Contrastive Learning Neuromotor Interface From Teacher

Kilian Freitag[*]
Chalmers University of Technology
Gothenburg, Sweden
tamino@chalmers.se

Ran Wei[*]
rw422@tamu.edu

Abstract

Adaptive neuromotor interfaces are poised to enhance human-computer interaction experience and increase human mobility and accessibility to diverse users. These interfaces are important in assisting human users with dynamic control tasks where direct, physical interactions may be undesirable or impossible by predicting user intent from neuro signals. Existing methods for developing neuromotor interfaces suffer from distribution shifts due to inter and intra-user variability and the requirement for large amounts of supervised training data. Towards implicitly adapting to streaming user behavior *without* intent labels, we propose an interactive contrastive fine-tuning method to address these limitations. We formulate pseudo-intent labeling as a Bayesian inference problem guided by an optimal task policy referred to as the "teacher" prior. Using a simulated robotic control task, we show that our method successfully aligns with user intent even when the teacher prior is misspecified against a diverse group of simulated users.

1 Introduction

Adaptive human-computer interfaces (HCI) are on the verge of challenging the status quo of how we interact with computers as well as increasing human mobility in scenarios where direct, physical interaction may be undesirable or impaired, such as in robot teleoperation and active prosthesis [Meeker and Ciocarlie, 2019, Luu et al., 2022]. A human-computer interface takes a human-provided signal or command as input and generates an output to interact with an environment. The goal of the output is typically to match human intent (e.g., an intended control action) to perform some tasks in the environment. In this work, we focus on surface electromyography (EMG) which is a type of neuromotor interface based on readings of electric signals from muscles using electrodes attached to the user's skin. These interfaces have a wide range of potential applications, particularly in gesture-based control [Li et al., 2021] and handwriting recognition [Beltrán Hernández et al., 2020].

Developing EMG interfaces typically involves collecting pairs of motor intent and EMG signals from users and training an EMG policy to decode motor intent. However, this approach often degrades due to distribution shifts introduced by inter and intra-user variability. Transfer learning has been explored to improve generalization across distributions by incorporating more data from diverse users and scenarios [Cote-Allard et al., 2021], but it requires large amounts of training data and data collection effort, particularly for large or continuous action spaces. Recent works have demonstrated online interactive training as a promising way to reduce data requirements and adapt to both the user and the task distributions [Gijsberts et al., 2014, Freitag et al., 2023]. However, these methods still require that (proxy) intent labels can be obtained from the online training environment. Alternatively, Reddy et al. [2022] demonstrated the feasibility of unsupervised interface policy alignment using mutual information objectives. However, the authors also found that the resulting interface policy

[*]Equal contribution.

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

may be entangled with task environment dynamics, which potentially limits user mobility and task transfer. Thus, a key challenge in applying the interactive fine-tuning paradigm to general control tasks is to *obtain motor intent labels from user behavior*.

In this work, we propose a semi-supervised contrastive learning approach to address the lack of intent-labeled data and distribution shift when learning adaptive EMG interfaces from diverse users. We formulate user intent extraction as a Bayesian inference problem to naturally leverage labeled pre-training data for unlabeled fine-tuning data. Then, we use contrastive mutual information estimation [Aitchison and Ganev, 2021, Walker et al., 2023] to compute the posterior over latent intent from input EMG signals. For the contrastive method to extract meaningful intents, we use a class of prior whose density is proportional to that of the optimal task policy which we refer to as the "teacher". Using simulated robotics control experiments, we show that the teacher prior is necessary to achieve both high task performance and intent prediction accuracy. More importantly, we show that the teacher prior is also not overly restrictive when adapting to diverse users for whom it may be misspecified due to varying task expertise and EMG policy quality awareness.

2 Contrastive Learning neuromotor Interface from Teacher (LIFT)

2.1 Problem setup

We consider the setting where a human user operates in a dynamic control task modeled as a Markov Decision Process (MDP) by the tuple $(\mathcal{S}, \mathcal{A}, P, R, \mu, \gamma)$ where S is a set of states, $\mathcal{A}$ a set of actions, $P : \mathcal{S} \times \mathcal{A} \to \Delta(S)$ a transition probability distribution, $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ a reward function, $\mu : \Delta(\mathcal{S})$ an initial state distribution, and $\gamma \in (0, 1)$ a discount factor. An optimal control policy in the MDP is a mapping $\pi^* : \mathcal{S} \to \Delta(\mathcal{A})$ which maximizes the expected discounted cumulative rewards defined as: $J(\pi) = \mathbb{E}_{\mu,P,\pi}[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t)]$.

In the setting of motor control through an EMG interface (e.g., in robot teleoperation), we assume the user has a control strategy $\pi(a|s)$ which maps environment states to a motor intent. However, the user cannot directly implement their intended actions in the environment but their intended actions can be sensed through EMG devices attached to their limbs. We assume the EMG signals x are sampled from a distribution $P(x|a)$. To facilitate the user with the control task, we wish to train an EMG policy $E(a|x)$ that takes human EMG signals x as input and outputs (or decodes) their intended action a to the environment. The goal for the EMG policy is *not* to maximize return in the control task per se but rather maximize user mobility by making accurate predictions of their motor intents.

2.2 Pseudo-labeling motor intent from prior teacher

Our main goal is to infer unobserved user motor intents from their behavior, which serve as pseudo labels for fine-tuning the EMG policy. The combination of user and the current EMG policy generates a dataset of trajectories in the form of sequences of environment states, EMG signals, and EMG policy predictions: $\mathcal{D} = \{\tau_{1:N}\}, \tau = (s_{0:T}, x_{0:T}, u_{0:T})$, where we use u to denote the control actions predicted by the data-collecting EMG policy. We formulate pseudo-labeling as computing a posterior distribution over the missing motor intents a:

$$\begin{aligned} &P(a_{0:T}|s_{0:T}, x_{0:T}, u_{0:T}) \\ &= \frac{\prod_{t=0}^{T} E(u_t|x_t)P(x_t|a_t)\pi(a_t|s_t)P(s_t|s_{t-1}, u_{t-1})}{\int_{a_{0:T}} \prod_{t=0}^{T} E(u_t|x_t)P(x_t|a_t)\pi(a_t|s_t)P(s_t|s_{t-1}, u_{t-1})} = \prod_{t=0}^{T} \frac{P(x_t|a_t)\pi(a_t|s_t)}{\int_{a_t} P(x_t|a_t)\pi(a_t|s_t)} \end{aligned} \tag{1}$$

where $P(s_0|s_{-1}, u_{-1}) = \mu(s_0)$ is the initial environment state distribution.

To overcome the intractability of computing the exact posterior, we instead train an approximate posterior $Q(a|x)$ using auto-encoding variational Bayes [Kingma and Welling, 2013]. However, we also do not want to train a generative model of realistic EMG signals in the form of $P(x|a)$ with our interactively collected dataset. We thus resort to an energy-based or recognition-based formulation from [Poole et al., 2019, Aitchison and Ganev, 2021, Walker et al., 2023] where the distribution $P(x|a)$ is jointly parameterized by an energy function $f(x, a)$ and the empirical observed data distribution $\hat{P}(x) = \frac{1}{N}\sum_i \delta(x - x_i), x_i \in \mathcal{D}$ as:

$$P_f(x|a) = \frac{\hat{P}(x)e^{f(x,a)}}{F_f(a)}, \quad F_f(a) = \int_x \hat{P}(x)e^{f(x,a)} \tag{2}$$

To ensure that the latent variable learned by the model corresponds to environment control actions rather than arbitrary permutations common in latent variable models [Murphy, 2012], we use a strong prior that the unknown user policy is the optimal control policy $\pi = \arg\max \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t)]$. We refer to this policy as the "teacher" since it contains a substantial amount of knowledge to align the latent variable a to environment controls. Following Burgess et al. [2018], we also introduce a temperature parameter $\beta \geq 0$ to control the strength of this assumption where a higher value encourages the encoder to copy the teacher and a lower value encourages the encoder to learn from EMG signals. The well-known evidence lower bound (ELBO) objective can be written as follows:

$$\begin{aligned}
&L(f, Q; \mathcal{D}) \\
&= \mathbb{E}_{(x,s)\sim\mathcal{D}, a\sim Q(\cdot|x)}[\log P_f(x|a) + \beta(\log \pi(a|s) - \log Q(a|x))] \\
&= \mathbb{E}_{(x,s)\sim\mathcal{D}, a\sim Q(\cdot|x)}[\log \hat{P}(x) + \log e^{f(x,a)} - \log F_f(a) + \beta(\log \pi(a|s) - \log Q(a|x))] \\
&= \underbrace{\mathbb{E}_{(x,s)\sim\mathcal{D}, a\sim Q(\cdot|x)}[f(x,a) - \log \mathbb{E}_{\tilde{x}\sim\hat{P}(\cdot)}[e^{f(\tilde{x},a)}]]}_{\text{MI loss}} - \beta \underbrace{\mathbb{E}_{(x,s)\sim\mathcal{D}}\mathbb{KL}[Q(a|x)||\pi(a|s)]}_{\text{Teacher loss}} + C
\end{aligned} \tag{3}$$

where $C = \mathbb{E}_{x\sim\mathcal{D}}[\log \hat{P}(x)]$. The first term is the contrastive loss from mutual information-based representation learning approaches [Poole et al., 2019]. We thus use the InfoNCE variation due to its low variance [Oord et al., 2018]. The second term encourages the encoded latent variables to align with environment actions via the teacher.

2.3 Contrastive pre-training and fine-tuning

We use the contrastive approach for both pre-training and fine-tuning so that the energy function $f(x, a)$ is warm-started for the fine-tuning stage. Given the pre-training dataset has ground truth intent labels (denoted with y), we adapt the objective in (3) by replacing the teacher prior $\pi(a|s)$ with a data-dependent prior $P(a|y) = \mathcal{N}(a|y, \sigma^2)$ which is a Gaussian distribution centered at y with fixed variance σ^2. We also add a log-likelihood loss to (3) for the observed label where the likelihood is defined as $P(y|a) = \mathcal{N}(y|a, \sigma^2)$ with the same variance.

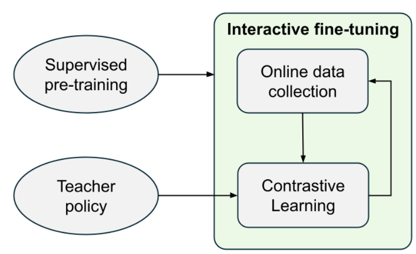


Figure 1: Proposed EMG policy training pipeline.

After pre-training, the iterative fine-tuning proceeds in multiple data collection sessions. Between each session, we initialize the energy function f and encoder $Q(a|x)$ from the previous session and optimize a weighted combination of pre-training (pt) and fine-tuning (ft) losses: $L(f, Q; \mathcal{D}^{ft}) + \lambda L(f, Q; \mathcal{D}^{pt})$ where $\lambda \geq 0$. We then set the latest EMG policy to the updated encoder $Q(a|x)$ to collect data in the next session. See Fig. 1 for our training pipeline.

3 Experiments

We perform simulated experiments in the Fetch-Reach environment, which is a 3 continuous DOF goal-reaching environment with varying goal locations in the Gymnasium-Robotics platform [Plappert et al., 2018]. For pre-training we use the dataset presented in [Côté-Allard et al., 2019] which contains single-DOF discrete movements. We use a K-nearest neighbor-based data augmentation method to generate additional synthetic continuous multi-DOF EMG feature-action pairs. The same method is used to simulate EMG features of new users during fine-tuning. We then train the encoder using the method described in section 2 and evaluate the EMG policy's action prediction accuracy in mean-absolute error (MAE) and achieved rewards when interfacing with simulated users. We repeat all experiments with 4 seeds and report the average metrics. See Appendix A.3 for details.

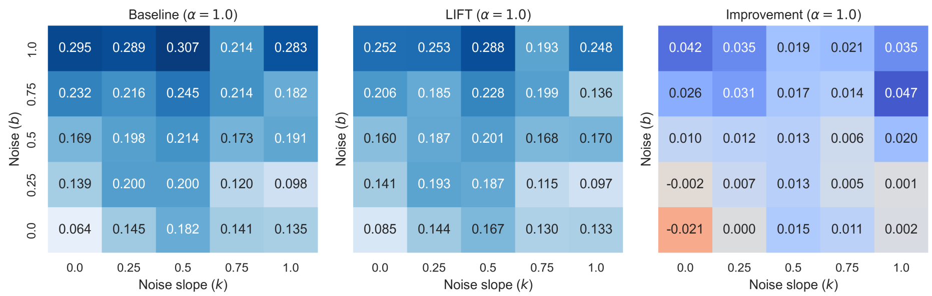


Figure 3: User intent prediction accuracy (MAE) achieved by a copy-the-teacher baseline and LIFT (ours) and the improvement upon the baseline for different user beliefs of EMG policy noises and slopes. LIFT achieved higher accuracy when the users were more misalignment with the teacher prior and believed the EMG policy was more noisy.

The importance of teacher prior To understand the importance of the teacher prior in (3) for accurately predicting user intent, we fine-tuned a set of EMG encoders with varying β from 0 to 1. Setting β to 0 corresponds to optimizing only the mutual information between EMG signals and encoder outputs, fully ignoring alignment with environment controls and the teacher. In order to rule out the possibility that the importance of the teacher prior reduces with the amount of training data, we sampled a large number of environment steps (10k) in a single data collection session and trained the EMG policy $Q(a|x)$ on this dataset. Fig. 2 shows that setting $\beta = 0$ failed to align with user intent as indicated by the high action prediction accuracy. However, only a small β value was sufficient to achieve low MAE. For subsequent experiments, we choose $\beta = 0.5$ and $\lambda = 0.5$.

Adapting to diverse users We then study whether our method can effectively adapt to diverse users, whose policies may deviate substantially from the optimal teacher policy. To simulate these diverse users, we built a user model that chooses actions based on 1) their subjective beliefs of the EMG policy quality and 2) their task expertise (see details in Appendix A.1). We model subjective beliefs using two parameters: the amount of noise it adds to the user's true intent (noise b) and how the noise level increases with the intended action magnitude (slope k). We model expertise using a user policy noise scaling parameter α. As shown in Fig. 5 in the Appendix, the teacher prior becomes strongly misspecified at higher noise, low slope, and high α. Our goal is to understand how misspecification affects the performance of our method, how it interacts with β, and most importantly, whether it offers benefit over a simple baseline copying the teacher prior by dropping the contrastive loss.

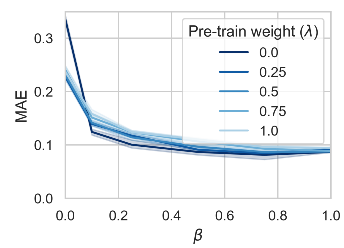


Figure 2: Intent prediction accuracy (MAE) for different weightings of teacher prior (β) and pre-train loss (λ).

Our main results are shown in Fig. 3 which plots the MAE for a range of user beliefs of EMG policy noises and slopes for $\alpha = 1$. As expected, the best MAEs are achieved at low noise levels where the teacher prior is well-specified. The benefit of our approach is most pronounced at higher noise levels where in the best case we achieved an improvement of 0.047 MAE ($\approx$ 25 %) over the copy-the-teacher baseline. As illustrated in Fig. 6 (Appendix), a similar pattern holds for nosier users with $\alpha = 3$, albeit at a smaller amount of improvements. Fig. 8 in the Appendix shows that both MAE and reward improves over the training sessions, highlighting the utility of interactive training. However, a limitation of the experiment is that, compared to the baseline, LIFT's final improvement in MAE did not lead to significant increase in task reward most likely due to the simplicity of the task (see Appendix Fig. 7).

4 Conclusion & Limitations

In this paper, we proposed a contrastive approach to aligning EMG interface policies with diverse user intent without direct supervision or reward. A limitation of this work is the simplicity of the control task and the lack of a user study. We hope to address these in future work.

Acknowledgment

The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725.

References

Laurence Aitchison and Stoil Ganev. Infonce is variational inference in a recognition parameterised model. *arXiv preprint arXiv:2107.02495*, 2021.

Mark Beliaev, Andy Shih, Stefano Ermon, Dorsa Sadigh, and Ramtin Pedarsani. Imitation learning by estimating expertise of demonstrators. In *International Conference on Machine Learning*, pages 1732–1748. PMLR, 2022.

José Guadalupe Beltrán Hernández, José Ruiz Pinales, Pedro López Rodríguez, José Luis López Ramírez, and Juan Gabriel Aviña Cervantes. Multi-stroke handwriting character recognition based on semg using convolutional-recurrent neural networks. *MATHEMATICAL BIOSCIENCES AND ENGINEERING, 17 (5), 5432-5448.*, 2020.

Christopher P Burgess, Irina Higgins, Arka Pal, Loic Matthey, Nick Watters, Guillaume Desjardins, and Alexander Lerchner. Understanding disentangling in β-vae. *arXiv preprint arXiv:1804.03599*, 2018.

Ulysse Côté-Allard, Cheikh Latyr Fall, Alexandre Drouin, Alexandre Campeau-Lecours, Clément Gosselin, Kyrre Glette, François Laviolette, and Benoit Gosselin. Deep learning for electromyographic hand gesture signal classification using transfer learning. *IEEE transactions on neural systems and rehabilitation engineering*, 27(4):760–771, 2019.

Ulysse Cote-Allard, Gabriel Gagnon-Turcotte, Angkoon Phinyomark, Kyrre Glette, Erik Scheme, François Laviolette, and Benoit Gosselin. A transferable adaptive domain adversarial neural network for virtual reality augmented emg-based gesture recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:546–555, 2021.

Kilian Freitag, Rita Laezza, Jan Zbinden, and Max Ortiz-Catalan. Improving bionic limb control through reinforcement learning in an interactive game environment. 2023.

Arjan Gijsberts, Rashida Bohra, David Sierra González, Alexander Werner, Markus Nowak, Barbara Caputo, Maximo A Roa, and Claudio Castellini. Stable myoelectric control of a hand prosthesis using non-linear incremental learning. *Frontiers in neurorobotics*, 8:8, 2014.

Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*, pages 1861–1870. PMLR, 2018a.

Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, et al. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*, 2018b.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Wei Li, Ping Shi, and Hongliu Yu. Gesture recognition using surface electromyography and deep learning for prostheses hand: state-of-the-art, challenges, and future. *Frontiers in neuroscience*, 15: 621885, 2021.

Diu Khue Luu, Anh Tuan Nguyen, Ming Jiang, Markus W Drealan, Jian Xu, Tong Wu, Wing-kin Tam, Wenfeng Zhao, Brian ZH Lim, Cynthia K Overstreet, et al. Artificial intelligence enables real-time and intuitive control of prostheses via nerve interface. *IEEE Transactions on Biomedical Engineering*, 69(10):3051–3063, 2022.

Cassie Meeker and Matei Ciocarlie. Emg-controlled non-anthropomorphic hand teleoperation using a continuous teleoperation subspace. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 1576–1582. IEEE, 2019.

Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

Markus Nowak and Claudio Castellini. The let procedure for prosthetic myocontrol: towards multi-dof control using single-dof activations. *PLoS One*, 11(9):e0161678, 2016.

Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

Matthias Plappert, Marcin Andrychowicz, Alex Ray, Bob McGrew, Bowen Baker, Glenn Powell, Jonas Schneider, Josh Tobin, Maciek Chociej, Peter Welinder, et al. Multi-goal reinforcement learning: Challenging robotics environments and request for research. *arXiv preprint arXiv:1802.09464*, 2018.

Ben Poole, Sherjil Ozair, Aaron Van Den Oord, Alex Alemi, and George Tucker. On variational bounds of mutual information. In *International Conference on Machine Learning*, pages 5171–5180. PMLR, 2019.

Sid Reddy, Anca Dragan, and Sergey Levine. Where do you think you're going?: Inferring beliefs about dynamics from behavior. *Advances in Neural Information Processing Systems*, 31, 2018.

Siddharth Reddy, Sergey Levine, and Anca Dragan. First contact: Unsupervised human-machine co-adaptation via mutual information maximization. *Advances in Neural Information Processing Systems*, 35:31542–31556, 2022.

Francisco R Ruiz, Titsias RC AUEB, David Blei, et al. The generalized reparameterization gradient. *Advances in neural information processing systems*, 29, 2016.

William I Walker, Hugo Soulat, Changmin Yu, and Maneesh Sahani. Unsupervised representation learning with recognition-parametrised probabilistic models. In *International Conference on Artificial Intelligence and Statistics*, pages 4209–4230. PMLR, 2023.

Ran Wei, Siliang Zeng, Chenliang Li, Alfredo Garcia, Anthony D McDonald, and Mingyi Hong. A bayesian approach to robust inverse reinforcement learning. In *Conference on Robot Learning*, pages 2304–2322. PMLR, 2023.

A Appendix

A.1 User model

Following a long tradition of modeling rational user behavior under incorrect beliefs [Reddy et al., 2018, Wei et al., 2023, Beliaev et al., 2022], we designed a set of user models (i.e., their true policies) for which the teacher policy may be strongly misspecified to test the performance of our approach in these settings. We assume all policies in the user model set optimize the true task reward, given users receive clear task instructions. The primary reasons why policies in the set could differ are 1) users have different beliefs about the quality of the EMG policy and 2) users have different levels of task expertise. We now introduce our model for these aspects respectively.

Modeling user belief In our prior exploration of the pre-trained EMG policy, we found that the intent prediction error the policy makes on average is positively correlated with the ground truth action magnitude and can be described with a linear trend and a positive bias. Inspired by this, we model the user's belief of the quality of the EMG policy using a positive slope parameter k and a bias b. These two parameters jointly determine the amount of Gaussian distributed noise $\epsilon \sim \mathcal{N}(0, k \cdot |a| + b)$ it adds to corrupt user-intended actions. The slope parameter increases noise when the action magnitude is high, forcing the user to take smaller magnitude actions. On the other hand, the base noise b forces the user to take higher magnitude actions to neutralize or "fight" EMG policy noise. We denote the concatenation of slope and base noise as $z = [k, b]$ and the stochastic mapping from user intent to decoded control from this process as $P(u|a, z)$.

We obtain the type-conditioned user policy $\pi(a|s, z)$ using meta reinforcement learning in a user simulator developed based on the above user model. At the beginning of each training episode, we first sample a base noise $b \sim P(b)$ and a slope $k \sim P(k)$. We then train the user policy to maximize the following criterion:

$$\max_{\pi} \mathbb{E}_{\substack{b \sim P(\cdot), k \sim P(\cdot), \pi(\cdot|s_t, z) \\ u_t \sim P(\cdot|a_t, z), s_{t+1} \sim P(s_{t+1}|s_t, u_t)}} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, u_t) \right] \tag{4}$$

where the reward is computed from the EMG policy-predicted actions rather than the user intended actions. The value functions for the optimal user policy are defined as:

$$\begin{aligned} Q(s, a, z) &= \mathbb{E}_{u \sim P(\cdot|a, z)}[R(s, u) + \gamma \mathbb{E}_{s' \sim P(\cdot|s, u)}[V(s', z)] \\ V(s, z) &= \max_a Q(s, a, z) \end{aligned} \tag{5}$$

Modeling user expertise Following Beliaev et al. [2022], we model user task expertise using state-dependent variance scaling of a policy, which is applied to the above type-conditioned user policy. Let $\alpha : \mathcal{S} \to \mathbb{R}^+$ denote a mapping from a state to a (negative) temperature scale, we define the following user policy:

$$\pi(a|s, z, \alpha) = \frac{\exp(\alpha(s) Q(s, a, z))}{\int_{\tilde{a}} \exp(\alpha(s) Q(s, \tilde{a}, z))} \tag{6}$$

where higher $\alpha(s)$ reduces the variance of the policy and concentrates the action probability density around the optimal actions; on the other hand, lower $\alpha(s)$ increases the variance of the policy.

For Gaussian policies (used in our experiments) where for each state a mean μ and a variance σ^2 over actions is predicted by the policy network, we can show that (6) simply corresponds to reducing the variance for higher $\alpha(s)$ (see blow). We thus apply variance scaling instead of fitting another policy network to achieve (6).

We design a variance scaling function $\rho : \mathcal{S} \to \mathbb{R}^+$ which computes the distance to the goal along each dimension from the environment state and outputs a scale by linearly interpolating the current distance on an interval of variance scale against a minimum and maximum applicable distance. The minimum of the variance scale interval is set to 1 so that the original (zero-temperature) policy variance is used. We then apply the scale to the optimal policy obtained from (4) and compute the scaled variance as $(\rho(s)\sigma)^2$.

Proposition A.1. *For Gaussian distributions with mean and variance parameters $\theta = [\mu, \sigma^2]$ fitted to approximate policies of the form:*

$$\pi_\theta(a) \approx \frac{\exp(\alpha Q(a))}{\int_{\tilde{a}} \exp(\alpha Q(\tilde{a}))} \tag{7}$$

where $Q(a)$ is the action value function and α is a temperature parameter, changing α is equivalent to scaling the variance of the Gaussian distribution.

Proof. Fitting the above policy is equivalent to solving the following optimization problem [Haarnoja et al., 2018a]:

$$\max_\theta \mathbb{E}_{\pi_\theta(a)}[Q(a)] + \frac{1}{\alpha}\mathbb{H}[\pi_\theta(a)] \tag{8}$$

where $\mathbb{H}[\cdot]$ denotes Shannon entropy.

To show that different α does not affect the mean parameter of the Gaussian distribution, recognize that:

$$\begin{aligned} &\nabla_\theta \mathbb{E}_{\pi_\theta(a)}[Q(a)] - \frac{1}{\alpha}\nabla_\theta \mathbb{H}[\pi_\theta(a)] \\ &= \mathbb{E}_{P(\epsilon)}[\nabla_a Q(a)|_{a=\mu+\sigma\epsilon} \nabla_\theta(\mu + \sigma\epsilon)] - \frac{1}{\alpha} 0.5 \nabla_\theta \log(2\pi e \sigma^2) \end{aligned} \tag{9}$$

where we have used the reparameterization trick [Kingma and Welling, 2013, Ruiz et al., 2016] with $P(\epsilon) = \mathcal{N}(0, 1)$ to express the gradient of the first term. The temperature parameter only affects the gradient of the variance parameter σ^2 through the second term and not the mean parameter μ which only exists in the first term. Thus, for different α, we can achieve (8) by fixing the mean parameter and adjusting the variance parameter. □

A.2 Motivation for weighted fine-tuning loss

The weighted fine-tuning loss $L(f, Q; \mathcal{D}^{ft}) + \lambda L(f, Q; \mathcal{D}^{pt}), \lambda \geq 0$ defined in section 2.3 was used to retain knowledge from the pre-training session and avoid catastrophic forgetting in the fine-tuning sessions. This could be motivated from the following Bayesian principle.

Let a prior distribution over the energy function be denoted as $P(f)$. The posterior distribution over f given the pre-training dataset can be written as follows:

$$\begin{aligned} P(f|\mathcal{D}^{pt}) &\propto \exp\big(\log P(\mathcal{D}^{pt}|f) + \log P(f)\big) \\ &\approx \exp\big(|\mathcal{D}^{pt}| L(f, Q; \mathcal{D}^{pt}) + \log P(f)\big) \end{aligned} \tag{10}$$

where the approximation is due to the loss $L(f, Q; \mathcal{D}^{pt})$ being a sample-based approximation of the ELBO, which is itself a lower bound on the marginal likelihood $\log P(\mathcal{D}^{pt}|f)$.

When adapting to new users, we assume their intent to EMG mapping $P_{f'}(x|a)$ is a perturbation of that of the pre-training users, i.e., the transition distribution $P(f'|f)$ is concentrated around f. This should result in a prior $P(f'|\mathcal{D}^{pt}) = \int_f P(f'|f)P(f|\mathcal{D}^{pt})$. Instead of modeling the transition distribution, we directly model the prior over f' using the pre-trained log-posterior and a temperature parameter $\lambda \geq 0$ as:

$$P(f'|\mathcal{D}^{pt}) \propto \exp(\lambda \log P(f'|\mathcal{D}^{pt})) \tag{11}$$

After observing the fine-tuning dataset, the posterior over f' can be written as:

$$\begin{aligned} P(f'|\mathcal{D}^{ft}) &\propto \exp\big(\log P(\mathcal{D}^{pt}|f') + \log P(f'|\mathcal{D}^{pt})\big) \\ &\approx \exp\big(|\mathcal{D}^{ft}| L(f', Q; \mathcal{D}^{ft}) + \lambda |\mathcal{D}^{pt}| L(f', Q; \mathcal{D}^{pt}) + \lambda \log P(f')\big) \end{aligned} \tag{12}$$

Omitting the initial prior, the approximate log-posterior after fine-tuning is proportional to:

$$\log P(f'|\mathcal{D}^{ft}) \propto L(f', Q; \mathcal{D}^{ft}) + \lambda \frac{|\mathcal{D}^{pt}|}{|\mathcal{D}^{ft}|} L(f', Q; \mathcal{D}^{pt}) \tag{13}$$

We subsume all weightings under a single hyperparameter to facilitate tuning.

A.3 Implementation detail

Our implementation is available at `https://github.com/KilianFt/LIFT`.

Dataset and features For supervised pre-training, we use the dataset presented in Côté-Allard et al. [2019] which contains 7 different classes of wrist movements that closely resemble movements in 3 DOF: Neutral, Radial and Ulnar Deviation, Wrist Flexion and Extension, and Hand Open and Close. The dataset was recorded using a Thalmic Labs Myo Armband, with 8 EMG channels and a frequency of 200 Hz from 40 participants (28 males and 12 females). Each movement was recorded for 20-80 seconds depending on the data acquisition protocol. Samples were collected in chunks of 5 seconds with pauses in between to avoid user fatigue. These movements were treated as discrete prediction targets in [Côté-Allard et al., 2019].

We split EMG signals into windows of size 200 with an overlap of 150. For each channel in a window, we extract the mean absolute value resulting in a vector x of length 8 used as input to the EMG policy.

Continuous supervised pre-training on discrete data While our dataset (and all available others) are collected on discrete actions, our control task has continuous action space. We solve this problem using data augmentation where we linearly interpolate discrete action EMG features following Nowak and Castellini [2016] to obtain continuous multi-DOF labels.

Specifically, we create new samples as follows. Let $\mathbf{a}_{\text{aug}}$ represent a new action, $\mathbf{A} = \{\mathbf{a}_i\}_{i=1}^{n}$ be the set of original actions with features $\mathbf{x}_i$, $d(\mathbf{a}_{\text{aug}}, \mathbf{a}_i)$ the Euclidean distance between the augmented and original action and k_{nn} the number of nearest neighbors to consider.

We first compute the distance between the new action $\mathbf{a}_{\text{aug}}$ and each original action $\mathbf{a}_i$ as

$$d(\mathbf{a}_{\text{aug}}, \mathbf{a}_i) = \|\mathbf{a}_{\text{aug}} - \mathbf{a}_i\|_2 \tag{14}$$

We then select the indices of the top k_{nn} smallest distances:

$$\{i_1, i_2, \ldots, i_{k_{nn}}\} = \text{argmin}_i\, d(\mathbf{a}_{\text{aug}}, \mathbf{a}_i) \tag{15}$$

which in turn are used to calculate the weights of each selected action as the inverse of the distances to the top-k nearest neighbors:

$$w_j = \frac{1}{d(\mathbf{a}_{\text{aug}}, \mathbf{a}_{i_j}) + \epsilon} \quad \text{for } j \in \{1, 2, \ldots, k_{nn}\} \tag{16}$$

where ϵ is a small value to prevent division by zero. The weights are normalized so that they sum to 1:

$$\hat{w}_j = \frac{w_j}{\sum_{j=1}^{k_{nn}} w_j} \tag{17}$$

Finally, the interpolated feature $\mathbf{x}_{\text{aug}}$ is computed as the weighted sum of the features corresponding to the top-k nearest actions:

$$\mathbf{x}_{\text{aug}} = \sum_{j=1}^{k_{nn}} \hat{w}_j \mathbf{x}_{i_j} \tag{18}$$

Thus the interpolated feature vector is a linear combination of the features of the nearest actions, weighted by their inverse distances.

For pre-training, we held out 3 participant's data as validation and used the rest to create the augmented dataset $\mathcal{D}^{aug} = \{(x^{aug}, a^{aug})_{1:M}\}$ where a^{aug} are sampled uniformly from the action space interval $[-1, 1]$. We then combine it with the original dataset and train the initial EMG policy using supervised contrastive learning. For all experiments, we set k_{nn} to 3.

EMG simulation To evaluate our method in simulation, we developed an EMG simulator using the same mechanism as in the previous section. For each user input intent a, we output EMG features computed using (18). Importantly, the EMG data used for interpolation were sampled from the participant data held out from the training set to mimic a real user study.

Task environment We used the Fetch-Reach environment in the Gymnasium-Robotics platform [Plappert et al., 2018], where each observation is a concatenation of the kinematic information about the end effector, the desired goal and the achieved goal. The action space for our experiments has 3 continuous DOF with a range $[-1, 1]$ with a dense reward given by the negative distance to the goal. The environment observation space was only available to the teacher and not the EMG decoding policy.

Teacher & user model training To train the teacher and user policies, we used Soft Actor Critic (SAC) [Haarnoja et al., 2018a] because of its stochastic nature. The policy's output is a tanh Normal distribution in the same range as the environment action space and consists of 3 layers, 256 hidden units each, and ReLU activation function. The critic networks use the same hidden configuration. We trained for a total of 150k frames with a batch size of 256 with both policy and critic learning rates equal to $3e-4$. Following [Haarnoja et al., 2018b], we set the initial SAC temperature parameter to 1 and update it to match an automatically set policy entropy. In practice, and our experiments as well, the final temperature at the end of training becomes approximately 0.

EMG policy training The EMG policy/encoder is a multi-layer perceptron (MLP) network with 3 layers each with 256 neurons and SiLU activation function and without dropout.

For pre-training we created 1000 augmentations per person using the method in (18). We set the magnitude of the original actions present in the dataset to 0.8 (with 1.0 being the maximum). We then combined the original and augmentation dataset to get a total of 66353 samples. The data from three users (7783 samples) were held out for validation while the rest (58570 samples) was used for training.

As described in section 2.3, we learn from the ground truth action labels using a likelihood $P(y|a) = \mathcal{N}(y|a, \sigma^2)$ and a data dependent prior $P(a|y) = \mathcal{N}(a|y, \sigma^2)$. We use $\sigma = 0.2$ for both.

For the KL divergence term in (3), we use a single sample approximation for the expected log likelihood under the prior distribution and compute the encoder entropy in closed form because it predicts the parameters of Gaussian distributions. In our experiments, we noticed that down-weighting the entropy term led to improved accuracy. We thus chose an entropy weight of 0.01 for all experiments.

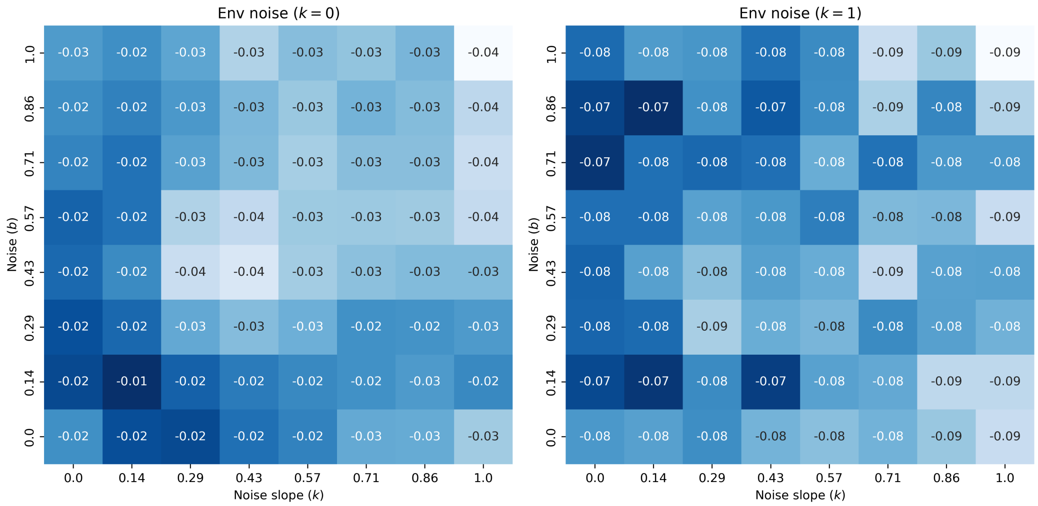
For pre-training, we trained for 50 epochs with a learning rate of $3e-4$ and a batch size of 128, which we found works well for convergence. For fine-tuning, we collect 2000 samples per session and train for a maximum of 2000 steps between each data collection session. This is repeated for 5 sessions in our experiments. Finally, in the sixth round, we again collected 2000 samples for validation purposes.

Hardware All experiments were conducted on a cluster of NVIDIA Tesla T4 GPUs. Pre-training was completed in approximately 3 minutes, while fine-tuning experiments required around 8 minutes per run, with a total of 520 fine-tuning experiments (120 for teacher importance and 400 for user adaptation).

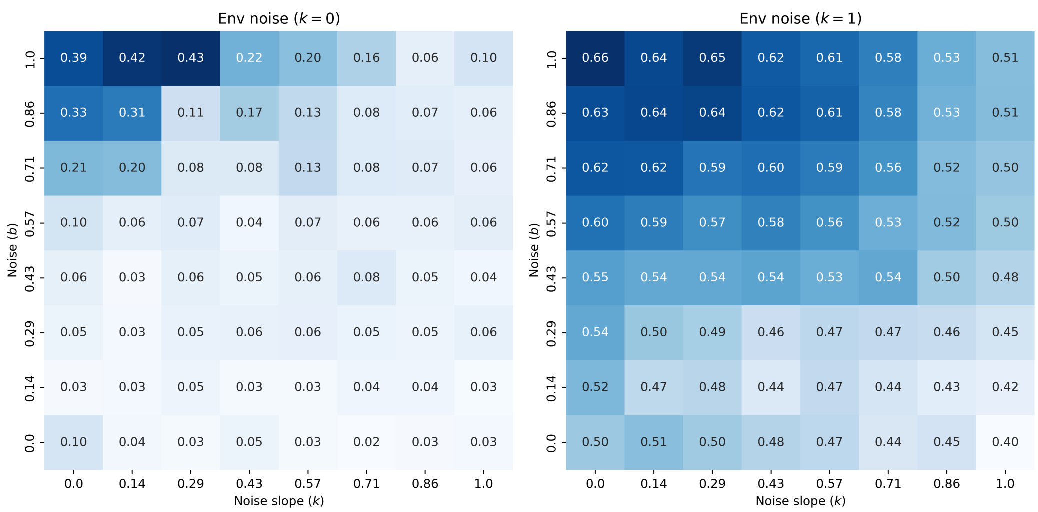
A.4 Additional results

User model In this section, we illustrate the behavior of the user model trained by meta RL. Fig. 4 shows the reward achieved and average action magnitude taken by the user model for varying user beliefs about EMG policy noise and slope under different true EMG policy noise. As expected, when the environment had low noise, users with the correct beliefs of low noise and low slope generally achieved higher reward. When the environment has higher noise, users who believed the noise slope was lower achieved higher reward. In both environment noise settings, the user model tends to take higher magnitude actions when it believes the noise level is higher and slope lower. This can be understood as "fighting against" environment noise.

In order to ensure our method can adapt to diverse users, we would also like the optimal control policy under inaccurate user beliefs make the teacher prior strongly misspecified. Fig. 5 shows that this is indeed the case. When users believe the EMG policy noise and slope are significantly higher than zero, the difference between the user policy and the teacher policy in terms of MAE of the most likely actions can be as high as 0.36.



(a) Reward achieved by simulated users with varying beliefs about EMG policy noise and slope under different true EMG policy noise.



(b) Average action magnitude taken by simulated users with varying beliefs about EMG policy noise and slope under different true EMG policy noise.

Figure 4: Illustration of simulated user behavior

Adapting to diverse users Here we include additional results for the fine-tuning experiment. Fig. 6 shows the user intent prediction MAE for the baseline and LIFT for users with $\alpha = 3$ (i.e., noisier behavior/lower expertise). The pattern is similar to Fig. 3 except that the improvement upon the baseline is smaller.

Fig. 7 shows that the improvement in reward by LIFT compared to the baseline is very small. We believe this is due to the simplicity of the Reach environment where even copying the incorrect teacher can achieve good reward.

Fig. 8 show the performance of LIFT and baseline over the fine-tuning sessions. In most cases, the highest improvements occurred in sessions 1 and 2 with smaller changes in following iterations. It can be seen that LIFT compensates for high noise more effectively than the copy-the-teacher baseline and improves the mean MAE and reward in all cases compared to pre-training.

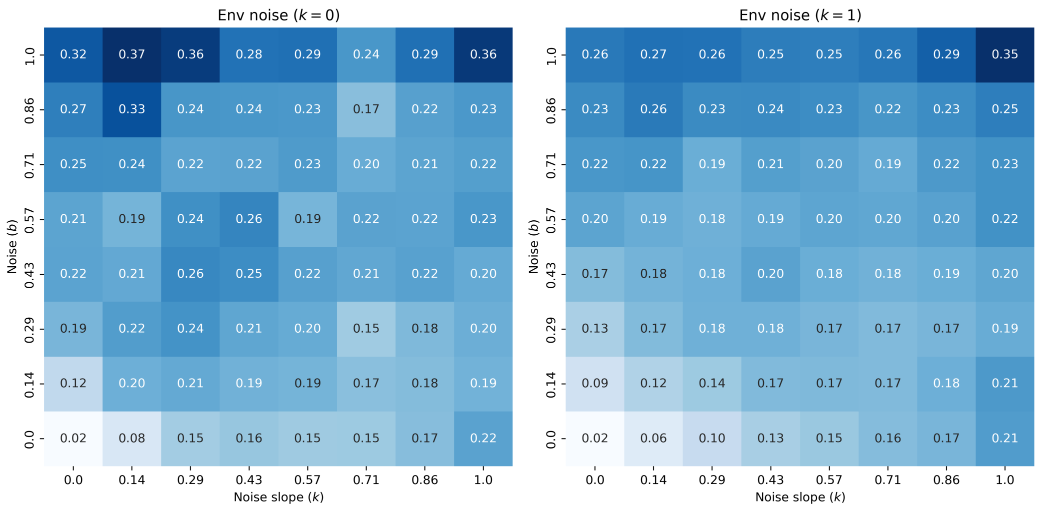


Figure 5: MAE of the most likely actions under user and teacher policies for varying beliefs about EMG policy noise and slope under different true EMG policy noise. This shows that the teacher policy can be strongly misspecified for users with incorrect beliefs.

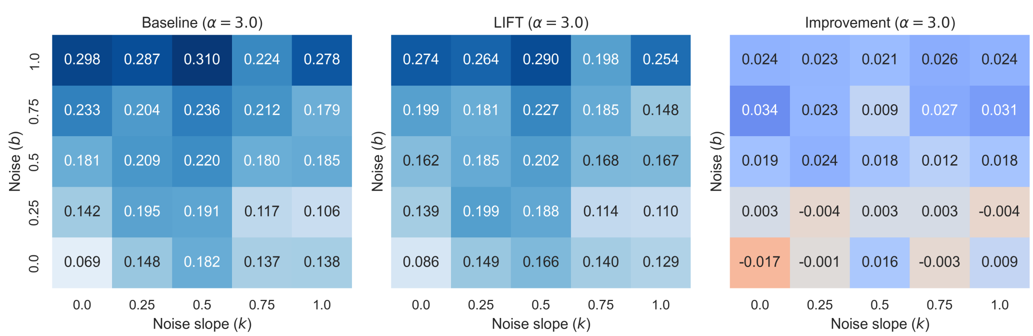


Figure 6: User intent prediction accuracy (MAE) achieved by a copy-the-teacher baseline and LIFT (ours) and the improvement upon the baseline for different user beliefs of EMG policy noises and slopes at $\alpha = 3$.

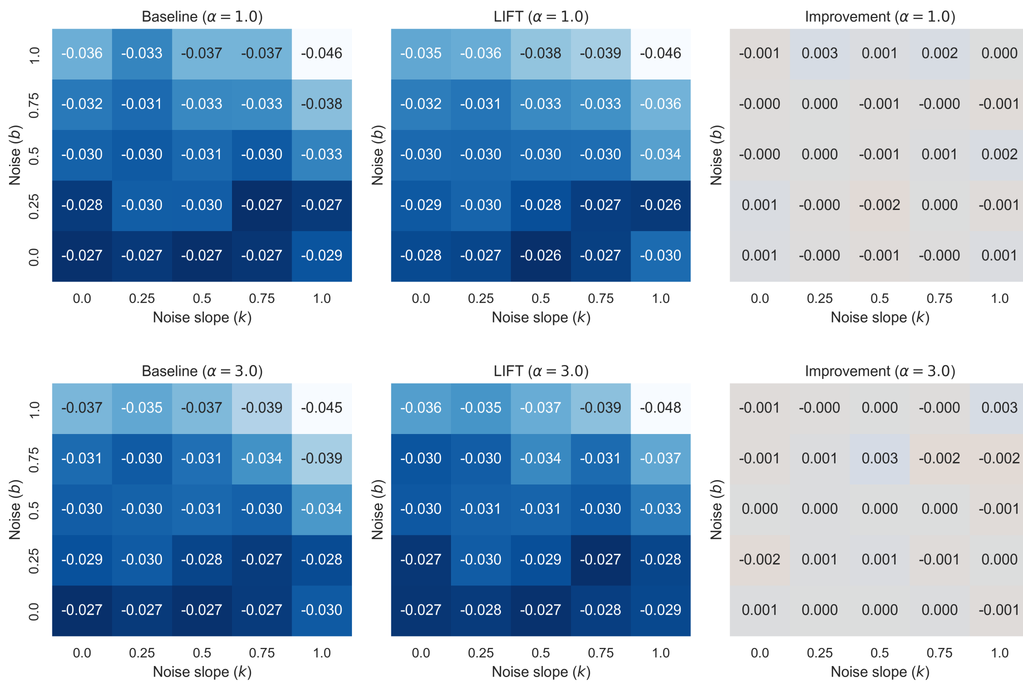


Figure 7: Task reward achieved by a copy-the-teacher baseline and LIFT (ours) and the improvement upon the baseline for different user beliefs of EMG policy noises and slopes at $\alpha = 1$ and $\alpha = 3$.

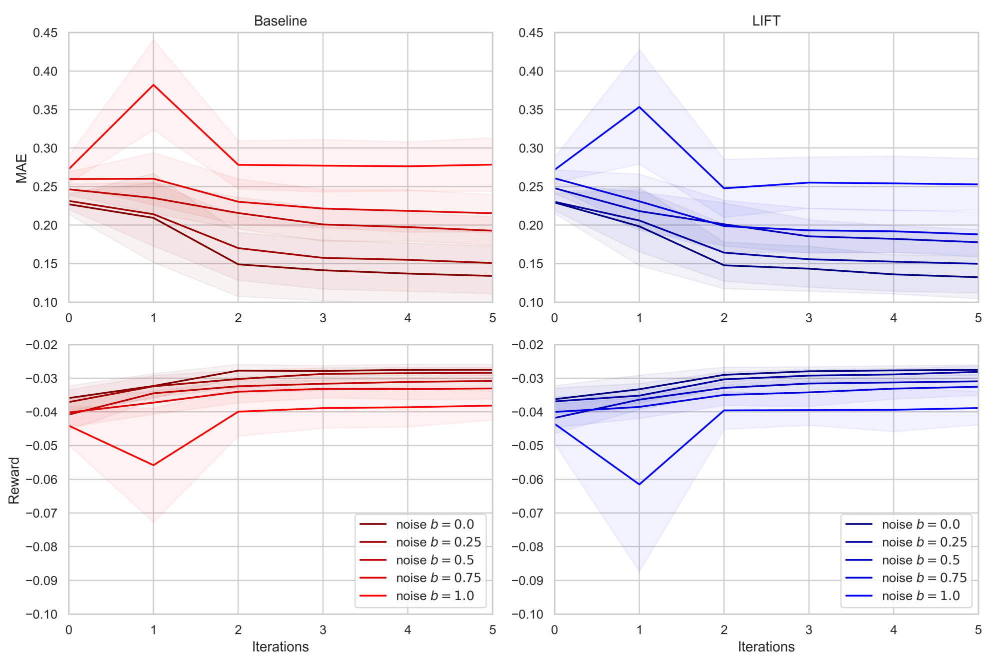


Figure 8: Mean validation MAE and reward over all interactive fine-tuning sessions after the first session for varying user beliefs of noise bias b. The average is taken over the seeds, noise slopes and policy variances. Higher noise leads to worse MAE but LIFT manages to retain better performance when noise increases compared to the copy-the-teacher baseline. Different to the MAE, reward trends between copy-the-teacher baseline and LIFT are relatively similar due to the simplicity of the task.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The goal of this paper is to propose a contrastive method for fine-tuning EMG interface policies, which is achieved by the body of the paper.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Yes, we discussed the limitations of the lack of user study and simplicity of the experiments in the final section.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: We provide all proofs in the appendix and have verified them to the best of our knowledge.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Yes, we have extensively documented implementation details in the appendix to the best we can.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: We will make the code available after the review process.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Yes, we explain our experimental details in the appendix and will open-source our code for full transparency.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Yes, we provide error bars in the appendix but rely on mean results in the main text due to space constraints.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).

- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Yes, this is documented in the appendix in the paragraph EMG policy training.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: Yes, this work conforms with the NeurIPS code of ethics to the best level we could determine.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: This paper conducts fundamental research and does not directly concern area with potential positive or negative societal impact.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: This paper poses no such risks.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: Creators and original owners of assets are properly cited.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.

- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: Tis paper does not release new assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: This paper does not involve human subjects

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: This paper does not involve human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.