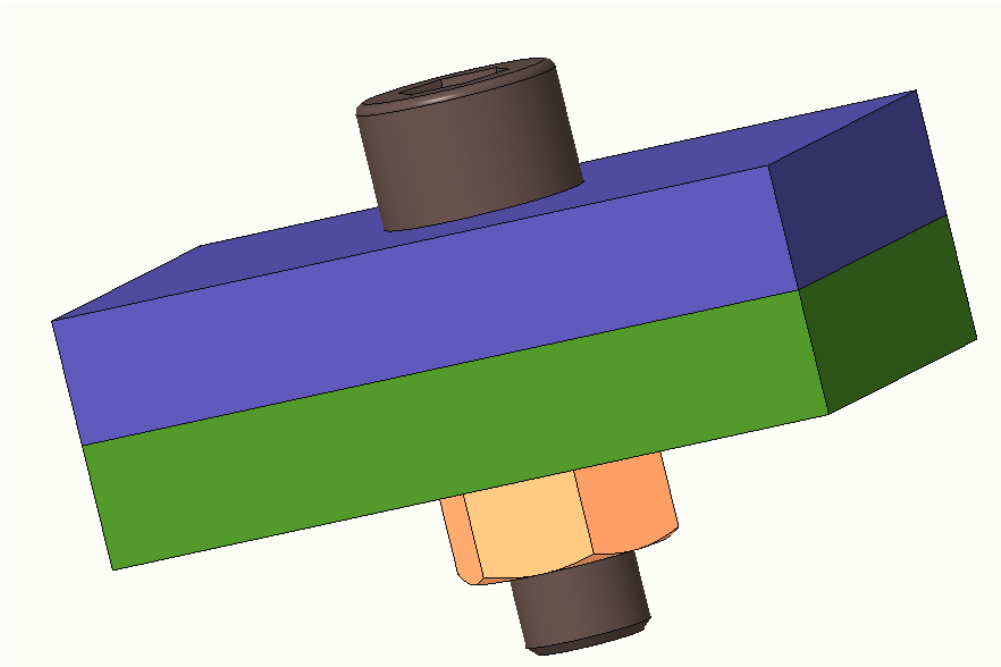


Model-Based Estimation of Clamp Force in Bolted-Joint Tightening

Master's thesis in Systems, Control and Mechatronics

ATHREYA TALLANJE

DEPARTMENT OF ELECTRICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

Master's thesis 2026

Model-Based Estimation of Clamp Force in Bolted-Joint Tightening

Athreya Tallanje

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Division of Automation and Mechatronics
Chalmers University of Technology
Gothenburg, Sweden 2026

Model-Based Estimation of Clamp Force in Bolted-Joint Tightening

Athreya Tallanje

Supervisor: Nils Dressler, Atlas Copco Industrial Technique
Examiner: Balázs Adam Kulcsár, Department of Electrical Engineering

Master's Thesis 2026
Department of Electrical Engineering
Division of Automation and Mechatronics
Chalmers University of Technology
SE-412 96 Gothenburg

Cover: Bolted Joint

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2026

Model-Based Estimation of Clamp Force in Bolted-Joint Tightening
Athreya Tallanje
Department of Electrical Engineering
Division of Automation and Mechatronics
Chalmers University of Technology

Abstract

Bolted joints are widely used in mechanical assemblies, and their reliability depends strongly on the clamp force generated during tightening. Direct measurement of clamp force in industrial applications is often impractical. Conventional torque-based tightening methods estimate clamp force by assuming a constant coefficient of friction (CoF). However, the accuracy of this assumption is influenced by surface conditions, lubrication, tightening speed, and other factors. Therefore, accurate clamp force estimation remains a challenging problem. Previous Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF)-based approaches achieved estimation errors of approximately 3–5%, but were found to be sensitive to initial state assumptions. In addition, the coupling between clamp force and CoF can produce similar torque outputs for different state values, resulting in non-unique estimates.

This thesis focuses on developing system formulations for tightening in the elastic region. The main states considered are clamp force and CoF, which are strongly coupled through the torque expression. To analyze the resulting non-uniqueness in state estimation, both local and global observability perspectives are considered. Multiple observer-based estimation approaches are then evaluated, including an LMI-based Luenberger observer using an LPV formulation, an output-inverted EKF, and a Rao-Blackwellized Particle Filter (RBPF). The EKF and RBPF formulations reduce the dimensionality of the directly estimated state space.

The results indicate that the system is theoretically observable from the local observability analysis, but the near-marginal values of the observability matrix suggest limitations in state estimation. The output-trajectory-based observability study further shows trajectory-dependent regions of practical unobservability, making it difficult to generalize a single unobservable region across different tightenings. The LMI-based Luenberger observer was highly sensitive to initial-state errors and relied strongly on the motion model, whereas the output-inverted EKF was mainly limited by derivative-induced noise amplification in the measurement model. The RBPF provided improved estimations by maintaining multiple hypotheses, but its performance depended strongly on both initialization and the nonlinearity of the output trajectory. Overall, the study shows that reliable reconstruction of clamp force and CoF is fundamentally constrained by weak observability, low effective CoF contribution to the measured output, and model mismatch under nonlinear tightening conditions.

Keywords: Bolted Joints, Tightening, Observability, Estimation, Observer, Leunberger, Dynamic Inversion, Kalman Filter, State Coupling, Rao-Blackwellized Filter

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Nils Dressler, for providing me the opportunity to carry out this thesis at Atlas Copco, and for the guidance, encouragement, and valuable feedback provided throughout this work. His continuous support, together with his background in tightening technology and research, has played an important role in shaping this thesis.

I am also grateful to my examiner for the constructive feedback on the approaches, insightful suggestions, and academic guidance provided during this study. The support with the methodological direction have been invaluable throughout the thesis process.

My sincere thanks go to Erik Persson, my manager at Atlas Copco, for giving me the opportunity to conduct this thesis in an industrial setting and for providing the support and resources needed to complete this work.

I would also like to thank the team members at Atlas Copco who contributed through technical discussions, ideas, and encouragement during the thesis process. Their support has been greatly appreciated. I am equally thankful to Atlas Copco for providing the industrial environment and opportunity to carry out this work.

Finally, I would like to express my deepest gratitude to my family members and friends for their constant support, patience, and belief in me throughout this journey. Their encouragement has been a constant source of strength and motivation.

I would like to acknowledge that, AI tools were used solely to improve the grammar, sentence structure, and readability of the report. They were not used to generate scientific ideas, technical analysis, results, or conclusions, unless otherwise explicitly stated. Responsibility for the content of this thesis remains entirely with the author.

Athreya Tallanje, Gothenburg, June 2026

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

COF	Coefficient of Friction
EKF	Extended Kalman Filter
GPR	Gaussian Process Regression
RBF	Radial Basis Function
LPV	Linear Parameter Varying
LMI	Linear Matrix Inequality
CNN	Convolutional Neural Network
RBPF	Rao-Blackwellized Particle Filter

Nomenclature

Below is the nomenclature of indices, sets, parameters, and variables that have been used throughout this thesis.

Indices

i	Indices for particles
k	Index for time step in discrete time
t	Time in continuous time

Parameters

Δt	Time discretization step (sampling interval)
Δ_{bar}	Upper bound on uncertainty in the output matrix
ℓ	Length scale of the RBF kernel
ϵ^{true}	Threshold used to define practical indistinguishability
$\kappa(T, T')$	Covariance kernel in Gaussian Process Regression
ρ	Scheduling parameter in the LPV formulation
σ_f^2	Signal variance in the RBF kernel
σ_n^2	Observation noise variance in Gaussian Process Regression
$\sigma_{K_j}^{\text{rel}}$	Relative uncertainty in the joint stiffness
$\mathbf{C}(\rho)$	Parameter-dependent output matrix
$\mathbf{L}$	Observer gain matrix
$\mathbf{P}$	Lyapunov matrix / error covariance matrix
$\mathbf{Z}$	Auxiliary decision variable in the LMI formulation
C_f	Friction-related coefficient in the torque expression
C_p	Pitch-related coefficient in the torque expression
K_j	Effective angular joint stiffness

$K_{j,\text{avg}}$	Average joint stiffness used in estimation
$m(T)$	Mean function in Gaussian Process Regression
N_p	Number of particles in the RBPF
Q_α	Process noise covariance associated with α
Q_μ	Process noise covariance associated with CoF particles
Q_{F_c}	Process noise covariance associated with clamp force
R	Measurement noise covariance in the RBPF
R_y	Measurement noise covariance associated with y_k
R_z	Measurement noise covariance associated with z_k

Variables

α	Auxiliary scalar state used in the output-inverted EKF
ΔT	Discrete-time increment in torque
$\mathcal{O}$	Observability matrix
$\mathcal{X}_u$	Practical unobservable region in the state space
$\hat{F}_{c,k}$	Estimated clamp force at time step k
$\hat{\mu}_k$	Estimated coefficient of friction at time step k
μ	Coefficient of friction (CoF)
$\mu_k^{(i)}$	CoF particle for particle i at time step k
ω	Angular speed of the fastener
$\mathbf{T}$	Vector of torque training inputs in GPR
$\mathbf{u}$	Input vector
$\mathbf{x}$	State vector
$\mathbf{y}$	Measurement/output vector
$\boldsymbol{\mu}$	Vector of CoF training outputs in GPR
$f(T)$	Latent function in Gaussian Process Regression
F_c	Clamp force
$F_{c,k}^{(i)}$	Clamp-force estimate associated with particle i
$h(\alpha)$	Nonlinear measurement function in the EKF formulation
K_{**}	Covariance matrix of the test data in GPR
K_{N*}	Cross-covariance matrix between training and test data in GPR
K_{NN}	Covariance matrix of the training data in GPR
N_{eff}	Effective sample size in the RBPF

$P_{F,k}^{(i)}$	Conditional variance of clamp force for particle i
T	Torque
$\dot{T}$	Time derivative of torque
$\dot{T}$	Time derivative of torque
v_k	Measurement noise at time step k
v_k^y	Measurement noise associated with output y_k
v_k^z	Measurement noise associated with derived measurement z_k
w_k	Process noise at time step k
w_k^α	Process noise associated with α
w_k^μ	Process noise associated with CoF dynamics
$w_k^{F_c}$	Process noise associated with clamp-force dynamics
$\mathrm{w}_k^{(i)}$	Weight of particle i at time step k
y_k	Measured output at time step k
z_k	Derived measurement used in the EKF formulation

Contents

List of Figures

List of Tables

1

Introduction

1.1 Bolted joints

Bolted joints are one of the most widely used methods for joining mechanical components due to their simplicity, reliability, and ease of assembly and disassembly. The clamp force plays a critical role in the structural integrity of the bolted joints. Clamp force can be defined as the compressive force acting between the mating components applied by the bolt. This force helps in ensuring that the joint members stay in continuous contact, while allowing external loads to be transmitted through the contact surfaces, preventing relative motion and loosening of the fastener. In practice, the clamp force is induced by applying torque to the bolt, which results in a compressive force acting on the joint members. This clamp force results in the tensile elongation of the bolt, and hence results in a compressive reaction force on the joint members. However, not all of the applied torque is converted into useful clamp force, a large portion of it is dissipated due to friction. This frictional loss occurs primarily at two interfaces: the contact between the threads of the bolt and nut/internal threads and the interface between the bolt underhead and the mating surface. The associated terms are thread friction loss for the friction between the threads and bearing loss which addresses the other surfaces in contact.

Bolt tightening can be carried out using several control strategies, such as the torque-controlled, angle-controlled, and yield-controlled tightening, each with its own advantages and limitations. The simplest and the most commonly used strategy is torque-controlled tightening. In this process the tightening is carried out until a predefined torque is applied to the bolt. This predefined torque is determined based on the assumed constant friction conditions and predefined design standards. Although these values are arrived based on the empirical knowledge and testing data, due to the unpredictable nature of friction, the actual Coefficient of Friction (CoF) can be different. This could be due to the facts such as the surface conditions, lubrication, tightening speed and many other parameters. As a result the torque controlled tightening often leads to an applied clamp force that may be different from the calculated one and this may impact the reliability and repeatability. Angle-controlled tightening tries to overcome the issue CoF variations. In this method, an initial snug torque is first applied to ensure that a complete contact has been established in the joint. This means all the gaps between the mating surfaces are eliminated. After this snug point is reached, the bolt is rotated through a predefined angle, which is expected to produce a predictable elongation of the bolt and, consequently, a more consistent clamp force. However, this method relies on accurate detection of

the snug point and assumes consistent material behavior and joint stiffness, both of which can introduce additional uncertainties in practical applications.
The present study is restricted to the elastic region. One of the reasons for this is that, within this range, the tightening behavior can still be represented using relatively simple and physically meaningful models. Beyond the elastic region, stiffness variation, plastic deformation, variation in CoF and other irreversible effects introduce significant nonlinearities that reduce repeatability and complicate both modeling and estimation.

1.2 Background

State estimation techniques such as Kalman filters are widely used to estimate the unmeasured states in dynamic systems. The nonlinear variants of this filter, such as Extended Kalman Filter(EKF), Unscented Kalman filter(UKF) and Cubature Kalman Filter (CKF) can be used for a nonlinear dynamic system. In the context of bolt tightening, these techniques can be used to estimate clamp force and the CoF, by using the torque-angle data. However, a major challenge arises because clamp force and CoF are not reflected independently in the measured torque. Instead, both states appear in a coupled manner in the torque expression, meaning that a change in one state can be compensated by a change in the other while producing a very similar torque response. As a result, different combinations of clamp force and CoF may result in the same measured trajectory, making the estimation problem ambiguous and, in some cases, non-unique. This reduces the reliability of observer-based estimation methods and makes the estimates highly sensitive to initialization, measurement noise, and model mismatch.
An attempt to address this issue was presented in [2], where the coefficient of friction (CoF) and clamp force were simultaneously estimated using EKF and UKF for two torque-controlled tightening strategies. A nonlinear dynamic model of the tightening process was formulated by defining the fastener geometry and treating clamp force and CoF as the system states, while torque and angular velocity were used as measurable signals. The study considered both Continuous Drive tightening, in which the angular velocity is kept approximately constant as torque increases, and TurboTight tightening, in which the speed starts high and decreases with increasing torque. The methods were experimentally validated for both strategies. The results showed that both filters significantly improved clamp force estimation compared to conventional torque-controlled methods. In Continuous Drive tightening, both EKF and UKF performed well, whereas under the more nonlinear TurboTight strategy the UKF clearly outperformed the EKF. Overall, the reported clamp force errors were approximately 3–5%.
Further internal investigations conducted at Atlas Copco Industrial Technique[3], building upon the above study, revealed additional challenges associated with the estimation process. It was observed that the performance of the filters is highly sensitive to the initial state assumptions, with accurate initial guesses leading to improved estimation performance, while large deviations in initial conditions resulted in increased estimation errors. Moreover, it was observed that multiple combinations of clamp force and CoF could produce similar torque outputs, indicating

non-uniqueness in the estimation problem. The study also included the evaluation of condition number and this study indicated that the output matrix was extremely ill-conditioned, highlighting potential observability limitations of the system. These observations motivate a deeper investigation into the observability of the tightening process and the development of robust observer-based estimation methods, which form the focus of the present work.

1.2.1 Identified Challenges

Based on the analysis conducted in this work, several challenges in the clamp-force estimation problem were identified.
First is the issue with the states occurring as pairs in the output expression, which was identified in the previous study. An observer may struggle to uniquely reconstruct the states as these are coupled. Second, the dependence between clamp force and CoF implies that variations in CoF directly influence the corresponding clamp-force estimate. Therefore, even when system dynamics are introduced, the ability to distinguish between the effects of these two states on the output may remain limited. Third, an additional issue observed from the experimental data is the stiffness of the joint. Since the focus is towards the tightening in the elastic region, an assumption is that the joint stiffness is constant, however, it was observed that this parameter is not constant, and some of the tightening exhibited a highly nonlinear curve.
These issues indicate that clamp-force estimation is fundamentally constrained not only by the observability but even from the perspective of modeling.

1.3 Thesis Scope

Considering the existing issues, the scope of the thesis is as follows:

1. Develop multiple formulations for the bolt-tightening, specifically within the elastic region.
2. To investigate the coupling between the clamp force and the CoF in the torque expression and analyze the impact of variation of each.
3. Evaluate the observability of the tightening system using local and output trajectory based observability analyses.
4. Design observers that could eliminate the existing issues with the sensitivity to the initial states.

The following are some of the research questions that can be answered with this study:

1. Is it sufficient to address the system with two internal states, which are correlated, with one measurement OR is there a requirement of additional measurement?
2. What are the possibilities and limitations of reconstructing unmeasured internal states that appear multiplicatively in the measurement equation when no prior information about the individual state values is available?
3. To what extent can non-Gaussian estimation methods, such as particle filters and related observer frameworks, reconstruct clamp force under large initial state uncertainties?

4. How does error in the assumed joint stiffness propagate through the estimation framework, and what impact does it have on clamp force estimation quality?

1.4 Thesis Limitations

The present study considers the estimation of clamp force only in the linear elastic region of tightening, since this is the region where knowledge of clamp force is most relevant for joint quality and where the tightening behavior can still be described using relatively simple and physically meaningful models. Beyond this region, significant stiffness variation, plastic deformation, and other irreversible effects introduce nonlinearities that complicate both modeling and estimation. The measurements used from the real-time hardware is available with a filtered version using IIR (Infinite Impulse Response) filter and the same has been used as the true measurements. The study is specific to a particular surface condition, which results in CoF to mostly in a certain range of values. Also, the analyses is carried out for a particular tightening speed. Moreover, every tightening speed has a limited test samples. So, any analyses and the performance metrics will be based on these samples. The current study only focuses on the estimation for tightenings carried out for the first time, eliminating the re-tightenings as the friction could have a different behavior. The proposed method is evaluated for a specific joint configuration, and its applicability to different bolt geometries, materials, or lubrication conditions may require further analyses. The results are limited to the test cases presented in the document, and does not address other possible scenarios.

2

Theoretical Background and Methodology

2.1 System Description

2.1.1 Bolt Tightening process

Bolt tightening consists of different stages: rundown, alignment, elastic, and post-yield, as shown in Fig. 2.1. The rundown stage corresponds to the initial phase of tightening, in which the fastener advances through the existing threads before full bearing contact is established between the under-head (or nut face) and the mating surface. In this stage, the clamp force is essentially zero, and the measured torque is mainly associated with frictional resistance in the threaded region, often referred to as prevailing torque. The prevailing torque may be affected by lubrication conditions, or thread fit and tolerance effects. [4]
The alignment zone is a nonlinear stage of tightening in which contact between the mating components has already begun, but the joint has not yet reached full seating. The behavior in this region is influenced by the process of bringing the mating parts into proper contact. This may involve macroscopic effects such as bolt bending caused by non-parallelism between the bearing surface and the under-head, as well as misalignment between the mating components themselves. In addition to these macroscopic effects, several microscopic phenomena also contribute to the nonlinear behavior. These include local deformation of surface coatings under contact stress, asperity interactions due to surface roughness, and local thread deformations. Throughout this stage, the joint is considered to progressively align itself with the mating surfaces. The alignment zone ends when the joint reaches the snug point, which is the stage in tightening where all initial gaps have been eliminated and full contact has been established at the mating interfaces. Beyond this point, the elastic region begins [5].
In the elastic zone, the bolted joint behaves approximately as a linear elastic system. After complete contact has been established at the snug point, the initial gaps between the mating components are eliminated, and any additional torque or angular displacement results primarily in elastic deformation of the bolt and joint members, rather than relative motion between the components. Moreover, due to the complete contact, the deformation of the bolt and the clamped members follows Hooke's law, which leads to a linear relationship between the applied load and the resulting deformation. So, the clamp force is considered to increase proportional to angular displacement. This proportionality constant is the effective joint stiffness.

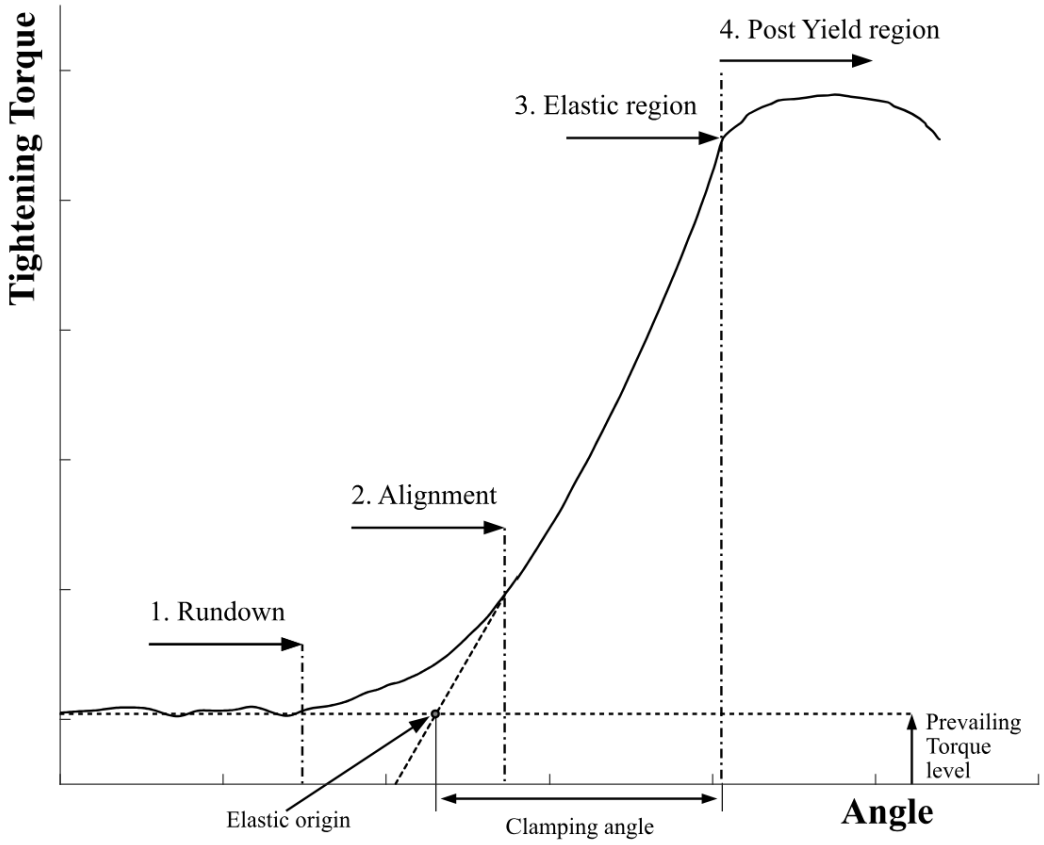


Figure 2.1: Different stages in typical tightening curve of bolted joint. This image is reproduced from [1].

Consequently, the torque–angle curve in this region is often approximated as locally linear, with an approximately constant slope, provided that the joint stiffness is nearly constant and friction varies slowly. This does not imply that most of the applied energy is stored elastically, since a significant portion of the tightening torque is dissipated in friction. Rather, it indicates that the relationship between torque, clamp force, and angular displacement is more consistent in this region, making it more suitable for modeling and estimation. Also, since there are simpler empirical relations that have been already established, so it is considered as the region of interest for the study. This region ends with the material reaching the yield point from where the linearity will not hold and begins the post yield region. The torque at the yield point is impacted by the CoF, an increase in friction increases the torque required to bring the bolt to yield also increases as observed in [5]. This point may be crucial to define the end of estimation.
The post-yield region occurs at the yield point of the material where the material plasticity and nonlinear stiffness dominates. The behavior of the joint becomes significantly complex, making it challenging for modeling.
It can be understood that the bolt tightening can be viewed as a dynamic system with variables such as clamp force, torque, angular change in tightening.
Accurate identification of the snug point is essential for distinguishing the onset of

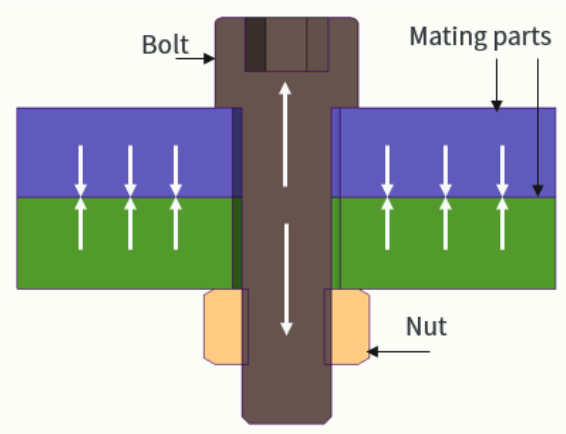


Figure 2.2: Cross-sectional view of standard bolted joint with the representation of the compressive and tensile forces

the elastic region. A study addressing the estimation and detection of the snug point was presented by Zhang [6], where data-driven and hybrid approaches were investigated for identifying the snug location in torque–angle traces. In this work, torque–angle signals were processed and resampled to a fixed length representation, and features such as torque, angle, and change between consecutive samples of torque and angle were extracted. A one-dimensional Convolutional Neural Network (CNN) was employed both to classify the presence of a snug point and to predict its location within the tightening trace. Additionally, a hybrid framework combining the CNN with an existing rule-based algorithm was proposed to incorporate prior mechanical knowledge. Experimental results demonstrated that the hybrid approach improved snug detection accuracy compared to purely rule-based methods, highlighting the potential of combining data-driven techniques with physics-informed heuristics for identifying critical tightening events. This can be a useful way to determine the initial start of linearity, although it may be expensive from the computational sense.

The relevance of snug-point detection becomes evident when examining the experimental tightening data in the present study. The torque-angle trajectories for multiple samples at the same tightening speed are shown in Fig. 2.3. These are carried out for a particular joint, with a M10 bolt of 1.5mm pitch. The fasteners are coated with the Zinc–Nickel alloy coating.

A key observation from these plots is that the beginning of the elastic region does not occur at the same torque or angle level across all samples. This indicates that the transition from the seating/alignment phase to the elastic deformation phase varies between tightenings, making the identification of the exact onset of linearity necessary. Another important observation concerns the slope of the torque-angle trajectory in the elastic region. The slope differs from sample to sample, which may be caused by variations in CoF and/or changes in effective stiffness. In particular, the trajectory of Sample-5 does not exhibit a clearly defined linear region.

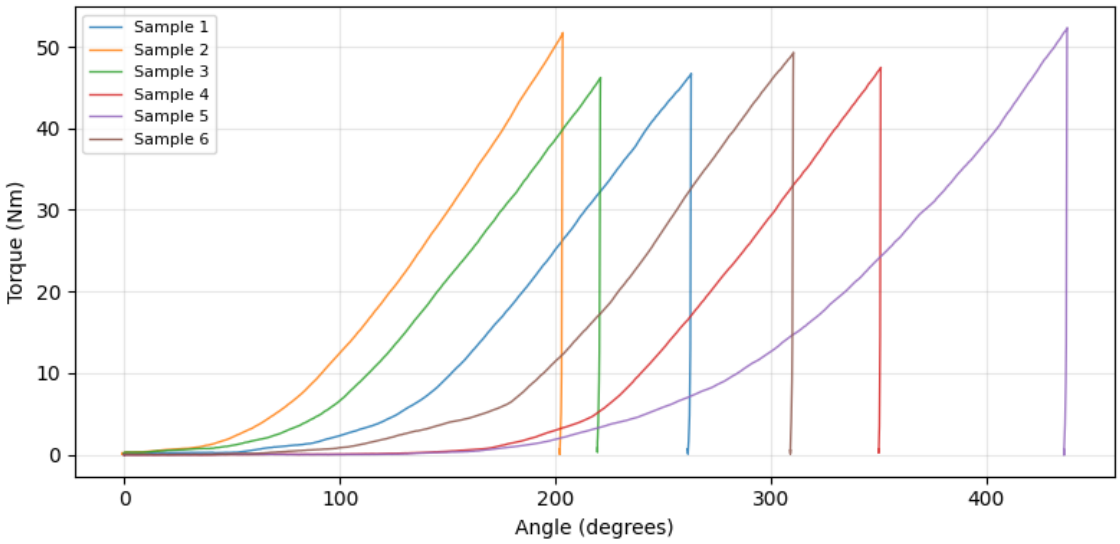


Figure 2.3: Torque-angle plot for bolted joint tightening based on the measurements

2.1.2 Tightening Variables

Coefficient of Friction

As discussed in the previous chapter, there are two locations which are considered to contribute to the frictional loss. The first one is the under-head region (μ_h), which is the friction due to the contact between the under-head of the bolt (or nut) and the bearing surface. The second location is at the interface of the threads (μ_{th}).

The magnitude of CoF at these interfaces is influenced by several factors, which include the surface roughness, material properties, lubrication, presence of external dust/abrasives and coating conditions. It has been observed by Kumar [1] that hard coating results in increased and more variable friction. Whereas softer coatings, behave more as solid lubricants, resulting in stable and reduced CoF values. Moreover, the cleanliness of the surface apart from the obvious surface roughness played a major role, as it can create unintended contact and resistance.

The coefficient of friction (CoF) is also affected by the rotational speed of the bolt during tightening. This relationship has been reported in an internal technical report at Atlas Copco [3], as well as in the studies by Kumar [1] and Blom [7]. It was observed that lower tightening speeds tend to produce higher CoF values, which is commonly attributed to increased adhesion effects and prolonged material interaction at the contact interfaces. As the tightening speed increases, the CoF tends to decrease or reach a more stable level, since the shorter contact duration reduces adhesion and material transfer between the mating surfaces.

The technical study at Atlas Copco [3], which includes the analysis of variation in the friction with the changes in the clamp force and the speed, indicated that the relation between the clamp force and the friction was nearly linear with a very low variance when the speed was around 60-100 RPM. However, at speeds below 60 RPM, the friction data becomes significantly more scattered, indicating higher variability and reduced predictability of the tightening process. This highlights

the importance of selecting appropriate operating conditions for both modeling and experimental analysis. It also indicates that a constant tightening speed can lead to relatively predictable CoF values.
In the present study, experimental tightening data is analyzed to understand the behavior of CoF under a fixed rotational speed. The observed trajectories show noticeable variation across multiple tightening samples, even under similar operating conditions. A detailed discussion is presented in the following section. This indicates that CoF cannot be treated as a constant parameter and motivates the need for estimation and modeling approaches that can account for its variability.

Clamp force and Stiffness
Clamp force, as defined earlier, can be described as the compressive force exerted by the bolt on the mating components. This force is generated by the tensile elongation of the bolt during tightening. From a mechanical perspective, the bolted joint can be modeled as an equivalent elastic system in which the bolt undergoes axial elongation while the mating components undergo compression.
This behavior is commonly represented using a spring model, where the bolt acts as a tensile spring characterized by its axial stiffness, and the mating components act as a compressive spring whose stiffness depends on material properties, geometry, and contact conditions. Since both elements carry the same load but deform differently, the overall joint behavior can be approximated by two springs arranged in series [8]. This means that the elastic portion of the applied tightening energy is stored in the equivalent spring system, being shared between tensile strain energy in the bolt and compressive deformation of the joint members. The distribution of this deformation depends on the relative stiffness of the two components: a stiffer bolt leads to more deformation in the joint members, whereas stiffer joint members lead to greater elongation of the bolt.
From this equivalent spring representation, the effective linear stiffness of the joint, K_{jl}, is expressed in [2] as,

$$K_{jl} = \frac{K_b K_m}{K_b + K_m} \tag{2.1}$$

where K_b, K_m are the stiffness of bolt and mating components respectively. The term 'linear' is used to specify that this stiffness is defined within the elastic region, where the deformation follows Hooke's law and it corresponds to a linear deformation.
However, in practical tightenings, the rotation of the fastener is to be considered in the stiffness, meaning the force generated per unit rotation is to be considered. Therefore it is necessary to relate the linear deformation of the joint to the angular domain. This conversion can be done by considering the thread pitch of the bolt. A rotation of the bolt by an angle $\Delta\theta$ produces an axial displacement proportional to the pitch P. In particular, a complete rotation of 360° corresponds to an axial displacement of the bolt equal to P. Therefore, the axial displacement corresponding to an angular increment $\Delta\theta$ is

$$\Delta l = \frac{P}{360}\,\Delta\theta. \tag{2.2}$$

where Δl is the linear displacement.

Using this expression in the force-displacement relation results in an equation for clamp-force and angular displacement:

$$\frac{\Delta F_c}{\Delta \theta} = K_{jl}\frac{P}{360}. \tag{2.3}$$

Defining the effective angular stiffness as $K_j = K_{jl}\frac{P}{360}$, the expression becomes,

$$\frac{\Delta F_c}{\Delta \theta} = K_j \tag{2.4}$$

Generally clamp force can be measured by using a sensor at the interface of the mating components. However, this may not be feasible on every joint and even if it could be, the process becomes expensive. In order to predict the clamp force using an indirect measurement, an alternative approach was evaluated by Dyberg and Troillet [9], where the clamp force was determined by natural frequencies of the joint. This eliminates the estimation of CoF, while just focusing on the stiffness and hence the clamp force. It is based on the concept that a change in the clamp force would result in change in the vibrational characteristics. Contact and non-contact methods were evaluated using accelerometer and a microphone sensor, respectively. The results indicated a correlation between the first natural frequency and clamp force. This study suggests that vibration-based approach can be an indirect method to estimate the clamp force without explicitly considering friction. However, the drawback with this approach is it relies on the offline excitation to be induced and frequency analysis rather than real-time dynamic estimation.

Although the vibration-based approach provides an alternative indirect method for clamp force estimation, the present study considers a different estimation framework based on torque-driven tightening data. In the current study, clamp force measurements are available for the tightening curves shown in figure-2.3. These measurements are not used as the primary estimation signal, but rather as reference values for assessing the quality of the observer-based estimates.
It is clearly evident from the torque trajectory that the slope of the trajectories in the linear region is not same for all the samples. Although the Samples-1-4 and 6 have almost a similar slope the Sample-5 does not seem to have a distinguishable linear region. The slope of clamp force and angle which is the joint stiffness is not the same for all the tightenings.

Torque Expression
In the elastic region of the tightening, the torque is considered to be directly proportional to the clamp force. The applied torque is converted to clamp force with some frictional loss. This has been reported in [1], based on the formulation of [10],

$$T_{applied} = F_c\left(\frac{P}{2\pi} + \frac{d_2}{2\cos(\alpha)}\mu_{th} + \frac{D_b}{2}\mu_h\right) \tag{2.5}$$

where d_2 is pitch diameter of the thread, μ_h is the CoF at the interface of underhead and bearing surface, D_b is the mean diameter of the underhead surface under the

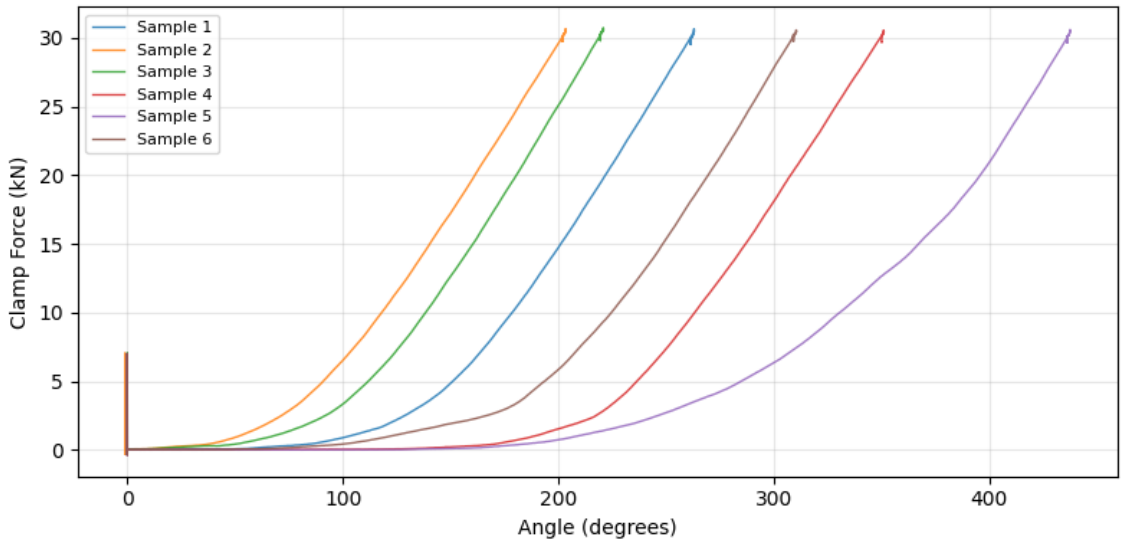


Figure 2.4: Clamp-force–angle plot for complete tightening based on the measured values

fastener head, $cos(\alpha)$ is the half angle thread and μ_{th} is the CoF at the thread contact region.
The physical interpretation of the above expression can be that the applied torque is distributed in two parts:

The term $F_c \frac{P}{2\pi}$ represents the useful torque with axial advancement of the bolt, while contributing to the development clamp force in the joint. This term corresponds to the torque that is required to convert the rotational motion to axial force.

The remaining term, $F_c \left(\frac{d_2}{2\cos(\alpha)} \mu_{th} + \frac{D_{km}}{2} \mu_h \right)$, represents the torque component that is dissipated in overcoming friction. In particular, the first part represents the friction in thread interface and the second part represents the friction under the bolt head or the nut interface. These frictional contributions do not generate additional clamp force but instead reduce the efficiency of torque-to-force conversion.

Simplifying expression (2.5),

$$T = F_c(C_p + C_{th}\mu_{th} + C_b\mu_h) \tag{2.6}$$

where $C_p = \frac{P}{2\pi}$, $C_{th} = \frac{d_2}{2\cos(\alpha)}$, $C_b = \frac{D_{km}}{2}$
A simpler version of this is considered, where $(C_{th}\mu_{th} + C_b\mu_h) \approx C_f\mu$. It has been noted that there can be an error of $1 - 2\%$ due to this approximation [11]

$$T = F_c(C_p + C_f\mu) \tag{2.7}$$

It should be noted that although the torque–force relationship appears linear in the elastic region, the proportionality constant is influenced by the coefficient of friction, which is time-varying. As a result, the relationship is not strictly constant and

introduces uncertainty in clamp force estimation.

Elastic region detection
Since the interest in the current study is in the linear region of tightening, so it is necessary to detect the point where the elastic region begins as mentioned in earlier sections. One way to do this could be by checking the point where the slope of the Torque-angle plot starts to become constant. This can be done easily by looking at the complete tightening curve. However, when this is being estimated during the tightening, it is required to check the slope at every instant and compare it with the previous as there is no information on the complete trajectory. So, the following method is used,

For a particular angle and torque at time instant k, $\theta_k, ; T_k, \quad k = 0, \ldots, N_{overall} - 1$, where $N_{overall}$ is the length of time samples of the overall tightening Point-wise slope is calculated using,

$$b_k \approx \frac{T_k - T_{k-1}}{\theta_k - \theta_{k-1}} \quad \forall\ k > 1 \tag{2.8}$$

It is required to check whether this is consistent over a long interval of tightening as it covers the elastic region,

$$|b_k - b_{k-j}| < \varepsilon, \quad \forall\ j = 1, \ldots, n-1 \tag{2.9}$$

where n is the number of data points chosen to consider the system has entered the elastic region and ε caters for allowable deviation in the linearity. The starting index of elastic region then is $k - j$.
With this information about the system, three formulations have been developed.

2.2 System Formulations

The primary aim of developing new system formulations is to improve the representation of the tightening process by capturing the underlying physical behavior more accurately. At the same time, the formulations are kept as compact as possible by minimizing the number of state variables. This is because a lower-dimensional state representation simplifies the observability analysis, observer design, and numerical implementation, while still having the key dynamics required for estimating the clamp force.
It is evident from the previous work and expression in (2.6) that the major variables are the CoF and clamp force. Although clamp force can be viewed as one of the primary variables, it is implicitly dependent on other system variables such as joint stiffness, angular displacement, tightening speed and other parameters. The study at Atlas Copco [3], considers parameters such as clamp force origin angle and stiffness as some additional state variables apart from the clamp force and CoF.

The CoF is a complex to model and difficult-to-measure parameter especially in the tightening process. It is influenced by multiple factors as mentioned earlier and some of these factors are not directly measurable. One possible approach with this

constraint is to use the previously collected experimental data and develop a data-driven model for the estimation. Measurable parameters can be used as inputs to such models to estimate the CoF.

To further examine the CoF behavior in the present study, it is calculated from the measured torque and clamp force using (2.7). The resulting trajectories in the elastic region are shown in Fig. 2.5.

$$\mu = \left(\frac{T}{figure - F_c} - C_p \right) C_f^{-1} \tag{2.10}$$

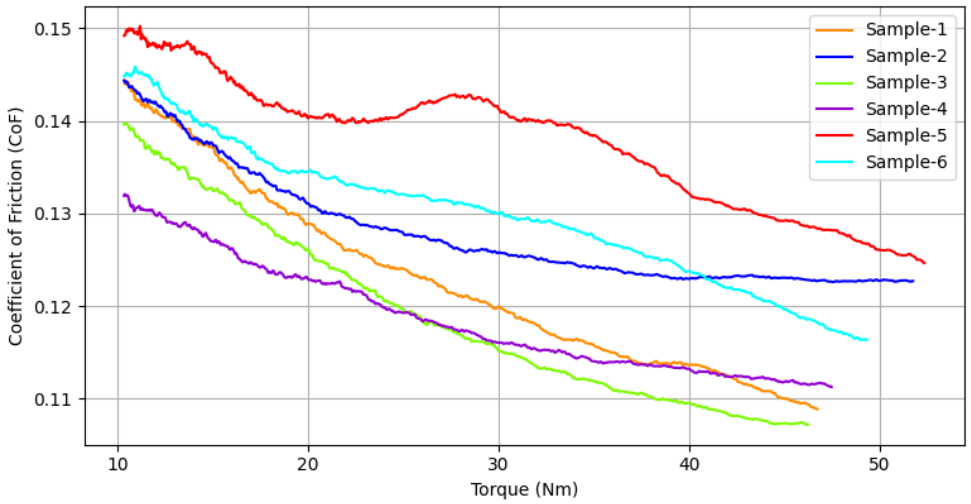


Figure 2.5: Calculated CoF-Torque plot in the elastic region based on the measurements

It can be observed from the above plot that the CoF is not linearly decreasing. Although it seems to decrease from a broader sense, but some of the samples exhibit a local increase. This may be due to multiple reasons, one of which could be the measurement error. The clamp force and torque are used to calculate CoF and these could be associated with a sensor noise and since the magnitude of CoF is small, the impact of noise on the CoF may be high. However these effects are not considered in the current study and all the measured and filtered values are considered as almost the true values.

2.2.1 Data-Driven Estimation of Coefficient of Friction using Gaussian Process Regression

In the current study, the non-constant measurable parameter is torque, so the CoF can be modeled as a function of this measured torque. Since the relationship between torque and CoF is nonlinear at the microscopic level, varies across samples, and the measurements are affected by noise, a flexible probabilistic approach may

is considered for modeling this behavior.

Gaussian Process Regression (GPR) is chosen for this purpose due to its ability to model complex nonlinear and nonparametric relationships. GPR model not only gives the mean value of the estimate, but also quantifies the confidence intervals for it. This can be particularly useful in the tightening process, where measurements are noisy and the underlying friction behavior is uncertain. In addition, multiple experimental tightening datasets corresponding to similar operating conditions can be combined to learn a probabilistic mapping.

GPR has been successfully applied in the context of bolt tightening for preload estimation, demonstrating high prediction accuracy and the ability to quantify uncertainty [12]. Additionally, GPR-based approaches have been shown to be effective in modeling nonlinear friction behavior in engineering systems [13].

The Gaussian Process Regression (GPR) method can be expressed mathematically as

$$f(T) \sim \mathcal{GP}\Big(m(T),\, \kappa(T, T')\Big), \tag{2.11}$$

where $m(T)$ is the mean function and $\kappa(T, T')$ is the covariance kernel [14]. T, T' are the two input points.
Let

$$\mathbf{T} = [T_1, \ldots, T_N]^T, \qquad \boldsymbol{\mu} = [\mu_1, \ldots, \mu_N]^T, \tag{2.12}$$

denote the measured data, where N is the total number of samples.
Assuming Gaussian measurement noise,

$$\mu_i = f(T_i) + \epsilon_i, \qquad \epsilon_i \sim \mathcal{N}(0, \sigma_n^2), \tag{2.13}$$

the covariance matrix of the observations is given by

$$K_{NN} = \kappa(\mathbf{T}, \mathbf{T}) + \sigma_n^2 I, \tag{2.14}$$

where K_{NN} is an $N \times N$ matrix containing the similarity between all training torque values [14]. The term $\sigma_n^2 I$ accounts for measurement noise.
For a set of test inputs $\mathbf{T}_*$, the covariance between training and test points, and among test points themselves, are given by

$$K_{N*} = \kappa(\mathbf{T}, \mathbf{T}_*), \qquad K_{**} = \kappa(\mathbf{T}_*, \mathbf{T}_*). \tag{2.15}$$

Using these covariance matrices, the GPR model predicts the CoF values for new inputs. The prediction follows a Gaussian distribution with mean [14]

$$\bar{\mathbf{f}}_* = m(\mathbf{T}_*) + K_{N*}^T K_{NN}^{-1}\Big(\boldsymbol{\mu} - m(\mathbf{T})\Big), \tag{2.16}$$

and covariance

$$\mathrm{Cov}[\mathbf{f}_*] = K_{**} - K_{N*}^T K_{NN}^{-1} K_{N*}. \tag{2.17}$$

The kernel function $\kappa(T, T')$ defines how similar two torque values are, and hence how similar the corresponding CoF values are expected to be. In this work, a Radial

Basis Function (RBF) kernel is used. A Radial Basis Function (RBF) kernel is a covariance function that defines the similarity between two input points based on their distance. It assigns high correlation to nearby inputs and progressively smaller correlation as the distance between them increases. Mathematically, the RBF kernel is written as

$$\kappa(T, T') = \sigma_f^2 \exp\left(-\frac{(T - T')^2}{2\ell^2}\right), \tag{2.18}$$

where ℓ controls how quickly the CoF changes with torque, and σ_f^2 defines the overall variation in the function.
The hyperparameters $\theta = \{\ell, \sigma_f^2, \sigma_n^2, m\}$ are learned by maximizing the log marginal likelihood:

$$\log p(\boldsymbol{\mu} \mid \mathbf{T}, \theta) = -\frac{1}{2}(\boldsymbol{\mu} - m)^T K_{NN}^{-1}(\boldsymbol{\mu} - m) - \frac{1}{2}\log|K_{NN}| - \frac{N}{2}\log(2\pi). \tag{2.19}$$

This optimization balances how well the model fits the data and how complex the model is.
The RBF kernel is chosen due to its ability to model smooth and continuous relationships between inputs and outputs. In the present problem, the coefficient of friction varies with respect to torque, without abrupt discontinuities. The RBF kernel captures this behavior effectively by assigning high similarity to nearby torque values and smoothly decreasing correlation as the distance between inputs increases. Additionally, the RBF kernel can model a wide range of nonlinear relationships without requiring a specific parametric form. This makes it suitable for problems where the underlying relationship between torque and CoF is unknown or complex. Furthermore, the kernel is controlled by a small number of interpretable hyperparameters, such as the length scale ℓ and signal variance σ_f^2, which allow flexibility in adjusting the smoothness and variability of the learned function.
It should be noted that GPR works on the entire dataset at once and does not explicitly consider time dynamics. Although it provides a flexible mapping between torque and CoF, the estimation accuracy is limited in this application due to the very small variation in CoF over time. The detailed results are discussed in the next chapter.
The GPR model is trained using the data in the elastic region, and a cross-evaluation is carried out, which means that the model is trained with all the samples, except the one that it is tested for.

2.2.2 Formulation-1: F_c as a state

The current formulation uses F_c as a state, however, since the exact evolution of the system maybe difficult to express mathematically, as it depends on the linearity of joint stiffness K_j and tightening speed, which is considered constant, so it is expressed as a random walk in discrete time. Another motivation to use this approach the use of this in [3].
The CoF is predicted using the torque value from the previous time step, which can be expressed mathematically in discrete time as follows:

$$\mu_k = f(T_{k-1} + \Delta T) \tag{2.20}$$

In the above expression , ΔT is the average change in torque per time step.
The continuous time motion model is as follows:

$$\dot{F}_c(t) = 0 \tag{2.21}$$

The measurement model, which is the torque is expressed based on equation (2.7),

$$T(t) = F_c(t)\left(C_p + C_f\mu(t)\right) \tag{2.22}$$

where μ is determined using expression (2.20).
The discrete time state space of the same can be expressed as,

$$F_{c,k+1} = F_{c,k} + w_k \tag{2.23}$$
$$T_k = F_{c,k}(C_p + C_f\mu) \tag{2.24}$$

where w_k is the process noise.

2.2.3 Formulation-2: Defining dynamics for F_c

Another approach that can be used to describe force evolution is to use the relation given in (2.4). However, the drawback with this approach would be that the joint stiffness K_j, which is assumed to be constant in the elastic region, may be slightly varying in real-time.
The continuous time motion model is,

$$\dot{F}_c(t) = K_j\omega \tag{2.25}$$

where ω is the tightening speed, which is constant as mentioned in the earlier sections.
Measurement model is,

$$T(t) = F_c(t)(C_p + C_f\mu(t)) \tag{2.26}$$

Similar to the previous formulation, the expression in (2.20) is used to estimate μ.
The discrete time state space of the same can be expressed as,

$$F_{c,k+1} = F_{c,k} + K_j\omega_k\Delta t + w_k \tag{2.27}$$
$$T_k = F_{c,k}(C_p + C_f\mu) \tag{2.28}$$

where w_k is the process noise.

2.2.4 Formulation-3: Defining dynamics for F_c and μ

Instead of predicting the values of μ from a data-driven model, an approach can be to approximately describe it's evolution. The following approach was considered to arrive at the approximate dynamics for the CoF:
Let,

$$\mu = h(T(t)) \tag{2.29}$$

where h is an arbitrary function expressing the relation between the CoF and torque.

It is observed from the measured data the CoF decreases over time and can be approximately defined as a linear function. If m is the slope of this linear function and μ_0 is the intercept then,

$$\mu(t) = m\,T(t) + \mu_0 \tag{2.30}$$

Differentiating with respect to time,

$$\dot{\mu} = m\dot{T} \tag{2.31}$$

Also, from the measured data, it is observed that the Torque varies linearly with tightening speed (which is a constant). Considering the slope of Torque against the tightening speed to be b,

$$\dot{\mu} = m(b\dot{\theta}) = m\,b\,\omega \tag{2.32}$$

So with the dynamics for CoF and force, a new motion model is developed,

$$\dot{\mu}(t) = m\,b\,\omega \tag{2.33}$$

$$\dot{F}_c(t) = K_j\omega \tag{2.34}$$

The measurement model is considered the same,

$$T(t) = F_c(t)(C_p + C_f\mu(t)) \tag{2.35}$$

From a state space representation, the formulation can be expressed in continuous time as,

$$\mathbf{x} = \begin{bmatrix} \mu \\ F_c \end{bmatrix} \tag{2.36}$$

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{\mu} \\ \dot{F}_c \end{bmatrix} = 0\mathbf{x} + \underbrace{\begin{bmatrix} mb \\ K_j \end{bmatrix}}_{\mathbf{B}} \underbrace{\omega}_{\mathbf{u}} \tag{2.37}$$

$$\mathbf{y} = T = F_c(C_p + C_f\mu) \tag{2.38}$$

The discrete time version sampled at Δt can be expressed as,

$$\begin{aligned} \mu_{k+1} &= \mu_k + m\,b\,\omega\Delta t + w_k^{\mu}) \\ F_{c,k+1} &= F_{c,k} + K_j\omega\Delta t + w_k^{F_c} \\ T_k &= F_{c,k}(C_p + C_f\mu_k) \end{aligned} \tag{2.39}$$

where $w_k^{\mu} \sim \mathcal{N}(0, Q_\mu)$ and $w_k^{F_c} \sim \mathcal{N}(0, Q_{F_c})$ are measurement noise for μ, F_c respectively with Q_μ and Q_{F_c} as the process noise variance.

2.3 Observability

Observability is a property of a dynamical system that determines whether its internal states can be uniquely reconstructed from available input and output measurements over a finite time interval. A system is considered observable if all internal states influence the outputs in a distinguishable manner, allowing them to be identified from the measured signals. This concept is particularly important in state estimation problems, as it defines whether an observer or estimator can reliably reconstruct unmeasured states from sensor data.

2.3.1 Local Observability

Local observability refers to the ability of a dynamical system to uniquely determine its internal states from input and output measurements within a neighborhood of a given operating point.
For a dynamic system,

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}, \mathbf{u}) \tag{2.40}$$
$$\mathbf{y} = h(\mathbf{x}) \tag{2.41}$$

Local observability can be interpreted in accordance with [15] as follows:
Let $\mathbf{x}^0$ and $\mathbf{x}^1$ be two initial states of the system such that $\mathbf{x}^0 \neq \mathbf{x}^1$. The system is said to be locally observable if the corresponding output trajectories $\mathbf{y}_{\mathbf{x}^0}$ and $\mathbf{y}_{\mathbf{x}^1}$ are different.
Hermann and Krener [16] extended the concept of observability to nonlinear systems by formulating it in terms of distinguishability of state trajectories. According to their definition, a nonlinear system is locally observable if any two nearby initial states produce distinct output trajectories over a finite time interval. In other words, small variations in the initial state must result in measurable differences in the system output.
Their work gave an insight that observability for nonlinear systems is not necessarily a global property, but depends on the specific operating point or the trajectory of the system. Consequently, a system may be locally observable in certain regions of the state space while becoming unobservable in others. If perturbations in the states produce indistinguishable outputs, then the corresponding states cannot be uniquely estimated.
While the Herman-Krener's work provide the foundation for the local observability in non-linear systems, Isidori [17] complements this by providing a systematic approach to representing and transforming nonlinear systems for practical implementation of Hermann-Krener's theory. Isidori's approach helps in bridging the gap between the theoretical observability and practical observer implementation.

A common approach used to check the observability of a nonlinear systems is through Lie-Derivative test [18]. A Lie Derivative of a dynamic system represents the rate of change of output along the system trajectories. It can be mathematically expressed as follows considering the dynamic system in (2.40),
The first Lie derivative of the output function $h(\mathbf{x})$ along the vector field $f(\mathbf{x}, \mathbf{u})$ is defined as

$$L_f h(\mathbf{x}) = \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{u}). \tag{2.42}$$

Higher-order Lie derivatives can be computed recursively as

$$L_f^n h(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} \left(L_f^{n-1} h(\mathbf{x}) \right) f(\mathbf{x}, \mathbf{u}), \tag{2.43}$$

with $L_f^0 h(\mathbf{x}) = h(\mathbf{x})$.

The successive Lie Derivatives capture the evolution of output and its impact by the internal states. The observability matrix is then evaluated as follows:

$$\mathcal{O}(\mathbf{x}, \mathbf{u}) = \frac{\partial}{\partial \mathbf{x}} \begin{bmatrix} L_f^0 h \\ L_f^1 h \\ L_f^2 h \\ \vdots \\ L_f^n h \end{bmatrix} \tag{2.44}$$

where n is the number of states. For the system to be observable locally, it is necessary that the above matrix is full-rank at the operating point.
From an observability perspective, it is essential that the variation in the initial value of clamp force and CoF results in a different torque. If the mapping between the states and the output is not sufficiently sensitive or if multiple state combinations produce similar torque responses, the system may suffer from local observability issues.

2.3.2 Output trajectory-based observability study

With the results in the previous study in [3], it has been observed that the accuracy of the initial state reflects on the estimation error. As the observer is not able to uniquely estimate the state due to the states appearing as products in the output expression as in (2.35), so an approach could be to find the best region in the possible values of the state pair, which leads to an indistinguishable outputs as compared to the measured ones. This observability analysis was carried out by Zeng in [19], where output trajectories generated from different initial state pairs were compared with the true output trajectory. Based on this comparison, a region of practical unobservability was defined, and the approach was referred to as global observability analysis. It was defined as the region in the initial state space $\mathcal{X}$ such that different initial state pairs produce output trajectories that are sufficiently close to the measured output trajectory, i.e.,

$$\mathcal{X}_u = \left\{ \mathbf{x}_0 \in \mathcal{X} \;\middle|\; \|\mathbf{y}(k; \mathbf{x}_0) - \mathbf{y}_{\text{true}}(k)\| \leq \epsilon,\ \forall\ k \in [0, N] \right\}, \tag{2.45}$$

where $\mathbf{y}(k; \mathbf{x}_0)$ is the open-loop output trajectory corresponding to the initial state $\mathbf{x}_0$, $\mathbf{y}_{\text{true}}(k)$ is the measured output trajectory, and ϵ is a predefined tolerance.
Here open-loop means the trajectories generated only by the system definition, i.e., the state evolution and does not include and predictions. The reason for choosing this is to consider the error in the formulation when compared to the actual. In theoretical sense, this means to consider all the trajectories within that lie within ϵ to be indistinguishable.
For every initial possible pairs of state values, state trajectory can be generated using the dynamics as mentioned in (2.37). Then the corresponding open-loop output trajectory can be generated and can be compared with the measured torque trajectory which is filtered. This process would require the comparison of outputs at every time sample for every incorrect initial state.

In order to simplify the analysis process, ϵ can be defined as

$$\epsilon^{\text{true}} = \sqrt{\frac{\sum_{k=0}^{N} \left(\mathbf{y}(k; \mathbf{x}_0^{true}) - \mathbf{y}_{\text{true}}(k)\right)^2}{N}} \tag{2.46}$$

where $\mathbf{x}_0^{true}$ is the true initial state matrix from the actual measurement. Then the unobservable region can be defined as,

$$\mathcal{X}_u = \left\{ \mathbf{x}_0 \in \mathcal{X} - \{\mathbf{x}_0^{\text{true}}\} \;\middle|\; \sqrt{\frac{\sum_{k=0}^{N} \left(\mathbf{y}(k; \mathbf{x}_0) - \mathbf{y}_{\text{true}}(k)\right)^2}{N}} \leq \epsilon^{\text{true}},\ \forall\ k \in [0, N] \right\} \tag{2.47}$$

In order to analyze this better, heat plot as in [19] is generated, considering a certain bounds on the initial state, where the state and the output trajectories are generated using the dynamics.

2.3.3 Sensitivity

With the observability and distinguishability being discussed, it is crucial to evaluate the sensitivity. The sensitivities can be analyzed for different parameters, however, in this case, the interest lies in the variation of output trajectories due to change in the states, so sensitivity can be evaluated for the torque for changes in the clamp force and CoF. The output sensitivity to the unit change in states is of interest as the sensitivity to the initial states has been viewed in the previous sections. To evaluate the output sensitivity to initial states, the following approach is used,
Sensitivity of torque to clamp force using (2.35),

$$\frac{\partial T(t)}{\partial F_c(t)} = C_p + C_f \mu(t) \tag{2.48}$$

And sensitivity of torque to CoF can be expressed as,

$$\frac{\partial T(t)}{\partial \mu(t)} = F_c(t) C_f \tag{2.49}$$

In order to relate the analytical sensitivity to the actual discrete-time to use the existing measured data, the partial derivatives are scaled using the per-time-step variations of the states. This results in an approximation of the output variation due to changes in the states:

$$\Delta T_\mu \approx \frac{\partial T}{\partial \mu} \cdot \Delta \mu \tag{2.50}$$

$$\Delta T_{F_c} \approx \frac{\partial T}{\partial F_c} \cdot \Delta F_c \tag{2.51}$$

This provides a more physically meaningful interpretation of sensitivity, as it reflects the actual contribution of state variations to the change in torque over each time step. One important consequence of this is the sign of resulting quantity that since μ value decreases in a broader sense, the $\Delta\mu$ value becomes negative, and hence the

value of sensitivity becomes negative. Therefore, just the magnitude of change in μ is considered.

An observation made from the existing data on the measurement error against the variation in CoF is made. The average increment in torque per time step and the corresponding error (with respect to the filtered data) are evaluated. The resulting impact on the coefficient of friction is summarized in Table-2.1.

	Avg Increment per time sample	Error
Torque values	0.01501 Nm	0.013545 Nm
Effect on μ values	9×10^{-5}	8.1×10^{-5}

Table 2.1: Comparison of signal variation and error propagation

It can be observed that the variation in the coefficient of friction is of the same order of magnitude as the error introduced due to measurement and signal processing. This indicates a very low signal-to-noise ratio, making it difficult to accurately resolve the evolution of μ from the measured torque.
This further explains the challenges in state estimation, as even small measurement errors can significantly influence the inferred state values.

2.4 Observer Design

2.4.1 LMI based Luenberger Observer

Since the value of COF varies in a very small band, so, an approach can be to view it as a scheduling parameter. The system in 2.2.4 can be then viewed as a Linear Parameter Varying (LPV) system and can be described as,

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{bmatrix} \dot{\mu} \\ \dot{F}_c \end{bmatrix} = \begin{bmatrix} m\, b \\ K_j \end{bmatrix} \omega \\ T = C'(\rho)F_c &= \underbrace{\begin{bmatrix} 0 & C'(\rho) \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} \mu \\ F_c \end{bmatrix} \end{aligned} \tag{2.52}$$

where $C'(\rho) = C_p + C_f\mu$ and ρ is the scheduling parameter μ
From the perspective of an observer design, a Luenberger observer is chosen as the basic estimation technique. A Luenberger Observer for a linear system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \tag{2.53}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} \tag{2.54}$$

is described in [18] as,

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{L}(\mathbf{y} - \hat{\mathbf{y}}) \tag{2.55}$$

where $\hat{\mathbf{x}}$ is the estimated state vector and $\mathbf{L}$ is the observer gain. The error in prediction is described as

$$\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}} \tag{2.56}$$

then the error dynamics is defined as,

$$\dot{\mathbf{e}} = (\mathbf{A} - \mathbf{LC})\mathbf{e} \tag{2.57}$$

The objective is to select $\mathbf{L}$ such that the error dynamics is asymptotically stable, in other words the error is to be minimized.
For the system in (2.52), the error dynamics become,

$$\dot{\mathbf{e}} = -\mathbf{L}C'(\rho)\mathbf{e} \tag{2.58}$$

Linear Matrix Inequality (LMI) Formulation:
There are many theories upon which an observer can be designed. A general and most commonly used observer is Gaussian estimator, however, since this has already been directly implemented on the system formulations in the previous studies, so different approaches are evaluated.
Linear Matrix Inequalities (LMIs) [20] provide a convenient framework for transforming the observer gain design problem into a convex optimization problem that can be solved efficiently and robustly.
To guarantee the convergence of the estimation error, Lyapunov-based methods are commonly employed. Consider a quadratic Lyapunov function of the form [18]

$$V(\mathbf{e}) = \mathbf{e}^T\mathbf{P}\mathbf{e}, \quad \mathbf{P} = \mathbf{P}^T \succ 0 \tag{2.59}$$

A sufficient condition for asymptotic stability is

$$\dot{V}(\mathbf{e}) < 0. \tag{2.60}$$

Using the error dynamics $\dot{\mathbf{e}} = -\mathbf{LC}(\rho)\mathbf{e}$, the derivative of the Lyapunov function becomes

$$\dot{V}(\mathbf{e}) = \dot{\mathbf{e}}^T\mathbf{P}\mathbf{e} + \mathbf{e}^T\mathbf{P}\dot{\mathbf{e}} \tag{2.61}$$

$$= \mathbf{e}^T\left[(-\mathbf{LC}(\rho))^T\mathbf{P} + \mathbf{P}(-\mathbf{LC}(\rho))\right]\mathbf{e}. \tag{2.62}$$

A sufficient condition for $\dot{V}(\mathbf{e}) < 0$ is

$$\mathbf{C}(\rho)^T\mathbf{L}^T\mathbf{P} + \mathbf{PLC}(\rho) \succ 0, \quad \mathbf{P} \succ 0, \tag{2.63}$$

where $\mathbf{C}(\rho) = [0, C'(\rho)]$.
The condition in (2.63) is bilinear in the unknowns $\mathbf{L}$ and $\mathbf{P}$ due to the product $\mathbf{PL}$. To obtain a convex formulation, a change of variables is introduced:

$$\mathbf{Z} = \mathbf{PL} \tag{2.64}$$

Substituting this into (2.63) yields the LMI

$$\mathbf{C}(\rho)^T\mathbf{Z}^T + \mathbf{ZC}(\rho) \succ 0. \tag{2.65}$$

Since the output matrix $\mathbf{C}(\rho)$ depends on CoF, which is highly uncertain, it is modeled as,

$$\mathbf{C}(\rho) = \mathbf{C}_j + \Delta(\rho) \tag{2.66}$$

where $\Delta(\rho)$ is the bounded uncertainty and satisfies,

$$\|\Delta(\rho)\| \leq \Delta_{\text{bar}} \tag{2.67}$$

where Δ_{bar} is the upper bound on the uncertainty in $\mathbf{C}(\rho)$ due to the variation in CoF and measurement noise.
To ensure the robustness with respect to this uncertainty, LMI conditions in (2.65) is update to,

$$\mathbf{C}_j^T\mathbf{Z}^T + \mathbf{Z}\mathbf{C}_j - 2\Delta_{\text{bar}}\mathbf{I} \succ 0, \quad j = 1, 2 \tag{2.68}$$

Since the scheduling parameter $\rho = \mu$ varies within a bounded interval $\rho \in [\mu_{\min}, \mu_{\max}]$, the system can be represented as a Linear Parameter-Varying (LPV) model [18]. A polytopic representation is adopted, and the vertex matrices are defined as:

$$\mathbf{C}_1 = [0,\ C_p + C_f\mu_{\min}] \tag{2.69}$$

$$\mathbf{C}_2 = [0,\ C_p + C_f\mu_{\max}] \tag{2.70}$$

By convexity, satisfaction at the vertices guarantees satisfaction for all $\rho \in [\mu_{\min}, \mu_{\max}]$.
The observer gain design problem can therefore be formulated as the following convex optimization problem:

$$\text{Find } \mathbf{P}, \mathbf{Z} \quad \text{such that:}$$

$$\mathbf{C}_j^T\mathbf{Z}^T + \mathbf{Z}\mathbf{C}_j - 2\Delta_{\text{bar}}\mathbf{I} \succ 0, \quad i = 1, 2 \tag{2.71}$$

$$\mathbf{P} \succ 0. \tag{2.72}$$

In order to obtain $\mathbf{P}$ and $\mathbf{Z}$ satisfying the above expressions, the observer design problem is formulated as a convex optimization problem.
The constraints in (2.71), (2.72) are Linear Matrix Inequalities (LMIs), which are affine in $\mathbf{P}$ and $\mathbf{Z}$. The problem is formulated as a semi-definite programming (SDP) feasibility problem.
In this work, the optimization problem is implemented using the CVXPY framework [21][22], where $\mathbf{P}$ and $\mathbf{Z}$ are defined as optimization variables and the LMI constraints are expressed directly as matrix inequalities. A suitable SDP solver is then used to compute a feasible solution, from which the observer gain is recovered as $\mathbf{L} = \mathbf{P}^{-1}\mathbf{Z}$.
Once $\mathbf{P}$ and $\mathbf{Z}$ are obtained, the observer gain is recovered as

$$\mathbf{L} = \mathbf{P}^{-1}\mathbf{Z}. \tag{2.73}$$

In the numerical implementation, the matrix $\mathbf{P}$ is parameterized as a diagonal positive definite matrix to reduce complexity. Lower bounds are imposed on its diagonal entries to ensure numerical stability. Additionally, constraints are imposed on the gain matrix to avoid trivial solutions (e.g., $\mathbf{L} = 0$), which satisfy the LMI conditions but do not yield a meaningful observer.
The state estimate then becomes,

$$\dot{\hat{\mathbf{x}}} = \mathbf{B}\mathbf{u} + \mathbf{L}(\mathbf{y} - \hat{\mathbf{y}}) \tag{2.74}$$

where $\hat{\mathbf{x}} = \begin{bmatrix} \hat{\mu} \\ \hat{F}_c \end{bmatrix}$

Initialization
Initialization can be performed either by assigning an initial value to the clamp force or to the coefficient of friction at the start of the tightening process. One of the state is guessed and the other is calculated analytically using the initial torque measurement.

2.4.2 Extended Kalman Filter for Inverted Dynamics

In the output expression in (2.2.4), due to the states appearing as products, the changes in the measurement may not be directly reflected in the estimated states, so a better approach would be to estimate only one state from the measurement. This could be achieved by eliminating a state in the output expression, and one way to do this could be by deriving the time derivative of the output.

2.4.2.1 Estimating CoF from Torque

One approach could be to choose CoF as a state to be extracted from the measurement. Here, a parameter, which is dependent on the CoF is considered as the state to be retained in the output expression. An implicit state space is generated, in order to estimate the parameter. The mathematical approach can be expressed as follows:
The output expression in (2.2.4) can be expressed as,

$$T = F_c \alpha \tag{2.75}$$

where $\alpha = C_p + C_f \mu$ time derivative of (2.75) is,

$$\dot{T} = \dot{F}_c \alpha + F_c \dot{\alpha} \tag{2.76}$$

Using (2.37) and $\dot{\alpha} = \dot{\mu}\, C_f = (m\, b\, \omega) C_f$,

$$\dot{T} = K_j \omega \alpha + F_c m\, b\, \omega C_f \tag{2.77}$$

Substituting $F_c = \frac{T}{\alpha}$ using (2.75)

$$\underbrace{K_j \alpha^2 - \frac{\dot{T}}{\omega}\alpha + T C_f m\, b}_{\phi(\alpha)} = 0 \tag{2.78}$$

or

$$\phi(\alpha) = 0 \tag{2.79}$$

ϕ is quadratic in α

A very simple approach at this stage would be to find α such that ϕ is 0 or minimized. However, due to the direct dependency on $\dot{T}$, which can be noisy, the estimate of

α may not be accurate. Additionally, since the dynamics in (2.37) are based on assumptions, even with a filtered $\dot{T}$, the estimates will be poor.

To avoid this a further analytical simplification is carried out,

Let $h(\alpha) = K_j\alpha + \frac{TC_f m\, b}{\alpha}$ and $z = \frac{\dot{T}}{\omega}$, then

$$\phi(\alpha) = \alpha(h(\alpha) - z) \tag{2.80}$$

from expression (2.79)

$$\alpha(h(\alpha) - z) = 0 \tag{2.81}$$

$$\implies \alpha = 0 \text{ or } z = h(\alpha) \tag{2.82}$$

Using the above expression a state space can be formulated in discrete time as follows:
Motion model,

$$\alpha_{k+1} = \alpha_k + w_k^\alpha, \quad w_k^\alpha \sim \mathcal{N}(0, Q_\alpha) \tag{2.83}$$

Measurement model,

$$z_k = h(\alpha_k) + v_k^z, \quad v_k^z \sim \mathcal{N}(0, R_z) \tag{2.84}$$

Extended Kalman filter(EKF) is chosen as an observer for estimation of α.

EKF Initialization:
The EKF [23] requires an initial estimate $\hat{\alpha}_0$ and an initial covariance P_0. Since $\alpha = C_p + C_f\mu$, the initial value is selected from an admissible friction range as

$$\hat{\alpha}_0 = C_p + C_f\hat{\mu}_0, \tag{2.85}$$

where $\hat{\mu}_0$ is the assumed initial CoF. The initial covariance P_0 is chosen to represent the uncertainty in this initial guess. A larger value of P_0 indicates lower confidence in the initial estimate and allows the filter to rely more strongly on the measurements during the early update steps.
Since $h(\alpha)$ is nonlinear in α, it is to be linearized,

$$H = \frac{\partial h}{\partial \alpha} = K_j - \frac{TC_f mb}{\alpha^2} \tag{2.86}$$

EKF recursion[23]:
Prediction step,

$$\hat{\alpha}_{k|k-1} = \hat{\alpha}_{k-1} \tag{2.87}$$

$$P_{k|k-1} = P_{k-1} + Q_\alpha \tag{2.88}$$

Innovation,

$$\tilde{T}_k = z_k - h(\hat{\alpha}_{k|k-1}) \tag{2.89}$$

Kalman Gain,

$$K_k = \frac{P_{k|k-1}H_k}{H_k^2 P_{k|k-1} + R_z} \tag{2.90}$$

Update step,

$$\hat{\alpha}_k = \hat{\alpha}_{k|k-1} + K_k \tilde{T}_k \tag{2.91}$$

$$P_k = (1 - K_k H_k) P_{k|k-1} \tag{2.92}$$

Using the estimated α, the clamp force can then be estimated using (2.75),

$$\hat{F}_{c,k} = T_k / \alpha_k \tag{2.93}$$

Since the measurement in this case is $z_k = \frac{\dot{T}}{\omega}$, the $\dot{T}$ requires conversion to discrete time, which can be ΔT, the time step increment in Torque. However, since in the hardware the per time step increment in torque is not always constant and due to the presence of measurement noise, the ΔT can be noisy. To reduce the effect of noise and variability in the discrete-time torque measurements, the derivative $\dot{T}$ is approximated using a moving average over a finite window.
The torque derivative $\dot{T}$ is approximated using a moving average of the discrete-time increments $\Delta T_k = T_k - T_{k-1}$. For $k \geq N_w$,

$$\dot{T}(k) \approx \frac{1}{N_w} \sum_{i=k-N_w+1}^{k} \Delta T(i), \tag{2.94}$$

while for $k < N_w$,

$$\dot{T}(k) \approx \frac{1}{k} \sum_{i=1}^{k} \Delta T(i), \tag{2.95}$$

where N_w is the averaging window length
The tuning of the process and measurement noise variances, Q_α and R_z, plays a critical role in the performance of the estimator. The choice of Q_α determines the extent to which the observer relies on the system dynamics, whereas the choice of R_z governs the reliance on the measurement information.
A larger value of Q_α allows the state estimate to adapt more rapidly by increasing the influence of the measurement updates, while a smaller value enforces stronger adherence to the model predictions. Similarly, a smaller R_z increases sensitivity to measurement variations, potentially leading to noise amplification, whereas a larger R_z results in smoother estimates by filtering out high-frequency fluctuations.
In the present case, the measurement z_k is derived from a discretized and averaged approximation of the torque derivative, which inherently contains noise due to differentiation. Therefore, the tuning objective is to balance responsiveness and robustness by selecting Q_α and R_z such that the estimator captures the overall trend of the measurement without overfitting high-frequency noise. This balance is essential to ensure stable and physically meaningful estimation of the states.

2.4.2.2 Estimating Clamp Force from Torque

Another approach can be to extract clamp force from the measured torque and its' derivative. This is done to evaluate if there are can be other behavior observed

during the estimation. In the expression (2.77), instead of replacing F_c, α can be replaced as $\alpha = \dfrac{T}{F_c}$, so (2.78) becomes,

$$F_c^2(m\,b\,\omega)\,C_f - \dot{T}F_c + K_j\omega T = 0 \tag{2.96}$$

Let

$$y = K_j\omega T \tag{2.97}$$

$$g(F_c) = -F_c^2(m\,b\,\omega)C_f + \dot{T}F_c \tag{2.98}$$

then equation (2.96) becomes,

$$-g(F_c) + y = 0 \implies y = g(F_c) \tag{2.99}$$

Then the state space using the dynamics in (2.37) becomes,
motion model,

$$F_{c,k+1} = F_{c,k} + K_j\,\omega\,\Delta t + w_k^{F_c} \tag{2.100}$$

Measurement model,

$$y_k = g(F_{c,k}) + v_k^y \tag{2.101}$$

where $w_k^{F_c} \sim \mathcal{N}(0, Q_{F_c})$ and $v_k^y \sim \mathcal{N}(0, R_y)$ are the process and measurement noise respectively.

EKF Initialization:
The EKF requires an initial estimate $\hat{F}_{c,0}$ and an initial covariance P_0. The choice of initialization is important, since the analysis shows that errors in the initial clamp-force guess have a significantly larger impact on the estimation accuracy than errors in the initial CoF guess. Therefore, $\hat{F}_{c,0}$ is obtained by first assigning an initial guess for the CoF and then using the initial torque measurement to calculate the corresponding clamp force.

EKF recursion:
Prediction step,

$$\hat{F}_{c,k|k-1} = \hat{F}_{c,k-1} + K_j\omega\Delta t \tag{2.102}$$

$$P_{k|k-1} = P_{k-1} + Q_{F_c} \tag{2.103}$$

Innovation,

$$\tilde{y}_k = y_k - g(\hat{F}_{c,k|k-1}) \tag{2.104}$$

Since the term $F_c\,g(F_c)$ is nonlinear, it requires linearization. This can be done by,

$$G = \frac{\partial(g(F_c))}{\partial F_c} = -2F_c(m\,b\,\omega)C_f + \dot{T} \tag{2.105}$$

Kalman Gain,

$$\mathbf{K}_k = \frac{P_{k|k-1}G_k}{G_k^2 P_{k|k-1} + R_y} \tag{2.106}$$

Update step,

$$\hat{F}_{c,k} = \hat{F}_{c,k|k-1} + \mathbf{K}_k \tilde{y}_k \tag{2.107}$$

$$P_k = (1 - \mathbf{K}_k H_k) P_{k|k-1} \tag{2.108}$$

In this case, the measurement $y_k = K_j \omega T_k$. So, it can be said that the measurements would not consist of the noise associated with the $\dot{T}$. However, since this term appears in the $g(F_c)$, so it's noisy impact may still can be reflected in the Innovation and Kalman gain calculations. As in the previous section, the tuning of Q_{F_c} and R_y becomes crucial.

2.4.3 Rao-Blackwellized Particle Filter

For the system in (2.39), the key difficulty is that the output is bilinear in $(\mu_k, F_{c,k})$. An approach could be to decompose the states by using a Rao-Blackwellized Particle Filter(RBPF). The initial idea for this approach was explored with the assistance of GitHub Copilot. The method was subsequently verified, adapted, and critically assessed by the author before being included in this work.
For the state matrix,

$$\mathbf{x}_k = \begin{bmatrix} \mu_k \\ F_{c,k} \end{bmatrix} \tag{2.109}$$

μ_k is modeled as a stochastic state to account for unmodeled effects and variability in friction behavior, while $F_{c,k}$ is treated as conditionally linear-Gaussian given each particle and is estimated analytically by a scalar Kalman filter for each particle.
The objective of the RBPF is to approximate the joint posterior distribution,

$$p(\mu_k, F_{c,k} \mid T_{1:k}) \tag{2.110}$$

where $T_{1:k}$ denotes the sequence of torque measurements up to time step k. Using the Rao-Blackwellization theory, this posterior can be factorized as [24],

$$p(\mu_k, F_{c,k} \mid T_{1:k}) = p(F_{c,k} \mid \mu_k, T_{1:k})\, p(\mu_k \mid T_{1:k}) \tag{2.111}$$

Here, $p(\mu_k \mid T_{1:k})$ is approximated using a set of weighted particles, while $p(F_{c,k} \mid \mu_k, T_{1:k})$ is represented by a Gaussian distribution conditioned on each particle.

The conditional clamp-force for every particle i with weight $\mathrm{w}_k^{(i)}$ for friction state $\mu_k^{(i)}$ is,

$$p(F_{c,k} \mid \mu_k^{(i)}, T_{1:k}) = \mathcal{N}\left(F_{c,k}^{(i)}, P_{F,k}^{(i)}\right) \tag{2.112}$$

where $F_{c,k}^{(i)}$ is the conditional mean and $P_{F,k}^{(i)}$ is the assocaited variance.
So, every particle carries,

$$\{\mu_k^{(i)},\ F_{c,k}^{(i)},\ P_{F,k}^{(i)},\ \mathrm{w}_k^{(i)}\} \tag{2.113}$$

The coefficient of friction is propagated using a dynamic model,

$$\mu_{k+1}^{(i)} = \mu_k^{(i)} + m\, b\, \omega_k\, \Delta t + w_k^{\mu} \tag{2.114}$$

where $w_k^{\mu} \sim \mathcal{N}(0, Q_{\mu})$ represents process noise capturing modeling uncertainty.

Initialization
The initial friction particles can be sampled from a Gaussian prior,

$$\mu_0^{(i)} \sim \mathcal{N}(\bar{\mu}_0, \sigma_{\mu,0}^2) \tag{2.115}$$

where $\bar{\mu}_0$,$\sigma_{\mu,0}$ are mean and standard deviation arrived from the experimental values.
Using the first filtered torque measurement T_0, the initial clamp force for each particle in initial CoF can be obtained as follows,

$$F_{c,0}^{(i)} = \frac{T_0}{C_p + C_f \mu_0^{(i)}} \tag{2.116}$$

Prediction
The prediction for the next time instant is carried out using the dynamic in (2.39)
The clamp-force mean is propagated using the stiffness relation

$$F_{c,k|k-1}^{(i)} = F_{c,k-1|k-1}^{(i)} + K_j \omega_k \Delta t \tag{2.117}$$

The associated covariance is inflated to account for uncertainty in the clamp-force propagation:

$$P_{F,k|k-1}^{(i)} = P_{F,k-1|k-1}^{(i)} + \mathrm{Q}_{F,k-1} \tag{2.118}$$

where $\mathrm{Q}_{F,k-1}$ is the process noise variance associated with the clamp force. One possible reasons for clamp force noise could be the nonlinearity in the clamp force increase, which is a varying joint stiffness. So, for this application, the process noise is defined as follows:

$$\mathrm{Q}_{F,k} = \left(\sigma_{K_j}^{\mathrm{rel}} K_{j,\mathrm{avg}} \omega_k \Delta t\right)^2 + Q_{F,\min} \tag{2.119}$$

Here, $\sigma_{K_j}^{\mathrm{rel}}$ denotes the relative uncertainty in the joint stiffness and $Q_{F,\min}$ is the minimum process noise limit

Conditional Kalman Update
For each particle, the innovation term from the prediction becomes,

$$\tilde{T}_k^{(i)} = T_k \; - \; F_{c,k|k-1}^{(i)}(C_p + C_f \mu_k^{(i)}) \tag{2.120}$$

Innovation covariance,

$$S_k^{(i)} = \left(C_p + C_f \mu_k^{(i)}\right)^2 P_{F,k|k-1}^{(i)} + R \tag{2.121}$$

where R is the measurement noise covariance.
Kalman gain,

$$K_k^{(i)} = \frac{(C_p + C_f \mu_k^{(i)}) P_{F,k|k-1}^{(i)}}{S_k^{(i)}} \tag{2.122}$$

The conditional clamp-force estimate is updated as

$$F_{c,k|k}^{(i)} = F_{c,k|k-1}^{(i)} + K_k^{(i)} \tilde{T}_k^{(i)} \tag{2.123}$$

and the corresponding variance becomes

$$P_{F,k|k}^{(i)} = \left(1 - K_k^{(i)}(C_p + C_f \mu_k^{(i)})\right) P_{F,k|k-1}^{(i)} \tag{2.124}$$

Weight update
Parallelly the weights are to be updated for the particles based on the residuals.

$$r_k^{(i)} = T_k - F_{c,k|k}^{(i)}(C_p + C_f \mu_k^{(i)}) \tag{2.125}$$

Assuming Gaussian measurement noise, the particle weights are updated according to

$$\mathrm{w}_k^{(i)} \propto \mathrm{w}_{k-1}^{(i)} \exp\left(-\frac{1}{2}\frac{\left(r_k^{(i)}\right)^2}{R}\right) \tag{2.126}$$

followed by normalization,

$$\sum_{i=1}^{N_p} \mathrm{w}_k^{(i)} = 1 \tag{2.127}$$

where N_p is the number of particles.

Resampling
Particle degeneracy is monitored using the effective sample size,

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N_p}\left(\mathrm{w}_k^{(i)}\right)^2} \tag{2.128}$$

When N_{eff} falls below a selected threshold, resampling is performed to remove particles with negligible weights and duplicate particles with high posterior probability. Finally, the state estimates are obtained as weighted ensemble averages:

$$\hat{\mu}_k = \sum_{i=1}^{N_p} \mathrm{w}_k^{(i)} \mu_k^{(i)} \tag{2.129}$$

$$\hat{F}_{c,k} = \sum_{i=1}^{N_p} \mathrm{w}_k^{(i)} F_{c,k}^{(i)} \tag{2.130}$$

2.5 Summary

This chapter presented the theoretical background on the tightening and the methodology for the estimation of clamp force and CoF in bolted joints. The tightening process and its key parameters were first described, followed by the system formulations representing the relationship between torque, clamp force and friction.

Methods to carry out observability is proposed with the inclusion of study on sensitivity of torque to the clamp force and CoF. Gaussian Process regression method was

introduced to estimate CoF from the measured torque. Apart from this, multiple observer designs were proposed to address the estimation problem. These include an LMI-based Luenberger observer, an Extended Kalman Filter applied to inverted dynamics, and a Rao–Blackwellized Particle Filter. Each approach addresses the estimation problem from a different perspective, highlighting various trade-offs in terms of robustness, computational complexity, and sensitivity to noise.

3
Results

3.1 Gaussian Process Regression Estimates

The GPR model was trained with 5 tightening samples with one left out for testing.

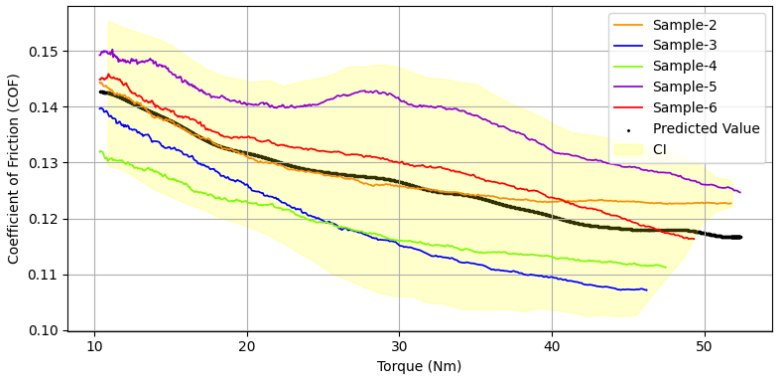


Figure 3.1: GPR predictions of CoF for sample-1 with Confidence Interval (CI) and training data

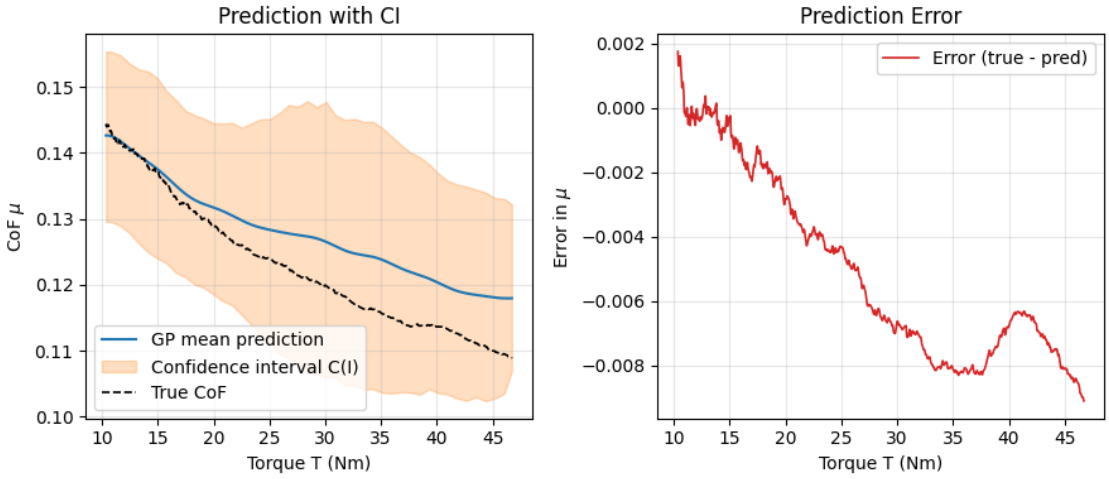


Figure 3.2: GPR predictions of CoF plotted with the actual measurements

At first glance, the estimation errors may appear to be relatively small. However, when compared to the actual variation in the coefficient of friction, which is on

the order of 7×10^{-6} per time step, these errors become significant. This indicates that the GPR model is not capturing the underlying behavior with the required resolution.
A key reason for this limitation is that GPR does not explicitly account for the temporal evolution of the coefficient of friction. Instead, it relies primarily on the density and distribution of the available data points. As a result, the model is unable to effectively track the very small, time-dependent changes in CoF, leading to reduced estimation accuracy in this application.

3.2 Local Observability

Using the Lie Derivatives approach observability matrix was developed and the observability is evaluated at initial measured state. This is carried out for all the formulations and the following is obtained:
For formulation in (2.23),(2.24),

$$\mathcal{O}_1 = 0.002007 \tag{3.1}$$

The same was observed for the system formulation in (2.25),(2.26)

$$\mathcal{O}_2 = 0.002007 \tag{3.2}$$

However since the third formulation in (2.37),(2.38) has two states, so the observability matrix then becomes,

$$\mathcal{O}_3 = \begin{bmatrix} 0.062 & 0.0018 \\ 1.6e+2 & -0.00016 \end{bmatrix} \tag{3.3}$$

$$det|\mathcal{O}_3| = -0.28800 \tag{3.4}$$

From the observability matrices it is clear that the rank for the respective formulations is equal to the number of states. This means that all the system formulations are observable. However, the determinant values are very small, which suggests that the system is weakly observable. In other words, although the states can be reconstructed mathematically, the estimation maybe highly sensitive to noise, model uncertainty, and numerical errors.

3.3 Output Trajectory based Observability

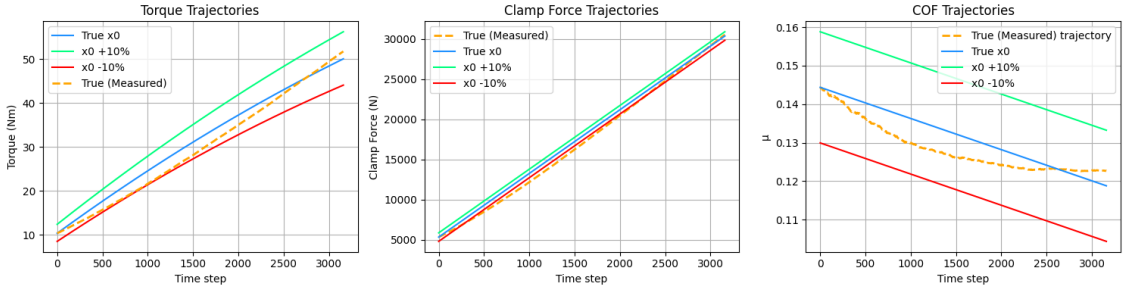
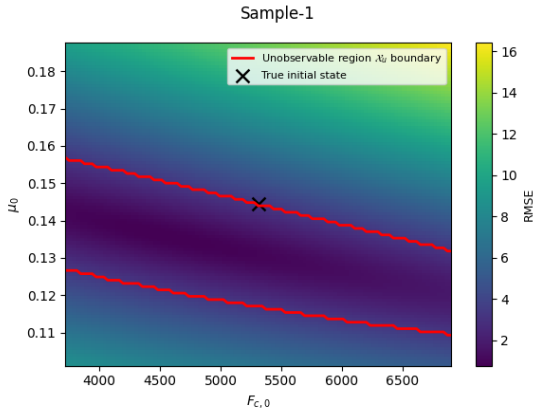


Figure 3.3: Trajectory comparison using open-loop

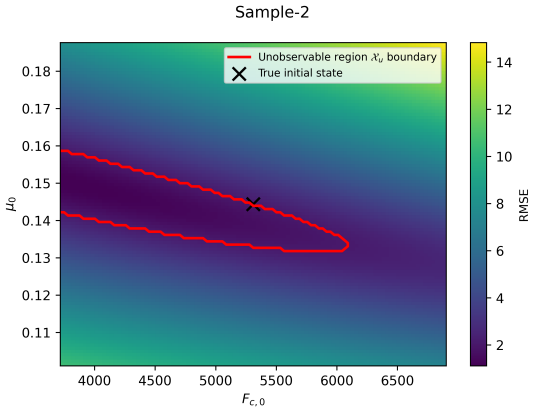
The above plot corresponds to a measured trajectory (dashed and orange color) and multiple generated trajectories. The blue colored curve corresponds to the trajectory with the initial state as is the measured trajectory, referred to as true initial state trajectory. The red and the green colored plots represent the trajectories for drifted initial states of about 10%.
It is clear from the above plot that there is a difference in the true initial state trajectory and the measured trajectory. The reasons for this is discussed in the previous chapter. It can also be seen that the output trajectories for drifted initial states are completely different. However, the attention here is towards any trajectory that lie in the region between the measured (orange) and the open-loop trajectory in blue color.
The plots in the figure 3.4, is generated using the error in the generated trajectory from different initial state pair against the true. These trajectories are then compared with the measured ones for every sample. The initial state pairs corresponding to the lower errors are the darker regions in the plot, whereas the brighter region corresponds to higher error, indicating that the output trajectories are distinguishable. The red-contour corresponds to the region where the error is less than ϵ^{true}, which is the unobservable region. Any output trajectory generated from the initial state pairs in this region is considered indistinguishable, within the chosen tolerance, from the measured trajectory. Therefore, it may not be possible to uniquely recover the true initial state from output measurements alone when the initial condition lies within this region.
A key observation from the plot is the shape and orientation of the unobservable region, which vary across different samples. This variation may be due to the process noise and/or measurement noise. As a result, it may become difficult to generalize the unobservable region.
Another observation is that the direction where there is a simultaneous increase in $F_{c,0}$ and reduction in μ_0 or vice versa, produces minimal change in the output trajectory, leading to low error values. On the other hand, simultaneous increase/decrease in both the parameters result in significant error, indicating that the system is highly sensitive to such variations. This basically indicates the coupling between the states, that is the effect of change of one state can be compensated by the variation of other state in the opposite direction.

3.4 Sensitivity Analysis

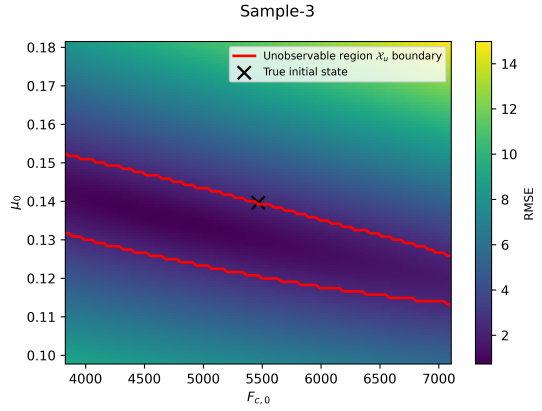
It can be observed from the sensitivity analysis that the contributions of the two states to the torque evolution differ significantly. Although the analytical sensitivity of torque with respect to the coefficient of friction increases over time, due to the increasing clamp force, the actual variation in CoF per time step is extremely small. As a result, the contribution of CoF to the torque variation remains limited.
In contrast, the sensitivity of torque to clamp force remains relatively consistent over time, while the variation in clamp force is comparatively larger. Consequently, the torque evolution is primarily driven by the changes in clamp force.
From an estimation perspective, this implies that the effect of CoF is weak and can be comparable to the measurement noise, making it difficult to reliably reconstruct.



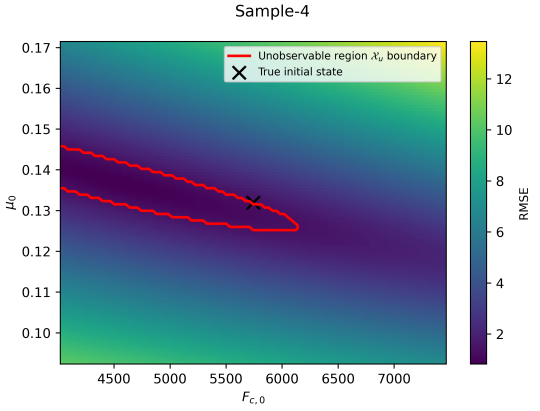
(a) Sample 1



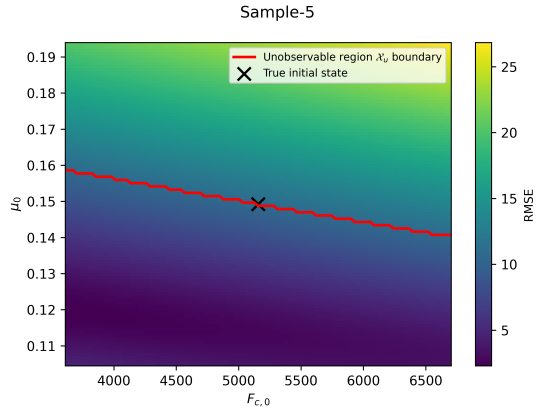
(b) Sample 2



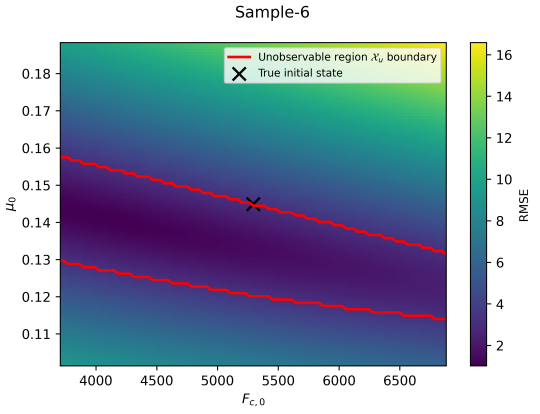
(c) Sample 3



(d) Sample 4



(e) Sample 5



(f) Sample 6

Figure 3.4: Observability plots for six different measurement trajectories

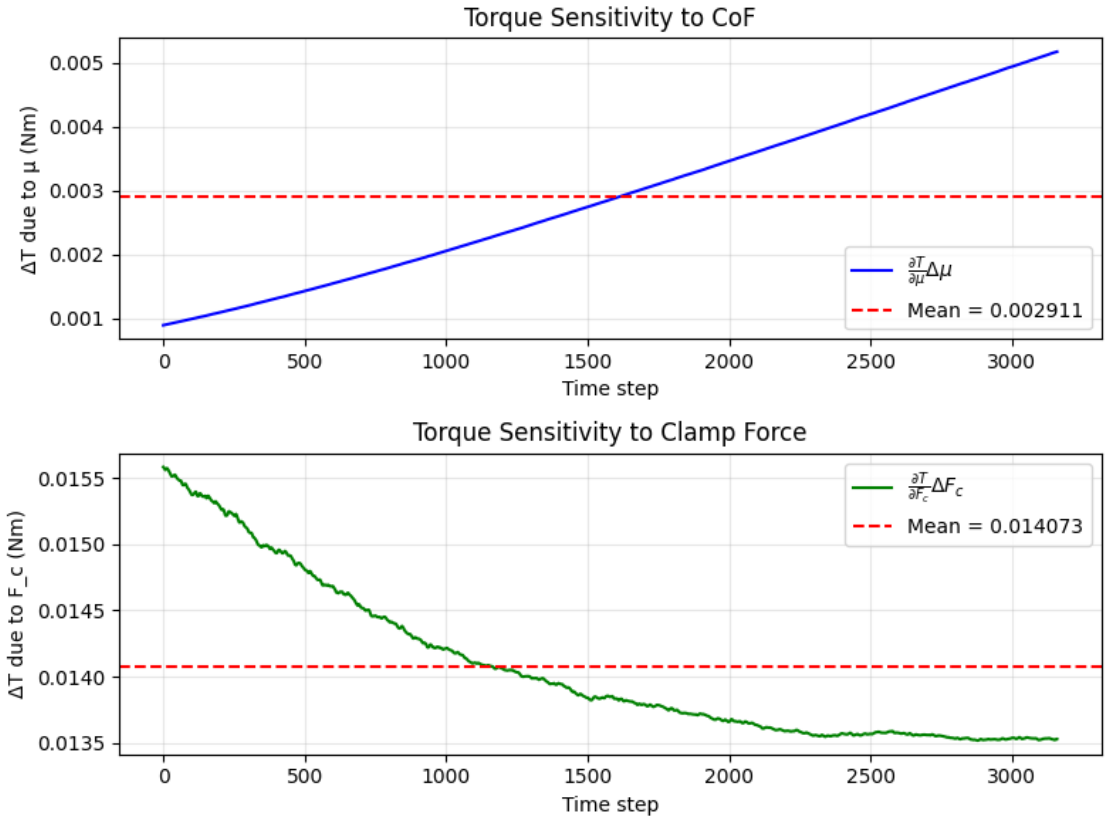


Figure 3.5: Output sensitivity to average per time sample change in states

On the other hand, clamp force has a dominant influence on torque and can be estimated with greater confidence.

3.5 LMI-Based Luenberger Observer

The observer performance is evaluated under multiple cases.

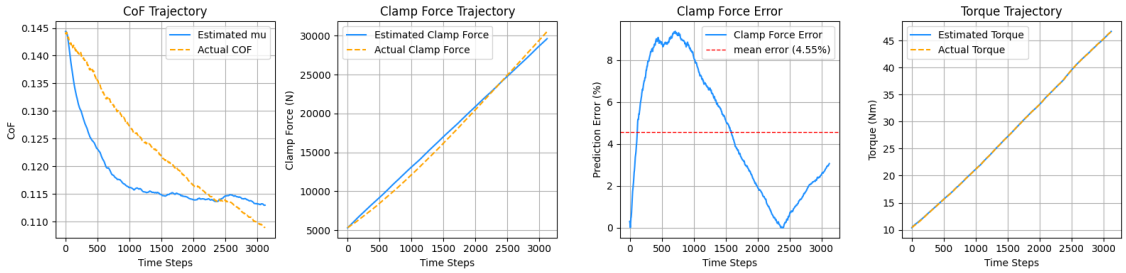


Figure 3.6: Estimations using the correct initial guess on both the states

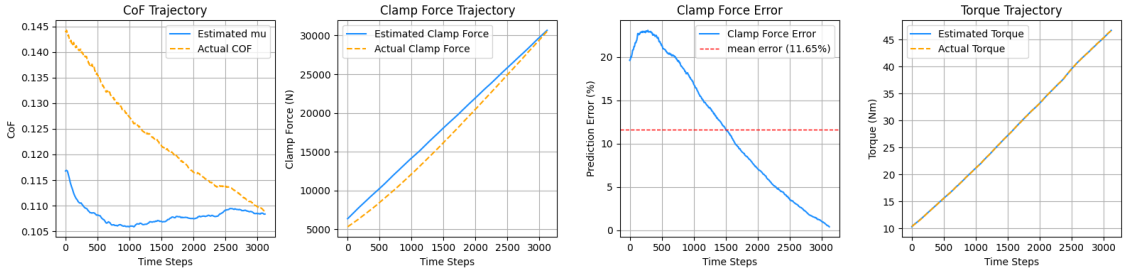


Figure 3.7: Estimations using the incorrect initial guess on clamp force of +20% error

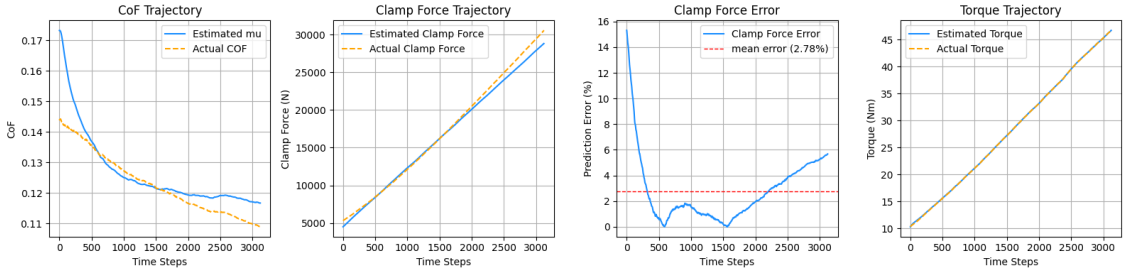


Figure 3.8: Estimations using the incorrect initial guess on CoF +20% error

Some observations from the above plots can be made.
The influence of the initial state estimates on the observer performance is significant. With the correct initial guess for both states, the average clamp force estimation error is approximately 4.55%, whereas an incorrect initial guess in F_c leads to a significantly higher error of 11.65%. In contrast, an incorrect initial guess in μ results in a trajectory similar to that of the correct initialization case. This indicates that the estimation accuracy is highly sensitive to the initial value of the clamp force.
Another key observation is that the clamp force is not sufficiently corrected by the observer during estimation. Instead, the estimated F_c primarily follows the system

dynamics, with only minor corrections from the measurement. This is evident in Fig. 3.7, where the estimated clamp force follows a nearly linear trajectory dictated by the model dynamics, with minimal influence from the measurement updates. Furthermore, due to the coupling between clamp force and CoF, errors in the clamp force estimation propagate into the CoF estimate. Since the observer minimizes the output error, it compensates for inaccuracies in F_c by adjusting μ. Due to sign This results in biased CoF estimates when the clamp force is not correctly estimated. The above plots correspond to a sample with an approximately linear torque trajectory. A similar analysis is performed for a nonlinear torque profile (see Sample-5 in Fig. 2.3), where deviations in estimation behavior are further observed.

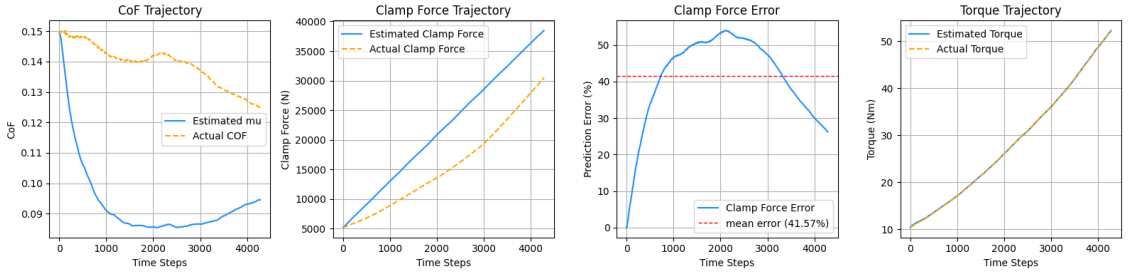


Figure 3.9: Estimations using the correct initial guess on both the states for nonlinear torque trajectory

This result further confirms that the clamp force estimates are only weakly corrected by the measurements and primarily follow the system dynamics. The limitation arises from the imbalance in sensitivity, where torque variation is dominated by clamp force while the contribution of CoF is very small and comparable to noise. Consequently, errors in clamp force estimation propagate into CoF, leading to biased state reconstruction.

3.6 Dynamic Inversion

3.6.1 Estimation of CoF from the output

As discussed in the previous chapter, the tuning of the covariance matrices Q_α and R_z plays a critical role in determining the observer performance. In the present study, the tuning was carried out with the objective of minimizing the clamp force estimation error. Consequently, the covariance values were selected such that the observer relies more on the system dynamics while capturing only the overall trend of the measurement trajectory rather than high-frequency variations. The estimation results are illustrated for two representative cases: one corresponding to a nearly linear torque–angle trajectory and the other exhibiting nonlinear behavior.

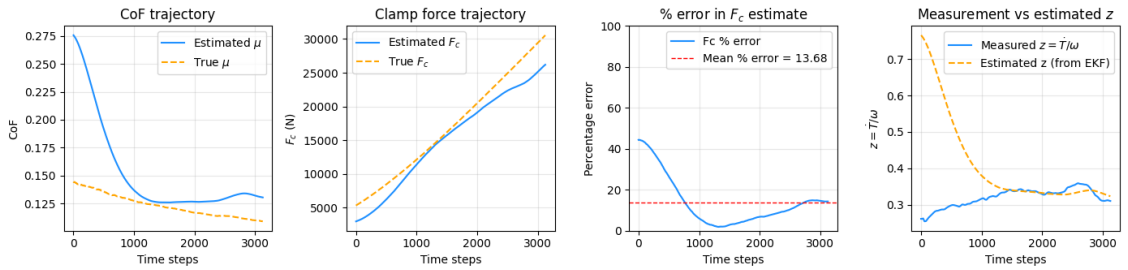


Figure 3.10: Estimations for approximately linear torque trajectory with an initial guess error of +20%

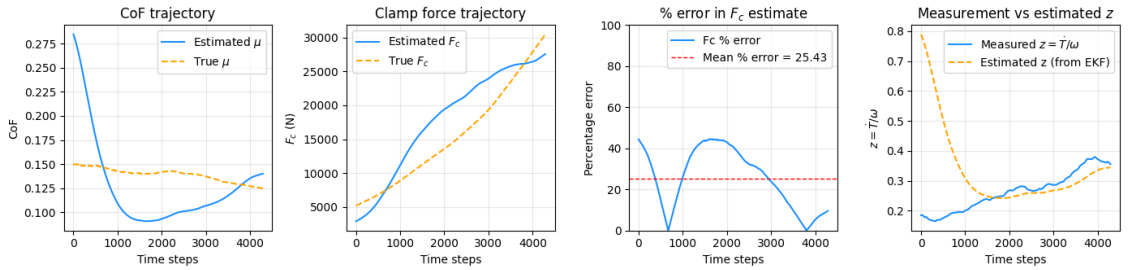


Figure 3.11: Estimations for nonlinear torque trajectory with an initial guess error of +20%

It is clearly evident from the plots in figures 3.10 and 3.11 that the clamp force estimations are better in case of linear torque trajectory when compared to the nonlinear variation in torque. This behavior can be attributed to the chosen covariance tuning. Since the observer is tuned to rely heavily on the model dynamics, it performs well when the underlying system behavior is close to the assumed model structure, i.e., in the case of a linear torque response. However, for nonlinear torque trajectory, deviations from the model which considers a constant joint stiffness are not sufficiently corrected due to the lower weighting assigned to measurement updates, resulting in larger estimation errors in the clamp force.

Furthermore, the clamp force exhibits limited correction during the estimation process, which can be observed particularly in the nonlinear case. This behavior is reinforced by the sensitivity characteristics of the system as discussed in the previous sections. Although the torque evolution is dominated by the clamp force, the observer relies weakly on measurement updates due to its tuning, resulting in limited correction of F_c. In contrast, the contribution of μ to the torque variation remains small due to its extremely low variation per time step. As a result, even when measurement updates are considered, the observer tends to adjust μ to compensate for output errors, while the clamp force remains largely influenced by its initial estimate and dynamics. In the case of nonlinear torque trajectories, this effect becomes more pronounced, as the model mismatch is not adequately corrected due to the low measurement weighting.

Moreover, it can be observed that the CoF does not accurately follow the true trajectory in both cases and shows a considerable deviation. This is a consequence of both the reduced reliance on measurements and the limited contribution of μ to the torque signal. Since μ is indirectly estimated from the measurement z_k, its estimation becomes highly sensitive to inaccuracies in F_c, leading to biased behavior.

In order to check the impact of initial state guess on the estimates, the observer is provided the measured initial guess on α, which can be seen in the plot of CoF in the following figure. It is observed that with a correct initial guess on the CoF, the clamp force estimation accuracy shows only marginal improvement, but the estimates deviate in the later stages as the clamp force continues to follow the assumed dynamics. This behavior is further influenced by the limited corrective contribution from measurements and the propagation of estimation errors between the coupled states.

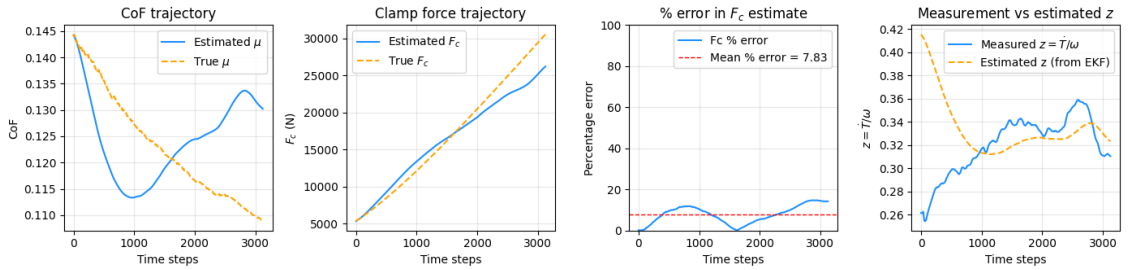


Figure 3.12: Estimations for approximately linear torque trajectory with accurate initial CoF guess

In order to improve the observer's ability to capture the nonlinear variations in the torque trajectory, the measurement noise covariance R_z was reduced, thereby increasing the reliance on the measurement signal. However, as shown in Fig. 3.13, this results in a deterioration of the estimation performance. This behavior can be attributed to the noisy nature of the measurement $z_k = \dot{T}/\omega$, which is obtained through discretization and averaging of torque increments.
Reducing R_z causes the observer to place excessive trust in this noisy measurement, leading to overfitting of high-frequency variations that do not represent the

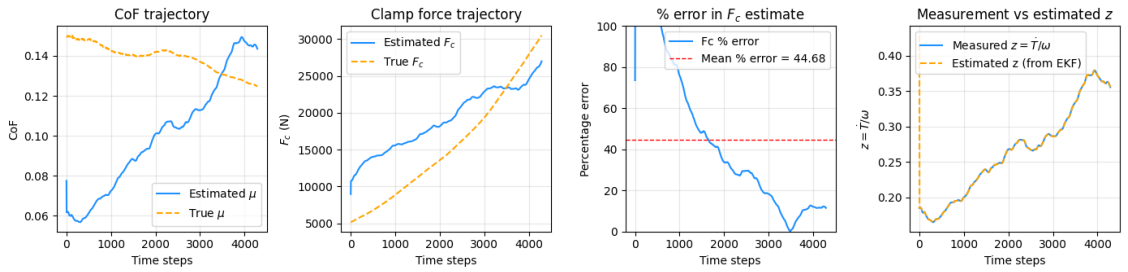


Figure 3.13: Estimation for nonlinear torque trajectory with high reliance on the measurements

true system behavior. As a consequence, the estimation becomes more sensitive to measurement noise, which adversely affects both the transient and steady-state performance. In particular, the initial estimates are distorted despite being initialized close to the true values, as the observer aggressively adjusts the states in response to noisy measurements.
Therefore, although reducing R_z enables the observer to respond more strongly to nonlinear variations, it also amplifies the impact of measurement noise, resulting in poorer overall estimation performance.

3.6.2 Estimation of Clamp force from the output

In this case the clamp force is estimated directly from the measurements derived from the measured torque. Similar to the previous case of CoF estimation, the priority is towards minimum error in the clamp force. So, the observer is designed to depend more on the dynamics, while considering the overall variation in the torque.

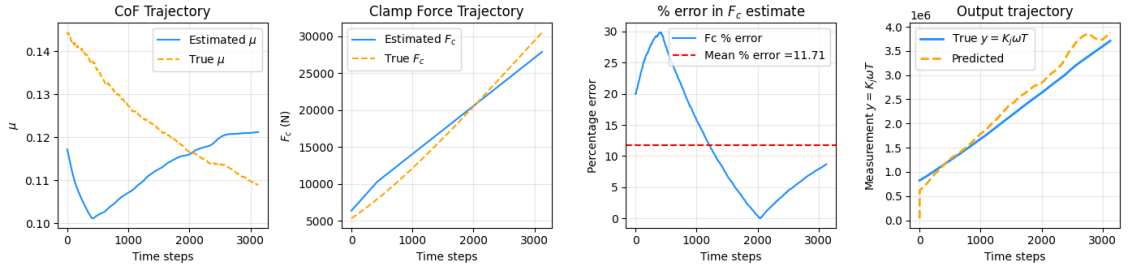


Figure 3.14: Estimations for approx. linear torque trajectory with +20% error in initial CoF guess

For the same covariance values, the estimations are carried out in case of nonlinear torque trajectory.

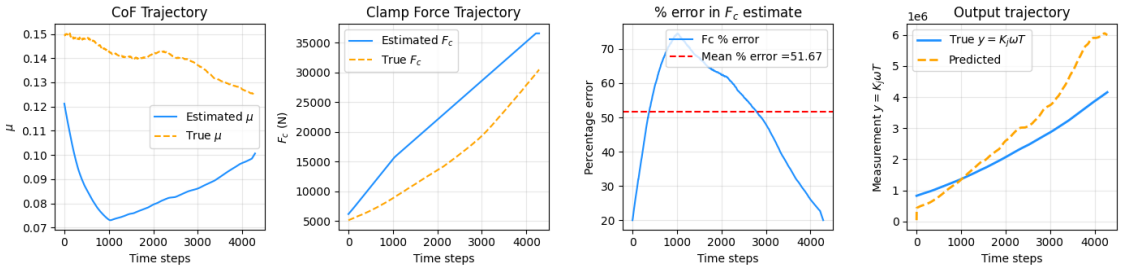


Figure 3.15: Estimations for nonlinear torque trajectory with +20% error in initial CoF guess

The above plots are considering the initial states have an error of around 20% on F_c. It can be inferred from both plots that the clamp force estimation quality has improved, in case of a linear torque trajectory. However, the estimation qualities in case of nonlinear curves have deteriorated. The estimations for linear torque trajectory works well because of almost constant increase in the measurement y_k which is evaluated as $K_j\omega T$. Moreover, since the value of K_j almost remains constant in the linear torque case, so the measurements act almost the true. However, due to the noisy $\dot{T}$ which is multiplied by the huge values of F_c in the output expression $g(F_c)$, so any minor noise in torque derivative will be amplified, which is also visible at the final region in the output plot. Another observation from the linear torque plot is that the observer still struggles to follow the actual clamp force trajectory.

The possible cause for the poor estimation in the nonlinear torque case is that $\dot{T}$ appears inside the nonlinear function $g(F_c)$, causing noise and approximation errors in the torque derivative to propagate into the measurement model. This affects the innovation term and leads to incorrect state updates. Furthermore, the nonlinearity in $g(F_c)$ amplifies the effect of errors in both $\dot{T}$ and F_c, resulting in unstable or biased estimates. This is further aggravated by the nonlinear joint stiffness present in the measurement, reducing the reliability of the measurement information for correcting the state estimates.

3.7 Rao-Blackwellized Particle Filter

The estimations are carried out using 1000 particles for the coefficient of friction. The covariance matrices are tuned to achieve reliable clamp force estimation for the different cases presented in this section.
Since the CoF particles are generated randomly based on a predefined distribution, multiple simulations are performed using different random realizations. The test cases for this observer include nonlinear and linear torque trajectories with correct and incorrect initial guess. The results obtained across these simulations exhibit consistent behavior, with similar estimation trends and error characteristics as discussed in the following paragraphs.
For the case with incorrect initialization, the RBPF demonstrates a partial ability to recover from initial errors. However, the estimation quality depends strongly on the torque trajectory. In the nearly linear case (Fig. 3.16), the clamp force error

remains moderate and relatively stable. The coefficient of friction follows the model, but the particle spread remains constant. This is because the linear torque trajectory provides insufficient variation in the measurement, leading to similar weights across particles and preventing effective state separation.
In contrast, the nonlinear case (Fig. 3.17) shows improved estimation accuracy when compared to the previous observers despite the incorrect initialization. The time-varying nature of the torque introduces stronger variations in the measurement, resulting in larger differences in particle likelihoods. This enables more effective resampling and concentration of particles around feasible solutions, leading to improved state reconstruction. However, this improvement arises from the increased ability to distinguish between particle hypotheses due to excitation.
Notably, in both cases, the predicted torque closely matches the measured torque despite bias in the estimated states. This indicates that the filter compensates for errors in one state by adjusting the other, highlighting the non-uniqueness of the solution.

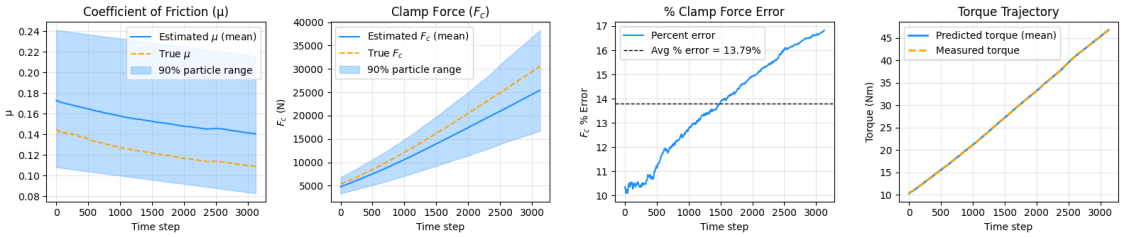


Figure 3.16: RBPF estimates for approx. linear torque trajectory with +20% error in initial CoF guess

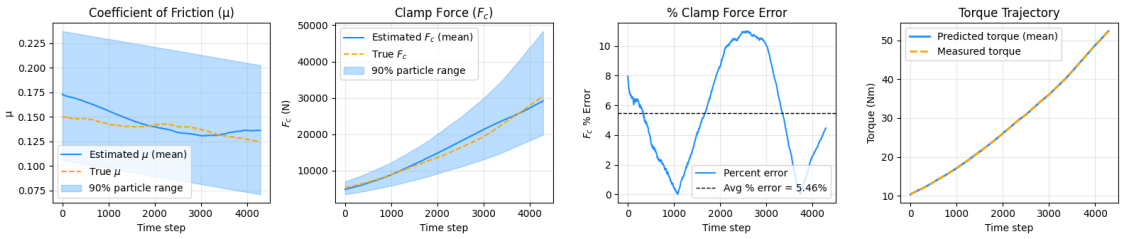


Figure 3.17: RBPF estimates for nonlinear torque trajectory with +20% error in initial CoF guess

For the case with correct initialization, the estimation performance improves significantly when the torque trajectory is approximately linear. The clamp force is accurately tracked with low error when compared to the previous cases and observers. Additionally, unlike the LMI-based Luenberger observer, the RBPF actively updates the CoF based on measurement information.
However, for nonlinear torque trajectories, even with correct initialization, the estimation accuracy is compromised when compared to the same trajectory with incorrect initial guess. This behavior suggests that while excitation improves discrimination among particles, it also increases sensitivity to model mismatch. As a result,

particles may diverge from the true state despite a good initial estimate, leading to degraded performance.

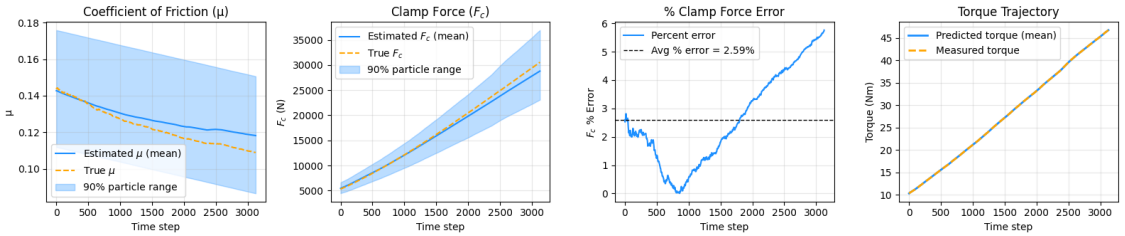


Figure 3.18: RBPF estimates for approx. linear torque trajectory with almost exact initial guess

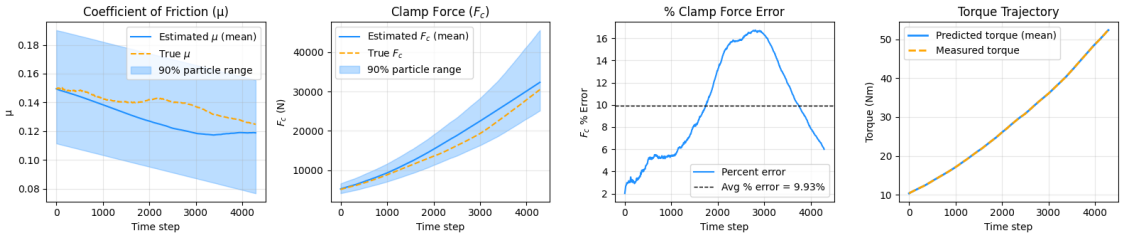


Figure 3.19: RBPF estimates for nonlinear linear torque trajectory with almost exact initial guess

Apart from the observations discussed above, an additional asymmetry was identified in the RBPF behavior for the nonlinear torque trajectory. In particular, when the initial CoF was chosen below the true initial value, the estimator was unable to recover the correct state and the resulting estimation error remained high. This behavior is consistent with the output trajectory based observability analysis shown in Fig. 3.4 for Sample-5. In that figure, it can be seen that values of CoF less than 0.14 has lower error, meaning the resulting output trajectory is almost indistinguishable. Due to this, the RBPF is unable to sufficiently distinguish between these initial hypotheses and therefore cannot drive the particles back toward the true state. This further suggests that the recovery from incorrect initialization is strongly dependent on whether the initial guess lies outside the trajectory-dependent unobservable region. It can be observed in all cases that the spread of clamp force particles increases over time, while the spread of the CoF particles remains relatively consistent. This behavior arises because the clamp force is obtained by scaling the torque using the CoF particles. As the torque increases during tightening, any uncertainty in CoF propagates and gets amplified in the clamp force, leading to a widening spread.
Overall, the results indicate that the performance of the RBPF is governed by the interaction between initialization accuracy, the information content of the torque trajectory, and the location of the initial guess in the state space. In particular, nearly linear trajectories preserve good initial estimates but are unable to correct large initialization errors, whereas nonlinear trajectories improve correction capability through increased excitation, but may also degrade accurate initial estimates due to model mismatch. In addition, the recovery from incorrect initialization is not

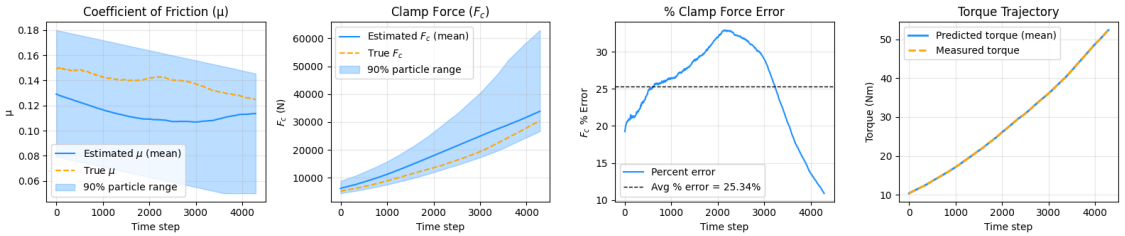


Figure 3.20: RBPF estimates for nonlinear linear torque trajectory with initial state guess lower than the actual

symmetric: for the nonlinear case, initial CoF values below the true state were not recoverable comparatively well, which is consistent with the trajectory-dependent unobservable region identified in the output trajectory based observability analysis. Furthermore, accurate torque prediction does not necessarily imply accurate state estimation, as multiple combinations of F_c and μ can produce similar outputs. Nevertheless, compared to the previous observer designs, the RBPF provides an improvement by maintaining multiple hypotheses through particles, allowing better adaptation in the presence of nonlinearity.

3.8 Benchmark Comparison

A benchmarking study is carried out by comparing the proposed observer-based methods against the conventional assumption of constant friction. The mean of the CoF values from the available data is used as a nominal constant CoF value, which provides a baseline clamp-force estimate for comparison with the observer-based methods. The same nominal CoF value is also used as the initial CoF guess for the observers.

Since earlier analysis showed that errors in the initial clamp-force guess have a considerably larger impact on the estimation accuracy than errors in the initial CoF guess, the observers that estimate clamp force as a direct state are initialized by first assigning this prior CoF value and then computing the corresponding initial clamp force from the measured torque. The resulting average percentage error in clamp-force estimation for each tightening sample is summarized in Table 3.1.

In Table 3.1, the second column represents the error in the initial CoF when mean value from the available data is used. This highlights the quality of the observer estimates when the prior CoF information is taken from a practical assumption, which may in some cases deviate substantially from the actual tightening condition. It is also important to note that the RBPF values may vary slightly across runs because the particle samples are generated randomly; an additional variation of approximately $\pm 1.5\%$ may therefore be expected.

The results show that the constant-friction assumption based on the standard leads to relatively large clamp-force errors for most of the samples, confirming that a fixed CoF is not sufficient for reliable estimation. Also, this errors are very sensitive to the true values, and may increase if the actual variation is very nonlinear. In contrast, the observer-based methods generally provide lower errors. However, these lower

Table 3.1: Comparison of average percentage error in clamp-force estimation

Sample	Initial CoF % error in the initial CoF	Average % Error in F_c over the trajectory				
		Constant Friction	LMI-Based Luen-berger	EKF with α as state	EKF with F_c as state	RBPF
1	3.03	11.94	5.35	7.85	9.00	5.36
2	3.03	8.63	5.04	8.74	8.40	3.74
3	0.28	14.08	3.75	8.15	7.89	6.14
4	6.12	14.46	2.17	5.52	6.06	3.96
5	6.18	3.44	44.02	30.87	46.70	18.47
6	3.73	7.67	10.24	9.35	12.76	4.98

errors should be interpreted with care. The apparent improvement is not only due to better state estimation, but also because the initialization error in CoF has a comparatively weaker influence on the clamp-force estimate than an error in the initial clamp force itself.
For the approximately linear tightening samples, the LMI-based Luenberger observer gives particularly low errors. Nevertheless, this does not necessarily imply strong measurement-based correction. As observed in the earlier results, the Luenberger observer relies heavily on the assumed motion model and only weakly on the measurement updates. Therefore, its low error in these cases is mainly a consequence of the tightening trajectory being close to the assumed model behavior.
A similar caution applies to the EKF-based formulations. Although some of the estimates work better than the constant friction model, the estimates are still strongly influenced by the model structure and initialization.
The RBPF generally yields lower errors for the approximately linear trajectories, since the limited variation in the measured torque provides insufficient information to strongly discriminate between particle hypotheses. However, in the nonlinear case (Sample-5), the RBPF performs poor as the initial state lies in the unobservable region. However, it can be seen from the previous chapter that RBPF can work well, in case of the nonlinear torque trajectory with the initial state in the observable region as defined in the output trajectory based observability.
Overall, the comparison indicates that no single observer performs best for all tightening samples. The LMI-based Luenberger observer and the EKF-based approaches can yield low errors when the tightening trajectory is close to the assumed model behavior, but these results are strongly model-driven and should not be interpreted as evidence of strong measurement-based correction. The RBPF works well if the initial states are close to the actual in case of a linear torque trajectory. In case of nonlinear torque trajectory, the initial state matters, as start from the unobservable region as defined in the output trajectory based observability results.

4
Conclusion

This thesis investigated the challenges associated with estimating clamp force and CoF from measured torque and tightening speed. The estimation problem was found to be inherently difficult due to the coupling between the states in the torque expression and the limited independent information available from the measurements. A comprehensive study was carried out, including system formulation, observability analysis, sensitivity evaluation, and the implementation of multiple estimation techniques.

The system was first analyzed through different mathematical formulations linking torque, clamp force, and CoF, with the objective of reducing model complexity while retaining physical interpretability. A data-driven approach using Gaussian Process Regression (GPR) was also explored to estimate CoF directly from torque measurements. However, the predicted CoF lacked sufficient accuracy, mainly because the underlying variations in CoF were extremely small and difficult to separate from noise. In addition, GPR does not explicitly model temporal evolution, which limits its ability to capture the fine-scale dynamics of the CoF and the variation introduced by nonlinear torque trajectories.

The observability of the system was evaluated using both local and trajectory-based approaches. The local observability analysis, based on Lie derivatives, indicated that while the system is theoretically observable, the observability matrix elements are close to zero, suggesting near-marginal observability. The trajectory-based analysis further showed that the practical unobservable region is not fixed, but varies with each measured torque trajectory. This makes it difficult to define a generalized unobservable region that remains valid across different tightening conditions and highlights the trajectory-dependent nature of the estimation problem.

To complement the observability study, a sensitivity analysis was performed to assess the influence of state variations on the output. By scaling the analytical sensitivities with the actual per-time-step variations of the states, it was shown that the torque evolution is predominantly governed by the clamp force, while the contribution of CoF is significantly smaller. This low effective contribution of CoF limits the information available for its estimation and increases the ambiguity in separating the two states.

Based on these insights, several observer-based estimation techniques were implemented. The LMI-based Luenberger observer was found to be highly sensitive to initial conditions and to rely strongly on the model dynamics, with only limited correction from measurements. As a result, its low clamp-force error in the nearly linear cases should be interpreted carefully, since the performance was mainly driven by the validity of the assumed model rather than by strong measurement-based adaptation.

The Extended Kalman Filter (EKF) was implemented using modified system formulations to reduce the explicit coupling between the states. Although this enabled the estimation of individual parameters, the observer performance remained limited by the measurement model. In particular, the use of torque derivatives introduced significant noise amplification, which degraded both the direct CoF estimation and the clamp-force estimation derived from the output model.
To better handle nonlinearity and uncertainty in CoF, a Rao-Blackwellized Particle Filter (RBPF) was implemented. The RBPF provided improved flexibility by maintaining multiple hypotheses through particle representations. Its performance depended strongly on initialization accuracy and the model mismatch. For nearly linear torque trajectories, good initial estimates worked well, but poor initial states were not effectively corrected due to weak particle discrimination. In contrast, nonlinear trajectories improved discrimination and enabled partial recovery from poor initialization, particularly when the initial CoF exceeded the true value. However, the recovery was not symmetric: initial CoF values below the true state were not recoverable, which is consistent with the trajectory-dependent practical unobservable region identified in the observability analysis.
A benchmark comparison against the standard constant-friction assumption showed that a nominal CoF value taken from the standard leads to relatively large clamp-force errors for all samples. Although some observer-based methods produced lower errors, these results should not always be interpreted as strong correction from measurements. In many of the nearly linear cases, the lower error mainly reflects the weak influence of CoF initialization error on clamp-force estimation and the strong reliance of the observers on the model. In this sense, the RBPF was the only method that showed a clear ability to make effective use of measurement information in the nonlinear case.
Overall, the study shows that the limitations in estimation performance are not solely due to the choice of observer, but are fundamentally rooted in the system characteristics. Weak practical observability, low effective CoF contribution to the measured output, measurement noise, trajectory-dependent non-identifiability, and model mismatch together impose inherent constraints on the achievable estimation accuracy.

4.1 Future Work

The findings of this thesis suggest several directions for future research.
A primary direction is the improvement of the system formulation itself. The current formulations are able to describe certain tightening cases reasonably well, but the results indicate that their validity is not consistent across all tightening samples. In particular, the system behavior, estimation performance, and trajectory-based observability vary significantly from one trajectory to another. Future work should therefore focus on developing a more general system formulation that is capable of representing the tightening dynamics across a wider range of tightening samples. This would require extensive testing over a larger number of trajectories with different frictional behavior, stiffness evolution, and torque-angle characteristics. The quality of such formulations can be evaluated based on the practical observability

region obtained from the output trajectory-based observability analysis. A formulation that yields smaller and more consistent unobservable regions across different tightening samples would be more suitable for better state estimation. Such an approach would help identify formulations that are not only physically meaningful, but also practically estimable under varying tightening conditions.

Another important direction is the improvement of model fidelity. The current formulations rely on simplified assumptions, such as constant or weakly varying joint stiffness in the elastic region. However, the experimental results show that the torque-angle behavior is not always linear, and the effective stiffness may vary between tightening samples. Future work should therefore consider nonlinear or state-dependent stiffness models that more accurately capture the actual joint behavior. Such an extension would increase the complexity of the estimation problem, as stiffness may need to be introduced as an additional state or varying parameter, thereby making the analysis and observer design more challenging. At the same time, this difficulty may be reduced by incorporating additional information about the joint stiffness into the estimation framework. Prior work in this direction suggests that stiffness-related information can provide an additional source of system information and this may improve both observability and estimation reliability while minimizing the coupling between the states.

Apart from improving the stiffness modeling, another important direction for future work is the development of formulations that reduce or break the bilinear dependency between clamp force and coefficient of friction in the output expression. Since this coupling is a major source of ambiguity in the estimation problem, reformulating the system in a way that weakens this dependency could significantly improve state identifiability. A possible extension is to incorporate thermal effects during tightening, as temperature changes at the contact interfaces may affect friction behavior. By simultaneously modeling and measuring relevant thermal variables, additional information may be introduced into the system, which could improve the estimation of CoF and hence clamp force.

A further important direction is the introduction of additional excitation into the tightening process. The RBPF results suggest that the filter performs better when the torque trajectory contains stronger nonlinear variation, as this improves particle discrimination and state separation. Future work could therefore investigate whether the tightening dynamics can be enriched by varying the angular speed, rather than assuming constant rotational speed. Such input variation may provide additional information in the measurements and improve distinguishability between the states. However, introducing speed variation may also increase uncertainty in the CoF evolution, and this trade-off should be investigated carefully.

Another approach could be to model friction more accurately using a combination of physical understanding and data-driven information. The current friction dynamics are simplified and may not fully capture the trajectory-dependent behavior observed in the experiments. A hybrid friction model that combines tribological

insight, physical constraints, and experimental data may provide a more realistic description of CoF evolution and reduce model mismatch.

In the present study, the CoF is indirectly calculated from torque and clamp force measurements, both of which are affected by measurement noise and estimation uncertainty. In addition, as discussed in the previous chapter, the torque equation used to compute the CoF is itself associated with an error of approximately 1-2%. Consequently, the resulting CoF trajectories may already contain propagated errors, which limits the ability to identify the true friction evolution. Future work could therefore investigate alternative means of measuring or calibrating CoF, for example through acoustic emission or other friction-sensitive sensing methods. Such approaches may provide more accurate friction measurements and help reveal consistent patterns in CoF evolution that are currently masked by indirect estimation errors.

A further direction could be the exploration of conditional data-driven models. For example, instead of learning CoF only as a function of torque, future work could investigate conditional Gaussian Process Regression or related models that use additional inputs such as angular speed, angular displacement, or short histories of the tightening signal. Such an approach may improve local prediction quality, although it would still require better CoF calibration and richer measurement information to avoid inheriting the same identifiability limitations observed in the present study.

Another important direction for future work is the improvement of measurement quality through the use of more accurate and less noisy sensors. Moreover, although the angular speed is treated as constant in the present study, practical hardware limitations introduce small variations during tightening. Since these deviations may affect both the system dynamics and the estimation results, their impact can be investigated explicitly in future studies.

Overall, future work could focus on improving formulation generality, introducing excitation, measurement richness and CoF calibration in order to overcome the fundamental limitations identified in this thesis.

Bibliography

[1] M. Kumar, *Friction in threaded fasteners: Influence of materials and tooling*. KTH Royal Institute of Technology, 2022.

[2] M. Al-Barghouthi, "Extended and Unscented Kalman Filtering for Estimating Friction and Clamping Force in Threaded Fasteners,"

[3] E. Wallbom, "Internal report on friction, Clamp force estimation, and Frequency measurement, Atlas Copco 2025,"

[4] R. S. Shoberg, "Engineering fundamentals of threaded fastener design and analysis." Available via VDI 2230 resources.

[5] PCB Piezotronics, Inc., "Threaded fastener design and analysis," tech. rep., PCB Piezotronics, Inc., 2011. Accessed: 2026-02-24.

[6] K. Zhang, "Machine Learning and Hybrid Model for Snug Detection in Advanced Tightening Optimization: A research on how rule-based algorithm helps to improve ML models for snug detection,"

[7] A. Blom, "Speed dependent friction in bolt joints,"

[8] "VDI 2230-1:2015 – Systematic Calculation of Load-Carrying Capacity of Bolted Joints – Part 1: Statics," 2015.

[9] M. Dyberg and E. Troillet, "Comparative analysis of contact and non- contact method for determining the natural frequencies of a bolted joint,"

[10] R. Kellermann and H.-C. Klein, "Investigations on the influence of friction on the clamp force and tightening torque of bolted joints," *Konstruktion*, vol. 7, no. 2, pp. –, 1955.

[11] "Fasteners – torque/clamp force testing," 2005.

[12] Y. Qiao, B. Zhao, D. Deng, and W. Ouyang, "Prediction model for nonlinear variation of bolt preload with tightening torque based on mechanism and data fusion," *Scientific Reports*, vol. 15, p. 8020, Mar. 2025.

[13] R. Dai, G. Evangelisti, and S. Hirche, "Physically Consistent Modeling & Identification of Nonlinear Friction with Dissipative Gaussian Processes," May 2024. arXiv:2405.17199 [eess].

[14] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[15] M. Anguelova, "Observability and identifiability of nonlinear systems with applications in biology,"

[16] R. Hermann, "Nonlinear Controllability and Observability,"

[17] A. Isidori, *Nonlinear Control Systems*. Communications and Control Engineering, London: Springer London, 1995.

[18] H. K. Khalil, "Nonlinear systems. Hauptbd.," Upper Saddle River, NJ: Prentice Hall, 3. ed ed., 2002. Num Pages: 750.

[19] S. Zeng, "Observability Measures for Nonlinear Systems," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 4668–4673, IEEE, Dec. 2018.
[20] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.
[21] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.
[22] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *Journal of Control and Decision*, vol. 5, no. 1, pp. 42–60, 2018.
[23] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley, 2006.
[24] A. Doucet, N. de Freitas, K. Murphy, and S. Russell, "Rao-blackwellised particle filtering for dynamic bayesian networks," *Proceedings of UAI*, 2000.
[25] Y. Qiao, B. Zhao, D. Wu, W. Ouyang, H. Li, and S. Zhang, "Nonlinear bonding analysis of bolt connection and secondary tightening strategy," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 19, pp. 5583–5599, Aug. 2025.
[26] N. Dressler and L. Eriksson, "Implementation of a bolted joint model in Modelica," pp. 175–183, Oct. 2023.
[27] A. Chambers, "Preloaded Joint Analysis Methodology for Space Flight Systems,"
[28] T. Held, J. Peth, and C. Friedrich, "Preload Stability of Modern Bolted Joints," *Athens Journal of echnology & Engineering*, vol. 9, pp. 197–212, Aug. 2022.
[29] G. R. Toth, "Controlled Tightening Over the Yield Point of a Screw: Based on Taylor's Series Expansions," *Journal of Pressure Vessel Technology*, vol. 125, pp. 460–466, Nov. 2003.
[30] G. R. Toth, "Torque and Angle Controlled Tightening Over the Yield Point of a Screw—Based on Monte-Carlo Simulations," *Journal of Mechanical Design*, vol. 126, pp. 729–736, July 2004.
[31] H. Miao, X. Xia, A. S. Perelson, and H. Wu, "On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics," *SIAM Review*, vol. 53, pp. 3–39, Jan. 2011.
[32] C. Kreutz, A. Raue, and J. Timmer, "Likelihood based observability analysis and confidence intervals for predictions of dynamic models," *BMC Systems Biology*, vol. 6, no. 1, p. 120, 2012.
[33] T. Mullari, Kotta, Z. Bartosiewicz, and E. Pawłuszewicz, "The concepts of Lie derivative for discrete-time systems; pp. 253–265," *Proceedings of the Estonian Academy of Sciences*, vol. 61, pp. 253–265, Nov. 2012.
[34] S. Di Gennaro, M. Cappelli, and B. Castillo–Toledo, "Advanced control system: theory and application to nuclear reactors," in *Instrumentation and Control Systems for Nuclear Power Plants*, pp. 393–524, Elsevier, 2023.

A

Appendix

A.1 GPR Estimates for the other samples

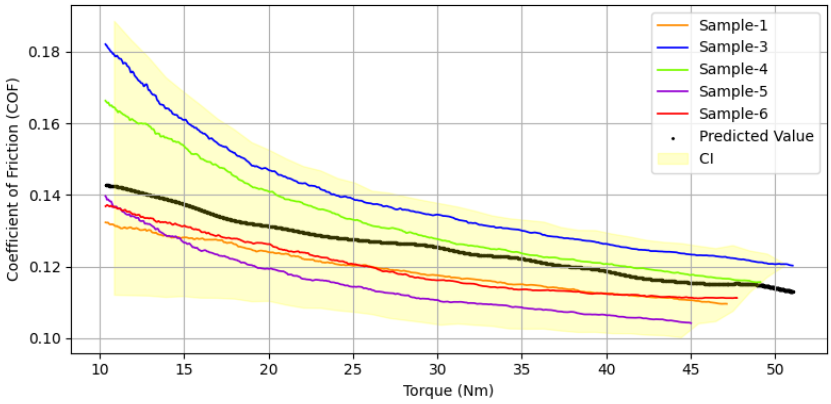


Figure A.1: GPR predictions of CoF for sample-2 with Confidence Interval (CI) and training data

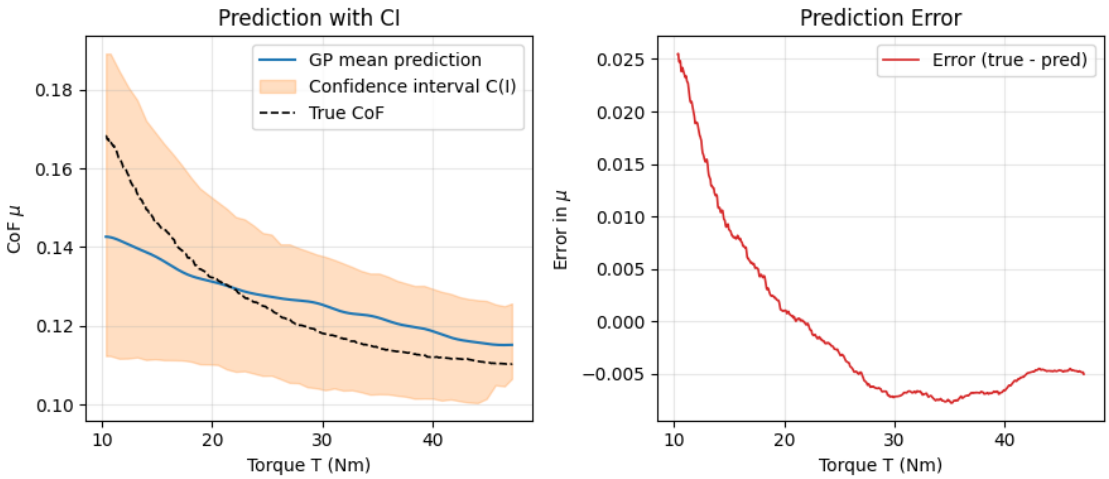


Figure A.2: GPR predictions of CoF plotted with the actual measurements for sample-2

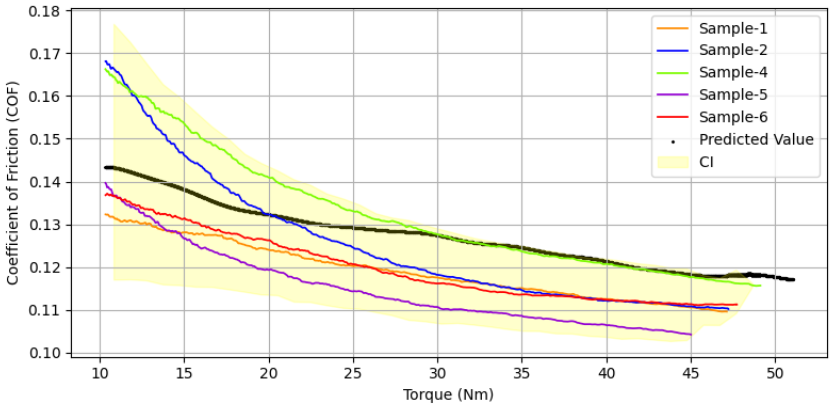


Figure A.3: GPR predictions of CoF for sample-3 with Confidence Interval (CI) and training data

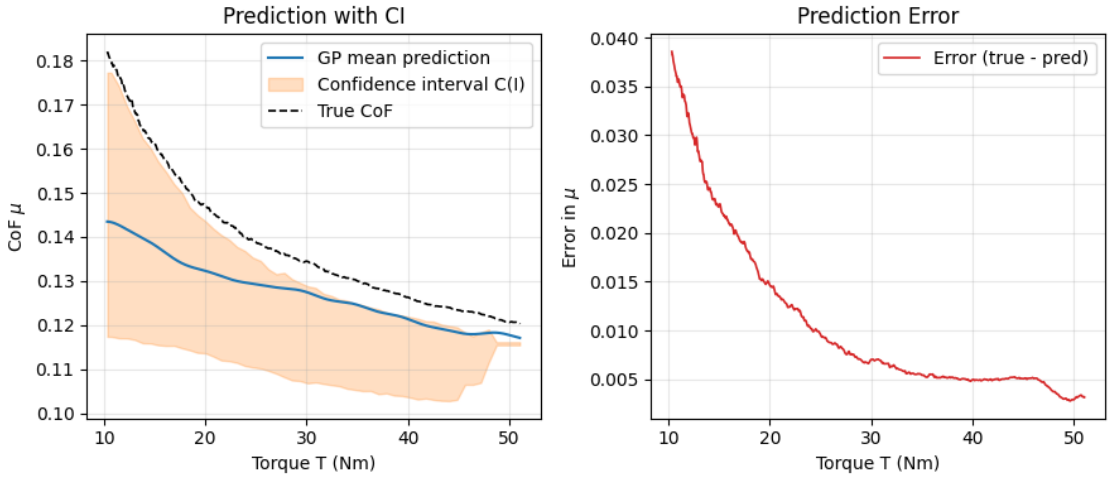


Figure A.4: GPR predictions of CoF plotted with the actual measurements for sample-3

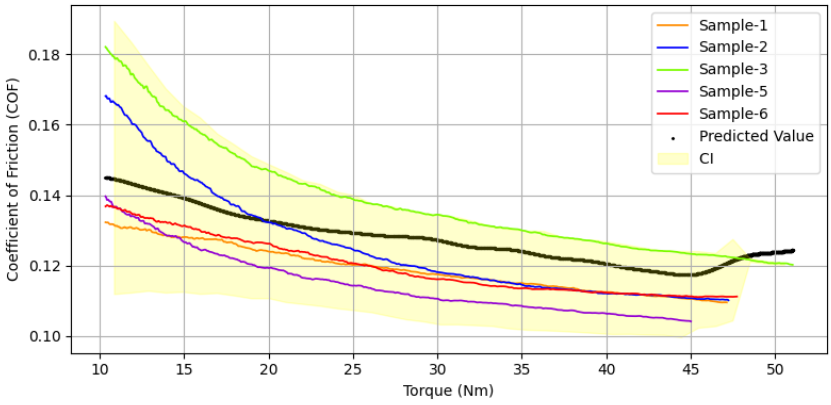


Figure A.5: GPR predictions of CoF for sample-4 with Confidence Interval (CI) and training data

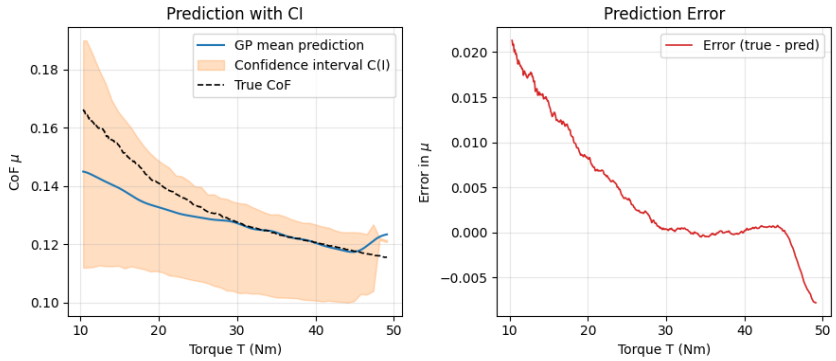


Figure A.6: GPR predictions of CoF plotted with the actual measurements for sample-4

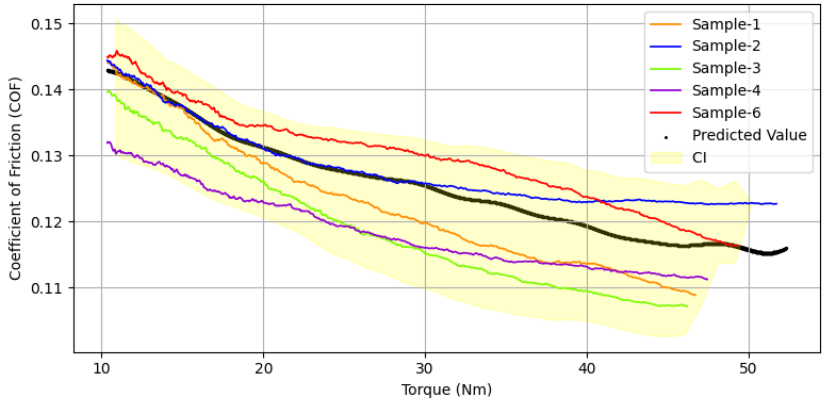


Figure A.7: GPR predictions of CoF for sample-5 with Confidence Interval (CI) and training data

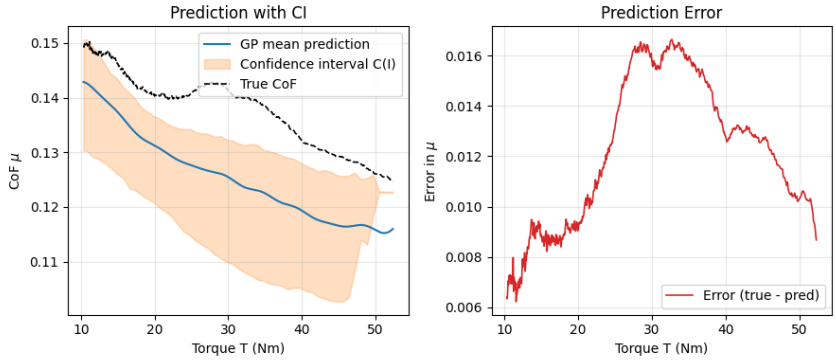


Figure A.8: GPR predictions of CoF plotted with the actual measurements for sample-5

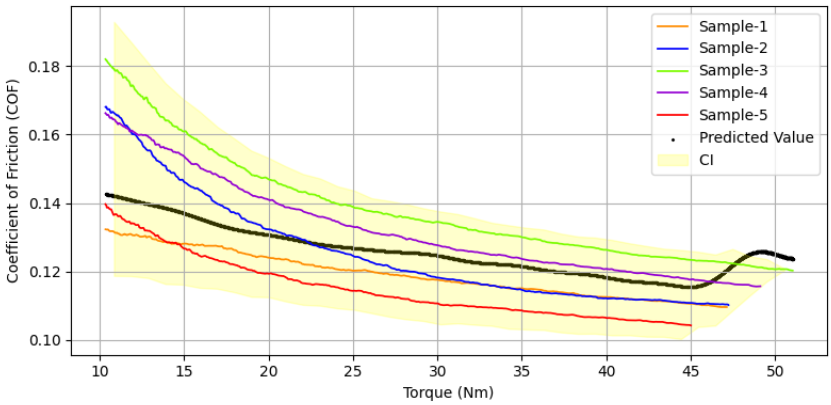


Figure A.9: GPR predictions of CoF for sample-6 with Confidence Interval (CI) and training data

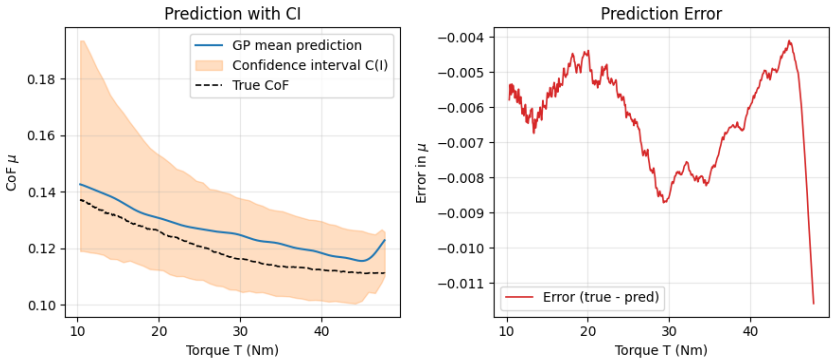


Figure A.10: GPR predictions of CoF plotted with the actual measurements for sample-6

A.2 RBPF Estimates for all the samples

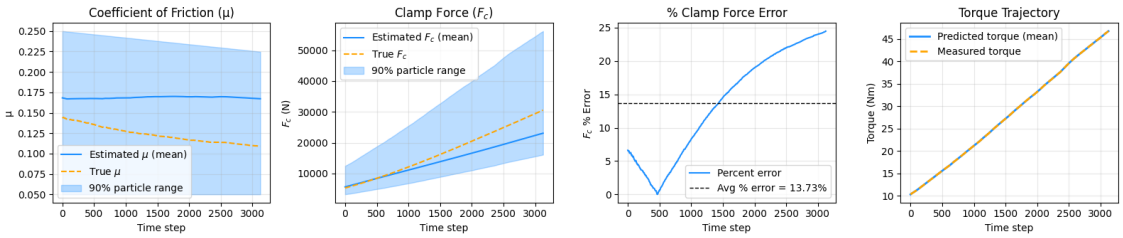


Figure A.11: Sample-1 with +20% error in initial CoF

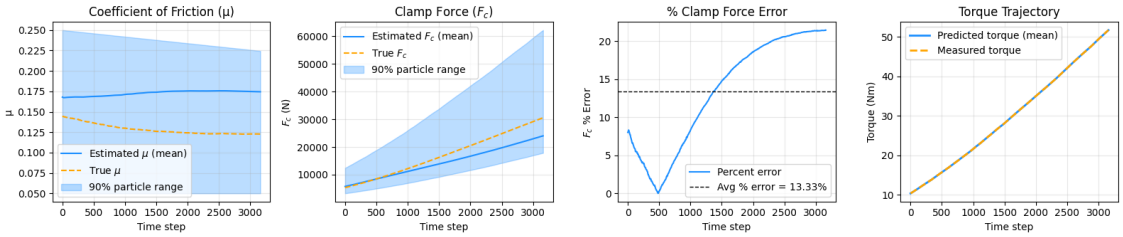


Figure A.12: Sample-2 with +20% error in initial CoF

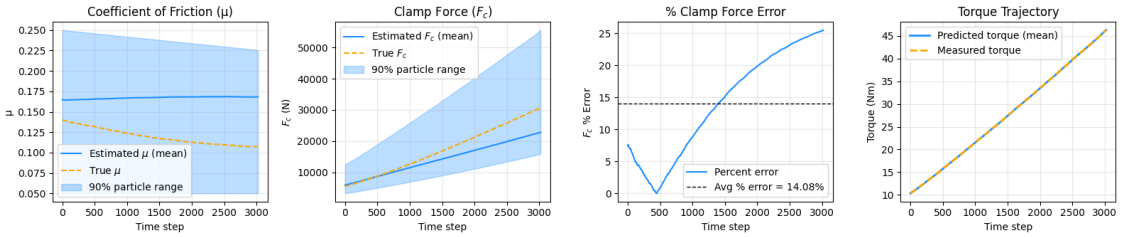


Figure A.13: Sample-3 with +20% error in initial CoF

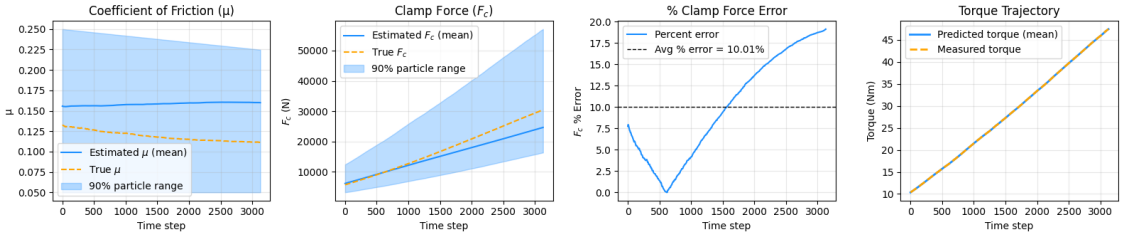


Figure A.14: Sample-4 with +20% error in initial CoF

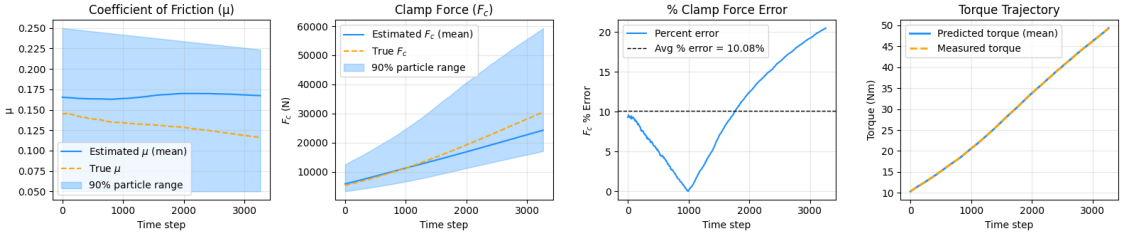


Figure A.15: Sample-6 with +20% error in initial CoF

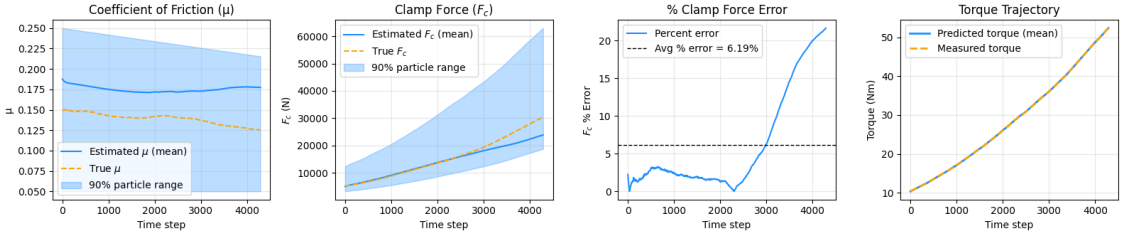


Figure A.16: Sample-5 with +**40**% error in initial CoF

DEPARTMENT OF SOME SUBJECT OR TECHNOLOGY
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY