THEORY FOR GROUP-ROBUST CONTINUAL LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

To support the continuous adaptation of agentic AI systems as new task data arrives, continual learning (CL) has become increasingly important. However, rehearsal-based methods, a dominant class of approaches in CL, typically treat all past tasks equally to mitigate catastrophic forgetting. In this paper, we prove that this practice can destabilize learning by causing poor retention of knowledge from rare or difficult task groups in terms of large excess risk. Accordingly, we develop *group-robust continual learning* (GRCL), a new framework for achieving stable and robust performance across task groups, with a particular focus on improving the worst-group performance. We design two rehearsal-based GRCL methods guided by excess risk bounds: group-balanced replay allocates memory uniformly across groups to minimize a simplified bound, while difficulty-adaptive replay incorporates group learning difficulty estimated during CL into its replay buffer allocation and minimizes a tighter, difficulty-aware bound. We further extend GRCL to the setting where group information is unknown throughout CL. In this case, we propose two group identification methods based on maximum mean discrepancy and online expectation-maximization, respectively, and establish theoretical guarantees for their identification performance. We also derive a small excess-risk bound for GRCL using the estimated group identities. Experiments on real-world datasets show that our rehearsal-based GRCL methods significantly improve robustness over classicalbaselines by enhancing worst-group performance.

1 INTRODUCTION

The recent rise of agentic AI systems, which autonomously plan, act, and interact with dynamic environments, marks a shift from static learning to continuous, closed-loop decision-making. To operate over long horizons, such systems must adapt to changing demands and expanding knowledge without forgetting previously learned information. However, continually acquiring new information can make them prone to catastrophic forgetting. For example, a smart-home customer-support agent trained mostly on routine requests such as password resets and device setup may forget how to handle rare but safety-critical cases including gas-leak sensor failures and emergency escalation. Continual learning (CL), which studies how models learn from sequential tasks while mitigating forgetting, is therefore essential for robust, self-evolving AI agents (Liu et al., 2023). Among CL approaches, rehearsal-based methods store a small set of samples from previous tasks and replay them during training, and have shown strong empirical effectiveness (Chaudhry et al., 2019; Rolnick et al., 2019; Garg et al., 2024; Hacohen & Tuytelaars, 2025; Aghasanli et al., 2025).

However, current rehearsal-based CL methods face a critical challenge: differences in arrival frequency and learning difficulty can lead to imbalanced catastrophic forgetting across task groups. Strong average performance can conceal systematic failures on rare or difficult groups, creating a false sense of reliability and potentially masking AI hallucinations in agentic AI systems. As observed in the Waterbirds dataset (Wah et al., 2011; Zhou et al., 2017; Sagawa et al., 2019), where waterbirds appear predominantly on water and landbirds on land, models may rely on spuriously correlated backgrounds to distinguish bird types, creating a misleading impression of model reliability. Such practice can undermine the robustness of CL in agentic AI systems. In the smart-home customer-support example, failing to diagnose a recurring smoke-detector malfunction from ambiguous user reports could directly compromise user safety.

This imbalance also raises serious fairness concerns, particularly when groups correspond to demographic or otherwise underrepresented subpopulations. Because average performance is dominated

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

by majority groups, it may fail to reflect poor performance on minority groups. For example, in smart-home customer support, failing to retain knowledge of minority customer preferences or underrepresented usage patterns may lead to consistently poorer service for these users. All these examples show that overlooking task-group imbalance can make traditional CL methods unreliable and unfair.

One related line of work mitigates model bias toward frequent classes in long-tailed or class-imbalanced CL by balancing replay buffers or reweighting gradients across classes (Chrysakis & Moens, 2020; Kim et al., 2020; Belouadah et al., 2020; Cano & Krawczyk, 2022; Liu et al., 2022; He, 2024). However, these approaches primarily address frequency imbalance across class labels, largely overlooking inherent differences in learning difficulty across groups corresponding to data features, which typically exhibit distinct learning or forgetting behavior. Another related line of work partially accounts for differences in learning difficulty through replay buffer heuristics that select representative or diverse samples (Rebuffi et al., 2017; Bang et al., 2021; Yoon et al., 2022; Wang et al., 2022; Tong et al., 2025). However, these methods do not explicitly account for group structure and are not designed to protect rare or difficult task groups or ensure robust performance across heterogeneous groups. Consequently, rare or difficult groups still remain underserved throughout CL. Beyond these methodological limitations, both lines of work are largely empirical, while existing CL theory primarily focuses on classical rehearsal methods without explicitly considering task-group imbalance (Deng et al., 2025). Thus, there remains a lack of both methods that ensure robust CL across groups with different arrival frequencies and learning difficulties and theoretical understanding of when and why rehearsal methods succeed or fail under such group imbalance.

Achieving group robustness is further complicated when group information is unobservable. Outside CL, group distributionally robust optimization (GDRO) addresses robustness and fairness under known group labels by shifting from average loss minimization to worst-group loss minimization, primarily in offline or stochastic approximation settings with full data access (Sagawa et al., 2019; Haghtalab et al., 2022; Zhang et al., 2023; 2021; Lei et al., 2021; Farnia & Ozdaglar, 2021; Qi et al., 2021). However, a fundamental gap remains between these settings and CL, where tasks arrive sequentially in a nonstationary, potentially adversarial manner, task-group information may be unobservable, access to past data is limited, and group imbalance can interact with catastrophic forgetting over time. Together, these limitations raise a fundamental question: **How can we achieve group-robust CL with rigorous theoretical guarantees?**

Our key novelties and main contributions are summarized as follows:

- In Section 2, we first prove that classicalrehearsal-based CL can lead to poor retention of rare or difficult task groups in terms of large excess risk. We then develop the *group-robust continual learning* (GRCL), a new framework for achieving balanced performance across task groups in CL, with a particular focus on improving the worst-group performance.
- In Section 3, we introduce a new CL model update rule for GRCL in rehearsal-based CL and establish an excess risk upper bound for it. We then propose two replay buffer allocation strategies in the rehearsal-based method to reduce this excess risk: group-balanced replay and difficulty-adaptive replay. The former minimizes a simplified upper bound by allocating replay memory uniformly across groups, while the latter minimizes a tighter, difficulty-aware bound by estimating group learning difficulty during CL and allocating more memory to harder groups.
- We further bridge the gap between the proposed replay buffer allocation strategies and practical algorithms in Section 4 by addressing two challenges: the replay buffer can only be updated using the current dataset, so the target allocation may not be achieved exactly, and group information may be unknown. To address the first challenge, we introduce Allocation-Track Reservoir Sampling, which tracks the target group allocation as closely as possible. To address the second, we propose two group identification methods based on maximum mean discrepancy (MMD) and online expectation-maximization (EM), both with theoretical guarantees. We also establish an excess-risk bound for GRCL under estimated group identities.
- We conduct extensive experiments on synthetic data and standard group-robust learning benchmarks, including Waterbirds (Sagawa et al., 2019) and CelebA (Liu et al., 2015) with multiple groups of tasks, in Section 5. The results show that our group identification methods achieve high identification accuracy, and our rehearsal-based GRCL significantly improves

worst-group performance compared with classicalrehearsal-based CL in both known and unknown group settings. In addition to strong worst-group performance, our methods also achieve high performance on standard CL metrics that measure average performance.

2 System Model and Problem Formulation

In this section, we first introduce the standard rehearsal-based CL framework. We then formulate the GRCL and theoretically show that classical rehearsal-based CL methods can incur large excess risk with respect to the worst group performance.

2.1 Rehearsal-based CL

In CL, tasks arrive sequentially over time steps $t = 1, \dots, T$. At each step t, the learner receives a dataset $D_t = \{z_i\}_{i=1}^{n_t} \subseteq \mathcal{Z}$, where $\mathcal{Z}$ denotes the instance space and n_t is the number of samples in task t. We assume that each dataset D_t belongs to a single group and assign group indices in the order of their first appearance. Let K_t denote the number of groups observed by time t, $[K_t] = \{1, \dots, K_t\}$ the set of observed group indices, and $I_t \in [K_t]$ the group index of dataset D_t. If a previously unseen group arrives at time $t+1$, we set $K_{t+1} = K_t + 1$, assign it the identity $I_{t+1} = K_{t+1}$, and update the observed group set as $[K_{t+1}] = [K_t] \cup \{K_t + 1\}$. Otherwise, $K_{t+1} = K_t$ and $I_{t+1} \in [K_t]$. The samples in D_t are drawn from the underlying distribution $\mathcal{D}_{I_t}$ associated with group I_t. The learner then incrementally updates the model parameter $w_t \in \mathcal{W} \subset \mathbb{R}^d$ using the newly arrived dataset and information retained from previous tasks.

To mitigate forgetting, rehearsal-based CL methods maintain a replay buffer M_t with fixed capacity M that stores representative samples from earlier tasks: $M_t \subseteq \bigcup_{s=1}^{t-1} D_s$ and $|M_t| \le M$. Typically, the replay buffer size is much smaller than the number of samples in each task, i.e., $M \ll n_t$ for all t. This memory serves as a proxy for past data, enabling the model to revisit earlier tasks while learning from newly arriving data. Rehearsal-based CL methods typically consist of two key components (Rolnick et al., 2019; Deng et al., 2025):

- **Replay buffer allocation**: the update rule that determines which samples from the current replay buffer M_t and the current dataset D_t are retained in the next replay buffer M_{t+1};
- **Model update rule**: the rule used to train the model $w_t \in \mathcal{W}$ based on the current dataset D_t and the replay buffer M_t.

2.2 GRCL Formulation and Limitations of Classical Rehearsal-based CL

In this section, we explain why classical rehearsal methods can perform poorly under task-group imbalance and formulate GRCL. Recent theoretical work on rehearsal-based CL (e.g., (Deng et al., 2025)) studies a basic method that is also commonly used in empirical studies (e.g., (Chaudhry et al., 2019; Rolnick et al., 2019)). This method adopts reservoir sampling as the replay buffer allocation strategy at each time step t, ensuring that every previously observed sample has an equal probability of being retained and thus yielding an unbiased approximation of the historical data distribution. The model w_t is then trained by minimizing the following empirical risk minimization (ERM) objective:

$$w_t \in \arg\min_{w \in \mathcal{W}} \frac{1}{|M_t| + n_t} \sum_{z \in M_t \cup D_t} \ell(w, z), \tag{1}$$

where $\ell(w, z)$ denotes the loss incurred by model w on sample z. This classical rehearsal-based CL aims to minimize the expected excess risk of the model w_t over all arrived tasks (Deng et al., 2025):

$$\mathbb{E}\left[\frac{1}{t}\sum_{\tau=1}^{t} L_{I_\tau}(w_t) - \min_{w \in \mathcal{W}} \frac{1}{t}\sum_{\tau=1}^{t} L_{I_\tau}(w)\right], \quad \text{where } L_{I_\tau}(w) := \mathbb{E}_{z \sim \mathcal{D}_{I_\tau}}[\ell(w, z)]. \tag{2}$$

However, this classical framework has two key limitations: reservoir sampling preserves samples according to arrival frequency, thereby underrepresenting rare groups, and ERM ignores heterogeneous learning difficulty across groups. As a result, groups that are rare or distributionally distinct may suffer disproportionate forgetting, leading to imbalanced and unfair CL performance.

To address this issue, we instead aim to minimize the group-robust objective in CL that extends the static GDRO objective (Sagawa et al., 2019; Qi et al., 2021; Zhang et al., 2023), treating the dynamic

K_t groups independently of their arrival frequencies and explicitly optimizing for the worst-group:

$$\mathcal{E}_{\mathrm{GRCL}}(w_t) := \mathbb{E}\left[\phi_t(w_t) - \min_{w\in\mathcal{W}} \phi_t(w)\right], \tag{3}$$

where

$$\phi_t(w) := \max_{q\in\Delta_{K_t}} f_t(w,q), \qquad f_t(w,q) := \sum_{k=1}^{K_t} q^k L_k(w) - \lambda D(q \,\|\, u_t). \tag{4}$$

Here $q = (q^1, \ldots, q^{K_t}) \in \Delta_{K_t}$ is an adversarial group-weighting that emphasizes groups with larger losses, and $D(q \,\|\, u_t)$ denotes the KL divergence that measures the distance between q and the reference distribution $u_t \in \Delta_{K_t}$, regularizing q toward u_t. u_t encodes prior group importance and defaults to the uniform distribution when no prior information is available, while $\lambda \geq 0$ controls the regularization strength. Thus, minimizing (4) directly reduces the worst-group excess risk, ensures consistently strong performance across all groups and mitigates the disparities induced by (2).

To improve the worst-group performance, we first show that classical rehearsal CL in (1) can perform poorly under objective (3). We make the following assumptions for our theoretical analysis.

Assumption 1 (Boundedness and Lipschitz continuity)**.** *The parameter space $\mathcal{W}$ is bounded, i.e., $\|w\| \leq W,\ \forall w \in \mathcal{W}$. Moreover, the loss is uniformly bounded and Lipschitz continuous in w: for any $w, w' \in \mathcal{W}$ and $z \in \mathcal{Z}$, $|\ell(w,z)| \leq B,\ |\ell(w,z) - \ell(w',z)| \leq L\|w - w'\|$.*

These assumptions are mild and standard in modern nonconvex deep learning theory, where model parameters are constrained to remain within a practically bounded region during training (e.g., (Jacot et al., 2018; Chizat et al., 2019; Oymak & Soltanolkotabi, 2019; Li & Liang, 2018)).

We establish the following excess risk bound $\mathcal{E}_{\mathrm{GRCL}}(w_t)$ for (1) under (3).

Lemma 1. *Suppose w_t is the model obtained by* (1) *with reservoir sampling replay buffer update. Let $q_t(w) := \arg\max_{q\in\Delta_{K_t}} f_t(w,q)$. For each group $k \in [K_t]$, denote by M_t^k the subset of replay buffer samples from group k, and let $m_t^k := |M_t^k|$. Then the excess risk $\mathcal{E}_{\mathrm{GRCL}}(w_t)$ for w_t at each time t is upper bounded as follows:*

$$\begin{aligned}\mathcal{E}_{\mathrm{GRCL}}(w_t) \leq \mathbb{E}\Bigg[&\sum_{k\in[K_t],\, k\neq I_t} 2B\left|q_t^k(w_t) - \frac{m_t^k}{n_t+M}\right| + 2B\left|q_t^{I_t}(w_t) - \frac{n_t + m_t^{I_t}}{n_t+M}\right| \\ &+ \sum_{k\in[K_t],\, k\neq I_t} \frac{2Bm_t^k}{n_t+M}\psi(m_t^k) + 2B\frac{n_t+m_t^{I_t}}{n_t+M}\psi(n_t+m_t^{I_t})\Bigg],\end{aligned} \tag{5}$$

where $\psi(m) := 4\sqrt{\frac{2d\ln(\frac{2LW\sqrt{m}}{B}+1)}{m}} + \frac{2}{\sqrt{m}},\ m > 0$. The above upper bound holds for any group arrival sequence including adversarial arrivals.

The proof is given in Appendix A. The bound in (5) has two main components. The first line on the right-hand side captures the mismatch between the ERM objective in (1) and the population group-robust objective in (3). This term can be large when a difficult group k receives a large adversarial weight $q_t^k(w_t)$ but is underrepresented in $M_t \cup D_t$ with a small $\frac{m_t^k}{M+n_t}$. This term captures the dominant error caused by the gap between the group-robust optimization and frequency-weighted training. The second line captures finite-sample statistical error, where $\psi(m)$ is the complexity term decreasing in the number of stored samples m. Since reservoir sampling retains samples roughly in proportion to arrival frequencies, imbalanced group arrivals can yield heterogeneous m_t^k across groups and amplify this error. We further establish an excess risk lower bound in Lemma 3 (Appendix A), showing that, in the worst case, the excess risk $\mathcal{E}_{\mathrm{GRCL}}(w_t)$ under classical rehearsal CL does not vanish as M grows. This highlights the need for a CL algorithm tailored to GRCL.

3 Rehearsal GRCL with New Model and Replay Buffer Update Design

In this section, we redesign both the model update and the replay buffer allocation to improve the GRCL objective (3). We first propose a group-robust model update and establish its excess-risk upper bound that depends on the replay buffer allocation. Motivated by this bound, we develop two replay buffer allocation strategies: group-balanced replay, which balances storage across groups, and

difficulty-adaptive replay, which further incorporates group-specific learning difficulty. We assume known group identities I_t in this section and address the unknown case in the next section.

3.1 NEW GRCL MODEL UPDATE

First, we replace the standard empirical objective in (1) with a group-robust empirical counterpart of (3). Specifically, define

$$\hat{f}_{D_t\cup M_t}(w,q)=\sum_{k\in[K_t],\,k\neq I_t} q^k \sum_{z\in M_t^k}\frac{\ell(w,z)}{m_t^k}+q^{I_t}\sum_{z\in M_t^{I_t}\cup D_t}\frac{\ell(w,z)}{m_t^{I_t}+n_t}-\lambda D(q\|u_t). \tag{6}$$

At each time t, we update the GRCL model by solving

$$w_t\in\arg\min_{w\in\mathcal{W}}\hat{\phi}_{D_t\cup M_t}(w),\qquad \hat{\phi}_{D_t\cup M_t}(w)=\max_{q\in\Delta_{K_t}}\hat{f}_{D_t\cup M_t}(w,q). \tag{7}$$

We establish the following upper bound for $\mathcal{E}_{\mathrm{GRCL}}(w_t)$ under this GRCL model update.

Proposition 1. *Let $w_t^*\in\arg\min_{w\in\mathcal{W}}\phi_t(w)$ be the optimal model of the group-robust objective in* (3)*. If w_t is obtained from* (7)*, then the excess risk $\mathcal{E}_{\mathrm{GRCL}}(w_t)$ at each time t satisfies*

$$\begin{aligned}\mathcal{E}_{\mathrm{GRCL}}(w_t)\le B\sum_{\substack{k=1\\k\neq I_t}}^{K_t}\mathbb{E}[q_t^k(w_t)+q^k_{D_t\cup M_t}]\,\psi(m_t^k)+B\mathbb{E}[q_t^{I_t}(w_t)+q^{I_t}_{D_t\cup M_t}]\,\psi(m_t^{I_t}+n_t)\\ +2B\sqrt{\frac{2\log K_t}{\min_{k\in[K_t]}\{m_t^k+n_t\mathbf{1}\{k=I_t\}\}}},\end{aligned} \tag{8}$$

where $q_{D_t\cup M_t}\in\arg\max_{q\in\Delta_{K_t}}\hat{f}_{D_t\cup M_t}(w_t^*,q)$ and $q_t(w_t):=\arg\max_{q\in\Delta_{K_t}}f_t(w_t,q)$.

The proof of Proposition 1 is provided in Appendix B. Compared with the classical rehearsal CL bound in (5), the bound in (8) for our GRCL update removes the dominant error term in the first line of (5). In (8), $q_{D_t\cup M_t}$ and $q_t(w_t)$ denote the empirical and population adversarial group weights evaluated at w_t^* and w_t, respectively. Hence, groups that are both difficult, with large $q^k_{D_t\cup M_t}$ and $q_t^k(w_t)$, and rare with large $\psi(m_t^k)$, contribute more to the upper bound in (8). This motivates designing the replay allocation that decides the value of $m_t=(m_t^k)_{k=1}^{K_t}$ at each time t to minimize (8). However, direct optimization over m_t is infeasible because the next task size n_t, group identity I_t, and adversarial weights $q_t(w_t)$ and $q_{D_t\cup M_t}$ are unknown. The weights depend on population quantities such as w_t^* and f_t, while m_t must be decided at the end of time $t-1$, before I_t, n_t are observed.

To handle the first uncertainty on I_t and n_t, we formulate allocation as a robust optimization problem over all possible realizations of I_t and n_t under arbitrary group shifts. We define the set of all admissible replay allocations as $\mathcal{A}_t(M):=\left\{m_t=(m_t^1,\dots,m_t^{K_t})\in\mathbb{Z}_+^{K_t}:\sum_{k=1}^{K_t}m_t^k=M\right\}$, and formulate the following optimization problem:

$$\begin{aligned}&\min_{m_t\in\mathcal{A}_t(M)}\max_{I_t\in[K_t]}U_t^{I_t}(w_t,m_t)\\ :=&\min_{m_t\in\mathcal{A}_t(M)}\max_{I_t\in[K_t]}\left\{B\sum_{\substack{k=1\\k\neq I_t}}^{K_t}\mathbb{E}\big[q_t^k(w_t)+q^k_{D_t\cup M_t}\big]\,\psi(m_t^k)\right.\\ &\left.+B\mathbb{E}\Big[q_t^{I_t}(w_t)+q^{I_t}_{D_t\cup M_t}\Big]\,\psi(m_t^{I_t}+\underline{n})+2B\sqrt{\frac{2\log K_t}{\min_{k\in[K_t]}\{m_t^k+\underline{n}\mathbf{1}\{k=I_t\}\}}}\right\}.\end{aligned} \tag{9}$$

where we replace n_t with $\underline{n}\le\min_t n_t$ to ensure that $\max_{I_t\in[K_t]}U_t^{I_t}(w_t,m_t)$ remains a strict upper bound on the excess risk in (8), since $\psi(m_t^{I_t}+n)$ is decreasing in n. Therefore, minimizing $\max_{I_t\in[K_t]}U_t^{I_t}(w_t,m_t)$ over m_t and allocating the replay buffer M_t accordingly yields a principled replay allocation criterion that controls a worst-case upper bound on $\mathcal{E}_{\mathrm{GRCL}}(w_t)$.

To address the second uncertainty in $q_t(w_t)$ and $q_{D_t\cup M_t}$, we next introduce two computable upper-bound surrogates for (9) that retain control over $\mathcal{E}_{\mathrm{GRCL}}(w_t)$. We minimize these surrogates over m_t to obtain two replay buffer allocation strategies: group-balanced replay and difficulty-adaptive replay.

3.2 New Replay buffer Allocation Design

The first simple upper-bound surrogate for (9) takes the worst case upper bound over all possible $q_t(w_t)$ and $q_{D_t \cup M_t}$:

$$\min_{m_t \in \mathcal{A}_t(M)} \max_{I_t \in [K_t]} \left[\max \left\{ \max_{\substack{k \in [K_t] \\ k \neq I_t}} 2B\psi(m_t^k),\ 2B\psi(m_t^{I_t} + \underline{n}) \right\} + 2B \sqrt{\frac{2 \log K_t}{\min_{k \in [K_t]} \{ m_t^k + \underline{n} \mathbf{1}\{k = I_t\} \}}} \right]. \tag{10}$$

Its optimal solution yields the following group-balanced allocation strategy.

Lemma 2. *An optimal solution to the problem in* (10) *satisfies* $m_t^j \in \{\lfloor M/K_t \rfloor, \lceil M/K_t \rceil\},\ \forall j \in [K_t]$. *Specifically,* $M - K_t \lfloor M/K_t \rfloor$ *groups receive* $\lceil M/K_t \rceil$ *samples, and the remaining groups receive* $\lfloor M/K_t \rfloor$ *samples. The corresponding optimal value of* (10) *is* $2B\psi(\lfloor M/K_t \rfloor) + 2B\sqrt{\frac{2 \log K_t}{\lfloor M/K_t \rfloor}}$.

The proof of Lemma 2 is provided in Appendix B. Although simple and robust, group-balanced replay ignores the adversarial weights $q_t(w_t)$ and $q_{D_t \cup M_t}$, which encode group-specific difficulty. This motivates a second surrogate that incorporates this information.

To preserve the information in $q_t(w_t)$ and $q_{D_t \cup M_t}$, we use the empirical adversarial weighting evaluated at w_t as a computable surrogate for both, obtained during the GRCL update in (7):

$$\hat{q}_{D_t \cup M_t} \in \arg \max_{q \in \Delta_{K_t}} \hat{f}_{D_t \cup M_t}(w_t, q). \tag{11}$$

To further analyze the excess risk upper bound under $\hat{q}_{D_t \cup M_t}$, we require the following assumption.

Assumption 2 (Local quadratic growth)**.** *Let* $w_t^* \in \arg\min_{w \in \mathcal{W}} \phi_t(w)$. *There exists* $\mu > 0$ *such that, for all* $w \in \mathcal{W}$, $\phi_t(w) - \phi_t(w_t^*) \geq \frac{\mu}{2} \|w - w_t^*\|^2$.

This is a standard local growth condition and can hold for some nonconvex models under suitable regularity conditions (Mu & Klabjan, 2025). We then prove that $\mathbb{E}[\hat{q}_{D_t \cup M_t}]$ approximates the theoretical quantities $\mathbb{E}[q_t(w_t)]$ and $\mathbb{E}[q_{D_t \cup M_t}]$ up to an error of order $\max_{k \in [K_t]} O((m_t^k)^{-1/4})$, indicating that the approximation error decreases as the replay buffer size M grows. The formal bounds on $\|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_{D_t \cup M_t}]\|_\infty$ and $\|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_t(w_t)]\|_\infty$ are provided in Lemma 5 in Appendix B.

As such, we aggregate empirical maximizers $\hat{q}_{D_t \cup M_t}$ across time and use the resulting average to approximate both $q_{D_t \cup M_t}$ and $q_t(w_t)$, yielding a reliable surrogate. Using the theoretical result in Lemma 5, we derive the following difficulty-adaptive upper-bound surrogate for (9).

Proposition 2. *Suppose that at time* t, K_t *task groups have arrived. Let* $t_0 < t$ *be the first time such that each group in* $[K_t]$ *has appeared at least once. Define* $\tilde{q}_t := \frac{1}{t - t_0} \sum_{\tau=t_0}^{t-1} \hat{q}_{D_\tau \cup M_\tau}$. *Let* $\varepsilon_\tau := \hat{q}_{D_\tau \cup M_\tau} - \mathbb{E}[\hat{q}_{D_\tau \cup M_\tau}]$ *be the centered empirical fluctuation, and assume* ε_i *satisfies, for every coordinate* $k \in [K_t]$, $\sum_{i=t_0}^{t-1} \sum_{j=t_0}^{t-1} \mathrm{Cov}(\varepsilon_{i,k}, \varepsilon_{j,k}) \leq C(t - t_0)$. *for some constant* $C > 0$ [1]. *Then, using* $\tilde{q}_t$ *as a surrogate for* $q_{D_t \cup M_t}$ *and* $q_t(w_t)$, *with probability at least* $1 - \delta$, *the following problem gives a strict upper-bound surrogate for* (9)*:*

$$\begin{aligned} \min_{m_t \in \mathcal{A}_t(M)} \max_{I_t \in [K_t]} \Bigg\{ & B \sum_{\substack{k=1 \\ k \neq I_t}}^{K_t} \left(2\tilde{q}_t^k + 5h_t + \frac{2}{t - t_0} \sum_{\tau=t_0}^{t-1} h_\tau + 2\sqrt{\frac{CK_t}{(t - t_0)\delta}} \right) \psi(m_t^k) \\ & + B \left(2\tilde{q}_t^{I_t} + 5h_t + \frac{2}{t - t_0} \sum_{\tau=t_0}^{t-1} h_\tau + 2\sqrt{\frac{CK_t}{(t - t_0)\delta}} \right) \psi(m_t^{I_t} + \underline{n}) \\ & + 2B \sqrt{\frac{2 \log K_t}{\min_{k \in [K_t]} \{ m_t^k + \underline{n} \mathbf{1}\{k = I_t\} \}}} \Bigg\}. \end{aligned} \tag{12}$$

where $h_\tau := \frac{B}{2\lambda} \max_{k \in [K_t]} \left\{ \psi(m_\tau^k) + \sqrt{\frac{2 \log K_t}{m_\tau^k}} \right\} + \frac{L}{2\lambda} \sqrt{\frac{4B}{\mu} \max_{k \in [K_t]} \left\{ \psi(m_\tau^k) + \sqrt{\frac{2 \log K_t}{m_\tau^k}} \right\}}$.

[1] The dependence condition on $\sum_{i,j=t_0}^{t-1} \mathrm{Cov}(\varepsilon_i, \varepsilon_j)$ accounts for the fact that $D_\tau \cup M_\tau$ overlaps across time. It is mild in our setting because each estimate also uses a newly arrived dataset D_τ, where $n_\tau \gg M$. As such, the averaged weight $\tilde{q}_t$ concentrates around its expectation, $\frac{1}{t - t_0} \sum_{\tau=t_0}^{t-1} \mathbb{E}[\hat{q}_{D_\tau \cup M_\tau}]$.

Solving (12) for m_t gives our difficulty-adaptive replay buffer allocation. The proof of Proposition 2 is provided in Appendix B. $\psi(\cdot)$ is log-convex, which ensures that the objective in (12) is convex in the continuous relaxation. We can therefore efficiently obtain a near-optimal solution on m_t of (12) in practice by first solving its continuous relaxation with standard numerical methods and then rounding the solution to a feasible integer allocation.

Compared to the difficulty-adaptive replay above, group-balanced replay is simple and stable, achieving the optimal surrogate solution in Lemma 2 without additional assumptions. In contrast, difficulty-adaptive replay requires Assumption 2, but can yield sharper guarantees when it holds. In particular, when M is large, group arrivals are stable and $t - t_0$ grows, h_τ becomes small, making the bound (12) tighter than the optimal surrogate solution in Lemma 2. In that case, optimizing (12) more directly targets $\mathcal{E}_{\mathrm{GRCL}}(w_t)$, allowing difficulty-adaptive replay to improve over group-balanced replay.

However, a gap remains between the proposed allocation and what the replay buffer can actually achieve. Since M_t is formed by updating M_{t-1} using only D_{t-1}, an under-allocated group cannot receive additional samples immediately if D_{t-1} contains no samples from that group. The next section introduces a practical rehearsal-based GRCL algorithm that addresses this gap.

4 Replay Algorithm Design for GRCL with Provable Guarantees

In this section, we translate the design in Section 3 into a practical rehearsal-based algorithm for GRCL to solve two practical challenges: the planned replay allocation cannot be enforced immediately, and the group identity I_t and number of groups K_t may be unknown. We address them with *Allocation-Track Reservoir Sampling*, which tracks the target group allocation, and two group-identification methods with different computation costs: nonparametric BUFFER-MMD rule, and structured EM-MMD rule.

4.1 Rehearsal GRCL Algorithm

We first present our complete rehearsal-based GRCL algorithm with unknown I_t in Alg. 1. At each time t, the learner receives D_t. Alg. 1 estimates the group identity $\hat{I}_t$ of D_t using BUFFER-MMD or EM-MMD. The model w_t is then updated by the GRCL model update in (7). Alg. 1 then computes a target replay allocation $\hat{m}_{t+1}$ for replay buffer at the next time step M_{t+1}, using either the group-balanced allocation in Lemma 2 or solving the difficulty-adaptive allocation in (12). Finally, ALLOC-TRACK RESERVOIR in Alg. 2 updates the buffer so that the actual allocation m_{t+1} tracks target replay allocation $\hat{m}_{t+1}$. The actual replay allocation can gradually align with the target allocation in a stable learning environment. We next introduce our two group identification methods.

Algorithm 1 Rehearsal GRCL

1: **for** $t = 1, \ldots, T$ **do**
2: Receive dataset D_t. Estimate group information:
$(\hat{I}_t, \hat{K}_t) \leftarrow$ BUFFER-MMD(M_t, D_t) or $\hat{I}_t \leftarrow$ EM-MMD$(D_t; \{\hat{\theta}^k_{t-1}\}_{k=1}^{K_T})$.
For EM-MMD, set $\hat{K}_t \leftarrow K_T$.
3: Update w_t on $D_t \cup M_t$ by solving (7) using $\hat{I}_t$.
4: Compute target allocation: Set $(\hat{m}^k_{t+1})_{k\in[\hat{K}_t]}$ by Lemma 2 or solving (12).
5: Update buffer: $M_{t+1} \leftarrow$ ALLOC-TRACK RESERVOIR $\left((\hat{m}^k_{t+1})_{k\in[\hat{K}_t]}, M_t, D_t, \hat{I}_t\right)$ in Alg. 2.
6: **end for**

4.2 BUFFER-MMD Group Identification in Line 2 of Alg. 1

We first introduce the nonparametric identification rule that requires no knowledge of the distribution family and group number. Our idea is to use each group-specific buffer M^k_t as an empirical prototype of group k. Given a new dataset D_t, we compare it with each existing buffer subset M^k_t, $k \in [\hat{K}_{t-1}]$, using maximum mean discrepancy (MMD), a metric for measuring discrepancies between distributions (Gretton et al., 2012). If all MMD distances exceed a predetermined threshold $\gamma/2$, we assign D_t to a new group $\hat{K}_{t-1} + 1$. Otherwise, we assign it to the closest existing group. Let $\mathrm{MMD}(M^k_t, D_t)$ denote the MMD between the two empirical datasets M^k_t and D_t, with the detailed computation provided in Appendix C.2. Define $k^\star_t := \arg\min_{k\in[\hat{K}_{t-1}]} \mathrm{MMD}(M^k_t, D_t)$, the

identification rule is

$$(\hat{I}_t, \hat{K}_t) = \begin{cases} (k_t^\star, \hat{K}_{t-1}), & \text{if } \mathrm{MMD}(M_t^{k_t^\star}, D_t) \le \gamma/2, \\ (\hat{K}_{t-1}+1, \hat{K}_{t-1}+1), & \text{otherwise.} \end{cases} \tag{BUFFER-MMD}$$

We then establish a new theoretical guarantee for this MMD-based rule in replay-based CL that accounts for limited replay buffers and sequential data arrivals. See Proposition 4 in Appendix C.3.

BUFFER-MMD is simple and distribution-free, requiring only $O(\hat{K}_t)$ additional MMD computations at each time t. However, its reliability depends on the quality of group buffers as empirical prototypes. Small m_t^k makes these prototypes M_t^k noisy and increases misclassification risk. Such errors may contaminate buffers and affect future identification. Thus, BUFFER-MMD is most reliable when each group has sufficient replay samples.

4.3 ONLINE EM-MMD GROUP IDENTIFICATION IN LINE 2 OF ALG. 1

We next consider a more structured setting where the number of groups is fixed at K_T over time, and groups arriving at each time t follow a stationary unknown distribution $(\pi_k)_{k\in[K_T]}$. The parametric family of group distributions is known, while each $\mathcal{D}_k$ is parameterized by an unknown θ^k, i.e., $\mathcal{D}_k = \mathcal{D}(\theta^k)$. During CL, we maintain lightweight estimates $(\hat{\theta}_t^k)_{k=1}^{K_T}$ for $(\theta^k)_{k=1}^{K_T}$ and update them using online EM (Cappé & Moulines, 2009) at each time t, which converges to the true parameters $(\theta^k)_{k=1}^{K_T}$ under standard regularity conditions as $t \to \infty$.

However, online EM alone does not directly yield a group-identification guarantee: EM responsibilities provide soft population assignments, but do not by themselves ensure a convergent hard-identification accuracy for the group label of D_t. To obtain a provable identification rule, we combine online EM with MMD as follows, so that the analysis can rely on distributional separation between groups and the estimation error of $(\hat{\theta}_{t-1}^k)_{k\in[K_T]}$.

$$\hat{I}_t = \arg\min_{k\in[K_T]} \mathrm{MMD}\Big(\mathcal{D}(\hat{\theta}_{t-1}^k), D_t\Big). \tag{EM-MMD}$$

This yields the following theoretical guarantee.

Proposition 3. *Under the assumptions stated in Appendix C.4, EM-MMD satisfies*

$$\liminf_{t\to\infty} \mathbb{P}(\hat{I}_t = I_t) \ge 1 - \delta.$$

Moreover, let $\underline{m} \ge 1$ denote a deterministic lower bound on the stabilized replay sizes, so that $\min_{k\in[K_T]} m_t^k \ge \underline{m}$ for all sufficiently large t almost surely. As $t \to \infty$,

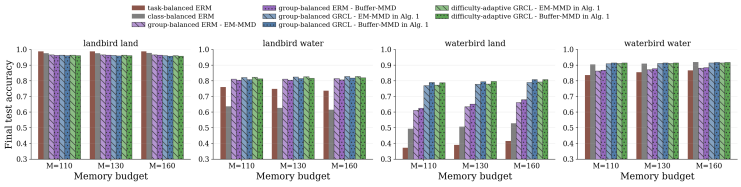
$$\mathcal{E}_{\mathrm{GRCL}}(w_t) \le \widetilde{O}\left(B\sqrt{\frac{d}{\underline{m}}}\right) + O\left(\frac{B\delta}{\min_k \pi_k}\right) + o(1). \tag{13}$$

The complete version and the proof of Proposition 3 are given in Appendix C.4. Proposition 3 establishes stable convergence of (EM-MMD) under the structured-stability condition, introducing only $O(\delta)$ penalty compared to the known group bound in (8). Compared with the Buffer-MMD rule, the EM-MMD rule accommodates weaker distributional separation, as indicated by the separation conditions in (30) and (28), and remains stable under occasional group misidentification.

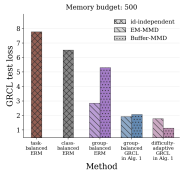
5 EXPERIMENTS ON REAL-WORLD DATASETS FOR GRCL EVALUATIONS

In this section, we evaluate GRCL under realistic continual learning settings using Waterbirds (Wah et al., 2011; Zhou et al., 2017; Sagawa et al., 2019) and CelebA (Liu et al., 2015), two widely used benchmarks for group robustness. We assess performance using both group-robustness and standard continual learning metrics. Implementation details and additional results including an additional classification experiment are provided in Appendix D. We further present experiments on synthetic data in Appendix D.2 to illustrate the learning behavior predicted by our theoretical analysis.

We first evaluate on Waterbirds (Sagawa et al., 2019) with 20 CL tasks. We use a pretrained ResNet-50 backbone (Deng et al., 2009; He et al., 2016), freeze its parameters, and train an additional 512-128 ReLU head. This dataset contains four groups: landbirds/waterbirds on land/water backgrounds. The prediction target is bird type, while the background is spuriously correlated with the label. Thus, landbirds on water backgrounds and waterbirds on land backgrounds become the most difficult groups. Assuming unknown group identities during replay, we use EM-MMD and Buffer-MMD



(a) Waterbirds per-group test accuracy.



(b) CelebA group-robust loss.

Figure 1: Final performance on Waterbirds and CelebA under task-balanced ERM update, class-balanced ERM update, and the methods in Alg. 1. (a): Final per-group test accuracy on Waterbirds. (b): Final group-robust loss on CelebA with replay budget $M = 500$.

Table 1: Performance comparison on the Waterbirds dataset. The best result for each metric on each dataset is shown in bold. ↑ (↓) indicates that higher (lower) values are better.

Method	**Overall Acc.** ↑	**WGA** ↑	**Task F** ↓	**BWT** ↑	**MGF** ↓
Ours: Group-balanced GRCL	88.25 ± 0.45	76.28 ± 0.87	4.44 ± 0.25	3.42 ± 0.48	7.64 ± 0.25
Ours: Difficulty-adaptive GRCL	$\mathbf{88.34 \pm 0.42}$	$\mathbf{76.55 \pm 0.84}$	4.34 ± 0.21	3.52 ± 0.47	7.47 ± 0.22
DER++ (Buzzega et al., NeurIPS 2020)	83.45 ± 0.37	27.08 ± 2.06	6.20 ± 0.32	1.99 ± 0.38	13.75 ± 0.57
MIR (Aljundi et al., NeurIPS 2019)	79.29 ± 0.48	46.79 ± 1.18	5.32 ± 0.17	4.54 ± 0.38	13.91 ± 0.37
LPR (Yoo et al., ICML 2024)	85.61 ± 0.22	28.27 ± 2.00	$\mathbf{1.70 \pm 0.12}$	$\mathbf{8.82 \pm 0.42}$	$\mathbf{4.23 \pm 0.36}$

group identification as in Alg. 1. We compare group-balanced GRCL replay and difficulty-adaptive GRCL replay with task-balanced ERM replay (Rolnick et al., 2019; Deng et al., 2025) and class-balanced ERM replay (Chrysakis & Moens, 2020; Kim et al., 2020). We additionally compare with ERM benchmark enhanced by our group-balanced replay including our group identification methods. Our results show that both identification methods achieve near-perfect group identification as in Proposition 4 and Proposition 3. Consequently, Fig. 5b in Appendix D shows that methods in Alg. 1 reduce the group-robust loss by approximately 50% compared with all benchmarks. Fig. 1a shows that this gain mainly comes from substantially improved accuracy on the two challenging groups: landbirds on water and waterbirds on land.

We then evaluate on CelebA (Liu et al., 2015), a more complex benchmark with eight groups defined by binary hair color, gender, and eyeglasses, with blond hair as the prediction target. We use a more expressive ImageNet-pretrained ResNet-18 (Deng et al., 2009; He et al., 2016) classifier trained end-to-end, which contains 11,177,025 trainable parameters. Fig. 1b shows that difficulty-adaptive replay with Buffer-MMD achieves the lowest loss, 1.121, compared with 7.261 for task-balanced ERM and 5.298 for group-balanced ERM under the same identification method. Figure 7 in the appendix shows the per-group accuracies, indicating that our lowest group-robust loss stems from substantial improvements in the worst-performing groups with blond hair.

We further compare GRCL with state-of-the-art continual learning methods, including DER++ (Buzzega et al., 2020), MIR (Aljundi et al., 2019), and LPR (Yoo et al., 2024). In addition to final worst-group accuracy (WGA), we report four standard CL metrics: average overall accuracy, task forgetting (Task F), backward transfer (BWT), and macro group forgetting (MGF), which assigns equal weight to each unique group. As shown in Table 1, GRCL achieves the best WGA. Importantly, its advantages extend beyond group robustness: GRCL also performs strongly on standard CL metrics that measure average-level performance, achieving the best results on several of them. The complete results, including the comparison on the CelebA dataset, are given in Table 2 in Appendix D.

6 Conclusion

In this paper, we propose GRCL to improve worst-group performance in CL. GRCL includes group-balanced and difficulty-adaptive replay methods, extends to settings with unknown group identities through MMD and EM-based identification, and provides corresponding theoretical guarantees. Experiments show that our GRCL consistently improves worst-group performance.

AI USE STATEMENT

In this work, we used generative AI tools to polish the manuscript, search for related work, and assist with running experiments. We did not use generative AI tools to develop the ideas. The remaining disclosure categories are not applicable to this work. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

To support reproducibility, we provide detailed descriptions of the datasets, data-stream construction, model architectures, training procedures, hyperparameters, and evaluation protocols in Appendix D. The setup and additional results for the synthetic experiments are provided in Appendix D.2. All assumptions and complete proofs of our theoretical results are included in the corresponding appendices. We also provide the source code, configuration files, and instructions required to reproduce our experiments in an anonymous code repository.

REFERENCES

Agil Aghasanli, Yi Li, and Plamen Angelov. Prototype-based continual learning with label-free replay buffer and cluster preservation loss. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 6545–6554, 2025.

Rahaf Aljundi, Min Lin, Baptiste Goujaud, and Yoshua Bengio. Gradient based sample selection for online continual learning. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/e562cd9c0768d5464b64cf61da7fc6bb-Paper.pdf.

Jihwan Bang, Heesu Kim, YoungJoon Yoo, Jung-Woo Ha, and Jonghyun Choi. Rainbow memory: Continual learning with a memory of diverse samples. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8218–8227, 2021.

Eden Belouadah, Adrian Popescu, Umang Aggarwal, and Léo Saci. Active class incremental learning for imbalanced datasets. In *European Conference on Computer Vision*, pp. 146–162. Springer, 2020.

Pietro Buzzega, Matteo Boschini, Angelo Porrello, Davide Abati, and SIMONE CALDERARA. Dark experience for general continual learning: a strong, simple baseline. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 15920–15930. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/b704ea2c39778f07c617f6b7ce480e9e-Paper.pdf.

Alberto Cano and Bartosz Krawczyk. Rose: robust online self-adjusting ensemble for continual learning on imbalanced drifting data streams. *Machine Learning*, 111(7):2561–2599, 2022.

Olivier Cappé and Eric Moulines. On-line expectation–maximization algorithm for latent data models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 71(3):593–613, 2009.

Arslan Chaudhry, Marc'Aurelio Ranzato, Marcus Rohrbach, and Mohamed Elhoseiny. Efficient lifelong learning with a-GEM. In *International Conference on Learning Representations*, 2019.

Lénaïc Chizat, Edouard Oyallon, and Francis Bach. On lazy training in differentiable programming. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Aristotelis Chrysakis and Marie-Francine Moens. Online continual learning from imbalanced data. In *International Conference on Machine Learning*, pp. 1952–1961. PMLR, 2020.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.

Junze Deng, Qinhang Wu, Peizhong Ju, Sen Lin, Yingbin Liang, and Ness Shroff. Unlocking the power of rehearsal in continual learning: A theoretical perspective. In *International Conference on Machine Learning*, pp. 13277–13326. PMLR, 2025.

Farzan Farnia and Asuman Ozdaglar. Train simultaneously, generalize better: Stability of gradient-based minimax learners. In *International Conference on Machine Learning*, pp. 3174–3185. PMLR, 2021.

Saurabh Garg, Mehrdad Farajtabar, Hadi Pouransari, Raviteja Vemulapalli, Sachin Mehta, Oncel Tuzel, Vaishaal Shankar, and Fartash Faghri. Tic-CLIP: Continual training of CLIP models. In *The Twelfth International Conference on Learning Representations*, 2024.

Arthur Gretton, Karsten M Borgwardt, Malte J Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *The journal of machine learning research*, 13(1):723–773, 2012.

Guy Hacohen and Tinne Tuytelaars. Predicting the susceptibility of examples to catastrophic forgetting. In *Forty-second International Conference on Machine Learning*, 2025.

Nika Haghtalab, Michael Jordan, and Eric Zhao. On-demand sampling: Learning optimally from multiple distributions. *Advances in Neural Information Processing Systems*, 35:406–419, 2022.

Jiangpeng He. Gradient reweighting: Towards imbalanced class-incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16668–16677, 2024.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

Arthur Jacot, Franck Gabriel, and Clement Hongler. Neural tangent kernel: Convergence and generalization in neural networks. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Chris Dongjoo Kim, Jinseo Jeong, and Gunhee Kim. Imbalanced continual learning with partitioning reservoir sampling. In *European conference on computer vision*, pp. 411–428. Springer, 2020.

Yunwen Lei, Zhenhuan Yang, Tianbao Yang, and Yiming Ying. Stability and generalization of stochastic gradient methods for minimax problems. In *International Conference on Machine Learning*, pp. 6175–6186. PMLR, 2021.

Yuanzhi Li and Yingyu Liang. Learning overparameterized neural networks via stochastic gradient descent on structured data. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Bing Liu, Sahisnu Mazumder, Eric Robertson, and Scott Grigsby. Ai autonomy: Self-initiated open-world continual learning and adaptation. *AI Magazine*, 44(2):185–199, 2023.

Xialei Liu, Yu-Song Hu, Xu-Sheng Cao, Andrew D Bagdanov, Ke Li, and Ming-Ming Cheng. Long-tailed class incremental learning. In *European Conference on Computer Vision*, pp. 495–512. Springer, 2022.

Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *Proceedings of the IEEE International Conference on Computer Vision*, 2015.

Siqiao Mu and Diego Klabjan. Rewind-to-delete: Certified machine unlearning for nonconvex functions. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Samet Oymak and Mahdi Soltanolkotabi. Overparameterized nonlinear learning: Gradient descent takes the shortest path? In Kamalika Chaudhuri and Ruslan Salakhutdinov (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4951–4960. PMLR, 09–15 Jun 2019.

Qi Qi, Zhishuai Guo, Yi Xu, Rong Jin, and Tianbao Yang. An online method for a class of distributionally robust optimization with non-convex objectives. *Advances in Neural Information Processing Systems*, 34:10067–10080, 2021.

Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H Lampert. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 2001–2010, 2017.

David Rolnick, Arun Ahuja, Jonathan Schwarz, Timothy Lillicrap, and Gregory Wayne. Experience replay for continual learning. *Advances in neural information processing systems*, 32, 2019.

Shiori Sagawa, Pang Wei Koh, Tatsunori B Hashimoto, and Percy Liang. Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. *arXiv preprint arXiv:1911.08731*, 2019.

Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

Ruilin Tong, Yuhang Liu, Javen Qinfeng Shi, and Dong Gong. Coreset selection via reducible loss in continual learning. In *The Thirteenth International Conference on Learning Representations*, 2025.

C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011.

Zhenyi Wang, Li Shen, Le Fang, Qiuling Suo, Tiehang Duan, and Mingchen Gao. Improving task-free continual learning by distributionally robust memory evolution. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 22985–22998. PMLR, 17–23 Jul 2022.

Jinsoo Yoo, Yunpeng Liu, Frank Wood, and Geoff Pleiss. Layerwise proximal replay: A proximal point method for online continual learning. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 57199–57216. PMLR, 21–27 Jul 2024. URL `https://proceedings.mlr.press/v235/yoo24a.html`.

Jaehong Yoon, Divyam Madaan, Eunho Yang, and Sung Ju Hwang. Online coreset selection for rehearsal-based continual learning. In *International Conference on Learning Representations*, 2022.

Junyu Zhang, Mingyi Hong, Mengdi Wang, and Shuzhong Zhang. Generalization bounds for stochastic saddle point problems. In *International Conference on Artificial Intelligence and Statistics*, pp. 568–576. PMLR, 2021.

Lijun Zhang, Peng Zhao, Zhen-Hua Zhuang, Tianbao Yang, and Zhi-Hua Zhou. Stochastic approximation approaches to group distributionally robust optimization. *Advances in Neural Information Processing Systems*, 36:52490–52522, 2023.

Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10 million image database for scene recognition. *IEEE transactions on pattern analysis and machine intelligence*, 40(6):1452–1464, 2017.

A MISSING DETAILS IN SECTION 2

Proof of Lemma 1. Let $p \in \Delta_{K_t}$ denote the empirical group distribution in $D_t \cup M_t$, with $p^k = \frac{m_t^k}{M+n_t}$ for $k \neq I_t$, and $p^{I_t} = \frac{m_t^{I_t}+n_t}{M+n_t}$. Then we have:

$$
\begin{aligned}
&\mathbb{E}[\phi_t(w_t) - \phi_t(w_t^*)] \\
&\leq \mathbb{E}[f_t(w_t, q_t(w_t)) - f_t(w_t^*, q_t(w_t))] \\
&= \mathbb{E}[f_t(w_t, q_t(w_t)) - f_t(w_t, p) + f_t(w_t^*, p) - f_t(w_t^*, q_t(w_t)) + f_t(w_t, p) - f_t(w_t^*, p)] \\
&\leq \mathbb{E}\left[f_t(w_t, q_t(w_t)) - f_t(w_t, p) + f_t(w_t^*, p) - f_t(w_t^*, q_t(w_t)) + 2\sup_w |\hat{f}_{D_t \cup M_t}(w, p) - f_t(w, p)|\right] \\
&\leq 2B\,\mathbb{E}\left[\sum_{k \in [K_t]} |q_t^k(w_t) - p^k|\right] + 2\,\mathbb{E}\left[\sup_w |\hat{f}_{D_t \cup M_t}(w, p) - f_t(w, p)|\right].
\end{aligned}
$$

The first inequality follows because $q_t(w_t^*)$ maximizes $f_t(w_t^*, \cdot)$. The second follows because w_t minimizes (1). The last inequality follows by cancellation of the KL terms and $|L_k(w_t) - L_k(w_t^*)| \leq 2B$. Conditioning on the reservoir group counts, applying the triangle inequality and Lemma 4 to each represented group, and then taking expectations completes the proof.

□

Lemma 3. *There exists a problem instance such that the model w_t obtained by solving* (1) *has the excess risk lower bounded as*

$$\mathcal{E}_{\mathrm{GRCL}}(w_t) \geq \frac{(1-\delta)B}{36}.$$

where $0 < \delta < 1$.

Proof of Lemma 3. Consider an instance with two groups, where the covariates are drawn from truncated Gaussian distributions on $[-2, 2]$: $\mathcal{D}_1 = \mathcal{N}(1, 0)$ and $\mathcal{D}_2 = \mathcal{N}(0, \sigma^2)$, with σ^2 chosen so that the second moment of the truncated distribution is 1. The labels are generated by $y = xw^*$ for $x \sim \mathcal{D}_1$ and by $y = -xw^*$ for $x \sim \mathcal{D}_2$. Suppose that $D(q\|u_t)$ in (4) is the KL divergence and that $u_t = u$ is the uniform distribution. The loss for a sample $z = (x, y)$ is $\ell(w, z) = (xw - y)^2$. We set $\mathcal{W} = [-\sqrt{B}/6, \sqrt{B}/6]$ and $w^* = \sqrt{B}/3$, so that $\ell(w, z) \leq B$ for all $w \in \mathcal{W}$ and all samples in the construction. Then, for $\phi_t(w)$, the adversarial group weights q admit the following closed-form solution:

$$q = \left(\frac{\exp((w-w^*)^2/\lambda)}{\exp((w-w^*)^2/\lambda) + \exp((w+w^*)^2/\lambda)}, \frac{\exp((w+w^*)^2/\lambda)}{\exp((w-w^*)^2/\lambda) + \exp((w+w^*)^2/\lambda)}\right).$$

For $\lambda > 0$, substituting the optimal weights into the regularized objective gives

$$
\begin{aligned}
\phi_t(w) &= \lambda \log\left(\frac{e^{(w-w^*)^2/\lambda} + e^{(w+w^*)^2/\lambda}}{2}\right) \\
&= w^2 + (w^*)^2 + \lambda \log\cosh\left(\frac{2ww^*}{\lambda}\right).
\end{aligned}
$$

This objective is minimized at $w = 0$. Hence

$$\phi_t(w) - \phi_t(0) = w^2 + \lambda \log\cosh\left(\frac{2ww^*}{\lambda}\right) \geq w^2.$$

The same inequality holds for $\lambda = 0$, when $\phi_t(w) = (|w| + w^*)^2$.

Suppose both groups have appeared before time t, and group 1 arrives at time t with $n_t \geq 12M$. Since $x = 1$ in group 1 and $x^2 \leq 4$ in group 2,

$$n_t + m_t^1 \geq 12M \geq 3 \sum_{(x,y) \in M_t^2} x^2.$$

Therefore, the unconstrained empirical minimizer satisfies

$$w^* \frac{n_t + m_t^1 - \sum_{(x,y)\in M_t^2} x^2}{n_t + m_t^1 + \sum_{(x,y)\in M_t^2} x^2} \geq \frac{w^*}{2}.$$

As $\mathcal{W} = [-w^*/2, w^*/2]$, the constrained ERM solution is $w_t = w^*/2$. Consequently,

$$\mathcal{E}_{\mathrm{GRCL}}(w_t) \geq \mathbb{E}[w_t^2] = \frac{(w^*)^2}{4} = \frac{B}{36} \geq \frac{(1-\delta)B}{36},$$

which proves the claimed lower bound. □

Lemma 4. *Let $D = \{z_1, \ldots, z_n\}$ be a dataset of n i.i.d. samples drawn from distribution $\mathcal{D}$. Then*

$$\mathbb{E}_{D\sim\mathcal{D}^n}\left[\sup_{w\in\mathcal{W}}\left|\frac{1}{n}\sum_{z\in D}\ell(w,z) - \mathbb{E}_{z\sim\mathcal{D}}[\ell(w,z)]\right|\right] \leq 4B\sqrt{\frac{2d\ln(\frac{2LW\sqrt{n}}{B}+1)}{n}} + \frac{2B}{\sqrt{n}}.$$

Proof of Lemma 4. Define

$$\hat{L}_D(w) := \frac{1}{n}\sum_{i=1}^{n}\ell(w,z_i), \qquad L(w) := \mathbb{E}_{z\sim\mathcal{D}}[\ell(w,z)].$$

If $2LW\sqrt{n}/B \leq 1/2$, boundedness and Lipschitz continuity give

$$\mathbb{E}_D\Big[\sup_{w\in\mathcal{W}}|\hat{L}_D(w) - L(w)|\Big] \leq \frac{B}{\sqrt{n}} + 4LW \leq \frac{2B}{\sqrt{n}},$$

which proves the claim. Henceforth assume $2LW\sqrt{n}/B > 1/2$.

By the standard symmetrization inequality (Lemma 26.2 in Shalev-Shwartz & Ben-David (2014)),

$$\mathbb{E}_{D\sim\mathcal{D}^n}\Big[\sup_{w\in\mathcal{W}}|\hat{L}_D(w) - L(w)|\Big] \leq 2\,\mathbb{E}_{D\sim\mathcal{D}^n}\big[\mathfrak{R}_D(\ell\circ\mathcal{W})\big], \tag{14}$$

where

$$\mathfrak{R}_D(\ell\circ\mathcal{W}) := \frac{1}{n}\mathbb{E}_\sigma\Big[\sup_{w\in\mathcal{W}}\Big|\sum_{i=1}^{n}\sigma_i\ell(w,z_i)\Big|\Big]$$

is the absolute empirical Rademacher complexity and $\sigma_1, \ldots, \sigma_n$ are i.i.d. Rademacher variables.

Fix a dataset $D = \{z_1, \ldots, z_n\}$ and define the set of loss vectors

$$A_D := \Big\{(\ell(w,z_1), \ldots, \ell(w,z_n)) \in \mathbb{R}^n : \ w \in \mathcal{W}\Big\}.$$

Since $|\ell(w,z_i)| \leq B$ for every i, each $a \in A_D$ satisfies

$$\|a\|_2 \leq \sqrt{n}\,B.$$

Let A'_D be a minimal ρ-cover of A_D in ℓ_2 norm, and let

$$\pi : A_D \to A'_D$$

be a nearest-point projection, so that $\|a - \pi(a)\|_2 \leq \rho$ for all $a \in A_D$. Then

$$\begin{aligned}\mathfrak{R}_D(\ell\circ\mathcal{W}) &= \frac{1}{n}\mathbb{E}_\sigma\Big[\sup_{a\in A_D}|\langle a,\sigma\rangle|\Big] \\ &\leq \frac{1}{n}\mathbb{E}_\sigma\Big[\sup_{a'\in A'_D}|\langle a',\sigma\rangle|\Big] + \frac{1}{n}\mathbb{E}_\sigma\Big[\sup_{a\in A_D}|\langle a-\pi(a),\sigma\rangle|\Big].\end{aligned}$$

For the second term, by Cauchy–Schwarz,

$$\sup_{a\in A_D}|\langle a-\pi(a),\sigma\rangle| \leq \sup_{a\in A_D}\|a-\pi(a)\|_2\,\|\sigma\|_2 \leq \rho\sqrt{n}.$$

Hence

$$\mathfrak{R}_D(\ell\circ\mathcal{W}) \leq \mathfrak{R}_D(A'_D) + \frac{\rho}{\sqrt{n}}.$$

Since A'_D is finite and every $a' \in A'_D$ satisfies $\|a'\|_2 \leq \sqrt{n}\,B$, Massart's lemma (Lemma 26.8 in Shalev-Shwartz & Ben-David (2014)) gives

$$\mathfrak{R}_D(A'_D) \leq \frac{\sqrt{2\ln(2|A'_D|)}}{n}\max_{a'\in A'_D}\|a'\|_2 \leq B\sqrt{\frac{2\ln(2N(\rho,A_D))}{n}}.$$

Therefore,

$$\mathfrak{R}_D(\ell \circ \mathcal{W}) \le B\sqrt{\frac{2\ln(2N(\rho, A_D))}{n}} + \frac{\rho}{\sqrt{n}}. \tag{15}$$

Next we bound the covering number through the parameter space. For any $w, w' \in \mathcal{W}$, since $\ell(\cdot, z)$ is L-Lipschitz for every z,

$$\|(\ell(w, z_1), \dots, \ell(w, z_n)) - (\ell(w', z_1), \dots, \ell(w', z_n))\|_2 \le L\sqrt{n}\,\|w - w'\|_2.$$

Hence

$$N(\rho, A_D) \le \mathcal{N}\left(\frac{\rho}{L\sqrt{n}}, \mathcal{W}, \|\cdot\|_2\right).$$

Since $\mathcal{W}$ is contained in the Euclidean ball of radius W in $\mathbb{R}^d$, the standard volumetric bound implies

$$\mathcal{N}\left(\frac{\rho}{L\sqrt{n}}, \mathcal{W}, \|\cdot\|_2\right) \le \left(1 + \frac{2LW\sqrt{n}}{\rho}\right)^d.$$

Substituting this into (15), we obtain

$$\mathfrak{R}_D(\ell \circ \mathcal{W}) \le B\sqrt{\frac{2\ln 2 + 2d\ln\left(1 + \frac{2LW\sqrt{n}}{\rho}\right)}{n}} + \frac{\rho}{\sqrt{n}}.$$

Since $2LW\sqrt{n}/B > 1/2$ and $d \ge 1$,

$$\ln 2 \le 3d\ln\left(1 + \frac{2LW\sqrt{n}}{B}\right).$$

Choosing $\rho = B$ yields

$$\mathfrak{R}_D(\ell \circ \mathcal{W}) \le 2B\sqrt{\frac{2d\ln(\frac{2LW\sqrt{n}}{B} + 1)}{n}} + \frac{B}{\sqrt{n}}.$$

Finally, applying (14),

$$\mathbb{E}_{D\sim\mathcal{D}^n}\Big[\sup_{w\in\mathcal{W}} |\hat{L}_D(w) - L(w)|\Big] \le 4B\sqrt{\frac{2d\ln(\frac{2LW\sqrt{n}}{B} + 1)}{n}} + \frac{2B}{\sqrt{n}}.$$

This completes the proof. □

B Missing details in Section 3

Proof of Proposition 1. Define $\hat{L}_k^t(w)$ as the average empirical loss over the samples from group k in $D_t \cup M_t$. The arrival sequence and replay counts are fixed, and each group's effective sample size $m_t^k + n_t\mathbf{1}\{k = I_t\}$ is positive. By empirical optimality and the definitions of the adversarial weights,

$$\begin{aligned}
&\mathbb{E}[\phi_t(w_t) - \phi_t(w_t^*)] \\
&\le \mathbb{E}\Big[\phi_t(w_t) - \hat{\phi}_{D_t\cup M_t}(w_t) + \hat{\phi}_{D_t\cup M_t}(w_t^*) - \phi_t(w_t^*)\Big] \\
&= \mathbb{E}\Big[\max_q f_t(w_t, q) - \max_{q_1} \hat{f}_{D_t\cup M_t}(w_t, q_1) + \max_{q_2} \hat{f}_{D_t\cup M_t}(w_t^*, q_2) - \max_{q_3} f_t(w_t^*, q_3)\Big] \\
&\le \mathbb{E}\Big[f_t(w_t, q_t(w_t)) - \hat{f}_{D_t\cup M_t}(w_t, q_t(w_t)) + \hat{f}_{D_t\cup M_t}(w_t^*, q_{D_t\cup M_t}) - f_t(w_t^*, q_{D_t\cup M_t})\Big] \\
&\le \mathbb{E}\Big[\sum_{k=1}^{K_t}\big(q_t^k(w_t) + q_{D_t\cup M_t}^k\big)\sup_w |L_k(w) - \hat{L}_k^t(w)|\Big].
\end{aligned}$$

Since the weights are nonnegative and $\sum_k (q_t^k(w_t) + q_{D_t\cup M_t}^k) = 2$, centering each group's uniform deviation gives

$$\begin{aligned}
\mathcal{E}_{\mathrm{GRCL}}(w_t) \le &\sum_{k=1}^{K_t} \mathbb{E}[q_t^k(w_t) + q_{D_t\cup M_t}^k]\,\mathbb{E}\Big[\sup_w |L_k(w) - \hat{L}_k^t(w)|\Big] \\
&+ 2\mathbb{E}\max_{k\in[K_t]}\Big\{\sup_w |L_k(w) - \hat{L}_k^t(w)| - \mathbb{E}\Big[\sup_w |L_k(w) - \hat{L}_k^t(w)|\Big]\Big\}.
\end{aligned}$$

By Lemma 4,

$$\mathbb{E}\Big[\sup_w |L_k(w) - \hat{L}_k^t(w)|\Big] \le B\psi\big(m_t^k + n_t \mathbf{1}\{k = I_t\}\big).$$

It remains to bound the centered maximum. Replacing one observation from group k changes its uniform deviation by at most $2B/(m_t^k + n_t\mathbf{1}\{k = I_t\})$. The bounded-differences exponential-moment inequality therefore gives, for every $a > 0$,

$$\mathbb{E}\exp\left(a\Big\{\sup_w |L_k(w) - \hat{L}_k^t(w)| - \mathbb{E}\Big[\sup_w |L_k(w) - \hat{L}_k^t(w)|\Big]\Big\}\right) \le \exp\left(\frac{a^2B^2}{2(m_t^k + n_t\mathbf{1}\{k = I_t\})}\right).$$

Using the log-sum-exp bound and Jensen's inequality,

$$\mathbb{E}\max_{k\in[K_t]}\Big\{\sup_w |L_k(w) - \hat{L}_k^t(w)| - \mathbb{E}\Big[\sup_w |L_k(w) - \hat{L}_k^t(w)|\Big]\Big\} \le \frac{\log K_t}{a} + \frac{aB^2}{2\min_{k\in[K_t]}\{m_t^k + n_t\mathbf{1}\{k = I_t\}\}}.$$

For $K_t > 1$, minimizing over $a > 0$ bounds this expression by

$$B\sqrt{\frac{2\log K_t}{\min_{k\in[K_t]}\{m_t^k + n_t\mathbf{1}\{k = I_t\}\}}}.$$

For $K_t = 1$, the centered expectation is zero, so the same bound holds. Combining the preceding inequalities yields

$$\begin{aligned}\mathcal{E}_{\mathrm{GRCL}}(w_t) \le{}& B\sum_{\substack{k=1\\k\neq I_t}}^{K_t}\mathbb{E}[q_t^k(w_t) + q_{D_t\cup M_t}^k]\,\psi(m_t^k)\\ &+ B\mathbb{E}[q_t^{I_t}(w_t) + q_{D_t\cup M_t}^{I_t}]\,\psi(m_t^{I_t} + n_t)\\ &+ 2B\sqrt{\frac{2\log K_t}{\min_{k\in[K_t]}\{m_t^k + n_t\mathbf{1}\{k = I_t\}\}}}. \end{aligned} \tag{16}$$

This completes the proof. □

Proof of Lemma 2. Define

$$\psi(x) = 4\sqrt{\frac{2d\ln(\frac{2LW\sqrt{x}}{B} + 1)}{x}} + \frac{2}{\sqrt{x}}, \qquad x > 0.$$

For simplicity, write $m_k = m_t^k$.

We first show that ψ is strictly decreasing on $(0,\infty)$. Since $x \mapsto x^{-1/2}$ is strictly decreasing, it remains to show that

$$x \mapsto \frac{\ln(\frac{2LW\sqrt{x}}{B} + 1)}{x}$$

is decreasing. Let $a = 2LW/B > 0$, and define

$$h(x) = \frac{\ln(1 + a\sqrt{x})}{x}.$$

Set $y = \sqrt{x}$. Then

$$h(x) = \frac{\ln(1 + ay)}{y^2}.$$

Differentiating with respect to y, we get

$$\frac{d}{dy}\left(\frac{\ln(1 + ay)}{y^2}\right) = \frac{\frac{ay}{1+ay} - 2\ln(1 + ay)}{y^3}.$$

Let $u = ay > 0$. It is enough to show that $\frac{u}{1+u} - 2\ln(1 + u) < 0$. Equivalently, define

$$q(u) = 2\ln(1 + u) - \frac{u}{1 + u}.$$

Then $q(0) = 0$, and

$$q'(u) = \frac{2}{1+u} - \frac{1}{(1+u)^2} = \frac{1+2u}{(1+u)^2} > 0$$

for all $u > 0$. Therefore $q(u) > 0$ for all $u > 0$. Hence

$$\frac{u}{1+u} - 2\ln(1+u) < 0,$$

which implies that $h(x)$ is strictly decreasing. Therefore $\psi(x)$ is strictly decreasing on $(0, \infty)$.

For $K_t \geq 2$, the arriving group can be chosen different from a group with minimum replay size. Therefore, the worst-case objective equals

$$2B\psi\left(\min_k m_t^k\right) + 2B\sqrt{\frac{2\log K_t}{\min_k m_t^k}}.$$

Both terms decrease with $\min_k m_t^k$. Since $\min_k m_t^k \leq \lfloor M/K_t \rfloor$, the stated balanced allocation attains the optimum.

□

Lemma 5. *Let $\hat{q}_{D_t \cup M_t}$, $q_{D_t \cup M_t}$, and $q_t(w_t)$ be defined in* (11)*, Proposition 1, and Lemma 1, respectively. Then,*

$$\begin{aligned}
&\|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_{D_t \cup M_t}]\|_\infty \\
\leq &\frac{1}{2\lambda}\left(2B \max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\} + L\sqrt{\frac{4B}{\mu}\max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\}}\right), \quad (17) \\
&\|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_t(w_t)]\|_\infty \\
\leq &\frac{1}{2\lambda}\left(B \max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\} + 2L\sqrt{\frac{4B}{\mu}\max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\}}\right). \quad (18)
\end{aligned}$$

Proof of Lemma 5. First, by Assumptions 1 and 2, for any $k^* \in \arg\max_{k\in[K_t]} |L_k(w_t) - L_k(w_t^*)|$, we have

$$\|L(w_t) - L(w_t^*)\|_\infty = |L_{k^*}(w_t) - L_{k^*}(w_t^*)| \leq L\|w_t - w_t^*\| \leq L\sqrt{\frac{2}{\mu}\big(\phi_t(w_t) - \phi_t(w_t^*)\big)}. \quad (19)$$

Let $q_t^* = \arg\max_q f_t(w_t^*, q)$. We next upper bound $\left\|\mathbb{E}\big[q_t(w_t) - q_t^*\big]\right\|_\infty$, $\left\|\mathbb{E}\big[q_{D_t\cup M_t} - q_t^*\big]\right\|_\infty$, and $\left\|\mathbb{E}\big[\hat{q}_{D_t\cup M_t} - q_t^*\big]\right\|_\infty$, respectively.

Since $D(\cdot\|\cdot)$ is the KL divergence, the maximizers admit the following closed-form expressions:

$$\begin{aligned}
q_{D_t\cup M_t} &= \left(\frac{\exp(\hat{L}_1^t(w_t^*)/\lambda)}{\sum_{j\in[K_t]}\exp(\hat{L}_j^t(w_t^*)/\lambda)}, \ldots, \frac{\exp(\hat{L}_{K_t}^t(w_t^*)/\lambda)}{\sum_{j\in[K_t]}\exp(\hat{L}_j^t(w_t^*)/\lambda)}\right), \\
q_t(w_t) &= \left(\frac{\exp(L_1(w_t)/\lambda)}{\sum_{j\in[K_t]}\exp(L_j(w_t)/\lambda)}, \ldots, \frac{\exp(L_{K_t}(w_t)/\lambda)}{\sum_{j\in[K_t]}\exp(L_j(w_t)/\lambda)}\right), \\
q_t^* &= \left(\frac{\exp(L_1(w_t^*)/\lambda)}{\sum_{j\in[K_t]}\exp(L_j(w_t^*)/\lambda)}, \ldots, \frac{\exp(L_{K_t}(w_t^*)/\lambda)}{\sum_{j\in[K_t]}\exp(L_j(w_t^*)/\lambda)}\right), \\
\hat{q}_{D_t\cup M_t} &= \left(\frac{\exp(\hat{L}_1^t(w_t)/\lambda)}{\sum_{j\in[K_t]}\exp(\hat{L}_j^t(w_t)/\lambda)}, \ldots, \frac{\exp(\hat{L}_{K_t}^t(w_t)/\lambda)}{\sum_{j\in[K_t]}\exp(\hat{L}_j^t(w_t)/\lambda)}\right). \quad (20)
\end{aligned}$$

The softmax mapping is $1/(2\lambda)$-Lipschitz in the infinity norm. By Lemma 4 and the centered-maximum bound established in the proof of Proposition 1,

$$\mathbb{E}\left[\max_{k\in[K_t]}\sup_w |\hat{L}_k^t(w) - L_k(w)|\right] \leq B\max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\}.$$

Moreover, since the two adversarial weight vectors each sum to one, (16) implies

$$\mathbb{E}[\phi_t(w_t) - \phi_t(w_t^*)] \le 2B \max_{k\in[K_t]} \left\{ \psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}} \right\}.$$

Together with (19) and Jensen's inequality, we obtain

$$\begin{aligned}
\|\mathbb{E}[q_t(w_t) - q_t^*]\|_\infty &\le \frac{1}{2\lambda}\mathbb{E}\left[\|L(w_t) - L(w_t^*)\|_\infty\right] \\
&\le \frac{L}{2\lambda}\sqrt{\frac{2}{\mu}\mathbb{E}[\phi_t(w_t) - \phi_t(w_t^*)]} \\
&\le \frac{L}{2\lambda}\sqrt{\frac{4B}{\mu}\max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\}}. \qquad (21)
\end{aligned}$$

Next, by the Lipschitz property of the softmax mapping,

$$\begin{aligned}
\|\mathbb{E}[q_{D_t\cup M_t} - q_t^*]\|_\infty &\le \frac{1}{2\lambda}\mathbb{E}\left[\|\hat{L}^t(w_t^*) - L(w_t^*)\|_\infty\right] \\
&\le \frac{1}{2\lambda}\mathbb{E}\left[\max_{k\in[K_t]}\sup_w |\hat{L}_k^t(w) - L_k(w)|\right] \\
&\le \frac{B}{2\lambda}\max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\}. \qquad (22)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|\mathbb{E}[\hat{q}_{D_t\cup M_t} - q_t^*]\|_\infty &\le \frac{1}{2\lambda}\mathbb{E}\left[\|\hat{L}^t(w_t) - L(w_t^*)\|_\infty\right] \\
&\le \frac{1}{2\lambda}\Big(\mathbb{E}[\|\hat{L}^t(w_t) - L(w_t)\|_\infty] + \mathbb{E}[\|L(w_t) - L(w_t^*)\|_\infty]\Big) \\
&\le \frac{1}{2\lambda}\Bigg(B\max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\} \\
&\qquad + L\sqrt{\frac{4B}{\mu}\max_{k\in[K_t]}\left\{\psi(m_t^k) + \sqrt{\frac{2\log K_t}{m_t^k}}\right\}}\Bigg). \qquad (23)
\end{aligned}$$

Finally, by the triangle inequality,

$$\begin{aligned}
\|\mathbb{E}\left[\hat{q}_{D_t\cup M_t} - q_{D_t\cup M_t}\right]\|_\infty &\le \|\mathbb{E}\left[\hat{q}_{D_t\cup M_t} - q_t^*\right]\|_\infty + \|\mathbb{E}\left[q_t^* - q_{D_t\cup M_t}\right]\|_\infty, \\
\|\mathbb{E}\left[\hat{q}_{D_t\cup M_t} - q_t(w_t)\right]\|_\infty &\le \|\mathbb{E}\left[\hat{q}_{D_t\cup M_t} - q_t^*\right]\|_\infty + \|\mathbb{E}\left[q_t^* - q_t(w_t)\right]\|_\infty. \qquad (24)
\end{aligned}$$

Combining the bounds in (21), (22), and (23) completes the proof.

□

Proof of Proposition 2. By Lemma 6, with probability at least $1-\delta$, we have

$$\begin{aligned}
\left\|\tilde{q}_t - \frac{1}{t-t_0}\sum_{\tau=t_0}^{t-1}\mathbb{E}[\hat{q}_{D_\tau\cup M_\tau}]\right\|_\infty &= \left\|\frac{1}{t-t_0}\sum_{\tau=t_0}^{t-1}\varepsilon_\tau\right\|_\infty \\
&\le \sqrt{\frac{CK_t}{(t-t_0)\delta}}.
\end{aligned}$$

Since the group set and the uniform reference distribution remain unchanged from t_0 to t, the population objective and its optimal adversarial weights q_t^* also remain unchanged. By Lemma 5, we

have

$$
\begin{aligned}
&\|\tilde{q}_t - \mathbb{E}[q_{D_t \cup M_t}]\|_\infty \\
&\le \left\| \tilde{q}_t - \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} \mathbb{E}[\hat{q}_{D_\tau \cup M_\tau}] \right\|_\infty \\
&\quad + \left\| \mathbb{E}\left[\hat{q}_{D_t \cup M_t} - \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} \hat{q}_{D_\tau \cup M_\tau} \right] \right\|_\infty + \|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_{D_t \cup M_t}]\|_\infty \\
&\le \left\| \tilde{q}_t - \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} \mathbb{E}[\hat{q}_{D_\tau \cup M_\tau}] \right\|_\infty + \|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_t^*]\|_\infty \\
&\quad + \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} \|\mathbb{E}[q_t^* - \hat{q}_{D_\tau \cup M_\tau}]\|_\infty + \|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_{D_t \cup M_t}]\|_\infty \\
&\le \sqrt{\frac{CK_t}{(t-t_0)\delta}} + h_t + \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} h_\tau \\
&\quad + \frac{1}{2\lambda} \left(2B \max_{k \in [K_t]} \left\{ \psi(m_t^k) + \sqrt{\frac{2 \log K_t}{m_t^k}} \right\} + L \sqrt{\frac{4B}{\mu} \max_{k \in [K_t]} \left\{ \psi(m_t^k) + \sqrt{\frac{2 \log K_t}{m_t^k}} \right\}} \right).
\end{aligned}
$$

The last inequality follows from (23), Lemma 5, and the definition of h_τ. Similarly,

$$
\begin{aligned}
&\|\tilde{q}_t - \mathbb{E}[q_t(w_t)]\|_\infty \\
&\le \left\| \tilde{q}_t - \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} \mathbb{E}[\hat{q}_{D_\tau \cup M_\tau}] \right\|_\infty \\
&\quad + \left\| \mathbb{E}\left[\hat{q}_{D_t \cup M_t} - \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} \hat{q}_{D_\tau \cup M_\tau} \right] \right\|_\infty + \|\mathbb{E}[\hat{q}_{D_t \cup M_t} - q_t(w_t)]\|_\infty \\
&\le \sqrt{\frac{CK_t}{(t-t_0)\delta}} + h_t + \frac{1}{t-t_0} \sum_{\tau=t_0}^{t-1} h_\tau \\
&\quad + \frac{1}{2\lambda} \left(B \max_{k \in [K_t]} \left\{ \psi(m_t^k) + \sqrt{\frac{2 \log K_t}{m_t^k}} \right\} + 2L \sqrt{\frac{4B}{\mu} \max_{k \in [K_t]} \left\{ \psi(m_t^k) + \sqrt{\frac{2 \log K_t}{m_t^k}} \right\}} \right).
\end{aligned}
$$

Therefore, for every $k \in [K_t]$,

$$
\begin{aligned}
\mathbb{E}[q_t^k(w_t) + q^k_{D_t \cup M_t}] - 2\tilde{q}_t^k &\le \|\tilde{q}_t - \mathbb{E}[q_t(w_t)]\|_\infty + \|\tilde{q}_t - \mathbb{E}[q_{D_t \cup M_t}]\|_\infty \\
&\le 5h_t + \frac{2}{t-t_0} \sum_{\tau=t_0}^{t-1} h_\tau + 2\sqrt{\frac{CK_t}{(t-t_0)\delta}}.
\end{aligned}
$$

Substituting this bound into (9) proves the result. □

Lemma 6. *Let $\xi_1, \dots, \xi_n$ be random vectors in $\mathbb{R}^d$ such that*

$$\mathbb{E}[\xi_i] = 0, \qquad i = 1, \dots, n.$$

Assume that for every coordinate $k = 1, \dots, d$,

$$\sum_{i=1}^n \sum_{j=1}^n \mathrm{Cov}(\xi_{i,k}, \xi_{j,k}) \le Cn$$

for some constant $C > 0$. Then

$$\mathbb{E}\left[\left\| \frac{1}{n} \sum_{i=1}^n \xi_i \right\|_\infty^2 \right] \le \frac{Cd}{n}.$$

Consequently, for any $\delta \in (0,1)$, with probability at least $1-\delta$,

$$\left\| \frac{1}{n} \sum_{i=1}^{n} \xi_i \right\|_\infty \leq \sqrt{\frac{Cd}{n\delta}}.$$

Proof of Lemma 6. Define $\bar{\xi}_n := \frac{1}{n}\sum_{i=1}^n \xi_i \in \mathbb{R}^d$. For each coordinate $k = 1, \ldots, d$, write $\bar{\xi}_{n,k} = \frac{1}{n}\sum_{i=1}^n \xi_{i,k}$. Since $\mathbb{E}[\xi_i] = 0$, we have

$$\mathbb{E}[\xi_{i,k}] = 0 \qquad \text{for all } i = 1, \ldots, n,\ k = 1, \ldots, d.$$

Hence $\mathbb{E}[\bar{\xi}_{n,k}] = 0$. Therefore, $\mathbb{E}[\bar{\xi}_{n,k}^2] = \mathrm{Var}(\bar{\xi}_{n,k})$. By bilinearity of covariance,

$$\begin{aligned}
\mathrm{Var}(\bar{\xi}_{n,k}) &= \mathrm{Var}\left(\frac{1}{n} \sum_{i=1}^{n} \xi_{i,k} \right) \\
&= \frac{1}{n^2} \mathrm{Var}\left(\sum_{i=1}^{n} \xi_{i,k} \right) \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}(\xi_{i,k}, \xi_{j,k}).
\end{aligned}$$

Using the assumed coordinate-wise covariance bound, we obtain

$$\mathbb{E}[\bar{\xi}_{n,k}^2] \leq \frac{Cn}{n^2} = \frac{C}{n}.$$

This holds for every $k = 1, \ldots, d$.

Now observe that

$$\|\bar{\xi}_n\|_\infty^2 = \max_{1 \leq k \leq d} \bar{\xi}_{n,k}^2 \leq \sum_{k=1}^{d} \bar{\xi}_{n,k}^2.$$

Taking expectations gives

$$\mathbb{E}\left[\|\bar{\xi}_n\|_\infty^2 \right] \leq \sum_{k=1}^{d} \mathbb{E}[\bar{\xi}_{n,k}^2] \leq \sum_{k=1}^{d} \frac{C}{n} = \frac{Cd}{n}.$$

This proves the expectation bound.

Next, by Markov's inequality applied to the nonnegative random variable $\|\bar{\xi}_n\|_\infty^2$, for any $r > 0$,

$$\mathbb{P}\left(\|\bar{\xi}_n\|_\infty \geq r \right) = \mathbb{P}\left(\|\bar{\xi}_n\|_\infty^2 \geq r^2 \right) \leq \frac{\mathbb{E}\left[\|\bar{\xi}_n\|_\infty^2 \right]}{r^2}.$$

Using the expectation bound above,

$$\mathbb{P}\left(\|\bar{\xi}_n\|_\infty \geq r \right) \leq \frac{Cd}{nr^2}.$$

Set

$$r = \sqrt{\frac{Cd}{n\delta}}.$$

Then

$$\mathbb{P}\left(\|\bar{\xi}_n\|_\infty \geq \sqrt{\frac{Cd}{n\delta}} \right) \leq \delta.$$

Equivalently, with probability at least $1-\delta$,

$$\left\| \frac{1}{n} \sum_{i=1}^{n} \xi_i \right\|_\infty \leq \sqrt{\frac{Cd}{n\delta}}.$$

This completes the proof. □

C MISSING DETAILS IN SECTION 4

C.1 ALLOCATION-TRACK RESERVOIR SAMPLING ALGORITHM

Algorithm 2 ALLOC-TRACK RESERVOIR

Require: replay buffer allocation plan $(\hat{m}_{t+1}^k)_{k\in[\hat{K}_t]}$, replay buffer M_t, current dataset $D_t = \{z_i\}_{i=1}^{n_t}$, estimated group $\hat{I}_t$, group counts $\{G_k\}_{k\in[\hat{K}_t]}$

Ensure: Updated replay buffer M_{t+1} and allocation $\{m_{t+1}^k\}_{k\in[\hat{K}_t]}$

$M_{t+1} \leftarrow M_t$
$m_{t+1}^k \leftarrow |M_t^k|$ for each $k \in [\hat{K}_t]$
for $i = 1, \dots, n_t$ **do**
  $G_{\hat{I}_t} \leftarrow G_{\hat{I}_t} + 1$
  if $|M_{t+1}| < M$ **then**
    Add z_i to $M_{t+1}^{\hat{I}_t}$.
    $m_{t+1}^{\hat{I}_t} \leftarrow m_{t+1}^{\hat{I}_t} + 1$.
  else if $m_{t+1}^{\hat{I}_t} < \hat{m}_{t+1}^{\hat{I}_t}$ **then**
    Find an over-allocated group

$$k' \leftarrow \arg\max_{k\in[\hat{K}_t]} \left(m_{t+1}^k - \hat{m}_{t+1}^k\right).$$

    Remove a uniformly selected sample from $M_{t+1}^{k'}$ and add z_i to $M_{t+1}^{\hat{I}_t}$.
    $m_{t+1}^{k'} \leftarrow m_{t+1}^{k'} - 1$.
    $m_{t+1}^{\hat{I}_t} \leftarrow m_{t+1}^{\hat{I}_t} + 1$.
  else
    Sample $v \sim \mathrm{Unif}(0, 1)$.
    if $v \le m_{t+1}^{\hat{I}_t}/G_{\hat{I}_t}$ **then**
      Randomly select a sample from $M_{t+1}^{\hat{I}_t}$ and replace it with z_i.
    end if
  end if
end for
return updated replay buffer M_{t+1}

C.2 DETAILS WITH MMD

MMD is defined with respect to a positive definite kernel $\kappa : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}$. With a slight abuse of notation, we write $\mathrm{MMD}(X, Y)$ for the MMD between X and Y, where X and Y may denote either distributions or empirical datasets, and write $\mathrm{MMD}^2(X, Y)$ for the squared MMD.

When X and Y are both distributions, the population MMD is

$$\mathrm{MMD}^2(X, Y) = \mathbb{E}_{x,x'\sim X}[\kappa(x, x')] + \mathbb{E}_{y,y'\sim Y}[\kappa(y, y')] - 2\mathbb{E}_{x\sim X,\, y\sim Y}[\kappa(x, y)]. \tag{25}$$

When X and Y are both empirical datasets, we use the biased empirical estimator

$$\mathrm{MMD}^2(X, Y) = \frac{1}{|X|^2}\sum_{x,x'\in X}\kappa(x, x') + \frac{1}{|Y|^2}\sum_{y,y'\in Y}\kappa(y, y') - \frac{2}{|X||Y|}\sum_{x\in X}\sum_{y\in Y}\kappa(x, y). \tag{26}$$

When X is a distribution and Y is an empirical dataset, their MMD is the distance between their RKHS mean embeddings:

$$\mathrm{MMD}^2(X,Y) = \mathbb{E}_{x,x'\sim X}\kappa(x,x') + \frac{1}{|Y|^2}\sum_{y,y'\in Y}\kappa(y,y') - \frac{2}{|Y|}\sum_{y\in Y}\mathbb{E}_{x\sim X}\kappa(x,y). \tag{27}$$

C.3 Theoretical Results for Buffer-MMD

Proposition 4. *Suppose that replay buffers constructed using true group identities have prescribed sizes $m_\tau^k \geq 1$ for every previously observed group $k \in [K_{\tau-1}]$. Each such buffer consists of i.i.d. samples from $\mathcal{D}_k$. Let κ be a fixed positive definite kernel satisfying $0 \leq \kappa(x,y) \leq F$, and use the biased empirical MMD.*

Suppose that any two distinct group distributions satisfy

$$\mathrm{MMD}(\mathcal{D}_i, \mathcal{D}_j) > \gamma, \qquad i \neq j,$$

where, for some $\varepsilon > 0$,

$$\gamma > 4\sqrt{\frac{F}{m_\tau^k}} + 4\sqrt{\frac{F}{n_\tau}} + \varepsilon, \qquad 2 \leq \tau \leq T, \quad k \in [K_{\tau-1}]. \tag{28}$$

Then (Buffer-MMD)*, with the first task initialized as group 1, correctly identifies all tasks up to time t with probability at least*

$$1 - 2\sum_{\tau=2}^{t}\sum_{k\in[K_{\tau-1}]}\exp\left(-\frac{\varepsilon^2 m_\tau^k n_\tau}{8F(m_\tau^k + n_\tau)}\right).$$

Let w_t^{or} denote the model obtained by Algorithm 1 using true group identities. Under Assumption 1,

$$\mathcal{E}_{\mathrm{GRCL}}(w_t) \leq \mathcal{E}_{\mathrm{GRCL}}(w_t^{\mathrm{or}}) + 4B\sum_{\tau=2}^{t}\sum_{k\in[K_{\tau-1}]}\exp\left(-\frac{\varepsilon^2 m_\tau^k n_\tau}{8F(m_\tau^k + n_\tau)}\right). \tag{29}$$

With exact empirical model updates, the oracle term satisfies the bound in Proposition 1.

Proof of Proposition 4. Consider the oracle execution of Algorithm 1, which uses true group identities, with the same initialization, random coins, and tie-breaking as the Buffer-MMD execution. Let M_τ^k denote its buffer for group k.

By Theorem 7 of Gretton et al. (2012), for each τ and each previously observed group $k \in [K_{\tau-1}]$,

$$\mathbb{P}\left(\left|\mathrm{MMD}(D_\tau, M_\tau^k) - \mathrm{MMD}(\mathcal{D}_{I_\tau}, \mathcal{D}_k)\right| > 2\sqrt{\frac{F}{m_\tau^k}} + 2\sqrt{\frac{F}{n_\tau}} + \frac{\varepsilon}{2}\right) \leq 2\exp\left(-\frac{\varepsilon^2 m_\tau^k n_\tau}{8F(m_\tau^k + n_\tau)}\right).$$

Taking a union bound over all $\tau = 2, \ldots, t$ and $k \in [K_{\tau-1}]$ gives a simultaneous deviation event with the probability stated in the proposition.

On this event, if I_τ is a previously observed group, then

$$\mathrm{MMD}(D_\tau, M_\tau^{I_\tau}) \leq 2\sqrt{\frac{F}{m_\tau^{I_\tau}}} + 2\sqrt{\frac{F}{n_\tau}} + \frac{\varepsilon}{2} < \frac{\gamma}{2}.$$

For every previously observed group $k \neq I_\tau$,

$$\begin{aligned}\mathrm{MMD}(D_\tau, M_\tau^k) &\geq \mathrm{MMD}(\mathcal{D}_{I_\tau}, \mathcal{D}_k) - 2\sqrt{\frac{F}{m_\tau^k}} - 2\sqrt{\frac{F}{n_\tau}} - \frac{\varepsilon}{2} \\ &> \frac{\gamma}{2}.\end{aligned}$$

Thus an existing group is correctly identified. If I_τ is a new group, the latter inequality holds for every existing prototype, so the rule correctly creates a new group.

Before the first identification error, the Buffer-MMD and oracle executions coincide. Starting from the correct initialization of the first task, induction therefore rules out a first error on the simultaneous deviation event. This proves the identification guarantee without conditioning the oracle samples on previous identification success.

On the same event, $w_t = w_t^{\mathrm{or}}$. Outside this event, each population excess risk lies in $[0, 2B]$, since $|\ell(w, z)| \leq B$. Consequently, the expected excess risk of the Buffer-MMD execution is at most the oracle expected excess risk plus $2B$ times the failure probability. Substituting the preceding probability bound proves the claimed risk inequality. □

C.4 Results with EM-MMD

We present the online EM algorithm in Alg. 3 with Gaussian group distributions $\mathcal{D}_k = \mathcal{N}(\mu^k, \Sigma^k)$, where $\theta^k = (\mu^k, \Sigma^k)$. The method can be extended to other distribution families through the corresponding online EM updates (Cappé & Moulines, 2009). In Alg. 3, $(x_{t-1}^k, y_{t-1}^k, n_{t-1}^k, \hat{\pi}_{t-1}^k)_{k=1}^K$ are auxiliary variables, and $\phi(z; \mu, \Sigma)$ denotes the Gaussian density. We first use $\hat{\theta}_{t-1}$ to identify D_t through (EM-MMD), and then update the estimates using D_t.

Algorithm 3 Online EM Algorithm

Require: Previous sufficient statistics $(x_{t-1}^k, y_{t-1}^k, n_{t-1}^k, \hat{\pi}_{t-1}^k)_{k=1}^K$, dataset D_t, step size η_t
Ensure: Updated parameters $(\hat{\mu}_t^k, \hat{\Sigma}_t^k, \hat{\pi}_t^k)_{k=1}^K$

1: **for** $k = 1, \ldots, K$ **do**
2: Compute the responsibility using the previous parameters:

$$r_t^k = \frac{\hat{\pi}_{t-1}^k \prod_{z \in D_t} \phi(z; \hat{\mu}_{t-1}^k, \hat{\Sigma}_{t-1}^k)}{\sum_{j=1}^K \hat{\pi}_{t-1}^j \prod_{z \in D_t} \phi(z; \hat{\mu}_{t-1}^j, \hat{\Sigma}_{t-1}^j)}.$$

3: **end for**
4: **for** $k = 1, \ldots, K$ **do**
5: Update the mixture weight:

$$\hat{\pi}_t^k = (1 - \eta_t)\hat{\pi}_{t-1}^k + \eta_t r_t^k.$$

6: Update the sufficient statistics:

$$n_t^k = (1 - \eta_t) n_{t-1}^k + \eta_t r_t^k n, \qquad x_t^k = (1 - \eta_t) x_{t-1}^k + \eta_t r_t^k \sum_{z \in D_t} z,$$

$$y_t^k = (1 - \eta_t) y_{t-1}^k + \eta_t r_t^k \sum_{z \in D_t} z z^\top.$$

7: Update the Gaussian parameters:

$$\hat{\mu}_t^k = x_t^k / n_t^k, \qquad \hat{\Sigma}_t^k = y_t^k / n_t^k - \hat{\mu}_t^k (\hat{\mu}_t^k)^\top.$$

8: **end for**
9: **return** $(\hat{\mu}_t^k, \hat{\Sigma}_t^k, \hat{\pi}_t^k)_{k=1}^K$

For this result, we consider a fixed number of groups $K = K_T$ and equal task sizes $n_t = n$. Suppose that I_t are drawn independently with $\mathbb{P}(I_t = k) = \pi_k > 0$. Conditional on $I_t = k$, the samples in D_t are drawn independently from $\mathcal{D}_k$, independently of previous tasks. The parametric family is correctly specified. We match the component labels to the true groups by a fixed permutation and use uniform reference weights in the group-robust objective. All groups appear in finite time almost surely, after which the objective contains all K groups.

We assume that the initial estimates lie in a neighborhood of the true parameters where the online EM iterates remain stable and converge almost surely to the true parameters. The distribution $\mathcal{D}(\theta)$ is continuous in θ under MMD in this neighborhood. This is a local consistency assumption in addition to the stationary-point convergence considered in Cappé & Moulines (2009). The initial sufficient statistics have positive weights and nonsingular covariances, and the step sizes satisfy $0 < \eta_t < 1$, $\sum_t \eta_t = \infty$, and $\sum_t \eta_t^2 < \infty$. The estimates depend only on incoming tasks and are independent of the random draws used to update the replay buffer.

Suppose that the positive definite MMD kernel satisfies $0 \le \kappa(z, z') \le F$ and that, for some $\delta \in (0, 1)$ and $\varepsilon > 0$,

$$\mathrm{MMD}(\mathcal{D}_i, \mathcal{D}_j) \ge 2\sqrt{\frac{F}{n}} + 2\sqrt{\frac{2F\log(1/\delta)}{n}} + \varepsilon, \qquad i \ne j. \tag{30}$$

For the excess-risk bound, we further assume Assumption 1, $M \ge K$, exact minimization of the empirical objective, and positive replay sizes for all groups from an almost surely finite time onward.

Let $H_{t,k}$ contain all samples in tasks assigned to group k before time t, and let $\hat{L}_{H_{t,k}}(w)$ be their average loss. To compare the actual buffer with uniform replay, consider a reference reservoir for each feasible capacity $r \le M$ in each group. Each reservoir is updated by standard reservoir sampling, using the same acceptance draws and replacement indices as the actual buffer whenever their capacities agree. Conditional on the incoming tasks, each fixed-capacity reservoir is a uniform subset of $H_{t,k}$. Let $\bar{M}_t^k$ be the reference reservoir with capacity m_t^k, and define

$$A_t = \max_{k\in[K]} \sup_{w\in\mathcal{W}} \left| \hat{L}_{M_t^k}(w) - \hat{L}_{\bar{M}_t^k}(w) \right|.$$

These reference reservoirs are used only in the proof. Empty-pool mean losses are set to zero. Before all replay sizes are positive, we set A_t and the two integrands in the first two terms of (31) to zero; the contribution of these rounds vanishes asymptotically. As before, $q_{D_t \cup M_t}$ is the empirical maximizer at the population optimizer, with empirical group losses computed using the estimated assignments.

Proposition 5. *(Complete version of Proposition 3.) Under the assumptions above, for each $k \in [K]$,*

$$\liminf_{t\to\infty} \mathbb{P}(\hat{I}_t = I_t \mid I_t = k) \ge 1 - \delta.$$

Moreover, as $t \to \infty$,

$$\begin{aligned}
\mathcal{E}_{\mathrm{GRCL}}(w_t) \le B\mathbb{E}\left[\sum_{k=1}^{K} \big(q_t^k(w_t) + q_{D_t\cup M_t}^k\big)\psi\big(m_t^k + n\mathbf{1}\{k = I_t\}\big)\right] \\
+ 2B\sqrt{\big(2\log(2K(M+1)) + 2\big)\mathbb{E}\left[\frac{1}{\min_{k\in[K]}\{m_t^k + n\mathbf{1}\{k = I_t\}\}}\right]} + 2\mathbb{E}A_t \\
+ 2B\delta + 4B\max_{k\in[K]}\left\{\delta + \frac{(1-\pi_k)\delta}{\pi_k(1-\delta) + (1-\pi_k)\delta}\right\} + o(1).
\end{aligned} \tag{31}$$

If the positive target replay sizes eventually become fixed and are attained, then $\mathbb{E}A_t \to 0$. For the capacity-dependent bound in Proposition 3, let $\underline{m} \ge 1$ be a deterministic lower bound on these stabilized sizes, satisfying

$$\mathbb{P}\left(\liminf_{t\to\infty} \min_{k\in[K]} m_t^k \ge \underline{m}\right) = 1.$$

We prove the resulting bound after the general excess-risk bound below.

Proof of Proposition 5. By the assumed convergence of online EM and continuity in MMD, we have

$$\max_{k\in[K]} \mathrm{MMD}\Big(\mathcal{D}(\hat{\theta}_{t-1}^k), \mathcal{D}_k\Big) \longrightarrow 0 \quad \text{almost surely.}$$

Then, conditional on previous tasks and $I_t = i$, Lemma 7 gives, with probability at least $1 - \delta$,

$$\mathrm{MMD}(D_t, \mathcal{D}_i) \le \sqrt{F/n} + \sqrt{2F\log(1/\delta)/n}.$$

When the maximum estimation error is at most $\varepsilon/4$, the triangle inequality gives

$$\begin{aligned}\mathrm{MMD}\Big(D_t, \mathcal{D}(\hat{\theta}^i_{t-1})\Big) &\le \sqrt{F/n} + \sqrt{2F\log(1/\delta)/n} + \varepsilon/4,\\ \mathrm{MMD}\Big(D_t, \mathcal{D}(\hat{\theta}^k_{t-1})\Big) &\ge \mathrm{MMD}(\mathcal{D}_i, \mathcal{D}_k) - \mathrm{MMD}(D_t, \mathcal{D}_i) - \mathrm{MMD}\Big(\mathcal{D}(\hat{\theta}^k_{t-1}), \mathcal{D}_k\Big)\\ &\ge \sqrt{F/n} + \sqrt{2F\log(1/\delta)/n} + 3\varepsilon/4, \qquad k \ne i.\end{aligned}$$

These inequalities hold simultaneously for all $k \ne i$ on the same event. Therefore,

$$\mathbb{P}(\hat{I}_t \ne I_t \mid I_t = k) \le \delta + \mathbb{P}\left(\max_j \mathrm{MMD}\Big(\mathcal{D}(\hat{\theta}^j_{t-1}), \mathcal{D}_j\Big) > \varepsilon/4\right).$$

The second term converges to zero, which proves the identification result.

Next, we bound the fraction of incorrectly assigned samples. Almost surely, all estimation errors are at most $\varepsilon/4$ after a finite time. Thereafter, conditional on previous tasks, the probability of assigning a group-k task to k is at least $\pi_k(1-\delta)$, while the probability of assigning a task from another group to k is at most $(1-\pi_k)\delta$. By the strong law for bounded martingale differences and $n_t = n$, we have

$$\limsup_{t\to\infty} \frac{\sum_{\tau<t} n\mathbf{1}\{I_\tau \ne k, \hat{I}_\tau = k\}}{\sum_{\tau<t} n\mathbf{1}\{\hat{I}_\tau = k\}} \le \frac{(1-\pi_k)\delta}{\pi_k(1-\delta) + (1-\pi_k)\delta} \quad \text{almost surely.} \tag{32}$$

The correctly assigned samples may also be biased because the assignment depends on their values. Once the estimation errors are at most $\varepsilon/4$, conditioning a group-k task on correct assignment changes its distribution in total variation by at most δ. Thus its expected average loss differs from $L_k(w)$ by at most $2B\delta$. Applying the strong law for bounded martingale differences to these task losses, and then using a finite cover of $\mathcal{W}$ and Lipschitz continuity, gives the same bound uniformly over w. The finite initial period has vanishing weight. Since incorrectly assigned samples contribute at most $2B$ times their fraction, (32) implies

$$\limsup_{t\to\infty} \max_{k\in[K]} \sup_{w\in\mathcal{W}} |\hat{L}_{H_{t,k}}(w) - L_k(w)| \le 2B \max_{k\in[K]} \left\{\delta + \frac{(1-\pi_k)\delta}{\pi_k(1-\delta) + (1-\pi_k)\delta}\right\} \quad \text{almost surely.} \tag{33}$$

We next bound the sampling error of the reference buffers. Conditional on previous tasks and I_t, consider each feasible reference size $r \le M$, adding D_t to group I_t. Let $\Delta_{t,k}(r)$ be the supremum over w of the absolute difference between this empirical mean and its conditional expectation, namely the size-weighted average of $\hat{L}_{H_{t,k}}(w)$ and $L_k(w)$.

For a uniform sample without replacement from a fixed finite pool, the exponential moment of its centered sum is bounded by the corresponding with-replacement moment.[2] Together with the independence of D_t, this gives a Hoeffding bound for each fixed w. A parameter net of radius $B/(L\sqrt{s})$ has at most $(1 + 2LW\sqrt{s}/B)^d$ points, where $s = r + n\mathbf{1}\{k = I_t\} > 0$; approximating both losses contributes at most $2B/\sqrt{s}$. A union bound over both signs and the net points, using the definition of ψ, therefore gives, for $v > 0$,

$$\mathbb{P}\left(\Delta_{t,k}(r) > B\psi(s) + \frac{Bv}{\sqrt{s}} \,\middle|\, \text{past incoming tasks}, I_t\right) \le 2e^{-v^2/2}.$$

If $L = 0$ or $W = 0$, one cover point suffices.

Let

$$Z_t := \max_{\substack{k,r \text{ feasible}\\ r+n\mathbf{1}\{k=I_t\}>0}} \frac{\sqrt{r + n\mathbf{1}\{k = I_t\}}}{B} \left[\Delta_{t,k}(r) - B\psi(r + n\mathbf{1}\{k = I_t\})\right]_+.$$

There are at most $K(M+1)$ candidates. Hence

$$\mathbb{P}(Z_t > v) \le \min\{1, 2K(M+1)e^{-v^2/2}\}, \qquad \mathbb{E}Z_t^2 \le 2\log(2K(M+1)) + 2.$$

The second inequality follows by integrating the tail bound. Taking the maximum over r allows us to use the bound at the data-dependent replay sizes m^k_t.

[2] This is the convex comparison in Theorem 4 of Hoeffding (1963), *Probability Inequalities for Sums of Bounded Random Variables*, doi:10.1080/01621459.1963.10500830.

When all replay sizes are positive and $\hat{I}_t = I_t$, the triangle inequality gives

$$\sup_w |\hat{L}_k^t(w) - L_k(w)| \le B\psi\big(m_t^k + n\mathbf{1}\{k = I_t\}\big) + \frac{BZ_t}{\sqrt{m_t^k + n\mathbf{1}\{k = I_t\}}} + A_t + \max_j \sup_w |\hat{L}_{H_{t,j}}(w) - L_j(w)|.$$

For group I_t, the last two terms can be multiplied by $m_t^k/(m_t^k + n) \le 1$. By empirical optimality and cancellation of the KL terms, we have

$$\phi_t(w_t) - \phi_t(w_t^*) \le \sum_{k=1}^{K} \big(q_t^k(w_t) + q_{D_t \cup M_t}^k\big) \sup_w |\hat{L}_k^t(w) - L_k(w)|. \tag{34}$$

Since $\sum_k (q_t^k(w_t) + q_{D_t \cup M_t}^k) = 2$, the contribution of Z_t is at most

$$2B\mathbb{E}\left[\frac{Z_t}{\sqrt{\min_k\{m_t^k + n\mathbf{1}\{k = I_t\}\}}}\right] \le 2B\sqrt{\big(2\log(2K(M+1)) + 2\big)\mathbb{E}\left[\frac{1}{\min_k\{m_t^k + n\mathbf{1}\{k = I_t\}\}}\right]}.$$

The inequality follows from Cauchy–Schwarz. The inverse-size terms are set to zero when any replay size is zero. If a replay size is zero or the current task is misidentified, we use the bound $2B$ on the excess loss. The probability of a zero replay size converges to zero, and the limsup probability of current-task misidentification is at most δ. Taking expectations in (34) and using (33) therefore gives (31). Here boundedness allows us to pass the almost-sure upper limit in (33) to expectations. The $o(1)$ term accounts for the vanishing asymptotic remainders.

Finally, suppose that the positive target replay sizes eventually become fixed and are attained. For a fixed time a, consider updates after a at each possible fixed capacity. If the actual and reference buffers use the same acceptance draws and replacement indices, a sample occupying a mismatched position at time a survives until t with probability $G_k(a)/G_k(t)$, conditional on the incoming tasks. A union bound over the finitely many possible capacities bounds the expected mismatch by the probability that the allocation has not stabilized by a, plus a finite multiple of $\mathbb{E}[\min\{1, G_k(a)/G_k(t)\}]$. This comparison does not condition the reference sampling law on the random stabilization time. By the positive frequency of correctly assigned tasks, $G_k(t) \to \infty$ almost surely. Letting $t \to \infty$ and then $a \to \infty$, the expected mismatch converges to zero. Since each mismatched fraction contributes at most $2B$ times that fraction to the mean-loss difference, $\mathbb{E}A_t \to 0$.

By the definition of $\underline{m}$, each group eventually retains at least $\underline{m}$ samples almost surely. Monotonicity of ψ and the sum of the two weight vectors give

$$\mathcal{E}_{\mathrm{GRCL}}(w_t) \le 2B\psi(\underline{m}) + 2B\sqrt{\frac{2\log(2K(M+1)) + 2}{\underline{m}}} + \frac{10B\delta}{\min_k \pi_k} + o(1).$$

Here we used

$$\frac{(1-\pi_k)\delta}{\pi_k(1-\delta) + (1-\pi_k)\delta} \le \frac{\delta}{\pi_k}.$$

The probability that the replay sizes have not yet stabilized tends to zero, so these rounds contribute only to the $o(1)$ term. Substituting the definition of ψ proves (13). For $\underline{m} = \Omega(M/K)$, the statistical term is $\widetilde{O}(B\sqrt{dK/M})$. □

Lemma 7. *Let $D = \{z_i\}_{i=1}^m$ consist of i.i.d. samples from $\mathcal{D}_k$, and let the fixed positive definite kernel satisfy $0 \le \kappa(z, z') \le F$. For every $\delta \in (0, 1)$, with probability at least $1 - \delta$,*

$$\mathrm{MMD}(D, \mathcal{D}_k) \le \sqrt{\frac{F}{m}} + \sqrt{\frac{2F\log(1/\delta)}{m}}.$$

The MMD is the distance between the empirical and population RKHS mean embeddings, consistently with (27).

Proof. Let $\mu_D = m^{-1}\sum_{i=1}^m \kappa(z_i, \cdot)$ and $\mu_{\mathcal{D}_k} = \mathbb{E}\kappa(z, \cdot)$. Independence and $\kappa(z, z) \le F$ imply

$$\mathbb{E}\|\mu_D - \mu_{\mathcal{D}_k}\|_{\mathcal{H}}^2 = \frac{1}{m}\mathbb{E}\|\kappa(z, \cdot) - \mu_{\mathcal{D}_k}\|_{\mathcal{H}}^2 \le F/m.$$

Jensen's inequality gives $\mathbb{E}\,\mathrm{MMD}(D, \mathcal{D}_k) \le \sqrt{F/m}$. Replacing one observation changes the embedding distance by at most $2\sqrt{F}/m$. McDiarmid's inequality therefore gives

$$\mathbb{P}(\mathrm{MMD}(D, \mathcal{D}_k) - \mathbb{E}\,\mathrm{MMD}(D, \mathcal{D}_k) > v) \le e^{-mv^2/(2F)}.$$

Choosing $v = \sqrt{2F\log(1/\delta)/m}$ proves the result. □

D Missing details in Section 5

D.1 Implementation Details

Synthetic Data. For the synthetic experiment, we construct a binary classification problem with three task groups. Each input feature is $x \in \mathbb{R}^{20}$. For group $k \in \{1, 2, 3\}$, we sample

$$x \mid g = k \sim \mathcal{N}(\mu_k, \sigma_k^2 I_{20}),$$

where

$$\mu_1 = (3, 0, 0, \ldots, 0), \quad \mu_2 = (0, 3, 0, \ldots, 0), \quad \mu_3 = (0, 0, 3, \ldots, 0),$$

and

$$\sigma_1 = 0.5, \qquad \sigma_2 = 0.5, \qquad \sigma_3 = 1.5.$$

Thus, the third group has substantially larger feature variance and is harder to learn.

The binary label is generated from a noisy linear rule. Let

$$w^\star = \frac{(4, 1, 0.05, 0, \ldots, 0)}{\|(4, 1, 0.05, 0, \ldots, 0)\|_2}.$$

For each generated sample x, we sample independent label noise

$$\epsilon \sim \mathcal{N}(0, 0.5^2),$$

and assign

$$y = \mathbf{1}\{x^\top w^\star + \epsilon > 0\}.$$

The arriving group distribution is $\pi = (0.76, 0.19, 0.05)$. We run 10 continual-learning rounds, and in each round draw 10,000 new examples from the arriving group. The test set contains 20,000 examples per group, so the reported test risk closely approximates the population group risk.

Our trained model is a two-hidden-layer MLP, with architecture $20 \to 16 \to 16 \to 1$ and ReLU activations. The output is a single logit and the training loss is binary cross-entropy with logits. The model is trained with AdamW using learning rate 3×10^{-3} and weight decay 10^{-3} for 10 epochs per continual-learning round. We use full-batch training on the current replay-training set in each round. We evaluate replay budgets $M \in \{300, 1500, 5000\}$ and average over 100 random seeds. For difficulty-adaptive replay, we approximately solve (12) by allocating storage across groups in proportion to the 4/5-th power of the average $\hat{q}_{D_\tau \cup M_\tau}$ for each group.

Waterbirds dataset For Waterbirds, we use the standard four groups defined by the bird label and background: landbird on land, landbird on water, waterbird on land, and waterbird on water. The training set is highly imbalanced across these groups, with 3498, 184, 56, and 1057 samples, respectively. The test set contains 2255, 2255, 642, and 642 samples for the four groups, respectively. In the CL stream, each task contains samples from only one hidden group. The learner does not observe this true group identity during replay buffer allocation or training; the true group labels are used only for evaluation.

We use an ImageNet-pretrained ResNet-50 as a frozen feature extractor. Each image is mapped to a 2048-dimensional feature vector, and the classifier is a two-hidden-layer MLP head trained on these frozen features. The head has architecture

$$2048 \to 512 \to 128 \to 1,$$

with ReLU activations and a single binary logit output. The classifier is trained with binary cross-entropy loss. We run 20 continual-learning rounds with replay budgets $M \in \{110, 130, 160\}$ and average over 400 random seeds. The classifier is trained using AdamW with learning rate 10^{-4}, weight decay 10^{-4}, batch size 256, and 6 epochs per round.

For online group identification, we use the same frozen ResNet-50 features. To reduce the dimensionality used by the identification module, we normalize the 2048-dimensional features and apply a

fixed random projection to 64 dimensions. The projection matrix is sampled once with entries from a Gaussian distribution with standard deviation $1/\sqrt{64}$ and is fixed throughout the experiment. Both EM-MMD and Buffer-MMD operate in this projected 64-dimensional feature space.

For Buffer-MMD, we use a new-group threshold $\gamma = 0.02$. Given an incoming batch in the projected identification space, we compute its MMD distance to each active replay buffer using the RBF kernel

$$k(u, v) = \exp\left(-\frac{\|u - v\|_2^2}{2h^2}\right),$$

where the bandwidth h is chosen by the median heuristic on the incoming batch (Gretton et al., 2012). For EM-MMD, we use the same new-group activation rule as Buffer-MMD to initialize at most four latent groups. Each latent group is modeled as a diagonal Gaussian in the projected identification space, and its parameters are updated online using Alg. 3. For difficulty-adaptive replay, we approximately solve (12) by allocating storage across groups in proportion to $\frac{1}{2}\tilde{q}_t^{2/3} + \frac{1}{2}u$, where u denotes the uniform allocation over groups.

CelebA dataset. Due to the high computational cost of end-to-end model training, we run the CelebA experiment once for 20 continual-learning rounds with replay budget $M = 500$. We train the model using AdamW with learning rate 10^{-4}, weight decay 10^{-4}, batch size 256, and 5 epochs per round. For difficulty-adaptive replay, we approximately solve (12) by allocating storage across groups in proportion to $\frac{1}{2}\tilde{q}_t^{2/3} + \frac{1}{2}u$, where u denotes the uniform allocation over groups.

For all experiments, we choose the prior distribution u_t to be the uniform distribution over groups, use the KL divergence on $D(q\|u_t)$, and set $\lambda = 0.1$.

Computational Resources. All experiments were run on a single workstation with 32 CPU logical cores and one NVIDIA GeForce RTX 5080 GPU with 16GB memory. For the synthetic and Waterbirds experiments, running 100 random seeds takes less than 10 minutes. The CelebA experiment is substantially more expensive because it trains the full ResNet-18 model end-to-end; one seed takes more than 6 hours.

Dataset licenses and terms. We use all existing datasets only for non-commercial research purposes and follow their respective license terms and usage policies. CelebA is available for non-commercial research purposes only.[3] Waterbirds[4] is a derived benchmark constructed from CUB-200-2011[5] and Places (Wah et al., 2011; Zhou et al., 2017; Sagawa et al., 2019). We cite the original data sources and follow the license terms and usage policies of the underlying datasets.

Additional Continual Learning Metrics Let $A_{t,g}$ denote the accuracy on the fixed test set of group g after training round t, and let K_T denote the total number of unique groups observed by round T. We first report the final sample-weighted overall accuracy. Let n_g denote the number of test samples belonging to group g. The overall accuracy after the final training round is defined as

$$\mathrm{OverallAcc} = \frac{\sum_{g=1}^{K_T} n_g A_{T,g}}{\sum_{g=1}^{K_T} n_g}.$$

Thus, each group contributes in proportion to its number of test samples, making $\mathrm{OverallAcc}$ equivalent to the standard accuracy computed over the entire test set. Higher overall accuracy indicates better average predictive performance, although it may be dominated by groups containing more samples.

[3] https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html

[4] https://github.com/kohpangwei/group_DRO

[5] https://www.vision.caltech.edu/datasets/cub_200_2011/

Table 2: Performance comparison on the Waterbirds and CelebA datasets. The best result for each metric on each dataset is shown in bold. ↑ (↓) indicates that higher (lower) values are better.

Method	Overall Acc. ↑	WGA ↑	Task F ↓	BWT ↑	MGF ↓
Waterbirds					
Task ER	81.35 ± 0.59	32.90 ± 1.42	7.09 ± 0.27	1.58 ± 0.48	17.37 ± 0.47
Class ER	78.06 ± 0.76	37.87 ± 1.50	7.90 ± 0.35	-0.16 ± 0.51	18.82 ± 0.43
Group ER	86.27 ± 0.73	59.24 ± 1.42	6.81 ± 0.52	2.22 ± 0.64	11.98 ± 0.60
Ours: Group-balanced GRCL	88.25 ± 0.45	76.28 ± 0.87	4.44 ± 0.25	3.42 ± 0.48	7.64 ± 0.25
Ours: Difficulty-adaptive GRCL	$\mathbf{88.34 \pm 0.42}$	$\mathbf{76.55 \pm 0.84}$	4.34 ± 0.21	3.52 ± 0.47	7.47 ± 0.22
DER++ (Buzzega et al., NeurIPS 2020)	83.45 ± 0.37	27.08 ± 2.06	6.20 ± 0.32	1.99 ± 0.38	13.75 ± 0.57
MIR (Aljundi et al., NeurIPS 2019)	79.29 ± 0.48	46.79 ± 1.18	5.32 ± 0.17	4.54 ± 0.38	13.91 ± 0.37
LPR (Yoo et al., ICML 2024)	85.61 ± 0.22	28.27 ± 2.00	$\mathbf{1.70 \pm 0.12}$	$\mathbf{8.82 \pm 0.42}$	$\mathbf{4.23 \pm 0.36}$
CelebA					
Task ER	93.57	6.25	20.00	-19.98	37.61
Class ER	93.34	0.00	17.39	-13.58	32.63
Group ER	93.23	12.50	16.91	-13.94	29.77
Ours: Group-balanced GRCL	91.81	18.75	12.89	-11.73	20.18
Ours: Difficulty-adaptive GRCL	92.84	**62.50**	**10.33**	-5.42	**14.27**
DER++ (Buzzega et al., NeurIPS 2020)	94.37	0.00	18.44	-11.56	35.16
MIR (Aljundi et al., NeurIPS 2019)	**94.67**	24.39	12.56	$\mathbf{-1.92}$	24.13
LPR (Yoo et al., ICML 2024)	87.69	0.00	25.48	-18.37	41.62

Let I_j denote the group arriving at round j. We compute task forgetting as

$$\text{Task Forgetting} = \frac{1}{T-1} \sum_{j=1}^{T-1} \left[\max_{t \in \{j,\ldots,T\}} A_{t,I_j} - A_{T,I_j} \right],$$

and backward transfer as

$$\text{Backward Transfer} = \frac{1}{T-1} \sum_{j=1}^{T-1} \left(A_{T,I_j} - A_{j,I_j} \right).$$

We also evaluate forgetting at the group level. Let s_g denote the first round in which group g appears. Macro group forgetting is defined as

$$\text{Macro Group Forgetting} = \frac{1}{K_T} \sum_{g=1}^{K_T} \left[\max_{t \in \{s_g,\ldots,T\}} A_{t,g} - A_{T,g} \right].$$

Lower task forgetting and macro group forgetting indicate less forgetting, whereas higher backward transfer indicates more positive backward transfer. Because groups may recur in our stream, task forgetting and backward transfer are arrival-weighted extensions of conventional CL metrics, whereas macro group forgetting assigns equal weight to each unique group.

D.2 Additional Results on Synthetic Data

We construct a synthetic binary classification problem with a simple case of three groups, denoted as easy, medium, and hard. The stream is imbalanced: the three groups arrive with frequencies $16 : 4 : 1$, respectively. We compare our proposed methods in Alg. 1, against three baselines: task-balanced replay with ERM model update (Rolnick et al., 2019; Deng et al., 2025), class-balanced replay with ERM model update (Chrysakis & Moens, 2020), and task-balanced replay enhanced by our GRCL model update. After 10 CL rounds, our methods in Alg. 1 improve both the worst-group test accuracy and the group-robust loss, see Figs. 3 and 4. We further visualize buffer allocation over groups for four replay methods in Fig. 2. Task-balanced replay overstores the frequent easy group, while class-balanced replay mainly stores the easy and hard groups because they dominate the two class labels, which fails to capture the difficult medium group. In contrast, our difficulty-adaptive replay uses the learned adversarial weights to shift memory toward hard and medium groups and achieves the lowest group-robust objective in Fig. 3.

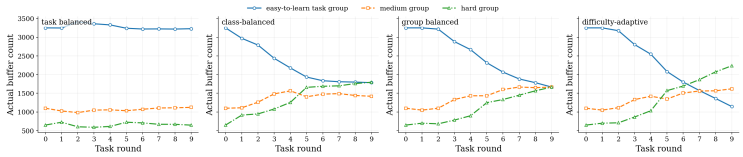


Figure 2: Actual replay buffer allocation on the synthetic dataset under budget $M = 5000$, corresponding to task-balanced, class-balanced, our group-balanced, and difficulty-adaptive replay.

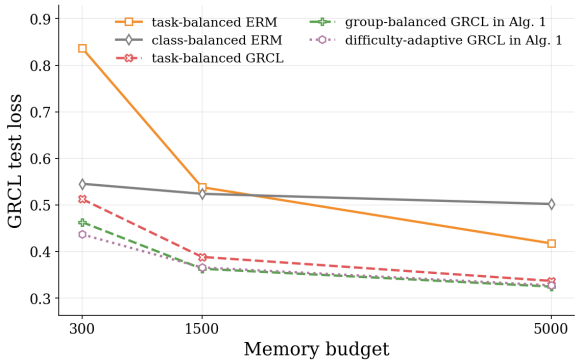


Figure 3: Final GRCL loss on the synthetic dataset under replay budgets $M \in \{300, 1500, 5000\}$. We compare group-balanced GRCL replay and difficulty-adaptive GRCL replay with task-balanced ERM replay (Deng et al., 2025), class-balanced ERM replay, group-balanced ERM replay, and task-balanced GRCL replay.

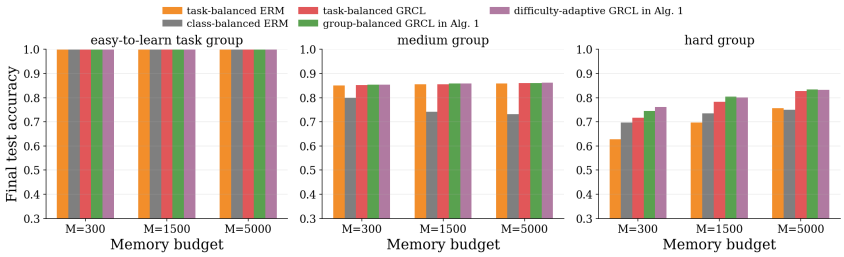


Figure 4: Final per-group test accuracy on the synthetic dataset under replay budgets $M \in \{300, 1500, 5000\}$.

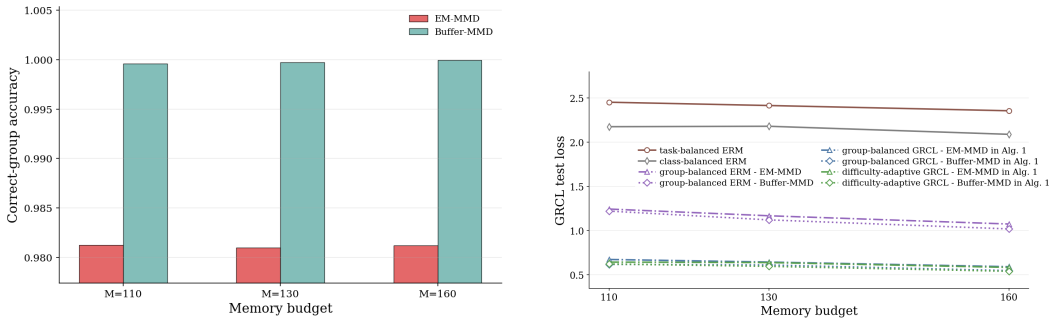
Fig. 3 shows that GRCL-based replay substantially reduces the final group-robust loss compared with ERM replay baselines. At budget $M = 1500$, task-balanced ERM replay and group-balanced ERM replay obtain GRCL losses of 0.538 and 0.463, respectively. In contrast, task-balanced GRCL replay reduces the loss to 0.388, while group-balanced GRCL replay and difficulty-adaptive replay further reduce it to 0.363 and 0.350, respectively. This confirms that the improvement is not only due to balancing the replay buffer, but also to the GRCL model update. Across budgets, difficulty-adaptive replay performs best at $M = 300$ and $M = 1500$, while group-balanced GRCL replay becomes slightly better at $M = 5000$. This phenomenon is intuitive: as the storage size increases, the benefit of replay allocation becomes smaller.

Fig. 4 explains the source of the GRCL-loss improvement. The easy group is already learned well by most methods, so the main difference comes from the medium and hard groups. At $M = 1500$, the worst-group accuracy increases from 0.696 under task-balanced ERM replay and 0.723 under group-balanced ERM replay to 0.801 under group-balanced GRCL replay and 0.809 under difficulty-adaptive GRCL replay. Thus, the lower GRCL objective is achieved by improving the underperforming groups rather than only improving average accuracy. The gap between task-balanced and group-aware replay also shows the effect of the imbalanced arrival distribution: task-balanced replay over-stores examples from the frequent easy group, while group-balanced and difficulty-adaptive replay allocate more memory to the medium and hard groups.

D.3 Additional Results on Waterbirds

Fig. 5 reports group-identification accuracy and final GRCL loss on Waterbirds. Both identification methods are reliable, with Buffer-MMD achieving nearly perfect replay buffer group identification. At budget $M = 160$, Buffer-MMD obtains buffer identification accuracy close to 1.0, while EM-MMD is around 0.98. This indicates that the projected ResNet-50 features contain enough information to separate the hidden Waterbirds groups, even though true group identities are not provided to the replay algorithm. Although Buffer-MMD has a worse theoretical guarantee, its empirical performance is better than EM-MMD.

The final GRCL loss in Fig. 5b shows that our methods in Alg. 1 outperform all baselines. At $M = 160$, task-balanced ERM replay obtains a GRCL loss of 2.356. Replacing task-balanced replay with class-balanced ERM replay only reduces the loss to 2.089, showing that balancing the bird label alone is not sufficient to address the spurious correlation between bird type and background. When we further replace class-balanced replay with group-balanced replay using Buffer-MMD, the group-balanced ERM loss decreases to 1.019, indicating the benefit of identifying and balancing the hidden label-background groups. Finally, replacing the ERM model update with the GRCL model update further reduces the loss to 0.548 for group-balanced GRCL replay, and difficulty-adaptive GRCL replay achieves the lowest loss of 0.540. This confirms that both components are important: group-aware replay mitigates the imbalance over spurious-correlation groups, while the GRCL update further emphasizes the high-loss minority groups during training.



(a) replay buffer group-identification accuracy under $M \in \{110, 130, 160\}$.

(b) Final GRCL loss under EM-MMD and Buffer-MMD group identification.

Figure 5: Additional Waterbirds results for group identification and final GRCL loss.

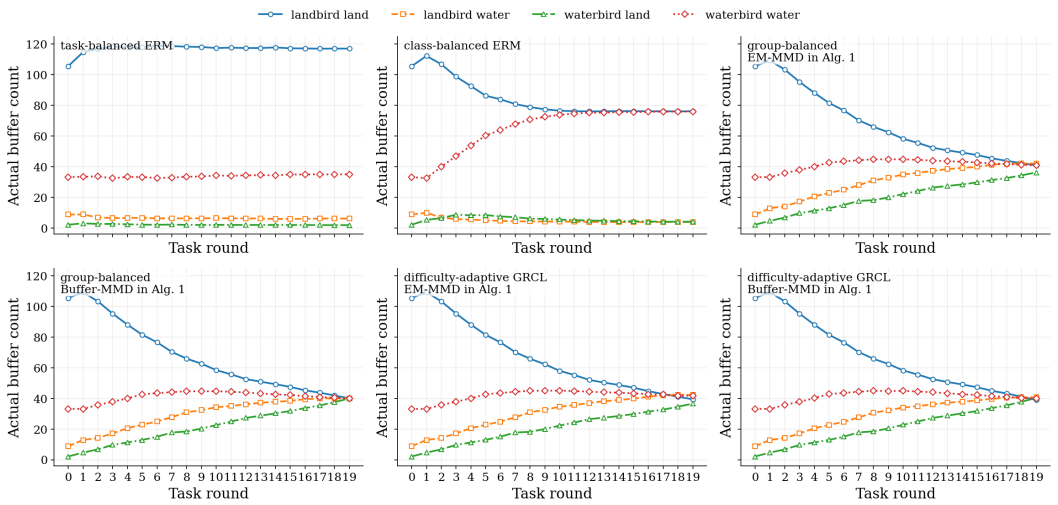


Figure 6: replay buffer allocation across the four Waterbirds groups during continual learning under budget $M = 160$.

Fig. 6 shows how replay storage evolves during continual learning. Task-balanced ERM stores mostly samples from the frequent majority group, landbirds on land, and allocates very little memory to the rare groups. Class-balanced replay balances the two bird labels, but it still largely overlooks the landbird-on-water and waterbird-on-land groups, which are the minority groups induced by the spurious background correlation. In contrast, the group-balanced and difficulty-adaptive methods allocate memory much more evenly across the hidden groups identified by EM-MMD or Buffer-MMD. This explains why their worst-group performance is much stronger: they preserve replay examples from the minority spurious-correlation groups that are most vulnerable to forgetting.

In Fig. 6, the replay allocations of group-balanced and difficulty-adaptive replay are nearly the same, due to our conservative approximation to (12), which averages the uniform allocation and the learned $\tilde{q}_t$.

D.4 Additional Results on CelebA

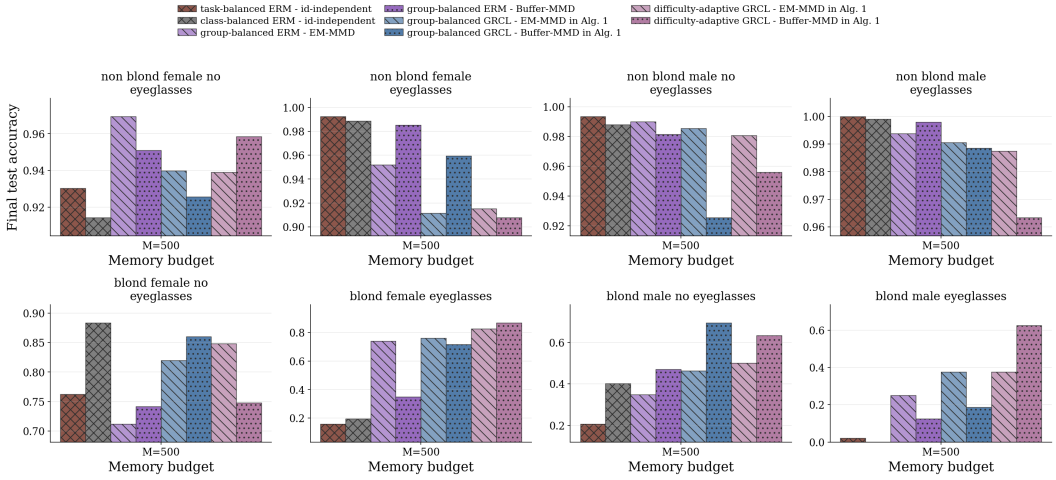


Figure 7: Final per-group test accuracy on CelebA across the eight groups defined by hair color, gender, and eyeglasses.

Fig. 7 reports per-group accuracy on CelebA. The results show that task-balanced and class-balanced ERM replay perform well on frequent or easy groups, but their accuracy drops substantially on rare groups involving blond male examples. In particular, the blond-male groups are the most challenging because they are both minority groups and strongly affected by attribute imbalance. Group-balanced GRCL and difficulty-adaptive GRCL improve these difficult groups, demonstrating that the proposed replay methods remain effective on a more complex dataset. The comparison with group-balanced ERM further shows that group-aware storage alone is not enough; the GRCL update is also important for improving worst-group robustness.

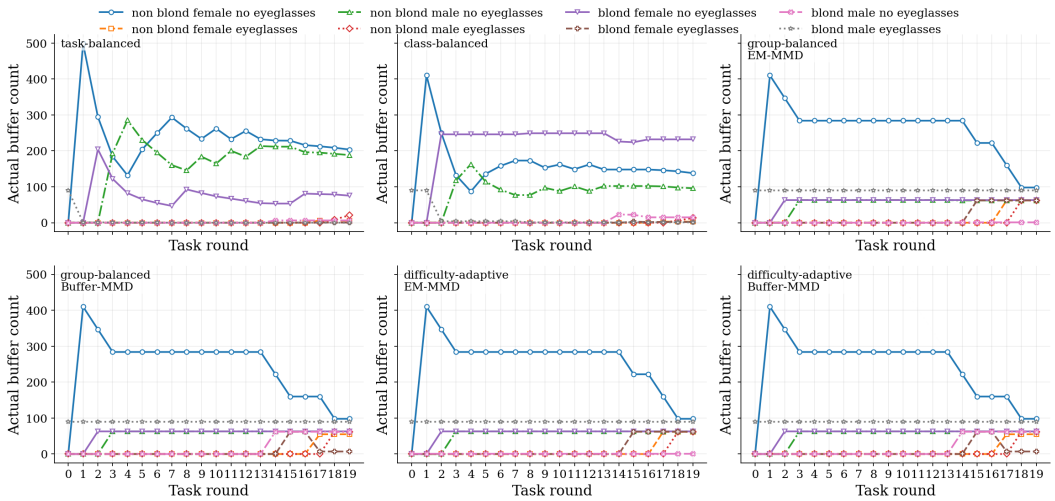


Figure 8: replay buffer allocation across the eight CelebA groups during continual learning under budget $M = 500$.

Fig. 8 shows the replay buffer allocation over the 20 continual-learning rounds on CelebA. Task-balanced replay follows the arrival frequencies and therefore stores more samples from the dominant groups. Class-balanced replay balances the blond and non-blond labels, but it does not distinguish gender or eyeglasses, so it can still underrepresent rare intersectional groups. By contrast, group-balanced and difficulty-adaptive replay allocate memory more evenly across the identified latent groups. This allocation behavior is consistent with the per-group accuracy results in Fig. 7. Since group-balanced and difficulty-adaptive replay use the same storage allocation due to our conservative approximation to (12), the better performance of difficulty-adaptive replay in Figs. 1b and 7 is due to one-time randomness.

CelebA contains 40 attribute labels, whereas our experiments use only 8 of them to define the prediction tasks and group structure. The remaining attributes are not considered in our experimental setup. In group-robust CL, however, the choice of group definition is critical: an inappropriate or incomplete grouping may fail to capture important sources of distributional imbalance and can potentially lead to fairness issues. Therefore, applying group-robust CL in practice requires careful consideration of which attributes are used to define groups and how these groups reflect the underlying fairness concerns.

D.5 Additional Evaluation on CIFAR-10

We further evaluate our method on CIFAR-10 under a recurring-class continual-learning setting. The 50,000 training images are divided into 30 class-homogeneous arrivals by partitioning each of the ten classes into three disjoint subsets. The arrival order is independently randomized for each seed, while every training example is observed exactly once. The model performs single-head 10-way classification without access to the arrival identity at test time. We use a memory budget of $M = 500$ and compare three replay strategies: (i) frequency-based reservoir replay trained with ERM, (ii) class-uniform memory allocation combined with Group DRO updates, and (iii) LPR-adapted, which applies the layerwise proximal replay mechanism to our multiclass protocol.

All methods use the same CIFAR-adapted ResNet-18, model initialization, data stream, augmentation, optimizer, and training budget. At each arrival, the model is trained for 10 epochs using balanced minibatches containing 128 current and 128 replay examples. We report results over ten matched seeds (52–61) as mean $\pm$ standard deviation. Worst-group accuracy is the minimum final test accuracy across the ten classes, GDRO accuracy is the final group-accuracy average weighted by the learned Group DRO weights, and average accuracy is the mean final test accuracy across all classes. Since CIFAR-10 has a balanced test set, average accuracy is also equal to sample-weighted overall accuracy.

Table 3: CIFAR-10 results under a recurring-class continual stream with $M = 500$. Results are mean $\pm$ standard deviation over ten matched seeds. Higher is better for all metrics.

Method	Worst-group Acc. ↑	GDRO Acc. ↑	Average Acc. ↑
Frequency + reservoir replay	12.70 ± 8.80	31.37 ± 9.67	$\mathbf{39.86 \pm 4.00}$
Group-balanced GRCL	$\mathbf{21.17 \pm 5.96}$	$\mathbf{38.66 \pm 4.58}$	38.75 ± 3.83
LPR-adapted	13.76 ± 8.75	32.89 ± 9.90	39.43 ± 3.39

Uniform memory allocation combined with Group DRO substantially improves robust performance. In particular, it improves worst-group accuracy by 8.47 percentage points over frequency-based reservoir replay and by 7.41 points over LPR-adapted. It also achieves the highest GDRO accuracy, outperforming the two baselines by 7.29 and 5.77 points, respectively. These improvements incur only a modest reduction in average accuracy: 1.11 points relative to frequency-based replay and 0.68 points relative to LPR-adapted. Moreover, its lower standard deviations on the two robust metrics indicate greater stability across different arrival orders. Overall, the results demonstrate that balancing replay memory across recurring groups and optimizing the robust objective can markedly improve worst-group performance while largely preserving average predictive accuracy.