Priors inspired by Speed-Accuracy Trade-Offs for Incremental Learning of Probabilistic Movement Primitives *

Daniel Schäle[1 [0000−0002−3437−4719]], Martin F. Stoelen[1 [0000−0002−2944−759X]], and Erik Kyrkjebø[1 [0000−0002−5487−6839]]

[1] Dept. of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Førde, Norway. dasc@hvl.no

Abstract. Probabilistic Movement Primitives (ProMPs) model robot motor skills by capturing the mean and variance of a set of demonstrations provided by a human teacher. Such a probabilistic representation of motor skills is beneficial in physical human-robot cooperation (pHRC) where robots have to respond to the inherent variance in human motion. However, learning ProMPs incrementally and from scratch, as it is desirable in pHRC, is difficult due to the large number of parameters required to model the distribution of a motor skill compared to the few demonstrations available at the beginning of training. In this paper we propose to predict the variance structure of a motor skill based on the speed found in the individual demonstrations and to incorporate this prediction into the prior parameter distribution of the ProMP. Our basic approach is taking inspiration from the speed-accuracy trade-off found in human motion. Experimental evaluation suggests that with the proposed prior parameter distributions, the true distribution is approached faster in incremental learning of a motor skill than with the priors previously proposed for batch learning.

Keywords: Speed-Accuracy Trade-off · Movement Primitives · Learning by Demonstration

1 Introduction

In physical human-robot cooperation (pHRC) robots are placed in unstructured environments where they have to deal with changing surroundings and, since human and robot are in continuous physical contact over longer periods, the variance inherent in the motions of their human partners. Some of the variations in the environment and human motion can be captured by probabilistic representations of motor skills, as for example probabilistic movement primitives (ProMPs) [5]. ProMPs represent probability distributions over trajectories, summarizing motor skills in terms of a mean trajectory and corresponding variance. ProMPs can be learned from human demonstrations which is a common approach in robotics, enabling humans to teach robots new motor skills fast and

* This work was funded by the Research Council of Norway through grant Nr. 280771.

without writing code. In learning by demonstration (LbD) a human demonstrates multiple instances of a movement to a robot by moving the robot's end effector around either directly or through a jointly gripped object. The robot then learns a ProMP by computing e.g. maximum likelihood estimates (MLE) of the ProMP parameters based on the provided examples.

This learning process can either happen incrementally or in batch. Learning in batch means that a set of demonstrations is recorded at the beginning of training. Thereafter, a learning algorithm computes the ProMP parameters on basis of all demonstrations in the set. Since the robot remains passive during the demonstration phase, any cooperation between human and robot is delayed and the human experiences no support in providing the demonstrations in cooperative tasks such as the joint manipulation of objects. A possibility to achieve human-robot cooperation with batch learning is to first train a ProMP covering all expected variations in a task and then use conditioning to adapt the ProMP to the preferences and circumstances at the time. However, this approach requires external sensor systems to detect changes in the environment, aggravating its use in practice. Furthermore, lay users have a limited understanding of how a robot learns motor skills and how it reacts to new demonstration inputs, making it difficult to deliver just the right demonstrations to the robot without any preview of the resulting motor skill. The training in batch mode has the advantage that a larger data set is available, making the estimation of the ProMP parameters easier, but is however unintuitive in cooperative tasks, where humans expect an incremental training progress of their partner. Batch learning is thus better suited for tasks outside the pHRC domain, where human and robot are not contributing actively and simultaneously to the same task, and where tasks can be broken down intuitively into a training and an execution phase.

In incremental learning, the ProMP parameters are updated sequentially – each time a new demonstration arrives [6]. With that, incremental learning is well suited to pHRC, where human and robot solve tasks together and a new demonstration naturally arrives with each execution of a task. After each execution, the motor skill can be slightly adjusted to the most recent preferences of how to execute the task. Incremental learning has the advantage that the human can observe the robot's training progress while they learn a new task together. Cooperation can emerge after the first demonstration and the human can focus on corrections and adaptations – compared to batch learning, the time in which the human has to take on the full leader role is minimized. Shaping motor skills over time with corrective demonstrations can be further facilitated by a forgetting factor which gives more weight to recent demonstrations over what was learned earlier [6]. Incremental learning using a forgetting factor is a step towards life long learning of motor skills.

A challenge in incremental learning is the low number of demonstrations available at the beginning of training. When learning a motor skill from scratch, the ProMP distribution initially has to be estimated on basis of a single demonstration, leading to a degenerate distribution with zero covariance. Also the subsequent estimates, as long as the number of demonstrations is small compared to

the dimension of the covariance matrix, suffer from numerical instability caused by rank deficient (singular) covariance matrices. Poor estimates of the covariance matrix are especially severe when the robot's compliance along the trajectory is adjusted according to the variance of the ProMP.

The problem of poor parameter estimates in face of a small sample of demonstrations can be countered with regularization in the form of prior parameter distributions, and hence the computation of maximum a posteriori estimates (MAP) of the parameters, which has been shown to improve the robustness of parameter estimates in batch [3] and incremental [6] learning of ProMPs. The authors in [3] propose to use an uninformative prior for the mean of the ProMP and an informative, data-dependent prior for its covariance matrix, where they use the (scaled) block-diagonal maximum likelihood estimate of the covariance matrix as an initial guess for the parameter. With this setting, the MAP of the mean is equal to its MLE. The block-diagonal prior for the covariance matrix favours the off-diagonal elements representing the correlations between the robot's joints to be zero in presence of a small number of demonstrations, yielding numerically more robust estimates while holding the variances of each joint close to their MLEs. Even though this prior has a positive effect on the robustness of the covariance estimates, it does not actually input prior knowledge into the system but instead utilizes information from all available demonstrations. Hence, it does not cope with the lack of information about the movement/task itself encountered when learning new motor skills incrementally.

It is of course difficult to make general prior assumptions about all arbitrary movements that could possibly be learnt, but it may be possible to make assumptions about the distribution of a specific motor skill based on individual demonstrations. Considering that the demonstrations come from a human manipulating the robot's end effector, characteristics/features of human motion can potentially serve as a source for such prior assumptions. One such feature is the speed accuracy trade-off found in human motion, implying that movement accuracy decreases as speed increases. The most renowned model of this trade-off may be Fitts' law [2]. Fitts' law, originally proposed for translational movements, states that the movement time in a pointing task is a function of the distance to the target and the target width. The smaller the target width, the greater the elapsed time to reach the target. Experiments show that also one-dimensional rotational movements and combined translational and rotational movements can be modelled well with a Fitts' law equivalent [7]. Fitts' law was later generalized to trajectory-based movements, resulting in the steering law [1]. The steering law proposes a linear relationship between the steering time and the "tunnel" width which imposes a spatial accuracy constraint on the movement. Further research has been devoted to the steering law, investigating the effects of combinations of spatial and temporal constraints [11], the effects of narrowing or widening tunnels on the steering time [9] and steering through sequential linear path segments [10]. The aforementioned research is rooted in the field of human computer interaction and the derived models are only verified for simple 1 or 2 dimensional movements in absence of force interactions with the environment.

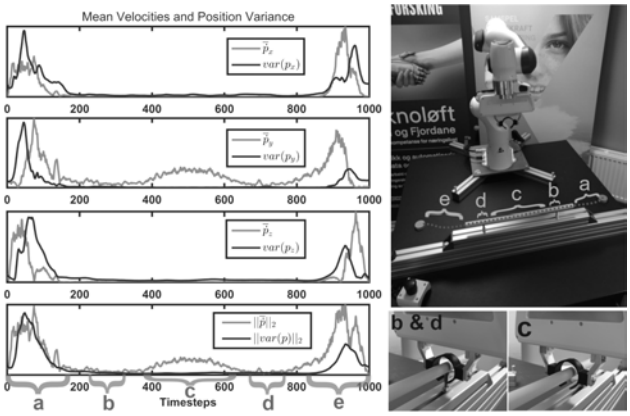

Fig. 1. Example data supporting our assumption of a speed-accuracy trade-off in kinesthetic teaching. The data are computed from 51 kinesthetic demonstrations recorded in the setup on the top-right. The plot shows in **light grey** the absolute mean task-space velocities in X,Y,Z as well as the 2-norm of the mean velocity vector. The sample variance of the mean task-space positions in X,Y,Z and the 2-norm of the mean position vector are shown in **black**. When the task-space velocity is high the variance of the task space position is high as well. The variance and thus the velocity is low at sections **b** and **d** and **c** where the movement is physically constraint by the environment. Note that the velocity peak in section **c** has no corresponding peak in the variance since we aligned the data with dynamic time warping (DTW). Without DTW the speed-accuracy relationship is even more distinct.

Only recently, motivated by the growth of virtual and mixed reality technology, the development of higher dimensional speed-accuracy models suited to describe 3D object interactions has gotten into focus [8].

Even though we did not find specific studies proving the speed-accuracy trade-off to be present while a human is manipulating a robot's end effector in 3D space by kinesthetic teaching, we suggest to exploit the basic idea of Fitts' law to make prior assumptions about the variance of a ProMP/motor skill based on individual demonstrations. We assume that while delivering demonstrations, the human is subject to a speed-accuracy trade-off limiting the human to guide the robot in a slower pace in directions in which the spatial constraints of the movement are stringent. Fig. 1 shows observations we found in kinesthetic demonstration data from a setup that imposes a spatial constraint on the robot's end effector that support this assumption. We propose, that by examining the task-space velocity along the path, we are able to make an estimate of the variance that can be expected in one section relative to other sections of the movement. These estimated variances can be incorporated into the prior of the covariance matrix, enhancing the ProMP with context about the task in early stages of training.

2 Probabilistic Movement Primitives

A ProMP represents a distribution over trajectories [5]. A trajectory $\boldsymbol{\tau} = \{\boldsymbol{y}_t\}_{t=1}^{T}$ is a time-series of vector-valued robot states $\boldsymbol{y}_t \in \mathcal{S}$ in a state space $\mathcal{S} \subseteq \mathbb{R}^D$, where D is the dimension of the state space. In this paper we encode trajectories in task space by recording the robot's end effector position in 3D Euclidean space. Trajectories are concisely represented as weight vectors $\boldsymbol{w} \in \mathbb{R}^{KD}$ in the basis function model $\boldsymbol{y}_t = \boldsymbol{\Phi}_t \boldsymbol{w} + \boldsymbol{\epsilon}_y$, where $\boldsymbol{\Phi}_t \in \mathbb{R}^{D \times KD}$ is a time dependent, block diagonal basis function matrix containing on its diagonal a row vector $\boldsymbol{\phi}_{d,t}^{\intercal} \in \mathbb{R}^K$ for each degree of freedom, which again contains the values of K normalized, evenly spaced, Gaussian basis functions $\phi_k(t)$ evaluated at time t. The last term $\boldsymbol{\epsilon}_y \in \mathbb{R}^D$ is a vector containing the observation noise which is assumed to be independent and identically distributed according to the normal distribution $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_y)$. Given a weight vector $\boldsymbol{w}$, it follows that a trajectory $\boldsymbol{\tau}$ consisting of T time steps is distributed according to

$$p(\boldsymbol{\tau}|\boldsymbol{w}) = \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{y}_t|\boldsymbol{\Phi}_t \boldsymbol{w}, \boldsymbol{\Sigma}_y) \ . \tag{1}$$

Multiple demonstrations of the same movement are expected to differ slightly. This implies that different weight vectors $\boldsymbol{w}_n$ are needed to represent the n different instances of a movement. The underlying mechanism generating the weight vector samples is assumed to be a Gaussian distribution

$$p(\boldsymbol{w}|\boldsymbol{\theta}_w) = \mathcal{N}(\boldsymbol{w}|\boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w) \ , \tag{2}$$

where $\boldsymbol{\theta}_w = \{\boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w\}$ are the distribution parameters. The mean vector $\boldsymbol{\mu}_w \in \mathbb{R}^{KD}$ summarizes the mean of the demonstrations in each degree of freedom. The covariance matrix $\boldsymbol{\Sigma}_w \in \mathbb{R}^{KD \times KD}$ represents the variances and covariances of the demonstrations in respectively between each degree of freedom.

Learning a ProMP from demonstration can be done by maximizing the likelihood of the N observed trajectories $\boldsymbol{Y} = \{\boldsymbol{\tau}_n\}_{n=1}^{N}$ with respect to the ProMP parameters i.e. computing the maximum likelihood estimate (MLE) $\boldsymbol{\theta}_w^{MLE} = \arg\max_{\boldsymbol{\theta}_w} p(\boldsymbol{Y}|\boldsymbol{\theta}_w)$, where the marginal likelihood is given by

$$p(\boldsymbol{Y}|\boldsymbol{\theta}_w) = \prod_{n=1}^{N} \int p(\boldsymbol{w}_n|\boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w) \prod_{t=1}^{T} p(\boldsymbol{y}_{nt}|\boldsymbol{w}_n) d\boldsymbol{w}_n \ . \tag{3}$$

For regularization of the MLE, a prior distribution $p(\boldsymbol{\theta}_w)$ over the ProMP parameters can be incorporated into the maximization problem which becomes $\boldsymbol{\theta}_w^{MAP} = \arg\max_{\boldsymbol{\theta}_w} p(\boldsymbol{Y}|\boldsymbol{\theta}_w) p(\boldsymbol{\theta}_w)$. Where $\boldsymbol{\theta}_w^{MAP}$ is the mode of the posterior distribution $p(\boldsymbol{\theta}_w|\boldsymbol{Y})$ known as the maximum a posteriori estimate (MAP). This maximization can be accomplished by means of the expectation-maximization (EM) algorithm in batch [3] and incremental [6] learning settings. Details about the learning algorithms can be found in the respective publications. Since the distribution in Eq. 1 is assumed to be a multivariate normal with unknown mean and variance its conjugate prior is a normal-inverse Wishart distribution [4]. Using the conjugate prior has the advantage that the computations in the EM can

be solved in closed form. The normal-inverse Wishart prior has the form

$$\begin{aligned} p(\boldsymbol{\theta}_w) &= \mathrm{NIW}(\boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w | \boldsymbol{m}_0, k_0, \boldsymbol{S}_0, v_0) \\ &= \mathcal{N}(\boldsymbol{\mu}_w | \boldsymbol{m}_0, \tfrac{1}{k_0}\boldsymbol{\Sigma}_w)\mathrm{IW}(\boldsymbol{\Sigma}_w | \boldsymbol{S}_0, v_0) \ , \end{aligned} \tag{4}$$

where $\mathcal{N}(\boldsymbol{\mu}_w | \boldsymbol{m}_0, \frac{1}{k_0}\boldsymbol{\Sigma}_w)$ is a normal distribution representing our prior belief about the ProMP mean with $\boldsymbol{m}_0$ being the prior mean and k_0 controlling the prior strength, and $\mathrm{IW}(\boldsymbol{\Sigma}_w | \boldsymbol{S}_0, v_0)$ is an inverse Wishart distribution representing our prior belief about the covariance matrix of the ProMP with $\boldsymbol{S}_0$ being (proportional to) the prior mean and v_0 controlling the prior strength [4].

3 Prior Parameters inspired by Speed-Accuracy Trade-off

Inspired by the speed-accuracy trade-off described in studies on human motor control, we investigate the design of the scale matrix $\boldsymbol{S}_0$ specifying our initial guess of the ProMP variance based on the velocities in the demonstrations. For our initial investigation of this idea in this paper we choose a straight-forward approach. We consider learning a new motor skill incrementally and from scratch, using the incremental learning algorithm for ProMPs presented in [6]. The computation steps are shown in Fig. 2. The training process begins with the human

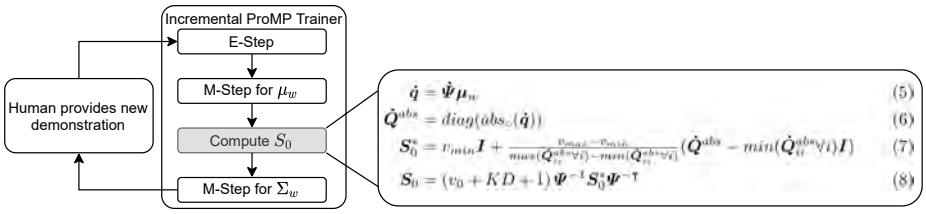

Fig. 2. Computation steps during incremental learning with the speed-accuracy trade-off (SAT) prior proposed in this paper. The SAT prior is computed in the grey block between the M-step of the mean $\boldsymbol{\mu}_w$ and the M-step of the covariance matrix $\boldsymbol{\Sigma}_w$.

providing the first demonstration to the robot via kinesthetic teaching. During the teaching, we record the Cartesian coordinates of the robot's end effector $\boldsymbol{\tau} = \{\boldsymbol{y}_t\}_{t=1}^{T} = \left\{\left(p_x^t \ p_y^t \ p_z^t\right)\right\}_{t=1}^{T}$. Each time a new demonstration is available, the learning algorithm is executed to incorporate the new demonstration into the ProMP. Between the M-step for the mean $\boldsymbol{\mu}_w$ and the M-step for the covariance matrix $\boldsymbol{\Sigma}_w$ we compute the prior parameter $\boldsymbol{S}_0$ as follows:

$$\dot{\boldsymbol{q}} = \dot{\boldsymbol{\Psi}}\boldsymbol{\mu}_w \tag{5}$$

$$\dot{\boldsymbol{Q}}^{abs} = diag(abs_\circ(\dot{\boldsymbol{q}})) \tag{6}$$

$$\boldsymbol{S}_0^* = v_{min}\boldsymbol{I} + \frac{v_{max}-v_{min}}{max(\dot{\boldsymbol{Q}}_{ii}^{abs}\forall i)-min(\dot{\boldsymbol{Q}}_{ii}^{abs}\forall i)}(\dot{\boldsymbol{Q}}^{abs} - min(\dot{\boldsymbol{Q}}_{ii}^{abs}\forall i)\boldsymbol{I}) \tag{7}$$

$$\boldsymbol{S}_0 = (v_0 + KD + 1)\,\boldsymbol{\Psi}^{-1}\boldsymbol{S}_0^*\boldsymbol{\Psi}^{-\intercal} \tag{8}$$

where $\boldsymbol{\Psi}, \dot{\boldsymbol{\Psi}} \in \mathbb{R}^{HD \times KD}$ are a block diagonal basis function matrices with each block containing K normalized Gaussian basis functions respectively the derivatives of K normalized Gaussian basis functions evaluated at $H = K$ evenly spaced time steps. $\dot{\boldsymbol{q}}$ is the concatenation of the velocity profiles in x, y, z. By setting $H = K$ the velocity profiles will be coarse and consist only of as many time steps as there are basis functions. In Eq. 6 the operator $abs_{\circ}(\cdot)$ computes the element wise absolute values of the velocity vector. The $diag(\cdot)$ operator forms a diagonal matrix from a given vector. In Eq. 7 the absolute velocities are rescaled to the range between the minimum and maximum desired variance $[v_{min}, v_{max}]$, where $0 < v_{min} \leq v_{max}$. The resulting diagonal matrix $\boldsymbol{S}_0^*$ contains the rescaled absolute velocities on its diagonal, thus when interpreted as a covariance matrix $\boldsymbol{S}_0^*$ has higher variance at sections of the movement where the velocity was high – representing a speed-accuracy trade-off. Multiplying $\boldsymbol{S}_0^*$ by the inverse of the basis function matrices has the effect that v_{min} and v_{max} correspond to actual minimum and maximum variance of the ProMP in task space, making it more intuitive to set the scaling parameters. So can we set them in task space units in terms of the smallest and biggest standard deviation we expect for the task. The scaling still has to be set manually but is at least limited to the maximum precision and maximum reach of the robot and can be estimated from the dimensions of the real world set up of the task.

Similar to [3], we multiply $\boldsymbol{S}_0^*$ by $(v_0 + KD + 1)$ such that the MAP estimate $\boldsymbol{\Sigma}_w$ becomes a convex combination of $\boldsymbol{S}_0$ and $\boldsymbol{\Sigma}_w^{MLE}$ with a coefficient dependent on the number of demonstrations N. We set the parameter $v_0 = KD+2$ to ensure that the expected value of the inverse Wishart distribution equals $\boldsymbol{S}_0$.

Instead of using the scaling matrix computed in Eq. 5-7 alone, we can also blend it with the block diagonal MLE of the covariance matrix, yielding

$$\boldsymbol{S}_0 = (v_0 + KD + 1)\left((1-\lambda)\,\boldsymbol{\Psi}^{-1}\boldsymbol{S}_0^*\boldsymbol{\Psi}^{-\intercal} + \lambda\,\mathrm{blockdiag}(\boldsymbol{\Sigma}_w^*)\right) \tag{9}$$

$$\lambda = \begin{cases} \frac{n-1}{\eta} & 1 \leq n \leq \eta \\ 1 & \text{otherwise} \end{cases}, \tag{10}$$

where n is the current number of demonstrations and η is the desired number of demonstrations in which the influence of $\boldsymbol{S}_0^*$ on $\boldsymbol{S}_0$ diminishes to zero. By setting $\eta \approx 15$, the prior assumption based on the speed-accuracy trade-off helps to bridge the initial phase of training where only a few demonstrations are available. As more and more demonstrations are accumulated the influence of the MLE on the prior can be increased.

4 Experimental Evaluation: Comparison of Prior Parameters in incremental Learning

For a brief experimental evaluation of our proposed approach, we compare the effects of different prior parameters for the covariance matrix on a data set generated from demonstrations recorded on a Franka Emika Panda manipulator.

The experimental setup is shown in Fig. 1. The goal of the task was to hand-guide the robot's end effector from the round mark on the right to the round mark on the left, while following along the tube with the circular gripper fingers without touching it. The tube poses a physical constraint in task space that could correspond to sliding/insertion movements or tasks like glueing, cutting or welding in industrial settings. During the tube section, the human has to be precise (diameter of the tube 10mm, diameter of the gripper 40mm), hence we argue that he/she is guiding the robot slowly. Our prior computed as described in section 3 incorporates this assumption and suggests a lower variance during the tube section than during the sections before and after the tube.

We recorded 51 demonstrations on the physical setup. Dynamic time warping (DTW) was used for temporal alignment of the demonstrations to rule out that our proposed prior predicts variance stemming from temporal misalignment of the data and not from the speed-accuracy trade-off of the human. Since DTW removes the variance tangential to the path, we removed the tangential variance from our prior by transforming $\boldsymbol{S}_0^*$ to the Frenet-Serret frames of the path represented by the mean vector $\boldsymbol{\mu}_w$, setting the tangential component to zero, and transforming it back to task space coordinates. From the aligned demonstrations, we compute MLE estimates of the parameters of a ProMP with the batch EM algorithm from [3]. We use this ProMP as a reference to compare the parameter estimates during the incremental training under the different priors. We compare following four priors with different prior parameter $\boldsymbol{S}_0$: *MLE*, *SAT*, *blended* and *const. tube*. *MLE* is computed from the block diagonal MLE estimate of $\boldsymbol{\Sigma}_w$ as proposed by [3]. *SAT* is computed as proposed in this paper in Eq. 5-8. *Blended* is computed as a blending of *SAT* and *blkdiag MLE* as proposed in Eq. 9-10 with $\eta = 15$. *Const. tube* is computed as in Eq. 8 with $\boldsymbol{S}_0^* = \frac{v_{min}+v_{max}}{2}\boldsymbol{I}$ which yields a constant tube with a radius equal to the mean of the minimum and maximum radius of the *SAT* prior. The *const. tube* prior serves as a benchmark to get an impression how much information is conveyed in the scaling of the *SAT* prior alone and whether the information about the variance structure based on the speed-accuracy trade-off is beneficial. All other prior parameters are set to same values: $\boldsymbol{m}_0 = \boldsymbol{0}$, $k_0 = 0$ and $v_0 = KD + 2 = 62$. To check the sensitivity of the performance of our proposed prior to the scaling parameters v_{min} and v_{max}, we tested and compared three different parameter settings. As training data for the incremental training, we sampled 500 demonstrations from the reference ProMP. These 500 demonstrations where split into 10 blocks of $N = 50$ demonstrations. The results presented in Fig. 3 are averaged results of the 10 blocks. All ProMPs in the experiment have $K = 20$ basis functions. We compare the effect of the different priors on the incremental training performance by means of the Kullback-Leibler divergence (KLD) between the reference ProMP and the ProMPs under test. In addition, and to emphasize the effect of the variances, we compute the KLD where we set all off-diagonal elements of the covariance matrices $\boldsymbol{\Sigma}_w$ of reference and ProMPs under test to zero. To monitor the numeric stability of the covariance estimates, we compute the matrix condition number of $\boldsymbol{\Sigma}_w$ during the course of training.

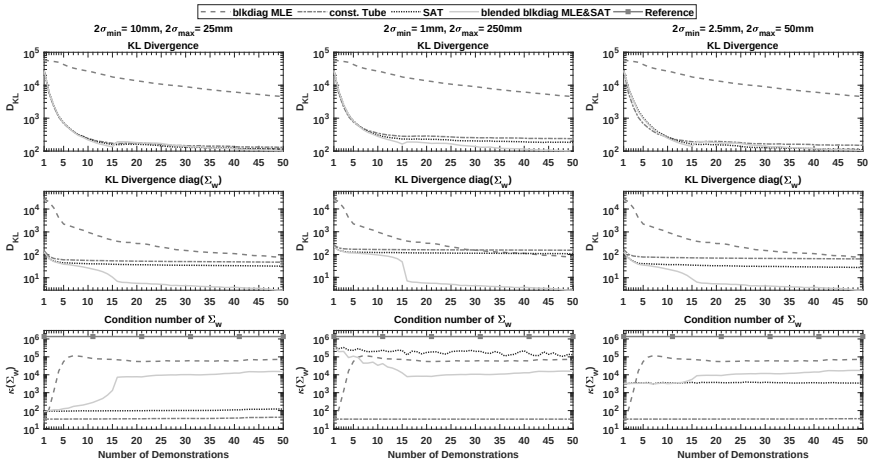

Fig. 3. Results of the experimental evaluation. Each column shows the results of a different scaling setting of the SAT prior and the const. tube. The first row shows the Kullback-Leibler divergence (KLD) of the ProMPs with different priors to the reference ProMP. The second row shows the KLD where only the diagonal elements of the covariance matrix $\boldsymbol{\Sigma}_w$ are considered. The last row shows the matrix condition number of $\boldsymbol{\Sigma}_w$. All y-axes are in logarithmic scale.

The KLD of the *MLE* prior from [3] is greater throughout the entire training than those of the other priors in all three scaling settings. The other three priors perform relatively similar during the first 15 demonstrations, towards the end of training the *SAT* and *blended* perform slightly better than the *const. tube*. The performance of all priors is better when we only consider the diagonalized covariance matrices. The *MLE* and hence also the *blended* show the biggest performance increase. The *blended* prior achieves the lowest final KLD in all experiments, followed by the *SAT* prior which has the second lowest final KLD except for the second scaling setting where the upper scaling bound is very large. For a qualitative comparison of the effects of the different priors on the initial training phase we show the ProMPs after training on only the first demonstration in Fig. 4. The ProMP trained with the *SAT* (*blended* and *SAT* are identical after one demonstration) resembles the reference best and already gives hints about the task constraints after just one demonstration.

5 Discussion and Conclusion

The experimental results in this paper show that the *MLE* prior from [3] is not the best choice for incremental learning settings. The three other priors tested approach the true distribution faster within 50 demonstrations of training. Measured in terms of the KLD, the difference between the *const. tube* and the *SAT* and *blended* is not large, but slightly more distinct when considering the diagonalized covariance matrices. We conclude that the proposed prior inspired by the

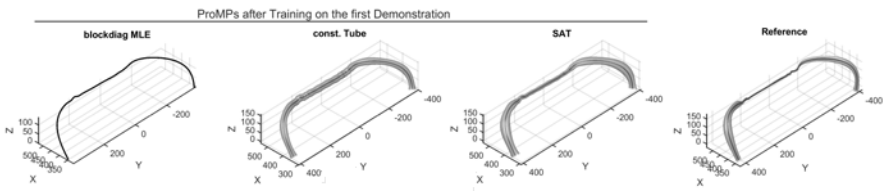

Fig. 4. Qualitative comparison: effects of the different priors on the variance structure of the ProMPs after training on the first demonstration only. The **black** lines show the mean and the grey **tubes** two standard deviations of the ProMPs, plotted without the observation noise $\boldsymbol{\Sigma}_y$. The ProMP trained with the **SAT** prior proposed in this paper resembles the **Reference** best and already gives hints about the task constraints after just one demonstration. The ProMP with the *blended* prior is identical to the **SAT** ProMP after the first demonstration and therefore not shown separately. The tube of the ProMP trained with the **MLE** prior appears black since its diameter is smaller than the thickness of the mean line.

speed-accuracy trade-off provides additional beneficial information compared to a constant variance structure along the movement. How much this additional information helps in practice, e.g. in the cooperative learning of a new task, has to be determined in real world experiments. Based on the qualitative comparison of the ProMPs which shows that our *SAT* prior reveals more information about the spatial constraints of the task than the others after the first demonstration, we speculate that our prior can have a noticeable effect in incremental, cooperative learning settings where the physical interaction between human and robot is controlled according to the variance of the ProMP. We speculate that when using the proposed prior, the robot would take over the leader role in precise sections of a task early, relieving the human from the tedious precise control in these sections. Since we only predict variance of the human in interaction with the environment but not variations in the environment itself, our approach fails in tasks that have precise sections (e.g. via points) that change position. In such cases either separate primitives for the via point positions have to be learnt, or the position of the via point has to be captured and encoded as a context variable of the primitive. More work has to be devoted to the proposed method for computing priors before it can be used in practice. At this stage, it is rather a tool to demonstrate the principle idea to estimate task constraints from movement speed. We have only tested our approach on a limited set of tasks and there are still some open parameters that have to be guessed when computing the prior, e.g. the scaling of minimal and maximal desired variance of the prior. Even though the experiments did not show a strong sensitivity towards the scaling, the influence of the parameters on the performance has to be studied in further experiments.

Regardless of the applicability and possible benefits of our prior, it is interesting that it is possible to predict, to some extent, the variance structure of a motor skill by means of the speed-accuracy trade-off. This paper did not aim to provide conclusive proof of a relationship between movement speed and spatial

accuracy in kinesthetic teaching, but we think it shows that further investigation of the underlying idea of this paper with regard to incremental learning of movement primitives may be worthwhile.

6 Future Work

To develop the idea of priors inspired by the speed-accuracy trade-off further and to establish them for use in practice, we need to test them on a wider range of tasks. We also need to investigate how to include orientations in addition to translations and how to determine the overall scaling of the variance structure automatically by some algorithm. Furthermore, it is interesting to explore whether there are differences in predicting the variance of point-to-point and trajectory based movements and if priors resembling Fitts' law or the steering law work better in either of them.

References

1. Accot, J., Zhai, S.: Beyond fitts' law. In: Pemberton, S. (ed.) Human factors in computing systems. pp. 295–302. ACM Press, Addison-Wesley (1997)
2. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of exp. psychology **47**(6), 381–391 (1954)
3. Gomez-Gonzalez, S., Neumann, G., Scholkopf, B., Peters, J.: Adaptation and robust learning of probabilistic movement primitives. IEEE Transactions on Robotics **36**(2), 366–379 (2020)
4. Murphy, K.P.: Machine learning: A probabilistic perspective. Adaptive computation and machine learning series, MIT Press, Cambridge, Mass. and London (2012)
5. Paraschos, A., Daniel, C., Peters, J.R., Neumann, G.: Probabilistic movement primitives. In: C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Q. Weinberger (eds.) Advances in Neural Information Processing Systems 26, pp. 2616–2624. Curran Ass., Inc (2013)
6. Schäle, D., Stoelen, M.F., Kyrkjebø, E.: Incremental learning of probabilistic movement primitives (promps) for human-robot cooperation (2021, preprint available on arXiv), https://arxiv.org/abs/2105.13775
7. Stoelen, M.F., Akin, D.L.: Assessment of fitts' law for quantifying combined rotational and translational movements. Human factors **52**(1), 63–77 (2010)
8. Triantafyllidis, E., Li, Z.: The challenges in modeling human performance in 3d space with fitts' law
9. Yamanaka, S., Miyashita, H.: Modeling the steering time difference between narrowing and widening tunnels. In: Kaye, J., Druin, A., Lampe, C., Morris, D., Hourcade, J.P. (eds.) CHI 2016. pp. 1846–1856. The Association for Computing Machinery, New York, New York (2016)
10. Yamanaka, S., Stuerzlinger, W., Miyashita, H.: Steering through sequential linear path segments. In: Mark, G., Fussell, S., Lampe, C., schraefel, m., Hourcade, J.P., Appert, C., Wigdor, D. (eds.) Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. pp. 232–243. ACM, New York, NY, USA (2017)
11. Zhou, X., Cao, X., Ren, X.: Speed-accuracy tradeoff in trajectory-based tasks with temporal constraint. In: Human-Computer Interaction – INTERACT 2009. Springer (2009)