Movement Primitive Diffusion: Learning Gentle Robotic Manipulation of Deformable Objects

Paul Maria Scheikl[1], Nicolas Schreiber[2], Christoph Haas[2], Niklas Freymuth[2], Gerhard Neumann[2], Rudolf Lioutikov[2], and Franziska Mathis-Ullrich[1]

***Abstract*—Policy learning in robot-assisted surgery (RAS) lacks data efficient and versatile methods that exhibit the desired motion quality for delicate surgical interventions. To this end, we introduce Movement Primitive Diffusion (MPD), a novel method for imitation learning (IL) in RAS that focuses on gentle manipulation of deformable objects. The approach combines the versatility of diffusion-based imitation learning (DIL) with the high-quality motion generation capabilities of Probabilistic Dynamic Movement Primitives (ProDMPs). This combination enables MPD to achieve gentle manipulation of deformable objects, while maintaining data efficiency critical for RAS applications where demonstration data is scarce. We evaluate MPD across various simulated and real world robotic tasks on both state and image observations. MPD outperforms state-of-the-art DIL methods in success rate, motion quality, and data efficiency. Project page: scheiklp.github.io/movement-primitive-diffusion**

***Index Terms*—Surgical Robotics: Laparoscopy; Imitation Learning; Score-based Diffusion Policies; Movement Primitives**

I. INTRODUCTION

ADVANCING the level of autonomy in Robot-Assisted Surgery (RAS) requires novel methods for training policies that satisfy the special requirements of surgical applications. RAS requires the policies to exhibit gentle manipulation of delicate tissue and perform with limited data as human demonstrations are costly. Additionally, human behavior is inherently multimodal [1], covering multiple distinct strategies for solving the same task. Imitation Learning (IL) methods that are unable to represent multimodal behavior may exhibit harmful behavior through mode averaging that is unacceptable in surgical settings, *e.g.*, by averaging over two distinct strategies of dissecting tissue and thus damaging healthy tissue. Diffusion-based Imitation Learning (DIL) has shown to perform well on high-dimensional action spaces, generate multimodal behaviors, and exhibit strong training stability [2], [3], making it a promising framework for application in RAS.

DIL methods train large neural networks to iteratively denoise action sequences drawn from a prior Gaussian distribution to generate motion conditioned on observations. We propose to add temporal correlations between actions during motion generation by utilizing Movement Primitives (MPs) to address both gentle manipulation of deformable objects and data efficiency in DIL. In other methods, neural networks output actions sequences directly [2], [3]. In our proposed method, Movement Primitive Diffusion (MPD), the neural network outputs parameters of a MP that encode a denoised action sequence. These parameters are decoded into smooth position trajectories to enable gentle manipulation of deformable objects.

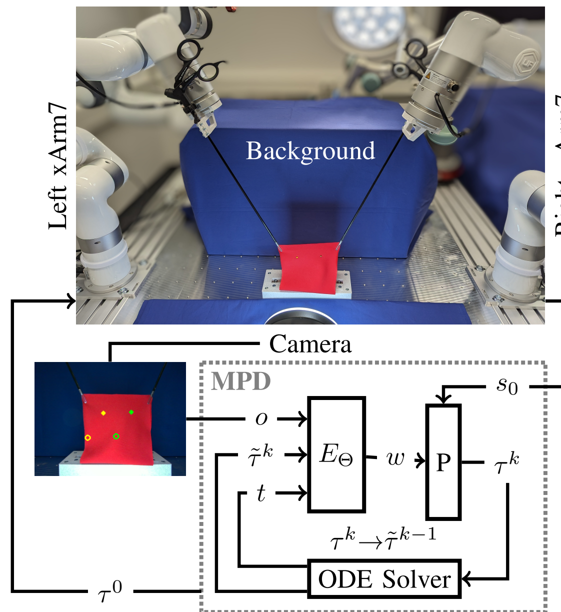

Fig. 1: Schematic for action sequence generation with MPD for bimanual tissue manipulation. Observations o and initial values s_0 for position and velocity are captured on the bimanual robotic setup. An ODE solver solves the Probability Flow ODE with learnable model E_Θ and ProDMP P by iteratively denoising an action sequence $\tilde{\tau}^k$ for diffusion step k and respective noise level t. The final denoised action sequence τ^0 is executed on the robots.

Leveraging both MPs and DIL, MPD increases data efficiency and generates smooth action sequences suitable for gentle deformable object manipulation in RAS. MPD outperforms state-of-the art DIL methods in terms of success rate, motion quality, and required training data. Further, it integrates a modern diffusion framework for real-time inference and guarantees initial conditions for position and velocity. MPD can be trained on both states and raw RGB image observations, making it applicable for RAS where images are the only readily available source of information [4]. Figure 1 illustrates the action sequence generation of MPD.

This work was supported by the Erlangen National High Performance Computing Center funded by the German Research Foundation (DFG), the HoreKa supercomputer funded by the Ministry of Science, Research and the Arts Baden-Württemberg and by the Federal Ministry of Education and Research, and the DFG – 448648559.

[1] P. M. Scheikl and F. Mathis-Ullrich are with the Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-University Erlangen-Nürnberg, 91054 Erlangen, Germany.
Corresponding author: franziska.mathis-ullrich@fau.de
[2] N. Schreiber, C. Haas, N. Freymuth, G. Neumann, and R. Lioutikov are with the Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany.

As baselines, Diffusion Policy [2] and BESO [5] present the current state of the art in robotic DIL. Both methods iteratively denoise action sequence samples to generate motion, conditioned on observations. Both works evaluate their methods against multiple state-of-the-art IL methods and find that DIL methods outperform non-diffusion-based methods in terms of success rate, and excel in learning multimodal behaviors. In this work, we investigate Diffusion Policy and BESO under the requirements of RAS and show that MPD addresses the shortcomings of these methods.

II. Methods

Problem Formulation: We predict action sequences $\tau = (\tau_i)_{i=0...n}$ that consist of k-dimensional values $\tau_i \in \mathbb{R}^k$ for the next n time steps relative to the current time. Depending on the task, the task space consists of k actuation Degree of Freedoms (DoFs) such as grasper articulation, and rotations and translations of surgical instruments in relation to a remote center of motion. The action sequences are predicted based on observations $o = (o_j)_{j=-m+1...0}$ from the previous m time steps. We follow an IL approach and train our models on a dataset $\mathcal{D}$ of human demonstrations d. Each demonstration is a sequence $(\tau_i, o_i)_{i=0..N}$ over one full task execution with N time steps. For training, the demonstrations are split into action and observation sub-sequences of lengths n and m, respectively. We focus on action sequences of length $1 < n < N$ instead of single actions or full trajectories to balance compounding errors and online adaptability in real-world scenarios.

Score-based Diffusion Models: Diffusion Policy [2] reverses the diffusion process at discrete noise levels, relying on probabilistic modeling of the process as a Markov chain. In contrast, BESO [3] builds on the Score-Based Generative Model (SGM) framework that describes the diffusion process as a time-continuous Stochastic Differential Equation (SDE) and learns the gradient of the log probability density, *i.e.*, the score, of the data distribution. This framework allows for modular selection of, e.g., the noise schedule and numerical solver [6] and is often computationally cheaper [5]. Both methods can represent multimodal distributions of action sequences [2], [3]. MPD adopts the SGM framework and closely follows the conventions proposed in [6] and [5].

Movement Primitive Diffusion: We propose MPD, which combines the advantages of SGMs and Probabilistic Dynamic Movement Primitives (ProDMPs). In MPD, the inner model consists of a trainable model E_Θ that outputs a weight vector $\boldsymbol{w}$. Combined with initial values s_0 for position and velocity, $\boldsymbol{w}$ is decoded into an action sequence τ using a ProDMP. Conceptually, the model E_Θ denoises an action sequence conditioned on observations and maps it into the ProDMP weight space. The ProDMP decodes the denoised weights back into action sequence space. The architecture of MPD is illustrated in Figure 1.

Utilizing ProDMP, MPD generates smooth, high-frequency action sequences with guaranteed initial conditions for position and velocity. ProDMP further helps modeling temporal correlations between actions, which increases data efficiency and generates motions that are suitable for gentle manipulation of deformable objects.

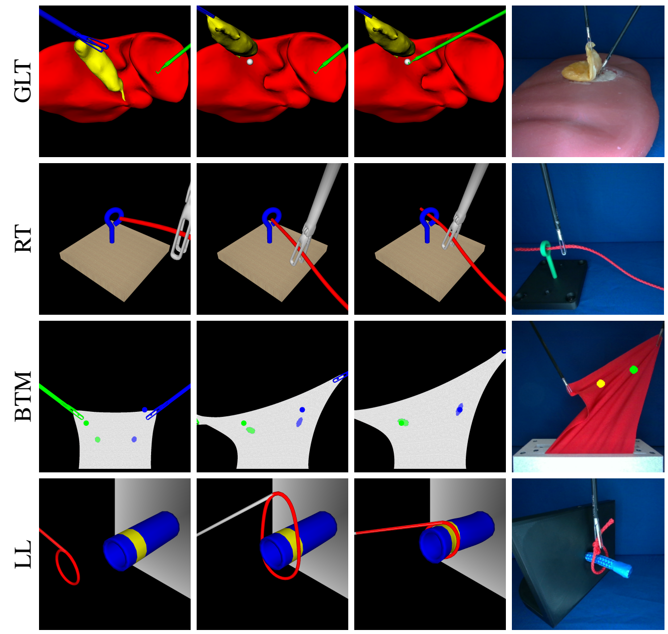

Fig. 2: Start, intermediate, and end state of the tasks in simulation. The final column shows the respective real world experiment. Grasp Lift Touch (GLT) requires sequential collaboration between instruments, Rope Threading (RT) and Ligating Loop (LL) depend on accurate alignment deformable ropes, and Bimanual Tissue Manipulation (BTM) requires concurrent collaboration between instruments to control the shape of a deformable tissue.

III. Experiments

Tasks: In our experiments, we evaluate how well MPD aligns with requirements for application in RAS, based on success rate and motion quality. MPD is evaluated on four different simulated LapGym [7] tasks and their respective real-world robotic setups, illustrated in Figure 2. The tasks represent different types of motion such as cooperation of instruments, grasping, and deformable object manipulation, all of which are crucial for successful application in RAS.

In Grasp Lift Touch (GLT), two laparoscopic instruments are controlled to successively grasp and lift a gallbladder, and touch a target point positioned below the bladder with an electrocautery hook. The task requires sequential instrument coordination as well as grasping and manipulation of deformable objects. Three additional tasks, specifically Bimanual Tissue Manipulation (BTM), Rope Threading (RT), and Ligating Loop (LL), are employed to assess the capability of MPD in controlling the overall shape of a deformable object, following waypoints, and indirectly manipulating a deformable object, respectively. MPD is evaluated against three baselines, namely BESO [3], and two variants of Diffusion Policy [2], DP-C and DP-T. BESO and DP-T are based on transformers, while DP-C employs a 1D temporal CNN model architecture.

Motion Metrics: Applications in RAS require specific motion behaviors in addition to raw success rate of task completion. Tissue acceleration quantifies surgical performance for gentle manipulation of delicate tissue [8] and should be minimized to reduce risk of tissue damage. We quantify instrument

TABLE I: Success rate mean and standard deviation across all simulation tasks based on 5 trained models and 100 rollouts. The best method is **bold**, the second best underlined.

	GLT	RT	BTM	LL	Average
			State Observations		
BESO	68.4 (6.8)	90.4 (1.2)	88.2 (3.9)	83.2 (4.8)	82.55
DP-T	**100** (0.0)	89.6 (1.9)	98.4 (0.5)	**100** (0.0)	97.00
DP-C	**100** (0.0)	82.0 (2.3)	93.2 (1.9)	**100** (0.0)	93.80
MPD	99.2 (0.7)	**93.8** (1.2)	**99.0** (0.6)	99.6 (0.5)	**97.90**
			Image Observations		
BESO	99.8 (0.4)	66.8 (2.3)	91.4 (2.4)	99.8 (0.4)	89.45
DP-T	**100** (0.0)	76.6 (4.5)	95.8 (0.7)	99.8 (0.4)	93.05
DP-C	**100** (0.0)	**83.2** (1.2)	85.2 (1.3)	**100** (0.0)	92.15
MPD	**100** (0.0)	78.6 (3.4)	**99.0** (1.1)	**100** (0.0)	**94.40**

trajectory smoothness based on minimizing instrument jerk to increase safety and reduce component wear. We further characterize the efficiency of the movements by *instrument energy*, the sum of accelerations over task execution, which should be minimized. The three metrics tissue acceleration, instrument jerk, and instrument energy are evaluated for all tasks, with the following task specific adaptations to represent the metrics. For RT, tissue acceleration is measured as the acceleration of points on the rope. For LL, the instrument consists of a rigid shaft and a deformable loop, so instrument jerk is examined for both parts individually. On the real world tasks, the full state of the tissue is not accessible, so tissue acceleration cannot be measured directly. However, for BTM-RW, marker acceleration is tracked as a surrogate.

IV. RESULTS

Success Rate: Table I reports the success rates for all methods evaluated on the simulation tasks. MPD outperforms the baseline methods with average success rates of $97.9\,\%$ and $94.4\,\%$ for state and image observations. DP-T is the second best, reaching $97.0\,\%$ and $93.05\,\%$, respectively. MPD and DP-T consistently improve over the baselines. Most methods work better on state observations, except for BESO.Yet in total performance, BESO is still outperformed by the other methods on image observations. We evaluate MPD and DP-C on real-world versions of all tasks, omitting the other methods as their generated trajectories are unfit for execution on the real robot without significant post-processing.

Motion Metrics: Figure 3 shows the results of the method on different motion metrics for image observations. The values are normalized by human demonstration data, so values below 1.0 indicate better-than-demonstrator motion quality. The transformer-based methods, namely BESO and DP-T, show high values for all motion quality metrics. Compared to BESO, DP-T differs noticeably across different runs, with min and max values that deviate far from the mean, see *e.g.*, tissue acceleration in Figure 3 b). We report min and max values for the motion metrics instead of the standard deviation to highlight best- and worst-case scenarios for the methods. The results across different runs are more bounded for MPD and DP-C, with min and max values closer to the mean. DP-C and especially MPD achieve gentle motions whose values are much closer to the human demonstrations on all metrics.

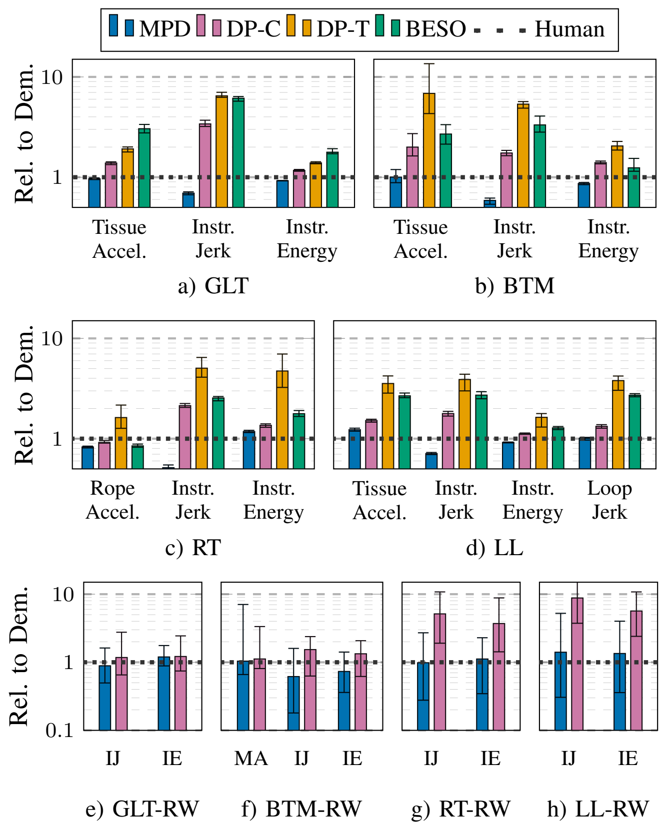

Fig. 3: Motion quality evaluation metrics relative to performance of human demonstrations. The bars show the mean normalized value, with min and max values as error bars. All methods were trained on image observations. MPD consistently outperforms the baselines and exhibits even less Instrument Jerk than the human demonstrations. Real world (RW) tasks e) to h) lack the Tissue Acceleration metric. BTM-RW uses the Marker Acceleration (MA) as a proxy for Tissue Acceleration. Abbreviations: Instrument Jerk (IJ), Instrument Energy (IE).

MPD performs best across all metrics. Using ProDMPs, MPD generates motions that have noticeably less instrument jerk, even compared to human demonstrations.

V. CONCLUSION

This work proposes MPD, a novel method for learning gentle robotic manipulation of deformable objects for RAS. MPD combines the versatility of DIL with the motion quality of ProDMPs, facilitating gentle manipulation of deformable objects and data efficient training that are crucial for surgical applications. The experiments show the superior performance of MPD over traditional DIL methods in terms of success rate, data efficiency, and motion quality. The integration of ProDMPs allows real-time generation of smooth, high-frequency action sequences with guaranteed initial conditions, as required for application in real-world robotic scenarios.

In summary, MPD's ability to learn accurate, high-quality motion from limited data makes it a promising approach for application in autonomous and semi-autonomous surgical systems. Future research may explore MPD in more diverse surgical scenarios and its integration with other surgical technologies, further pushing the boundaries of robotic assistance in complex medical procedures.

REFERENCES

[1] D. Blessing, *et al.*, "Information maximizing curriculum: A curriculum-based approach for learning versatile skills," in *Adv. Neural. Inf. Process. Syst.*, 2023.

[2] C. Chi, *et al.*, "Diffusion policy: Visuomotor policy learning via action diffusion," in *Robot.: Sci. Syst.*, 2023.

[3] M. Reuss and R. Lioutikov, "Multimodal diffusion transformer for learning from play," in *Conf. Robot. Learn. Workshop*, 2023.

[4] P. M. Scheikl, *et al.*, "Sim-to-real transfer for visual reinforcement learning of deformable object manipulation for robot-assisted surgery," *IEEE Robot. Autom. Lett.*, vol. 8, no. 2, pp. 560–567, 2023.

[5] M. Reuss, M. Li, X. Jia, and R. Lioutikov, "Goal conditioned imitation learning using score-based diffusion policies," in *Robot.: Sci. Syst.*, 2023.

[6] T. Karras, M. Aittala, T. Aila, and S. Laine, "Elucidating the design space of diffusion-based generative models," in *Adv. Neural. Inf. Process. Syst.*, 2022.

[7] P. M. Scheikl, *et al.*, "Lapgym - an open source framework for reinforcement learning in robot-assisted laparoscopic surgery," *Journal of Machine Learning Research*, vol. 24, no. 368, pp. 1–42, 2023.

[8] T. Sugiyama, *et al.*, "Tissue acceleration as a novel metric for surgical performance during carotid endarterectomy," *Oper. Neurosurg.*, pp. 10–1227, 2022.