a typo in the closed-form solution to min-regularized-norm when $m = 2$ in our main paper (line 277): "$\eta_t = \max\{\min\{\gamma_t'', 0\}, 1\}$" should be "$\eta_t = \max\{\min\{\gamma_t'', 1\}, 0\}$" (see the strict proof of Proposition D.1).

### D.1 More Details of the Calculation of Min-Regularized-Norm

Similar to [27], we first consider the setting of two objectives, namely $m = 2$. In this case, for any $\lambda = (\gamma, 1 - \gamma), \lambda_0 = (\gamma_0, 1 - \gamma_0) \in \mathcal{S}_2$, the L1-regularization $\|\lambda - \lambda_0\|_1$ equals to $2|\gamma - \gamma_0|$. Hence min-regularized-norm at round $t$ reduces to $\lambda_t = (\gamma_t, 1-\gamma_t)$ where

$$\gamma_t \in \operatorname*{arg\,min}_{0\le\gamma\le1} \|\gamma g_1 + (1-\gamma) g_2\|^2 + 2\alpha|\gamma - \gamma_0|.$$

Interestingly, we find that the above problem has a closed-form solution.

**Proposition D.1.** *Set $\gamma_L = (g_2^\top(g_2-g_1)-\alpha)/\|g_2-g_1\|^2$, and $\gamma_R = (g_2^\top(g_2-g_1)+\alpha)/\|g_2-g_1\|^2$. Then min-regularized-norm produces weights $\lambda_t = (\gamma_t, 1-\gamma_t)$ where*

$$\gamma_t = \max\{\min\{\gamma_t'', 1\}, 0\}, \quad \textit{where } \gamma_t'' = \max\{\min\{\gamma_0, \gamma_R\}, \gamma_L\}.$$

*Proof.* We solve the following two quadratic sub-problems, i.e.,

$$\min_{0\le\gamma\le\gamma_0} h_1(\gamma) = \|\gamma g_1 + (1-\gamma) g_2\|^2 + 2\alpha(\gamma_0 - \gamma),$$

as well as

$$\min_{\gamma_0\le\gamma\le1} h_2(\gamma) = \|\gamma g_1 + (1-\gamma) g_2\|^2 + 2\alpha(\gamma - \gamma_0).$$

It can be checked that in the former sub-problem, $h_1$ monotonously decreases on $(-\infty, \gamma_R]$ and increases on $[\gamma_R, +\infty)$; in the latter sub-problem, $h_2$ monotonously decreases on $(-\infty, \gamma_L]$ and increases on $[\gamma_L, +\infty)$. Since each sub-problem has its constraint ($[0, \gamma_0]$ or $[\gamma_0, 1]$), the solution to the original optimization problem can then be derived by comparing the optimal values of the two sub-problems with their constraints. Specifically, notice that $\gamma_L \le \gamma_R$ and $0 \le \gamma_0 \le 1$, and we can consider the following three cases.

(i) When $0 \le \gamma_0 \le \gamma_L \le \gamma_R$, then $h_1$ monotonously decreases on $[0, \gamma_0]$ and its minimum on $[0, \gamma_0]$ is $h_1(\gamma_0)$. Notice that $h_1(\gamma_0) = h_2(\gamma_0)$. For the sub-problem of $h_2$, we further consider two situations:
(i-a) If $\gamma_L \le 1$, then $\gamma_L \in [\gamma_0, 1]$, hence the minimum of $h_2$ on $[\gamma_0, 1]$ is $h_2(\gamma_L)$. Since $h_2(\gamma_L) \le h_2(\gamma_0) = h_1(\gamma_0)$, the minimal point of the original problem is $\gamma_L$, and hence $\gamma_t = \gamma_L$.
(i-b) If $\gamma_L > 1$, then $h_2$ monotonously decreases on $[\gamma_0, 1]$, and we surely have $h_2(1) \le h_2(\gamma_0) = h_1(\gamma_0)$. Hence $\gamma_t = 1$ in this situation.
Combining the above two situations, we have $\gamma_t = \min\{\gamma_L, 1\}$ in this case.

(ii) When $\gamma_L \le \gamma_R \le \gamma_0 \le 1$, then $h_2$ monotonously increases on $[\gamma_0, 1]$ and its minimum on $[\gamma_0, 1]$ is $h_2(\gamma_0)$. Notice that $h_1(\gamma_0) = h_2(\gamma_0)$. For the sub-problem of $h_1$, similar to the first case, we also consider two situations:
(ii-a) If $\gamma_R \ge 0$, then $\gamma_R \in [0, \gamma_0]$, hence the minimum of $h_1$ on $[0, \gamma_0]$ is $h_1(\gamma_R)$. Since $h_1(\gamma_R) \le h_1(\gamma_0) = h_2(\gamma_0)$, the minimal point of the original problem is $\gamma_R$, and hence $\gamma_t = \gamma_R$.
(ii-b) If $\gamma_R < 0$, then $h_1$ monotonously increases on $[0, \gamma_0]$. Hence we have $h_1(0) \le h_1(\gamma_0) = h_2(\gamma_0)$. Hence the solution to the original problem $\gamma_t = 0$.
Combining the above two situations, we have $\gamma_t = \max\{\gamma_R, 0\}$ in this case.

(iii) When $\gamma_L < \gamma_0 < \gamma_R$, then $h_1$ monotonously decreases on $[0, \gamma_0]$ and $h_2$ monotonously increases on $[\gamma_0, 1]$. Hence each sub-problem attains its minimum at $\gamma_0$, and thus $\gamma_t = \gamma_0$.

Summarizing the above three cases gives

$$\gamma_t = \begin{cases} \min\{\gamma_L, 1\}, & \gamma_0 \le \gamma_L; \\ \max\{\gamma_R, 0\}, & \gamma_0 \ge \gamma_R; \\ \gamma_0, & \text{otherwise.} \end{cases}$$

We can further rewrite the above formula into a compact form as follows, which can be checked case-by-case.

$$\gamma_t = \max\{\min\{\gamma_t'', 1\}, 0\}, \quad \text{where } \gamma_t'' = \max\{\min\{\gamma_0, \gamma_R\}, \gamma_L\},$$

This gives the closed-form solution of min-regularized-norm when $m = 2$. ■

Now that we have derived the closed-form solution to the min-regularized-norm solver with any two gradients, in principle, we can apply [27]'s technique to efficiently compute the solution to the solver with more than two gradients. We provide the full procedure in Algorithm 1, which is an extension of [27]. By following the exact line search technique [14] in MGDA, we get our line search oracle as line 5 in Algorithm 1. The first term is the same as that in MGDA, and the second term is an extra

**Algorithm 1** Frank-Wolfe Solver for Line 6 in Algorithm 1

**Initialize:** $\lambda_t = (\gamma_t^1, \dots, \gamma_t^m) = (\frac{1}{m}, \dots, \frac{1}{m})$.
Compute the matrix $\mathbf{U} = \nabla F_t(x_t)^\top \nabla F_t(x_t)$, i.e., $\mathbf{U}_{ij} = \nabla f_t^i(x_t)^\top \nabla f_t^j(x_t), \forall i, j \in \{1, \dots, m\}$.
**repeat**
  Select an index $k \in \arg\max_{i \in \{1,\dots,m\}} \{\sum_{j=1}^m \gamma_t^j \mathbf{U}_{ij} + \alpha \operatorname{sgn}(\gamma_t^i - \gamma_0^i)\}$.
  Compute $\delta \in \arg\min_{0 \le \delta \le 1} \left\|\delta \nabla f_t^k(x_t) + (1-\delta)\nabla F_t(x_t)\lambda_t\right\|_2^2 + \alpha\|\delta(e_k - \lambda_t) + \lambda_t - \lambda_0\|_1$.
  Update $\lambda_t = (1-\delta)\lambda_t + \delta e_k$.
**until** $\delta \sim 0$ **or** Number of Iteration Limits
**return** $\lambda_t$.

$l_1$-regularization term related to the design in Algorithm 1. Unlike the oracle of MGDA that has a closed-form solution by a reduction to the case of two gradients, the extra $l_1$-norm term makes our oracle difficult to get a closed-form solution. The reason is that, such an extra term is the $l_1$-norm of a $m$-dimension vector, hence it can not simply reduce to the case of two gradients. To proceed, we can directly apply numerical methods to get the solution (e.g. similar to the implementation in [19]).

### D.2 An Adaptive Version of DR-OMMD

We discuss an alternative version of DR-OMMD with adaptive $\eta_t$ and $\alpha_t$, termed A-DR-OMMD. The algorithm is presented in Algorithm 2. The analysis of A-DR-OMMD is very similar to the original DR-OMMD algorithm. We provide a static regret bound of A-DR-OMMD as below.

**Theorem D.2.** *Assume the assumptions of Theorem 4.4 hold. Then for any $\lambda_0 \in \mathcal{S}_m$ and non-increasing $\{\eta_t\}_{t=1}^T$, A-DR-OMMD attains the following static regret*

$$R_{\text{MOS}}(T) \le \frac{\gamma D}{\eta_T} + \sum\nolimits_{t=1}^T \frac{\eta_t}{2}(\|\nabla F_t(x_t)\lambda_t\|_2^2 + \frac{4F}{\eta_t}\|\lambda_t - \lambda_0\|_1).$$

**Corollary D.3.** *When $\eta_t = \frac{\sqrt{2\gamma D}}{G\sqrt{t}}, \alpha_t = \frac{4F}{\eta_t}$, the above bound is in the order of $O(\sqrt{T})$.*

Then we give a dynamic regret bound of A-DR-OMMD as follows.

**Theorem D.4.** *Assume the assumptions of Theorem 4.6 hold. Then for any $\lambda_0 \in \mathcal{S}_m$ and non-increasing $\{\eta_t\}_{t=1}^T$, A-DR-OMMD attains the following dynamic regret*

$$R_{\text{MOD}}(T) \le \frac{\gamma D V_T}{G^2\eta_T^2} + \sum_{t=1}^T \frac{\eta_t}{2}(\|\nabla F_t(x_t)\lambda_t\|_*^2 + \frac{8G^2FT}{V_T}(\frac{\eta_T}{\eta_t})\sum_{t=1}^T \|\lambda_t - \lambda_0\|_1) + 2\eta_T G^2 T.$$

**Corollary D.5.** *When $\eta_t = \frac{1}{G}(\frac{\gamma D V_T}{Gt})^{1/3}, \alpha_t = \frac{8G^2FT}{V_T}(\frac{\eta_T}{\eta_t})$, the above bound is in the order of $O(T^{2/3}V_T^{1/3})$.*

## E More Details of The Theoretical Analysis

In this section, we provide more details and highlight technical lemmas in theoretical analysis, which are omitted in our main paper due to space limitation. Note that there is a typo in the main paper: in the remark below Theorem 4.4 (line 304), "$\eta_t = \frac{\sqrt{2\gamma D}}{G\sqrt{t}}, \alpha_t = \frac{4F}{\eta_t}$" should be "$\eta = \frac{\sqrt{2\gamma D}}{G\sqrt{T}}, \alpha = \frac{4F}{\eta}$" (see the strict proof of Corollary E.1), which does not affect the following remarks.

### E.1 The Tightness of DR-OMMD's Bounds

Recall that in the remarks below Theorem 4.4 and 4.6 in our main paper, we give the order of the static regret bound $O(\sqrt{T})$ and the dynamic regret bound $O(V_T^{1/3}T^{2/3})$ for DR-OMMD with proper choices of $\eta$ and $\alpha$. They are actually directly derived from the main theorems. We present these results in formal corollaries as follows.

**Algorithm 2** An Adaptive Version of DR-OMMD (A-DR-OMMD)

1: **Input:** Convex set $\mathcal{X}$, time horizon $T$, regularization parameters $\alpha_t$, learning rates $\eta_t$, regularization function $R$, user preference $\lambda_0$.
2: **Initialize:** $x_1 \in \mathcal{X}$.
3: **for** $t = 1, \ldots, T$ **do**
4: Predict $x_t$ and receive a loss function $F_t : \mathcal{X} \to \mathbb{R}^m$.
5: Compute the multiple gradients $\nabla F_t(x_t) = [\nabla f_t^1(x_t), \ldots, \nabla f_t^m(x_t)] \in \mathbb{R}^{n\times m}$.
6: Determine the weights for the gradient composition via **min-regularized-norm**

$$\lambda_t = \arg\min_{\lambda\in\mathcal{S}_m} \|\nabla F_t(x_t)\lambda\|_2^2 + \alpha_t\|\lambda - \lambda_0\|_1.$$

7: Compute the composite gradient $g_t = \nabla F_t(x_t)\lambda_t$.
8: Perform online mirror descent using $g_t$

$$x_{t+1} = \arg\min_{x\in\mathcal{X}} \eta_t\langle g_t, x\rangle + B_R(x, x_t).$$

9: **end for**

**Corollary E.1.** *With $\eta = \frac{\sqrt{2\gamma D}}{G\sqrt{T}}$ and $\alpha = \frac{4F}{\eta}$, for any $\lambda_0 \in \mathcal{S}_m$, DR-OMMD achieves the following static regret*

$$R_{\text{MOS}}(T) \le O(\sqrt{T}).$$

**Corollary E.2.** *With $\eta = \frac{2}{G}(\frac{\gamma D V_T}{GT})^{1/3}$ and $\alpha = \frac{8FG^2T}{V_T}$, for any $\lambda_0 \in \mathcal{S}_m$, DR-OMMD achieves the following dynamic regret*

$$R_{\text{MOD}}(T) \le O(V_T^{1/3}T^{2/3}).$$

Next we show that both of the derived bounds are tight w.r.t. $m$ regarding any gradient-based algorithm. Specifically, we follow the standard worst-case analysis of deriving lower bounds and construct a special case in which any gradient-based algorithm will incur a static or dynamic regret in the order of $\Omega(\sqrt{T})$ or $\Omega(V_T^{1/3}T^{2/3})$.

Consider the case in which $f_t^1 = f_t^2 = \cdots = f_t^m$ at each round $t$. In this case, the instantaneous gradients of all the objectives are identical, i.e., $g_t^i = \nabla f_t^i(x_t) \equiv \nabla f_t^1(x_t) = g_t^1, \forall i \in \{1, ..., m\}$. For any gradient-based algorithm, since it can only utilize the gradient information of the objectives, it cannot distinguish the objective to which a certain gradient belongs. Alternatively speaking, in this case, any multiple gradient algorithm will **treat all gradients in the same way** and thus behave like a single-objective algorithm using the single gradient $g_t^1$. Hence, in intuition, for any gradient-based algorithm, the worst-case bounds are **at least independent** of $m$. In particular, the worst-case bounds of gradient-based algorithms cannot decrease as $m$ increases; otherwise, the above case will be violated.

In the following, we take the static regret as an example and give a detailed proof of the tightness of the $O(\sqrt{T})$ bound. In the above case, since $f_t^1 = f_t^2 = \cdots = f_t^m, \forall t$, the cumulative losses of all the objectives are also identical, i.e., $\sum_{t=1}^T f_t^1 = \sum_{t=1}^T f_t^2 = \cdots = \sum_{t=1}^T f_t^m$. Therefore, the Pareto set $\mathcal{X}^*$ of the cumulative vector loss $\sum_{t=1}^T F_t$ coincides with the optimal decision set of the cumulative loss $\sum_{t=1}^T f_t^1$ of the first objective, i.e., $\mathcal{X}^* = \operatorname{argmin}_{x\in\mathcal{X}} \sum_{t=1}^T f_t^1(x)$. From our definition of the multi-objective static regret, since $\lambda^\top F_t(x) = f_t^1(x)$ for any $\lambda \in \mathcal{S}_m$, we have

$$R_{\text{MOS}}(T) = \sup_{x^*\in\mathcal{X}^*} (\sum_{t=1}^T f_t^1(x_t) - \sum_{t=1}^T f_t^1(x^*)) = \sum_{t=1}^T f_t^1(x_t) - \min_{x^*\in\mathcal{X}} \sum_{t=1}^T f_t^1(x^*),$$

which exactly reduces to the static regret $R_S(T)$ defined by the losses $\{f_t^1\}_{t=1}^T$ of the first objective. Hence we have $R_{\text{MOS}}(T) = R_S(T)$ in this case. Since the losses $\{f_t^1\}_{t=1}^T$ of the first objective can be chosen adversarially, we can follow Section 3.2 in [13] to construct a certain sequence $\{f_t^1\}_{t=1}^T$ that admits a lower bound of $\Omega(\sqrt{T})$. Hence in this certain case, any multiple gradient algorithm will admit $R_{\text{MOS}}(T) = \Omega(\sqrt{T})$ w.r.t. $T$ and $m$, matching our derived static regret bound for DR-OMMD in terms of both $T$ and $m$.

As for the multi-objective dynamic regret, the analysis is similar. Specifically, since the losses of all the objectives are identical, the Pareto set $\mathcal{X}_t^*$ of the instantaneous loss $F_t$ coincides with the optimal decision set of the scalar loss $f_t^1$ of the first objective, i.e., $\mathcal{X}_t^* = \operatorname{argmin}_{x \in \mathcal{X}} f_t^1(x)$. Since $\lambda^\top F_t(x) = f_t^1(x)$ for any $\lambda \in \mathcal{S}_m$, from Proposition 3.1, we have that the multi-objective dynamic regret

$$R_{\text{MOD}}(T) = \sup_{x_t^* \in \mathcal{X}_t^*, 1 \le t \le T} \sum_{t=1}^{T} (f_t^1(x_t) - f_t^1(x_t^*)) = \sum_{t=1}^{T} (f_t^1(x_t) - \min_{x_t^* \in \mathcal{X}} f_t^1(x_t^*)),$$

which exactly reduces to the single-objective dynamic regret $R_D(T)$ defined by the losses $\{f_t^1\}_{t=1}^T$ of the first objective. Hence we have $R_{\text{MOD}}(T) = R_D(T)$ in this case. Recall that in single-objective dynamic online learning [3, 33], the best attainable bound is $R_D(T) = O(V_T^{1/3} T^{2/3})$. Therefore, the best attainable bound w.r.t. $T$ and $m$ we may expect in this case is $R_{\text{MOD}}(T) = O(V_T^{1/3} T^{2/3})$, and our derived dynamic regret bound for DR-OMMD matches the order in terms of $T$ and $m$.

Some readers may suspect it unreasonable that in the multi-objective setting, the derived regret bounds do not increase as $m$ increases. Now we explicate the rationality of such independence in the following.

For the dynamic version $R_{\text{MOD}}(T)$, the independence of $m$ lies in the adoption of PSG in the formulation of the regret. Recall that, in the definition of PSG, "$\exists i \in \{1, \ldots, m\}$" means that it just needs to pick one coordinate $i$ to satisfy $f_t^i(x_t) - \epsilon < f_t^i(x'')$, which omits the dependency of $m$. We can see this point from another perspective. Recall that in Proposition 1, PSG at each round $t$ has an equivalent form, namely $\sup_{x_t^*} \inf_{\lambda_t^* \in \mathcal{S}_m} (\lambda_t^*)^\top (F_t(x_t) - F_t(x_t^*))$, or equivalently $\sup_{x_t^*} \min_{i \in \{1,\ldots,m\}} (f_t^i(x_t) - f_t^i(x_t^*))$. In particular, PSG takes a minimum operation over all objectives, and thus it does not necessarily increase as $m$ increases. For the static version $R_{\text{MOS}}(T)$, it similarly takes an infimum over $\lambda^* \in \mathcal{S}_m$, or equivalently a minimum operation over all objectives $i \in \{1, \ldots, m\}$, and thus it does not necessarily grow with $m$ either.

There is another intuitive way that can help understand the rationality of the independence of $m$. As is well recognized in existing research in multi-objective optimization [12], the proportion of the Pareto optimal solutions (or more precisely, non-dominated solutions) in the decision domain tends to increase rapidly as the number of objectives increases. As a consequence, it might not be harder to reach the Pareto optimal set when $m$ turns larger, hence intuitively, the regret bound does not necessarily increase as $m$ increases.

### E.2 More Details in the Comparison with Linearization

Recall that in the remark below Theorem 4.4, we show that our derived bound for DR-OMMD is smaller than that of linearization. We here remark that the two bounds are basically in the same order. Note that this theoretical result is also very commonly seen in the offline setting, where multiple gradient algorithms often have the same (convergence) rate as linearization [32, 19]. The benefit of multiple gradient algorithms is mainly due to the implementation of gradient composition. For example, the concept of common descent [27, 32] eliminates the gradient conflicting issue; the resulting algorithm achieves substantial performance improvements compared to linearization in their experiments. In this paper, we also consider the concept of common descent in the design of DR-OMMD, which has the intrinsic superiority over linearization. We further show that DR-OMMD attains smaller bounds than linearization in theory, and significantly outperforms linearization in experiments, which are the best we can do for now. More-refined regret bound comparisons are left as an open question.

### E.3 A More Detailed Dynamic Regret Bound for DR-OMMD

We first introduce an important lemma for our analysis of dynamic regret.

**Lemma E.3.** *Suppose the diameter of the decision set $\mathcal{X}$ is bounded by $D$. Assume $F_t$ is bounded, i.e., $|f_t^i(x)| \le F$ for any $x \in \mathcal{X}, t \in \{1, \ldots, T\}, i \in \{1, \ldots, m\}$. Then for any $\lambda_0 \in \mathcal{S}_m, \Delta \in$*

$\{1,\ldots,T\}$*, DR-OMMD attains the following dynamic regret*

$$R_{\text{MOD}}(T) \leq 2\Delta \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |f_t^i(x) - f_{t+1}^i(x)| + \frac{\gamma D}{\eta}\lceil \frac{T}{\Delta} \rceil + \frac{\eta}{2}\sum_{t=1}^{T} \|\nabla F_t(x_t)\lambda_t\|_*^2 + 4\Delta F T \sum_{t=1}^{T} \|\lambda_t - \lambda_0\|_1.$$

In Corollary E.2, we only give the order of DR-OMMD's regret bound w.r.t. the core factors $m, T$ and $V_T$. In the following theorem, we supplement a more detailed regret bound for DR-OMMD which explicates the dependency of other factors on the regret. The proof is presented in Appendix I.

**Theorem E.4.** *Set* $\eta = \frac{2}{G}(\frac{\gamma D V_T}{GT})^{1/3}$ *and* $\alpha = \frac{8FG^2T}{V_T}$*. Then for any* $\lambda_0 \in \mathcal{S}_m, \Delta \in \{1,\ldots,T\}$*, DR-OMMD attains the following dynamic regret*

$$R_{\text{MOD}}(T) \leq (\frac{\gamma D}{G^4})^{1/3}\frac{(V_T)^{1/3}}{T^{1/3}} \sum_{t=1}^{T} \inf_{\lambda\in\mathcal{S}_m} (\|\nabla F_t(x_t)\lambda\|_*^2 + \frac{8FG^2T}{V_T}\|\lambda - \lambda_0\|_1) + 2(\gamma D G^2)^{1/3}(V_T)^{1/3}T^{2/3}.$$

Note that, as we have discussed in our main paper, just like its simplified version Corollary E.2, the above theorem is derived under the assumption of "regular" temporal variability, namely $\Omega(1) \leq V_T \leq o(T)$, which is implicitly assumed in other works for dynamic online learning [3, 31, 6]. To give a more general analysis, in the following subsection, we also provide strict regret bounds that are valid for arbitrary $V_T \in [0, \infty)$.

### E.4 Dynamic Regret Bounds for DR-OMMD Under Arbitrary Temporal Variability

We here provide a detailed dynamic regret bound for DR-OMMD under arbitrary temporal variability $V_T \in [0, \infty)$. Specifically, we first give a detailed regret bound in analogy to Theorem E.4. The proof of this theorem is given in Appendix I.

**Theorem E.5.** *In DR-OMMD, set* $\eta = \min\{\max\{\frac{2}{G}(\frac{\gamma D V_T}{GT})^{1/3}, \frac{4V_T}{G^2T}\}, \frac{4V_T}{G^2}\}$ *and* $\alpha = \frac{8FG^2T}{V_T}$*. Then it attains the following dynamic regret*

$$R_{\text{MOD}}(T) \leq \max\{(\frac{\gamma D}{G^4})^{1/3}\frac{(V_T)^{1/3}}{T^{1/3}} \sum_{t=1}^{T} \inf_{\lambda\in\mathcal{S}_m} \left(\frac{8FG^2T}{V_T}\|\lambda - \lambda_0\|_1 + \|\nabla F_t(x_t)\lambda\|_*^2\right) + 2(\gamma D G^2)^{1/3}(V_T)^{1/3}T^{2/3}, \quad 6V_T, \quad \frac{3G}{2}(2\gamma D)^{1/2}T^{1/2}\}.$$

The above bound can be rewritten into a simpler form, if we are only interested in its order w.r.t. $m, T$ and $V_T$, just as in Corollary E.2.

**Corollary E.6.** *In DR-OMMD, set* $\eta = \min\{\max\{\frac{2}{G}(\frac{\gamma D V_T}{GT})^{1/3}, \frac{4V_T}{G^2T}\}, \frac{4V_T}{G^2}\}$ *and* $\alpha = \frac{8FG^2T}{V_T}$*. Then it attains the following dynamic regret*

$$R_{\text{MOD}}(T) \leq \max\{O({V_T}^{1/3}T^{2/3}), O(V_T), O(T^{1/2})\}.$$

## F More Experimental Results

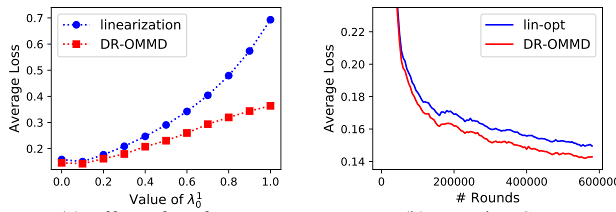
### F.1 More Details of the Experimental Setup

The *protein* and *covtype* datasets used in our experiments are publicly available in [11]. The *MultiMNIST* dataset is acquired by the code provided by [27].

All runs are deployed on Xeon(R) E5-2699 @ 2.2GHz.

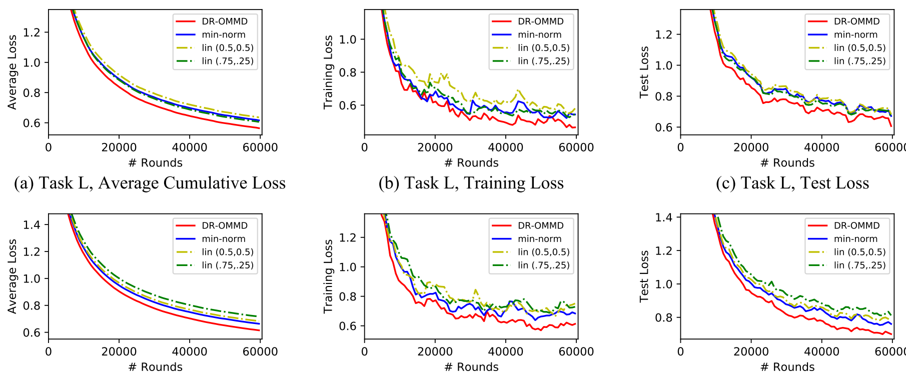
### F.2 More Results for Adaptive Regularization

We supplement the empirical results on *covtype* in Figure 1, which have been omitted from our main paper due to the lack of space. These results are consistent with the results on *protein* as presented in our main paper.




(a) Effect of Preference (b) Learning Curve

Figure 1: Results to verify the effectiveness of adaptive regularization on *covtype*. (a) Performance of DR-OMMD and linearization under varying $\lambda_0 = (\lambda_0^1, 1 - \lambda_0^1)$. (b) Performance using the optimal weights $\lambda_0 = (0.1, 0.9)$.




(a) Task L, Average Cumulative Loss (b) Task L, Training Loss (c) Task L, Test Loss

(d) Task R, Average Cumulative Loss (e) Task R, Training Loss (f) Task R, Test Loss

Figure 2: Results to verify the effectiveness of DR-OMMD for online deep learning. The plots show the average cumulative loss, training loss and test loss for both tasks (*task L* and *task R*) on MultiMNIST.

### F.3 Non-Convex Experiments: Deep Multi-Task Learning via Multi-Objective Optimization

We evaluate DR-OMMD in the online non-convex setting. We adopt MultiMNIST [26], which is a multi-task version of the MNIST dataset for image classification and commonly used in deep multi-task learning [27, 18]. In MultiMNIST, each sample is constructed by putting a random image of some digit at the top-left and another image at the bottom-right. The goal to classify the digit at the top-left (*task L*) and that at the bottom-right (*task R*) at the same time.

We follow [27]'s setup and adopt the LeNet architecture. For linearization, we examine two choices of weights $(0.5, 0.5)$ and $(0.75, 0.25)$ [18]. For all the examined algorithms, learning rates $\eta$ are selected via a grid search over $\{0.0001, 0.001, 0.01, 0.1\}$. For DR-OMMD, the parameter $\alpha$ is set according to Theorem 4.4, and the initial weights are simply set as $\lambda_0 = (0.5, 0.5)$. Note that in online experiments, samples arrive at the learner one after another in a sequential manner, which is largely different from offline experiments where batches of samples are randomly sampled from the training set.

Figure 2 compares the average cumulative loss, training loss and test loss of all the examined algorithms. Note that the first metric is often used in online experiments and the last two are commonly used in offline experiments [25]. The results show that DR-OMMD outperform counterpart algorithms using min-norm or linearization in all metrics on both tasks, which verifies the effectiveness of DR-OMMD in the online non-convex setting.

### F.4 More Results for Simulation Experiment: Handling Unknown $V_T$

Recall that our analysis of the multi-objective dynamic regret $R_{\mathrm{MOD}}(T)$ follows the dynamic setting of [3], where the temporal variability $V_T$ is assumed to be known in advance. However, in many real-world application scenarios, the value of $V_T$ is often unknown to the learner. Since the step size

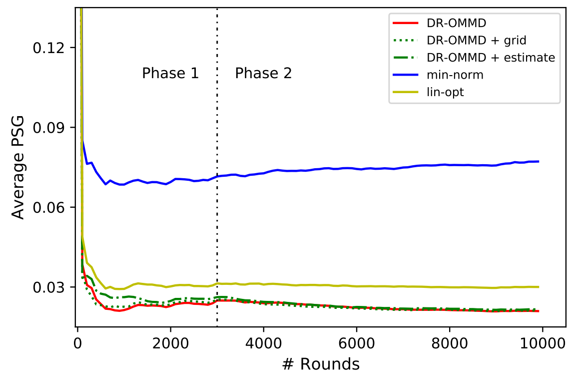



Figure 3: The plot compares the performance of DR-OMMD using grid search, DR-OMMD using estimated $V_T$, DR-OMMD using true $V_T$, and the two baselines.

$\eta$ and the regularization strength $\alpha$ in DR-OMMD depend on $V_T$, we need to handle unknown $V_T$ when applying DR-OMMD to these scenarios.

In fact, in online learning, it is ubiquitous that the choice of hyperparameters depends on the value of some parameters that are hard to know in advance [24, 3]. Typically, there are many approaches to handle unknown parameters. Here we discuss three of them in the existing literature.

The first way is to conduct a grid search on the unknown parameter and choose the best setting from the grid [22]. Since the only unknown parameter in our dynamic setting is $V_T$, similar to the grid search conducted in [22], we can directly perform a grid search on $V_T$ and then derive $\eta$ and $\alpha$ from $V_T$. Specifically, we set the grid of $V_T$ as $\{2^i \mid 1 \le i \le N_1\}$. Then the hyperparameters can be set as $\eta = \frac{2}{G}(\frac{\gamma D v}{GT})^{1/3}$ and $\alpha = \frac{8FG^2T}{v}$ for $v \in \{2^i \mid 1 \le i \le N_1\}$. In the simulation experiment, since the time horizon $T = 10000$ and the maxima of the loss function $F \le 9$, we set $N_1 = 17$ so that any possible $V_T$ lies in the range of our grid $[2, 2^{N_1}]$. We conduct DR-OMMD using the above grid, and find that the best value of $v$ is 13.

The second way is to fit the unknown parameter into some growth pattern w.r.t. $T$. For example, [3] assume $V_T = T^{\beta}$ and estimate the value of $\beta$. Following [3], we can estimate $\beta$ according to the first $T'$ rounds. In the simulation experiment, we set $T' = 1000$, which is considerably small compared to the total time horizon $T = 10000$. We first simulate for $T'$ rounds and derive the estimate $\hat{\beta} = 0.95$, then use $V_T = T^{0.95}$ to decide the hyperparameters $\eta$ and $\alpha$.

The third way is to combine the above grid search with a meta-algorithm [34]. For example, [34] runs multiple dynamic algorithms with varying hyperparameters and uses a meta-algorithm called VariationHedge to integrate the decision of each algorithm. Following [34], in practice, we can run multiple DR-OMMDs where $\eta$ and $\alpha$ are decided by the above grid $V_T \in \{2^i \mid 1 \le i \le N_1\}$.

We realize the first two approaches in the simulation experiment, and compare their performance with DR-OMMD using the true $V_T$ and *min-norm* and *lin-opt* in Figure 3. The results show that DR-OMMD equipped with the above approaches perform comparably to DR-OMMD using the true $V_T$ and substantially outperforms the two examined baselines. Specifically, the average regret of DR-OMMD using grid search is 0.0211, and the average regret of DR-OMMD using estimated $V_T$ is 0.0216. They are slightly worse than DR-OMMD using the true $V_T$ (0.0209) but still substantially better than the *min-norm* (0.0772) and *lin-opt* (0.0302) baselines.

# G Omitted Proofs of Proposition 3.1 (The Equivalent Form of Multi-Objective Dynamic Regret)

*Proof.* We first analyze the PSG measurement $\Delta_t(x_t)$ at each round $t \in \{1, \dots, T\}$. Specifically, for any comparator $x \in \mathcal{X}$, we first define the **pair-wise suboptimality gap** between decisions $x_t$ and $x$, i.e.,

$$\delta_t(x_t; x) = \inf_{\epsilon \ge 0} \{\epsilon \mid F_t(x_t) - \epsilon \mathbf{1} \nprec F_t(x)\}.$$

Then PSG can be expressed via $\Delta_t(x_t) = \sup_{x \in \mathcal{P}_X(F_t)} \delta_t(x_t, x)$.

We then focus on the pair-wise gap $\delta_t(x_t; x)$ w.r.t. any Pareto optimal decision $x \in \mathcal{X}_t^* \equiv \mathcal{P}_X(F_t)$. Since $x$ is a Pareto optimal decision of $F_t$, from the definition of Pareto optimality, there must exist some $i \in \{1, \ldots, m\}$ such that $f_t^i(x_t) - f_t^i(x) \geq 0$.

The pair-wise suboptimality gap has an equivalent expression, namely,

$$\delta_t(x_t; x) = \min_{k \in \{1,\ldots,m\}} \{(f_t^k(x_t) - f_t^k(x))_+\},$$

where $(l)_+ = \max\{l, 0\}, l \in \mathbb{R}$ is the truncation operator. Denote $\mathcal{U}_m = \{e_k \mid 1 \leq k \leq m\}$ as the set of all unit vector in $\mathbb{R}^m$, then we equivalently have

$$\delta_t(x_t; x) = \min_{\lambda \in \mathcal{U}_m} \lambda^\top (F_t(x_t) - F_t(x))_+.$$

Note that, we here slightly abuse the truncation operator $(l)_+$ to allow $l$ to be a vector in $\mathbb{R}^m$, which represents the vector whose $i$-th coordinate equals to $\max\{l^i, 0\}$ for any $i \in \{1, \ldots, m\}$.

Now the calculation of $\delta_t(x_t; x)$ becomes a minimization problem over $\lambda \in \mathcal{U}_m$. Since $\mathcal{U}_m$ is a discrete set, we can apply a linear relaxation trick. Specifically, we now turn to minimize the quantity $p(\lambda) = \lambda^\top (F_t(x_t) - F_t(x))_+$ over the convex curvature of $\mathcal{U}_m$, which is exactly the probability simplex $\mathcal{S}_m = \{\lambda \in \mathbb{R}^m \mid \lambda \succ \mathbf{0}, \|\lambda\|_1 = 1\}$. Note that, $\mathcal{U}_m$ contains all vertexes of $\mathcal{S}_m$. Since $\inf_{\lambda \in \mathcal{S}_m} p(\lambda)$ is a linear optimization problem, the minimal point ${\lambda_t^*}^\top$ must be a vertex of the simplex, i.e., ${\lambda_t^*}^\top \in \mathcal{U}_m$. Thus the relaxed problem is equivalent to the original problem, namely,

$$\min_{\lambda \in \mathcal{U}_m} \lambda^\top (F_t(x_t) - F_t(x))_+ = \inf_{\lambda \in \mathcal{S}_m} \lambda^\top (F_t(x_t) - F_t(x))_+.$$

For now, we have transformed the calculation of pair-wise suboptimality gap $\delta_t(x_t; x)$ into a optimization problem of finding the minimal linear scalarization of $(F_t(x_t) - F_t(x))_+$. Hence, the PSG at each round $t$ can be expressed as

$$\Delta_t(x_t) = \sup_{x \in \mathcal{X}_t^*} \inf_{\lambda \in \mathcal{S}_m} \lambda^\top (F_t(x_t) - F_t(x))_+.$$

In the above expression, the existence of truncation operator $(\cdot)_+$ will incur irregularity (i.e., nonlinearity) when we try to optimize the loss $F_t$. Suprisingly, we find that such operator **can be dropped** when we compute $\Delta_t(x_t)$ w.r.t. the Pareto set $\mathcal{X}_t^*$, as shown in the following.

**Lemma G.1.** *Let $\mathcal{X}_t^*$ be the Pareto set of $F_t : \mathcal{X} \to \mathbb{R}^m$. Then for any $x_t \in \mathcal{X}$, it holds that*

$$\sup_{x \in \mathcal{X}^*} \inf_{\lambda \in \mathcal{S}_m} \lambda^\top (F_t(x_t) - F_t(x))_+ = \sup_{x \in \mathcal{X}_t^*} \inf_{\lambda \in \mathcal{S}_m} \lambda^\top (F_t(x_t) - F_t(x)).$$

*Proof.* Define the alternative "gap" metric without the truncation operator as $\delta_t'(x_t; x) = \inf_{\lambda \in \mathcal{S}_m} \lambda^\top (F_t(x_t) - F_t(x))$. Moreover, define the supremum of $\delta_t'(x_t; x)$ over $x \in \mathcal{X}_t^*$ as $\Delta_t'(x_t) = \sup_{x \in \mathcal{X}_t^*} \delta_t'(x_t; x)$. Then from the definition of truncation operator $(\cdot)_+$, we have $\delta(x_t; x) \geq \delta'(x_t; x)$ and $\Delta_t(x_t) \geq \Delta_t'(x_t)$.

It then suffices to prove that, for any given $x_t \in \mathcal{X}$, there exists some certain $x_t^* \in \mathcal{X}_t^*$ such that the value of $\delta'(x_t; x_t^*)$ can be as large as $\Delta_t(x_t)$. Indeed, if this is the case, then $\Delta_t(x_t) \leq \Delta_t'(x_t)$, and hence the two quantities $\Delta_t(x_t)$ and $\Delta_t'(x_t)$ are equal. We consider the following two cases:

(i) $x_t$ is already a Pareto optimal point of $F_t$, i.e., $x_t \in \mathcal{X}_t^*$. Then from the definition of PSG, we directly have $\Delta_t(x_t) = 0$. Notice that $\delta'(x_t; x_t) = 0$, and hence $\Delta_t(x_t) = \sup_{x \in \mathcal{X}_t^*} \delta'(x_t; x) \geq \delta'(x_t; x_t) = 0$. Consequently, the relation $\Delta_t(x_t) \leq \Delta_t'(x_t)$ holds in this case.

(ii) $x_t$ is not a Pareto optimal point of $F_t$, i.e., $x_t \notin \mathcal{X}_t^*$. Then we have $\Delta_t(x_t) > 0$. Set $\epsilon = \Delta_t(x_t)$, and denote $x_t^* \in \arg\max_{x \in \mathcal{X}_t^*} \delta_t(x_t; x)$, then $\delta_t(x_t; x_t^*) = \epsilon > 0$. Therefore, from the definition of $\delta_t(x_t; x_t^*)$ we know that, for any $i \in \{1, \ldots, m\}$, we have $f_t^i(x_t) - f_t^i(x_t^*) \geq \epsilon$. Thus all entries of $F_t(x_t) - F_t(x^*)$ are positive, and we have $(F_t(x_t) - F_t(x^*))_+ = F_t(x_t) - F_t(x^*)$. Consequently, we have $\Delta_t'(x_t) = \sup_{x \in \mathcal{X}_t^*} \delta_t'(x_t; x) \geq \delta'(x_t; x_t^*) = \delta(x_t; x_t^*) = \Delta_t(x_t)$. ■

From the above lemma, the PSG measurement at round $t$ has an equivalent form as

$$\Delta_t(x_t) = \sup_{x \in \mathcal{X}_t^*} \inf_{\lambda \in \mathcal{S}_m} \lambda^\top (F_t(x_t) - F_t(x)),$$

and correspondingly, the multi-objective dynamic regret becomes

$$R_{\mathrm{MOD}}(T) = \sum_{t=1}^{T} \Delta_t(x_t) = \sum_{t=1}^{T} \sup_{x \in \mathcal{X}_t^*} \inf_{\lambda \in \mathcal{S}_m} \lambda^\top (F_t(x_t) - F_t(x)).$$

Since the time horizon $T$ is finite, we can first swap the summation over $t$ and the supremum over $x$, then swap the summation over $t$ and the infimum over $\lambda$. Then the multi-objective dynamic regret further equals to

$$R_{\mathrm{MOD}}(T) = \sup_{x_t^* \in \mathcal{X}_t^*, 1 \le t \le T} \inf_{\lambda_1^* \ldots, \lambda_t^* \in \mathcal{S}_m} \sum_{t=1}^{T} ({\lambda_t^*}^\top F_t(x_t) - {\lambda_t^*}^\top F_t(x_t^*)),$$

which proves the proposition. ■

# H Omitted Proofs of Proposition 4.1 (Min-Norm May Incur Linear Regrets)

*Proof.* As we have described in our main paper, we consider the following two-objective optimization problem. Decision domain is set as $\mathcal{X} = \{(u, v) \mid u + v \le \frac{1}{2}, v - u \le \frac{1}{2}, v \ge 0\}$. At each round $t$, the loss function $F_t : \mathcal{X} \to \mathbb{R}^2$ takes

$$F_t(x) = \begin{cases} (\|x-a\|^2, \|x-b\|^2), & t = 2k-1, k = 1, 2, ...; \\ (\|x-b\|^2, \|x-c\|^2), & t = 2k, k = 1, 2, ..., \end{cases}$$

where $a = (-2, -1), b = (0, 1), c = (2, -1)$. For simplicity of analysis, we first consider the case when the total time horizon $T$ is an even number. Then it can be checked that the cumulative loss function takes

$$\begin{aligned} \sum_{t=1}^{T} F_t(x) &= \frac{T}{2} \cdot (\|x-a\|^2 + \|x-b\|^2, \|x-b\|^2 + \|x-c\|^2) \\ &= T \cdot ((u+1)^2 + v^2 + 2, (u-1)^2 + v^2 + 2), \end{aligned}$$

for any $x = (u, v) \in \mathcal{X}$. Obviously the Pareto optimal set $\mathcal{X}^*$ of the cumulative loss coincides with the line segment between $(-1, 0)$ and $(1, 0)$, i.e., $\mathcal{X}^* = \{(u, v) \mid -\frac{1}{2} \le u \le \frac{1}{2}, v = 0\}$ (note that $\mathcal{X}^*$ is the intersection of the line segment and $\mathcal{X}$).

Now consider equipping OMD with vanilla min-norm, where the composite gradients are produced by the min-norm method. Suppose the learning process starts at any $x_1 = (u_1, v_1) \in \mathcal{X}$ such that $v_1 > 0$. Note that this is true if and only if $x_1 \notin \mathcal{X}^*$. Then for the iterate $x_t = (u_t, v_t)$ at each round $t$, we can directly calculate the gradients as

$$g_t^1 = \begin{cases} 2(x_t - a) = (2u_t + 4, & 2v_t + 2), & t = 2k-1; \\ 2(x_t - b) = (2u_t, & 2v_t - 2), & t = 2k. \end{cases}$$

$$g_t^2 = \begin{cases} 2(x_t - b) = (2u_t, & 2v_t - 2), & t = 2k-1; \\ 2(x_t - c) = (2u_t - 4, & 2v_t + 2), & t = 2k. \end{cases}$$

The min-norm weights can be computed as $\lambda_t = (\gamma_t, 1 - \gamma_t)$ where

$$\gamma_t = \begin{cases} \dfrac{(x_t - b)^\top (a - b)}{\|a - b\|^2} = \dfrac{1 - u_t - v_t}{4}, & t = 2k-1; \\ \dfrac{(x_t - c)^\top (b - c)}{\|b - c\|^2} = \dfrac{3 - u_t + v_t}{4}, & t = 2k. \end{cases}$$

The composite gradient

$$g_t^{comp} = \begin{cases} \gamma_t \cdot 2(x - a) + (1 - \gamma_t) \cdot 2(x - b) = (u_t - v_t + 1, & -u_t + v_t - 1), & t = 2k-1; \\ \gamma_t \cdot 2(x - b) + (1 - \gamma_t) \cdot 2(x - c) = (-u_t - v_t - 1, & -u_t - v_t - 1), & t = 2k. \end{cases}$$

Recall that the update form of OMD takes

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \eta_t g_t^{comp}),$$

where $\eta_t > 0$ is the learning rate and $\Pi_{\mathcal{X}}$ is the projection operation onto $\mathcal{X}$. Denote the iterate $x_t = (u_t, v_t)$ at each round. Now we can investigate the relation between $x_t$ and $x_{t+1}$ by considering the following two cases:

(i) If $x_t - \eta_t g_t^{comp} \in \mathcal{X}$, then we do not need projection, and directly have $x_{t+1} = x_t - \eta_t g_t^{comp}$.

(ii) If $x_t - \eta_t g_t^{comp} \notin \mathcal{X}$, then we need to project $x_t - \eta_t g_t^{comp}$ back to $\mathcal{X}$. Denote $x'_{t+1} = x_t - \eta_t g_t^{comp}$. For simplicity we consider the projection based on the Euclidean distance, namely $\Pi_{\mathcal{X}}(x) = \arg\min_{x' \in \mathcal{X}} \|x - x'\|_2^2$. Since the composite gradient is orthogonal to the boundary on which the iterate after projection $x_{t+1} = \Pi_{\mathcal{X}}(x'_{t+1})$ is located, it can be checked that $x_{t+1}$ lies on the line segment linking $x_t$ and $x'_{t+1}$. Alternatively speaking, $x_{t+1}$ can be expressed as $x_t - \eta'_t g_t^{comp}$ for some $0 \le \eta'_t < \eta_t$.

Combining the above two cases, we know that at each round $t$, there exists some $\eta'_t \in [0, \eta_t]$ such that $x_{t+1} = x_t - \eta'_t g_t^{comp}$. Now we can analyze the relation between each entry of $x_t$ and $x_{t+1}$. Specifically, since the second entry of the composite gradient is always non-positive, namely $-u_t + v_t - 1 \le 0$ and $-u_t - v_t - 1 \le 0$, we have $v_{t+1} \ge v_t$ for any $t$. Moreover, since the first entry of $g_t^{comp}$ is non-negative when $t = 2k - 1$, namely $u_{2k-1} - v_{2k-1} + 1 \ge 0$, we have $u_{2k} \le u_{2k-1}$ for any $k$; since the first entry of $g_t^{comp}$ is non-positive when $t = 2k$, namely $-u_{2k} - v_{2k} - 1 \le 0$, we have $u_{2k+1} \ge u_{2k}$ for any $k$.

Now we can go back to analyze the gap between the composite weights at any two consecutive rounds. It is easy to verify that $\gamma_{2k-1} < \gamma_{2k}$ and $\gamma_{2k} > \gamma_{2k+1}$, hence we have

$$\|\lambda_{2k} - \lambda_{2k-1}\|_1 = 2(\gamma_{2k} - \gamma_{2k-1}) = \frac{2 - (u_{2k} - u_{2k-1}) + (v_{2k} + v_{2k-1})}{2} \ge 1 + v_1,$$
$$\|\lambda_{2k+1} - \lambda_{2k}\|_1 = 2(\gamma_{2k} - \gamma_{2k+1}) = \frac{2 - (u_{2k} - u_{2k+1}) + (v_{2k} + v_{2k+1})}{2} \ge 1 + v_1.$$

Therefore, the composite weights $\lambda_t$ indeed change radically at any two consecutive rounds.

The above analysis on $v_t$ also implies the failure of min-norm in this problem. Recall that any Pareto optimal solution $x^* = (u^*, v^*) \in \mathcal{X}^*$ must satisfy $v^* = 0$. Suppose the initial iterate $x_1 = (u_1, v_1)$ does not lie in $\mathcal{X}^*$, i.e., $v_1 > 0$, which is almost sure for random initialization $x_1 \in \mathcal{X}$. Then we iteratively have $0 < v_1 \le v_2 \le ... \le v_T$, which means that $x_t$ moves away from the Pareto set $\mathcal{X}^*$.

In the following, we strictly prove that min-norm indeed incurs a **linear** multi-objective static regret. To calculate $R_{\text{MOS}}(T)$, we first investigate the quantity $R(x^*, \lambda) = \lambda^\top \sum_{t=1}^T (F_t(x_t) - F_t(x^*))$ for any fixed weights $\lambda = (\gamma, 1 - \gamma) \in \mathcal{S}_2$ and best fixed decision $x^* = (u^*, 0) \in \mathcal{X}^*$. Specifically, recall the form of $\sum_{t=1}^T F_t$ derived above, then we have

$$\lambda^\top \sum_{t=1}^T F_t(x^*) = (\gamma(u^* + 1)^2 + (1 - \gamma)(u^* - 1)^2 + 2)T.$$

Denote the cumulative loss $\sum_{t=1}^T F_t(x_t) = (L_1, L_2)$, we now consider the loss of each objective $L_1$ and $L_2$ separately. Specifically, for the first objective, we have

$$L_1 = \sum_{k=1}^{T/2} ((u_{2k-1} + 2)^2 + u_{2k}^2 + (v_{2k-1} + 1)^2 + (v_{2k} - 1)^2).$$

Since $0 < v_1 \le v_2 \le ... \le v_T \le 1$, for the term regarding $v_t$ we have

$$\begin{aligned}
\sum_{k=1}^{T/2} ((v_{2k-1} + 1)^2 + (v_{2k} - 1)^2) &= (v_1 + 1)^2 + (v_T - 1)^2 + \sum_{k=1}^{T/2-1} ((v_{2k} - 1)^2 + (v_{2k+1} + 1)^2) \\
&\ge \sum_{k=1}^{T/2-1} ((v_{2k} - 1)^2 + (v_{2k} + 1)^2) = \sum_{k=1}^{T/2-1} (2v_{2k}^2 + 2) \\
&\ge \sum_{k=1}^{T/2-1} (2v_1^2 + 2) = (2v_1^2 + 2)(\frac{T}{2} - 1) \ge v_1^2 T + T - 2.
\end{aligned}$$

For the $k$-th term regarding $u_t$, we have

$$(u_{2k-1}+2)^2+u_{2k}^2=(u_{2k-1}+1)^2+(u_{2k}+1)^2+2(u_{2k-1}-u_{2k})+2.$$

Recall that we have derived $u_{2k}\le u_{2k-1}$, thus we have

$$\sum_{k=1}^{T/2}(u_{2k-1}+2)^2+u_{2k}^2\ge\sum_{k=1}^{T/2}((u_{2k-1}+1)^2+(u_{2k}+1)^2+2)\ge\sum_{t=1}^{T}(u_t+1)^2+T\ge(\bar{u}+1)^2T+T,$$

where $\bar{u}=\frac{1}{T}\sum_{t=1}^{T}u_t$ and the last inequality is derived from Jensen's inequality. In summary, for the cumulative loss $L_1$ of the first objective, we have

$$L_1\ge(\bar{u}+1)^2T+v_1^2T+2T-2.$$

Similarly, we can analyze the cumulative loss $L_2$ of the second objective

$$L_2=\sum_{k=1}^{T/2}(u_{2k-1}^2+(u_{2k}-2)^2+(v_{2k-1}-1)^2+(v_{2k}+1)^2).$$

Since $0<v_1\le v_2\le\ldots\le v_T\le 1$, for the term regarding $v_t$ we have

$$\sum_{k=1}^{T/2}((v_{2k-1}-1)^2+(v_{2k}+1)^2)\ge\sum_{k=1}^{T/2}((v_{2k-1}-1)^2+(v_{2k-1}+1)^2)\ge v_1^2T+T.$$

For the term regarding $u_t$, we also have

$$\begin{aligned}\sum_{k=1}^{T/2}(u_{2k-1}^2+(u_{2k}-2)^2)&=\sum_{k=1}^{T/2}((u_{2k-1}-1)^2+(u_{2k}-1)^2+2(u_{2k-1}-u_{2k})+2)\\&\ge\sum_{t=1}^{T}(u_t-1)^2+T\ge(\bar{u}-1)^2T+T,\end{aligned}$$

where the last inequality is derived from Jensen's inequality. Therefore, we have

$$L_2\ge(\bar{u}-1)^2T+v_1^2T+2T.$$

Combining the above inequalities, we have

$$\begin{aligned}R(x^*,\lambda)&=\gamma L_1+(1-\gamma)L_2-\lambda^\top\sum_{t=1}^{T}F_t(x^*)\\&\ge\gamma((\bar{u}+1)^2-(u^*+1)^2)+(1-\gamma)((\bar{u}-1)^2-(u^*-1)^2)+v_1^2T-2\gamma.\end{aligned}$$

For any $\lambda\in\mathcal{S}_2$ (i.e., $\gamma\in[0,1]$), set $x'=(\bar{u},0)\in\mathcal{X}^*$, then it holds that

$$R(x',\lambda)\ge v_1^2T-2.$$

Equivalently, the multi-objective static regret satisfies

$$R_{\text{MOS}}(T)=\sup_{x^*\in\mathcal{X}^*}\inf_{\lambda\in\mathcal{S}_2}R(x^*,\lambda)\ge\inf_{\lambda\in\mathcal{S}_2}R(x',\lambda)\ge v_1^2T-2,$$

which is linear w.r.t. $T$ for any $x_1=(u_1,v_1)\in\mathcal{X}$ such that $v_1>0$.

We now investigate the case when $T$ is an odd number. Since the calculation of the composite weights $\lambda_t$ and the composite gradient $g_t^{comp}$ at each round is independent of the total time horizon $T$, we still have $\|\lambda_{t+1}-\lambda_t\|_1\ge v_1+1$ for any $t$. Hence the first desired property also holds for any odd $T$.

It remains to prove that OMD with min-norm still incurs a linear regret when $T$ is odd. In this case, the Pareto optimal set $\mathcal{X}^*$ does not lie in the $x$-axis anymore, hence it is difficult to directly compute $R_{\text{MOS}}(T)$. However, we can still use our derived $R(x^*,\lambda)$ for any even $T$ to estimate the regret. Specifically, set $x'=(\frac{1}{T-1}\sum_{t=1}^{T-1}u_t,0)$; from the above derivation with even $T$, for any $\lambda\in\mathcal{S}_2$, we still have (note that now $T-1$ is an even number)

$$\lambda^\top\sum_{t=1}^{T-1}F_t(x_t)-\lambda^\top\sum_{t=1}^{T-1}F_t(x')\ge v_1^2T-2.$$

Since for any $x \in \mathcal{X}$, we have $0 \le \|x-a\|^2, \|x-b\|^2, \|x-c\|^2 \le 10$, we have

$$R(x', \lambda) = \lambda^\top \sum_{t=1}^T F_t(x_t) - \lambda^\top \sum_{t=1}^T F_t(x') \ge v_1^2 T - 12.$$

Furthermore, from the definition of Pareto optimality, there exists some $x'' \in \mathcal{X}^*$ that Pareto dominates $x'$ regarding the cumulative loss $\sum_{t=1}^T F_t$, namely $\sum_{t=1}^T F_t(x'') \preceq \sum_{t=1}^T F_t(x')$. Hence

$$R(x'', \lambda) = \lambda^\top \sum_{t=1}^T F_t(x_t) - \lambda^\top \sum_{t=1}^T F_t(x'') \ge R(x', \lambda),$$

for any $\lambda \in \mathcal{S}_2$. Therefore, the multi-objective static regret

$$R_{\text{MOS}}(T) = \sup_{x^* \in \mathcal{X}^*} \inf_{\lambda \in \mathcal{S}_2} R(x^*, \lambda) \ge \inf_{\lambda \in \mathcal{S}_2} R(x'', \lambda) \ge \inf_{\lambda \in \mathcal{S}_2} R(x', \lambda) \ge v_1^2 T - 12,$$

which is also linear w.r.t. $T$ for any $x_1 = (u_1, v_1) \in \mathcal{X}$ such that $v_1 > 0$. ■

# I Omitted Proofs of the Regret Bounds for DR-OMMD

## I.1 Proof of Theorem 4.4

*Proof.* We start from the definition of $R_{\text{MOS}}(T)$. Specifically, for any $\lambda \in \mathcal{S}_m$ and $\lambda_1 \dots, \lambda_T \in \mathcal{S}_m$, it holds that

$$\begin{aligned}
R_{\text{MOS}}(T) &= \sup_{x^* \in \mathcal{X}^*} \inf_{\lambda^* \in \mathcal{S}_m} \sum_{t=1}^T {\lambda^*}^\top (F_t(x_t) - F_t(x^*)) \le \sup_{x^* \in \mathcal{X}^*} \sum_{t=1}^T \lambda^\top (F_t(x_t) - F_t(x^*)) \\
&= \sum_{t=1}^T \Big( (\lambda^\top F_t(x_t) - {\lambda_t}^\top F_t(x_t)) + {\lambda_t}^\top (F_t(x_t) - F_t(x^*)) + ({\lambda_t}^\top F_t(x^*) - \lambda^\top F_t(x^*)) \Big) \\
&\le \sum_{t=1}^T F \|\lambda - \lambda_t\|_1 + \sum_{t=1}^T {\lambda_t}^\top (F_t(x_t) - F_t(x^*)) + \sum_{t=1}^T F \|\lambda - \lambda_t\|_1 \\
&= 2F \sum_{t=1}^T \|\lambda - \lambda_t\|_1 + \sum_{t=1}^T {\lambda_t}^\top (F_t(x_t) - F_t(x^*)).
\end{aligned}$$

To tackle the second term in the above inequality, we set $u_1 = u_2 = \ldots = u_T = x^*$ in Lemma I.2, which results in

$$\sum_{t=1}^T {\lambda_t}^\top (F_t(x_t) - F_t(x^*)) \le \frac{1}{\eta} B_R(x^*, x_1) + \frac{\eta}{2} \sum_{t=1}^T \|\nabla F_t(x_t) \lambda_t\|_*^2.$$

Consequently, the static regret can be bounded as

$$\begin{aligned}
R_{\text{MOS}}(T) &\le 2F \sum_{t=1}^T \|\lambda_t - \lambda_0\|_1 + \frac{1}{\eta} B_R(x^*, x_1) + \frac{\eta}{2} \sum_{t=1}^T \|\nabla F_t(x_t) \lambda_t\|_*^2 \\
&= 2F \sum_{t=1}^T \|\lambda_t - \lambda_0\|_1 + \frac{1}{\eta} B_R(x^*, x_1) + \frac{\eta}{2} \sum_{t=1}^T \|\nabla F_t(x_t) \lambda_t\|_*^2 \\
&= \frac{1}{\eta} B_R(x^*, x_1) + \frac{\eta}{2} \sum_{t=1}^T (\|\nabla F_t(x_t) \lambda_t\|_*^2 + \frac{4F}{\eta} \|\lambda_t - \lambda_0\|_1),
\end{aligned}$$

which proves the theorem. ■

### I.2 Proof of Corollary E.1

*Proof.* We now set $\alpha = \frac{4F}{\eta}$ in DR-OMMD, and specify $\lambda_t$ to be the composition weights generated by the algorithm at round $t$. The min-regularized-norm solver guarantees that $\lambda_t \in \arg\min_{\lambda \in \Delta_t} \|\nabla F_t(x_t)\lambda\|_*^2 + \frac{4F}{\eta}\|\lambda - \lambda_0\|_1$. In particular, we have $\|\nabla F_t(x_t)\lambda_t\|_*^2 + \frac{4F}{\eta}\|\lambda_t - \lambda_0\|_1 \le \|\nabla F_t(x_t)\lambda_0\|_*^2$. Therefore, from Theorem 4.4 we know that

$$R_{\text{MOS}}(T) \le \frac{1}{\eta} B_R(x^*, x_1) + \frac{\eta}{2} \sum_{t=1}^{T} \|\nabla F_t(x_t)\lambda_0\|_*^2.$$

Recall that the domain of $\mathcal{X}$ is bounded by $D$. From Assumption 4.2 we further have $B_R(x^*, x_1) \le \gamma\|x^*, x_1\| \le \gamma D$. Utilize Assumption 4.3, and set $\eta = \frac{\sqrt{2\gamma D}}{G\sqrt{T}}$, then we have

$$R_{\text{MOS}}(T) \le \frac{\gamma D}{\eta} + \frac{\eta}{2} G^2 T = G\sqrt{2\gamma D T},$$

which proves the reduced bound. ■

### I.3 Proof of Lemma E.3

*Proof.* Before analyzing the dynamic regret bound of DR-OMMD, we first introduce two useful lemmas.

**Lemma I.1.** *For the standard OMD algorithm with learning rate $\eta$ and loss ${\lambda_t}^\top F_t(x)$, we have the following recursion*

$${\lambda_t}^\top F_t(x_t) - {\lambda_t}^\top F_t(x_t^*) \le \frac{1}{\eta}(B_R(x_t^*; x_t) - B_R(x_t^*; x_{t+1})) + \frac{\eta}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2, \tag{1}$$

*for any $t \in \{1, \dots, T\}$.*

*Proof.* Our proof is similar to the analysis of OMD in the single-objective setting [28, 7]. Specifically, fix $f_t = {\lambda_t}^\top F_t$ and $g_t = {\lambda_t}^\top F_t(x_t)$. From the convexity of $f_t$, we have

$$f_t(x_t) - f_t(x_t^*) \le g_t^\top (x_t - x_t^*) = g_t^\top (x_{t+1} - x_t^*) + g_t^\top (x_t - x_{t+1}).$$

From the first-order optimal condition of $x_{t+1}$, for any $x' \in \mathcal{X}$, we have

$$(\eta \nabla F_t(x_t)\lambda_t + \nabla R(x_{t+1}) - \nabla R(x_t))^\top (x' - x_{t+1}) \ge 0.$$

We set $x' = x_t^*$ in the above inequality, and consequently derive

$$f_t(x_t) - f_t(x_t^*) \le \frac{1}{\eta}(\nabla R(x_{t+1}) - \nabla R(x_t))^\top (x_t^* - x_{t+1}) + g_t^\top (x_t - x_{t+1}).$$

Recall the definition of Bregman divergence $B_R$. We can check that (also see [2])

$$B_R(x_t^*, x_t) - B_R(x_t^*, x_{t+1}) - B_R(x_{t+1}, x_t) = (\nabla R(x_{t+1}) - \nabla R(x_t))^\top (x_t^* - x_{t+1}).$$

Since $R$ is 1-strongly convex, we have $B_R(x_{t+1}, x_t) \ge \|x_{t+1} - x_t\|^2/2$. Hence

$$f_t(x_t) - f_t(x_t^*) \le \frac{1}{\eta}(B_R(x_t^*, x_t) - B_R(x_t^*, x_{t+1}) - \frac{1}{2}\|x_{t+1} - x_t\|^2) + g_t^\top (x_t - x_{t+1}).$$

Moreover, from the Cauchy-Schwartz inequality we have

$$g_t^\top (x_t - x_{t+1}) \le \frac{\eta}{2}\|g_t\|_*^2 + \frac{1}{2\eta}\|x_t - x_{t+1}\|^2.$$

Combining the above two inequalities, we can prove the lemma. ■

**Lemma I.2.** *For an arbitrary comparator sequence $u_1, \dots, u_T \in \mathcal{X}$, we have*

$$\sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(u_t)) \le \frac{\gamma}{\eta} \sum_{t=1}^{T-1} \|u_t - u_{t+1}\| + \frac{1}{\eta} B_R(u_1, x_1) + \frac{\eta}{2} \sum_{t=1}^{T} \|\nabla F_t(x_t)\lambda_t\|_*^2$$

*Proof.* Summing the inequality in Lemma I.1 over $t \in \{1, \dots, T\}$, we have

$$\sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(u_t)) \le \sum_{t=1}^{T} \frac{1}{\eta}(B_R(u_t, x_t) - B_R(u_t, x_{t+1})) + \frac{\eta}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2$$
$$\le \frac{1}{\eta}\sum_{t=1}^{T-1}(B_R(u_{t+1}, x_{t+1}) - B_R(u_t, x_{t+1})) + \frac{1}{\eta}B_R(u_1, x_1) + \frac{\eta}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2.$$

Recall the assumption that $B_R(x, z) - B_R(y, z) \le \gamma\|x - y\|, \forall x, y, z \in \mathcal{X}$. We then obtain

$$\sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(u_t)) \le \frac{1}{\eta}\sum_{t=1}^{T-1}\gamma\|u_t - u_{t+1}\| + \frac{1}{\eta}B_R(u_1, x_1) + \frac{\eta}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2,$$

which proves the lemma. ■

We can now return to prove Lemma E.3. Notice that, for an arbitrary comparator sequence $u_1, \dots, u_T \in \mathcal{X}$, the multi-objective dynamic regret can be decomposed as

$$\sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t) - \sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t^*)$$
$$= \underbrace{\sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t) - \sum_{t=1}^{T} {\lambda_t}^\top F_t(u_t)}_{\text{term } A} + \underbrace{\sum_{t=1}^{T} {\lambda_t}^\top F_t(u_t) - \sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t^*)}_{\text{term } B}.$$

We now instantiate the comparator sequence to be a piece-wise stationary sequence, i.e.,

$$\{u_1, \dots, u_T\} = \left\{ \underbrace{w^\star_{\mathcal{I}_1}, \dots, w^\star_{\mathcal{I}_1}}_{\Delta \text{ times}}, \dots, \underbrace{w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil - 1}}, \dots, w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil - 1}}}_{\Delta \text{ times}}, \underbrace{w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil}}, \dots, w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil}}}_{(1 + \frac{T}{\Delta} - \lceil \frac{T}{\Delta}\rceil)\Delta \text{ times}} \right\},$$

which starts with $w_1^*$ and only changes for every $\Delta$ steps ($\Delta$ is an integer such that $\Delta \le T$). More specifically, for any $i \in \{1, \dots, \lceil T/\Delta\rceil\}$, denote $p_i = (i-1)\Delta + 1$ and $q_i = i\Delta$, and then $\mathcal{I}_i = [p_i, q_i]$ is exactly the $i$-th stationary piece of the comparator sequence.

For any piece $i \in \{1, \dots, \lceil T/\Delta\rceil\}$, we set all $u_t, t \in \mathcal{I}_i$ to be the best fixed decision $x^\star_{\mathcal{I}_i}$ regarding the cumulative linearized losses during interval $\mathcal{I}_i$, i.e., $u_t \equiv x^\star_{\mathcal{I}_i} = \arg\min_{x \in \mathcal{X}} \sum_{t \in \mathcal{I}_i} {\lambda_t}^\top F_t(x)$, for any $t \in \mathcal{I}_i, i \in \{1, \dots, \lceil T/\Delta\rceil\}$. Then we can apply Lemma I.2 to such comparator sequence and bound the term $A$ as

$$A \le \frac{\gamma}{\eta}\sum_{i=1}^{\lceil T/\Delta\rceil - 1}\|u_{\mathcal{I}_i} - u_{\mathcal{I}_{i+1}}\| + \frac{1}{\eta}B_R(u_{\mathcal{I}_1}, x_1) + \frac{\eta}{2}\sum_{t=1}^{T}\|\nabla F_t(x_t)\lambda_t\|_*^2$$
$$\le \frac{\gamma}{\eta}(\lceil \frac{T}{\Delta}\rceil - 1)D + \frac{\gamma D}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T}\|\nabla F_t(x_t)\lambda_t\|_*^2 = \frac{\gamma D}{\eta}\lceil \frac{T}{\Delta}\rceil + \frac{\eta}{2}\sum_{t=1}^{T}\|\nabla F_t(x_t)\lambda_t\|_*^2.$$

Notice that, the second term $B$ measures the difference between the cumulative linearized loss of the best decisions $x^*_{\mathcal{I}_i}$ regarding each interval $\mathcal{I}_i$ and that of the comparators $\{x_t^*\}$. To analyze this term, we consider such difference $B_i$ restricted to any interval $\mathcal{I}_i, i \in \{1, \dots, \lceil T/\Delta\rceil\}$:

$$B_i = \sum_{t \in \mathcal{I}_i} {\lambda_t}^\top F_t(u_t) - \sum_{t \in \mathcal{I}_i} {\lambda_t}^\top F_t(x_t^*)$$
$$= \sum_{t \in \mathcal{I}_i} {\lambda_t}^\top F_t(u_t) - \sum_{t \in \mathcal{I}_i} {\lambda_{p_i}}^\top F_{p_i}(x_{p_i}^*) + \sum_{t \in \mathcal{I}_i} {\lambda_{p_i}}^\top F_{p_i}(x_{p_i}^*) - \sum_{t \in \mathcal{I}_i} {\lambda_t}^\top F_t(x_t^*).$$

Recall our definition of $u_t$ and $x_t^*$, then we have $\sum_{t\in\mathcal{I}_i} {\lambda_t}^\top F_t(u_t) \le \sum_{t\in\mathcal{I}_i} {\lambda_t}^\top F_t(x_{p_i}^*)$ and ${\lambda_{p_i}}^\top F_{p_i}(x_{p_i}^*) \le {\lambda_{p_i}}^\top F_{p_i}(x_t^*)$. Hence we further have

$$B_i \le \sum_{t\in\mathcal{I}_i} {\lambda_t}^\top F_t(x_{p_i}^*) - \sum_{t\in\mathcal{I}_i} {\lambda_{p_i}}^\top F_{p_i}(x_{p_i}^*) + \sum_{t\in\mathcal{I}_i} {\lambda_{p_i}}^\top F_{p_i}(x_t^*) - \sum_{t\in\mathcal{I}_i} {\lambda_t}^\top F_t(x_t^*).$$

Moreover, for any $t \in \mathcal{I}_i, x \in \mathcal{X}$, we have

$$|{\lambda_t}^\top F_t(x) - {\lambda_{p_i}}^\top F_{p_i}(x)| = |\sum_{k=p_i}^{t-1} ({\lambda_{k+1}}^\top F_{k+1}(x) - {\lambda_k}^\top F_k(x))| \le \sum_{k=p_i}^{q_i} \sup_{x\in\mathcal{X}} |{\lambda_k}^\top F_k(x) - {\lambda_{k+1}}^\top F_{k+1}(x)|.$$

Recall that $\mathcal{I}_i$ has at most $\Delta$ elements, we further have

$$B_i \le 2\Delta \sum_{k=p_i}^{q_i} \sup_{x\in\mathcal{X}} |{\lambda_k}^\top F_k(x) - {\lambda_{k+1}}^\top F_{k+1}(x)|.$$

Hence the term $B$ can be bounded as

$$B = \sum_{i=1}^{\lceil T/\Delta \rceil} B_i \le 2\Delta \sum_{i=1}^{\lceil T/\Delta \rceil} \sum_{k=p_i}^{q_i} \sup_{x\in\mathcal{X}} |{\lambda_k}^\top F_k(x) - {\lambda_{k+1}}^\top F_{k+1}(x)|.$$

From the definition of $p_i$ and $q_i$, we further have

$$B \le 2\Delta \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top F_t(x) - {\lambda_{t+1}}^\top F_{t+1}(x)|.$$

Combining the above two inequalities on terms $A$ and $B$, we derive

$$\begin{aligned} R_{\text{MOD}}(T) \le \frac{\gamma}{\eta} &\sum_{i=1}^{\lceil T/\Delta \rceil - 1} \|u_{\mathcal{I}_i} - u_{\mathcal{I}_{i+1}}\| + \frac{1}{\eta} B_R(u_{\mathcal{I}_1}, x_1) + \frac{\eta}{2} \sum_{t=1}^{T} \|\nabla F_t(x_t)\lambda_t\|_*^2 \\ &+ 2\Delta \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top F_t(x) - {\lambda_{t+1}}^\top F_{t+1}(x)|. \end{aligned}$$

We note that, term $\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top F_t(x) - {\lambda_{t+1}}^\top F_{t+1}(x)|$ in the above regret bound is analogous to the term $\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |f_t(x) - f_{t+1}(x)|$ in the dynamic regret analysis for the single-objective setting. Note that, in the single-objective setting, from the definition of temporal variability, we directly have $\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |f_t(x) - f_{t+1}(x)| \le V_T$. However, things become much more complex in the multi-objective setting, since **the vector loss $F_t(x)$ is now coupled with the composite weights $\lambda_t$**. As a result, we cannot directly use the temporal variability of each loss component to bound this term. To tackle this term, we introduce the following lemma which is highly non-trivial.

**Lemma I.3.** *For DR-OMMD with any initial weights $\lambda_0 \in \mathcal{S}_m$, it holds that*

$$\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top F_t(x) - {\lambda_{t+1}}^\top F_{t+1}(x)| \le 2F \sum_{t=1}^{T} \|\lambda_t - \lambda_0\|_1 + V_T.$$

*Proof.* We decompose the supremum operation as

$$\begin{aligned} &\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top F_t(x) - {\lambda_{t+1}}^\top F_{t+1}(x)| \\ \le &\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} (|(\lambda_t - \lambda_{t+1})^\top F_{t+1}(x)| + |{\lambda_t}^\top (F_t(x) - F_{t+1}(x))|) \\ \le &\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |(\lambda_t - \lambda_{t+1})^\top F_{t+1}(x)| + \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top (F_t(x) - F_{t+1}(x))|. \end{aligned}$$

Then for any $x \in \mathcal{X}$, $|f_t^i(x)| \le F$, and consequently

$$\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |(\lambda_t - \lambda_{t+1})^\top F_{t+1}(x)| \le \sum_{t=1}^{T-1} \|\lambda_t - \lambda_{t+1}\|_1 \|F_{t+1}(x)\|_\infty \le 2F \sum_{t=1}^{T-1} \|\lambda_t - \lambda_{t+1}\|_1.$$

It then remains to tackle the term $\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |\lambda_t^\top (F_t(x) - F_{t+1}(x))|$. For any $\lambda_0 = (\lambda_0^1, ..., \lambda_0^m) \in \mathcal{S}_m$, we can decompose this term as

$$\begin{aligned}
&\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top (F_t(x) - F_{t+1}(x))| \\
=& \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |(\lambda_t - \lambda_0)^\top (F_t(x) - F_{t+1}(x)) + {\lambda_0}^\top (F_t(x) - F_{t+1}(x))| \\
\le& \underbrace{\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |(\lambda_t - \lambda_0)^\top (F_t(x) - F_{t+1}(x))|}_{\text{term A}} + \underbrace{\sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_0}^\top (F_t(x) - F_{t+1}(x))|}_{\text{term B}}.
\end{aligned}$$

Since we have assumed that $|f_t^i(x)| \le F$ for any $x \in \mathcal{X}, t \in \{1, \ldots, T\}, i \in \{1, \ldots, m\}$, it holds that $|f_t^i(x) - f_{t+1}^i(x)| \le 2F$. Consequently, we can bound the term A as

$$\begin{aligned}
\text{term A} &= \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |\sum_{i=1}^{m} (\lambda_t^i - \lambda_0^i)(f_t^i(x) - f_{t+1}^i(x))| \\
&\le \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} \sum_{i=1}^{m} |\lambda_t^i - \lambda_0^i| \cdot |f_t^i(x) - f_{t+1}^i(x)| \\
&\le \sum_{t=1}^{T-1} \sum_{i=1}^{m} |\lambda_t^i - \lambda_0^i| \cdot \sup_{x\in\mathcal{X}} |f_t^i(x) - f_{t+1}^i(x)| \\
&\le \sum_{t=1}^{T-1} \sum_{i=1}^{m} |\lambda_t^i - \lambda_0^i| \cdot 2F = \sum_{t=1}^{T-1} 2F \|\lambda_t - \lambda_0\|_1,
\end{aligned}$$

where $\lambda_t^i$ and $\lambda_0^i$ represents the $i$-th entry of $\lambda_t$ and $\lambda_0$, respectively.

As for term B, we have

$$\begin{aligned}
\text{term B} &= \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |\lambda_0^\top (F_t(x) - F_{t+1}(x))| \le \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} \sum_{i=1}^{m} \lambda_0^i \cdot |f_t^i(x) - f_{t+1}^i(x)| \\
&\le \sum_{i=1}^{m} \lambda_0^i \cdot \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |f_t^i(x) - f_{t+1}^i(x)| \le \sum_{i=1}^{m} \lambda_0^i \cdot V_T = V_T,
\end{aligned}$$

where the last inequality is derived from the assumption of temporal variability. Combining the above bounds for term A and term B, we finally prove the lemma. ■

Assume that the diameter of the decision domain $\mathcal{X}$ is upper bounded by some $D > 0$, then $\|u_{\mathcal{I}_i} - u_{\mathcal{I}_{i+1}}\| \le D$ and $B_R(u_{\mathcal{I}_1}, x_1) \le \gamma \|u_{\mathcal{I}_1} - x_1\| \le \gamma D$. Hence we have

$$\frac{\gamma}{\eta} \sum_{i=1}^{\lceil T/\Delta \rceil - 1} \|u_{\mathcal{I}_i} - u_{\mathcal{I}_{i+1}}\| + \frac{1}{\eta} B_R(u_{\mathcal{I}_1}, x_1) \le \frac{\gamma D}{\eta}(\lceil T/\Delta \rceil - 1) + \frac{\gamma D}{\eta} = \frac{\gamma D}{\eta} \lceil \frac{T}{\Delta} \rceil.$$

Plugging the above result and lemma into the above bound for $R_{\text{MOD}}(T)$, and replace the quantity $\Delta \in \{1, \ldots, T\}$ by $\delta$, we have

$$R_{\text{MOD}}(T) \le 2\delta V_T + 4\delta F \sum_{t=1}^{T} \|\lambda_t - \lambda_0\|_1 + \frac{\eta}{2} \sum_{t=1}^{T} \|\nabla F_t(x_t)\lambda_t\|_*^2 + \frac{\gamma D}{\eta} \lceil \frac{T}{\delta} \rceil,$$

for any $\delta \in \{1, \ldots, T\}$. We thus prove the lemma. ■

### I.4 Proof of Theorem 4.6

*Proof.* This theorem can be directly derived from Lemma E.3.

Specifically, when $\frac{4V_T}{G^2T} \le \eta \le \frac{4V_T}{G^2}$, we can set $\delta = \frac{\eta G^2 T}{V_T}$, which satisfies $1 \le \delta \le T$. Plugging it into Lemma E.3 and rearranging the inequality, we can directly derive the theorem. ■

### I.5 Proof of Theorem E.4

*Proof.* Since this theorem is a special case of its following Theorem E.5 when $\Omega(1) \le V_T \le o(T)$, it can be proved as we derive Theorem E.5 in the following subsection.

In fact, as we assume $\Omega(1) \le V_T \le o(T)$, in the following derivation of Theorem E.5, we are always in the case of (i). In addition, when $\Omega(1) \le V_T \le o(T)$ in Theorem E.5 we exactly have $\eta = \frac{2}{G}(\frac{\gamma D V_T}{GT})^{1/3}$. Hence this theorem can be directly derived from Theorem E.5. ■

### I.6 Proof of Theorem E.5

*Proof.* Denote $\eta_0 = \frac{2}{G}(\frac{\gamma D V_T}{GT})^{1/3}$. We consider the following three cases:

(i) When $\frac{4V_T}{G^2T} \le \eta_0 \le \frac{4V_T}{G^2}$, we can directly apply the above lemma.

(ii) When $\eta_0 < \frac{4V_T}{G^2T}$, or equivalently $V_T > (\frac{\gamma D}{8})^{1/2} GT$, we have $\eta = \frac{4V_T}{G^2T}$. Set $\delta = 1$ in Lemma E.3, then it can be verified that

$$\begin{aligned}
R_{\text{MOD}}(T) &\le 2V_T + \frac{2V_T}{G^2T}\sum_{t=1}^{T}(\alpha\|\lambda_t - \lambda_0\|_1 + \|\nabla F_t(x_t)\lambda_t\|_*^2) + \frac{\gamma D G^2 T^2}{4V_T} \\
&\le 4V_T + \frac{2V_T}{G^2T}\sum_{t=1}^{T}\|\nabla F_t(x_t)\lambda_0\|_*^2 \\
&\le 4V_T + \frac{2V_T}{G^2T}\sum_{t=1}^{T} G^2 = 6V_T.
\end{aligned}$$

(iii) When $\eta_0 > \frac{4V_T}{G^2}$, or equivalently $V_T < (\frac{\gamma D}{8T})^{1/2} G$, we have $\eta = \frac{4V_T}{G^2}$. Set $\delta = T$ in Lemma E.3, then it can be verified that

$$\begin{aligned}
R_{\text{MOD}}(T) &\le 2TV_T + \frac{2V_T}{G^2}\sum_{t=1}^{T}(\alpha\|\lambda_t - \lambda_0\|_1 + \|\nabla F_t(x_t)\lambda_t\|_*^2) + \frac{\gamma D G^2}{4V_T} \\
&\le G(2\gamma D)^{1/2}T^{1/2} + \frac{2V_T}{G^2}\sum_{t=1}^{T}\|\nabla F_t(x_t)\lambda_0\|_*^2 \\
&\le G(2\gamma D)^{1/2}T^{1/2} + \frac{2V_T}{G^2}\sum_{t=1}^{T} G^2 \le \frac{3G}{2}(2\gamma D)^{1/2}T^{1/2}.
\end{aligned}$$

Combining (i)-(iii), we prove the theorem. ■

### I.7 Proof of Corollary E.2

*Proof.* Since this corollary is a special case of its following Corollary E.6 when $\Omega(1) \le V_T \le o(T)$, it can be directly derived from Corollary E.6.

Specifically, since $V_T \ge \Omega(1)$, we have $V_T^{1/3}T^{2/3} \ge O(T^{1/2})$. Moreover, since $V_T \le o(T)$, we have $V^{1/3}T^{2/3} \ge o(V_T)$. Therefore, the dominating term in the bound of Corollary E.6 is $V^{1/3}T^{2/3}$, which proves the corollary. ■

### I.8 Proof of Corollary E.6

*Proof.* We start from the general bound derived in Theorem E.5. Specifically, in the first regret term, since $\lambda_t$ is selected to minimize $\alpha\|\lambda_t - \lambda_{t-1}\|_1 + \|\nabla F_t(x_t)\lambda_t\|_2^*$ at each step $t$, we further have

$$\min_{\lambda \in \mathcal{S}_m} \{\alpha\|\lambda - \lambda_{t-1}\|_1 + \|\nabla F_t(x_t)\lambda\|_*^2\} \leq \|\nabla F_t(x_t)\lambda_{t-1}\|_*^2 \leq (\|\lambda_{t-1}\|_1 \|\nabla F_t(x_t)\|_\infty)^2 \leq G^2.$$

Plugging it into the bound in Theorem E.5 directly proves the corollary. ■

## J Omitted Proofs of the Regret Bounds for A-DR-OMMD

The proofs of A-DR-OMMD are very similar to the proofs of DR-OMMD in Appendix I, except that we replace the fixed learning rate $\eta$ by the adaptive ones $\eta_t$. For conciseness, we here only highlight the different steps from the proofs of DR-OMMD.

### J.1 Proof of Theorem D.2

*Proof.* For any $\lambda \in \mathcal{S}_m$ and $\lambda_1 \dots, \lambda_T \in \mathcal{S}_m$, by the same deduction in the proof of Theorem 4.4 in Appendix I.1, we have

$$R_{\text{MOS}}(T) \leq 2F \sum_{t=1}^{T} \|\lambda - \lambda_t\|_1 + \sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(x^*)).$$

To proceed, we introduce two lemmas, which are the analogies of Lemma I.1 and I.2 in Appendix I.3.

**Lemma J.1.** *For the standard OMD algorithm with non-increasing learning rates $\eta_t$ and loss ${\lambda_t}^\top F_t(x)$, we have the following recursion*

$$ {\lambda_t}^\top F_t(x_t) - {\lambda_t}^\top F_t(x_t^*) \leq \frac{1}{\eta_t}(B_R(x_t^*; x_t) - B_R(x_t^*; x_{t+1})) + \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2, \tag{2}$$

*for any $t \in \{1, \dots, T\}$.*

*Proof.* Since the lemma is only regarding any single round $t$, the proof is nearly the same as the proof of Lemma I.1 in Appendix I.3. In fact, we just need to replace all $\lambda$ in the proof of Lemma I.1 by $\lambda_t$, then we recovers the proof of this lemma. ■

**Lemma J.2.** *For an arbitrary comparator sequence $u_1, \dots, u_T \in \mathcal{X}$, we have*

$$\sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(u_t)) \leq \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T-1} \frac{\gamma}{\eta_t}\|u_t - u_{t+1}\| + \sum_{t=1}^{T} \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2.$$

*Proof.* Summing the inequality in Lemma J.1 over $t \in \{1, \dots, T\}$, we have

$$\sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(u_t)) \leq \sum_{t=1}^{T} \frac{1}{\eta_t}(B_R(u_t, x_t) - B_R(u_t, x_{t+1})) + \sum_{t=1}^{T} \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2.$$

Since the Bregman divergence is always non-negative, we have

$$\begin{aligned}
&\sum_{t=1}^{T} \frac{1}{\eta_t}(B_R(u_t, x_t) - B_R(u_t, x_{t+1})) \\
&\qquad \leq \sum_{t=1}^{T-1} (\frac{1}{\eta_{t+1}} B_R(u_{t+1}, x_{t+1}) - \frac{1}{\eta_t} B_R(u_t, x_{t+1})) + \frac{1}{\eta_1} B_R(u_1, x_1).
\end{aligned}$$

Now we consider each summation term. Since $\{\eta_t\}_{t=1}^T$ is non-increasing, we have $\frac{1}{\eta_t} \leq \frac{1}{\eta_{t+1}}$, hence

$$\begin{aligned}
&\frac{1}{\eta_{t+1}} B_R(u_{t+1}, x_{t+1}) - \frac{1}{\eta_t} B_R(u_t, x_{t+1}) \\
&= (\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}) B_R(u_{t+1}, x_{t+1}) + \frac{1}{\eta_t}(B_R(u_{t+1}, x_{t+1}) - B_R(u_t, x_{t+1}))
\end{aligned}$$

Recall the assumption that $B_R(x,z) - B_R(y,z) \leq \gamma\|x-y\|, \forall x,y,z \in \mathcal{X}$. We then obtain $B_R(u_{t+1},x_{t+1}) - B_R(u_t,x_{t+1}) \leq \gamma\|u_{t+1}-u_t\|$ and $B_R(u_{t+1},x_{t+1}) \leq \gamma D$. Hence

$$\frac{1}{\eta_{t+1}} B_R(u_{t+1},x_{t+1}) - \frac{1}{\eta_t} B_R(u_t,x_{t+1}) \leq (\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t})\gamma D + \frac{1}{\eta_t}\gamma\|u_t - u_{t+1}\|.$$

Plugging it into the desired formula, we have

$$\sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(u_t)) \leq \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T-1} \frac{\gamma}{\eta_t}\|u_t - u_{t+1}\| + \sum_{t=1}^{T} \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2,$$

which proves the lemma. ∎

Now set $u_1 = u_2 = \ldots = u_T = x^*$ in Lemma J.2, we have

$$\sum_{t=1}^{T} {\lambda_t}^\top (F_t(x_t) - F_t(x^*)) \leq \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T} \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2.$$

Consequently, the static regret can be bounded as

$$\begin{aligned} R_{\text{MOS}}(T) &\leq 2F \sum_{t=1}^{T} \|\lambda_t - \lambda_0\|_1 + \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T} \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2 \\ &= \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T} \frac{\eta_t}{2}(\|\nabla F_t(x_t)\lambda_t\|_*^2 + \frac{4F}{\eta_t}\|\lambda_t - \lambda_0\|_1), \end{aligned}$$

which proves the theorem. ∎

### J.2 Proof of Corollary D.3

*Proof.* In A-DR-OMMD, when $\alpha_t = \frac{4F}{\eta_t}$, from the formulation of $\lambda_t$ we have

$$\begin{aligned} R_{\text{MOS}}(T) &= \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T} \frac{\eta_t}{2} \min_{\lambda \in \mathcal{S}_m} \{\|\nabla F_t(x_t)\lambda\|_*^2 + \alpha_t\|\lambda - \lambda_0\|_1\} \\ &\leq \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T} \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_0\|_*^2 \leq \frac{\gamma D}{\eta_T} + \frac{G^2}{2}\sum_{t=1}^{T} \eta_t. \end{aligned}$$

When $\eta_t = \frac{\sqrt{2\gamma D}}{G\sqrt{t}}$, we further have

$$R_{\text{MOS}}(T) \leq G\sqrt{2\gamma D T},$$

which proves corollary. ∎

### J.3 Proof of Theorem D.4

Similar to the proof of Theorem 4.6, before analyzing the dynamic regret of A-DR-OMMD, we first introduce a lemma, which is an analogy to Lemma E.3 with adaptive $\eta_t$ and $\alpha_t$.

**Lemma J.3.** *Suppose the assumptions of Lemma E.3 hold. Then for any $\lambda_0 \in \mathcal{S}_m, \Delta \in \{1,\ldots,T\}$, and non-increasing $\{\eta_t\}_{t=1}^T$, DR-OMMD attains the following dynamic regret*

$$R_{MOD}(T) \leq \lceil \frac{T}{\Delta} \rceil \frac{\gamma D}{\eta_T} + \sum_{t=1}^{T} \frac{\eta_t}{2}\|\nabla F_t(x_t)\lambda_t\|_*^2 + 4\Delta F \sum_{t=1}^{T} \|\lambda_t - \lambda_0\|_1 + 2\Delta V_T.$$

*Proof.* Similar to the proof of Lemma E.3, we decompose the regret as

$$\begin{aligned} &\sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t) - \sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t^*) \\ &= \underbrace{\sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t) - \sum_{t=1}^{T} {\lambda_t}^\top F_t(u_t)}_{\text{term } A} + \underbrace{\sum_{t=1}^{T} {\lambda_t}^\top F_t(u_t) - \sum_{t=1}^{T} {\lambda_t}^\top F_t(x_t^*)}_{\text{term } B}, \end{aligned}$$

and construct a comparator sequence

$$\{u_1,\dots,u_T\} = \left\{ \underbrace{w^\star_{\mathcal{I}_1},\dots,w^\star_{\mathcal{I}_1}}_{\Delta \text{ times}},\dots,\underbrace{w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil -1}},\dots,w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil -1}}}_{\Delta \text{ times}},\underbrace{w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil}},\dots,w^\star_{\mathcal{I}_{\lceil \frac{T}{\Delta}\rceil}}}_{(1+\frac{T}{\Delta}-\lceil \frac{T}{\Delta}\rceil)\Delta \text{ times}} \right\}.$$

For any $i \in \{1,\dots,\lceil T/\Delta\rceil\}$, denote $p_i = (i-1)\Delta + 1$ and $q_i = i\Delta$. Then $\mathcal{I}_i = [p_i, q_i]$ is the $i$-th stationary piece of the comparator sequence. For any $i \in \{1,\dots,\lceil T/\Delta\rceil\}$, we set all $u_t, t \in \mathcal{I}_i$ to be $x^\star_{\mathcal{I}_i} = \arg\min_{x\in\mathcal{X}} \sum_{t\in\mathcal{I}_i} {\lambda_t}^\top F_t(x)$. Then we can apply Lemma J.2 to bound the term $A$ as

$$\begin{aligned} A &\le \sum_{i=1}^{\lceil T/\Delta\rceil -1} \frac{\gamma}{\eta_{p_i}} \|u_{\mathcal{I}_i} - u_{\mathcal{I}_{i+1}}\| + \frac{\gamma D}{\eta_T} + \sum_{t=1}^T \frac{\eta_t}{2} \|\nabla F_t(x_t)\lambda_t\|_*^2 \\ &\le \gamma D\Big(\frac{1}{\eta_T} + \sum_{i=1}^{\lceil T/\Delta\rceil -1} \frac{1}{\eta_{p_i}}\Big) + \sum_{t=1}^T \frac{\eta_t}{2} \|\nabla F_t(x_t)\lambda_t\|_*^2. \end{aligned}$$

Since $\{\eta_t\}_{t=1}^T$ is non-increasing, we further have

$$A \le \lceil \frac{T}{\Delta} \rceil \frac{\gamma D}{\eta_T} + \sum_{t=1}^T \frac{\eta_t}{2} \|\nabla F_t(x_t)\lambda_t\|_*^2.$$

The analysis of the term $B$ is the same as that in the proof of Lemma E.3, i.e.,

$$B \le 2\Delta \sum_{t=1}^{T-1} \sup_{x\in\mathcal{X}} |{\lambda_t}^\top F_t(x) - {\lambda_{t+1}}^\top F_{t+1}(x)|.$$

We can further apply Lemma I.3 to the term $B$ and derive

$$B \le 4\Delta F \sum_{t=1}^T \|\lambda_t - \lambda_0\|_1 + 2\Delta V_T.$$

Combining the above two inequalities, we have

$$R_{\text{MOD}}(T) \le \lceil \frac{T}{\Delta} \rceil \frac{\gamma D}{\eta_T} + \sum_{t=1}^T \frac{\eta_t}{2} \|\nabla F_t(x_t)\lambda_t\|_*^2 + 4\Delta F \sum_{t=1}^T \|\lambda_t - \lambda_0\|_1 + 2\Delta V_T,$$

which proves the lemma. ■

Now we can go back to prove Theorem D.4.

*Proof of Theorem D.4.* We start from the general bound in Lemma J.3. Set $\Delta = \frac{\eta_T G^2 T}{V_T}$, then

$$R_{\text{MOD}}(T) \le \frac{\gamma D V_T}{G^2 \eta_T^2} + \sum_{t=1}^T \frac{\eta_t}{2} \|\nabla F_t(x_t)\lambda_t\|_*^2 + \frac{4\eta_T G^2 F T}{V_T} \sum_{t=1}^T \|\lambda_t - \lambda_0\|_1 + 2\eta_T G^2 T,$$

which proves the theorem. ■

### J.4 Proof of Corollary D.5

*Proof.* When $\alpha_t = \frac{8G^2FT}{V_T}(\frac{\eta_T}{\eta_t})$, from the formulation of $\lambda_t$, we have

$$\begin{aligned} R_{\text{MOD}}(T) &\le \frac{\gamma D V_T}{G^2 \eta_T^2} + \sum_{t=1}^T \frac{\eta_t}{2} \operatorname*{arg\,min}_{\lambda\in\mathcal{S}_m} \{\|\nabla F_t(x_t)\lambda\|_*^2 + \alpha_t \|\lambda - \lambda_0\|_1\} + 2\eta_T G^2 T \\ &\le \frac{\gamma D V_T}{G^2 \eta_T^2} + \sum_{t=1}^T \frac{\eta_t}{2} \|\nabla F_t(x_t)\lambda_0\|_*^2 + 2\eta_T G^2 T \\ &\le \frac{\gamma D V_T}{G^2 \eta_T^2} + \frac{G^2}{2} \sum_{t=1}^T \eta_t + 2\eta_T G^2 T. \end{aligned}$$

When $\eta_t = \frac{1}{G}(\frac{\gamma D V_T}{Gt})^{1/3}$, since by induction we can prove that $\sum_{t=1}^{T}(\frac{1}{t})^{1/3} \leq 2T^{2/3}$, we have

$$R_{\text{MOD}}(T) \leq O(T^{2/3}V_T^{1/3}),$$

which proves the corollary. ■

# References


[1] Jacob Abernethy, Peter L Bartlett, and Elad Hazan. Blackwell approachability and no-regret learning are equivalent. In *Annual Conference on Learning Theory*, pages 27–46, 2011.

[2] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

[3] Omar Besbes, Yonatan Gur, and Assaf Zeevi. Non-stationary stochastic optimization. *Operations Research*, 63(5):1227–1244, 2015.

[4] David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.

[5] Róbert Busa-Fekete, Balázs Szörényi, Paul Weng, and Shie Mannor. Multi-objective bandits: Optimizing the generalized gini index. In *International Conference on Machine Learning*, pages 625–634. PMLR, 2017.

[6] Nicolò Campolongo and Francesco Orabona. A closer look at temporal variability in dynamic online learning. *arXiv preprint arXiv:2102.07666*, 2021.

[7] Nicolò Cesa-bianchi, Pierre Gaillard, Gabor Lugosi, and Gilles Stoltz. Mirror descent meets fixed share (and feels no regret). In *Advances in Neural Information Processing Systems*, pages 980–988, 2012.

[8] Zhao Chen, Jiquan Ngiam, Yanping Huang, Thang Luong, Henrik Kretzschmar, Yuning Chai, and Dragomir Anguelov. Just pick a sign: Optimizing deep multitask models with gradient sign dropout. *Advances in Neural Information Processing Systems*, 2020.

[9] Rémy Degenne, Thomas Nedelec, Clément Calauzènes, and Vianney Perchet. Bridging the gap between regret minimization and best arm identification, with application to a/b tests. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1988–1996, 2019.

[10] Jean-Antoine Désidéri. Multiple-gradient descent algorithm for multiobjective optimization. *Comptes Rendus Mathematique*, 350(5-6):313–318, 2012.

[11] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.

[12] Michael TM Emmerich and André H Deutz. A tutorial on multiobjective optimization: Fundamentals and evolutionary methods. *Natural Computing*, 17(3):585–609, 2018.

[13] Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

[14] Martin Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *International Conference on Machine Learning*, pages 427–435, 2013.

[15] Adrián Javaloy and Isabel Valera. Rotograd: Gradient homogenization in multitask learning. *arXiv preprint arXiv:2103.02631*, 2021.

[16] Wouter Koolen. The pareto regret frontier. *Advances in Neural Information Processing Systems*, pages 1–9, 2013.

[17] Wouter M Koolen and Tim Van Erven. Second-order quantile methods for experts and combinatorial games. In *Conference on Learning Theory*, pages 1155–1175, 2015.

[18] Xi Lin, Hui-Ling Zhen, Zhenhua Li, Qing-Fu Zhang, and Sam Kwong. Pareto multi-task learning. In *Advances in Neural Information Processing Systems*, pages 12060–12070, 2019.

[19] Bo Liu, Xingchao Liu, Xiaojie Jin, Peter Stone, and Qiang Liu. Conflict-averse gradient descent for multi-task learning. *arXiv preprint arXiv:2110.14048*, 2021.

[20] Shiyin Lu, Guanghui Wang, Yao Hu, and Lijun Zhang. Multi-objective generalized linear bandits. In *International Joint Conference on Artificial Intelligence*, pages 3080–3086, 2019.

[21] Shie Mannor, Vianney Perchet, and Gilles Stoltz. Approachability in unknown games: Online learning meets multi-objective optimization. In *Annual Conference on Learning Theory*, pages 339–355, 2014.

[22] Brendan McMahan and Matthew Streeter. Delay-tolerant algorithms for asynchronous distributed online learning. *Advances in Neural Information Processing Systems*, 27, 2014.

[23] Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.

[24] Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *Advances in Neural Information Processing Systems*, pages 3066–3074, 2013.

[25] Sashank J Reddi, Satyen Kale, and Sanjiv Kumar. On the convergence of Adam and beyond. In *International Conference on Learning Representations*, 2018.

[26] Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. Dynamic routing between capsules. In *Advances in Neural Information Processing Systems*, pages 3859–3869, 2017.

[27] Ozan Sener and Vladlen Koltun. Multi-task learning as multi-objective optimization. In *Advances in Neural Information Processing Systems*, pages 525–536, 2018.

[28] Nati Srebro, Karthik Sridharan, and Ambuj Tewari. On the universality of online mirror descent. In *Advances in Neural Information Processing Systems*, pages 2645–2653, 2011.

[29] Cem Tekin and Eralp Turğay. Multi-objective contextual multi-armed bandit with a dominant objective. *IEEE Transactions on Signal Processing*, 66(14):3799–3813, 2018.

[30] Eralp Turgay, Doruk Oner, and Cem Tekin. Multi-objective contextual bandit problem with similarity information. In *International Conference on Artificial Intelligence and Statistics*, pages 1673–1681, 2018.

[31] Tianbao Yang, Lijun Zhang, Rong Jin, and Jinfeng Yi. Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient. In *International Conference on Machine Learning*, pages 449–457, 2016.

[32] Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. In *Advances in Neural Information Processing Systems*, pages 5824–5836, 2020.

[33] Lijun Zhang, Tianbao Yang, Zhi-Hua Zhou, et al. Dynamic regret of strongly adaptive methods. In *International Conference on Machine Learning*, pages 5882–5891, 2018.

[34] Peng Zhao, Yu-Jie Zhang, Lijun Zhang, and Zhi-Hua Zhou. Dynamic regret of convex and smooth functions. *Advances in Neural Information Processing Systems*, 33:12510–12520, 2020.