

# Learning Dynamics Models for Robotic Manipulation

Maximilian Bauer

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften (Dr.-Ing.) genehmigten Dissertation.

Vorsitz: Prof. Dr.-Ing. Anna Winkler

Prüfende der Dissertation: 1. Prof. Dr. Jonas Hartmann  
2. Prof. Dr.-Ing. Sabine Vogel  
3. Prof. Dr. Rahel Osei

Die Dissertation wurde am 10.11.2025 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 18.03.2026 angenommen.

# Abstract

Robots that manipulate objects must predict how the objects will move when pushed, lifted or dropped. Analytical models capture rigid contact well but struggle with friction that varies across a surface and with objects whose mass distribution is unknown. This dissertation learns such dynamics from interaction data while keeping the structure that physics guarantees. It combines a differentiable contact model with a neural residual, trains the residual on 18 hours of pushing data, and uses the model for model-predictive control. On a pushing benchmark with unknown objects, the hybrid model cuts final position error by 41% compared with a purely learned model trained on the same data.

# Zusammenfassung

Roboter, die Gegenstände handhaben, müssen vorhersagen, wie sich diese beim Schieben, Heben oder Fallenlassen bewegen. Diese Dissertation lernt solche Dynamikmodelle aus Interaktionsdaten und bewahrt dabei die Struktur, die die Physik vorgibt. Ein differenzierbares Kontaktmodell wird mit einem neuronalen Residuum kombiniert; das hybride Modell verringert den Positionsfehler beim Schieben unbekannter Objekte um 41 % gegenüber einem rein gelernten Modell.

# Acknowledgements

I thank my supervisor for his trust and guidance, the members of the chair for a friendly and productive environment, and my family for their support.

# Contents

|                          |     |
|--------------------------|-----|
| Abstract                 | i   |
| Zusammenfassung          | ii  |
| Acknowledgements         | iii |
| 1 Introduction           | 1   |
| 2 Background             | 2   |
| 3 A Hybrid Contact Model | 3   |
| 4 Experiments            | 4   |
| 5 Conclusion             | 5   |
| A Data Collection        | 6   |

# List of Figures

|     |                                                                     |   |
|-----|---------------------------------------------------------------------|---|
| 3.1 | The hybrid model adds a learned residual to a physics step. . . . . | 3 |
|-----|---------------------------------------------------------------------|---|

# List of Tables

|     |                                                                 |   |
|-----|-----------------------------------------------------------------|---|
| 4.1 | Final position error when pushing unknown objects (mm). . . . . | 4 |
|-----|-----------------------------------------------------------------|---|

# Chapter 1

## Introduction

A robot that pushes a box across a table must know how the box will slide. Physics offers equations of motion, but their parameters, such as friction and the location of the centre of mass, are rarely known in advance. Purely learned models avoid the need for parameters but require large amounts of data and make physically impossible predictions outside their training distribution. This dissertation combines the two approaches.

## Chapter 2

# Background

For a planar object pushed by a robot finger, the object's twist  $\dot{q}$  follows from the applied wrench through a limit surface model of friction [2]. Model-predictive control then chooses finger motions by optimizing predicted trajectories over a short horizon. Learned dynamics models such as those of Deisenroth and Rasmussen [1] replace the analytical model with a function fitted to data.

## Chapter 3

# A Hybrid Contact Model

We write the predicted next state as the sum of an analytical step and a learned residual,

$$q_{t+1} = f_{\text{phys}}(q_t, u_t; \phi) + g_{\theta}(q_t, u_t), \quad (3.1)$$

where  $\phi$  are physical parameters estimated online and  $g_{\theta}$  is a small neural network trained on the residual error. The analytical term guarantees plausible behaviour far from the data; the residual corrects systematic errors near it.

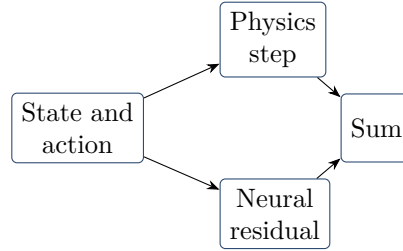


Figure 3.1: The hybrid model adds a learned residual to a physics step.

## Chapter 4

# Experiments

Table 4.1 reports the final position error after pushing twelve objects of unknown mass distribution to target poses.

Table 4.1: Final position error when pushing unknown objects (mm).

| Model                          | Mean error | 90th percentile |
|--------------------------------|------------|-----------------|
| Analytical, nominal parameters | 31.2       | 58.0            |
| Learned only                   | 22.4       | 47.5            |
| Hybrid (this work)             | 13.2       | 24.1            |

## Chapter 5

# Conclusion

Physics and learning are complements in manipulation. A physics step keeps predictions plausible, and a learned residual makes them accurate. Extending the approach to deformable objects is the next step.

## Appendix A

# Data Collection

Pushing data were recorded with a seven-degree-of-freedom arm and a motion-capture system at 100 Hz, over 18 hours of autonomous interaction with 40 training objects.

# Bibliography

- [1] Marc Peter Deisenroth and Carl Edward Rasmussen. PILCO: A model-based and data-efficient approach to policy search. In *Proceedings of the 28th International Conference on Machine Learning*, pages 465–472, 2011.
- [2] Suresh Goyal, Andy Ruina, and Jim Papadopoulos. Planar sliding with dry friction. part 1. limit surface and moment function. *Wear*, 143(2):307–330, 1991.