Safe Policy Search for Lifelong Reinforcement Learning with Sublinear Regret
Online Appendix

Haitham Bou Ammar HAITHAMB@SEAS.UPENN.EDU
Rasul Tutunov TUTUNOV@SEAS.UPENN.EDU
Eric Eaton EEATON@CIS.UPENN.EDU
University of Pennsylvania, Computer and Information Science Department, Philadelphia, PA 19104 USA

This is the online appendix to the ICML 2015 paper "Safe Policy Search for Lifelong Reinforcement Learning with Sublinear Regret." Appendix A includes derivations of the update equations for $\boldsymbol{L}$ and $\boldsymbol{S}$ in the special case of Gaussian policies. Appendix B provides detailed proofs of all lemmas from the main paper, leading to the proof of sublinear regret (Theorem 1).

A. Update Equations Derivation

In this appendix, we derive the update equations for $\boldsymbol{L}$ and $\boldsymbol{S}$ in the special case of Gaussian policies. Please note that these derivations can be easily extended to other policy forms in higher dimensional action spaces.

For a task t_j, the policy $\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k,t_j)}|\boldsymbol{x}_m^{(k,t_j)}\right)$ is given by:

$$\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k,t_j)}|\boldsymbol{x}_m^{(k,t_j)}\right) = \frac{1}{\sqrt{2\pi\sigma^2_{t_j}}}\exp\left(-\frac{1}{2\sigma^2_{t_j}}\left(\boldsymbol{u}_m^{(k,t_j)} - \left(\boldsymbol{L}\boldsymbol{s}_{t_j}\right)^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)^2\right).$$

Therefore, the safe lifelong reinforcement learning optimization objective can be written as:

$$\boldsymbol{e}_r(\boldsymbol{L},\boldsymbol{S}) = \sum_{j=1}^{r}\frac{\eta_{t_j}}{2\sigma^2_{t_j}n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\left(\boldsymbol{u}_m^{(k,t_j)} - \left(\boldsymbol{L}\boldsymbol{s}_{t_j}\right)^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)^2 + \mu_1||\boldsymbol{S}||^2_{\mathsf{F}} + \mu_2||\boldsymbol{L}||^2_{\mathsf{F}} \ . \tag{13}$$

To arrive at the update equations, we need to derive Eq. (13) with respect to each $\boldsymbol{L}$ and $\boldsymbol{S}$.

A.1. Update Equations for $\boldsymbol{L}$

Starting with the derivative of $\boldsymbol{e}_r(\boldsymbol{L},\boldsymbol{S})$ with respect to the shared repository $\boldsymbol{L}$, we can write:

$$\begin{aligned}\nabla_{\boldsymbol{L}}\boldsymbol{e}_r(\boldsymbol{L},\boldsymbol{S}) &= \nabla_{\boldsymbol{L}}\left[\sum_{j=1}^{r}\frac{\eta_{t_j}}{2\sigma^2_{t_j}n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\left(\boldsymbol{u}_m^{(k,t_j)} - \left(\boldsymbol{L}\boldsymbol{s}_{t_j}\right)^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)^2 + \mu_1||\boldsymbol{S}||^2_{\mathsf{F}} + \mu_2||\boldsymbol{L}||^2_{\mathsf{F}}\right] \\ &= -\sum_{j=1}^{r}\left[\frac{\eta_{t_j}}{\sigma^2_{t_j}n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\left(\boldsymbol{u}_m^{(k,t_j)} - \left(\boldsymbol{L}\boldsymbol{s}_{t_j}\right)^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}^{\mathsf{T}}_{t_j}\right] + 2\mu_2\boldsymbol{L} \ .\end{aligned}$$

To acquire the minimum, we set the above to zero:

$$\sum_{j=1}^{r}\left[\frac{\eta_{t_j}}{\sigma^2_{t_j}n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\left(\boldsymbol{u}_m^{(k,t_j)} - \left(\boldsymbol{L}\boldsymbol{s}_{t_j}\right)^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}^{\mathsf{T}}_{t_j}\right] + 2\mu_2\boldsymbol{L} = 0$$

$$\sum_{j=1}^{r}\left[\frac{\eta_{t_j}}{\sigma^2_{t_j}n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{s}^{\mathsf{T}}_{t_j}\boldsymbol{L}^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}^{\mathsf{T}}_{t_j}\right] + 2\mu_2\boldsymbol{L} = \sum_{j=1}^{r}\frac{\eta_{t_j}}{\sigma^2_{t_j}n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{u}_m^{(k,t_j)}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}^{\mathsf{T}}_{t_j} \ .$$

Noting that $\boldsymbol{s}_{t_j}^{\mathsf{T}}\boldsymbol{L}^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right) \in \mathbb{R}$, we can write:

$$\sum_{j=1}^{r}\left[\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\boldsymbol{\Phi}^{\mathsf{T}}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{L}\boldsymbol{s}_{t_j}\right] + 2\mu_2\boldsymbol{L} = \sum_{j=1}^{r}\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{u}_m^{(k,t_j)}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}} \quad . \tag{14}$$

To solve Eq. (14), we introduce the standard vec$(\cdot)$ operator leading to:

$$\text{vec}\left(\sum_{j=1}^{r}\left[\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\boldsymbol{\Phi}^{\mathsf{T}}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{L}\boldsymbol{s}_{t_j}\right] + 2\mu_2\boldsymbol{L}\right)$$
$$= \text{vec}\left(\sum_{j=1}^{r}\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{u}_m^{(k,t_j)}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\right)$$

$$\sum_{j=1}^{r}\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\text{vec}\left(\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\right)\text{vec}\left(\boldsymbol{\Phi}^{\mathsf{T}}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{L}\boldsymbol{s}_{t_j}\right) + 2\mu_2\text{vec}(\boldsymbol{L})$$
$$= \sum_{j=1}^{r}\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\text{vec}\left(\boldsymbol{u}_m^{(k,t_j)}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\right) \quad .$$

Knowing that for a given set of matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{X}$, $\text{vec}(\boldsymbol{AXB}) = \left(\boldsymbol{B}^{\mathsf{T}} \otimes \boldsymbol{A}\right)\text{vec}(\boldsymbol{X})$, we can write

$$\sum_{j=1}^{r}\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\text{vec}\left(\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\right)\left(\boldsymbol{s}_{t_j}^{\mathsf{T}} \otimes \boldsymbol{\Phi}^{\mathsf{T}}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)\text{vec}(\boldsymbol{L}) + 2\mu_2\text{vec}(\boldsymbol{L})$$
$$= \sum_{j=1}^{r}\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\text{vec}\left(\boldsymbol{u}_m^{(k,t_j)}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\right) \quad .$$

By choosing $\boldsymbol{Z}_{\boldsymbol{L}} = 2\mu_2\boldsymbol{I}_{dk\times dk} + \sum_{j=1}^{r}\frac{\eta_{t_j}}{n_{t_j}\sigma_{t_j}^2}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\text{vec}\left(\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\right)\left(\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right) \otimes \boldsymbol{s}_{t_j}^{\mathsf{T}}\right)$, and $\boldsymbol{v}_{\boldsymbol{L}} = \sum_{j=1}^{r}\frac{\eta_{t_j}}{n_{t_j}\sigma_{t_j}^2}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\text{vec}\left(\boldsymbol{u}_m^{(k,t_j)}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{s}_{t_j}^{\mathsf{T}}\right)$, we can update $\boldsymbol{L} = \boldsymbol{Z}_{\boldsymbol{L}}^{-1}\boldsymbol{v}_{\boldsymbol{L}}$.

A.2. Update Equations for $\boldsymbol{S}$

To derive the update equations with respect to $\boldsymbol{S}$, similar approach to that of $\boldsymbol{L}$ can be followed. The derivative of $\boldsymbol{e}_r(\boldsymbol{L}, \boldsymbol{S})$ with respect to $\boldsymbol{S}$ can be computed column-wise for all tasks observed so far:

$$\nabla_{\boldsymbol{s}_{t_j}}\boldsymbol{e}_r(\boldsymbol{L},\boldsymbol{S}) = \nabla_{\boldsymbol{s}_{t_j}}\left[\sum_{j=1}^{r}\frac{\eta_{t_j}}{2\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\left(\boldsymbol{u}_m^{(k,t_j)} - \left(\boldsymbol{L}\boldsymbol{s}_{t_j}\right)^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)^2 + \mu_1||\boldsymbol{S}||_{\mathsf{F}}^2 + \mu_2||\boldsymbol{L}||_{\mathsf{F}}^2\right]$$
$$= -\sum_{t_k=t_j}\left[\frac{\eta_{t_j}}{\sigma_{t_j}^2 n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\left(\boldsymbol{u}_m^{(k,t_j)} - \left(\boldsymbol{L}\boldsymbol{s}_{t_j}\right)^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right)\boldsymbol{L}^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\right] + 2\mu_2\boldsymbol{s}_{t_j} \quad .$$

Using a similar analysis to the previous section, choosing

$$\boldsymbol{Z}_{\boldsymbol{s}_{t_j}} = 2\mu_1\boldsymbol{I}_{k\times k} + \sum_{t_k=t_j}\frac{\eta_{t_j}}{n_{t_j}\sigma_{t_j}^2}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{L}^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{\Phi}^{\mathsf{T}}\left(\boldsymbol{x}_m^{(k,t_j)}\right)\boldsymbol{L} \quad ,$$
$$\boldsymbol{v}_{\boldsymbol{s}_{t_j}} = \sum_{t_k=t_j}\frac{\eta_{t_j}}{n_{t_j}\sigma_{t_j}^2}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\boldsymbol{u}_m^{(k,t_j)}\boldsymbol{L}^{\mathsf{T}}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,t_j)}\right) \quad ,$$

we can update $\boldsymbol{s}_{t_j} = \boldsymbol{Z}_{\boldsymbol{s}_{t_j}}^{-1}\boldsymbol{v}_{\boldsymbol{s}_{t_j}}$.

B. Proofs of Theoretical Guarantees

In this appendix, we prove the claims and lemmas from the main paper, leading to sublinear regret (Theorem 1).

Lemma 1. *Assume the policy for a task t_j at a round r to be given by* $\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k,\,t_j)}|\boldsymbol{x}_m^{(k,\,t_j)}\right)\Big|_{\hat{\boldsymbol{\theta}}_r} = \mathcal{N}\left(\boldsymbol{\alpha}^{\mathsf{T}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right),\boldsymbol{\sigma}_{t_j}\right)$*, for* $\boldsymbol{x}_m^{(k,\,t_j)}\in\mathcal{X}_{t_j}$ *and* $\boldsymbol{u}_m^{(k,\,t_j)}\in\mathcal{U}_{t_j}$ *with* $\mathcal{X}_{t_j}$ *and* $\mathcal{U}_{t_j}$ *representing the state and action spaces, respectively. The gradient* $\nabla_{\boldsymbol{\alpha}_{t_j}}l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}$*, for* $l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)=-1/n_{t_j}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\log\left[\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k,\,t_j)}|\boldsymbol{x}_m^{(k,\,t_j)}\right)\right]$ *satisfies*

$$\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}}l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2\le\frac{M_{t_j}}{\sigma^2_{t_j}}\left[\left(u_{max}+\max_{t_k\in\mathcal{I}_{r-1}}\left\{||\boldsymbol{A}^{+}_{t_k}||_2\left(||\boldsymbol{b}_{t_k}||_2+\boldsymbol{c}_{max}\right)\right\}\boldsymbol{\Phi}_{max}\right)\boldsymbol{\Phi}_{max}\right],$$

with $u_{max}=\max_{k,m}\left\{\left|\boldsymbol{u}_m^{(k,\,t_j)}\right|\right\}$ *and* $\boldsymbol{\Phi}_{max}=\max_{k,m}\left\{\left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\}$ *for all trajectories and all tasks.*

Proof. The proof of the above lemma will be provided as a collection of claims. We start with the following:

<u>Claim:</u> Given $\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k)}|\boldsymbol{x}_m^{(k)}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}=\mathcal{N}\left(\boldsymbol{\alpha}^{\mathsf{T}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right),\boldsymbol{\sigma}_{t_j}\right)$, for $\boldsymbol{x}_m^{(k,\,t_j)}\in\mathcal{X}_{t_j}$ and $\boldsymbol{u}_m^{(k,\,t_j)}\in\mathcal{U}_{t_j}$, and $l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)=-1/n_{t_j}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\log\left[\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k,\,t_j)}|\boldsymbol{x}_m^{(k,\,t_j)}\right)\right]$, $\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}}l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2$ satisfies

$$\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}}l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2\le\frac{M_{t_j}}{\sigma^2_{t_j}}\left[\left(u_{\text{max}}+\left|\left|\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2\boldsymbol{\Phi}_{\text{max}}\right)\boldsymbol{\Phi}_{\text{max}}\right].\tag{15}$$

<u>Proof:</u> Since $\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k,\,t_j)}|\boldsymbol{x}_m^{(k,\,t_j)}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}=\mathcal{N}\left(\boldsymbol{\alpha}^{\mathsf{T}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right),\boldsymbol{\sigma}_{t_j}\right)$, we can write

$$\log\left[\pi^{(t_j)}_{\boldsymbol{\alpha}_{t_j}}\left(\boldsymbol{u}_m^{(k,\,t_j)}|\boldsymbol{x}_m^{(k,\,t_j)}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right]=-\log\left[\sqrt{2\pi\sigma^2_{t_j}}\right]-\frac{1}{2\sigma^2_{t_j}}\left(\boldsymbol{u}_m^{(k,\,t_j)}-\boldsymbol{\alpha}^{\mathsf{T}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right)^2.$$

Therefore:

$$\begin{aligned}
\nabla_{\boldsymbol{\alpha}_{t_j}}l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}&=-\frac{1}{n_{t_j}}\sum_{k=1}^{n_{t_j}}\sum_{m=0}^{M_{t_j}-1}\frac{1}{\sigma^2_{t_j}}\left(\boldsymbol{u}_m^{(k,\,t_j)}-\boldsymbol{\alpha}^{\mathsf{T}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right)\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\\
\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}}l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2&\le\frac{M_{t_j}}{\sigma^2_{t_j}}\left[\max_{k,m}\left\{\left|\boldsymbol{u}_m^{(k,\,t_j)}-\boldsymbol{\alpha}^{\mathsf{T}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\times\left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\}\right]\\
&\le\frac{M_{t_j}}{\sigma^2_{t_j}}\left[\max_{k,m}\left\{\left|\boldsymbol{u}_m^{(k,\,t_j)}\right|\times\left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\}\right.\\
&\qquad\left.+\max_{k,m}\left\{\left|\boldsymbol{\alpha}^{\mathsf{T}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\times\left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\}\right]\\
&\le\frac{M_{t_j}}{\sigma^2_{t_j}}\left[\max_{k,m}\left\{\left|\boldsymbol{u}_m^{(k,\,t_j)}\right|\right\}\max_{k,m}\left\{\left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\}\right.\\
&\qquad\left.+\max_{k,m}\left\{\left|\left\langle\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r},\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right\rangle\right|\right\}\max_{k,m}\left\{\left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\}\right].
\end{aligned}$$

Denoting $\max_{k,m}\left\{\left|\boldsymbol{u}_m^{(k,\,t_j)}\right|\right\}=u_{\text{max}}$ and $\max_{k,m}\left\{\left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\}=\boldsymbol{\Phi}_{\text{max}}$ for all trajectories and all tasks, we can write

$$\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}}l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2\le\frac{M_{t_j}}{\sigma^2_{t_j}}\left[\left(u_{\text{max}}+\max_{k,m}\left\{\left|\left\langle\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r},\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right\rangle\right|\right\}\right)\boldsymbol{\Phi}_{\text{max}}\right].$$

Using the Cauchy-Shwarz inequality (Horn & Mathias, 1990), we can upper bound $\max_{k,m}\left\{\left|\left\langle\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r},\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right\rangle\right|\right\}$

as

$$\max_{k,m}\left\{\left|\left\langle \boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}, \boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right\rangle\right|\right\} \leq \max_{k,m}\left\{\left|\left|\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \left|\left|\boldsymbol{\Phi}\left(\boldsymbol{x}_m^{(k,\,t_j)}\right)\right|\right|_2\right\} \leq \max_{k,m}\left\{\left|\left|\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2\right\}\boldsymbol{\Phi}_{\max}$$
$$\leq \left|\left|\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \boldsymbol{\Phi}_{\max} \; .$$

Finalizing the statement of the claim, the overall bound on the norm of the gradient of $l_{t_j}(\boldsymbol{\alpha}_{t_j})$ can be written as

$$\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \leq \frac{M_{t_j}}{\sigma_{t_j}^2}\left[\left(u_{\max} + \left|\left|\boldsymbol{\alpha}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \boldsymbol{\Phi}_{\max}\right)\boldsymbol{\Phi}_{\max}\right] \; . \tag{16}$$

■

Claim: The norm of the gradient of the loss function satisfies:

$$\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \leq \frac{M_{t_j}}{\sigma_{t_j}^2}\left[\left(u_{\max} + \max_{t_k \in \mathcal{I}_{r-1}}\left\{\|\boldsymbol{A}_{t_k}^{+}\|_2\left(\|\boldsymbol{b}_{t_k}\|_2 + \boldsymbol{c}_{\max}\right)\right\}\boldsymbol{\Phi}_{\max}\right)\boldsymbol{\Phi}_{\max}\right] \; .$$

Proof: As mentioned previously, we consider the linearization of the loss function l_{t_j} around the constraint solution of the previous round, $\hat{\boldsymbol{\theta}}_r$. Since $\hat{\boldsymbol{\theta}}_r$ satisfies $\boldsymbol{A}_{t_k}\boldsymbol{\alpha}_{t_k} = \boldsymbol{b}_{t_k} - \boldsymbol{c}_{t_k}, \forall t_k \in \mathcal{I}_{r-1}$. Hence, we can write

$$\begin{aligned}
&\boldsymbol{A}_{t_k}\boldsymbol{\alpha}_{\boldsymbol{t_k}} + \boldsymbol{c}_{t_k} = \boldsymbol{b}_{t_k} \quad \forall t_k \in \mathcal{I}_{r-1} \\
&\quad \Longrightarrow \; \boldsymbol{\alpha}_{t_k} = \boldsymbol{A}_{t_k}^{+}\left(\boldsymbol{b}_{t_k} - \boldsymbol{c}_{t_k}\right) \quad \text{with } \boldsymbol{A}_{t_k}^{+} = \left(\boldsymbol{A}_{t_k}^{\mathsf{T}}\boldsymbol{A}_{t_k}\right)^{-1}\boldsymbol{A}_{t_k}^{\mathsf{T}} \text{ being the left pseudo-inverse.}
\end{aligned}$$

Therefore

$$\begin{aligned}
||\boldsymbol{\alpha}_{t_k}||_2 &\leq \left|\left|\boldsymbol{A}_{t_k}^{+}\right|\right|_2\left(||\boldsymbol{b}_{t_k}||_2 + ||\boldsymbol{c}_{t_k}||_2\right) \\
&\leq \left|\left|\boldsymbol{A}_{t_k}^{+}\right|\right|_2\left(||\boldsymbol{b}_{t_k}||_2 + \boldsymbol{c}_{\max}\right) \; .
\end{aligned}$$

Combining the above results with those of Eq. (16) we arrive at

$$\left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}\left(\boldsymbol{\alpha}_{t_j}\right)\right|\right|_2 \leq \frac{M_{t_j}}{\sigma_{t_j}^2}\left[\left(u_{\max} + \max_{t_k \in \mathcal{I}_{r-1}}\left\{\left|\left|\boldsymbol{A}_{t_k}^{+}\right|\right|_2\left(||\boldsymbol{b}_{t_k}||_2 + \boldsymbol{c}_{\max}\right)\right\}\boldsymbol{\Phi}_{\max}\right)\boldsymbol{\Phi}_{\max}\right] \; .$$

■

The previous result finalizes the statement of the lemma, bounding the gradient of the loss function in terms of the *safety* constraints.

□

Lemma 2. *The norm of the gradient of the loss function evaluated at $\hat{\boldsymbol{\theta}}_r$ satisfies*

$$\left|\left|\nabla_{\boldsymbol{\theta}} l_{t_j}\left(\boldsymbol{\theta}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2^2 \leq \left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}\left(\boldsymbol{\theta}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2^2\left(q \times d\left({}^{2d}/{p^2}\max_{t_k \in \mathcal{I}_{r-1}}\left\{\left|\left|\boldsymbol{A}_{t_j}^{\dagger}\right|\right|_2^2\left(\left|\left|\boldsymbol{b}_{t_j}\right|\right|_2^2 + \boldsymbol{c}_{max}^2\right)\right\} + 1\right)\right) \; .$$

Proof. The derivative of $l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}$ can be written as

$$\nabla_{\boldsymbol{\theta}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r} = \begin{bmatrix} \nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}^{\mathsf{T}}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r} \begin{bmatrix} \frac{\partial \boldsymbol{\alpha}_{t_j}^{(1)}}{\partial \boldsymbol{\theta}_1}\Big|_{\hat{\boldsymbol{\theta}}_r} \\ \vdots \\ \frac{\partial \boldsymbol{\alpha}_{t_j}^{(d)}}{\partial \boldsymbol{\theta}_1}\Big|_{\hat{\boldsymbol{\theta}}_r} \end{bmatrix} \\ \vdots \\ \nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}^{\mathsf{T}}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r} \begin{bmatrix} \frac{\partial \boldsymbol{\alpha}_{t_j}^{(1)}}{\partial \boldsymbol{\theta}_{dk+k|\mathcal{T}|}}\Big|_{\hat{\boldsymbol{\theta}}_r} \\ \vdots \\ \frac{\partial \boldsymbol{\alpha}_{t_j}^{(d)}}{\partial \boldsymbol{\theta}_{dk+k|\mathcal{T}|}}\Big|_{\hat{\boldsymbol{\theta}}_r} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}^{\mathsf{T}}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r} \begin{bmatrix} \boldsymbol{\theta}_{dk+1}\Big|_{\hat{\boldsymbol{\theta}}_r} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ \vdots \\ \nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}^{\mathsf{T}}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r} \begin{bmatrix} 0 \\ \vdots \\ \boldsymbol{\theta}_{(d+1)k+1}\Big|_{\hat{\boldsymbol{\theta}}_r} \end{bmatrix} \\ \vdots \\ \nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}^{\mathsf{T}}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r} \begin{bmatrix} \boldsymbol{\theta}_{d(k+1)+1}\Big|_{\hat{\boldsymbol{\theta}}_r} \\ \vdots \\ \boldsymbol{\theta}_{dk}\Big|_{\hat{\boldsymbol{\theta}}_r} \end{bmatrix} \end{bmatrix}$$

$$\implies \left\|\nabla_{\boldsymbol{\theta}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2^2 \leq \left\|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}(\boldsymbol{\alpha}_{t_j})\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2^2 \left[d\left\|\boldsymbol{s}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2^2 + \left\|\boldsymbol{L}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_{\mathsf{F}}^2\right] .$$

The results of Lemma 1 bound $\left\|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2^2$.

Now, we target to bound each of $\left|\left|\boldsymbol{s}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2^2$ and $\left|\left|\boldsymbol{L}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_{\mathsf{F}}^2$.

Bounding $\left\|\boldsymbol{s}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2^2$ and $\|\boldsymbol{L}\Big|_{\hat{\boldsymbol{\theta}}_r}\|_{\mathsf{F}}^2$: Considering the constraint $\boldsymbol{A}_{t_j}\boldsymbol{L}\boldsymbol{s}_{t_j} + \boldsymbol{c}_{t_j} = \boldsymbol{b}_{t_j}$ for a task t_j, we realize that $\boldsymbol{s}_{t_j} = \boldsymbol{L}^{+}\left(\boldsymbol{A}_{t_j}^{+}\left(\boldsymbol{b}_{t_j} - \boldsymbol{c}_{t_j}\right)\right)$. Therefore,

$$\left\|\boldsymbol{s}_{t_j}\Big|_{\hat{\theta}_r}\right\|_2 \leq \left|\left|\boldsymbol{L}^{+}\left(\boldsymbol{A}_{t_j}^{+}\left(\boldsymbol{b}_{t_j} - \boldsymbol{c}_{t_j}\right)\right)\right|\right|_2 \leq \left|\left|\boldsymbol{L}^{+}\right|\right|_2 \left|\left|\boldsymbol{A}_{t_j}^{+}\right|\right|_2 \left(\left|\left|\boldsymbol{b}_{t_j}\right|\right|_2 + \left|\left|\boldsymbol{c}_{t_j}\right|\right|_2\right). \tag{17}$$

Noting that

$$\begin{aligned} \left|\left|\boldsymbol{L}^{+}\right|\right|_2 = \left|\left|\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\boldsymbol{L}^{\mathsf{T}}\right|\right|_2 &\leq \left|\left|\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right|\right|_2 \left|\left|\boldsymbol{L}^{\mathsf{T}}\right|\right|_2 \leq \left|\left|\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right|\right|_2 \left|\left|\boldsymbol{L}^{\mathsf{T}}\right|\right|_{\mathsf{F}} \\ &= \left|\left|\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right|\right|_2 \left|\left|\boldsymbol{L}\right|\right|_{\mathsf{F}} \ . \end{aligned}$$

To relate $||\boldsymbol{L}^{+}||_2$ to $||\boldsymbol{L}||_{\mathsf{F}}$, we need to bound $\left|\left|\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right|\right|_2$ in terms of $\|\boldsymbol{L}\|_{\mathsf{F}}$. Denoting the spectrum of $\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}$ as $\text{spec}\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right) = \{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_k\}$ such that $0 < \boldsymbol{\lambda}_1 \leq \cdots \leq \boldsymbol{\lambda}_k$, then $\text{spect}\left(\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right) = \{1/\boldsymbol{\lambda}_1, \ldots, 1/\boldsymbol{\lambda}_k\}$ such that $1/\boldsymbol{\lambda}_k \leq \cdots \leq 1/\boldsymbol{\lambda}_k$. Hence, $\left|\left|\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right|\right|_2 = \max\left\{\text{spec}\left(\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right)\right\} = 1/\boldsymbol{\lambda}_1 = 1/\lambda_{\min}\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)$. Noticing that $\text{spec}\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right) = \text{spec}\left(\boldsymbol{L}\boldsymbol{L}^{\mathsf{T}}\right)$, we recognize $\left|\left|\left(\boldsymbol{L}^{\mathsf{T}}\boldsymbol{L}\right)^{-1}\right|\right|_2 = 1/\boldsymbol{\lambda}_{\min}\left(\boldsymbol{L}\boldsymbol{L}^{\mathsf{T}}\right) \leq 1/p$. Therefore

$$\left|\left|\boldsymbol{L}^{+}\right|\right|_2 \leq \frac{1}{p}\left|\left|\boldsymbol{L}\right|\right|_{\mathsf{F}} \ . \tag{18}$$

Plugging the results of Eq. (18) into Eq. (17), we arrive at

$$\left\|\boldsymbol{s}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2 \leq 1/p \left\|\boldsymbol{L}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_{\mathsf{F}} \max_{t_k \in \mathcal{I}_{r-1}} \left\{\left|\left|\boldsymbol{A}_{t_k}^{+}\right|\right|_2 \left(\left|\left|\boldsymbol{b}_{t_k}\right|\right|_2 + \boldsymbol{c}_{\max}\right)\right\} \ . \tag{19}$$

Finally, since $\hat{\boldsymbol{\theta}}_r$ satisfies the constraints, we note that $\left\|\boldsymbol{L}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_{\mathsf{F}}^2 \leq q \times d$. Consequently,

$$\left|\left|\nabla_{\boldsymbol{\theta}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2^2 \leq \left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2^2 \left(q \times d\left(\frac{2d}{p^2} \max_{t_k \in \mathcal{I}_{r-1}} \left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2 \left(||\boldsymbol{b}_{t_k}||_2^2 + \boldsymbol{c}_{\text{max}}^2\right)\right\} + 1\right)\right) .$$

□

Lemma 3. *The L_2 norm of the constraint solution at round $r-1$, $\|\hat{\boldsymbol{\theta}}_r\|_2^2$ is bounded by*

$$\|\hat{\boldsymbol{\theta}}_r\|_2^2 \leq q \times d\left[1 + |\mathcal{I}_{r-1}| \frac{1}{p^2} \max_{t_k \in \mathcal{I}_{r-1}} \left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2 \left(||\boldsymbol{b}_{t_k}||_2 + \boldsymbol{c}_{max}\right)^2\right\}\right] .$$

with $|\mathcal{I}_{r-1}|$ being the cardinality of $\mathcal{I}_{r-1}$ representing the number of different tasks observed so-far.

Proof. Noting that $\hat{\boldsymbol{\theta}}_r = \Big[\underbrace{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{dk}}_{\boldsymbol{L}|_{\hat{\boldsymbol{\theta}}_r}}, \underbrace{\boldsymbol{\theta}_{dk+1}, \dots,}_{\boldsymbol{s}_{i_1}|_{\hat{\theta}_r}} \underbrace{\dots, \dots,}_{\boldsymbol{s}_{i_{r-1}}|_{\hat{\theta_r}}} \underbrace{\dots, \boldsymbol{\theta}_{dk+kT^\star}}_{\boldsymbol{0}\text{'s: unobserved tasks}}\Big]^{\mathsf{T}}$, it is easy to see

$$\begin{aligned}
\|\hat{\boldsymbol{\theta}}_r\|_2^2 &\leq \left\|\boldsymbol{L}\Big|_{\hat{\theta}_r}\right\|_{\mathsf{F}}^2 + |\mathcal{I}_{r-1}| \max_{t_k \in \mathcal{I}_{r-1}} \left\{\left\|\boldsymbol{s}_{t_k}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2^2\right\} \\
&\leq q \times d + |\mathcal{I}_{r-1}| \max_{t_k \in \mathcal{I}_{r-1}} \left[\frac{q \times d}{p^2}\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2 \left(||\boldsymbol{b}_{t_k}||_2 + \boldsymbol{c}_{\text{max}}\right)^2\right] \\
&\leq q \times d\left[1 + |\mathcal{I}_{r-1}| \frac{1}{p^2} \max_{t_k \in \mathcal{I}_{r-1}} \left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2 \left(||\boldsymbol{b}_{t_k}||_2 + \boldsymbol{c}_{\text{max}}\right)^2\right\}\right].
\end{aligned}$$

□

Lemma 4. *The L_2 norm of the linearizing term of $l_{t_j}(\boldsymbol{\theta})$ around $\hat{\boldsymbol{\theta}}_r$, $\left\|\hat{\boldsymbol{f}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2$, is bounded by*

$$\left\|\hat{\boldsymbol{f}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2 \leq \left\|\nabla_{\boldsymbol{\theta}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right\|_2 \left(1 + \|\hat{\boldsymbol{\theta}}_r\|_2\right) + \left|l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right| \leq \boldsymbol{\gamma}_1(r)\left(1 + \boldsymbol{\gamma}_2(r)\right) + \boldsymbol{\delta}_{l_{t_j}},$$

with $\boldsymbol{\delta}_{l_{t_j}}$ being the constant upper-bound on $\left|l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right|$, and

$$\begin{aligned}
\boldsymbol{\gamma}_1(r) &= \frac{1}{n_{t_j}\sigma_{t_j}^2}\left[\left(u_{max} + \max_{t_k \in \mathcal{I}_{r-1}} \left\{||\boldsymbol{A}_{t_k}^{+}||_2 \left(||\boldsymbol{b}_{t_k}||_2 + \boldsymbol{c}_{max}\right)\right\} \boldsymbol{\Phi}_{max}\right) \boldsymbol{\Phi}_{max}\right] \\
&\qquad \times \left({}^{d}/_{p}\sqrt{2q}\sqrt{\max_{t_k \in \mathcal{I}_{r-1}} \left\{\|\boldsymbol{A}_{t_k}^{\dagger}\|_2^2 \left(\|\boldsymbol{b}_{t_k}\|_2^2 + \boldsymbol{c}_{max}^2\right)\right\}} + \sqrt{qd}\right) . \\
\boldsymbol{\gamma}_2(r) &\leq \sqrt{q \times d} + \sqrt{|\mathcal{I}_{r-1}|}\sqrt{\left[1 + \frac{1}{p^2} \max_{t_k \in \mathcal{I}_{r-1}} \left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2 \left(||\boldsymbol{b}_{t_k}||_2 + \boldsymbol{c}_{max}\right)^2\right\}\right]} .
\end{aligned}$$

Proof. We have previously shown that $\left|\left|\hat{\boldsymbol{f}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \leq \left|\left|\nabla_{\boldsymbol{\theta}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 + \left|l_{t_j}\left(\hat{\theta}_r\right)\right| + \left|\left|\nabla_{\boldsymbol{\theta}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \times \left|\left|\hat{\boldsymbol{\theta}}_r\right|\right|_2$. Using the previously derived lemmas we can upper-bound $\left|\left|\hat{\boldsymbol{f}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2$ as follows

$$\left|\left|\nabla_{\boldsymbol{\theta}} l_{t_j}\left(\boldsymbol{\theta}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2^2 \leq \left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}\left(\boldsymbol{\theta}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2^2 \left(q\times d\left(\frac{2d}{p^2}\max_{t_k\in\mathcal{I}_{r-1}}\left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2\left(||\boldsymbol{b}_{t_k}||_2^2+\boldsymbol{c}_{\max}^2\right)\right\}+1\right)\right)$$

$$\begin{aligned}\left|\left|\nabla_{\boldsymbol{\theta}} l_{t_j}\left(\boldsymbol{\theta}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 &\leq \left|\left|\nabla_{\boldsymbol{\alpha}_{t_j}} l_{t_j}\left(\boldsymbol{\theta}\right)\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2 \left({}^{d}\!/_{p}\sqrt{2q}\sqrt{\max_{t_k\in\mathcal{I}_{r-1}}\left\{\|\boldsymbol{A}_{t_k}^{\dagger}\|_2^2\left(\|\boldsymbol{b}_{t_k}\|_2^2+\boldsymbol{c}_{\max}^2\right)\right\}}+\sqrt{qd}\right)\\ &\leq \frac{1}{n_{t_j}\sigma_{t_j}^2}\left[\left(u_{\max}+\max_{t_k\in\mathcal{I}_{r-1}}\left\{||\boldsymbol{A}_{t_k}^{+}||_2\left(||\boldsymbol{b}_{t_k}||_2+\boldsymbol{c}_{\max}\right)\right\}\boldsymbol{\Phi}_{\max}\right)\boldsymbol{\Phi}_{\max}\right]\\ &\quad\times\left({}^{d}\!/_{p}\sqrt{2q}\sqrt{\max_{t_k\in\mathcal{I}_{r-1}}\left\{\|\boldsymbol{A}_{t_k}^{\dagger}\|_2^2\left(\|\boldsymbol{b}_{t_k}\|_2^2+\boldsymbol{c}_{\max}^2\right)\right\}}+\sqrt{qd}\right) \quad .\end{aligned}$$

Further,

$$\left|\left|\hat{\boldsymbol{\theta}}_r\right|\right|_2^2 \leq q\times d+|\mathcal{I}_{r-1}|\max_{t_k\in\mathcal{I}_{r-1}}\left[1+\frac{1}{p^2}\max_{t_k\in\mathcal{I}_{r-1}}\left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2\left(||\boldsymbol{b}_{t_k}||_2+\boldsymbol{c}_{\max}\right)^2\right\}\right]$$

$$\implies \left|\left|\hat{\boldsymbol{\theta}}_r\right|\right|_2 \leq \sqrt{q\times d}+\sqrt{|\mathcal{I}_{r-1}|}\sqrt{\left[1+\frac{1}{p^2}\max_{t_k\in\mathcal{I}_{r-1}}\left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2\left(||\boldsymbol{b}_{t_k}||_2+\boldsymbol{c}_{\max}\right)^2\right\}\right]} \quad .$$

Therefore

$$\begin{aligned}\left|\left|\hat{\boldsymbol{f}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right| &\leq \left|\left|\nabla_{\boldsymbol{\theta}} l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right|\right|_2\left(1+\left|\left|\hat{\boldsymbol{\theta}}_r\right|\right|_2\right)+\left|l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right| \\ &\leq \boldsymbol{\gamma}_1(r)\left(1+\boldsymbol{\gamma}_2(r)\right)+\boldsymbol{\delta}_{l_{t_j}},\end{aligned} \tag{20}$$

with $\boldsymbol{\delta}_{l_{t_j}}$ being the constant upper-bound on $\left|l_{t_j}(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}_r}\right|$, and

$$\begin{aligned}\boldsymbol{\gamma}_1(r) &= \frac{1}{n_{t_j}\sigma_{t_j}^2}\left[\left(u_{\max}+\max_{t_k\in\mathcal{I}_{r-1}}\left\{||\boldsymbol{A}_{t_k}^{+}||_2\left(||\boldsymbol{b}_{t_k}||_2+\boldsymbol{c}_{\max}\right)\right\}\boldsymbol{\Phi}_{\max}\right)\boldsymbol{\Phi}_{\max}\right]\\ &\quad\times\left({}^{d}\!/_{p}\sqrt{2q}\sqrt{\max_{t_k\in\mathcal{I}_{r-1}}\left\{\|\boldsymbol{A}_{t_k}^{\dagger}\|_2^2\left(\|\boldsymbol{b}_{t_k}\|_2^2+\boldsymbol{c}_{\max}^2\right)\right\}}+\sqrt{qd}\right).\\ \boldsymbol{\gamma}_2(r) &\leq \sqrt{q\times d}+\sqrt{|\mathcal{I}_{r-1}|}\sqrt{\left[1+\frac{1}{p^2}\max_{t_k\in\mathcal{I}_{r-1}}\left\{\left|\left|\boldsymbol{A}_{t_k}^{\dagger}\right|\right|_2^2\left(||\boldsymbol{b}_{t_k}||_2+\boldsymbol{c}_{\max}\right)^2\right\}\right]} \quad .\end{aligned}$$

□

Theorem 1 (Sublinear Regret; restated from the main paper)**.** *After R rounds and choosing $\eta_{t_1}=\cdots=\eta_{t_j}=\eta=\frac{1}{\sqrt{R}}$, $\boldsymbol{L}\Big|_{\hat{\boldsymbol{\theta}}_1}=diag_{\boldsymbol{k}}(\zeta)$, with $diag_{\boldsymbol{k}}(\cdot)$ being a diagonal matrix among the $\boldsymbol{k}$ columns of $\boldsymbol{L}$, $p\leq\zeta^2\leq q$, and $\boldsymbol{S}\Big|_{\hat{\boldsymbol{\theta}}_1}=\boldsymbol{0}_{k\times|\mathcal{T}|}$, for any $\boldsymbol{u}\in\mathcal{K}$ our algorithm exhibits a sublinear regret of the form*

$$\sum_{j=1}^{R} l_{t_j}\left(\hat{\boldsymbol{\theta}}_r\right)-l_{t_j}(\boldsymbol{u})=\mathcal{O}\left(\sqrt{R}\right) \quad .$$

Proof. Given the ingredients of the previous section, next we derive the sublinear regret results which finalize the statement of the theorem. First, it is easy to see that

$$\nabla_{\boldsymbol{\theta}}\boldsymbol{\Omega}_0\left(\tilde{\boldsymbol{\theta}}_j\right)-\nabla_{\boldsymbol{\theta}}\boldsymbol{\Omega}_0\left(\tilde{\boldsymbol{\theta}}_{j+1}\right)=\eta_{t_j}\hat{\boldsymbol{f}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_j} \quad .$$

Further, from strong convexity of the regularizer we obtain:

$$\boldsymbol{\Omega}_0\left(\hat{\boldsymbol{\theta}}_j\right)\geq\boldsymbol{\Omega}_0\left(\hat{\boldsymbol{\theta}}_{j+1}\right)+\left\langle\nabla_{\boldsymbol{\theta}}\boldsymbol{\Omega}_0\left(\hat{\boldsymbol{\theta}}_{j+1}\right),\hat{\boldsymbol{\theta}}_j-\hat{\boldsymbol{\theta}}_{j+1}\right\rangle+\frac{1}{2}\left|\left|\hat{\boldsymbol{\theta}}_j-\hat{\boldsymbol{\theta}}_{j+1}\right|\right|_2^2 \quad .$$

It can be seen that

$$\left\|\hat{\boldsymbol{\theta}}_j-\hat{\boldsymbol{\theta}}_{j+1}\right\|_2\leq\eta_{t_j}\left\|\hat{\boldsymbol{f}}_{t_j}\Big|_{\hat{\boldsymbol{\theta}}_j}\right\|_2 \quad .$$

Finally, for any $\boldsymbol{u} \in \mathcal{K}$, we have:

$$\sum_{j=1}^{r} \eta_{t_j} \left(l_{t_j} \left(\hat{\boldsymbol{\theta}}_j \right) - l_{t_j}(\boldsymbol{u}) \right) \leq \sum_{j=1}^{r} \left[\eta_{t_j} \left(\left\| \hat{\boldsymbol{f}}_{t_j} \Big|_{\hat{\boldsymbol{\theta}}_j} \right\|_2 \right)^2 \right] + \boldsymbol{\Omega}_0(\boldsymbol{u}) - \boldsymbol{\Omega}_0(\hat{\boldsymbol{\theta}}_1) \ .$$

Assuming $\eta_{t_1} = \cdots = \eta_{t_j} = \eta$, we can derive

$$\sum_{j=1}^{r} \left(l_{t_j} \left(\hat{\boldsymbol{\theta}}_j \right) - l_{t_j}(\boldsymbol{u}) \right) \leq \eta \sum_{j=1}^{r} \left(\left\| \hat{\boldsymbol{f}}_{t_j} \Big|_{\hat{\boldsymbol{\theta}}_j} \right\|_2 \right)^2 + {}^{1}\!/\!_{\eta} \left(\boldsymbol{\Omega}_0(\boldsymbol{u}) - \boldsymbol{\Omega}_0(\hat{\boldsymbol{\theta}}_1) \right) \ .$$

The following lemma finalizes the statement of the theorem:

Lemma 5. *After T rounds and for $\eta_{t_1} = \cdots = \eta_{t_j} = \eta = \frac{1}{\sqrt{R}}$, our algorithm exhibits, for any $\boldsymbol{u} \in \mathcal{K}$, a sublinear regret of the form*

$$\sum_{j=1}^{R} l_{t_j}(\hat{\boldsymbol{\theta}}_j) - l_{t_j}(\boldsymbol{u}) \leq \mathcal{O}\left(\sqrt{R}\right) \ .$$

Proof. It is then easy to see

$$\begin{aligned} \left\| \hat{\boldsymbol{f}}_{t_j} \Big|_{\hat{\boldsymbol{\theta}}_r} \right\|_2^2 &\leq \boldsymbol{\gamma}_3(R) + 4\boldsymbol{\gamma}_1^2(R)\boldsymbol{\gamma}_2^2(R) \quad \text{with} \quad \boldsymbol{\gamma}_3(R) = 4\boldsymbol{\gamma}_1^2(R) + 2 \max_{t_j \in \mathcal{I}_{R-1}} \boldsymbol{\delta}_{t_j}^2 \\ &\leq \boldsymbol{\gamma}_3(R) + 8\frac{d}{p^2}\boldsymbol{\gamma}_1^2(R)qd + 8\frac{d}{p^2}\boldsymbol{\gamma}_1^2(R)qd \left|\mathcal{I}_{R-1}\right| \max_{t_k \in \mathcal{I}_{R-1}} \left\{ \|\boldsymbol{A}_{t_k}^{\dagger}\|_2 \left(\|\boldsymbol{b}_{t_k}\|_2 + \boldsymbol{c}_{\max}\right)^2 \right\} \ . \end{aligned}$$

Since $|\mathcal{I}_{R-1}| \leq |\mathcal{T}|$ with $|\mathcal{T}|$ being the total number of tasks available, then we can write

$$\left\| \hat{\boldsymbol{f}}_{t_j} \Big|_{\hat{\theta}_r} \right\|_2^2 \leq \boldsymbol{\gamma}_5(R)|\mathcal{T}| \ ,$$

with $\boldsymbol{\gamma}_5 = 8{}^{d}\!/\!_{p^2} q \boldsymbol{\gamma}_1^2(R) \max_{t_k \in \mathcal{I}_{R-1}} \left\{ \|\boldsymbol{A}_{t_k}^{\dagger}\|_2^2 \left(\|\boldsymbol{b}_{t_k}\|_2 + \boldsymbol{c}_{\max}\right)^2 \right\}$. Further, it is easy to see that $\boldsymbol{\Omega}_0(\boldsymbol{u}) \leq qd + \boldsymbol{\gamma}_5(R)|\mathcal{T}|$ with $\boldsymbol{\gamma}_5(R)$ being a constant, which leads to

$$\sum_{j=1}^{r} \left(l_{t_j} \left(\hat{\boldsymbol{\theta}}_j \right) - l_{t_j}(\boldsymbol{u}) \right) \leq \eta \sum_{j=1}^{r} \boldsymbol{\gamma}_5(R)|\mathcal{T}| + {}^{1}\!/\!_{\eta} \left(qd + \boldsymbol{\gamma}_5(R)|\mathcal{T}| - \boldsymbol{\Omega}_0(\hat{\boldsymbol{\theta}}_1) \right) \ .$$

Initializing $\boldsymbol{L}$ and $\boldsymbol{S}$: We initialize $\boldsymbol{L}\Big|_{\hat{\boldsymbol{\theta}}_1} = \text{diag}_{\boldsymbol{k}}(\zeta)$, with $p \leq \zeta^2 \leq q$ and $\boldsymbol{S}\Big|_{\hat{\boldsymbol{\theta}}_1} = \boldsymbol{0}_{k \times |\mathcal{T}|}$ ensures the invertability of $\boldsymbol{L}$ and that the constraints are met. This leads us to

$$\sum_{j=1}^{r} \left(l_{t_j} \left(\hat{\boldsymbol{\theta}}_j \right) - l_{t_j}(\boldsymbol{u}) \right) \leq \eta \sum_{j=1}^{r} \boldsymbol{\gamma}_5(R)|\mathcal{T}| + {}^{1}\!/\!_{\eta} \left(qd + \boldsymbol{\gamma}_5(R)|\mathcal{T}| - \mu_2 k \zeta \right) \ .$$

Choosing $\eta_{t_1} = \cdots = \eta_{t_j} = \eta = {}^{1}\!/\!\sqrt{R}$, we acquire sublinear regret, finalizing the statement of the theorem:

$$\begin{aligned} \sum_{j=1}^{r} \left(l_{t_j} \left(\hat{\boldsymbol{\theta}}_j \right) - l_{t_j}(\boldsymbol{u}) \right) &\leq {}^{1}\!/\!\sqrt{R}\, \boldsymbol{\gamma}_5(R)|\mathcal{T}|R + \sqrt{R} \left(qd + \boldsymbol{\gamma}_5(R)|\mathcal{T}| - \mu_2 k \zeta \right) \\ &\leq \sqrt{R} \left(\boldsymbol{\gamma}_5(R)|\mathcal{T}| + qd\boldsymbol{\gamma}_5(R)|\mathcal{T}| - \mu_2 k \zeta \right) \leq \mathcal{O}\left(\sqrt{R}\right) \ , \end{aligned}$$

with $\boldsymbol{\gamma}_5(R)$ being a constant. □

□