SIAM J. MATH. DATA SCI.
Vol. 7, No. 3, pp. 884–905

Online Machine Teaching under Learner Uncertainty: Gradient Descent Learners of a Quadratic Loss*

Belen Martin-Urcelay†, Christopher J. Rozell†, and Matthieu R. Bloch†

Abstract. We revisit the framework of online machine teaching, a special case of active learning in which a teacher with full knowledge of a model attempts to train a learner by adaptively presenting examples. While online machine teaching example selection strategies are typically designed assuming omniscience, i.e., the teacher has absolute knowledge of the learner state, we show that efficient machine teaching is possible even when the teacher is uncertain about the learner initialization. Specifically, we consider the case of learners that perform gradient descent of a quadratic loss to learn a linear classifier, and we propose an online machine teaching algorithm in which the teacher simultaneously learns the learner state while teaching the learner. We theoretically show that the learner's mean square error decreases exponentially with the number of examples, thus achieving a performance similar to the omniscient case and outperforming two stage strategies that first attempt to make the teacher omniscient before teaching. We empirically illustrate our approach in the context of a cross-lingual sentiment analysis problem.

Key words. machine teaching, active learning, online example selection, unknown initialization

MSC codes. 68W27, 68W40, 93C55

DOI. 10.1137/24M1657997

1. Introduction.

The size of datasets used in modern machine learning has grown manyfold over the last decade, making the training of models on entire datasets frequently impractical [11], either because of the associated training time, training cost or incurred energy consumption, and environmental cost. To circumvent these constraints, it is now common to only train models on a subset of examples. Using naive data selection strategies, such as randomly sampling a dataset, typically requires more examples than intentional strategies, such as active learning, by which the machine learning algorithm adaptively requests the labels of certain data points from a large pool of unlabeled examples [31]. Active learning has been successfully applied to a wide variety of settings, such as natural language processing [39, 4], data embedding [34, 7], or source localization [22, 25]. Machine teaching (MT) considers a variation of the setup in which a knowledgeable expert knowing the ground truth model, the *teacher,* selects the examples fed to the machine learning algorithm, the *learner.* The aim of machine teaching is to exploit the teacher's knowledge and identify the smallest set of examples to train the learner [40].

*Received by the editors May 29, 2024; accepted for publication (in revised form) March 14, 2025; published electronically July 1, 2025.
https://doi.org/10.1137/24M1657997

Funding: This work was funded by the Rafael del Pino Foundation, the Fulbright Association, and the International Peace Scholarship by P.E.O.

†Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (burcelay3@gatech.edu, crozell@gatech.edu, matthieu.bloch@ece.gatech.edu).

Machine teaching has proved useful in a variety of settings, ranging from an illustrative one-dimensional threshold classifier [41] to complex vocabulary learning platforms [36]. A crucial requirement in early machine teaching algorithms has been the need for consistent learners [8, 3], which directly discard all the hypotheses that do not agree with any training example. Therefore, these algorithms do not perform well in the presence of noisy labels. The consistency requirement has been relaxed in recent literature [20, 18, 19] by introducing the concept of *omniscient teaching.* An omniscient teacher possesses full knowledge of the learner, i.e., it is able to observe the state and dynamics of the learner during training. Under certain smoothness assumptions, the selection of examples reduces to a constrained convex optimization problem, for which a greedy machine teaching algorithm as in [20] achieves an exponential speedup compared to random example selection. Nevertheless, the omniscience requirement may pose practical implementation challenges [8].

First, we note that the initial state of an algorithm is often unknown. This is the case in adversarial attacks, such as training-state poisoning [13], in which attackers lack precise knowledge about the initial state of the targeted system, such as a spam filter. Unknown initial states also result from warm-starts [28], a technique by which pretrained models are used to accelerate the learning process or transfer knowledge from related tasks. Second, we note that teacher and learner may operate in different feature spaces. For example, words may be embedded in different spaces for different languages and the mapping between language spaces may be unknown.

An existing approach to address the lack of omniscience is learning for omniscience, which consists in introducing a preliminary probing phase during which the teacher queries the learner until enough feedback is gathered to accurately approximate the unknown information about the learner, such as the learner initial state [21, 18] or the learner dynamics [35]. Unfortunately, this strategy requires many interactions between the teacher and learner during which the learner does not improve its model.

The present work aims to tackle the above limitations by developing an efficient machine teaching algorithm capable of boosting the convergence speed of learners even when the teacher is not fully omniscient. Our algorithm addresses the challenges related to unknown learner starting states and unknown orthogonal mappings between the learner and teacher feature spaces. Our main contribution is in realizing that jointly teaching the learner while estimating its parameters may offer significant and previously not identified gains. In particular, we have the following:

1. We develop a non-omniscient machine teaching algorithm for gradient descent learners of a quadratic loss with unknown initializations. We prove that our algorithm achieves an exponential speedup compared to random example selection, without an explicit probing phase to estimate the learner initialization. Additionally, the exponential convergence guarantees to hold under unknown orthonormal mappings between learner and teacher.
2. We draw connections between control theory and machine teaching under learner uncertainty. These connections allow us to leverage well-studied techniques, such as Kalman filters and Riccati recursions, to obtain theoretical guarantees on learning performance.

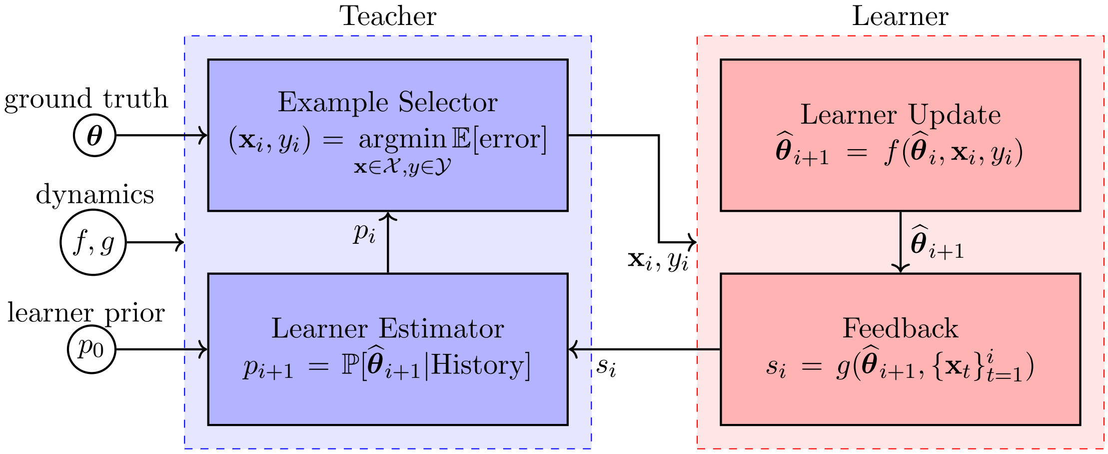


Figure 1. *Block diagram of the online machine teaching framework. The goal of the teacher is to steer the learner towards the ground truth $\boldsymbol{\theta}$, while simultaneously learning about the learner state $\widehat{\boldsymbol{\theta}}_{i+1}$.*

3. We empirically demonstrate the advantages of our framework over random sampling and probing based techniques, using the teaching of a binary sentiment classifier across languages as an example.

2. Framework. We now detail the framework of the machine teaching problem and introduce simplifying assumptions to make analytical progress in the non-omniscient setting. As illustrated in Figure 1, let the learner be a machine learning model parameterized by $\widehat{\boldsymbol{\theta}}$. For instance, $\widehat{\boldsymbol{\theta}}$ could represent an effective decision boundary. Machine teaching aims to guide the learner's learned parameter, $\widehat{\boldsymbol{\theta}}$, towards the ground truth, $\boldsymbol{\theta}$. Let the teacher be an entity with knowledge of the ground truth $\boldsymbol{\theta}$ and selecting the examples presented to the learner. At each time-step i, the teacher first presents an example and label pair $(\mathbf{x}_i, y_i)$ from a predetermined pool $(\mathcal{X}, \mathcal{Y})$ to the learner. The learner then uses the example to update its model $\widehat{\boldsymbol{\theta}}_{i+1} = f(\widehat{\boldsymbol{\theta}}_i, \mathbf{x}_i, y_i)$ for some known function f. The learner may also provide some feedback with information about its current state to the teacher $s_i = g(\widehat{\boldsymbol{\theta}}_{i+1}, \{\mathbf{x}_t\}_{t=1}^i)$, where g is some known function. In the case of an omniscient teacher, this feedback provides the exact learner state.

Although we assume that the teacher knows the function f that the learner uses to update its state, we emphasize that the teacher is not omniscient. Namely, the teacher does not know the starting point of the learner, $\widehat{\boldsymbol{\theta}}_0$. Instead, we assume the teacher starts with a prior Gaussian probability distribution $\mathbf{p}_0$ for $\widehat{\boldsymbol{\theta}}_0$. We shall also consider the case in which the teacher and learner do not share the same feature spaces: when the teacher selects an example $\mathbf{x}$, the learner observes $\widehat{\mathbf{x}} = \mathcal{G}(\mathbf{x})$, where $\mathcal{G}$ is an unknown orthonormal mapping between the teacher and learner feature spaces.

For analytical tractability, we restrict our attention to a learner that performs gradient descent to minimize the quadratic loss $l(\widehat{\boldsymbol{\theta}}) := \frac{1}{2}\|\widehat{\boldsymbol{\theta}}^T \mathbf{x} - y\|_2^2$. At each iteration the learner updates its state according to

$$\widehat{\boldsymbol{\theta}}_{i+1} = \widehat{\boldsymbol{\theta}}_i - \tau \left(\widehat{\boldsymbol{\theta}}_i^T \mathbf{x}_i - y_i\right) \mathbf{x}_i, \tag{2.1}$$

where $\tau \in \mathbb{R}^+$ is the learning rate, assumed known to the teacher. While this assumption restricts us to linear learners, it allows us to provide theoretical insights and to design a tractable algorithm. We denote the maximum norm of the states by P, i.e., $\max_i \|\widehat{\boldsymbol{\theta}}_i\|_2^2 \leq P$. We specifically look at teaching a linear binary classifier $\boldsymbol{\theta}$, s.t. $\|\boldsymbol{\theta}\|_2^2 \leq P$. The classifier labels any example $\mathbf{x} \in \mathcal{X}$ as $y = \text{sign}(\boldsymbol{\theta}^T \mathbf{x})$. In principle, one could attempt to extend the linear classifier to nonlinear problems by mapping the original nonlinear space into a higher-dimensional feature space in which the data is linearly separable, though this mapping is often hard to find in practice.

We consider synthesis based teaching [20] by which the teacher may provide any example within a ball $\mathcal{X} : \{\mathbf{x} = [1, x_1, \ldots, x_{d-1}]^T \in \mathbb{R}^d; \|\mathbf{x}\|_2^2 \leq P_{\mathbf{x}}\}$, together with any binary label in $\mathcal{Y} : \{-1, 1\}$. Following standard practice, the first coordinate of the examples is set to 1 to allow for the parameter $\boldsymbol{\theta}$ to account for both the direction and the offset of the hyperplane characterizing the classifier. The freedom to synthetically generate examples may lead to non-semantically-meaningful examples. To maintain interpretability, one can restrict the examples space $\mathcal{X}$ to data points that a teacher generates with a variational autoencoder (VAE) trained from a predefined dataset of meaningful examples. This restriction forces synthetic examples to resemble the original training dataset and thus be interpretable [29].

3. Theoretical guarantees.

Existing online teachers base their example selection criteria on their knowledge of the learner state, which naturally prompts a number of questions: How does a teacher handle learner uncertainty? Are there any convergence guarantees in that case? We tackle these questions under two different settings: when the teacher receives no information from the learner, and when the teacher receives some noisy feedback from the learner at each iteration.

3.1. Simultaneous Machine Teaching and Learning (SMTL) without feedback.

As a baseline, we first consider the situation in which the teacher receives no feedback from the learner. At each iteration, the teacher only communicates with the learner via a single example-label pair. We propose a greedy algorithm that chooses the example-label pair that most reduces the expected error of the learned parameter from one iteration to the next. The algorithm is motivated by the decomposition of the mean-square error (MSE) of the learned parameter as

$$\begin{aligned}\mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_{i+1} - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_i\right] &= \mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_i - \tau\left(\widehat{\boldsymbol{\theta}}_i^T \mathbf{x}_i - y_i\right)\mathbf{x}_i - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_i\right]\\ &= \mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_i\right] - \tau T(\mathbf{x}_i, y_i, \boldsymbol{\mu}_i, \mathbf{C}_i),\end{aligned}$$

where $H_i := \{\mathbf{p}_0, (\mathbf{x}_t, y_t)_{t=1}^i\}$ refers to the history of past examples and labels, as well as the prior distribution of $\widehat{\boldsymbol{\theta}}_0$ known by the teacher. We set $\boldsymbol{\mu}_i := \mathbb{E}[\widehat{\boldsymbol{\theta}}_i \mid H_i]$ and $\mathbf{C}_i := \mathbb{E}[\widehat{\boldsymbol{\theta}}_i \widehat{\boldsymbol{\theta}}_i^T - \boldsymbol{\mu}_i \boldsymbol{\mu}_i^T \mid H_i]$ to represent the expectation and covariance matrix of the learner state, respectively. We let $T(\mathbf{x}_i, y_i, \boldsymbol{\mu}_i, \mathbf{C}_i) = \mathbb{E}[2(\widehat{\boldsymbol{\theta}}_i^T \mathbf{x}_i - y_i)\langle \widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}, \mathbf{x}_i\rangle \; - \tau(\widehat{\boldsymbol{\theta}}_i^T \mathbf{x}_i - y_i)^2 \|\mathbf{x}_i\|_2^2 \mid H_i]$ represent the expected improvement, i.e., how much the teacher expects the MSE to reduce from time-step i to $i+1$.

The proposed policy selects the example-label pair that most reduces the error from one step to the next. Specifically, at time i, the teacher selects

$$(\widehat{\mathbf{x}}_i, \widehat{y}_i) = \underset{\mathbf{x}\in\mathcal{X}, y\in\mathcal{Y}}{\operatorname{argmax}} \quad T(\mathbf{x}, y, \boldsymbol{\mu}_i, \mathbf{C}_i). \tag{3.1}$$

Lemma 3.1. *The objective function in* (3.1) *is equivalent to*

$$T(\mathbf{x}, y, \boldsymbol{\mu}_i, \mathbf{C}_i) = \underbrace{\left(2 - \tau\|\mathbf{x}\|_2^2\right)\mathbf{x}^T\mathbf{C}_i\mathbf{x}}_{exploration} + \underbrace{2\left(\boldsymbol{\theta}^T - \boldsymbol{\mu}_i^T\right)\left(y - \boldsymbol{\mu}_i^T\mathbf{x}\right)\mathbf{x}}_{exploitation} \underbrace{-\tau\|\mathbf{x}\|_2^2\left(y - \boldsymbol{\mu}_i^T\mathbf{x}\right)^2}_{regularization}. \tag{3.2}$$

Lemma 3.1 follows from algebraic manipulations that are detailed in section SM1.1 of the supplementary materials, linked from the main article webpage. Note that T is a fourth degree polynomial with d unknowns: $y \in \{-1, 1\}$ and all but the first coordinate of $\mathbf{x}$. The unconstrained absolute maximum of T may be calculated with standard software such as the *fmincon* function of MATLAB. Additionally, the teacher does not need to track the probability distribution of the learner. The teacher only needs to track the first and second order moments to compute (3.1) and select the appropriate example.

The maximization of (3.2) implicitly accounts for the trade-off between estimating the learner state and teaching the ground truth to the learner. Under high uncertainty, corresponding to large values in the covariance $\mathbf{C}_i$, the first term in (3.2) dominates. The first term is an *exploration* component that promotes examples aligned with the direction of highest covariance, i.e., the examples that are most likely to decrease the teacher uncertainty about the learner state. On the other hand, the second term promotes examples that steer the estimated learner towards the ground truth, so the second term may be interpreted as an *exploitation* component. As the distance between the estimated learner state and the ground truth decreases, so does the relative weight of the exploitation term. The transition between phases focused on exploitation and exploration is further analyzed in section SM3.1, which examines the evolution of different sources of error. Last, the third term in (3.2) acts as a regularizer that discourages the norm of the gradient from being too large. This *regularization* term avoids abrupt and overly large updates in the learner state.

After sending the example and label pair to the learner, the teacher updates its estimation of the learner state following the known dynamical model of the learner. The mean and covariance are updated as

$$\begin{aligned}
\boldsymbol{\mu}_{i+1} &:= \mathbb{E}\left[\widehat{\boldsymbol{\theta}}_{i+1} \;\middle|\; H_{i+1}\right] = \mathbb{E}\left[\widehat{\boldsymbol{\theta}}_i - \tau\left(\widehat{\boldsymbol{\theta}}_i^T\mathbf{x}_i - y_i\right)\mathbf{x}_i \;\middle|\; H_i\right] = \boldsymbol{\mu}_i - \tau\left(\boldsymbol{\mu}_i^T\mathbf{x}_i - y_i\right)\mathbf{x}_i,\\
\mathbf{C}_{i+1} &:= \mathbb{E}\left[\widehat{\boldsymbol{\theta}}_{i+1}\widehat{\boldsymbol{\theta}}_{i+1}^T - \boldsymbol{\mu}_{i+1}\boldsymbol{\mu}_{i+1}^T \;\middle|\; H_{i+1}\right] = \mathbb{E}\left[\widehat{\boldsymbol{\theta}}_{i+1}\widehat{\boldsymbol{\theta}}_{i+1}^T - \boldsymbol{\mu}_{i+1}\boldsymbol{\mu}_{i+1}^T \;\middle|\; H_i\right]\\
&= \mathbb{E}\left[\left(\widehat{\boldsymbol{\theta}}_i - \tau\left(\widehat{\boldsymbol{\theta}}_i^T\mathbf{x}_i - y_i\right)\mathbf{x}_i\right)\left(\widehat{\boldsymbol{\theta}}_i - \tau\left(\widehat{\boldsymbol{\theta}}_i^T\mathbf{x}_i - y_i\right)\mathbf{x}_i\right)^T \;\middle|\; H_i\right]\\
&\quad - \mathbb{E}\left[\left(\boldsymbol{\mu}_i - \tau\left(\boldsymbol{\mu}_i^T\mathbf{x}_i - y_i\right)\mathbf{x}_i\right)\left(\boldsymbol{\mu}_i - \tau\left(\boldsymbol{\mu}_i^T\mathbf{x}_i - y_i\right)\mathbf{x}_i\right)^T \;\middle|\; H_i\right]\\
&= \mathbf{C}_i - \tau\mathbf{C}_i\mathbf{x}_i\mathbf{x}_i^T - \tau\mathbf{x}_i\mathbf{x}_i^T\mathbf{C}_i + \tau^2\mathbf{x}_i^T\mathbf{C}_i\mathbf{x}_i\mathbf{x}_i\mathbf{x}_i^T.
\end{aligned}$$

The first equality holds because, given the past history H_i, the teacher selects the next example-label pair in a deterministic way: in the absence of feedback, H_{i+1} is completely determined by H_i. We outline this approach, which we call Simultaneous Machine Teaching and Learning (SMTL), in Algorithm 3.1.

Algorithm 3.1. SMTL.

1: $\boldsymbol{\mu}_0, \mathbf{C}_0 \leftarrow p_0$
2: **for** $i = 0, 1, 2, \ldots$ **do**
3: Select example:
$(\mathbf{x}_i, y_i) \leftarrow \underset{\mathbf{x}\in\mathcal{X}, y\in\mathcal{Y}}{\operatorname{argmax}} \quad T(\mathbf{x}, y, \boldsymbol{\mu}_i, \mathbf{C}_i)$
4: Update estimations about learner:
$\boldsymbol{\mu}_{i+1} \leftarrow \boldsymbol{\mu}_i - \tau(\boldsymbol{\mu}_i^T \mathbf{x}_i - y_i)\mathbf{x}_i$
$\mathbf{C}_{i+1} \leftarrow \mathbf{C}_i - \tau \mathbf{C}_i \mathbf{x}_i \mathbf{x}_i^T - \tau \mathbf{x}_i \mathbf{x}_i^T \mathbf{C}_i + \tau^2 \mathbf{x}_i^T \mathbf{C}_i \mathbf{x}_i \mathbf{x}_i \mathbf{x}_i^T$
5: **end for**

Next, we characterize the convergence rate that SMTL provides. We recall the guarantees for omniscient teaching as a baseline against the proposed algorithm.

Theorem 3.2 (adapted from [20, Theorem 4]). *Consider a synthesis based omniscient teacher and a learner with updates given by* (2.1). *If* $\forall \widehat{\boldsymbol{\theta}}_i, \exists \gamma \in \mathbb{R}$ *with* $|\gamma| \leq \frac{\sqrt{P}}{\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2}$, $\nu(\gamma) \in \mathbb{R}$, *and* $y' \in \{-1, 1\}$ *s.t.* $0 < \tau(\widehat{\boldsymbol{\theta}}_i^T \mathbf{x}' - y')\mathbf{x}' \leq \nu(\gamma) < \frac{1}{\tau}$ *for* $\mathbf{x}' = \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})$, *then*

$$\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2 \leq (1 - \tau\nu)^{2i} \|\widehat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}\|_2^2.$$

Theorem 3.2 applies to the specific case of our framework in which $\forall i$ $\boldsymbol{\mu}_i = \widehat{\boldsymbol{\theta}}_i$ and $\mathbf{C}_i = \mathbf{0}$. The theorem guarantees that an omniscient teacher teaches a classifier to a gradient descent learner exponentially fast with the number of examples, thereby offering a significant improvement compared to the linear convergence obtained when randomly selecting examples [27]. The auxiliary variables γ and $\nu(\gamma)$ are related to the convergence speed. The guarantees for an omniscient teacher provide a baseline for the MSE of non-omniscient teachers. The following theorem offers a convergence rate guarantee in the non-omniscient scenario without feedback.

Theorem 3.3. *Consider a synthesis based teacher following SMTL and a learner with updates given by* (2.1). *If* $\forall \widehat{\boldsymbol{\theta}}_i, \exists \gamma \in \mathbb{R}$ *with* $|\gamma| \leq \frac{\sqrt{P}}{\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2}$, $\nu(\gamma) \in \mathbb{R}$, *and* $y' \in \{-1, 1\}$ *s.t.*

$$0 < \left(\widehat{\boldsymbol{\theta}}_i^T \gamma (\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - y' - \frac{1}{\tau\gamma} \right)^2 \leq \nu^2 < \frac{1}{\tau^2\gamma^2}, \tag{3.3}$$

then

$$\mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_{i-1} \right] \leq (\tau\gamma\nu)^{2i} \mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_0 \right].$$

Proof. We base our proof on [20, Theorem 4]. The expected evolution of the MSE from iteration i to iteration $i + 1$ is described by

$$\mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_{i+1} - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_i \right] = \mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_i \right] - \tau T(\widehat{\mathbf{x}}_i, \widehat{y}_i, \boldsymbol{\mu}_i, \mathbf{C}_i), \tag{3.4}$$

where

$$T(\widehat{\mathbf{x}}_i, \widehat{y}_i, \boldsymbol{\mu}_i, \mathbf{C}_i) = \mathbb{E}\left[2\left(\widehat{\boldsymbol{\theta}}_i^T \widehat{\mathbf{x}}_i - \widehat{y}_i\right)\langle \widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}, \widehat{\mathbf{x}}_i\rangle \; - \tau\left(\widehat{\boldsymbol{\theta}}_i^T \widehat{\mathbf{x}}_i - \widehat{y}_i\right)^2 \|\widehat{\mathbf{x}}_i\|_2^2 \;\middle|\; H_i\right]$$

represents the expected MSE improvement at the ith iteration when selecting the example-label pair $(\widehat{\mathbf{x}}_i, \widehat{y}_i)$. We analyze the objective function T at $(\mathbf{x}' = \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}), y')$, for some auxiliary parameter $\gamma \in \mathbb{R}$, to obtain the following lower bound:

(3.5)

$$\begin{aligned}
&T\left(\mathbf{x}', y', \boldsymbol{\mu}_i, \mathbf{C}_i\right) \\
&\quad = \mathbb{E}\left[2\left(\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - y'\right)\langle \widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}, \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\rangle - \tau\left(\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - y'\right)^2 \|\gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\|_2^2 \;\middle|\; H_i\right] \\
&\quad = \gamma\mathbb{E}\left[\|(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\|_2^2\left(2\left(\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - y'\right) - \tau\gamma\left(\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - y'\right)^2\right) \;\middle|\; H_i\right] \\
&\quad = \tau\gamma^2\mathbb{E}\left[\|(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\|_2^2\left(\frac{1}{\tau^2\gamma^2} - \left(\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - y' - \frac{1}{\tau\gamma}\right)^2\right) \;\middle|\; H_i\right] \\
&\quad = \frac{1}{\tau}\mathbb{E}\left[\|(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\|_2^2 \;\middle|\; H_i\right] - \tau\gamma^2\mathbb{E}\left[\|(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\|_2^2\left(\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - y' - \frac{1}{\tau\gamma}\right)^2 \;\middle|\; H_i\right] \\
&\quad \geq \frac{1}{\tau}\mathbb{E}\left[\|(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\|_2^2 \;\middle|\; H_i\right](1 - (\tau\gamma\nu)^2),
\end{aligned}$$

where the last inequality holds because of assumption (3.3).

The teacher selects the example-label pair that maximizes the expected improvement in MSE. By the definition of argmax in (3.1), $T(\widehat{\mathbf{x}}_i, \widehat{y}_i, \boldsymbol{\mu}_i, \mathbf{C}_i) \geq T(\mathbf{x}', y', \boldsymbol{\mu}_i, \mathbf{C}_i)\ \forall \mathbf{x}' \in \mathcal{X}, \forall y' \in \mathcal{Y}$. Combining this inequality with (3.5) and (3.4), we obtain

$$\begin{aligned}
\mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_{i+1} - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_i\right] &\leq \mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_{i-1}\right] - \tau T(\mathbf{x}', y', \boldsymbol{\mu}_i, \mathbf{C}_i) \\
&\leq \mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_{i-1}\right] - \tau\frac{1}{\tau}\mathbb{E}\left[\|(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})\|_2^2 \;\middle|\; H_{i-1}\right](1 - (\tau\gamma\nu)^2) \\
&\leq (\tau\gamma\nu)^2\mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_{i-1}\right] \\
&\leq (\tau\gamma\nu)^{2(i+1)}\mathbb{E}\left[\|\widehat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}\|_2^2 \;\middle|\; H_0\right],
\end{aligned}$$

where the shifts in the history index hold because, without feedback, the example selection criteria are deterministic given the prior distribution of the learner $\mathbf{p}_0 = H_0$. ■

Corollary 3.4. *Let the learning rate be* $0 < \tau < \frac{2P}{3}$. *Any learner with updates given by* (2.1) *converges exponentially with the number of examples when taught by a synthesis based teacher following the SMTL algorithm.*

Proof. To guarantee exponential convergence, it is sufficient to show that Theorem 3.3 is applicable for a learning rate $\tau \in (0, \frac{2P}{3})$, i.e., that assumption (3.3) holds. The following three inequalities are sufficient conditions for assumption (3.3) to hold:

$$\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) > y, \tag{3.6}$$

$$-\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) + \frac{2}{\tau\gamma} > -y, \tag{3.7}$$

$$\gamma^2 \widehat{\boldsymbol{\theta}}_i^T (\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) - \gamma y - \frac{1}{\tau} \neq 0. \tag{3.8}$$

Recall that $\max\{\|\boldsymbol{\theta}\|_2^2, \max_i \|\widehat{\boldsymbol{\theta}}_i\|_2^2\} \leq P$. Selecting $y' = -1$ and $0 < \gamma < \min\{\frac{1}{P}, \frac{\sqrt{P}}{\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2}\}$, we show that all requirements (3.6–3.8) hold.

We fulfill (3.6) because

$$\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) = \gamma \|\widehat{\boldsymbol{\theta}}_i\|_2 \left(\|\widehat{\boldsymbol{\theta}}_i\|_2 - \|\boldsymbol{\theta}\|_2 \cos\left(\angle \widehat{\boldsymbol{\theta}}_i, \boldsymbol{\theta}\right)\right) > -\gamma P > -1 = y,$$

where the operator $\angle \cdot, \cdot$ refers to the angle between two vectors. Next, we note that

$$-\widehat{\boldsymbol{\theta}}_i^T \gamma(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}) + \frac{2}{\tau\gamma} > -2\gamma P + \frac{2}{\tau\gamma} > -2 + \frac{2P}{\tau}. \tag{3.9}$$

As we restrict the step-size, $\tau \in (0, \frac{2P}{3})$, we may further lower bound (3.9) as

$$-2 + \frac{2P}{\tau} > -2 + 3 = 1 = -y,$$

so (3.7) is also fulfilled.

The left-hand side in (3.8) is a nondegenerate quadratic equation with respect to γ, with at most two roots. As the interval $(0, \frac{1}{P})$ is continuous, it must contain nonroot values, so there must exist a $\gamma \in (0, \min\{\frac{1}{P}, \frac{\sqrt{P}}{\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2}\})$ for which (3.8) also holds. Since (3.3) holds, we may directly apply Theorem 3.3 to conclude the proof. ∎

Theorem 3.3 shows that SMTL achieves an exponential behavior similar to omniscient teaching. To guarantee the desired exponential convergence of the learner to the ground truth with respect to the number of examples, we require $\mathbb{E}[\|\widehat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}\|_2^2 \mid H_0] < \infty$. This requirement is a characteristic of most machine learning models, as in general, the starting point of learning algorithms is bounded. A sufficient condition for this to hold in our system is $P < \infty$. In addition, Corollary 3.4 asserts that the assumptions of Theorem 3.3 are fulfilled as long as the learning rate is not too large.

Following SMTL, a gradient descent learner described by (2.1) needs $\mathcal{O}(\log \frac{1}{\epsilon} \mathbb{E}[\|\widehat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}\|_2^2])$ example-label pairs to learn an ϵ-approximation of the ground truth model. This convergence rate is of the same order as the one achieved by omniscient teaching, while relaxing the assumption about knowledge of the exact learner initialization.

The performance guarantees also to hold in the case of rescalable pool based teaching with a rich enough example set. The approach and its analysis are detailed in section SM2. Additionally, Lemma 3.5 below extends the problem to settings in which the example space of the learner suffers an unknown orthonormal transformation with respect to the example space of the teacher.

Lemma 3.5. *Let $\mathcal{G}$ be an unknown orthonormal transformation describing the mapping from the feature space of the teacher to the learner. For every example $\widetilde{\mathbf{x}}$ selected by the teacher* *according to SMTL, the learner observes* $\widehat{\mathbf{x}} = \mathcal{G}(\widetilde{\mathbf{x}})$ *and updates its state according to* (2.1)*. If* $\forall \widetilde{\boldsymbol{\theta}}_i, \exists \gamma \in \mathbb{R}$ *with* $|\gamma| \le \frac{\sqrt{P}}{\|\widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}\|_2}$, $\nu(\gamma) \in \mathbb{R}$, *and* $y' \in \{-1, 1\}$ *s.t.* $0 < (\widetilde{\boldsymbol{\theta}}_i^T \gamma(\widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}) - y' - \frac{1}{\tau\gamma})^2 \le \nu^2 < \frac{1}{\tau^2\gamma^2}$, *then*

$$\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_{i-1}\right] \le (\tau\gamma\nu)^{2(i+1)} \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_0 - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_0\right],$$

where $\widetilde{\boldsymbol{\theta}}_i = \mathcal{G}^T(\widehat{\boldsymbol{\theta}}_i)$ *and* $\widetilde{\boldsymbol{\theta}}$ *represent the learner state and ground truth, respectively, in the teacher feature space.*

Proof. Let $\mathcal{G}$ be an orthonormal transformation from the teacher feature space, whose elements are identified by $\widetilde{\cdot}$, to the learner feature space, whose elements are identified by $\widehat{\cdot}$. Let $\mathcal{G}^T$ denote the inverse mapping from the learner to the teacher feature space. By the definition of an orthonormal transformation, $\mathcal{G}$ preserves the inner product, i.e., $\langle \widehat{\boldsymbol{\theta}}_i, \widehat{\mathbf{x}} \rangle = \langle \widetilde{\boldsymbol{\theta}}_i, \widetilde{\mathbf{x}} \rangle$. Thus, we write the learner updates from iteration i to $i+1$ as

$$\widehat{\boldsymbol{\theta}}_{i+1} = \widehat{\boldsymbol{\theta}}_i - \tau\left(\widehat{\boldsymbol{\theta}}_i^T \widehat{\mathbf{x}}_i - y_i\right)\widehat{\mathbf{x}}_i = \widehat{\boldsymbol{\theta}}_i - \tau\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}}_i - y_i\right)\mathcal{G}\left(\widetilde{\mathbf{x}}_i\right).$$

The error metric is given by the expected squared distance between the ground truth $\widetilde{\boldsymbol{\theta}}$ and the teacher's estimation about the learner state $\widetilde{\boldsymbol{\theta}}_{i+1}$ in the teacher feature space. As the mapping is invertible $\mathcal{G}^T(\mathcal{G}(\mathbf{x})) = \mathbf{x}$, we may decompose the MSE as

$$\begin{aligned}
\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_{i+1} - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_i\right] &= \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_i - \tau\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}}_i - y_i\right)\mathcal{G}^T\mathcal{G}(\widetilde{\mathbf{x}}_i) - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_i\right] \\
&= \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}} - \tau\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}}_i - y_i\right)\widetilde{\mathbf{x}}_i\right\|_2^2 \;\middle|\; H_i\right] \\
&= \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_i\right] \\
&\quad + \mathbb{E}\left[-2\left\langle \widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}, \tau\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}}_i - y_i\right)\widetilde{\mathbf{x}}_i\right\rangle + \left\|\tau\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}}_i - y_i\right)\widetilde{\mathbf{x}}_i\right\|_2^2 \;\middle|\; H_i\right] \\
&= \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_i\right] \\
&\quad - \tau\mathbb{E}\left[2\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}}_i - y_i\right)\langle \widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}, \widetilde{\mathbf{x}}_i\rangle - \tau\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}}_i - y_i\right)^2 \|\widetilde{\mathbf{x}}_i\|_2^2 \;\middle|\; H_i\right].
\end{aligned} \tag{3.10}$$

The SMTL algorithm selects the example-label pair in the teacher feature space as

$$(\widetilde{\mathbf{x}}_i, y_i) = \underset{\mathbf{x}\in\mathcal{X}, y\in\mathcal{Y}}{\operatorname{argmax}} \;\; \mathbb{E}\left[2\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}} - y\right)\langle \widetilde{\boldsymbol{\theta}}_i - \widetilde{\boldsymbol{\theta}}, \widetilde{\mathbf{x}}\rangle - \tau\left(\widetilde{\boldsymbol{\theta}}_i^T \widetilde{\mathbf{x}} - y\right)^2 \|\widetilde{\mathbf{x}}\|_2^2 \;\middle|\; H_i\right],$$

such that the MSE is greedily minimized. This is equivalent to the teacher's behavior when the teacher and the learner share the same feature space. Therefore, we apply the inequality (3.5) to upper-bound (3.10) as

$$\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_{i+1} - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_i\right] \le (\tau\gamma\nu)^{2(i+1)} \mathbb{E}\left[\left\|\widetilde{\boldsymbol{\theta}}_0 - \widetilde{\boldsymbol{\theta}}\right\|_2^2 \;\middle|\; H_0\right]. \qquad \blacksquare$$

Lemma 3.5 shows that SMTL is invariant to rotations and reflections. Simultaneously teaching and learning provides an exponential speedup even when the learner and teacher do not share a representation space, but there exists an unknown orthonormal transformation between the teacher and learner feature spaces. This result extends the applicability of SMTL to various real-world problems, such as the cross-lingual sentiment analysis discussed in section 4.2.

3.2. Simultaneous Machine Teaching and Learning with Noisy Feedback (SMTL-F). We now analyze the situation in which the teacher receives some feedback from the learner. Without knowledge of the exact learner state, previous approaches [21] propose a dedicated probing phase in which the teacher exploits the feedback to obtain an accurate estimation of the learner state and then allowing the teacher to proceed as if it were omniscient. We show that the teacher may instead simultaneously learn the learner state and teach the ground truth to the learner, thereby avoiding an explicit probing phase that improves the learner's estimate without teaching.

For analytical tractability, we consider the case in which the feedback from the learner is given by

$$s_i = \widehat{\boldsymbol{\theta}}_{i+1}^T \mathbf{x}_i + w_i, \tag{3.11}$$

where $w_i \sim \mathcal{N}(0, \sigma^2)$ represents some random noise that accounts for imperfections in the communication channel between learner and teacher. The feedback is a noisy measurement of the learner certainty regarding the latest example classification. Specifically, the learner returns a noisy function of the distance and direction from the latest example to its current classifier. A large positive value of s_i suggests that the learner probably classifies the latest example $\mathbf{x}_i$ as class 1. Similarly, a large negative value of s_i suggests that a classification of $\mathbf{x}_i$ in class –1 is more probable. On the other hand, a value of s_i around 0 suggests that the example lies close to the learner classification boundary. Note that recovering the high-dimensional true parameter $\widehat{\boldsymbol{\theta}}_{i+1} \in \mathbb{R}^d$ from this noisy scalar $s_i \in \mathbb{R}$ is not straightforward.

At each time step, the teacher has access to two sources of information about the learner state. First, the teacher directly observes the noisy feedback. Second, the teacher knows the dynamical model of the learner and may predict its future state based on its current estimate. Kalman filtering is a well-known approach to optimally leverage these two sources of information.

The proposed Simultaneous Machine Teaching and Learning algorithm with noisy Feedback (SMTL-F) is summarized in Algorithm 3.2. The teacher interleaves the greedy example selection strategy given by (3.1), with a Kalman filter to achieve optimal tracking. Lines 4, 5, and 6 of Algorithm 3.2 outline the computations required to track the mean and covariance of the learner state.

Theorem 3.6. *Consider a learner that updates according to* (2.1) *and provides some feedback according to* (3.11)*. The estimator in SMTL-F then is the* optimal estimator*. Additionally, when* $\tau \le \frac{2}{P_{\mathbf{x}}}$*, the covariance of the teacher estimation about the learner state is monotonically nonincreasing:*

$$\|\mathbf{C_{i+1}}\|_\infty \le \|\mathbf{C_i}\|_\infty \,,$$

where $\|\mathbf{C}\|_\infty = \lim_{k\to\infty} \|\mathbf{C}^k\|^{1/k}$.

Algorithm 3.2. SMTL-F.

1: $\boldsymbol{\mu}_0, \mathbf{C}_0 \leftarrow p_0$.
2: **for** $i = 0, 1, 2, \ldots$ **do**
3: Select example:
$$(\mathbf{x}_i, y_i) \leftarrow \underset{\mathbf{x}\in\mathcal{X}, y\in\mathcal{Y}}{\operatorname{argmax}} \quad T(\mathbf{x}, y, \boldsymbol{\mu}_i, \mathbf{C}_i)$$
4: Estimator - Predict:
$$\boldsymbol{\mu}_{i+1|i} \leftarrow \boldsymbol{\mu}_i - \tau(\boldsymbol{\mu}_i^T \mathbf{x}_i - y_i)\mathbf{x}_i$$
$$\mathbf{C}_{i+1|i} \leftarrow \mathbf{C}_i - \tau \mathbf{C}_i \mathbf{x}_i \mathbf{x}_i^T - \tau \mathbf{x}_i \mathbf{x}_i^T \mathbf{C}_i + \tau^2 \mathbf{x}_i^T \mathbf{C}_i \mathbf{x}_i \mathbf{x}_i \mathbf{x}_i^T$$
5: Estimator - Observe feedback:
$$s_i \leftarrow \widehat{\boldsymbol{\theta}}_{i+1}^T \mathbf{x}_i + w_i$$
6: Estimator - Update estimation:
$$\mathbf{K}_{i+1} \leftarrow \mathbf{C}_{i+1|i} \mathbf{x}_i (\mathbf{x}_i^T \mathbf{C}_{i+1|i} \mathbf{x}_i + \sigma^2)^{-1}$$
$$\boldsymbol{\mu}_{i+1} \leftarrow \boldsymbol{\mu}_{i+1|i} + \mathbf{K}_{i+1}(s_i - \boldsymbol{\mu}_{i+1|i}^T \mathbf{x}_i)$$
$$\mathbf{C}_{i+1} \leftarrow (\mathbf{I} - \mathbf{K}_{i+1}\mathbf{x}_i^T)\mathbf{C}_{i+1|i}(\mathbf{I} - \mathbf{K}_{i+1}\mathbf{x}_i^T)^T + \sigma^2 \mathbf{K}_{i+1}\mathbf{K}_{i+1}^T$$
7: **end for**

Proof. Our connection to control theory enables us to employ established control-theoretic tools to prove Theorem 3.6. Specifically, we utilize the Bayesian optimality of the Kalman filter [24] and its closed-form solution for Gauss–Markov models [17] to design the optimal estimator in SMTL-F. Additionally, we derive stability guarantees by analyzing the Riccati equations, a standard approach in discrete-time linear-quadratic-Gaussian control. While prior work [19] already noted the connection between machine teaching and control theory, the focus was on omniscient teachers and long-term planning. Our contribution differs by addressing online, interactive teaching scenarios in which the teacher operates under uncertainty about the learner. Our approach leverages control theory to incorporate new information dynamically and optimally. Specifically, to prove Theorem 3.6, we must prove that the learner state estimator in SMTL-F is both optimal in the Bayesian sense and that it exhibits stable behavior with monotonically nonincreasing covariance.

Optimality of the estimator in SMTL-F. We define the system state $Z_i = \{\mathbf{x}_i, y_i, \boldsymbol{\mu}_{i+1}, \mathbf{C}_{i+1}, \widehat{\boldsymbol{\theta}}_{i+1}\}$. The functional dependence graph in Figure 2 shows that the state Z_i d-separates [5, Definition 2.14] the latest feedback s_i from the ground truth, past learner states, and past feedback. Therefore, the current feedback is conditionally independent of the history given the system state

$$\begin{aligned}\mathbb{P}\left[s_i, \boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}_0, \boldsymbol{\mu}_0, \mathbf{C}_0, \{Z_t\}_{t=0}^{i-1} \,\middle|\, Z_i\right] &= \mathbb{P}\left[s_i \,\middle|\, Z_i\right] \mathbb{P}\left[\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}_0, \boldsymbol{\mu}_0, \mathbf{C}_0, \{Z_t\}_{t=0}^{i-1} \,\middle|\, Z_i\right] \\ &= \mathbb{P}\left[s_i \,\middle|\, Z_i\right] \mathbb{P}\left[H_{i-1} \,\middle|\, Z_i\right].\end{aligned}$$

Figure 2 also shows that Z_i d-separates Z_{i-1} from Z_{i+1}, and therefore the state is Markovian: $Z_{i-1} \to Z_i \to Z_{i+1}$. We also observe that any state is independent of past feedback given the previous state, so that

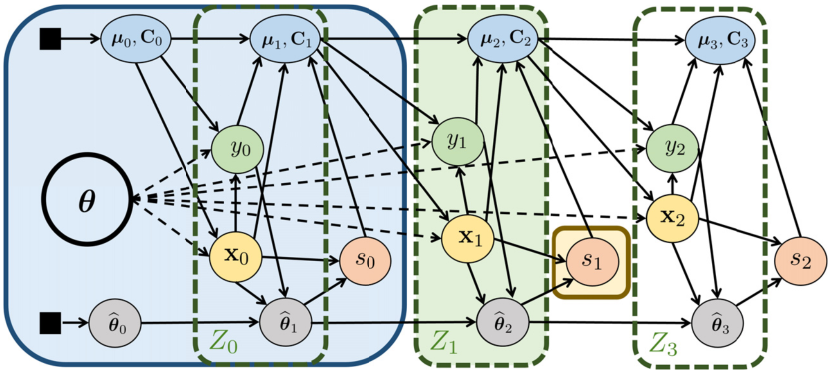


Figure 2. *Functional dependence graph showing causal relationships between the teacher estimators about the learner $\boldsymbol{\mu}, \mathbf{C}$, the true learner state $\widehat{\boldsymbol{\theta}}$, the ground truth $\boldsymbol{\theta}$, the example-label pairs $\{\mathbf{x}, y\}$, and the feedback s. We observe that the system state $Z_i = \{\mathbf{x}_i, y_i, \boldsymbol{\mu}_{i+1}, \mathbf{C}_{i+1}, \widehat{\boldsymbol{\theta}}_{i+1}\}$ is Markovian and that the feedback is conditionally independent of the past given the current state.*

$$\mathbb{P}[Z_{i+1} \;\Big|\; \{Z_t, s_t\}_{t=0}^{i}] = \mathbb{P}[Z_{i+1} \;\Big|\; Z_i].$$

Combining the conditional independence with the fact that both learner state and feedback are Gaussian random variables shows that the system follows a Gauss–Markov model. Consequently, the Kalman filter is the Bayesian optimal filter [24]. Moreover, the distributions are jointly Gaussian, so we only need to keep track of the mean and covariance matrices to obtain the optimal estimator of the learner state. SMTL-F implements the known closed-form solution of the Kalman filter for Gauss–Markov models [17]. Hence, SMTL-F obtains the optimal posterior probability density function of the learner state in a tractable way.

Stability of the estimator in SMTL-F. Next, we show that the estimation of the learner state derived by SMTL-F is stable, in the sense that the uncertainty about the learner state is monotonically nonincreasing. The detailed proofs of all auxiliary lemmas are in section SM1. We start by deriving the discrete-time algebraic Riccati recursion of the following system.

Lemma 3.7. *The dynamic Riccati equation describing the evolution of the teacher's covariance about the learner state is given by*

$$\mathbf{C_{i+1}} = \mathbf{F}_i \mathbf{C}_i \mathbf{F}_i \mathbf{T}_i, \tag{3.12}$$

where $\mathbf{F}_i = \mathbf{I} - \mathbf{x}_i \mathbf{x}_i^T$ is the Hermitian state transition matrix at the ith iteration and $\mathbf{T}_i = \mathbf{I} - (\mathbf{x}_i^T \mathbf{F}_i \mathbf{C}_i \mathbf{F}_i \mathbf{x}_i + \sigma^2)^{-1} \mathbf{x}_i \mathbf{x}_i^T \mathbf{F}_i \mathbf{C}_i \mathbf{F}_i$, where $\mathbf{I}$ represents the identity matrix.

As a stepping stone towards proving the stability of SMTL-F, we analyze the spectral radius of the factors in the Riccati equation (3.12).

Lemma 3.8. *The spectral radius of $\mathbf{F}_i$ is 1 for $\tau \leq \frac{2}{P_{\mathbf{x}}}$.*

Lemma 3.9. *The spectral radius of $\mathbf{T}_i$ is 1.*

Last, we take the submutiplicative matrix norm $\|\cdot\|_\infty := \lim_{k\to\infty} \|\cdot^k\|^{1/k}$ on both sides of the Riccati recursion (3.12),

$$\|\mathbf{C_{i+1}}\|_\infty \leq \|\mathbf{C_i}\|_\infty \|\mathbf{F}_i\|_\infty^2 \|\mathbf{T}_i\|_\infty. \tag{3.13}$$

Gelfand's formula guarantees that $\rho(\mathbf{A}) = \|\mathbf{A}\|_\infty$ [12], where the operator $\rho(\cdot)$ represent the spectral radius of a matrix. Applying this result together with Lemmas 3.8 and 3.9 to (3.13), we obtain

$$\|\mathbf{C_{i+1}}\|_\infty \leq \|\mathbf{C_i}\|_\infty \rho\left(\mathbf{F}_i\right)^2 \rho\left(\mathbf{T}_i\right) \leq \|\mathbf{C_i}\|_\infty,$$

which proves that $\|\mathbf{C}_i\|_\infty$ is monotonically nonincreasing. ■

In the presence of feedback, the estimation of the learner state derived by SMTL-F is both optimal (it achieves the smallest expected error) and stable (the uncertainty about the learner state is monotonically nonincreasing).

4. Empirical performance. We now analyze the empirical performance of the algorithms in a synthetic two-dimensional binary classification problem as well as in a real cross-lingual sentiment analysis problem. Additional experiments with images are evaluated in the supplementary materials (supplement.pdf [local/web 1.27MB]). The code with the algorithms to replicate the experiments is available online.[1]

4.1. Synthetic dataset. We first compare the performance of the SMTL and SMTL-F algorithms against the state-of- the-art online machine teaching methods with a synthetic dataset. We generate a standard two-dimensional binary dataset, shown in Figure 3, following the procedure outlined in [38].

We validate the proposed online algorithms against the baseline omniscient teaching algorithm. Figure 4(a) shows the evolution of the learner error $\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2$ as more examples are presented. We observe that the error decreases exponentially fast for both online algorithms as well as for the omniscient teacher, hence offering a significant improvement compared to the rate of traditional Stochastic Gradient Descent (SGD) in which examples are chosen randomly.

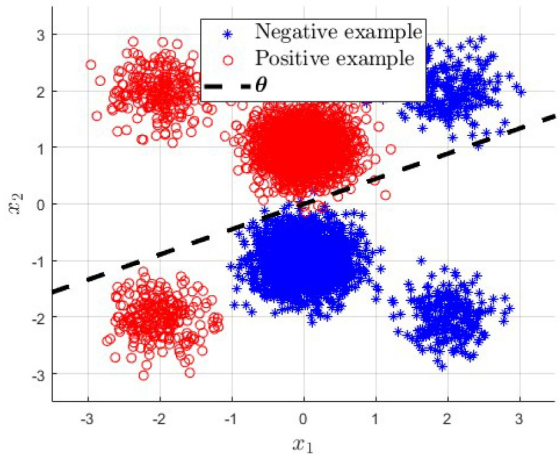
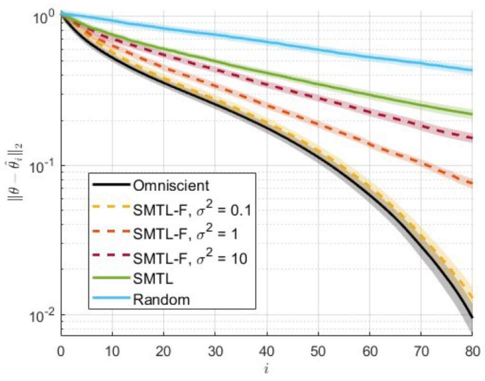
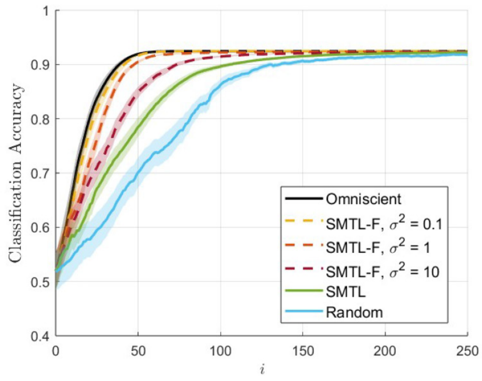


Figure 3. *Synthetic dataset synth2* [38].

[1]https://github.com/BelenMU/SMTL/tree/main



(a) MSE



(b) Accuracy

Figure 4. *Performance comparison between algorithms on the synthetic dataset. All online machine teaching algorithms achieve an exponential speedup w.r.t. randomly selecting examples. Within the exponential convergence of the MSE, the lower the noise level of the feedback the faster the MSE decreases and the classification accuracy increases.*

However, within the exponential rates, the omniscient teacher performs the best because it has the most information about the learner state.

In the presence of feedback, tracking the learner is a good strategy to bridge the gap in performance between the omniscient teacher and the no-feedback case. The MSE of SMTL-F is lower bounded by the MSE of omniscient teaching and upper bounded by the MSE of SMTL. As the feedback noise level decreases, SMTL-F approaches the omniscient teacher performance. In fact, as Figure 4(a) shows, under feedback with very low noise levels, SMTL-F rapidly achieves a precise estimation of the learner state, becoming a de facto omniscient teacher.

Although we use the squared distance between the learner and the ground truth as a performance metric, the ultimate objective is to achieve a good classification accuracy. Figure 5 shows how these two metrics are intertwined: a learner close to the ground truth, i.e., a low $\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2$, implies a good classification accuracy. The same relationship holds for different datasets, as analyzed in section SM3.2. This justifies a posteriori why the proposed online algorithms focus on nonincreasing $\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2$, as this is a good heuristic for classification accuracy improvement. The relationship between both metrics is highly nonlinear, meaning that an improvement on the learner state can strongly improve the classification accuracy when the state is far from the ground truth. Once the learner is sufficiently close to the ground truth, fine-tuning the learner's state yields a much less significant change in classification accuracy. This behavior highlights the benefits of SMTL-F: for sufficiently low noise levels on the feedback, teachers following SMTL-F are able to keep up with the omniscient teacher until a high enough accuracy is reached, at which point fine-tuning of the learner state no longer has a significant impact on classification accuracy.

The graphs in Figure 5 and Figure SM3 show that all algorithms exhibit similar relationships between MSE and classification accuracy. This behavior suggests that there are implicit *trajectories* that all the online machine teaching algorithms approximately follow and that the speed at which learners travel along the *trajectories,* measured in terms of number examples,

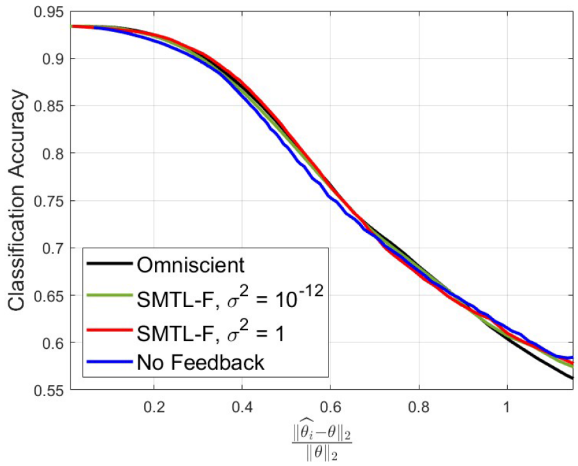


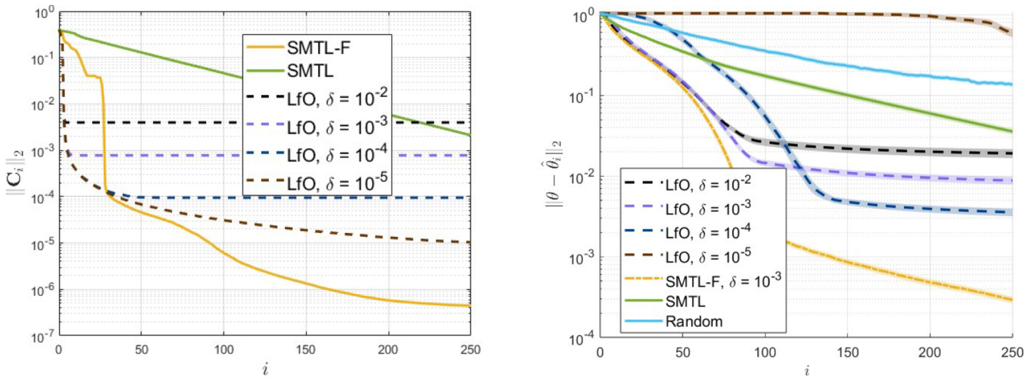
Figure 5. *Correspondence between classification accuracy and the learner's distance to the ground truth for different algorithms.*

strongly depends on the teacher's knowledge about the learner. Said differently, the feedback provided by the learner does not seem to provide advantages in terms of trajectory: it only seems to affect how fast the learner reaches a low $\|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2$ value.

Figure 4(b) summarizes the performance of the online algorithms, as measured by the classification accuracy. Machine teaching outperforms random example selection. With more information about the learner, the classification accuracy of the learner improves faster with respect to the number of examples.

We explore how learner initializations impact algorithm performance. We randomly initialize 50 learners and compare the resulting variation in performance. The shaded regions in Figure 4 represent the standard error between initializations. Notably, online machine teaching not only outperforms random example selection but also enhances robustness as SMTL and SMTL-F exhibit significantly lower variance. This finding suggests that the proposed algorithms offer more consistent and stable results under different starting conditions, making them a favorable choice for various applications. Online machine teaching mitigates the impact of learner initializations on performance, and this effect is further diminished as feedback noise decreases.

We further validate SMTL-F against the Learning for Omniscience (LfO) algorithm [21, 18]. There are two distinct phases of the LfO algorithm corresponding to the probing and teaching phases. At first, the teacher focuses solely on decreasing its uncertainty about the learner state, i.e., $(\widehat{\mathbf{x}}, \widehat{y}) = \operatorname{argmin}_{\mathbf{x}\in\mathcal{X}, y\in\mathcal{Y}} \quad \|\mathbf{C}_i\|_2$. As Figure 6(a) shows, at first teachers following LfO reduce their uncertainty about the learner much faster than SMTL and SMTL-F. However, getting an accurate estimation of the learner is done at the expense of teaching the ground truth. As Figure 6(b) shows, the MSE of the learner remains constant during the first iterations, as the learner does not update its state immediately [21]. The teaching phase of the LfO algorithm starts once the uncertainty about the learner state is sufficiently low, i.e. $\|\mathbf{C}_i\|_2 < \delta$, for a given threshold $\delta \in \mathbb{R}^+$. Then the teacher proceeds as if it were omniscient using its latest estimation.



(a) Uncertainty (b) MSE

Figure 6. *Performance of the proposed online machine teaching algorithms against the state of the art. SMTL-F outperforms LfO with* $\sigma^2 = 10^{-3}$ *by continuously updating its estimation about the learner while teaching, avoiding an explicit probing phase.*

Figure 6(b) shows the performance of SMTL-F against LfO when $\sigma^2 = 10^{-3}$. We observe that having separate learning and teaching phases negatively impacts the overall performance of the algorithm. If the probing phase is too short, the teacher does not have an accurate estimation of the learner, so it is not able to teach it efficiently and the error decreases much slower than with SMTL-F, which continuously improves its estimation of the learner. On the other hand, a longer probing phase leads to an accurate estimation of the learner state but requires many iterations without teaching in which the error does not decrease. In practice, LfO with a long probing phase, i.e., low δ, is unable to catch up with the online algorithm that has been teaching all along. The proposed algorithm with noisy feedback avoids the costly probing phase, while still obtaining an accurate and ever-improving estimation of the learner state.

These experiments confirm that jointly teaching the learner while estimating its parameters offers significant gains.

4.2. Cross-lingual sentiment analysis. Language can be harnessed to understand the attitude of individuals [30]. Towards this goal, binary sentiment word classification aims to accurately label words according to their connotation as positive (e.g., love) or negative (e.g., death). Traditionally, research on lingual sentiment analysis has focused on a few languages that have a large amount of annotated data [9]. To tackle this resource imbalance, cross-lingual adaptation [1, 15, 33] aims to transfer the knowledge of languages with plentiful resources to languages with few resources. In this section, we apply SMTL and SMTL-F to tackle the cross-lingual sentiment analysis problem. We assume that the teacher has access to a linear sentiment classifier in the word space created from a Spanish dictionary. The teacher aims to teach a learner working on the word space created from an Italian dictionary to accurately classify Italian words.

We use existing monolingual word embeddings[2] [2] and normalize each word vector. Previous work [37] empirically shows that the mapping of normalized word vectors between

[2] http://ixa2.si.ehu.es/martetxe/vecmap/es.emb.txt.gz

languages is accurately described by an orthonormal transformation. Hence, following Lemma 3.5, SMTL is suitable for cross-lingual knowledge transfer, even if the explicit mapping between the Spanish and Italian word embeddings is unknown.

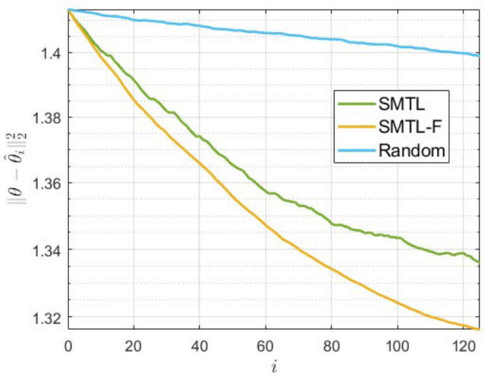
The teacher works in the Spanish word embedding. At each iteration, the teacher selects the example-label pair according to (3.1), where $\mathcal{X}$ is the set of embedded Spanish words. We limit the examples to a finite dataset by selecting the 10000 most common words. This extension of synthesis-based teaching to a pool-based setting is detailed in section SM2.1. We use Google Translate[3] to translate each example from Spanish to Italian. The learner only sees the embedding corresponding to the translated word in the Italian vector space. Figure 7(a) shows the evolution of the MSE when a teacher working in the Spanish word space teaches a word sentiment classifier to a learner in the Italian word space. Machine teaching decreases the error significantly faster than random example selection.

The performance further improves when the learner provides feedback about its state to the teacher. As orthonormal transformations preserve inner products, we follow the framework described in section 3.2. The feedback from the learner to the teacher is described as

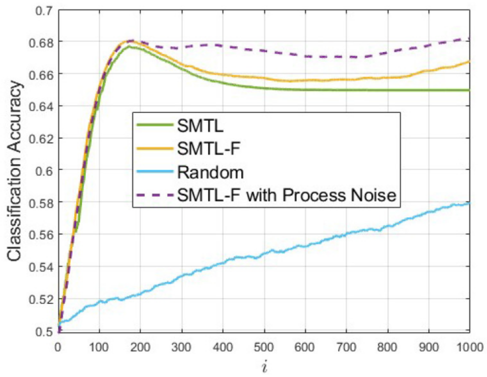
$$s_i = \widehat{\boldsymbol{\theta}}_{i+1}^T \mathcal{G}\left(\mathbf{x}_i\right) = \mathcal{G}^{-1}\left(\widehat{\boldsymbol{\theta}}_{i+1}\right)^T \mathbf{x}_i + w_i,$$

where $\mathcal{G}$ is the unknown orthonormal mapping of word embeddings from the teacher to the learner language space. As the real mapping is not exactly an orthonormal transformation, we introduce $w_i \sim \mathcal{N}(0, \sigma^2)$ to account for the deviations from the perfect orthonormality assumption.

We estimate the noise level σ^2 from the information exchanged between teacher and learner. The teacher samples N random pairs of words $(\widetilde{\mathbf{x}}_a, \widetilde{\mathbf{x}}_b)$, the learner observes the corresponding word pairs in the learner word space $(\widehat{\mathbf{x}}_a, \widehat{\mathbf{x}}_b)$, computes each pair's inner product,



(a) MSE



(b) Accuracy

Figure 7. *Performance on the cross-lingual sentiment analysis problem. Online machine teaching algorithms speed up the teaching. Adding process noise in the Kalman update reduces the drop in performance caused by nonorthogonalities in the mapping between Spanish and Italian words.*

[3] https://translate.google.com

and transmits the resulting products to the teacher. The teacher then calculates the differences in inner products between the pairs of words in the teacher language and the learner language. The variance among these differences becomes the estimator for σ^2,

$$\sigma^2 \approx \frac{1}{N}\sum_{n=1}^{N}\left(\widetilde{\mathbf{x}}_{a,n}^T\widetilde{\mathbf{x}}_{b,n} - \widehat{\mathbf{x}}_{a,n}^T\widehat{\mathbf{x}}_{b,n}\right)^2,$$

where $(\widetilde{\mathbf{x}}_{a,n}, \widetilde{\mathbf{x}}_{b,n})$ is the nth pair of words sampled by the teacher. As Figure 7(a) shows, incorporating learner feedback with this estimator further improves the rate at which the MSE decreases.

We also test the learner accuracy for classifying a pre-existing sentiment lexicon in Italian.[4] The results are shown in Figure 7(b). Online machine teaching algorithms are superior to random selection of examples. In fact, 50 examples selected by SMTL or SMTL-F achieve the same classification accuracy as 1000 randomly selected examples.

4.2.1. Deviations from orthogonal mappings. As the mapping between languages is not perfectly orthonormal, the teacher model of the learner dynamical system is slightly inaccurate. This could lead to instances in which the teacher is certain of its learner state estimation, but this estimation is inaccurate. This would explain the dip in accuracy observed in Figure 7(b). In this section, we further analyze this conjecture; i.e., we investigate how deviations from the orthonormality assumption in Lemma 3.5 affect the performance of SMTL-F. We also propose an extension of the algorithm to account for the deviations and diminish the performance dips they cause.

To empirically understand how SMTL-F performs under nonorthogonal transformations, we modify the synthetic experiments in section 4.1. We create a new learner example space by rotating each example

$$\widehat{\mathbf{x}}_i = \mathrm{Rotate}(\widetilde{\mathbf{x}}_i, \phi + z_i),$$

where the degrees of rotation $\phi + z_i$ are composed of a deterministic amount, unknown to the teacher, along with an additional random rotation. The deterministic rotation, denoted by ϕ, is sampled from a uniform distribution $\phi \sim \mathcal{U}(0, 2\pi)$ and remains constant for all examples. On the other hand, the random rotation z_i is sampled independently for each example from a Gaussian distribution $z_i \sim \mathcal{N}(0, z^2)$, which adds an extra random degree of rotation to each instance.

Figure 8 shows that as the examples deviate further from the perfect orthonormal transformation, a dip in accuracy appears. This behavior gives credence to our conjecture that the drop in performance in Figure 7(b) is caused by deviations from the assumption of orthogonal mapping between languages.

The SMTL-F algorithm assumes a perfect knowledge of the dynamical system of the learner. However, the examples from the teacher to the learner space do not always experience the same rotation, so, in practice, the teacher may not be able to exactly determine the evolution of the learner state. The teacher overcomes the estimation error when observing

[4] https://www.kaggle.com/datasets/rtatman/sentiment-lexicons-for-81-languages

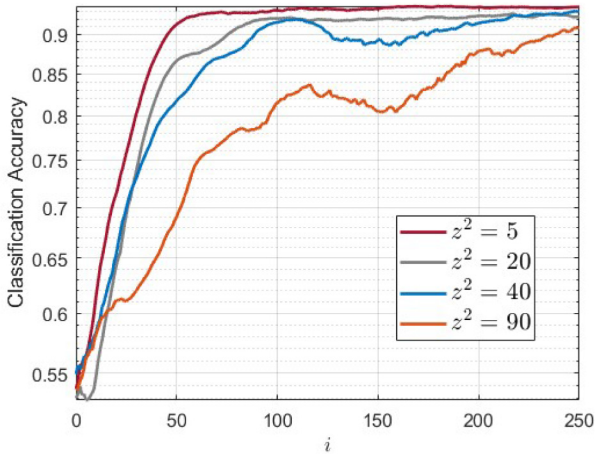


Figure 8. *Performance of SMTL-F for examples deviated by $\mathcal{N}(0, z^2)$ from perfect orthonormal mapping between teacher and learner feature spaces. Deviations lead to a performance dip.*

more feedback from the learner, which is consistent with previous works [23, 14] showing that interactivity mitigates the impact of imperfect knowledge and mismatches.

Another approach is to account for the mapping imperfections by introducing process noise in the dynamical model of the learner. Let $\mathbf{r}_i$ denote the difference between the teacher's example mapped in a perfectly orthogonal way $\mathcal{G}(\widetilde{\mathbf{x}})$ and the corresponding example in the learner space $\widehat{\mathbf{x}}$; i.e., $\mathbf{r}_i = \widehat{\mathbf{x}}_i - \mathcal{G}(\widetilde{\mathbf{x}}_i)$. Then the evolution of the learner state from iteration i to $i+1$ is given by

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}_{i+1} &= \widehat{\boldsymbol{\theta}}_i - \tau\left(\widehat{\boldsymbol{\theta}}_i^T \widehat{\mathbf{x}}_i - y_i\right)\widehat{\mathbf{x}}_i = \left(\mathbf{I} - \tau\widehat{\mathbf{x}}_i\widehat{\mathbf{x}}_i^T\right)\widehat{\boldsymbol{\theta}}_i + \tau y \widehat{\mathbf{x}}_i \\
&= \left(\mathbf{I} - \tau(\mathcal{G}(\widetilde{\mathbf{x}}_i) + \mathbf{r}_i)(\mathcal{G}(\widetilde{\mathbf{x}}_i) + \mathbf{r}_i)^T\right)\widehat{\boldsymbol{\theta}}_i + \tau y_i(\mathcal{G}(\widetilde{\mathbf{x}}_i) + \mathbf{r}_i) \\
&= \left(\mathbf{I} - \tau\mathcal{G}(\widetilde{\mathbf{x}}_i)\mathcal{G}(\widetilde{\mathbf{x}}_i)^T - \tau\mathcal{G}(\widetilde{\mathbf{x}}_i)\mathbf{r}_i^T - \tau\mathbf{r}_i\mathcal{G}(\widetilde{\mathbf{x}}_i)^T - \tau\mathbf{r}_i\mathbf{r}_i^T\right)\widehat{\boldsymbol{\theta}}_i + \tau y_i(\mathcal{G}(\widetilde{\mathbf{x}}_i) + \mathbf{r}_i) \\
&= \left(\mathbf{I} - \tau\mathcal{G}(\widetilde{\mathbf{x}}_i)\mathcal{G}(\widetilde{\mathbf{x}}_i)^T\right)\widehat{\boldsymbol{\theta}}_i + \tau y_i\mathcal{G}(\widetilde{\mathbf{x}}_i) - \tau\left(\mathcal{G}(\widetilde{\mathbf{x}}_i)\mathbf{r}_i^T + \mathbf{r}_i\mathcal{G}(\widetilde{\mathbf{x}}_i)^T + \mathbf{r}_i\mathbf{r}_i^T\right)\widehat{\boldsymbol{\theta}}_i + \tau y_i\mathbf{r}_i \\
&= \widehat{\boldsymbol{\theta}}_i - \tau\left(\widehat{\boldsymbol{\theta}}_i^T\mathcal{G}(\widetilde{\mathbf{x}}_i) - y_i\right)\mathcal{G}(\widetilde{\mathbf{x}}_i) \underbrace{-\tau\left(\mathcal{G}(\widetilde{\mathbf{x}}_i)\mathbf{r}_i^T + \mathbf{r}_i\mathcal{G}(\widetilde{\mathbf{x}}_i)^T + \mathbf{r}_i\mathbf{r}_i^T\right)\widehat{\boldsymbol{\theta}}_i + \tau y_i\mathbf{r}_i}_{\mathbf{v}_i}.
\end{aligned}$$

From a control perspective, the deviations from perfect orthogonal mappings create unknowns in the dynamical system: these unknowns $\mathbf{v}_i$ are random variables referred as process noise.

Dealing with process noise is a known and well-investigated problem in control theory [32, Chapter 7]. We leave the best modeling of this process noise for future work. For now, we model the deviations from orthogonality in a naive way by assuming that the noise is independent and identically distributed (i.i.d.) Gaussian, namely, $\mathbf{v}_i \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{v}})$. Under this assumption, the covariance extrapolation for the Kalman update becomes

$$\mathbf{C}_{i+1|i} = \left(\mathbf{I} - \tau\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^T\right)\mathbf{C}_{i|i}\left(\mathbf{I} - \tau\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^T\right)^T + \boldsymbol{\Sigma}_{\mathbf{v}}.$$

Despite the simplicity of the process noise model, we observe a significant improvement in performance. The dashed purple line in Figure 7(b) shows that assuming Gaussian process noise smooths the performance curve. We diminish the drop in performance in the cross-lingual experiment by accounting for the deviations from the orthogonal mapping between Italian and Spanish words with Gaussian process noise.

5. Discussion.

The present work addresses the challenge of online machine teaching with non-omniscient teachers, i.e., scenarios in which the teacher operates under uncertainty about the learner starting state or in a feature space distinct from the learner feature space. We propose two algorithms, SMTL and SMTL-F, which simultaneously guide the learner toward the correct model while estimating its state. SMTL operates without feedback, whereas SMTL-F leverages noisy feedback from the learner. The algorithms manage uncertainty by balancing an exploration-exploitation trade-off. Our experiments on cross-language binary sentiment classification demonstrate the algorithms' practical utility, while our theoretical guarantees confirm the significant performance benefits. Notably, SMTL achieves an exponential speedup compared to random example selection, and SMTL-F ensures optimal estimation with proven stability.

While we offer new insights into online machine teaching, our current results have limitations that provide avenues for future work. For example, although Theorem 3.6 proves the stability of SMTL-F, how to establish stronger convergence guarantees remains an open question. Our framework's connection to control theory opens opportunities for leveraging advanced control-theoretic tools; however, certain challenges remain unresolved. For instance, the optional stopping time theorem has been applied to bound sample complexity in [6] but relies on the existence of a supermartingale incompatible with our setup. The online feedback introduces historical dependencies, preventing the MSE from maintaining the supermartingale property. Similarly, the absence of process noise and the nonuniform structure of the observation matrix invalidate standard assumptions required for Lyapunov's stability theorems. Finally, the lack of persistence excitation condition [26, 16] leads to a lack of Schur stability at the equilibrium points. Addressing these challenges, potentially by incorporating process noise into the convergence framework, would certainly offer a more robust theoretical foundation.

Building on prior work that examines mismatched parameters between teacher and learner, such as different learning rates [10], future research could explore relaxing the assumption of fully known learner dynamics. Moreover, our approach of simultaneously guiding and estimating the learner could be adapted to scenarios in which there exists uncertainty about the learner dynamics but the learner state is perfectly observed. This approach may offer performance improvements compared to the state-of-the-art two-phase strategy, which first approximates the learner dynamics and subsequently teaches [35].

These challenges and opportunities highlight the potential for continued refinement and broader applications of our framework, particularly in scenarios requiring robust teaching under uncertainty.

Acknowledgments. We sincerely thank the associate editor and the reviewers for their valuable feedback and suggestions, which have helped improve this work.

REFERENCES

[1] M. Abdalla and G. Hirst, *Cross-lingual sentiment analysis without (good) translation*, in Proceedings of the International Joint Conference on Natural Language Processing, Vol. 7, 2017.

[2] M. Artetxe, G. Labaka, and E. Agirre, *Learning principled bilingual mappings of word embeddings while preserving monolingual invariance*, in Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2016, pp. 2289–2294, https://doi.org/10.18653/v1/D16-1250.

[3] F. J. Balbach and T. Zeugmann, *Recent developments in algorithmic teaching*, in Proceedings of the International Conference on Language and Automata Theory and Applications, Vol. 5457, Springer, Heidelberg, 2009, pp. 1–18, https://doi.org/10.1007/978-3-642-00982-2.

[4] M. Bayer and C. Reuter, *ActiveLLM: Large Language Model-Based Active Learning for Textual Few-Shot Scenarios*, preprint, http://arxiv.org/abs/2405.10808, 2024.

[5] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*, Cambridge University Press, Cambridge, UK, 2011.

[6] M. V. Burnašev, *Sequential discrimination of hypotheses with control of observations*, Mathematics of the USSR-Izvestija, 15 (1980), pp. 419–440, https://doi.org/10.1070/IM1980v015n03ABEH001255.

[7] G. Canal, S. Fenu, and C. Rozell, *Active ordinal querying for tuplewise similarity learning*, in Proceedings of the 34th AAAI Conference on Artificial Intelligence, 2020, pp. 3332–3340, https://doi.org/10.1609/aaai.v34i04.5734.

[8] S. Dasgupta, D. Hsu, S. Poulis, and X. Zhu, *Teaching a black-box learner*, in Proceedings of the International Conference on Machine Learning, PMLR, 2019, pp. 1547–1555.

[9] K. Dashtipour, S. Poria, A. Hussain, E. Cambria, A. Y. Hawalah, A. Gelbukh, and Q. Zhou, *Multilingual sentiment analysis: State of the art and independent comparison of techniques*, Cogn. Comput., 8 (2016), pp. 757–771, https://doi.org/10.1007/s12559-016-9415-7.

[10] R. Devidze, F. Mansouri, L. Haug, Y. Chen, and A. Singla, *Understanding the power and limitations of teaching with imperfect knowledge*, in Proceedings of the International Joint Conference on Artificial Intelligence, 2021, pp. 2647–2654, https://doi.org/10.24963/ijcai.2020/367.

[11] S. Freitas, E. Laber, P. Lazera, and M. Molinaro, *Time-constrained learning*, Pattern Recognition, 142 (2023), 109672, https://doi.org/10.1016/j.patcog.2023.109672.

[12] I. Gelfand, *Normierte Ringe*, Rec. Math. [Mat. Sbornik] N.S., 9 (1941), pp. 3–24.

[13] I. J. Goodfellow, J. Shlens, and C. Szegedy, *Explaining and harnessing adversarial examples*, in Proceedings of the International Conference on Learning Representations, 2015, pp. 1–11.

[14] C. Guerra, F. S. Melo, and M. Lopes, *FIT: Using feature importance to teach classification tasks to unknown learners*, in Proceedings of Progress in Artificial Intelligence, Springer, Lisbon, 2022, pp. 440–451, https://doi.org/10.1007/978-3-031-16474-3.

[15] X. He, H. Zhang, W. Chao, and D. Wang, *Semi-supervised learning on cross-lingual sentiment analysis with space transfer*, in Proceedings of the International Conference on Big Data Computing Service and Applications, Vol. 8, 2015, pp. 371–377, https://doi.org/10.1109/BigDataService.2015.57.

[16] B. M. Jenkins, A. M. Annaswamy, E. Lavretsky, and T. E. Gibson, *Convergence properties of adaptive systems and the definition of exponential stability*, SIAM J. Control Optim., 56 (2018), pp. 2463–2484, https://doi.org/10.1137/15M1047805.

[17] R. E. Kalman, *A new approach to linear filtering and prediction problems*, J. Basic Eng., 82 (1960), pp. 35–45, https://doi.org/10.1115/1.3662552.

[18] P. Kamalaruban, R. Devidze, V. Cevher, and A. Singla, *Interactive teaching algorithms for inverse reinforcement learning*, in Proceedings of the International Joint Conference on Artificial Intelligence, Vol. 5, AAAI Press, Macao, China, 2019, pp. 2692–2700.

[19] L. Lessard, X. Zhang, and X. Zhu, *An optimal control approach to sequential machine teaching*, in Proceedings of the Conference on Artificial Intelligence and Statistics, 2019, https://proceedings.mlr.press/v89/lessard19a.html.

[20] W. Liu, B. Dai, A. Humayun, C. Tay, C. Yu, L. B. Smith, J. M. Rehg, and L. Song, *Iterative machine teaching*, in Proceedings of the International Conference on Machine Learning, Vol. 5, JMLR, 2017, pp. 3390–3412.

[21] W. Liu, B. Dai, X. Li, Z. Liu, J. M. Rehg, and L. Song, *Towards black-box iterative machine teaching*, in Proceedings of the International Conference on Machine Learning, J. Dy and A. Krause, eds., 2018, pp. 3141–3149.
[22] Y. Lu, T. Javidi, and S. Lazebnik, *Adaptive object detection using adjacency and zoom prediction*, in Proceedings of the Conference on Computer Vision and Pattern Recognition, 2016, pp. 2351–2359, https://doi.org/10.1109/CVPR.2016.258.
[23] F. S. Melo, C. Guerra, and M. Lopes, *Interactive optimal teaching with unknown learners*, in Proceedings of the International Joint Conference on Artificial Intelligence, 2018, pp. 2567–2573, https://doi.org/10.24963/ijcai.2018/356.
[24] J. B. Moore and B. Anderson, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ, 1979.
[25] P. Narayanamurthy and U. Mitra, *Uncertainty-based non-parametric active peak detection*, in Proceedings of the International Symposium on Information Theory, 2022, pp. 1–10, https://ieeexplore.ieee.org/document/9834468.
[26] K. S. Narendra and A. M. Annaswamy, *Persistent excitation in adaptive systems*, Int. J. Control, 45 (1987), pp. 127–160, https://doi.org/10.1080/00207178708933715.
[27] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, *Robust stochastic approximation approach to stochastic programming*, SIAM J. Optim., 19 (2009), pp. 1574–1609, https://doi.org/10.1137/070704277.
[28] S. J. Pan and Q. Yang, *A survey on transfer learning*, IEEE Trans. Knowl. Data Eng., 22 (2010), pp. 1345–1359, https://doi.org/10.1109/TKDE.2009.191.
[29] Z. Qiu, W. Liu, T. Z. Xiao, Z. Liu, U. Bhatt, Y. Luo, A. Weller, and B. Schölkopf, *Iterative teaching by data hallucination*, in Proceedings of the International Conference on Artificial Intelligence and Statistics, F. Ruiz, J. Dy, and J.-W. van de Meent, eds., PMLR, 2023, pp. 9892–9913.
[30] M. D. Rocklage and R. H. Fazio, *The evaluative lexicon: Adjective use as a means of assessing and distinguishing attitude valence, extremity, and emotionality*, J. Exp. Soc. Psychol., 56 (2015), pp. 214–227, https://doi.org/10.1016/j.jesp.2014.10.005.
[31] B. Settles, *Active Learning*, Synth. Lect. Artif. Intell. Mach. Learn. 6, Springer, Cham, 2012, https://doi.org/10.1007/978-3-031-01560-1.
[32] D. Simon, *Kalman filter generalizations*, in Optimal State Estimation: Kalman, H-∞, and Nonlinear Approaches, John Wiley & Sons, Hoboken, NJ, 2006, pp. 183–227, https://doi.org/10.1002/0470045345.
[33] Q. Sun, M. S. Amin, B. Yan, C. Martell, V. Markman, A. Bhasin, and J. Ye, *Transfer learning for bilingual content classification*, in Proceedings of the ACM Conference on Knowledge Discovery and Data Mining, 2015, pp. 2147–2156, https://doi.org/10.1145/2783258.
[34] O. Tamuz, C. Liu, S. Belongie, O. Shamir, and A. T. Kalai, *Adaptively learning the crowd kernel*, in Proceedings of the International Conference on Machine Learning, 2011, pp. 673–680.
[35] C. Wang, S. Zilles, A. Singla, and Y. Chen, *Iterative machine teaching for black-box Markov learners*, in Workshop on Theory of Mind in Communicating Agents at ICML, 2023, https://openreview.net/forum?id=cmuVJMRWEK.
[36] J. Whitehill and J. Movellan, *Approximately optimal teaching of approximately optimal learners*, IEEE Trans. Learn. Technol., 11 (2018), pp. 152–164, https://doi.org/10.1109/TLT.2017.2692761.
[37] C. Xing, D. Wang, C. Liu, and Y. Lin, *Normalized word embedding and orthogonal transform for bilingual word translation*, in Proceedings of the Conference of Human Language Technologies, 2015, pp. 1006–1011, https://doi.org/10.3115/v1/N15-1104.
[38] Y. Yang and M. Loog, *A benchmark and comparison of active learning for logistic regression*, Pattern Recognit., 83 (2018), pp. 401–415, https://doi.org/10.1016/j.patcog.2018.06.004.
[39] Z. Zhang, E. Strubell, and E. Hovy, *A survey of active learning for natural language processing*, in Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP, 2022, pp. 6166–6190.
[40] X. Zhu, *Machine teaching: An inverse problem to machine learning and an approach toward optimal education*, in Proceedings of the International Joint Conference on Artificial Intelligence, Vol. 29, 2015, https://doi.org/10.1609/aaai.v29i1.9761.
[41] X. Zhu, A. Singla, S. Zilles, and A. N. Rafferty, *An Overview of Machine Teaching*, preprint, http://arxiv.org/abs/1801.05927, 2018.