

Table 1: Notation checklist.

Notation	Definition
t	Round index
i	Constraint index ($1 \leq i \leq m$)
$\mathcal{T}_t$	Task at round t
$\mathcal{D}_{0,t}$	Data distribution of loss function for task $\mathcal{T}_t$
$\mathcal{D}_{i,t}$ ($1 \leq i \leq m$)	Data distribution of constraint functions for task $\mathcal{T}_t$
$\mathcal{D}_t = \{\mathcal{D}_{0,t}, \mathcal{D}_{1,t}, \dots, \mathcal{D}_{m,t}\}$	Data distribution for task $\mathcal{T}_t$
$\mathcal{D}_{0,t}^{tr}$	Training dataset of loss function for task $\mathcal{T}_t$
$\mathcal{D}_{i,t}^{tr}$ ($1 \leq i \leq m$)	Training dataset of constraint functions for task $\mathcal{T}_t$
$\mathcal{D}_t = \{\mathcal{D}_{0,t}^{tr}, \mathcal{D}_{1,t}^{tr}, \dots, \mathcal{D}_{m,t}^{tr}\}$	Training dataset for task $\mathcal{T}_t$
$\mathcal{D}_{0,t}^{val}$	Validation dataset of loss function for task $\mathcal{T}_t$
$\mathcal{D}_{i,t}^{val}$ ($1 \leq i \leq m$)	Validation dataset of constraint functions for task $\mathcal{T}_t$
$\mathcal{D}_t^{val} = \{\mathcal{D}_{0,t}^{val}, \dots, \mathcal{D}_{m,t}^{val}\}$	Validation dataset for task $\mathcal{T}_t$
$ \mathcal{D}_0^{tr} $	Number of data points in $\mathcal{D}_{0,t}^{tr}$ (same for any task $\mathcal{T}_t$)
$ \mathcal{D}_+^{tr} $	Number of data points in $\mathcal{D}_{i,t}^{tr}$ (same for any task $\mathcal{T}_t$ and any i)
$ \mathcal{D}_0^{val} $	Number of data points in $\mathcal{D}_{0,t}^{val}$ (same for any task $\mathcal{T}_t$)
$\mathcal{L}^{val}(\theta, \mathcal{D}_{0,t}^{val})$	Loss of θ on $\mathcal{D}_{0,t}^{val}$
$p(\mathcal{T})$	Task distribution
θ_t^*	Optimal solution parameter for task $\mathcal{T}_t$
θ_t	Task-specific parameter for task $\mathcal{T}_t$
$D(\phi, \mathcal{T}_{1:T})$	Distance between ϕ and optimal parameters of $\mathcal{T}_{1:T}$
$\mathcal{S}^*(p(\mathcal{T}))$	Task similarity of $p(\mathcal{T})$
λ	regularization weight
$\mathcal{Alg}$	Within-task algorithm
ϕ / ϕ_t	Meta-parameter
$R_{0,t}$	Optimality gap
$R_{i,t}$	The i -th constraint violation
$\bar{R}_{0,[1:T]}$	Task-averaged optimality gap (TAOG)
$\bar{R}_{i,[1:T]}$	Task-averaged constraint violatio (TACV)