QDP: Learning to Sequentially Optimise Quasi-Static and Dynamic Manipulation Primitives for Robotic Cloth Manipulation

David Blanco-Mulero[1], Gokhan Alcan[1], Fares J. Abu-Dakka[2], Ville Kyrki[1]

I. Introduction

The broad variety of fabric materials that are handled by humans everyday presents several challenges when manipulated by robotic systems. As an example, the variability of cloth materials in house-hold objects in terms of shape, stiffness, elasticity, and mass [1], requires humans to perform a set of diverse manipulation actions such as those used for dressing assistance [2] or flattening, folding and twisting cloths [3]. However, it is not trivial to transfer such skills to robot manipulators, as manipulation primitives are often manually designed for a specific application [4], or a set of pre-tuned primitives is used [5]. Thus, to cope with a wide range of cloths, robotic systems should integrate learning algorithms that autonomously decide the parameter values of these manipulation primitives.

The manipulation primitives used in robotic cloth manipulation fall into two categories: *quasi-static*, such as the pick-and-place primitive [6]; and *dynamic*, which involve the forces of acceleration of the manipulator [7], such as the fling primitive [4]. Quasi-static primitives have been extensively used for cloth folding and unfolding [5], [6], [8]–[11], where different algorithms have been proposed for deciding the end-effector pick-and-place positions. However, parameters such as the specific trajectory height or velocity of the primitive have been neglected. This is crucial in applications such as cloth manipulation in-contact with a surface, where the size of the cloth will drastically affect the manipulation result, *e.g.,* bigger cloths will have more contact if they follow a trajectory with lower height (see Fig 1). Hence, these parameters should be taken into account to cope with a diverse set of cloth materials and sizes.

In this abstract, we propose a method to learn a visual policy that can determine the optimal parameters of manipulation primitives for cloth manipulation. We refer to our method as QDP, short for the Quasi-Dynamic Parameterisable manipulation method. Our key idea is to find the optimal parameters by following a sequential decision approach. We take inspiration from the Sequential RL (SRL) framework [12] to sequentially assign an action space to each primitive parameter. Here, each parameter of the primitive is informed by the previous proposed parameter. This enables learning the relationship between the primitive parameters and their impact on the manipulation performance. Our method works in a joint action space: the spatial action space of pick-and-place locations; and the parameter action space, that defines parameters such as the velocity of the manipulation primitive. To evaluate the effectiveness of the sequential decision of parameter values, we perform simulation and real-world experiments in the task of unfolding a cloth using a single robotic arm. The full paper [13], as well as supplemental material can be found at the project website[1]. Our contributions include:

- Introducing QDP, a novel approach that can optimise the parameters of manipulation primitives, decoupling the pick-and-place decisions as well as additional primitive parameters, without supervision or hand-labeled data during training,
- An analysis of the performance of QDP on the cloth unfolding task, showing the superior performance of the proposed method compared to baselines, as well as its ability to find optimal parameters for both quasi-static and dynamic manipulation primitives,
- For the first time, a real-world evaluation of different manipulation primitives on a public cloth unfolding benchmark [1].

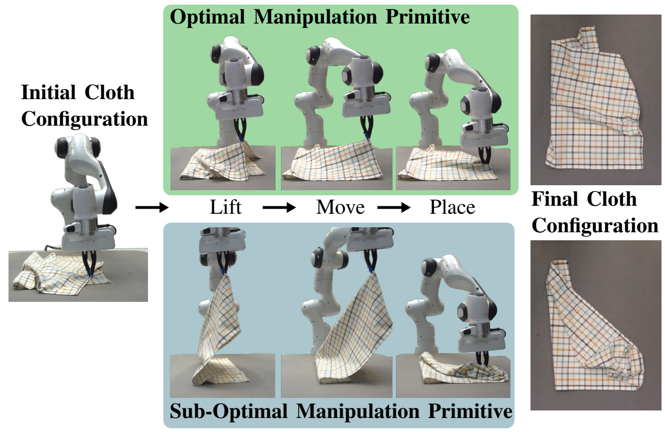

Fig. 1: QDP sequentially optimises the manipulation primitive parameter values to achieve better cloth configurations (*green*) compared to using sub-optimal parameter values (*blue*) for a manipulation primitive such as pick-and-place.

[1]https://sites.google.com/view/qdp-srl

This work was financially supported by the Academy of Finland grant numbers 317020 and 328399. Abu-Dakka, F. is supported by the European project euROBIN under grant agreement No 101070596.

[1]David Blanco-Mulero, Gokhan Alcan and Ville Kyrki are with School of Electrical Engineering, Aalto University, Finland.) (e-mail: david.blancomulero@aalto.fi

[2]Munich Institute of Robotics and Machine Intelligence, Technische Universität München, 80992 München, Germany. Part of the research presented in this work has been conducted when Abu-Dakka, F. was at Intelligent Robotics Group, EEA, Aalto University, 02150 Espoo, Finland

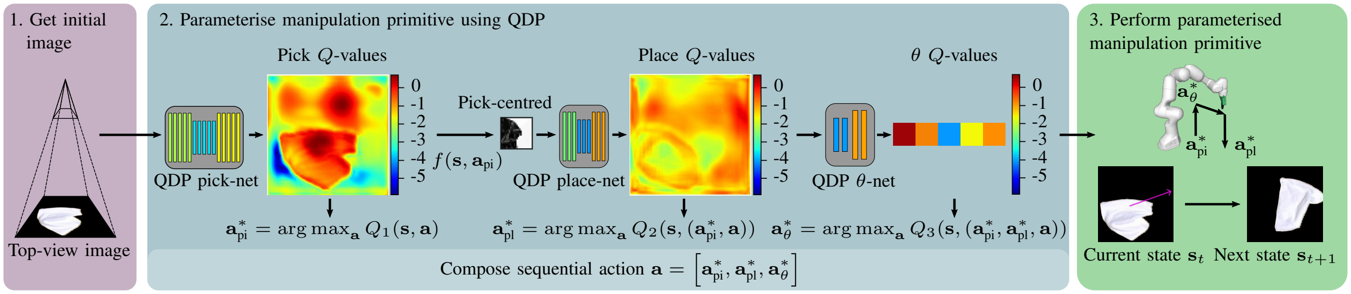

Fig. 2: The proposed Quasi-Dynamic Parameterisable approach starts by getting a top-view image of the cloth. Then, to find the optimal manipulation primitive parameters, a sequential action $\mathbf{a}$ is composed from the sub-action output of three different networks: QDP pick-net, predicts the optimal pick position $\mathbf{a}_{\text{pi}}^*$; QDP place-net, predicts the place location $\mathbf{a}_{\text{pl}}^*$; and QDP θ-net, predicts additional primitive parameters $\mathbf{a}_{\theta}^*$, such as the primitive velocity. Each sub-action takes into account the previous information via encodings, e.g. the place sub-action accounts for the pick location using a pick-centred image. Finally, the manipulation primitive is executed with the optimal parameters placing the cloth on a new state.

II. Sequential Parameter Choice via QDP

QDP learns a visual policy that can determine the optimal parameters of a manipulation primitive for cloth manipulation (see Figure 2). Given an initial top-view image, our proposed approach determines the optimal pick, $\mathbf{a}_{\text{pi}}^* \in A_1$; place, $\mathbf{a}_{\text{pl}}^* \in A_2$; and other parameters of the manipulation primitive, $\mathbf{a}_{\theta}^* \in A_3$. These sub-actions are determined sequentially following the SRL augmented state approach [12], decomposing the action space A into N sequential sub-actions $A = A_1 \times \cdots \times A_N$ for computational feasibility.

The state is defined as a gray-scale image $\mathbf{s} \in \mathbb{R}^{D\times D}$, where D is the height and width dimension of the image. By decomposing the pick and place into multiple decisions we can decide the place position based on the pick action. Thus, the first decision in our sequential RL setting is the pick position which is determined as

$$\mathbf{a}_{\text{pi}}^* = \arg\max_{\mathbf{a}_{\text{pi}}} Q_1(\mathbf{s}, \mathbf{a}_{\text{pi}}), \tag{1}$$

where the optimal pick $\mathbf{a}_{\text{pi}}^* \in \mathbb{N}^2$ is the pixel position in the image space. The action-value function Q_1 is approximated by a neural network that also provides an encoding $g(\mathbf{s})$ of the state-space. Then, for deciding the place position we use the augmented state $S_2 = S \times A_1$. In order to encode the sub-action $\mathbf{a}_{\text{pi}}$ we use a mapping $f : (\mathbf{s}, \mathbf{a}_{\text{pi}}) \mapsto \mathcal{I}_{\text{pick}}$ that creates a pick-centred image $\mathcal{I}_{\text{pick}} \in \mathbb{R}^{E\times E}$, which is a cropped image centred in the pick position, similar to [12]. Instead of using the state $\mathbf{s}$ as input to the action-value function Q_2 we use an encoding $g(\mathbf{s})$, which is part of the output from the neural network that approximates Q_1. The place sub-action is then computed as

$$\mathbf{a}_{\text{pl}}^* = \arg\max_{\mathbf{a}_{\text{pl}}} Q_2(g(\mathbf{s}), f(\mathbf{s}, \mathbf{a}_{\text{pi}}), \mathbf{a}_{\text{pl}}), \tag{2}$$

where $\mathbf{a}_{\text{pl}}^* \in \mathbb{N}^2$, same as the pick sub-action.

Similarly, the last parameter sub-action augmented state $S_3 = S \times A_1 \times A_2$ re-uses information from both previous sub-actions. Here, the action-value function Q_2 is taken as input to provide information of the place sub-action. In addition, the third augmented state is given the encoding $g(\mathbf{s})$ as well as an encoding of the pick-centred image. The parameter sub-action is thus computed as

$$\mathbf{a}_{\theta}^* = \arg\max_{\mathbf{a}_{\theta}} Q_3(g(\mathbf{s}), f(\mathbf{s}, \mathbf{a}_{\text{pi}}), Q_2(\mathbf{s}_2, \mathbf{a}_{\text{pl}}), \mathbf{a}_{\theta}), \tag{3}$$

where $\mathbf{a}_{\theta}^* \in \mathbb{N}^1$, and the parameter values are defined in a different range of values for each manipulation primitive [13]. Finally, the action is composed by the three sub-actions $\mathbf{a} = \left[\mathbf{a}_{\text{pi}}^*, \mathbf{a}_{\text{pl}}^*, \mathbf{a}_{\theta}^*\right]$ that have been sequentially computed re-using previous information throughout the augmented state space. More details about the training proccedure and the network structure are available in the full paper [13].

III. Experiments

Our experiments evaluate the impact of learning to sequentially optimise parameters such as the height or velocity of pre-defined manipulation primitives in simulation and real-world experiments. We evaluate two quasi-static manipulation primitives: Pick-and-Place (P-n-P) and drag; and one dynamic manipulation primitive. The performance of the proposed QDP method is compared against the Max Value Map (MVP) approach proposed by [4].

A. Optimal and Sub-Optimal Primitive Parameter Values

We start by analysing in simulation the effect of using optimal and sub-optimal velocities and height values for the manipulation primitives. The results, in Figure 3, show the performance of QDP against MVP using as fixed parameter, the median of the proposed values by QDP; as well as an experimentally selected sub-optimal value. In Figure 3 a), the dynamic manipulation primitive velocity has been set to an optimal value of $v_{\text{mid}} = 0.1$, and a sub-optimal value of $v_{\text{mid}} = 0.2$. The results show that using a sub-optimal value for the dynamic manipulation primitive is detrimental, as the performance drops more than 20% for both normal and large cloths. We attribute this poor performance to constantly using a high velocity, as opposed to adapting it, for cloth sizes and configurations that do not require such acceleration to improve their coverage.

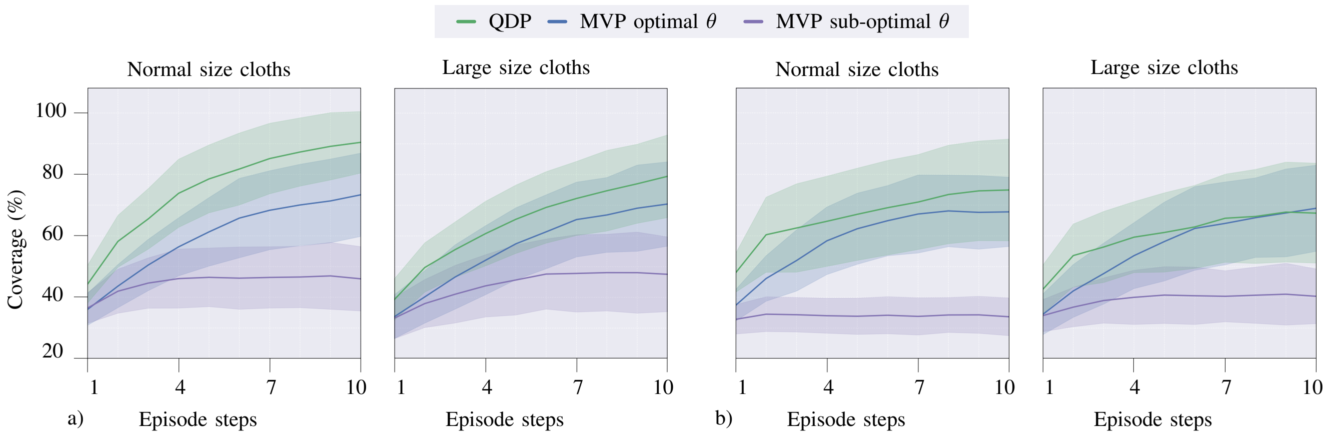

Fig. 3: Quantitative comparison of coverage percentage for unfolding cloth in simulation using the a) Pick-and-Place, and b) Dynamic manipulation primitive. The results compare the proposed QDP (*green*) against MVP using different velocity and height values. The additional parameter values are set as the median of the proposed values by QDP, denoted as MVP optimal θ (*blue*), and a sub-optimal value (*purple*); for normal rectangular cloths, and large rectangular cloths.

The results in Figure 3 b) compare the P-n-P primitive using an optimal height value of $h_\theta = 0.2$ and a suboptimal value of $h_\theta = 0.5$. The suboptimal height value leads to significant performance drops in both normal and large size cloths. This is a result of lifting the cloth from a single point and losing contact with the surface, which results in a crumpled configuration when dropped onto its place location.

By using QDP to adapt the primitives the coverage performance improves at least 10% for most of the cases. We note that the P-n-P performance for large cloths slightly drops compared to the optimal height, which can be a result of the lack of large cloths in the training data set.

B. Real-World Experiments

We evaluate in the real-world three manipulation primitives, transferring the networks in a zero-shot manner. The performance is reported over 3 test episodes, with 10 episode steps each, resulting in at least 30 interactions per cloth, where we discard action steps in which the grasp was unsuccessful and the cloth configuration was not altered by the grasp attempt. Table I shows the results of the normalised coverage improvement after 10 interactions with the cloth. The results on the dynamic manipulation primitive show that QDP can increase the coverage improvement up to 11.99% compared to MVP. In addition, the P-n-P primitive where the optimal height is determined by QDP outperforms all the other manipulation primitives, for both the towel and napkin cloths, increasing the mean coverage more than 4.46% compared to MVP P-n-P, which is the second best primitive. These results show that modifying the velocity and height of the manipulation primitives based on the cloth state is beneficial, following the results achieved in simulation. In addition, the results for each primitive on the chequered rag using QDP show superior performance compared to the baseline. We hypothesise that the transitions of the rag when performing the manipulation are closer to the training data, and thus more in line with the simulation results.

TABLE I: Quantitative results of coverage improvement percentage after 10 interactions with the cloth in the real-world. The results show the performance of QDP and MVP for the the pick-and-place (P-n-P), drag and dynamic primitives.

		Towel	Napkin	Chequered Rag
MVP	P-n-P	24.93 ± 8.64	17.69 ± 11.57	-0.30 ± 4.70
	Drag	-1.21 ± 6.34	3.72 ± 4.02	-6.41 ± 2.96
	Dynamic	9.89 ± 1.16	1.57 ± 1.01	-6.04 ± 8.45
QDP (Ours)	P-n-P	**29.39** ± 16.51	**20.43** ± 11.13	-4.97 ± 3.69
	Drag	-4.58 ± 4.89	3.69 ± 8.56	-1.35 ± 4.22
	Dynamic	9.88 ± 10.50	11.36 ± 4.23	**11.99** ± 11.96

IV. Conclusion

We presented QDP, a novel approach for sequentially choosing parameter values of manipulation primitives for cloth manipulation. While prior work has overlooked the effect of parameters such as the velocity or height of manipulation primitives, the proposed sequential decision process allows a greater variety and complexity of primitives to be used. This variety makes it possible to handle diverse fabric materials and sizes. Our experimental results show that, compared to previous work, QDP can improve up to a 20% coverage in simulation for the task of cloth unfolding. Furthermore, real-world experiments demonstrate the effectiveness of finding the optimal velocity and height for dynamic and quasi-static manipulation primitives.

This work paves the way to a broader range of complex manipulation primitives, eliminating the human effort of fine-tuning or designing primitives, while reducing computational requirements due to the sequential decision process. This gives promise to exploring complex parameterised manipulation skills for shaping other deformable materials such as visco-elastic or elasto-plastic ones, which are present in many industrial and house-hold environments.

REFERENCES

[1] I. Garcia-Camacho, J. Borràs, B. Calli, A. Norton, and G. Alenyà, "Household cloth object set: Fostering benchmarking in deformable object manipulation," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 5866–5873, 2022.

[2] R. P. Joshi, N. Koganti, and T. Shibata, "Robotic cloth manipulation for clothing assistance task using dynamic movement primitives," in *Proceedings of the Advances in Robotics*, ser. AIR '17. New York, NY, USA: Association for Computing Machinery, 2017.

[3] B. Jia, Z. Pan, Z. Hu, J. Pan, and D. Manocha, "Cloth manipulation using random-forest-based imitation learning," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 2086–2093, 2019.

[4] H. Ha and S. Song, "Flingbot: The unreasonable effectiveness of dynamic manipulation for cloth unfolding," in *Conference on Robotic Learning (CoRL)*, 2021.

[5] Y. Avigal, L. Berscheid, T. Asfour, T. Kröger, and K. Goldberg, "Speedfolding: Learning efficient bimanual folding of garments," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 1–8.

[6] Y. Wu, W. Yan, T. Kurutach, L. Pinto, and P. Abbeel, "Learning to Manipulate Deformable Objects without Demonstrations," in *Proceedings of Robotics: Science and Systems*, Corvalis, Oregon, USA, July 2020.

[7] M. T. Mason, *Mechanics of Robotic Manipulation*. Cambridge, MA, USA: MIT Press, 2001.

[8] R. Lee, D. Ward, V. Dasagi, A. Cosgun, J. Leitner, and P. Corke, "Learning arbitrary-goal fabric folding with one hour of real robot experience," in *Proceedings of the 2020 Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, J. Kober, F. Ramos, and C. Tomlin, Eds., vol. 155. PMLR, 16–18 Nov 2021, pp. 2317–2327.

[9] X. Lin, Y. Wang, Z. Huang, and D. Held, "Learning visible connectivity dynamics for cloth smoothing," in *Conference on Robot Learning*, 2021.

[10] Z. Huang, X. Lin, and D. Held, "Mesh-based Dynamics with Occlusion Reasoning for Cloth Manipulation," in *Proceedings of Robotics: Science and Systems*, New York City, NY, USA, June 2022.

[11] A. Canberk, C. Chi, H. Ha, B. Burchfiel, E. Cousineau, S. Feng, and S. Song, "Cloth funnels: Canonicalized-alignment for multi-purpose garment manipulation," *arXiv preprint arXiv:2210.09347*, 2022.

[12] D. Wang, C. Kohler, and R. Platt, "Policy learning in se(3) action spaces," in *Proceedings of the 2020 Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, J. Kober, F. Ramos, and C. Tomlin, Eds., vol. 155. PMLR, 16–18 Nov 2021, pp. 1481–1497.

[13] D. Blanco-Mulero, G. Alcan, F. J. Abu-Dakka, and V. Kyrki, "Qdp: Learning to sequentially optimise quasi-static and dynamic manipulation primitives for robotic cloth manipulation," 2023.