Adversarially Robust Multitask Adaptive Control

Kasra Fallah*, Leonardo F. Toso*, and James Anderson

Department of Electrical Engineering, Columbia University

November 2025

Abstract

We study *adversarially robust* multitask adaptive linear quadratic control; a setting where multiple systems collaboratively learn control policies under model uncertainty and adversarial corruption. We propose a clustered multitask approach that integrates clustering and system identification with *resilient aggregation* to mitigate corrupted model updates. Our analysis characterizes how clustering accuracy, intra-cluster heterogeneity, and adversarial behavior affect the expected regret of certainty-equivalent (CE) control across LQR tasks. We establish non-asymptotic bounds demonstrating that the regret decreases inversely with the number of honest systems per cluster and that this reduction is preserved under a bounded fraction of adversarial systems within each cluster.

1 Introduction

Adaptive control seeks to design controllers that adapt to uncertain or unknown system dynamics. Rooted in early work on self-tuning regulators for flight and aerospace applications [Åström and Wittenmark, 1973, Åström, 1983], it remains central to modern control. Among its formulations, the linear quadratic regulator (LQR) serves as a canonical benchmark due to its tractability and theoretical appeal. Extensive research over the last five or so years has established *non-asymptotic* performance guarantees for adaptive LQR through regret analysis [Abbasi-Yadkori and Szepesvári, 2011, Dean et al., 2018, Cohen et al., 2019, Simchowitz and Foster, 2020, Hazan et al., 2020, Ziemann and Sandberg, 2022], proving that in the single-system setting the optimal expected regret scales as $\mathcal{O}(\sqrt{dT})$, with $d = d_u^2 d_x$, where T is the time horizon and (d_x, d_u) denote the state and input dimensions [Simchowitz and Foster, 2020]. This lower bound reveals a fundamental limitation: certainty-equivalent (CE) control is inherently data-inefficient in high dimensions, as accurate model estimation demands extensive data collection.

To circumvent this limitation, recent work has investigated *multitask* system identification, where multiple systems collaboratively estimate their dynamics. When the participating systems are "similar", collaboration reduces the sample complexity required for accurate model estimation, with gains that scale proportionally with the number of participating systems [Xin et al., 2022,

*K. Fallah and L. F. Toso share first-authorship.
Correspondence to: {kasra.fallah, leonardo.toso}@columbia.edu.

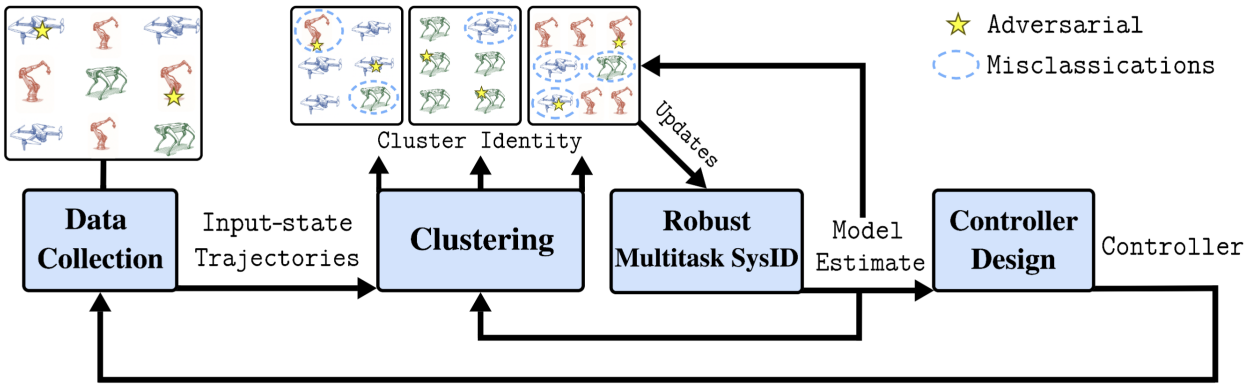

Figure 1: Workflow for adversarially robust multitask adaptive control[†].

Wang et al., 2023a, Toso et al., 2023, Keçeci et al., 2025a]. When the systems are homogeneous or share some model structure (e.g., a model basis), joint estimation yields improvements in sample complexity [Zhang et al., 2024]. On the other hand, when systems are only approximately similar, an additive *heterogeneity* bias emerges [Wang et al., 2023a]. As the primary source of regret in adaptive control stems from identification, multitask identification offers a natural path to surpass the $\mathcal{O}(\sqrt{T})$ scaling limit from the single-task setting.

Recent work has extended the idea of multitask system identification [Wang et al., 2023a, Zhang et al., 2024] to the control synthesis setting [Wang et al., 2023b, Toso et al., 2024, Wang et al., 2023c, Lee et al., 2025]. Lee et al. [2025] prove that learning a shared model basis across systems allows the expected regret to scale as $\mathcal{O}(\sqrt{T/\#\text{systems}})$, reducing the regret by the number of systems. However, this result relies on *structural* homogeneity, namely, the existence of a shared model representation across all participating systems.

In practice, multitask control systems, such as fleets of drones, autonomous vehicles, and distributed robotic platforms [Wang et al., 2023c] often exhibit diverse dynamics without a shared representation. In such settings, malfunctioning or compromised systems may transmit corrupted updates, undermining collaboration. This motivates the need for adversarially robust multitask adaptive control, where learning remains effective despite heterogeneous and adversarial systems.

This work studies clustered multitask adaptive control in the presence of *heterogeneous* and *adversarial* systems (see Figure 1), when a common representation across systems *may not exist*. We analyze how malicious systems can bias the collaborative learning step and design a robust approach to mitigate such behavior. Our analysis characterizes the interplay among adversarial behavior, intra-cluster heterogeneity, and clustering accuracy on the regret. We prove that the regret scales favorably with the number of systems per cluster. This benefit dominates even under a small fraction of adversarial systems per cluster. This provides the first non-asymptotic regret bounds for clustered multitask adaptive control under heterogeneous and adversarial systems (see Table 1). We now present an informal statement of our main result.

Theorem 1.1 (Informal)**.** *Consider the multitask adaptive linear quadratic control pipeline illustrated in Figure 1. For an appropriate choice of exploration and a sufficiently large amount of data per system, let T denote the time horizon, m_j the number of honest systems in cluster $\mathcal{C}_j$, and λ the resilient aggregation coefficient. Then, the expected regret of any system in $\mathcal{C}_j$ satisfies:*

[†] Illustrations of drones, quadruped robots, and robotic arms were created with assistance from ChatGPT (OpenAI).

$$\texttt{Regret} \lesssim \underbrace{\sqrt{\frac{dT}{m_j}}}_{\text{benefit of multitask}} + \underbrace{\texttt{cluster error} \times T}_{\text{clustering effect}} + \underbrace{\lambda\sqrt{dT}}_{\text{adversarial effect}} + \underbrace{\texttt{heterogeneity} \times T}_{\text{heterogeneity effect}},$$

where $\texttt{cluster error} \lesssim \exp(\texttt{-\# data samples per system})$.

Theorem 1.1 highlights the benefits of multitask adaptive control and quantifies the effects of clustering errors, adversarial systems, and heterogeneity. The first term captures the benefit of *collaboration*, reducing regret proportionally to the number of honest systems m_j. The second term reflects the clustering misclassification rate that decays exponentially with data size. The third term represents the effect of adversarial updates, governed by the resilient aggregation parameter λ, which scales with the ratio of adversarial to honest systems [Farhadkhani et al., 2022], when this ratio is small, the collaboration benefit is preserved. The final term captures the *heterogeneity* effect, which is negligible under small intra-cluster heterogeneity. We characterize the first two terms in Theorem 4.1, the adversarial effect in Theorem 4.2, and the heterogeneity effect in Corollary 4.1.

1.1 Contributions

• **Heterogeneous Systems.** We study multitask adaptive LQR under heterogeneous dynamics where global representation across system models are not assumed. We propose a clustered system identification approach that groups similar systems and performs multitask identification within each cluster. We derive non-asymptotic estimation error bounds demonstrating that sample complexity improves proportionally to the number of systems per cluster, up to an exponentially small cluster misclassification rate (Lemma 3.1 and Proposition 3.1).

• **Adversarial Systems.** Our approach handles adversarial systems that may contribute with corrupted model updates. Robustness is ensured with a *resilient* aggregation step that may employ any resilient aggregation rule [Farhadkhani et al., 2022]. Under a bounded fraction of adversarial systems, our approach preserves the asymptotic gains as in the fully honest setting (Lemma 3.2).

• **Regret Bounds.** We provide the first non-asymptotic regret bounds for adversarially robust multitask adaptive LQR. The analysis quantifies the impact of clustering misclassifications, intra-cluster heterogeneity, and adversarial behavior, demonstrating that with sufficient data, small intra-cluster heterogeneity, and bounded fraction of adversarial systems these effects are negligible, preserving the regret reduction by the number of honest systems (Theorems 4.1 and 4.2, and Corollary 4.1).

1.2 Related Work

Recent work in multitask system identification have explored collaborative learning to improve sample complexity. Wang et al. [2023a] first proved that the sample complexity of learning linear dynamical systems decreases with the number of collaborating systems, up to a heterogeneity bias. This bias is replaced by a misclassification term in clustered system identification [Toso et al., 2023, Keçeci et al., 2025b], where systems are assumed identical within clusters and the misclassification rate decays exponentially with the amount of local data. Zhang et al. [2024] instead assumes a shared latent representation across system models, implicitly requiring strong structural similarity. In contrast, our work considers multitask clustered system identification under intra-cluster heterogeneity, allowing for settings where no global representation exists.

Within adaptive control, the most relevant work is Lee et al. [2025], which learns a shared basis across systems to accelerate policy adaptation. Our framework removes this structural assumption by learning cluster-specific models that capture local similarities while remaining robust to adversarial systems. Adversarially resilient learning has been well studied in distributed and federated settings [Blanchard et al., 2017, Chen et al., 2018, Farhadkhani et al., 2022, Dong et al., 2023], where resilient aggregation ensures robustness against malicious agents. We extend these ideas to the more challenging domain of system identification and control, providing, to the best of our knowledge, the first analysis of adversarially robust multitask adaptive control.

Table 1: Expected regret for single-and multitask online learning. In the regret bound reported by Lee et al. [2023], d_w denotes the dimension of the task-specific weight vector, as the model is decomposed into a task weight and a shared representation. Moreover, $\delta_{\texttt{dist}}$ quantifies the discrepancy between the ground-truth representation of the system model and a given pre-trained representation. The setting considered in Lee et al. [2025] assumes that the M tasks are structurally similar so that a common representation exists.

Work	Setting	Heterogeneity	Adversarially Robust	Regret
Simchowitz and Foster [2020]	Single task	✗	✗	$\mathcal{O}(\sqrt{dT})$
Lee et al. [2023]	Single task	✗	✗	$\mathcal{O}(\sqrt{d_w d_x T}) + \delta_{\texttt{dist}} T$
Lee et al. [2025]	Multitask	Structurally Homogeneous	✗	$\mathcal{O}\left(\sqrt{\frac{\texttt{poly}(d_x,d_u)T}{M}}\log^2(TM)\right)$
Dong et al. [2023]	Distributed	Heterogeneous	✓	$\mathcal{O}(\sqrt{T} + \epsilon_{\text{het}} T)$
This work	Multitask	Heterogeneous	✓	$\mathcal{O}\left(\sqrt{\frac{dT}{m_j}} + \lambda\sqrt{dT} + \epsilon_{\text{het}} T\right)$

1.3 Notation

The norm $\|\cdot\|$ denotes the Euclidean norm for vectors and the spectral norm for matrices, while $\|\cdot\|_F$ is the Frobenius norm. The operators $\texttt{dlyap}(A, Q)$ and $\texttt{DARE}(A, B, Q, R)$ denotes the solutions to the discrete Lyapunov and Riccati equations, respectively. We use $\mathcal{O}(\cdot)$ to hide constant factors and $\widetilde{\mathcal{O}}(\cdot)$ to hide logarithmic terms.

2 Problem Setup

We consider M discrete-time linear time-invariant (LTI) systems, among which an unknown f systems may behave adversarially. The remaining $m = M - f$ honest systems evolve according to

$$x_{t+1}^{(i)} = \Theta_\star^{(i)} z_t^{(i)} + w_t^{(i)},\ \forall t = 0, 1, \ldots,\ z_t^{(i)} = \begin{bmatrix} x_t^{(i)} \\ u_t^{(i)} \end{bmatrix} \in \mathbb{R}^{d'},\ \text{with } d' = d_x + d_u, \tag{1}$$

where $\Theta_\star^{(i)} = [A_\star^{(i)}\ B_\star^{(i)}]$ denotes system i's model, and $\{w_t^{(i)}\}_t$ has elements that are i.i.d. mean-zero σ_w^2-sub-Gaussian, for some positive variance proxy σ_w^2 [Vershynin, 2018].

Clusters. Each honest system belongs to one of N_c clusters $\{\mathcal{C}_j\}_{j=1}^{N_c}$, grouping systems with similar dynamics. Systems within the same cluster share model parameters that are "close" in the Frobenius norm (Assumption 1). The separation between clusters is characterized by the minimum and maximum distances, $\Delta_{\min} \triangleq \min_{j\neq j'} \|\Theta_j - \Theta_{j'}\|$ and $\Delta_{\max} \triangleq \max_{j\neq j'} \|\Theta_j - \Theta_{j'}\|$, respectively.

Control Input. The control input is composed of a stabilizing and an exploratory component, $u_t^{(i)} = u_{\text{stab}}^{(i)} + u_{\text{exp}}^{(i)}$, where $u_{\text{stab}}^{(i)} = K^{(i)} x_t^{(i)}$ is the stabilization term and $u_{\text{exp}}^{(i)} = \sigma_u g_t^{(i)}$ provides

exploration, with $g_t^{(i)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_{d_u})$. For every honest system, the *exploration* term $\sigma_u g_t^{(i)}$ ensures persistent excitation, guaranteeing $\lambda_{\min}\left(\mathbb{E}[z_t^{(i)} z_t^{(i)\top}]\right) \geq \sigma_u^2$, where $z_t^{(i)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Sigma_{z_t}^{(i)})$. We refer the reader to [Wang et al., 2023a, Lemma 1] for the definition of the covariance matrix $\Sigma_{z_t}^{(i)}$.

Dataset. We denote the local trajectory data up to time τ by

$$X_\tau^{(i)} = [x_1^{(i)}, \ldots, x_\tau^{(i)}], \text{ and } Z_\tau^{(i)} = \left[[x_0^{(i)}, u_0^{(i)}]^\top, \ldots, [x_{\tau-1}^{(i)}, u_{\tau-1}^{(i)}]^\top\right],$$

and define system i's dataset $\mathcal{D}_\tau^{(i)} = \{X_\tau^{(i)}, Z_\tau^{(i)}\}$.

Assumption 1 (Intra-cluster heterogeneity)**.** *Let $\mathcal{H}_j \subseteq \mathcal{C}_j$ denote the index set of honest systems inside cluster $\mathcal{C}_j$. There exists a scalar $\epsilon_{\text{het}} \geq 0$ such that*

$$\max_{j \in [N_c],\, i, \ell \in \mathcal{H}_j} \left\| [A_\star^{(i)}\ B_\star^{(i)}] - [A_\star^{(\ell)}\ B_\star^{(\ell)}] \right\|_F \leq \epsilon_{\text{het}}. \tag{2}$$

2.1 Certainty-Equivalent Linear Quadratic Control

We fix an honest system $i \in \mathcal{C}_j$ with $u_{\text{exp}}^{(i)} = 0$ (for now). The local control objective is to design a linear feedback controller that minimizes the infinite-horizon quadratic cost

$$J^{(i)}(K) \triangleq \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_K \left[\sum_{t=0}^{T-1} c_t^{(i)} \right], \text{ with } c_t^{(i)} \triangleq x_t^{(i)\top} (Q + K^\top R K) x_t^{(i)},$$

where $Q \succeq 0$ and $R \succ 0$. When the ground-truth dynamics $(A_\star^{(i)}, B_\star^{(i)})$ are known, the optimal controller $K_\star^{(i)} = K(A_\star^{(i)}, B_\star^{(i)}) \triangleq -\left(R + B_\star^{(i)\top} P_\star^{(i)} B_\star^{(i)}\right)^{-1} B_\star^{(i)\top} P_\star^{(i)} A_\star^{(i)}$ is obtained by solving the discrete algebraic Riccati equation (DARE), $P_\star^{(i)} = \texttt{DARE}(A_\star^{(i)}, B_\star^{(i)}, Q, R)$.

In practice, the true model parameters $(A_\star^{(i)}, B_\star^{(i)})$ are typically *unknown* and must be inferred from data. A standard approach to designing a near-optimal controller is certainty-equivalent control [Mania et al., 2019], which first estimates the system model and then computes the controller as if the estimate were exact. In particular, by using trajectory data $\mathcal{D}_\tau^{(i)}$, each system performs ordinary least-squares (OLS) estimation as

$$\widehat{\Theta}^{(i)} \triangleq [\hat{A}^{(i)}\ \hat{B}^{(i)}] = \operatorname*{argmin}_{\Theta \in \mathbb{R}^{d_x \times d'}} \left\| X_\tau^{(i)} - \Theta Z_\tau^{(i)} \right\|_F^2, \tag{3}$$

and designs a controller $\hat{K}^{(i)} = K(\hat{A}^{(i)}, \hat{B}^{(i)})$ with the estimated model.

While conceptually simple, applying CE control to a single system, suffers from poor sample efficiency. This motivates *adaptive* LQR, where learning and control are performed jointly, i.e., data are collected under the current controller, the model is updated, and the control policy is refined iteratively. Simchowitz and Foster [2020] demonstrate that a simple greedy strategy, balancing exploration through $u_{\text{exp}}^{(i)} \neq 0$ and exploitation through the data size τ, achieves the optimal expected regret scaling of $\mathcal{O}(\sqrt{dT})$. Nevertheless, as the dominant source of regret arises from system identification, doing so on a single system still requires extensive data. This limitation makes CE control particularly challenging in data-scarce scenarios [Topcu and Leve, 2022].

2.2 Multitask Adaptive Control

To overcome the limitations of system identification using data from a single system, we consider a setting in which multiple systems collaborate, with $f = 0$ adversarial systems (for now), to learn a common model that best fits their collective data. When the true cluster identities are known, multitask system identification can be performed within each cluster $\mathcal{C}_j$ by replacing (3) with

$$\widehat{\Theta}_j \triangleq [\hat{A}_j \ \hat{B}_j] = \operatorname*{argmin}_{\Theta \in \mathbb{R}^{d_x \times d'}} \frac{1}{|\mathcal{C}_j|} \sum_{i \in \mathcal{C}_j} \left\| X_\tau^{(i)} - \Theta Z_\tau^{(i)} \right\|_F^2 .$$

The resulting estimate $\widehat{\Theta}_j$ defines a shared model for the cluster, from which a common controller $\widehat{K}^{(i)} = K(\hat{A}_j, \hat{B}_j)$ is designed for all systems $i \in \mathcal{C}_j$. As proved in prior work on federated system identification [Wang et al., 2023a], when local gradient updates $G_\ell(\widehat{\Theta}_j)$ are shared and aggregated by a trusted server using simple averaging, namely, $\widehat{\Theta}_j \leftarrow \widehat{\Theta}_j + \eta F\left(\{G_\ell(\widehat{\Theta}_j)\}_{\ell \in \mathcal{C}_j}\right)$, where F denotes the averaging operator and $\eta > 0$ is the stepsize, the sample complexity improves proportionally to the number of systems in the cluster.

In practice, however, the systems' cluster identities are typically unknown. To address this, a clustering step is introduced to group systems based on similarities in their locally estimated models. This clustered system identification approach proposed in Toso et al. [2023] incurs an additional but exponentially decaying misclassification error, as clustering accuracy improves with the amount of collected data. The complete multitask adaptive control pipeline comprising data collection, clustering, multitask system identification, and control design is illustrated in Figure 1.

2.3 Adversarial Systems

Finally, we are interested in the setting where some systems may behave adversarially, transmitting corrupted or malicious model updates to the server during aggregation. To guarantee recoverability of the true cluster models, we assume an *honest majority* within each cluster, i.e., $f_j < M_j/2$, where M_j denotes the total number of systems in cluster $\mathcal{C}_j$, otherwise, reconstruction of the true model becomes information-theoretically impossible. In such cases, simple averaging of model updates fails, as a fraction of corrupted gradients can arbitrarily skew the aggregate and compromise convergence. To mitigate this, we leverage resilient aggregators that are widely studied in adversarial federated and distributed learning [Guerraoui et al., 2024].

Definition 2.1 ((f, λ)-Resilient aggregation - Adapted from Farhadkhani et al. [2022])**.** *Given a scalar $\lambda \geq 0$, an aggregation rule F is said to be (f, λ)-resilient if, for any collection of matrices $G_1, \ldots, G_M \in \mathbb{R}^{d_x \times d'}$ and any subset of honest systems $\mathcal{H} \subseteq \{1, \ldots, M\}$ with $|\mathcal{H}| = m$,*

$$\left\| F(G_1, \ldots, G_M) - \bar{G} \right\| \leq \lambda \max_{i,j \in \mathcal{H}} \|G_i - G_j\|, \ \text{with } \bar{G} := \frac{1}{|\mathcal{H}|} \sum_{i \in \mathcal{H}} G_i.$$

Remark 2.1. *(f, λ)-resilient aggregation ensures that the aggregated update remains close to the mean of the honest systems, up to a factor λ capturing their dispersion. This guarantees robustness even when up to $f_j < M_j/2$ systems act adversarially. Many classical robust aggregation rules satisfy this property, including the* coordinate-wise trimmed mean *(CWTM),* coordinate-wise median *(CWMed),* geometric median *(GM),* minimum diameter averaging *(MDA), and* mean-around-median *(MeaMed) [Farhadkhani et al., 2022]. Typically, MDA and CWTM achieve $\lambda = \mathcal{O}(f_j/m_j)$, while GM and CWMed yield $\lambda = \mathcal{O}(1)$; the latter can be adjusted via pre-aggregation techniques such as nearest-neighbor mixing [Allouah et al., 2023] or bucketing [Karimireddy et al., 2020].*

Remark 2.2. *We assume that each cluster satisfies the honest-majority condition, i.e., $f_j < M_j/2$ for all $j \in [N_c]$. This standard assumption in robust aggregation [Farhadkhani et al., 2022] guarantees convergence to the true cluster model even when a subset of systems transmit corrupted updates. Extending the framework to cases where adversarial systems can also misreport their cluster identities, thereby violating the honest-majority assumption, is left for future work. Promising directions include (i) randomized warm-up clustering steps to prevent adversarial concentration within specific clusters and (ii) robust clustering methods such as iterative filtering [Ghosh et al., 2019].*

Goal. The aim of adversarially robust multitask adaptive control is to cluster systems and design controllers that remain performant under heterogeneity and adversarial behavior.

Our analysis quantifies the benefit of collaboration and the effects arising from clustering errors, intra-cluster heterogeneity, and adversarial behavior. We evaluate performance using the standard notion of cumulative regret from online learning [Abbasi-Yadkori and Szepesvári, 2011], defined as the difference between the cumulative cost incurred by the adaptive controller and that of the optimal controller under the true model: $\mathcal{R}_T^{(i)} = \sum_{t=1}^{T} \left(c_t^{(i)} - J^{(i)}(K_\star^{(i)})\right)$, where $c_t^{(i)}$ denotes the immediate cost for playing a suboptimal controller $\hat{K}^{(i)}$.

Algorithm 1 Multitask Certainty-Equivalent Control

Initialize: $\hat{K}_1^{(i)} \leftarrow K_0^{(i)}\ \forall i \in [M]$, $T \leftarrow \tau_1 2^{k_{\mathsf{fin}}-1}$
for $k = 1, 2, \ldots, k_{\mathsf{fin}}$
  for all systems $i = 1, \ldots, M$ **(in parallel)** `// Data collection`
    for $t = \tau_{k-1}, \tau_{k-1}+1, \ldots, \tau_k$
      If $\|x_t^{(i)}\|^2 \geq x_b^2 \log T$ or $\|\hat{K}_k^{(i)}\| \geq K_b$
        Abort and play $K_0^{(i)}$ forever
      Play $u_t^{(i)} = \hat{K}_k^{(i)} x_t^{(i)} + \sigma_k g_t^{(i)}$ `// Exploration`
    end for
  end for
  $\widehat{\Theta}_k^{(1:M)} \leftarrow \texttt{RCSI}(\widehat{\Theta}_{k-1}^{(1:M)}, \mathcal{D}_{\tau_k}^{(1:M)}, N_c, N, \eta)$ `// Robust SysID`
  Update the controller: $\hat{K}_{k+1}^{(i)} \leftarrow K(\widehat{\Theta}_k^{(1:M)}), \forall i \in [M]$ `// Controller update`
  $\tau_{k+1} \leftarrow 2\tau_k$
end for

Algorithm 2 `RCSI`: Robust Clustered System Identification

for $n = 0, 1, \ldots, N-1$ **do**
  Systems receive **all** the current cluster estimated models $\{\widehat{\Theta}_j\}_{j\in[N_c]}$
    for $i = 1, \ldots, M$ **(in parallel) do** `// Cluster identity estimation`
      $\hat{j} = \operatorname{argmin}_{j\in[N_c]} \|X^{(i)} - \widehat{\Theta}_j Z^{(i)}\|^2$
    end for
  Construct the **cluster identity** set $\mathcal{C}_{\hat{j}}$
  $G_\ell(\widehat{\Theta}^{(i)}) \leftarrow (X^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}) Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1}, \forall \ell \in \mathcal{C}_{\hat{j}}$ `// Update`
    for $i = 1, \ldots, M$ **(in parallel) do** `// Model estimation`
      $\widehat{\Theta}^{(i)} \leftarrow \widehat{\Theta}^{(i)} + \eta F\left(\left\{G_\ell(\widehat{\Theta}^{(i)})\right\}_{\ell\in\mathcal{C}_{\hat{j}}}\right)$ `// Aggregation`
    end for
end for

Algorithm 1 summarizes the adversarially robust multitask adaptive control procedure. Each system alternates between three phases: (i) data collection under the current controller, (ii) clustered robust system identification via resilient aggregation, and (iii) controller update using the estimated model. At each epoch k, all systems collect local trajectory data $\mathcal{D}_{\tau_k}^{(i)}$ in parallel. If the state or controller norm exceeds a threshold (i.e., x_b and K_b), the system switches to a fallback stabilizing controller $K_0^{(i)}$ to ensure bounded trajectories and regret. The identification step, implemented by `RCSI` (Algorithm 2), jointly performs clustering and robust aggregation: systems estimate their cluster identities, send local gradients $G_\ell(\widehat{\Theta}^{(i)})$, and the server aggregates them via a resilient rule F. The estimated model then defines the CE controller $\hat{K}_{k+1}^{(i)} = K(\hat{A}_k^{(i)}, \hat{B}_k^{(i)})$ for the next iteration.

We emphasize that a stabilizing initial controller $K_0^{(i)}$ is required for each system to ensure the initial trajectories remain bounded. Such controllers can be obtained using data-driven stabilization methods [Perdomo et al., 2021, Toso et al., 2025]. Moreover, Algorithm 1 operates under the following two assumptions. First, we assume that the initial model estimate is sufficiently accurate to ensure consistent clustering [Toso et al., 2023].

Assumption 2. *The initial model satisfies* $\|\widehat{\Theta}_0^{(i)} - \Theta_\star^{(i)}\| \le C_\alpha \Delta_{min}$, *for some constant* $C_\alpha \in (0, \frac{1}{2})$.

Second, we impose the following bounds on the state and controller norms.

Assumption 3. *We assume that the state and controller bounds are*

$$x_b \ge 400 \left(P_0^\vee\right)^2 \Psi_B^\vee \sigma_w \sqrt{d'}, \text{ and } K_b \ge \sqrt{P_0^\vee}.$$

where $\Psi_B^{(i)} \triangleq \max\left\{1, \|B_\star^{(i)}\|\right\}$, $\Psi_B^\vee \triangleq \max_i \Psi_B^{(i)}$, $P_0^\vee \triangleq \max_i \left\|P_{K_0^{(i)}}^{(i)}\right\|$, *with* $P_K^{(i)}$ *denoting the solution to the discrete Lyapunov equation,* $P_K^{(i)} \triangleq \texttt{dlyap}(A_\star^{(i)} + B_\star^{(i)} K, Q + K^\top R K)$

We present our theoretical analysis progressively across three settings, building from the simplest to the most general case: (i) intra-cluster homogeneity ($f = 0$, $\epsilon_{\text{het}} = 0$), (ii) bounded heterogeneity ($f = 0$, $\epsilon_{\text{het}} > 0$), and (iii) adversarial systems ($f > 0$, $\epsilon_{\text{het}} > 0$). For each setting, we first establish estimation error bounds and subsequently the corresponding regret bounds.

3 Error Bounds for Multitask System Identification

We are now ready to establish error bounds for the clustered multitask identification procedure described in Algorithm 2. This section characterizes how collaboration, heterogeneity, and adversarial behavior influence the estimation error of the local system models. Complete proofs and constant definitions are provided in Appendix E. For clarity of presentation, we introduce the following key quantities: $r^\vee \triangleq \max_{t\in[T],\ell\in\mathcal{C}_j} \operatorname{tr}\left(\Sigma_{z_t}^{(\ell)}\right)/\|\Sigma_{z_t}^{(\ell)}\|$, $P_\star^\wedge \triangleq \min_{i\in[M]} \|P_\star^{(i)}\|$, $C_\tau = \texttt{poly}(d_x, d_u)$, and the contraction rate $\rho = 1 - \eta$, for some step-size $\eta \in (0, 1)$.

Lemma 3.1 (Intra-cluster homogeneity)**.** *Consider a cluster* $\mathcal{C}_j$ *composed of* M_j *identical systems. Suppose that Assumption 2 holds and that the initial epoch length is set sufficiently large, such that*

$$\tau_1 \ge C_\tau \max\left\{\frac{d'\Delta_{\min}^2 \sigma_w \log\left(\frac{(2d_x+d_u)M_j}{\delta}\right)}{dM_j}, r^\vee + \log\left(\frac{2M_j}{\delta}\right), \log(4T^2), \frac{\log\frac{1}{P_\star^\wedge}}{\log\left(1-\frac{1}{P_\star^\wedge}\right)}\right\},$$

for a small $\delta \in (0,1)$. Then, after $N \geq \log\left(C_\alpha \Delta_{min} \tau_1^2\right) / \log(1/\rho)$ iterations of `RCSI`*, with probability $1-\delta$, the model estimate $\widehat{\Theta}^{(i)}$ at epoch k satisfies:*

$$\left\|\widehat{\Theta}_k^{(i)} - \Theta_\star^{(i)}\right\|_F^2 \leq C_{\text{stat}} \frac{\sigma_w^2 (d_x^2 + d_x d_u) \log(M_j/\delta)}{\sigma_k^2 M_j \tau_k} + C_{\text{mis},1}\, e^{-C_{\text{mis},2}\sigma_k^2 \tau_k}.$$

Proposition 3.1 (Intra-cluster heterogeneity)**.** *Consider a cluster $\mathcal{C}_j$ consisting of M_j similar but non-identical systems with bounded intra-cluster heterogeneity as in Assumption 1. Suppose the initial epoch length τ_1 and the number of iterations N in* `RCSI` *are chosen as in Lemma 3.1. Then, for a small $\delta \in (0,1)$, the model estimate $\widehat{\Theta}^{(i)}$ at epoch k satisfies:*

$$\left\|\widehat{\Theta}_k^{(i)} - \Theta_\star^{(i)}\right\|_F^2 \leq C_{\text{stat}} \frac{\sigma_w^2 (d_x^2 + d_x d_u) \log(M_j/\delta)}{\sigma_k^2 M_j \tau_k} + C_{\text{het}} \epsilon_{\text{het}}^2 + C_{\text{mis},1} e^{-C_{\text{mis},2}\sigma_k^2 \tau_k},$$

with probability at least $1-\delta$.

Lemma 3.2 (Adversarial systems)**.** *Consider a cluster $\mathcal{C}_j$ containing M_j similar but non-identical systems, among which at most $f_j < M_j/2$ are adversarial and $m_j = M_j - f_j$ are honest. Assume the aggregation rule is (f_j, λ)-resilient as defined in Definition 2.1, and that Assumption 1 holds. Suppose further that the initial epoch length τ_1 and the number of iterations N in* `RCSI` *are chosen as in Lemma 3.1. Then, for a small $\delta \in (0,1)$, the model estimate $\widehat{\Theta}^{(i)}$ at epoch k satisfies:*

$$\left\|\widehat{\Theta}_k^{(i)} - \Theta_\star^{(i)}\right\|_F^2 \leq C_{\text{stat}} \frac{\sigma_w^2 (d_x^2 + d_x d_u) \log(M_j/\delta)}{\sigma_k^2 \tau_k} \left(\frac{1}{m_j} + \lambda^2 d_x\right) + C_{\text{het}} (1+\lambda)^2 \epsilon_{\text{het}}^2 + C_{\text{mis},1} e^{-C_{\text{mis},2}\sigma_k^2 \tau_k}, \ \ w.p.\ 1-\delta.$$

Discussion. Taken together, Lemma 3.1, Proposition 3.1, and Lemma 3.2 provide a unified characterization of the error bounds for the multitask system identification step under adversarial and heterogeneous systems. Under intra-cluster homogeneity, the estimation error scales inversely with both the number of systems and the data length (epoch size), achieving $\mathcal{O}(\sigma_w^2/(M_j\tau_k))$ consistency. Introducing heterogeneity adds a fixed bias proportional to ϵ_{het}^2, while adversarial systems introduce an additional λ-scaled term determined by the resilience coefficient of the aggregation rule. The proofs of these results are provided in Appendix E, where the estimation error is controlled using the matrix Hoeffding inequality [Tropp, 2011]. These results underpin the regret analysis in Section 4, where improved estimation accuracy translates into a reduction in the expected regret.

4 Regret Analysis

We are now in place to establish the regret bounds for Algorithm 1 under the three settings of interest: (i) intra-cluster homogeneity, (ii) intra-cluster heterogeneity, and (iii) the presence of adversarial systems. We quantify how clustering accuracy, intra-cluster similarity, and adversarial robustness jointly affect the regret of the CE controller for any honest system within a given cluster $\mathcal{C}_j$. Complete derivations are provided in Appendix G. To clarify exposition, we introduce the following quantities: $\Omega_1 \triangleq 142 C_{\text{stat}} \|P_\star^{(i)}\|^8 \sigma_w^2 \log((M_j T)/\delta) + 2d_u(1 + 2\|P_\star^{(i)}\| \Psi_B^{(i)2})$, $\Omega_2 \triangleq 3d_x \left\|P_{K_0^{(i)}}^{(i)}\right\| \Psi_B^{(i)2} + 2x_b^2 \left\|P_\star^{(i)}\right\|$, $\Omega_3 \triangleq 142 C_{\text{mis},1} \|P_\star^{(i)}\|^8$, and $\Omega_4 \triangleq 142 C_{\text{het}} \|P_\star^{(i)}\|^8$.

Theorem 4.1 (Intra-cluster homogeneity). *Fix a system i belonging to a homogeneous cluster $\mathcal{C}_j$ of size M_j. Let Assumption 2 hold and consider the doubling-epoch schedule $\tau_k = 2^{k-1}\tau_1$ with initial epoch length τ_1 and number of* `RCSI` *iterations N as in Lemma 3.1, with exploration sequence $\sigma_k^2 = \frac{\sqrt{d}}{d_x^2 + d_x d_u}\frac{1}{\sqrt{\tau_k M_j}}$. Then the expected regret of any system $i \in \mathcal{C}_j$ under Algorithm 2 satisfies:*

$$\mathbb{E}\left[\mathcal{R}_T^{(i)}\right] \leq \Omega_1\sqrt{\tfrac{dT}{M_j}} + \Omega_2(\log T)^2 + \Omega_3 T e^{-C_{\mathrm{mis},2}\sqrt{\frac{\tau_1}{M_j}}}.$$

Corollary 4.1 (Intra-cluster heterogeneity). *Consider a system i belonging to a heterogeneous cluster $\mathcal{C}_j$ of size M_j. Suppose Assumptions 1 and 2 hold. Then, under the same conditions on the epoch length τ_1, number of* `RCSI` *iterations N, and exploration sequence σ_k^2 in Theorem 4.1, the expected regret for any system $i \in \mathcal{C}_j$ satisfies:*

$$\mathbb{E}\left[\mathcal{R}_T^{(i)}\right] \leq \Omega_1\sqrt{\tfrac{dT}{M_j}} + \Omega_2(\log T)^2 + \Omega_3 T e^{-C_{\mathrm{mis},2}\sqrt{\frac{\tau_1}{M_j}}} + \Omega_4 T\epsilon_{\mathrm{het}}^2.$$

Theorem 4.2 (Adversarial systems). *Let each cluster $\mathcal{C}_j$ satisfy honest-majority $f_j < \frac{M_j}{2}$. Suppose that the aggregation rule is (f_j, λ)-resilient as defined in Definition 2.1. Under the conditions on the epoch length τ_1 and number of* `RCSI` *iterations N as in Lemma 3.1, with exploration sequence $\sigma_k^2 = \frac{\sqrt{d_u^2 d_x}}{d_x^2 + d_x d_u}\sqrt{\frac{1+\lambda^2 d_x m_j}{\tau_k m_j}}$. Then, for any honest system in $\mathcal{C}_j$, the expected regret satisfies:*

$$\mathbb{E}\left[\mathcal{R}_T^{(i)}\right] \leq \Omega_1\sqrt{\tfrac{dT(1+\lambda^2 d_x m_j)}{m_j}} + \Omega_2(\log T)^2 + \Omega_3 T\, e^{-C_{\mathrm{mis},2}\sqrt{\frac{(1+\lambda^2 d_x m_j)\tau_1}{m_j}}} + \Omega_4 T(1+\lambda)^2\epsilon_{\mathrm{het}}^2.$$

Discussion. Theorem 4.1 demonstrates that, under intra-cluster homogeneity, multitask adaptive control achieves an $\mathcal{O}(1/\sqrt{M_j})$ improvement in the leading term of the regret compared to the single-system case [Simchowitz and Foster, 2020]. This scaling confirms that jointly estimating a common model for identical systems effectively reduces the regret of adaptive LQR control. Corollary 4.1 extends this result to heterogeneous clusters, where a residual bias term proportional to $\epsilon_{\mathrm{het}}^2$ appears in the regret. This term reflects the unavoidable model mismatch within each cluster, introducing an additional component linear in T. When ϵ_{het} is sufficiently small, the regret remains dominated by the $\widetilde{\mathcal{O}}(\sqrt{T/M_j})$ term, thus preserving the benefit of collaboration.

Finally, Theorem 4.2 demonstrates that the collaboration benefit persists even in the presence of adversarial systems, provided the honest-majority condition holds and the aggregation rule is (f_j, λ)-resilient. The additional factor $(1 + \lambda^2 d_x m_j)$ quantifies the robustness cost introduced by resilient aggregation. As discussed earlier, the resilience coefficient λ typically scales as $\mathcal{O}(f_j/m_j)$, thereby preserving the regret reduction under adversarial systems. Taken together, these results highlight the robustness of multitask adaptive control under heterogeneous and adversarial systems.

4.1 Proof Idea

The proof of Theorem 4.2 follows a standard regret decomposition [Cassel et al., 2020] into three parts: (i) regret under the success event, when Algorithm 1 does not abort and Lemma 3.2 holds; (ii) regret under failure; and (iii) regret from the first epoch. We prove that the success event occurs with high probability, at least $1 - T^{-2}$ (see Appendix F), ensuring that (i) dominates.

Under success, the regret scales as $\mathcal{O}\big((\tau_k - \tau_{k-1})(\|\widehat{\Theta}_k^{(i)} - \Theta_\star^{(i)}\|_F^2 + \sigma_k^2)\big)$, where estimation errors propagate via the Lyapunov bound on $P_\star^{(i)}$, linking model inaccuracy to cost difference [Simchowitz and Foster, 2020, Theorem 3]. Balancing exploration (σ_k) and exploitation (τ_k) in the error bound of Lemma 3.2, and summing over exponentially growing epochs, yields the dominant $\widetilde{\mathcal{O}}(\sqrt{T/m_j})$ term, with additive effects from resilient aggregation, heterogeneity, and clustering misclassification. The $\mathcal{O}((\log T)^2)$ term arises from the initial epoch. The regret under failure is negligible.

5 Numerical Validation

We consider multiple unicycle robots with details on the dynamics and implementation[2] deferred to Appendix B. Figure 2 summarizes the performance of Algorithm 1 across six panels, highlighting the effects of collaboration, heterogeneity, clustering accuracy, and adversarial behavior. In the top-left panel, homogeneous clusters show decreasing regret with more systems per cluster, matching the $\widetilde{\mathcal{O}}(\sqrt{T/M_j})$ scaling in Theorem 4.1. The top-middle panel introduces intra-cluster heterogeneity, where regret increases due to the $\mathcal{O}(\epsilon_{\text{het}}^2 T)$ term (Corollary 4.1); estimation accuracy initially improves with collaboration but saturates at a bias floor proportional to ϵ_{het}^2.

The top-right panel reports clustering accuracy, showing exponentially decaying misclassification with more samples and faster convergence for larger clusters, consistent with the $e^{-C_{\text{mis},2}\sigma_k^2\tau_k}$ rate. The bottom-left panel presents the adversarial setting, where a fraction $\rho_{\text{byz}} = f_j/M_j$ of systems send corrupted updates (see Appendix B). Increasing $\rho_{\text{byz}} \in \{10\%, 20\%, 30\%\}$ preserves sublinear regret for $\rho_{\text{byz}} < 0.5$, in line with the $\mathcal{O}(1 + \lambda^2 d_x)$ inflation predicted by Theorem 4.2.

The bottom-middle panel compares two resilient aggregation rules under identical adversarial contamination. We evaluate the trimmed mean and geometric median aggregators, which differ in their resilience coefficients λ. Both converge, confirming the (f_j, λ)-resilience property; consistent with Allouah et al. [2023], the trimmed mean achieves slightly lower regret due to a smaller λ. A broader empirical study across alternative aggregation schemes is left for future work.

The bottom-right panel compares Algorithm 1 with the multitask representation learning method of Lee et al. [2025] (denoted here by `RepL`). We evaluate both under (i) structurally homogeneous systems, where a shared representation exists, and (ii) a heterogeneous configuration with an added system violating this assumption. While `RepL` performs well in case (i), it deteriorates sharply without a shared model representation (ii). In contrast, Algorithm 1, though affected by intra-cluster heterogeneity, confines learning to similar systems and achieves lower regret under heterogeneity.

6 Future Work

This work provides the first analysis of multitask adaptive control under heterogeneous and adversarial systems, opening several promising directions. Our regret bounds reveal a non-vanishing heterogeneity bias ϵ_{het}, scaling as $\mathcal{O}(\epsilon_{\text{het}}^2 T)$, which may limit collaboration when intra-cluster heterogeneity is large. A natural extension is to integrate representation learning within clusters to mitigate this bias and remove the interplay between heterogeneity and adversarial effects $\mathcal{O}(\lambda^2 \epsilon_{\text{het}}^2)$. Another direction is developing robust clustering methods to handle cases where adversarial systems corrupt cluster identities. Extending this work to nonlinear dynamics and non-quadratic objectives also remains an important avenue for future work.

[2] Code available at https://github.com/jd-anderson/multi_task_adaptive_control

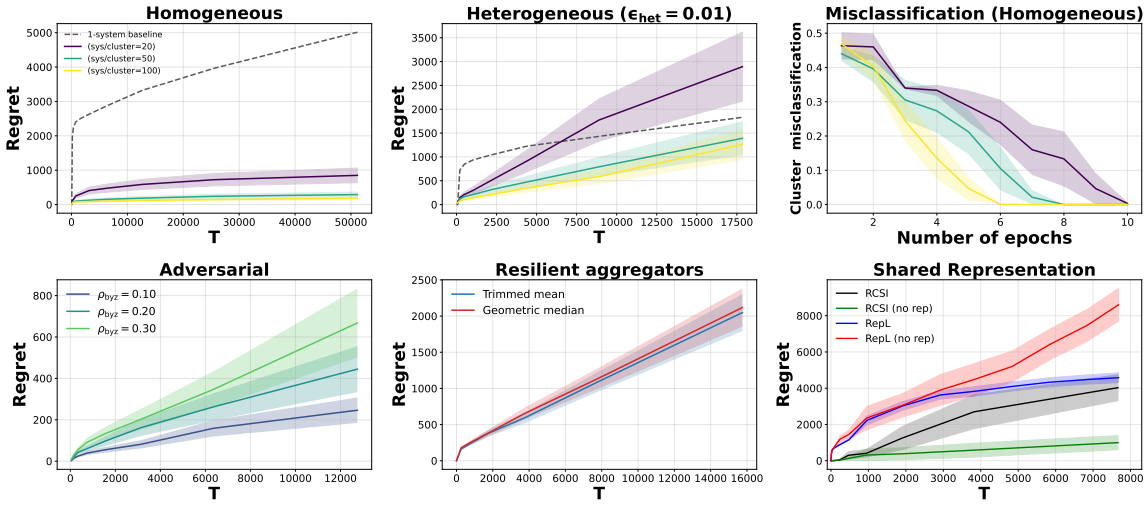

Figure 2: Illustration of main results. Top row–(left) homogeneous clusters; (middle) intra-cluster heterogeneity; (right) misclassification rate. Bottom row–(left) adversarial systems; (middle) robust aggregation comparison; (right) comparison with multitask representation learning.

7 Acknowledgments

Leonardo F. Toso thanks Rafael Pinot and Nirupam Gupta for instructive discussions on adversarial machine learning. Leonardo F. Toso is funded by the Center for AI and Responsible Financial Innovation (CAIRFI) Fellowship and by the Columbia Presidential Fellowship. James Anderson is partially funded by NSF grants ECCS 2144634 and 2231350 and the Columbia Center of AI Technology in collaboration with Amazon.

References

Karl Johan Åström and Björn Wittenmark. On Self Tuning Regulators. *Automatica*, 9(2):185–199, 1973.

Karl Johan Åström. Theory and Applications of Adaptive Control—a Survey. *Automatica*, 19(5): 471–486, 1983.

Yasin Abbasi-Yadkori and Csaba Szepesvári. Regret Bounds for the Adaptive Control of Linear Quadratic Systems. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 1–26. JMLR Workshop and Conference Proceedings, 2011.

Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. Regret Bounds for Robust Adaptive Control of the Linear Quadratic Regulator. *Advances in Neural Information Processing Systems*, 31, 2018.

Alon Cohen, Tomer Koren, and Yishay Mansour. Learning Linear-Quadratic Regulators Efficiently with Only $\sqrt{T}$ Regret. In *International Conference on Machine Learning*, pages 1300–1309. PMLR, 2019.

Max Simchowitz and Dylan Foster. Naive Exploration is Optimal for Online LQR. In *International Conference on Machine Learning*, pages 8937–8948. PMLR, 2020.

Elad Hazan, Sham Kakade, and Karan Singh. The Nonstochastic Control Problem. In *Algorithmic Learning Theory*, pages 408–421. PMLR, 2020.

Ingvar Ziemann and Henrik Sandberg. Regret Lower Bounds for Learning Linear Quadratic Gaussian Systems. *arXiv preprint arXiv:2201.01680*, 2022.

Lei Xin, Lintao Ye, George Chiu, and Shreyas Sundaram. Identifying the Dynamics of a System by Leveraging Data from Similar Systems. In *2022 American Control Conference (ACC)*, pages 818–824. IEEE, 2022.

Han Wang, Leonardo F. Toso, and James Anderson. FedSysID: A Federated Approach to Sample-Efficient System Identification. In *Learning for Dynamics and Control Conference*, pages 1308–1320. PMLR, 2023a.

Leonardo F Toso, Han Wang, and James Anderson. Learning Personalized Models with Clustered System Identification. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 7162–7169. IEEE, 2023.

Ertuğrul Keçeci, Müjde Güzelkaya, and Tufan Kumbasar. FedAlign: Federated Learning with State Alignment for System Identification. In *2025 International Conference on Control, Automation and Diagnosis (ICCAD)*, pages 1–6. IEEE, 2025a.

Thomas TCK Zhang, Leonardo Felipe Toso, James Anderson, and Nikolai Matni. Sample-Efficient Linear Representation Learning from Non-IID Non-Isotropic Data. In *The Twelfth International Conference on Learning Representations*, 2024.

Han Wang, Leonardo F. Toso, Aritra Mitra, and James Anderson. Model-free Learning with Heterogeneous Dynamical Systems: A Federated LQR Approach. *arXiv preprint arXiv:2308.11743*, 2023b.

Leonardo Felipe Toso, Donglin Zhan, James Anderson, and Han Wang. Meta-Learning Linear Quadratic Regulators: A Policy Gradient MAML Approach for Model-Free LQR. In *6th Annual Learning for Dynamics & Control Conference*, pages 902–915. PMLR, 2024.

Lirui Wang, Kaiqing Zhang, Allan Zhou, Max Simchowitz, and Russ Tedrake. Fleet Policy Learning via Weight Merging and An Application to Robotic Tool-Use. *arXiv preprint arXiv:2310.01362*, 2023c.

Bruce D Lee, Leonardo F Toso, Thomas T Zhang, James Anderson, and Nikolai Matni. Regret Analysis of Multi-Task Representation Learning for Linear-Quadratic Adaptive Control. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 18062–18070, 2025.

Sadegh Farhadkhani, Rachid Guerraoui, Nirupam Gupta, Rafael Pinot, and John Stephan. Byzantine Machine Learning Made Easy by Resilient Averaging of Momentums. In *International Conference on Machine Learning*, pages 6246–6283. PMLR, 2022.

Ertuğrul Keçeci, Müjde Güzelkaya, and Tufan Kumbasar. Redefining Clustered Federated Learning for System Identification: The Path of ClusterCraft. *arXiv preprint arXiv:2505.16857*, 2025b.

Peva Blanchard, El Mahdi El Mhamdi, Rachid Guerraoui, and Julien Stainer. Machine Learning with Adversaries: Byzantine Tolerant Gradient Descent. *Advances in Neural Information Processing Systems*, 30, 2017.

Lingjiao Chen, Hongyi Wang, Zachary Charles, and Dimitris Papailiopoulos. Draco: Byzantine-Resilient Distributed Training via Redundant Gradients. In *International Conference on Machine Learning*, pages 903–912. PMLR, 2018.

Xingrong Dong, Zhaoxian Wu, Qing Ling, and Zhi Tian. Byzantine-Robust Distributed Online Learning: Taming Adversarial Participants in an Adversarial Environment. *IEEE Transactions on Signal Processing*, 72:235–248, 2023.

Bruce D Lee, Anders Rantzer, and Nikolai Matni. Nonasymptotic Regret Analysis of Adaptive Linear Quadratic Control with Model Misspecification. *arXiv preprint arXiv:2401.00073*, 2023.

Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47. Cambridge University Press, 2018.

Horia Mania, Stephen Tu, and Benjamin Recht. Certainty Equivalence is Efficient for Linear Quadratic Control. *Advances in Neural Information Processing Systems*, 32, 2019.

Ufuk Topcu and Frederick Leve. Learning for Dynamical Systems when Data are Scarce. *Collections*, 55(09), 2022.

Rachid Guerraoui, Nirupam Gupta, and Rafael Pinot. *Robust Machine Learning*. Springer, 2024.

Youssef Allouah, Sadegh Farhadkhani, Rachid Guerraoui, Nirupam Gupta, Rafaël Pinot, and John Stephan. Fixing by Mixing: A Recipe for Optimal Byzantine ML under Heterogeneity. In *International Conference on Artificial Intelligence and Statistics*, pages 1232–1300. PMLR, 2023.

Sai Praneeth Karimireddy, Lie He, and Martin Jaggi. Byzantine-Robust Learning on Heterogeneous Datasets via Bucketing. *arXiv preprint arXiv:2006.09365*, 2020.

Avishek Ghosh, Justin Hong, Dong Yin, and Kannan Ramchandran. Robust Federated Learning in a Heterogeneous Environment. *arXiv preprint arXiv:1906.06629*, 2019.

Juan Perdomo, Jack Umenberger, and Max Simchowitz. Stabilizing Dynamical Systems via Policy Gradient Methods. *Advances in Neural Information Processing Systems*, 34:29274–29286, 2021.

Leonardo F Toso, Lintao Ye, and James Anderson. Learning Stabilizing Policies via an Unstable Subspace Representation. *arXiv preprint arXiv:2505.01348*, 2025.

Joel A. Tropp. User-Friendly Tail Bounds for Sums of Random Matrices. *Foundations of Computational Mathematics*, 12(4):389–434, aug 2011.

Asaf Cassel, Alon Cohen, and Tomer Koren. Logarithmic Regret for Learning Linear Quadratic Regulators Efficiently. In *International Conference on Machine Learning*, pages 1328–1337. PMLR, 2020.

Bin Yu. Rates of Convergence for Empirical Processes of Stationary Mixing Sequences. *The Annals of Probability*, pages 94–116, 1994.

Appendix

Table of Contents

A Appendix Roadmap

The appendix is organized as follows. Appendix B provides additional experiments and details on the experimental setup used in Section 5 to validate and illustrate our theoretical guarantees. Appendix C presents the detailed regret decomposition into three components: (i) the contribution corresponding to the success of Algorithm 1 (i.e., no abortion), (ii) the contribution associated with its failure, and (iii) the term arising from initialization with a suboptimal stabilizing controller. In Appendix D, we revisit key auxiliary results, including matrix concentration inequalities from Tropp [2011], regret bounds from Lee et al. [2023], and the assumption that the state-input covariates are distributed according to a β-mixing stationary processes, we refer the reader to Yu [1994] for the definition of such geometric processes. Appendix E presents error bounds for multitask system identification under three settings: (i) intra-cluster homogeneity, (ii) intra-cluster heterogeneity, and (iii) adversarial systems. Appendix F establishes that the success event under which Algorithm 1 does not abort holds with high probability. Finally, Appendix G provides the regret analysis, leveraging the estimation error bounds and success-event probability to derive the regret bounds.

B Numerical Implementation Details

System dynamics. All experiments were conducted on discrete and linear time-invariant (LTI) systems of the form

$$x_{t+1}^{(i)} = A_\star^{(i)} x_t^{(i)} + B_\star^{(i)} u_t^{(i)} + w_t^{(i)}, \text{ with noise } w_t^{(i)} \sim \mathcal{N}(0, \sigma_w^2 I),$$

where $x_t \in \mathbb{R}^3$ is the system state, $u_t \in \mathbb{R}^2$ is the control input, and $(A_\star^{(i)}, B_\star^{(i)})$ denote the dynamics of the i-th subsystem. Each subsystem is derived from a discrete-time unicycle model with kinematics $\dot{p}_x^{(i)} = v^{(i)} \cos\theta^{(i)}$, $\dot{p}_y^{(i)} = v^{(i)} \sin\theta^{(i)}$, $\dot{\theta}^{(i)} = \omega^{(i)}$, where the state is $x^{(i)} = (p_x^{(i)}, p_y^{(i)}, \theta^{(i)})$ and the control input is $u^{(i)} = (v^{(i)}, \omega^{(i)})$. Here, $(p_x^{(i)}, p_y^{(i)})$ denotes the position, $\theta^{(i)}$ the orientation, $v^{(i)}$ the forward velocity, and $\omega^{(i)}$ the yaw rate of the i-th robot. The dynamics are linearized around different nominal operating points $(v_{0,j}, \theta_{0,j}) \in \{(1.0, 0°), (1.0, 45°), (0.8, 90°)\}$ and discretized (with Euler step $\Delta t = 0.1$) to obtain the cluster models:

$$A_j = \begin{bmatrix} 1 & 0 & -\Delta t v_{0,j} \sin\theta_{0,j} \\ 0 & 1 & \Delta t v_{0,j} \cos\theta_{0,j} \\ 0 & 0 & 1 \end{bmatrix}, \quad B_j = \begin{bmatrix} \Delta t \cos\theta_{0,j} & 0 \\ \Delta t \sin\theta_{0,j} & 0 \\ 0 & \Delta t \end{bmatrix},$$

for each cluster $j \in \{1, 2, 3\}$, heterogeneous systems are created by adding Gaussian perturbations to the nominal parameters:

$$A_\star^{(i)} = A_j + \varepsilon_{ij}, \; B_\star^{(i)} = B_j + \delta_{ij},$$

where ε_{ij} and δ_{ij} are elementwise Gaussian perturbations with zero mean and variance ϵ_{het}. Unless otherwise stated, $\epsilon_{\text{het}} = 0.01$, which corresponds to a relatively strong heterogeneity level since the smallest inter-system coefficient difference among the nominal cluster models is on the order of 0.04. The process noise variance was fixed to $\sigma_w^2 = 0.01$. We set $Q = I_{d_x}$ and $R = I_{d_u}$.

Experimental configurations. In the first three panels (*homogeneous*, *heterogeneous*, and *misclassification*) in Figure 2, we used the above three cluster models and their perturbed systems

with $\epsilon_{\text{het}} = 0.01$. The homogeneous case is obtained by setting $\epsilon_{\text{het}} = 0$ for all systems. In the the misclassification experiment we use the same configuration as the homogeneous case.

For the fourth experiment (adversarial setting), we considered the same system clusters and heterogeneity level ($\epsilon_{\text{het}} = 0.01$) with 20 systems per cluster. Adversarial systems were randomly selected with corruption ratio $\rho_{\text{byz}} \in \{0, 0.15, 0.35\}$ to evaluate robustness.

In the fifth experiment (aggregation rules comparison), the same cluster models were used with 50 systems per cluster, and $\epsilon_{\text{het}} = 0.01$, and adversarial ratio $\rho_{\text{byz}} = 0.15$. Two aggregation rules were compared: (i) trimmed-mean with $\alpha_{\text{trim}} = \rho_{\text{byz}} + 0.05$, and (ii) geometric-median aggregation. Both methods were implemented using the `RCSI` update under identical simulation conditions.

The final experiment corresponds to the shared-representation setting. Here, we used 26 systems belonging to two clusters (see dataset files `A_norep.pkl` and `B_norep.pkl`). The first 25 systems share a common representation and were treated as belonging to a single cluster, while the 26-th system was constructed to violate this assumption, representing a structurally distinct system. We ran both Algorithm 1 and the multitask representation learning method from Lee et al. [2025] (denoted here by `RepL`) using identical algorithm hyperparameters:

$$\tau_1 = 15,\ k_{\text{fin}} = 9,\ x_b = 25,\ K_b = 15,\ \text{with } \epsilon_{\text{het}} = 0.004.$$

For both methods, the exploration sequence was geometrically decaying as

$$\sigma_{u,k} \in \{0.663, 0.411, 0.124, 0.037, 0.011, 0.0034, 0.0010, 0.0010\}.$$

Adversarial implementation. To simulate adversarial systems, each system i is assigned a binary corruption flag with probability ρ_{byz}. Corrupted systems replace their empirical regression statistic (XZ_i, ZZ_i) with an adaptive convex combination:

$$XZ_i^{(\text{byz})} = (1 - \beta)XZ_i + \beta(\Theta_j ZZ_i) + \varepsilon, \tag{4}$$

where Θ_j is the parameter matrix of an incorrect cluster, $\beta \in [0, 1]$ controls the corruption strength ($\beta = 0.6$), and ε is small Gaussian perturbation ensuring numerical non-degeneracy. When $\rho_{\text{byz}} = 0$, the algorithm defaults to `CSI` aggregation (simple averaging). When $\rho_{\text{byz}} > 0$, robust aggregation uses either a trimmed-mean rule (`alpha_trim`$= \rho_{\text{byz}} + 0.05$) or geometric-median consensus depending on the flag `use_geom_median`.

C Regret Decomposition

To synthesize our regret bounds, we begin by defining the events $\mathcal{E}_{\text{success}}$ and $\mathcal{E}_{\text{failure}}$. These events characterize the high-probability regimes under which the system trajectories remain bounded and the clustered system identification produces sufficiently accurate estimates of the dynamics. Each of the following settings–(i) intra-cluster homogeneity, (ii) intra-cluster heterogeneity, and (iii) adversarial systems–are associated with a distinct estimation event.

Success and failure events. We first introduce the event which ensures that the state and designed controller norms remain uniformly bounded throughout the epoch length for all epochs:

$$\mathcal{E}_{\text{bound}} := \left\{ \|x_t\|^2 \le x_b^2 \log T,\ \forall t = 1, \dots, T \right\} \cap \left\{ \|\hat{K}_k\| \le K_b,\ \forall k = 1, \dots, k_{\text{fin}} \right\}. \tag{5}$$

We recall that the true and estimated system models are denoted by

$$\Theta_\star^{(i)} = [A_\star^{(i)} \;\; B_\star^{(i)}], \;\; \widehat{\Theta}^{(i)} = [\widehat{A}^{(i)} \;\; \widehat{B}^{(i)}],$$

and introduce a the estimation events, each corresponding to one of the three principal cases considered in this work:

• **Case 1: Intra-cluster homogeneity.** When all systems within a cluster share identical dynamics, we denote by $\mathcal{E}_{\text{est},1}^{(k)}$ the event which the estimation error admits the following bound:

$$\left\|\widehat{\Theta}_k^{(i)} - \Theta_\star^{(i)}\right\|_F^2 \le \frac{C_{\text{stat}}\sigma_w^2(d_x^2 + d_x d_u)\log(1/\delta)}{\sigma_k^2 M_j \tau_k} + C_{\text{mis},1}\exp\big(-C_{\text{mis},2}\sigma_k^2\tau_k\big), \tag{6}$$

for all epochs $k \in [k_{\mathsf{fin}}]$, and systems $i \in [m]$. The constants $C_{\text{stat}}, C_{\text{mis},1}, C_{\text{mis},2}$ are positive and universal, and they are defined in Lemma E.1, along with their derivations.

• **Case 2: Intra-cluster heterogeneity.** When systems within a cluster are similar but not identical with bounded heterogeneity ϵ_{het}, we denote by $\mathcal{E}_{\text{est},2}^{(k)}$ the event when the estimation error satisfies:

$$\left\|\widehat{\Theta}_k^{(i)} - \Theta_\star^{(i)}\right\|_F^2 \le \frac{C_{\text{stat}}\sigma_w^2(d_x^2 + d_x d_u)\log(1/\delta)}{\sigma_k^2 M_j \tau_k} + C_{\text{mis},1}\exp\big(-C_{\text{mis},2}\sigma_k^2\tau_k\big) + C_{het}\epsilon_{het}^2, \tag{7}$$

for all epochs $k \in [k_{\mathsf{fin}}]$, and systems $i \in [m]$. Here ϵ_{het} denotes the maximum heterogeneity level across the systems inside the clusters, quantifying the worst-case intra-cluster deviation from the nominal dynamics (see Assumption 1).

• **Case 3: Adversarial systems.** When clusters contain both heterogeneous and adversarial systems, we denote by $\mathcal{E}_{\text{est},3}^{(k)}$ the event when the estimation error is bounded as follows:

$$\begin{aligned}\left\|\widehat{\Theta}_k^{(i)} - \Theta_\star^{(i)}\right\|_F^2 \le{}& \frac{C_{\text{stat}}\sigma_w^2(d_x^2 + d_x d_u)\log(1/\delta)}{\sigma_k^2\tau_k}\left(\frac{1}{m_{\hat{j}}} + \lambda^2 d_x\right) + C_{\text{het}}(1+\lambda)^2\epsilon_{\text{het}}^2 \\ &+ C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_k^2\tau_k),\end{aligned} \tag{8}$$

for all epochs $k \in [k_{\mathsf{fin}}]$, and all systems $i \in [M]$. The latter case will be elaborated in a subsequent section, where we integrate the effects of both heterogeneity and adversarial systems into our system identification guarantees.

The success event in each regime is then defined by

$$\mathcal{E}_{\text{success}}^{(j)} := \mathcal{E}_{\text{bound}} \cap \mathcal{E}_{\text{est},j}, \;\text{ for } j = 1,2,3,$$

and corresponding failure events $\mathcal{E}_{\text{failure}}^{(j)} := \big(\mathcal{E}_{\text{success}}^{(j)}\big)^c$.

Regret decomposition. With these events in place, as in Cassel et al. [2020], the expected regret for system $i \in [m]$ decomposes as follows:

$$\mathbb{E}\left[\mathcal{R}_T^{(i)}\right] = R_1^{(i)} + R_2^{(i)} + R_3^{(i)} - T\mathcal{J}^{(i)}(K_\star^{(i)}), \;\text{ where}$$

$$R_1^{(i)} = \mathbb{E}\left[\mathbf{1}\left(\mathcal{E}_{\text{success}}^{(j)}\right)\sum_{k=2}^{k_{\mathsf{fin}}} J_k^{(i)}\right], \quad R_2^{(i)} = \mathbb{E}\left[\mathbf{1}\left(\mathcal{E}_{\text{failure}}^{(j)}\right)\sum_{t=\tau_1+1}^{T} c_t^{(i)}\right], \quad R_3^{(i)} = \mathbb{E}\left[\sum_{t=1}^{\tau_1} c_t^{(i)}\right], \quad (9)$$

with $J_k^{(i)} := \sum_{t=\tau_k}^{\tau_{k+1}-1} c_t^{(i)}$ denoting the cost for epoch $k \in [k_{\mathsf{fin}}]$. The interpretation for each term is standard:

1. $R_1^{(i)}$ quantifies the regret under the success event.
2. $R_2^{(i)}$ captures the cost contribution under the failure event.
3. $R_3^{(i)}$ accounts for the one-time exploration cost in the initial epoch.

A key step for synthesizing the regret bounds is to establish that $\mathbb{P}(\mathcal{E}_{\text{success}}^{(j)}) \geq 1 - T^{-2}$ in each setting $j = 1, 2, 3$ (see Appendix F). This renders the contribution of $R_2^{(i)}$ negligible, and ensures that the dominant terms are $R_1^{(i)}$ and $R_3^{(i)}$, which we bound explicitly in terms of noise variance, heterogeneity, and adversarial contamination.

D Auxiliary Results

We now turn to the auxiliary concentration inequalities and matrix norm bounds that play a central role in analyzing our robust multitask adaptive LQR design. In particular, we invoke the matrix Hoeffding's inequality to derive high-probability bounds on the estimation error under sub-Gaussian noise. We further provide derivations of the error decomposition into variance, misclassification bias, and heterogeneity-induced components.

These results are fundamental for establishing the design rationale and theoretical guarantees of the clustering-based adaptive estimation approach. The proofs are structured around a unified estimator framework and rely on matrix concentration techniques along with the assumptions on the data distribution and honest-majority condition for the adversarial setting.

Lemma D.1 (Matrix Hoeffding [Tropp, 2011])**.** *Let $\{X_\ell\}_{\ell=1}^m$ be independent, random and symmetric matrices in $\mathbb{R}^{d\times d}$ with $\mathbb{E}[X_\ell] = 0$ and almost-sure bounds $X_\ell^2 \preceq B_\ell^2$ for fixed symmetric matrix $B_\ell \succeq 0$. Define $\sigma^2 := \left\|\sum_{\ell=1}^m B_\ell^2\right\|$. Then, for all $t \geq 0$,*

$$\mathbb{P}\left(\lambda_{\max}\left(\sum_{\ell=1}^m X_\ell\right) \geq t\right) \leq d\exp\left(-\frac{t^2}{8\sigma^2}\right). \quad (10)$$

In addition, for general rectangular $\{M_\ell\}_{\ell=1}^m \subset \mathbb{R}^{d_1\times d_2}$, we define $X_\ell := \begin{bmatrix} 0 & M_\ell \\ M_\ell^\top & 0 \end{bmatrix}$ and assume each M_ℓ satisfies $\mathbb{E}[M_\ell] = 0$. Then for all $t \geq 0$,

$$\mathbb{P}\left[\sigma_{\max}\left(\sum_{t=1}^T M_\ell\right) \geq t\right] \leq (d_1 + d_2)\exp\left(-\frac{t^2}{8\sigma^2}\right). \quad (11)$$

Proof. The proof for this lemma is detailed in Tropp [2011]. The idea is to apply a Laplace-transform method with the fact that $\log \mathbb{E}\exp(\theta X_\ell) \preceq \frac{\theta^2}{2} B_\ell^2$ for θ small enough when $X_\ell^2 \preceq B_\ell^2$ and finish with a matrix Chernoff bound plus a trace trick. □

Lemma D.2 (Bound on $R_1^{(i)}$ from Lee et al. [2023]). *Let the $\mathcal{E}_{success}^{(j)} = \mathcal{E}_{\text{bound}} \cap \mathcal{E}_{\text{est},j}$ for the settings $j = 1, 2, 3$. Fix a system $i \in [m]$. Then, the $R_1^{(i)}$ as defined in 9 is bounded by:*

$$R_1^{(i)} \leq \sum_{k=2}^{k_{\text{fin}}} \Bigg(\mathbb{E}\Big[\mathbf{1}(\mathcal{E}_{\text{est},j}^{(k-1)})\, 142\, (\tau_k - \tau_{k-1})\, \|P_\star^{(i)}\|^8 \left\| [\widehat{A}_{k-1}^{(i)}\ \widehat{B}_{k-1}^{(i)}] - [A_\star^{(i)}\ B_\star^{(i)}] \right\|_F^2 \Big] \\ + (\tau_k - \tau_{k-1})\, J^{(i)}(K_\star^{(i)}) + 4(\tau_k - \tau_{k-1})\, d_u\, \|P_\star^{(i)}\|\, \sigma_k^2\, \Psi_B^{(i)2} \\ + 2x_b^2 \log T\, \|P_\star^{(i)}\| \Bigg), \tag{12}$$

where $\Psi_B^{(i)} = \max\{1, \|B_\star^{(i)}\|\}$.

Lemma D.3 (Bound on $R_2^{(i)}$ adapted from Lee et al. [2023]). *Fix a system $i \in [m]$. Then, in the contribution of the failure event to the regret is bounded as:*

$$R_2^{(i)} \leq T^{-1}\big(\|Q\| + 2K_b^2\big)x_b^2 \log T + T^{-1} J^{(i)}(K_0^{(i)}) + 24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x + d_u)\sigma_w^2 T^{-2}\log(3T) \\ + 2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\, \|\Theta_\star^{(i)}\|_F^2\, K_b^2 x_b^2 \log T + \sum_{k=1}^{k_{\text{fin}}} 2(\tau_k - \tau_{k-1}) d_u\, \sigma_k^2, \tag{13}$$

where K_b and x_b denote controller and state norm bounds, and σ_k^2 is the variance of the exploration noise in epoch k.

Lemma D.4 (Bound on $R_3^{(i)}$ from Lee et al. [2023]). *The exploration cost during the first epoch with length τ_1 satisfies:*

$$R_3^{(i)} \leq 3\tau_1\, \max\{d_x, d_u\}\, \|P_{K_0^{(i)}}^{(i)}\|\, \Psi_B^{(i)2}. \tag{14}$$

We also revisit two additional results from Lee et al. [2023]. The first result in Lemma D.5 controls the largest norm of the process and exploration noises, $w_t^{(i)}$ and $x_t^{(i)}$, respectively, for any system $i \in [m]$, with high probability. The second result in Lemma D.6 bounds the norm of the state and the norm of the solution of the Lyapunov equation P_K, for a sufficiently large horizon length. In particular, the bound for the state norm scales with the largest norm of the process and exploration noise. Later, we see that Lemmas D.5 and D.6 can be used to demonstrate that if the initial epoch length is sufficiently large, the state and controller norm boundedness requirement in Algorithm 1 are satisfied.

Lemma D.5. *(Lee et al. [2023]) Let $\delta \in (0, 1)$. For any system $i \in [m]$, it holds that*

$$\max_{0 \leq t \leq T-1} \left\| \begin{bmatrix} w_t^{(i)} \\ g_t^{(i)} \end{bmatrix} \right\| \leq 4\sigma_w \sqrt{(d_x + d_u) \log \frac{T}{\delta}},$$

with probability at least $1 - \delta$.

Lemma D.6. *(Lee et al. [2023]) Consider the discrete LTI system $x_{s+1} = A_\star x_s + B_\star u_s + w_s$ with initial state x_0. Suppose that we play this system with the control action $u_s = Kx_s + \sigma_u g_s$ where K is stabilizing and $\sigma_u \leq 1$, for t time steps. Moreover, suppose that*

- $\|x_1\| \leq 16 \|P_{K_0}\|^{3/2} \Psi_{B^\star} \max_{0\leq t\leq T-1} \left\| \begin{bmatrix} w_t \\ g_t \end{bmatrix} \right\|$
- $\|P_K\| \leq 2 \|P_{K_0}\|$
- $t \geq \log_{\left(1-\frac{1}{\|P_K\|}\right)} \left(\frac{1}{4\|P_K\|}\right) + 1.$

Then for $s = 0, \ldots, t-1$, it holds that

$$\|x_s\| \leq 40 \|P_{K_0}\|^2 \Psi_{B^\star} \max_{1\leq t\leq T} \left\| \begin{bmatrix} w_t \\ g_t \end{bmatrix} \right\|.$$

In addition, we have

$$\|x_t\| \leq 16 \|P_{K_0}\|^{3/2} \Psi_{B^\star} \max_{1\leq t\leq T} \left\| \begin{bmatrix} w_t \\ g_t \end{bmatrix} \right\|.$$

Assumption 4. *(Geometric mixing) For any system $i \in [m]$, assume the state evolution -input process $\left\{z_t^{(i)}\right\}_{t\geq 0}$ is a mean-zero stationary β-mixing process, with stationary covariance $\Sigma_{z_t}^{(i)}$ and $\beta(s) \leq C_\beta \rho^s$, for some $C_\beta \geq 0$ and $\rho \in (0,1)$.*

E Multitask System Identification

We now turn our attention to characterizing the error bounds for clustered system identification under three different settings: 1) without intra-cluster heterogeneity and without adversarial systems; 2) with intra-cluster heterogeneity but still without adversarial systems; and 3) in the presence of adversarial systems. For the first two settings, the aggregation function used in Algorithm 1 is a simple average, whereas for the third, the adversarial setting, we employ a resilient aggregation scheme as defined in Definition 2.1. For completeness, we restate here the key lemmas presented in the main body of the paper.

E.1 System Identification with Intra-cluster Homogeneity

We now characterize the system estimation error in epoch k (of length τ_k) for clustered system identification when all systems within a cluster have identical local models.

Lemma E.1 (Estimation error under intra-cluster homogeneity)**.** *Fix a cluster $\mathcal{C}_j$ with M_j systems. For each system $i \in \mathcal{C}_j$, data generation follow (1). Moreover, suppose Assumption 2 holds. Then, given a small probability of failure $\delta \in (0,1)$, it holds that*

$$\left\| \widehat{\Theta}^{(i)} - \Theta_\star^{(i)} \right\|_F^2 \leq \frac{C_{\text{stat}} \sigma_w^2 (d_x^2 + d_x d_u) \log(1/\delta)}{\sigma_k^2 M_j \tau_k} + C_{\text{mis},1} \exp\left(-C_{\text{mis},2} \sigma_k^2 \tau_k\right), \tag{15}$$

for any system $i \in \mathcal{C}_j$, with probability at least $1-\delta-\exp(-C_{\text{mis},2}\sigma_k^2\tau_k)$, where $C_{\text{mis},2} = \mathcal{O}(\Delta_{\min}^2/\Delta_{\max}^2\sigma_w^2)$ quantifies the decay rate of the misclassification at epoch k, and $C_{\text{mis},1} = \mathcal{O}(\Delta_{\max}^2)$ corresponds is the misclassification constant.

Proof. For a given epoch k of length τ_k, each system $i \in [M]$ evolves according to the dynamics in (1). Each system uses its collected dataset $\mathcal{D}^{(i)} = \{X^{(i)}, Z^{(i)}\}$, consisting of state–input and next-state trajectories, to estimate its model and identify its cluster membership. Suppose that system i is assigned to cluster $\hat{j}$, where $\mathcal{C}_{\hat{j}}$ denotes the set of indices of systems belonging to cluster $\hat{j} \in [N_c]$. Each system computes its gradient using its local data and the common model of the cluster to which it is assigned. The gradients are then transmitted to the server, which performs the following aggregation step:

$$\widehat{\Theta}_{n+1}^{(i)} = \widehat{\Theta}_n^{(i)} + \frac{\eta}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in C_{\hat{j}}} G_\ell(\widehat{\Theta}_n^{(i)}),$$

where $G_\ell(\widehat{\Theta}^{(i)}) = (X^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}) Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1}$, as specified in Algorithm 2. We can further decompose the average over the set of systems assigned to cluster $\hat{j}$ into two parts: the average over systems that are correctly identified, i.e., those belonging to $\mathcal{C}_j \cap \mathcal{C}_{\hat{j}}$, and the average over systems that are misclassified to cluster $\hat{j}$, i.e., those belonging to $\mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}$, where $\mathcal{C}_j^c$ denotes the complement of $\mathcal{C}_j$. That is, we obtain

$$\widehat{\Theta}_{n+1}^{(i)} = \widehat{\Theta}_n^{(i)} + \frac{\eta}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} G_\ell(\widehat{\Theta}_n^{(i)}) + \frac{\eta}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}} G_\ell(\widehat{\Theta}_n^{(i)}),$$

where we can write

$$\begin{aligned}
\frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} G_l(\widehat{\Theta}_n^{(i)}) &= \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} (X^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}) Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1} \\
&= \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} (\Theta_\star^{(i)} Z^{(\ell)} + W^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}) Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1} \\
&= \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} \left(\Theta_\star^{(i)} Z^{(\ell)} + W^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}\right) Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1} \\
&= \Theta_\star^{(i)} - \widehat{\Theta}_n^{(i)} + \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} W^{(\ell)} Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1},
\end{aligned}$$

and the following expression for the average over the misclassified systems:

$$\frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}} G_\ell(\widehat{\Theta}_n^{(i)}) = \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}} (\Theta_j - \widehat{\Theta}^{(i)}) + \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}} W^{(\ell)} Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1}$$

where Θ_j denotes the common model that system q would use if it were correctly classified to its true cluster, which is different from $\hat{j}$. Therefore, we obtain

$$\widehat{\Theta}_{n+1}^{(i)} = \widehat{\Theta}_n^{(i)} + \eta \left(\Theta_\star^{(i)} - \widehat{\Theta}_n^{(i)}\right) + \underbrace{\frac{\eta}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_{\hat{j}}} W^{(\ell)} Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1}}_{\text{statistical error}} + \underbrace{\frac{\eta}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}} (\Theta_j - \widehat{\Theta}^{(i)})}_{\text{misclassification error}}, \quad (16)$$

where we subtract $\Theta_\star^{(i)}$ from both sides to obtain

$$\|\widehat{\Theta}_{n+1}^{(i)} - \Theta_\star^{(i)}\|_F \le (1-\eta)\|\widehat{\Theta}_n^{(i)} - \Theta_\star^{(i)}\|_F + \left\| \frac{\eta}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_{\hat{j}}} W^{(\ell)} Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1} \right\|_F + \left\| \frac{\eta}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}} (\Theta_j - \widehat{\Theta}_\star^{(i)}) \right\|_F,$$

• **Statistical Error:** For each system $\ell \in \mathcal{C}_{\hat{j}}$, we have trajectory data of length $\tau_k = \tau_{k,1} + \tau_{k,2}$ and the block split $\mathcal{I}_1$ (size $\tau_{k,1}$) and $\mathcal{I}_2$ (size $\tau_{k,2}$), with $\tau_{k,1}, \tau_{k,2} \ge 1$. We recall that $z_t^{(\ell)} = \begin{bmatrix} x_t^{(\ell)} \\ u_t^{(\ell)} \end{bmatrix} \in \mathbb{R}^d$ denotes the state-input vector, and $w_t^{(\ell)} \in \mathbb{R}^{d_x}$ is the process noise, independent across t and ℓ, with zero mean. Define

$$Z_{\mathcal{I}_1}^{(\ell)} := \left[z_t^{(\ell)}\right]_{t \in \mathcal{I}_1} \in \mathbb{R}^{d \times \tau_{k,1}}, \quad Z_{\mathcal{I}_2}^{(\ell)} := \left[z_t^{(\ell)}\right]_{t \in \mathcal{I}_2} \in \mathbb{R}^{d \times \tau_{k,2}}, \quad P^{(\ell)} := \left(Z_{\mathcal{I}_1}^{(\ell)} Z_{\mathcal{I}_1}^{(\ell)\top}\right)^{-1} \in \mathbb{R}^{d \times d}.$$

We also define the (decoupled) cross term on the second block

$$M^{(\ell)} := \sum_{t \in \mathcal{I}_2} w_t^{(\ell)} z_t^{(\ell)\top} P^{(\ell)} \in \mathbb{R}^{d_x \times d}, \quad \bar{M} := \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_{\hat{j}}} M^{(\ell)}.$$

Sample-split independence: By construction (i.e., sample-split in two independent blocks), $P^{(\ell)}$ is independent of $\{(w_t^{(\ell)}, z_t^{(\ell)}) : t \in \mathcal{I}_2\}$, and thus $\mathbb{E}[M^{(\ell)} \mid P^{(\ell)}] = 0$. We can see this by fixing a system ℓ and a time $t \in \mathcal{I}_2$, and by considering the σ–field $\mathcal{F}_{t-1} := \sigma\left(P^{(\ell)}, \{z_s^{(\ell)} : s \le t\}\right)$.

Therefore, the sample–split construction, $P^{(\ell)}$ depends only on the block $\mathcal{I}_1$ and is then independent of the pair $\left(w_t^{(\ell)}, z_t^{(\ell)}\right)$ from block $\mathcal{I}_2$. In the process–noise model, $w_t^{(\ell)}$ is zero–mean and it only affects $x_{t+1}^{(\ell)}$ (hence $z_{t+1}^{(\ell)}$), but is independent of the present regressor $z_t^{(\ell)}$ and of $\mathcal{F}_{t-1}$. Consequently,

$$\mathbb{E}\left[w_t^{(\ell)} z_t^{(\ell)\top} P^{(\ell)} \mid \mathcal{F}_{t-1}\right] = \left(\mathbb{E}[w_t^{(\ell)} \mid \mathcal{F}_{t-1}]\right) z_t^{(\ell)\top} P^{(\ell)} = 0,$$

where we used that $z_t^{(\ell)\top} P^{(\ell)}$ is $\mathcal{F}_{t-1}$–measurable and $\mathbb{E}[w_t^{(\ell)} \mid \mathcal{F}_{t-1}] = 0$. Taking expectations again and summing over $t \in \mathcal{I}_2$ yields

$$\mathbb{E}\left[M^{(\ell)} \mid P^{(\ell)}\right] = \sum_{t \in \mathcal{I}_2} \mathbb{E}\left[w_t^{(\ell)} z_t^{(\ell)\top} P^{(\ell)} \mid P^{(\ell)}\right] = 0.$$

Note that conditioning must be on $P^{(\ell)}$ (or on the past $\mathcal{F}_{t-1}$). If one conditioned on the entire future block $Z_{\mathcal{I}_2}^{(\ell)}$, then $w_t^{(\ell)}$ would correlate with $z_{t+1}^{(\ell)}$ and the conditional mean need not be zero.

We proceed by writing

$$\bar{M} = \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_{\hat{j}}} \left(\sum_{t \in \mathcal{I}_2} w_t^{(\ell)} z_t^{(\ell)\top} P^{(\ell)} \right).$$

We fix a system ℓ, and by the independence from the sample split we have that $\mathbb{E}[M^{(\ell)} \mid P^{(\ell)}] = 0$. In addition, we for each $t \in \mathcal{I}_2$, we obtain

$$\|w_t^{(\ell)} z_t^{(\ell)\top} P^{(\ell)}\| \leq \|w_t^{(\ell)}\|_2 \|z_t^{(\ell)}\|_2 \|P^{(\ell)}\|,$$

As we assume $w_t^{(\ell)}$ is mean-zero sub-Gaussian with variance σ_w^2, and thus $z_t^{(\ell)}$ is mean-zero sub-Gaussian with covariance $\Sigma_{z_t}^{(\ell)} := \mathbb{E}[z_t^{(\ell)} z_t^{(\ell)\top}]$ in the sense that $\|\langle v, z_t^{(\ell)} \rangle\|_{\psi_2} \leq C\sqrt{v^\top \Sigma_{z_t}^{(\ell)} v}$ for all v and some constant $C > 0$. Let $\kappa^{(\ell)} := \lambda_{\max}\left(\Sigma_{z_t}^{(\ell)}\right)$. In particular, by [Wang et al., 2023a, Lemma 1], we have

$$\Sigma_{z_t}^{(\ell)} \triangleq \begin{bmatrix} \sigma_k^2 G_t^{(\ell)} \left(G_t^{(\ell)}\right)^\top + \sigma_w^2 F_t^{(\ell)} \left(F_t^{(\ell)}\right)^\top & 0 \\ 0 & \sigma_k^2 I_{d_u} \end{bmatrix},$$

with $G_t \triangleq \begin{bmatrix} A^{(\ell)t-1}B^{(\ell)} & A^{(\ell)t-2}B^{(\ell)} & \cdots & B^{(\ell)} \end{bmatrix}$ and $F_t^{(\ell)} \triangleq \begin{bmatrix} A^{(\ell)t-1} & A^{(\ell)t-2} & \cdots & I_{d_x} \end{bmatrix}$, for any $t \geq 1$. Then, by standard inequalities for sub-Gaussian vectors, on a high-probability event, with probability at least $1 - \delta$,

$$\max_{t \in \mathcal{I}_2} \|w_t^{(\ell)}\|_2 \leq C_w \sigma_w \sqrt{\log \frac{\tau_{k,2}}{\delta}} := B_w, \quad \max_{t \in \mathcal{I}_2} \|z_t^{(\ell)}\|_2 \leq C_z \sqrt{\kappa^{(\ell)}} \left(\sqrt{d'} + \sqrt{\log \frac{\tau_{k,2}}{\delta}}\right) := B_z.$$

for some constants C_w and C_z. Hence, we obtain

$$\|w_t^{(\ell)} z_t^{(\ell)\top} P^{(\ell)}\| \leq B_w B_z \|P^{(\ell)}\|,$$

We can now apply matrix Hoeffding inequality (Lemma D.1) to $M^{(\ell)}$:

$$\mathcal{X}_\ell = \begin{bmatrix} 0 & M^{(\ell)} \\ M^{(\ell)\top} & 0 \end{bmatrix}, \quad \mathbb{E}[\mathcal{X}_\ell \mid P^{(\ell)}] = 0, \quad \mathcal{X}_\ell^2 \preceq B_w^2 B_z^2 \|P^{(\ell)}\|^2 I_{2d_x + d_u},$$

Therefore, we have that

$$\mathbb{P}\left(\left\| \sum_{\ell \in C_{\hat{j}}} M^{(\ell)} \right\| \geq t\right) \leq (2d_x + d_u) \exp\left(-\frac{t^2}{8\sigma^2}\right),$$

where $\sigma^2 = B_w^2 B_z^2 \sum_{\ell \in C_{\hat{j}}} \|P^{(\ell)}\|^2$. By setting $t = \sqrt{\tau_k} \sigma_k B_w B_z \sqrt{\sum_{\ell \in C_{\hat{j}}} \|P^{(\ell)}\|^2} \sqrt{8 \log\left(\frac{2d_x + d_u}{\delta}\right)}$, we obtain the following expression:

$$\left\|\bar{M}\right\| \leq \frac{\sqrt{\tau_k} \sigma_k B_w B_z}{|C_{\hat{j}}|} \sqrt{\sum_{\ell \in C_{\hat{j}}} \|P^{(\ell)}\|^2} \sqrt{8 \log\left(\frac{2d_x + d_u}{\delta}\right)},$$

with probability at least $1 - \delta$.

Controlling $\|P^{(\ell)}\|$: We now proceed to prove that $\|P^{(\ell)}\| \leq \frac{C_P}{\sigma_k^2 \tau_k}$, for some constant C_P, with probability $1 - \delta_P$. To prove this bound, we first assume that $\{z_t^{(\ell)}\}_{t \in \mathcal{I}_1}$ is a strictly stationary,

mean-zero Gaussian process in $\mathbb{R}^d$ with covariance $\Sigma_z^{(\ell)} := \mathbb{E}\left[z_t^{(\ell)} z_t^{(\ell)\top}\right] \succeq \sigma_k^2 I_d$ and geometric β-mixing, i.e., $\beta(s) \le C_\beta \rho^s$ for some $C_\beta > 0$ and $\rho \in (0,1)$ (see [Yu, 1994]).

Let $\Sigma_t^{1/2} := (\Sigma_{z_t}^{(\ell)})^{1/2}$. As we know that $\{z_t\}_t$ are Gaussian with covariance $\Sigma_{z_t}^{(\ell)}$, we can write $z_t = \Sigma^{1/2} y_t$ where $\{y_t\}_t$ is a stationary Gaussian process in $\mathbb{R}^d$ with the same β-mixing rate. Hence, we can write

$$S - \Sigma_{z_t}^{(\ell)} = \Sigma_t^{1/2} \left(\frac{1}{\tau_{k,1}} \sum_t y_t y_t^\top - I_d \right) \Sigma_t^{1/2}, \text{ and thus } \|S - \Sigma_{z_t}^{(\ell)}\| \le \|\Sigma_{z_t}^{(\ell)}\| \left\| \frac{1}{\tau_{k,1}} \sum_t y_t y_t^\top - I_d \right\|.$$

As $\{z_t\}_t$ are dependent over time t, we use a blocking technique to construct "independent blocks" of data in $\mathcal{I}_1$. Therefore, let us partition $\mathcal{I}_1$ into q kept blocks of length m_1 separated by gaps of length m_2. In addition, we denote $m = m_1 + m_2$, $q = \lfloor \tau_{k,1}/m \rfloor$. For the kept blocks, we define the block averages $\overline{Y}_j := \frac{1}{m_1} \sum_{t \in \mathcal{B}_j} (y_t y_t^\top - I_d)$. Therefore, there exist i.i.d. copies $\{\overline{Y}'_j\}_{j=1}^q$ with the same marginals such that $\mathbb{P}\{\exists j : \overline{Y}_j \neq \overline{Y}'_j\} \le q\beta(m_2) \le qC_\beta \rho^{m_2}$. This implies that, with probability at least $1 - qC_\beta \rho^{m_2}$, the kept blocks behave as independent.

Conditional on the coupling event, we apply an effective-rank covariance deviation bound to $\frac{1}{q}\sum_{j=1}^q \overline{Y}'_j$, where each $\overline{Y}'_j$ is an average of m_1 i.i.d. variables, where its sub-exponential norm is uniformly bounded, and its second moment has effective rank at most $r_{\text{eff}}^{(\ell)} = \text{tr}(\Sigma_t)/\|\Sigma_t\|$. A standard Gaussian covariance concentration (e.g., matrix Bernstein [Vershynin, 2018]) yields, for all $s > 0$,

$$\mathbb{P}\left\{ \left\| \frac{1}{q} \sum_{j=1}^q \overline{Y}'_j \right\| \ge C_Y \left(\sqrt{\frac{r_{\text{eff}}^{(\ell)} + s}{qm_1}} + \frac{r_{\text{eff}}^{(\ell)} + u}{qm_1} \right) \right\} \le 2e^{-s}.$$

As the contribution of discarded intervals of samples dilutes the kept sample size by at most a factor $\frac{m_1}{m_1+m_2}$, we conclude that

$$\left\| \frac{1}{\tau_{k,1}} \sum_t y_t y_t^\top - I_d \right\| \le C_Y \left(\sqrt{\frac{r_{\text{eff}}^{(\ell)} + \log(2/\delta)}{\tau_{k,1}}} + \frac{r_{\text{eff}}^{(\ell)} + \log(2/\delta)}{\tau_{k,1}} \right)$$

with probability at least $1 - \delta - qC_\beta \rho^{m_2}$, for some constant C_Y.

By setting $m_2 = \left\lceil \frac{\log(2qC_\beta/\delta)}{|\log \rho|} \right\rceil$ we have that $qC_\beta \rho^{m_2} \le \delta/2$, and by taking $m_1 = m_2$ so that $qm_1 \asymp \tau_{k,1}$. Then the bound above holds with probability at least $1 - \delta$ and becomes

$$\|S - \Sigma_{z_t}^{(\ell)}\| \le C_Y \|\Sigma_{z_t}^{(\ell)}\| \left(\sqrt{\frac{r_{\text{eff}}^{(\ell)} + \log(2/\delta)}{\tau_{k,1}}} + \frac{r_{\text{eff}}^{(\ell)} + \log(2/\delta)}{\tau_{k,1}} \right).$$

Therefore, by setting $\tau_{k,1} \ge C_Y(r_{\text{eff}}^{(\ell)} + \log(2/\delta))$ with a large enough C_Y, we have

$$\|P^{(\ell)}\| \le \frac{2}{\tau_{k,1} \lambda_{\min}(\Sigma_z^{(\ell)})} \le \frac{2}{\tau_{k,1} \sigma_k^2},$$

as we know that $\Sigma_{z_t}^{(\ell)} \succeq \sigma_k^2 I_d$, by persistency of excitation, and thus $\lambda_{\min}(\Sigma_z^{(\ell)}) \ge \sigma_k^2$.

Therefore, by using the above bound for $\|P^{(\ell)}\|$, we guarantee that the statistical error term in (16) is upper bounded as follows:

$$\left\| \frac{1}{|\mathcal{C}_{\hat{j}}|} \sum_{\ell \in \mathcal{C}_{\hat{j}}} W^{(\ell)} Z^{(\ell)\top} (Z^{(\ell)} Z^{(\ell)\top})^{-1} \right\|_F \leq \frac{\sqrt{d_x}\sqrt{\tau_{k,1}}\sigma_k B_w B_z}{|C_{\hat{j}}|} \sqrt{|C_{\hat{j}}|} \frac{C_P}{\sigma_k^2 \tau_k} \sqrt{8 \log\left(\frac{2d_x + d_u}{\delta}\right)}$$

$$= \frac{\sqrt{d_x} B_w B_z C_P}{\sigma_k} \sqrt{\frac{8}{|\mathcal{C}_{\hat{j}}| \tau_{k,1}} \log\left(\frac{2d_x + d_u}{\delta}\right)}. \tag{17}$$

where the extra $\sqrt{d_x}$ term is due to upper bounding the Frobenius norm with the spectral norm.

• **Misclassification Error:** Now let us control the misclassification error term in (16). For this purpose, we define the misclassification event for agent $\ell \in \mathcal{C}_{\hat{j}}$ as follows:

$$\mathcal{E}_{\text{mis}}^{(\ell)} := \left\{ \exists j' \neq j : \left\| X^{(\ell)} - \widehat{\Theta}_j Z^{(\ell)} \right\|_F^2 > \left\| X^{(\ell)} - \widehat{\Theta}_{j'} Z^{(\ell)} \right\|_F^2 \right\},$$

where we recall that $X^{(\ell)}$ and $Z^{(\ell)}$ denote the data matrices corresponding to system ℓ, and $\widehat{\Theta}_{j'}$ is the estimated model for cluster $j' \neq j$. Substituting $X^{(\ell)} = \Theta_j Z^{(\ell)} + W^{(\ell)}$, where $\Theta_j = \Theta_\star^{(\ell)}$, we obtain

$$\left\| (\Theta_j - \widehat{\Theta}_j) Z^{(\ell)} + W^{(\ell)} \right\|_F^2 > \left\| (\Theta_j - \widehat{\Theta}_{j'}) Z^{(\ell)} + W^{(\ell)} \right\|_F^2.$$

Rearranging terms, this event is equivalent to

$$D := \left\| \Delta_{j'} Z^{(\ell)} + W^{(\ell)} \right\|_F^2 - \left\| \Delta_j Z^{(\ell)} + W^{(\ell)} \right\|_F^2 < 0,$$

where we define $\Delta_{j'} := \Theta_j - \widehat{\Theta}_{j'}$ and $\Delta_j := \Theta_j - \widehat{\Theta}_j$ as the estimation residuals for the incorrect and correct cluster models, respectively. We proceed to analyze this term by defining

$$D_1 := \left\| \Delta_{j'} Z^{(\ell)} \right\|_F^2 - \left\| \Delta_j Z^{(\ell)} \right\|_F^2, \text{ and } D_2 := 2\text{tr}\left(W^{(\ell)\top} (\Delta_{j'} - \Delta_j) Z^{(\ell)} \right),$$

where $D = D_1 + D_2$. Then the probability of misclassification probability becomes

$$\mathbb{P}(\mathcal{E}_{\text{mis}}^{(\ell)}) = \mathbb{P}(D < 0) = \mathbb{P}(D_1 < -D_2).$$

As D_1 is a linear with respect to the assumed Gaussian noise $W^{(\ell)}$ matrix, it is sub-Gaussian with variance bounded as follows:

$$\text{Var}(D_2) \leq C_{D_2} \sigma_w^2 \mathbb{E} \left\| (\Delta_{j'} - \Delta_j) Z^{(\ell)} \right\|_F^2.$$

for some sufficiently large constant C_{D_2}. Moreover, by bounding $\|\Delta_{j'} - \Delta_j\|_F \leq 2\Delta_{\max}$ with the maximum cluster separation, we have that

$$\text{Var}(D_2) \leq 2C_{D_2} \sigma_w^2 \Delta_{\max}^2 \mathbb{E} \left\| Z^{(\ell)} \right\|_F^2.$$

In addition, we have that $\|Z^{(\ell)}\|_F^2 \le C_{Z,\text{spec}}\tau_k$, where $C_{Z,\text{spec}}$ is a uniform upper bound for the spectral norm of $Z^{(\ell)}$. Therefore, we obtain

$$\text{Var}(D_2) \le 2C_{D_2}\Delta_{\max}^2 C_{Z,\text{spec}}\sigma_w^2\tau_k,$$

To lower bound D_1, observe that by persistency of excitation, $\lambda_{\min}(\mathbb{E}[z_t^{(i)}z_t^{(i)\top}]) \ge \sigma_k^2$, and Assumption 2, we obtain

$$D_1 \ge \tau_k\lambda_{\min}\left(\mathbb{E}[z_t z_t^\top]\right)\left(\|\Delta_{j'}\|_F^2 - \|\Delta_j\|_F^2\right) \ge \sigma_k^2\tau_k\left(\Delta_{\min}^2 - \delta_{\text{stat}}^2\right).$$

where δ_{stat} denotes the statistical error from (17). Therefore, by setting $\tau_{k,1}$ such that

$$\frac{8d_xB_wB_zC_P}{\sigma_k^2|C_{\hat{j}}|\tau_{k,1}}\log\left(\frac{2d_x+d_u}{\delta}\right) \le \frac{\Delta_{\min}^2}{2} \to \tau_{k,1} \ge \frac{8d_xB_wB_zC_P\Delta_{\min}^2}{2\sigma_k^2M_{\hat{j}}}\log\left(\frac{2d_x+d_u}{\delta}\right),$$

which is guaranteed by making the initial epoch length sufficiently large. Therefore, we obtain

$$D_1 \ge \frac{\sigma_k^2\tau_k\Delta_{\min}^2}{2} \ge \frac{\sigma_k\tau_k\Delta_{\min}}{2},$$

where the second inequality follows from assuming that the clusters are sufficiently well separated such that $\Delta_{\min} \ge \frac{1}{\sigma_k}$. By substituting into a Gaussian tail bound yields the following misclassification probability:

$$\mathbb{P}(\mathcal{E}_{\text{mis}}^{(\ell)}) \le \exp\left(-\frac{D_1^2}{2D_2}\right) \le \exp\left(-\tau_k\frac{\sigma_k^2\Delta_{\min}^2}{8C_{D_2}C_{Z,\text{spec}}\Delta_{\max}^2\sigma_w^2}\right) := \exp\left(-C_2\sigma_k^2\tau_k\right).$$

with $C_2 = \frac{\Delta_{\min}^2}{8C_{D_2}C_{Z,\text{spec}}\Delta_{\max}^2\sigma_w^2}$. Therefore, the misclassification rate is bounded by

$$\frac{|\mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}|}{|\mathcal{C}_{\hat{j}}|} \le \exp(-C_2\sigma_k^2\tau_k),$$

which implies that

$$\left\|\frac{1}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell\in\mathcal{C}_j^c\cap\mathcal{C}_{\hat{j}}}(\Theta_j - \widehat{\Theta}_\star^{(i)})\right\|_F \le \Delta_{\max}\exp\left(-C_2\sigma_k^2\frac{\tau_k}{2}\right), \tag{18}$$

Combining the high-probability bound on the statistical error (17) with the misclassification error (18), we obtain the following error bound at iteration n:

$$\begin{aligned}\|\widehat{\Theta}_{n+1}^{(i)} - \Theta_\star^{(i)}\|_F &\le (1-\eta)\|\widehat{\Theta}_n^{(i)} - \Theta_\star^{(i)}\|_F + \frac{\eta\sqrt{d_x}B_wB_zC_P}{\sigma_k}\sqrt{\frac{8}{|\mathcal{C}_{\hat{j}}|\tau_{k,1}}\log\left(\frac{2d_x+d_u}{\delta}\right)} \\ &\quad + \eta\Delta_{\max}\exp\left(-C_2\frac{\tau_k}{2}\right) \\ &= (1-\eta)\|\widehat{\Theta}_n^{(i)} - \Theta_\star^{(i)}\|_F + \frac{\eta C_{\text{stat}}\sigma_w\sqrt{(d_x^2+d_xd_u)\log(1/\delta)}}{\sigma_k\sqrt{M_{\hat{j}}\tau_{k,1}}}\end{aligned}$$

$$+ \eta C_{\text{mis},1} \exp\left(-C_{\text{mis},2}\sigma_k^2 \tau_k\right),$$

where C_{stat} depends on $C_w, \log(\tau_{k,2}), C_z, \kappa^{(\ell)}, C_P$, and $\log(d)$. In addition, for clarity in our bounds, we use $C_{\text{mis},1} = \Delta_{\max}$ and $C_{\text{mis},2} = \frac{C_2}{2}$.

To conclude the proof we unroll the above expression over N iterations and write

$$\|\widehat{\Theta}_N^{(i)} - \Theta_\star^{(i)}\|_F \leq \rho^N \|\widehat{\Theta}_0^{(i)} - \Theta_\star^{(i)}\|_F + \frac{C_{\text{stat}}\sigma_w\sqrt{(d_x^2 + d_x d_u)\log(1/\delta)}}{\sigma_k\sqrt{M_{\hat{j}}\tau_{k,1}}} + C_{\text{mis},1}\exp\left(-C_{\text{mis},2}\sigma_k^2\tau_k\right),$$

with $\rho = 1 - \eta$. The proof is complete by setting $N \geq \log\left(C_\alpha \Delta_{\min}\tau_1^2\right)/\log(1/\rho)$, and noting that for a sufficiently large initial epoch length, $M_{\hat{j}} \approx M_j$, which guarantees that the systems are correctly classified and the error bound decays with the true total number of systems inside the cluster. Finally, we also note that $\tau_{k,1} = \mathcal{O}(\tau_k)$. In addition, we omit the contraction term of order $\mathcal{O}(1/\tau_k^2)$ by setting the total number of iterations N as before, since this term becomes negligible in the subsequent regret analysis. □

E.2 System Identification with Intra-cluster Heterogeneity

We now turn our attention to the setting in which the systems within each cluster are similar but not identical, exhibiting bounded heterogeneity characterized by ϵ_{het}.

Proposition E.1 (Estimation error under intra-cluster heterogeneity)**.** *Fix a cluster $\mathcal{C}_j$ with M_j systems. For each system $i \in \mathcal{C}_j$, data generation follow (1). Suppose Assumption 2 holds and that the intra-cluster heterogeneity is bounded by ϵ_{het}. Then, given a small probability of failure $\delta \in (0,1)$, it holds that*

$$\left\|\widehat{\Theta}^{(i)} - \Theta_\star^{(i)}\right\|_F^2 \leq \frac{C_{\text{stat}}\sigma_w^2(d_x^2 + d_x d_u)}{\sigma_k^2 M_j \tau_k} + C_{\text{het}}\epsilon_{\text{het}}^2 + C_{\text{mis},1}\exp\left(-C_{\text{mis},2}\sigma_k^2\tau_k\right), \tag{19}$$

for any system $i \in \mathcal{C}_j$, with probability at least $1 - \delta - \exp(-C_{\text{mis},2}\sigma_k^2\tau_k)$.

Proof. The proof for this lemma follows from using the of Lemma E.1 and writing

$$\widehat{\Theta}_{n+1}^{(i)} = \widehat{\Theta}_n^{(i)} + \frac{\eta}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} G_\ell(\widehat{\Theta}_n^{(i)}) + \frac{\eta}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell \in \mathcal{C}_j^c \cap \mathcal{C}_{\hat{j}}} G_\ell(\widehat{\Theta}_n^{(i)}),$$

where we can write for the statistical error term

$$\begin{aligned}
\frac{1}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} G_l(\widehat{\Theta}_n^{(i)}) &= \frac{1}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} (X^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}) Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} \\
&= \frac{1}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} \left(\left(\Theta_\star^{(i)} + \Theta_\star^{(\ell)} - \Theta_\star^{(i)}\right) Z^{(\ell)} + W^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}\right) Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} \\
&= \frac{1}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} \left(\Theta_\star^{(i)} Z^{(\ell)} + W^{(\ell)} - \widehat{\Theta}^{(i)} Z^{(\ell)}\right) Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} \\
&+ \underbrace{\frac{\eta}{|\mathcal{C}_{\hat{j}}|}\sum_{\ell \in \mathcal{C}_j \cap \mathcal{C}_{\hat{j}}} \left(\Theta_\star^{(\ell)} - \Theta_\star^{(i)}\right)}_{\text{System heterogeneity}},
\end{aligned}$$

where we note the presence of the system heterogeneity term which is further upper bounded using the bounded heterogeneity condition in (2). The remaining of the proof follows exactly as in Lemma E.1, where the bounds for the statistical error (17) and misclassification error (18) are leveraged to obtain the error bound presented in (19). □

E.3 Adversarially Robust System Identification

We now focus on the setting where adversarial systems may be present within the cluster of interest. To mitigate their effect in the multitask system identification process, we adopt an aggregation scheme that is (f, λ)-resilient, as defined in Definition 2.1.

Lemma E.2 (Estimation error under (f_j, λ)-resilient aggregation)**.** *Fix a cluster $\mathcal{C}_j$ with M_j systems, among which at most $f_j < M_j/2$ are adversarial and $m_j := M_j - f_j$ are honest. For each honest system $i \in \mathcal{C}_j$, data generation follow (1). Suppose that the aggregation function $F\left(\left\{G_\ell(\widehat{\Theta}^{(i)})\right\}_{\ell\in\mathcal{C}_j}\right)$ is (f, λ)-resilient and that the intra-cluster heterogeneity is bounded by ϵ_{het}. Moreover, suppose Assumption 2 holds. Then, given a small probability $\delta \in (0, 1)$, it holds that*

$$\|\widehat{\Theta}_N^{(i)} - \Theta_\star^{(i)}\|_F^2 \le \frac{C_{\text{stat}}\sigma_w^2(d_x^2 + d_x d_u)\log(1/\delta)}{\sigma_k^2\tau_k}\left(\frac{1}{m_{\hat{j}}} + \lambda^2 d_x\right) + C_{\text{het}}(1+\lambda)^2\epsilon_{\text{het}}^2 + C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_k^2\tau_k), \tag{20}$$

for any system $i \in \mathcal{C}_j$, with probability at least $1 - \delta - \exp(-C_{\text{mis},2}\sigma_k^2\tau_k)$.

Proof. We begin the proof by recalling that, during epoch k (of length τ_k), each system $i \in [M]$ uses its collected dataset $\mathcal{D}^{(i)} = \{X^{(i)}, Z^{(i)}\}$ to perform model identification through clustered system identification. As described in Algorithm 2, the state–input data are utilized both for estimating the cluster identity and for learning the model parameters. In the latter step, in particular, we store the gradient preconditioned by the inverse of the empirical state–input covariance matrix:

$$G_\ell(\widehat{\Theta}^{(i)}) \leftarrow (X^{(\ell)} - \widehat{\Theta}^{(i)}Z^{(\ell)})Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1}$$

of each system that has the same cluster identity as the i system. The model parameter os system i is then updated as follows:

$$\widehat{\Theta}_{n+1}^{(i)} = \widehat{\Theta}_n^{(i)} + \eta F(G_\ell(\widehat{\Theta}_n^{(i)}), \dots, G_i(\widehat{\Theta}_n^{(i)}), \dots, G_s(\widehat{\Theta}_n^{(i)})), \tag{21}$$

where $\mathcal{C}_{\hat{j}} = \{\ell, \dots, i, \dots, s\}$ is the set of indices of all systems that belong to the identified cluster $\hat{j}$ of system $i \in [M]$. From (21) we can write

$$\begin{aligned}
\widehat{\Theta}_{n+1}^{(i)} &= \widehat{\Theta}_n^{(i)} + \eta\left(\bar{G} + F\left(\left\{G_\ell(\widehat{\Theta}_n^{(i)})\right\}_{\ell\in\mathcal{C}_{\hat{j}}}\right) - \bar{G}\right) \\
&= \widehat{\Theta}_n^{(i)} + \eta\bar{G} + \eta\left(F\left(\left\{G_\ell(\widehat{\Theta}_n^{(i)})\right\}_{\ell\in\mathcal{C}_{\hat{j}}}\right) - \bar{G}\right) \\
&= \widehat{\Theta}_n^{(i)} + \underbrace{\frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}} G_l(\widehat{\Theta}_n^{(i)})}_{\text{Correct classified honest systems}} + \underbrace{\frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}} G_l(\widehat{\Theta}_n^{(i)})}_{\text{Misclassified honest systems}} + \underbrace{\eta\left(F\left(\left\{G_\ell(\widehat{\Theta}_n^{(i)})\right\}_{\ell\in\mathcal{C}_{\hat{j}}}\right) - \bar{G}\right)}_{\text{resilient aggregation error}},
\end{aligned}$$

where $\bar{G} = \frac{1}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_{\hat{j}}} G_l(\widehat{\Theta}_n^{(i)})$ denotes the average of honest system's gradient updates. In addition, $|\mathcal{H}_j \cap \mathcal{H}_{\hat{j}}|$ and $\mathcal{H}_j^c \cap \mathcal{H}_{\hat{j}}$ denote the number of honest systems that are correct classified to cluster $\hat{j}$ and misclassified to cluster $\hat{j}$, respectively. Here, $\mathcal{H}_j^c$ denotes the complement of the set of honest systems in cluster $j \in [N_c]$. We proceed our analysis, by controlling the estimation error in the correct classified honest systems. For this purpose, let us define S_1 = misclassified honest systems term + resilient aggregation error term, and write the following:

$$
\begin{aligned}
\widehat{\Theta}_{n+1}^{(i)} &= \widehat{\Theta}_n^{(i)} + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}} G_l(\widehat{\Theta}_n^{(i)}) + S_1 \\
&= \widehat{\Theta}_n^{(i)} + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}} (X^{(\ell)} - \widehat{\Theta}^{(i)}Z^{(\ell)})Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_1 \\
&= \widehat{\Theta}_n^{(i)} + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}} (\Theta_\star^{(\ell)}Z^{(\ell)} + W^{(\ell)} - \widehat{\Theta}^{(i)}Z^{(\ell)})Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_1 \\
&= \widehat{\Theta}_n^{(i)} + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}} \left(\left(\Theta_\star^{(i)} + \Theta_\star^{(\ell)} - \Theta_\star^{(i)}\right)Z^{(\ell)} + W^{(\ell)} - \widehat{\Theta}^{(i)}Z^{(\ell)}\right)Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_1 \\
&= \widehat{\Theta}_n^{(i)} + \eta\left(\Theta_\star^{(i)} - \widehat{\Theta}_n^{(i)}\right) + \underbrace{\frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}}\left(\Theta_\star^{(\ell)} - \Theta_\star^{(i)}\right)}_{\text{System heterogeneity}} \\
&+ \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}} W^{(\ell)}Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_1.
\end{aligned}
$$

We now proceed to control S_1. To do so, let us first denote by S_2 the resilient aggregation error term, and write the following:

$$
\begin{aligned}
S_1 &= \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}} G_l(\widehat{\Theta}_n^{(i)}) + S_2 \\
&= \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}} (\Theta_j Z^{(\ell)} + W^{(\ell)} - \widehat{\Theta}^{(i)}Z^{(\ell)})Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_2 \\
&= \underbrace{\frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}} (\Theta_j - \widehat{\Theta}^{(i)})}_{\text{Misclassification error}} + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}} W^{(\ell)}Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_2
\end{aligned}
$$

where Θ_j denotes the common model that system q would use if it were correctly classified to its true cluster, which is different from $\hat{j}$. Therefore, we obtain

$$
\begin{aligned}
\widehat{\Theta}_{n+1}^{(i)} = \widehat{\Theta}_n^{(i)} + \eta\left(\Theta_\star^{(i)} - \widehat{\Theta}_n^{(i)}\right) + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_{\hat{j}}}\left(\Theta_\star^{(\ell)} - \Theta_\star^{(i)}\right) + \frac{\eta|\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}|}{|\mathcal{H}_{\hat{j}}|}(\Theta_j - \widehat{\Theta}^{(i)}) \\
+ \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_{\hat{j}}} W^{(\ell)}Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_2
\end{aligned}
$$

where we can subtract $\Theta_\star^{(i)}$ from both sides to write

$$\widehat{\Theta}_{n+1}^{(i)} - \Theta_\star^{(i)} = \widehat{\Theta}_n^{(i)} - \Theta_\star^{(i)} + \eta\left(\Theta_\star^{(i)} - \widehat{\Theta}_n^{(i)}\right) + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_{\hat{j}}}\left(\Theta_\star^{(\ell)} - \Theta_\star^{(i)}\right) + \frac{\eta|\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}|}{|\mathcal{H}_{\hat{j}}|}(\Theta_j - \widehat{\Theta}^{(i)}) + \frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_{\hat{j}}} W^{(\ell)}Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1} + S_2$$

which implies

$$\|\widehat{\Theta}_{n+1}^{(i)} - \Theta_\star^{(i)}\|_F \le (1-\eta)\,\|\widehat{\Theta}_n^{(i)} - \Theta_\star^{(i)}\|_F + \eta\epsilon_{\text{het}} + \underbrace{\frac{\eta|\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}|}{|\mathcal{H}_{\hat{j}}|}\|\Theta_j - \widehat{\Theta}_\star^{(i)}\|_F}_{\text{misclassification error (18)}} + \underbrace{\frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_{\hat{j}}}\|W^{(\ell)}Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1}\|_F}_{\text{statistical error bound (17)}} + S_2. \tag{22}$$

As discussed previously in the proof of Lemma E.1, we have that

$$\frac{\eta}{|\mathcal{H}_{\hat{j}}|}\sum_{\ell\in\mathcal{H}_{\hat{j}}}\|W^{(\ell)}Z^{(\ell)\top}(Z^{(\ell)}Z^{(\ell)\top})^{-1}\|_F \le \frac{\eta C_{\text{stat}}\sigma_w\sqrt{(d_x^2+d_xd_u)\log(1/\delta)}}{\sigma_k\sqrt{\tau_k m_{\hat{j}}}}, \tag{23}$$

$$\frac{\eta|\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}|}{|\mathcal{H}_{\hat{j}}|}\|\Theta_j - \widehat{\Theta}_\star^{(i)}\|_F \le \eta C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_k^2\tau_k). \tag{24}$$

We now proceed to control the resilient aggregation error. For this purpose, we will use Definition 2.1 to write

$$\|S_2\|_F \le \eta\left\|F\left(\left\{G_\ell(\widehat{\Theta}_n^{(i)})\right\}_{\ell\in\mathcal{C}_{\hat{j}}}\right) - \bar{G}\right\|_F \le \eta\lambda\sqrt{d_x}\max_{s,p\in\mathcal{H}_{\hat{j}}}\left\|G_s(\widehat{\Theta}_n^{(i)}) - G_p(\widehat{\Theta}_n^{(i)})\right\|_F.$$

where the dimensionality factor $\sqrt{d_x}$ is due to upper bounding the Frobenius norm with the spectral norm of the corresponding matrix. We then proceed by examining the sets of system indices that are correctly and incorrectly assigned to cluster $\hat{j}$.

$$\|S_2\|_F \le \eta\lambda\sqrt{d_x}\left(\max_{s,p\in\mathcal{H}_j\cap\mathcal{H}_{\hat{j}}}\left\|G_s(\widehat{\Theta}_n^{(i)}) - G_p(\widehat{\Theta}_n^{(i)})\right\|_F + \max_{s,p\in\mathcal{H}_j^c\cap\mathcal{H}_{\hat{j}}}\left\|G_s(\widehat{\Theta}_n^{(i)}) - G_p(\widehat{\Theta}_n^{(i)})\right\|_F\right),$$

where we note that

$$\begin{aligned}\left\|G_s(\widehat{\Theta}_n^{(i)}) - G_p(\widehat{\Theta}_n^{(i)})\right\|_F &\le \left\|\Theta_\star^{(s)} - \Theta_\star^{(p)}\right\|_F + \left\|W^{(s)}Z^{(s)\top}(Z^{(s)}Z^{(s)\top})^{-1}\right\|_F \\ &\quad + \left\|W^{(p)}Z^{(p)\top}(Z^{(p)}Z^{(p)\top})^{-1}\right\|_F \\ &\le \epsilon_{\text{het}} + \frac{2\sqrt{d_x}C_{\text{stat}}\sigma_w\sqrt{(d_x^2+d_xd_u)\log(1/\delta)}}{\sigma_k\sqrt{\tau_k}}.\end{aligned} \tag{25}$$

It is worth noting that the first term corresponds to the intra-cluster system heterogeneity. This term arises in the adversarial aggregation error, as the server cannot distinguish between deviations caused by adversarial attacks and those due to natural heterogeneity. It reflects a fundamental limit of adversarially robust learning under heterogeneous systems, consistent with the lower bound established in Karimireddy et al. [2020] for the Byzantine federated learning setting.

Therefore, by plugging (25) into (22), we obtain

$$\|\widehat{\Theta}_{n+1}^{(i)} - \Theta_\star^{(i)}\|_F \leq (1-\eta)\,\|\widehat{\Theta}_n^{(i)} - \Theta_\star^{(i)}\|_F + \eta(1+2\lambda)\epsilon_{\text{het}} + \eta C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_k^2\tau_k) + \frac{\eta C_{\text{stat}}\sigma_w\sqrt{(d_x^2+d_xd_u)\log(1/\delta)}}{\sigma_k\sqrt{\tau_k m_{\hat{j}}}} + \frac{4\eta\lambda\sqrt{d_x}C_{\text{stat}}\sigma_w\sqrt{(d_x^2+d_xd_u)\log(1/\delta)}}{\sigma_k\sqrt{\tau_k}}.$$

Therefore, by unrolling the above expression over N iterations we obtain

$$\|\widehat{\Theta}_N^{(i)} - \Theta_\star^{(i)}\|_F \leq \rho^N\|\widehat{\Theta}_n^{(i)} - \Theta_\star^{(i)}\|_F + (1+2\lambda)\epsilon_{\text{het}} + C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_k^2\tau_k) + \frac{C_{\text{stat}}\sigma_w\sqrt{(d_x^2+d_xd_u)\log(1/\delta)}}{\sigma_k\sqrt{\tau_k m_{\hat{j}}}} + \frac{\lambda\sqrt{d_x}C_{\text{stat}}\sigma_w\sqrt{(d_x^2+d_xd_u)\log(1/\delta)}}{\sigma_k\sqrt{\tau_k}}.$$

with $\rho = 1-\eta$. Note that any additional constant factors are absorbed by C_{stat}, $C_{\text{mis},1}$ and $C_{\text{mis},2}$. Moreover, we can set the total number of iterations as $N \geq \log\left(C_\alpha\Delta_{\min}\tau_1^2\right)/\log(1/\rho)$, since the initial model estimate is as in Assumption 2. Therefore, the contraction term is of order $\mathcal{O}(1/\tau_k^2)$ and becomes negligible in the subsequent regret analysis. We then obtain the following estimation error bound:

$$\|\widehat{\Theta}_N^{(i)} - \Theta_\star^{(i)}\|_F^2 \leq \frac{C_{\text{stat}}\sigma_w^2(d_x^2+d_xd_u)\log(1/\delta)}{\sigma_k^2\tau_k}\left(\frac{1}{m_{\hat{j}}} + \lambda^2 d_x\right) + C_{\text{het}}(1+\lambda)^2\epsilon_{\text{het}}^2 + C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_k^2\tau_k),$$

for some universal constant C_{het}. We complete the proof by noting that the misclassification rate decays exponentially with the epoch length. By choosing a sufficiently large initial epoch length τ_0, and doubling the epoch length for subsequent epochs, we ensure that the misclassification rate becomes negligible, guaranteeing that the systems are correctly classified. Consequently, we have $\hat{j} \approx j$, and therefore $m_{\hat{j}} \approx m_j$. □

F Characterizing the Probability of the Success Event

We now demonstrate that the success event holds with high probability, that is, the event in which Algorithm 1 does not abort and the estimation error bounds for Cases 1, 2, and 3 hold, as given in (6), (7), and (8), respectively. Below, we establish the probability of success for Case 1 and note that the proof can be readily extended to Cases 2 and 3 by incorporating the additional heterogeneity bias and adversarial-aware terms that appear in those settings.

Lemma F.1. *Running Algorithm 1 with the arguments defined in Lemma E.1, the event* $\mathcal{E}_{\mathsf{success}}^{(j)}$, *for* $j = 1, 2, 3$ *holds with probability at least* $1 - T^{-2}$.

Proof. We show that the success event $\mathcal{E}^{(1)}_{\text{success}}$ holds with probability at least $1-T^{-2}$ by induction. In particular, we prove that for every epoch $k \in [k_{\text{fin}}]$, Algorithm 1 does not abort and the estimation error remains bounded as specified in $\mathcal{E}_{\text{est},1}$. The base case corresponds to the first epoch.

Base case: For convenience we assume that $x_0^{(i)} = 0$, for all systems $i \in [M]$[3].Note that we can use Lemma D.5 to obtain

$$\max_{0\le t\le T-1} \left\| \begin{bmatrix} w_t^{(i)} \\ g_t^{(i)} \end{bmatrix} \right\| \le 4\sigma_w \sqrt{3(d_x+d_u)\log(2MT)}, \tag{26}$$

with probability $1-\frac{1}{2}T^{-2}$, for all systems $i \in [M]$. As our initial state is zero, we can guarantee that

$$\left\| x_1^{(h)} \right\| \le 16 (P_0^{\vee})^{3/2} \Psi_B^{\vee} \max_{0\le t\le T-1} \left\| \begin{bmatrix} w_t^{(i)} \\ g_t^{(i)} \end{bmatrix} \right\|,$$

where $P_0^{\vee} \triangleq \max_{i\in[M]} \left\| P^{(i)}_{K_0^{(i)}} \right\|$, $P_\star^{\wedge} \triangleq \min_{i\in[M]} \left\| P_\star^{(i)} \right\|$, and $\Psi_B^{\vee} \triangleq \max_{i\in[M]} \Psi_B^{(i)}$. Therefore, by setting the initial epoch length according to $\tau_1 \ge \frac{c \log \frac{1}{P_\star^{\wedge}}}{\log\left(1-\frac{1}{P_\star^{\wedge}}\right)}$, for a sufficiently large constant c, we can use Lemma D.6 to obtain

$$\left\| x_t^{(i)} \right\| \le 40 (P_0^{\vee})^{2} \Psi_B^{\vee} \max_{0\le t\le T-1} \left\| \begin{bmatrix} w_t^{(i)} \\ g_t^{(i)} \end{bmatrix} \right\|, \forall t = \{0,1,\dots,\tau_1\}, \tag{27}$$

where we use (26) to obtain

$$\left\| x_t^{(i)} \right\|^2 \le 76800 (P_0^{\vee})^{4} (\Psi_B^{\vee})^2 \sigma_w^2 (d_x+d_u)\log(2MT), \quad \forall t = \{0,1,\dots,\tau_1\}$$

with probability $1-\frac{1}{2}T^{-2}$, for all systems $i \in [M]$, which implies that $\left\| x_t^{(i)} \right\|^2 \le x_b^2 \log T$ which satisfies the state norm requirement to not abort as described in Algorithm 1. On the other hand, to verify that the controller norm requirement is satisfied we note that $\|K_0^{(i)}\|^2 \le P_0^{\vee} \le 2P_0^{\vee}$, which leads to $\|K_0^{(i)}\| \le K_b$. Therefore, we have that $\mathcal{E}_{\text{bound}}$ holds with probability $1-\frac{1}{2}T^{-2}$.

We now proceed to control the estimation error at the first epoch. Note that by making τ_1 sufficiently large such that $\tau_1 \ge c(1+\log(4T^2))$, for a sufficiently large constant c, we can guarantee the following upper bound for the model estimation:

$$\left\| \begin{bmatrix} \hat{A}_1^{(i)} & \hat{B}_1^{(i)} \end{bmatrix} - \begin{bmatrix} A_\star^{(i)} & B_\star^{(i)} \end{bmatrix} \right\|_F^2 \le \frac{C_{\text{stat}} \sigma_w^2 (d_x^2 + d_x d_u)\log(2T^2)}{\sigma_1^2 M_j \tau_1} + C_{\text{mis},1} \exp\left(-C_{\text{mis},2}\sigma_1^2 \tau_1\right)$$

for any system $i \in C_j$ as discussed in Lemma E.1. Our induction step follows by considering the following inductive hypothesis:

$$\textbf{Bounded state:} \left\| x_{\tau_k}^{(i)} \right\| \le 16 (P_0^{\vee})^{3/2} \Psi_B^{\vee} \max_{0\le t\le T-1} \left\| \begin{bmatrix} w_t^{(i)} \\ g_t^{(i)} \end{bmatrix} \right\|, \tag{28}$$

[3] This proof can also be extended to bounded non-zero initial states.

and

Estimation error:

$$\left\| \begin{bmatrix} \hat{A}_k^{(i)} & \hat{B}_k^{(i)} \end{bmatrix} - \begin{bmatrix} A_\star^{(i)} & B_\star^{(i)} \end{bmatrix} \right\|_F^2 \leq \frac{C_{\text{stat}} \sigma_w^2 (d_x^2 + d_x d_u) \log(2T^2)}{\sigma_k^2 M_j \tau_k} + C_{\text{mis},1} \exp\left(-C_{\text{mis},2} \sigma_k^2 \tau_k\right), \quad (29)$$

where we demonstrate that the estimation error holds for the next epoch $k+1$ by balancing the exploration and exploitation as $\tau_k \sigma_k^2 \geq \frac{1}{2} \tau_{k+1} \sigma_{k+1}^2$. In addition, we guarantee that the state and controller norm bounds are not violated by combining the fact that $\left\| P_{\hat{K}_{k+1}}^{(i)} \right\| \leq 2(P_0^\vee)$ and $\tau_k \geq \tau_1 \geq \frac{c \log \frac{1}{P_\star^\wedge}}{\log\left(1 - \frac{1}{P_\star^\wedge}\right)}$, for a sufficiently large constant c, along with Lemma D.6 to obtain

$$\left\| x_t^{(i)} \right\| \leq 40 (P_0^\vee)^2 (\Psi_B^\vee) \max_{1 \leq t \leq T} \left\| \begin{bmatrix} w_t^{(i)} \\ g_t^{(i)} \end{bmatrix} \right\|, \quad \forall t = \{\tau_k + 1, \ldots, \tau_{k+1}\}, \quad (30)$$

which guarantees that the state norm requirement is satisfied. For the controller norm, we have that $\left\| \hat{K}_{k+1}^{(i)} \right\|^2 \leq \left\| P_{\hat{K}_{k+1}} \right\| \leq 2P_0^\vee$, and thus $\left\| \hat{K}_{k+1}^{(i)} \right\| \leq K_b$. We conclude the proof by noting that for epoch k the bounded state and controller norm requirements hold with probability $1 - \frac{1}{2} T^{-2}$, then for epoch $k+1$ also holds with at least the same probability. Then, by union bounding for all the epochs, we have that $\mathcal{E}_{\mathsf{success}}$ holds under probability of at least $1 - T^{-2}$. □

G Regret Analysis

With the estimation error bounds and the probability of the success event in place, we are now ready to synthesize the regret bounds for the three settings under consideration: (i) no intra-cluster heterogeneity, (ii) intra-cluster heterogeneity, and (iii) the presence of adversarial systems. In this section, we leverage the adapted regret bounds from Lee et al. [2023], revisited in Lemmas D.2, D.3, and D.4, and derive conditions on the exploration sequence $\{\sigma_k\}_k$ and the epoch length τ_k (i.e., exploitation phase) that ensure a reduction of the leading term in the regret proportional to the number of honest systems participating in the collaboration. For completeness, we restate here the key theorems presented in the main body of the paper.

G.1 Regret with Intra-cluster Homogeneity

We begin with the setting in which the models within each cluster $\mathcal{C}_j$, for any $j \in [N_c]$, are identical, and each cluster contains M_j models.

Theorem G.1 (Intra-cluster homogeneity)**.** *Fix a system $i \in [M]$ that belongs to a homogeneous cluster $\mathcal{C}_j$ of size M_j. Let the assumptions of Algorithm (2) hold. Suppose the exploration sequence satisfies $\sigma_k^2 = \frac{\sqrt{d_u^2 d_x}}{(d_x^2 + d_x d_u)\sqrt{\tau_k M_j}}$ and the epoch length doubles, i.e., $\tau_k = 2^{k-1}\tau_1$ with $T = \tau_{k_{\text{fin}}}$. Then the expected regret of system i satisfies*

$$\mathbb{E}\left[\mathcal{R}_T^{(i)}\right] \leq \Omega_1 \sqrt{\frac{d_u^2 d_x T}{M_j}} + \Omega_2 (\log T)^2 + \Omega_3 T \exp\left(-\frac{C_{\text{mis},2}\sqrt{\tau_1}}{\sqrt{M_j}}\right), \quad (31)$$

where

$$\begin{aligned}\Omega_1 &= 142C_{stat}\|P_\star^{(i)}\|^8\sigma_w^2\log(1/\delta) + 2d_u + 4d_u\|P_\star^{(i)}\|\Psi_B^{(i)2},\\ \Omega_2 &= 3\,\max\{d_x,d_u\}\,\|P_{K_0^{(i)}}^{(i)}\|\,\Psi_B^{(i)2} + 2x_b^2\,\|P_{K_\star}^{(i)}\|,\\ \Omega_3 &= 142C_{mis,1}\|P_\star^{(i)}\|^8.\end{aligned}$$

All constants above are as defined in the lemmas referenced in the proof.

Proof. By Lemma D.2, Lemma D.3, Lemma D.4, and the regret decomposition in (9), we have

$$\begin{aligned}\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le \sum_{k=2}^{k_{\mathrm{fin}}}\Bigg(&\mathbb{E}\Big[\mathbf{1}(\mathcal{E}_{\mathrm{est},1}^{(k-1)})\;142\,(\tau_k-\tau_{k-1})\,\|P_\star^{(i)}\|^8\,\big\|\,[\widehat{A}_{k-1}^{(i)}\;\widehat{B}_{k-1}^{(i)}]-[A_\star^{(i)}\;B_\star^{(i)}]\big\|_F^2\Big]\\ &+(\tau_k-\tau_{k-1})\,J^{(i)}(K_\star^{(i)})+4(\tau_k-\tau_{k-1})\,d_u\,\|P_\star^{(i)}\|\sigma_k^2\Psi_B^{(i)2}+2x_b^2\log T\,\|P_\star^{(i)}\|\Bigg)\\ &+T^{-1}\big(\|Q\|+2K_b^2\big)x_b^2\log T+T^{-1}J(K_0)\\ &+24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x+d_u)\,\sigma_w^2T^{-2}\log(3T)+2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\|\Theta_\star^{(i)}\|_F^2K_b^2x_b^2\log T\\ &+\sum_{k=1}^{k_{\mathrm{fin}}}2(\tau_k-\tau_{k-1})d_u\,\sigma_k^2+3\tau_1\max\{d_x,d_u\}\|P_{K_0^{(i)}}^{(i)}\|\Psi_{B_\star}^2-TJ(K_\star^{(i)}).\end{aligned}$$

In the homogeneous case and by (6), for each epoch $k\ge 2$,

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}(\mathcal{E}_{\mathrm{est},1}^{(k-1)})\,\big\|\,[\widehat{A}_{k-1}^{(i)}\;\widehat{B}_{k-1}^{(i)}]-[A_\star^{(i)}\;B_\star^{(i)}]\,\big\|_F^2\right] \le{}& \frac{C_{\mathrm{stat}}\sigma_w^2(d_x^2+d_xd_u)\log(1/\delta)}{\sigma_{k-1}^2M_j\tau_{k-1}}\\ &+C_{\mathrm{mis},1}\exp\left(-C_{\mathrm{mis},2}\,\sigma_{k-1}^2\,\tau_{k-1}\right).\end{aligned}$$

Substituting this bound yields

$$\begin{aligned}\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le \sum_{k=2}^{k_{\mathrm{fin}}}\Bigg(&142(\tau_k-\tau_{k-1})\|P_\star^{(i)}\|^8\Big(\frac{C_{\mathrm{stat}}\sigma_w^2(d_x^2+d_xd_u)\log(1/\delta)}{\sigma_{k-1}^2M_j\tau_{k-1}}\\ &+C_{\mathrm{mis},1}\exp\left(-C_{\mathrm{mis},2}\,\sigma_{k-1}^2\tau_{k-1}\right)\Big)+(\tau_k-\tau_{k-1})\,J(K^\star)+4(\tau_k-\tau_{k-1})\,d_u\,\|P_\star^{(i)}\|\,\sigma_k^2\,\Psi_B^{(i)2}\\ &+2x_b^2\log T\,\|P_\star^{(i)}\|\Bigg)+T^{-1}\big(\|Q\|+2K_b^2\big)x_b^2\log T+T^{-1}J^{(i)}(K_0^{(i)})\\ &+24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x+d_u)\,\sigma_w^2T^{-2}\log(3T)+2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\|\Theta_\star^{(i)}\|_F^2K_b^2x_b^2\log T\\ &+\sum_{k=1}^{k_{\mathrm{fin}}}2(\tau_k-\tau_{k-1})d_u\,\sigma_k^2+3\tau_1\max\{d_x,d_u\}\|P_{K_0^{(i)}}^{(i)}\|\Psi_{B_\star^{(i)}}^{(i)2}-TJ^{(i)}(K_\star).\end{aligned}$$

Assume $\sigma_k^2=\frac{\sqrt{d_u^2d_x}}{(d_x^2+d_xd_u)\sqrt{\tau_kM_j}}$ and $\tau_k=2^{k-1}\tau_1$. Since we double the epoch length over the epochs, i.e., τ_k increasing, we have that $e^{-C_{\mathrm{mis},2}\sqrt{\tau_k}}\le e^{-C_{\mathrm{mis},2}\sqrt{\tau_1}}$. Using these conditions, we

obtain

$$\begin{aligned}\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le \sum_{k=2}^{k_{\mathrm{fin}}} \Bigg(& 142(\tau_k-\tau_{k-1})\|P_\star^{(i)}\|^8 \Big(\frac{C_{\mathrm{stat}}\sigma_w^2\sqrt{d_u^2 d_x}\log(1/\delta)}{\sqrt{M_j\tau_{k-1}}} + C_{\mathrm{mis},1}\exp\left(-\frac{C_{\mathrm{mis},2}\sqrt{\tau_1}}{\sqrt{M_j}}\right) \\ & + (\tau_k-\tau_{k-1})J^{(i)}(K_\star^{(i)}) + 4(\tau_k-\tau_{k-1})d_u\sqrt{d_u^2 d_x}\|P_\star^{(i)}\|\frac{\Psi_B^{(i)2}}{\sqrt{\tau_k M_j}} + 2x_b^2\log T\,\|P_\star^{(i)}\| \Bigg) \\ & + T^{-1}\big(\|Q\|+2K_b^2\big)x_b^2\log T + T^{-1}J^{(i)}(K_0^{(i)}) \\ & + 24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x+d_u)\sigma_w^2 T^{-2}\log(3T) + 2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\|\Theta_\star^{(i)}\|_F^2 K_b^2 x_b^2\log T \\ & + \sum_{k=1}^{k_{\mathrm{fin}}} 2(\tau_k-\tau_{k-1})d_u\sqrt{d_u^2 d_x}\,\frac{1}{\sqrt{\tau_k M_i}} + 3\tau_1\max\{d_x,d_u\}\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2} - TJ^{(i)}(K_\star^{(i)}).\end{aligned}$$

Note that $\sum_{k=2}^{k_{\mathrm{fin}}}(\tau_k-\tau_{k-1}) = \tau_{k_{\mathrm{fin}}}-\tau_1 = T-\tau_1$ and $\sum_{k=2}^{k_{\mathrm{fin}}} 1 = k_{\mathrm{fin}}-1 \asymp \log T$. Moreover, a standard estimate for the doubling epoch length gives

$$\sum_{k=2}^{k_{\mathrm{fin}}} \frac{\tau_k-\tau_{k-1}}{\sqrt{\tau_{k-1}}} \le 2\big(\sqrt{T}-\sqrt{\tau_1}\big) \;\le\; 2\sqrt{T}.$$

By using these facts and separating terms, we obtain

$$\begin{aligned}\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le\; & 142\|P_\star^{(i)}\|^8 C_{\mathrm{stat}}\sigma_w^2\sqrt{d_u^2 d_x}\log(1/\delta)\sum_{k=2}^{k_{\mathrm{fin}}}\frac{\tau_k-\tau_{k-1}}{\sqrt{\tau_{k-1}M_j}} + 2d_u\sqrt{d_u^2 d_x}\sum_{k=1}^{k_{\mathrm{fin}}}\frac{\tau_k-\tau_{k-1}}{\sqrt{\tau_k M_j}} \\ & + 4d_u\sqrt{d_u^2 d_x}\|P_\star^{(i)}\|\Psi_B^{(i)2}\sum_{k=2}^{k_{\mathrm{fin}}}\frac{\tau_k-\tau_{k-1}}{\sqrt{\tau_k M_j}} + 2x_b^2\|P_\star^{(i)}\|(\log T)\sum_{k=2}^{k_{\mathrm{fin}}}1 \\ & + 142\|P_\star^{(i)}\|^8 C_{\mathrm{mis},1}\exp\left(-\frac{C_{\mathrm{mis},2}\sqrt{\tau_1}}{\sqrt{M_j}}\right)\sum_{k=2}^{k_{\mathrm{fin}}}(\tau_k-\tau_{k-1}) \\ & + T^{-1}\big(\|Q\|+2K_b^2\big)x_b^2\log T + T^{-1}J^{(i)}(K_0^{(i)}) \\ & + 24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x+d_u)\,\sigma_w^2 T^{-2}\log(3T) + 2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\|\Theta_\star^{(i)}\|_F^2 K_b^2 x_b^2\log T \\ & + \big(T-\tau_1\big)J^{(i)}(K_\star^{(i)}) - TJ(K_\star^{(i)}) + 3\tau_1\max\{d_x,d_u\}\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}.\end{aligned}$$

Therefore, since $(T-\tau_1)J(K^\star) - TJ(K^\star) = -\tau_1 J(K^\star) \le 0$, we may drop this (non-positive) term. By applying $\sum_k(\tau_k-\tau_{k-1})/\sqrt{\tau_k} \le 2\sqrt{T}$, $\sum_{k=2}^{k_{\mathrm{fin}}}1 \lesssim \log T$, and $\sum_{k=2}^{k_{\mathrm{fin}}}(\tau_k-\tau_{k-1}) = T-\tau_1 \le T$, we obtain

$$\begin{aligned}\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le\; & \Big(142\|P_\star^{(i)}\|^8 C_{\mathrm{stat}}\sigma_w^2(d_x^2+d_x d_u)\log(1/\delta)\sigma^2 + 2d_u + 4d_u\|P_\star^{(i)}\|\Psi_{B^\star}^2\Big)\sqrt{\frac{T}{M_i}} \\ & + \Big(3\max\{d_x,d_u\}\|P_{K_0}\|\Psi_{B_\star}^2 + 2x_b^2\|P_\star^{(i)}\|\Big)(\log T)^2 \\ & + 142\,\|P_\star^{(i)}\|^8 C_{\mathrm{mis},1}\exp\left(-\frac{C_{\mathrm{mis},2}\sqrt{\tau_1}}{\sqrt{M_j}}\right)T,\end{aligned}$$

which is precisely the claimed bound with the stated Ω_1, Ω_2,and Ω_3. We also emphasize that by choosing the initial epoch length as $\tau_1 \geq M_j \log(C_{\text{mis},1}T^2)/C^2_{\texttt{mis},2}$, so that $Te^{-C_{\text{mis},2}\tau_1}$ is negligible (i.e., $\mathcal{O}(1)$) and the T^{-1} and T^{-2} terms vanish as T grows. Absorbing the remaining τ_1-dependent constant into the $(\log T)^2$ term yields (31) without the misclassification term. □

G.2 Regret with Intra-cluster Heterogeneity

Next, we consider the setting where the models within the cluster $\mathcal{C}_j$ are similar but not identical, with bounded heterogeneity quantified by ϵ_{het} (see Assumption 1).

Corollary G.1 (Intra-cluster heterogeneity)**.** *Fix a system i that belongs to a heterogeneous cluster $\mathcal{C}_j$ of size M_j. Let the assumptions of Algorithm (2) hold. Suppose the exploration sequence satisfies $\sigma_k^2 = \frac{\sqrt{d_u^2 d_x}}{(d_x^2+d_x d_u)\sqrt{\tau_k M_j}}$ and the epoch length doubles, i.e., $\tau_k = 2^{k-1}\tau_1$ with $T = \tau_{k_{\text{fin}}}$. Then the expected regret of system i satisfies*

$$\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \leq \Omega_1\sqrt{\frac{d_u^2 d_x T}{M_j}} + \Omega_2\,(\log T)^2 + \Omega_3 T\exp\left(-\frac{C_{\text{mis},2}\sqrt{\tau_1}}{\sqrt{M_j}}\right) + \Omega_4 T\epsilon_{\text{het}}^2, \tag{32}$$

with additional constant $\Omega_4 = 142C_{het}\|P_\star^{(i)}\|^8$.

Proof. The proof follows directly from the derivations established in Theorem G.1. We recall that the expected regret can be decomposed as follows

$$\begin{aligned}
\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \leq \sum_{k=2}^{k_{\text{fin}}} \Bigg(& \mathbb{E}\Big[\mathbf{1}\big(\mathcal{E}_{\text{est},2}^{(k-1)}\big)\; 142\,(\tau_k - \tau_{k-1})\,\|P_\star^{(i)}\|^8 \,\big\|\,[\widehat{A}_{k-1}^{(i)}\;\widehat{B}_{k-1}^{(i)}] - [A_\star^{(i)}\;B_\star^{(i)}]\,\big\|_F^2\Big] \\
& + (\tau_k - \tau_{k-1})\,J^{(i)}(K_\star^{(i)}) + 4(\tau_k - \tau_{k-1})\,d_u\,\|P_\star^{(i)}\|\,\sigma_k^2\,\Psi_B^{(i)2} + 2x_b^2\log T\,\|P_\star^{(i)}\|\Bigg) \\
& + T^{-1}\big(\|Q\| + 2K_b^2\big)x_b^2\log T + T^{-1}J^{(i)}(K_0^{(i)}) \\
& + 24\|P^{(i)}_{K_0^{(i)}}\|\Psi_B^{(i)2}(d_x+d_u)\sigma_w^2 T^{-2}\log(3T) + 2T^{-2}\|P^{(i)}_{K_0^{(i)}}\|\|\Theta_\star^{(i)}\|_F^2 K_b^2 x_b^2\log T \\
& + \sum_{k=1}^{k_{\text{fin}}} 2(\tau_k - \tau_{k-1})d_u\sigma_k^2 + 3\tau_1\max\{d_x,d_u\}\|P^{(i)}_{K_0^{(i)}}\|\Psi_B^{(i)2} - TJ^{(i)}(K_\star^{(i)}).
\end{aligned}$$

We note that in the heterogeneous setting, the only additional contribution arises from the deviation between the true system parameters within each cluster (i.e., intra-cluster heterogeneity). This effect appears in the first term through the heterogeneity bound, which introduces an additional bias term proportional to ϵ_{het}^2. Consequently, the regret bound becomes

$$\begin{aligned}
\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \leq{}& \Omega_1\sqrt{\frac{d_u^2 d_x T}{M_j}} + \Omega_2\,(\log T)^2 + \Omega_3 T\exp\left(-\frac{C_{\text{mis},2}\sqrt{\tau_1}}{\sqrt{M_j}}\right) \\
& + \sum_{k=2}^{k_{\text{fin}}} 142\,(\tau_k - \tau_{k-1})\,\|P_\star^{(i)}\|^8 C_{\text{het}}\epsilon_{\text{het}}^2.
\end{aligned}$$

By evaluating the summation over epochs and applying the doubling schedule $\tau_k = 2^{k-1}\tau_1$, we obtain the final bound

$$\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \leq \Omega_1 \sqrt{\frac{d_u^2 d_x T}{M_j}} \; + \; \Omega_2 \left(\log T\right)^2 + \Omega_3 T \exp\left(-\frac{C_{\mathrm{mis},2}\sqrt{\tau_1}}{\sqrt{M_j}}\right) + 142 \left\|P_\star^{(i)}\right\|^8 C_{\mathrm{het}} T \epsilon_{\mathrm{het}}^2,$$

which completes the proof. □

G.3 Regret with Adversarial Systems

We now consider the setting where a given cluster $\mathcal{C}_j$ consists of f_j adversarial systems and $m_j = M_j - f_j$ honest systems. The effect of the adversarial systems on the regret bounds manifests through the resilient coefficient λ of the underlying resilient aggregation rule (see Definition 2.1). Well-known resilient aggregation rules such as the coordinate-wise trimmed mean (CWTM) and the mean-around-median (MeaMed) typically exhibit a resilient coefficient on the order of $\mathcal{O}(f_j/m_j)$. We refer the reader to Farhadkhani et al. [2022] for more details on such resilient aggregators. In Section 5, we further illustrate our adversarially robust adaptive control approach using GM, whose resilience coefficient scales as $\mathcal{O}\left(1 + \frac{m_j}{\sqrt{M_j - 2f_j}\, M_j}\right)$. In this case, the improvement in regret with respect to the number of honest systems within each cluster may be attenuated.

Theorem G.2 (Adversarial and intra-cluster heterogeneity)**.** *Consider a system i belonging to a heterogeneous cluster C_j of size M_j that may contain a fraction of adversarial systems. Let the assumptions of Algorithm (2) hold. Suppose that the exploration sequence satisfies $\sigma_k^2 = \frac{\sqrt{d_u^2 d_x}}{d_x^2 + d_x d_u}\sqrt{\frac{1+\lambda^2 d_x m_j}{\tau_k m_j}}$ and that the epoch length is double every epoch, i.e., $\tau_k = 2^{k-1}\tau_1$, with total horizon $T = \tau_{k_{\mathrm{fin}}}$. Then, the expected regret of system i satisfies*

$$\begin{aligned}\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \leq \Omega_1 \sqrt{\frac{d_u^2 d_x T(1+\lambda^2 d_x m_j)}{m_j}} + \Omega_2 (\log T)^2 + \Omega_3 T \exp\left(-C_{\mathrm{mis},2}\sqrt{\frac{1+\lambda^2 d_x m_j \tau_1}{m_j}}\right) \\ + \Omega_4 T(1+\lambda^2)\epsilon_{\mathrm{het}}^2.\end{aligned}$$

Proof. We begin the proof by rewriting the regret decomposition as follows:

$$\begin{aligned}\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \leq \sum_{k=2}^{k_{\mathrm{fin}}} \Bigg(& \mathbb{E}\Big[\mathbf{1}(\mathcal{E}_{\mathrm{est},3}^{(k-1)})\; 142\,(\tau_k - \tau_{k-1})\,\|P_\star^{(i)}\|^8 \,\big\|\,[\widehat{A}_{k-1}^{(i)}\;\widehat{B}_{k-1}^{(i)}] - [A_\star^{(i)}\;B_\star^{(i)}]\,\big\|_F^2\Big] \\ & + (\tau_k - \tau_{k-1})\,J^{(i)}(K_\star^{(i)}) + 4(\tau_k - \tau_{k-1}) d_u \|P_\star^{(i)}\| \sigma_k^2\, \Psi_B^{(i)2} + 2x_b^2 \log T \|P_\star^{(i)}\|\Bigg) \\ & + T^{-1}\big(\|Q\| + 2K_b^2\big)x_b^2 \log T + T^{-1} J^{(i)}(K_0^{(i)}) \\ & + 24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x + d_u)\sigma_w^2 T^{-2}\log(3T) + 2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\|\Theta_\star^{(i)}\|_F^2 K_b^2 x_b^2 \log T \\ & + \sum_{k=1}^{k_{\mathrm{fin}}} 2(\tau_k - \tau_{k-1}) d_u\, \sigma_k^2 + 3\tau_1 \max\{d_x, d_u\}\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2} - T J^{(i)}(K_\star^{(i)}),\end{aligned}$$

where for the adversarial setting we have

$$\mathbb{E}\left[\mathbf{1}(\mathcal{E}_{\text{est},3}^{(k-1)}) \left\| [\widehat{A}_{k-1}^{(i)} \ \widehat{B}_{k-1}^{(i)}] - [A_\star^{(i)} \ B_\star^{(i)}] \right\|_F^2\right] \le \frac{C_{\text{stat}}\sigma_w^2(d_x^2+d_xd_u)\log(1/\delta)}{\sigma_{k-1}^2\tau_{k-1}}\left(\frac{1}{m_j}+\lambda^2 d_x\right) + C_{\text{het}}(1+\lambda)^2\epsilon_{\text{het}}^2 + C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_{k-1}^2\tau_{k-1}).$$

Applying this bound in the regret bound yields

$$\begin{aligned}
\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le \sum_{k=2}^{k_{\text{fin}}} \Bigg(& 142(\tau_k-\tau_{k-1})\|P_\star^{(i)}\|^8\Big(\frac{C_{\text{stat}}\sigma_w^2(d_x^2+d_xd_u)\log(1/\delta)}{\sigma_{k-1}^2\tau_{k-1}}\left(\frac{1}{m_j}+\lambda^2 d_x\right) \\
& + C_{\text{het}}(1+\lambda)^2\epsilon_{\text{het}}^2 + C_{\text{mis},1}\exp(-C_{\text{mis},2}\sigma_{k-1}^2\tau_{k-1})\Big)\Bigg) \\
& + (\tau_k-\tau_{k-1})\,J(K^\star) + 4(\tau_k-\tau_{k-1})\,d_u\,\|P_\star^{(i)}\|\,\sigma_k^2\,\Psi_B^{(i)2} + 2x_b^2\log T\,\|P_\star^{(i)}\|\Bigg) \\
& + T^{-1}\big(\|Q\|+2K_b^2\big)x_b^2\log T + T^{-1}J(K_0) \\
& + 24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x+d_u)\,\sigma_w^2T^{-2}\log(3T) + 2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\|\Theta_\star^{(i)}\|_F^2K_b^2x_b^2\log T \\
& + \sum_{k=1}^{k_{\text{fin}}} 2(\tau_k-\tau_{k-1})d_u\,\sigma_k^2 + 3\tau_1\max\{d_x,d_u\}\|P_{K_0^{(i)}}^{(i)}\|\Psi_{B_\star^{(i)2}}^{(i)} - TJ^{(i)}(K_\star^{(i)}).
\end{aligned}$$

Therefore by setting the exploration sequence as $\sigma_k^2 = \frac{\sqrt{d_u^2d_x}}{d_x^2+d_xd_u}\sqrt{\frac{1+\lambda^2d_xm_j}{\tau_km_j}}$, we obtain

$$\begin{aligned}
\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le \sum_{k=2}^{k_{\text{fin}}} \Bigg(& 142(\tau_k-\tau_{k-1})\|P_\star^{(i)}\|^8\Big(\frac{C_{\text{stat}}\sigma_w^2\sqrt{d_u^2d_x}\log(1/\delta)}{\sqrt{\tau_{k-1}}}\sqrt{\frac{1}{m_j}+\lambda^2d_x} \\
& + C_{\text{het}}(1+\lambda)^2\epsilon_{\text{het}}^2 + C_{\text{mis},1}\exp\left(-C_{\text{mis},2}\sqrt{\frac{1+\lambda^2d_xm_j\tau_1}{m_j}}\right)\Big)\Bigg) \\
& + (\tau_k-\tau_{k-1})J^{(i)}(K_\star^{(i)}) + 4(\tau_k-\tau_{k-1})d_u\sqrt{d_u^2d_x}\|P_\star^{(i)}\|\sqrt{\frac{1+\lambda^2d_xm_j}{\tau_{k-1}m_j}}\,\Psi_B^{(i)2}\Bigg) \\
& + 2x_b^2\log T\|P_\star^{(i)}\| + T^{-1}\big(\|Q\|+2K_b^2\big)x_b^2\log T + T^{-1}J^{(i)}(K_0^{(i)}) \\
& + 24\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2}(d_x+d_u)\,\sigma_w^2T^{-2}\log(3T) + 2T^{-2}\|P_{K_0^{(i)}}^{(i)}\|\|\Theta_\star^{(i)}\|_F^2K_b^2x_b^2\log T \\
& + \sum_{k=1}^{k_{\text{fin}}} 2(\tau_k-\tau_{k-1})d_u\sqrt{d_u^2d_x}\sqrt{\frac{1+\lambda^2d_xm_j}{\tau_{k-1}m_j}} + 3\tau_1\max\{d_x,d_u\}\|P_{K_0^{(i)}}^{(i)}\|\Psi_B^{(i)2} - TJ^{(i)}(K_\star^{(i)}),
\end{aligned}$$

and by using the fact that $k_{\text{fin}} \asymp \log T$ and $\sum_{k=2}^{k_{\text{fin}}}\frac{\tau_k-\tau_{k-1}}{\sqrt{\tau_{k-1}}} \le 2\big(\sqrt{T}-\sqrt{\tau_1}\big) \le 2\sqrt{T}$, we obtain

$$\mathbb{E}\big[\mathcal{R}_T^{(i)}\big] \le \Big(142\|P_\star^{(i)}\|^8C_{\text{stat}}\sigma_w^2\log(1/\delta) + 2d_u + 4d_u\|P_\star^{(i)}\|\Psi_B^{(i)2}\Big)\sqrt{\frac{d_u^2d_xT(1+\lambda^2d_xm_j)}{m_j}}$$

$$
\begin{aligned}
&+ \left(3\max\{d_x, d_u\}\|P^{(i)}_{K_0^{(i)}}\|\Psi_B^{(i)2} + 2x_b^2\|P_\star^{(i)}\|\right)(\log T)^2 \\
&+ 142\|P_\star^{(i)}\|^8 C_{\mathrm{mis},1} \exp\left(-C_{\mathrm{mis},2}\sqrt{\frac{1+\lambda^2 d_x m_j \tau_1}{m_j}}\right) T \\
&+ 142\|P_\star^{(i)}\|^8 C_{\mathrm{het}}(1+\lambda)^2\epsilon_{\mathrm{het}}^2 T,
\end{aligned}
$$

which completes the proof. □