



Delft University of Technology

Model Predictive Control with Gaussian Processes for Flexible Multi-Modal Physical Human Robot Interaction

Haninger, Kevin; Hegeler, Christian; Peternel, L.

DOI

[10.1109/ICRA46639.2022.9811590](https://doi.org/10.1109/ICRA46639.2022.9811590)

Publication date

2022

Document Version

Final published version

Published in

Proceedings of the 2022 IEEE International Conference on Robotics and Automation (ICRA)

Citation (APA)

Haninger, K., Hegeler, C., & Peternel, L. (2022). Model Predictive Control with Gaussian Processes for Flexible Multi-Modal Physical Human Robot Interaction. In *Proceedings of the 2022 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 6948-6955). IEEE.
<https://doi.org/10.1109/ICRA46639.2022.9811590>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Model Predictive Control with Gaussian Processes for Flexible Multi-Modal Physical Human Robot Interaction

Kevin Haninger, Christian Hegeler, and Luka Peternel

Abstract—Physical human-robot interaction can improve human ergonomics, task efficiency, and the flexibility of automation, but often requires application-specific methods to detect human state and determine robot response. At the same time, many potential human-robot interaction tasks involve discrete modes, such as phases of a task or multiple possible goals, where each mode has a distinct objective and human behavior. In this paper, we propose a novel method for multi-modal physical human-robot interaction that builds a Gaussian process model for human force in each mode of a collaborative task. These models are then used for Bayesian inference of the mode, and to determine robot reactions through model predictive control. This approach enables optimization of robot trajectory based on the belief of human intent, while considering robot impedance and human joint configuration, according to ergonomic- and/or task-related objectives. The proposed method reduces programming time and complexity, requiring only a low number of demonstrations (here, three per mode) and a mode-specific objective function to commission a flexible online human-robot collaboration task. We validate the method with experiments on an admittance-controlled robot, performing a collaborative assembly task with two modes where assistance is provided in full six degrees of freedom. It is shown that the developed algorithm robustly re-plans to changes in intent or robot initial position, achieving online control at 15 Hz.

I. INTRODUCTION

Facilitated by new compliant manipulators, robots are gradually moving from caged cells separated from humans into production environments involving mixed human-robot teams. One of the crucial elements to enable such collaboration is the ability to control a safe and meaningful physical human-robot interaction (HRI) in a dynamic working environment. Such collaboration can improve ergonomics and efficiency but, today, most industrial applications of physical HRI are offline position teaching and workspace sharing with parallel tasks. While safety challenges contribute to this gap, the complexity of implementing flexible control of physical HRI still remain one of the key issues.

Several recent works in physical HRI have made an important step to account for this gap. For example, adaption of a collaborative robot to the human state can significantly improve task-related performance [1], ergonomics [2], and user satisfaction [3]. To do so, the robot should be able to perceive the human state, which presents challenges for both design of sensing [4] and inference methods [5]. On

K. Haninger and C. Hegeler are with the Department of Automation at Fraunhofer IPK, Berlin, Germany. L. Peternel is with the Cognitive Robotics Department, Delft University of Technology, Delft, The Netherlands. Corresponding email: kevin.haninger@ipk.fraunhofer.de

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820689 — SHERLOCK.

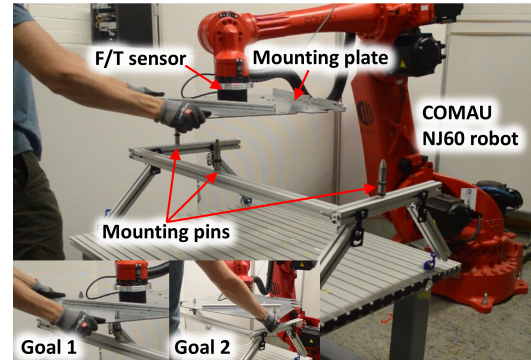


Fig. 1: Experimental setup for human-robot collaborative assembly. The task involved the plate to be mounted on one of two sets of pins on the base frame. Pictures in left bottom corner show two possible modes represented by two different assembly goal locations (i.e., Goal 1 and Goal 2). This emulates tasks where the same part is mounted in different locations, such as tires on a car.

the other hand, some human states or preferences can be more generally modelled, such as the desired motion of the robot/payload [6]–[9], physical fatigue [2], and preferred interaction force range [10].

Nevertheless, a major part of prompt robot adaption to the human is its ability to infer the human intent during the collaborative task execution [2], [11]. While human intent is often considered as a continuous variable [1], [9], it can also reflect discrete changes in the task [12], for example a collection of possible goals. Discrete human state has been considered for virtual fixtures [13], Dynamic Movement Primitives (DMPs) [12], and impedance control [14].

When the robot has the knowledge of human state and intent, it must respond promptly with appropriate actions to facilitate the collaborative task execution. This typically means generation of motion trajectories and often simultaneous impedance adjustments in order to govern a proper interaction behaviour. The trajectories can be chosen to optimize task-related objectives: reducing trajectory jerk [15], [16], minimize positioning error [17]–[19], or render appropriate velocity response [20], [21]. In addition, human operator-related objectives can optimised as well, such as minimizing interaction forces [22], or metabolic cost [23]. A common rule is that the robot can take over aspects of the task that require precision, while the human can handle adaption variations. The ‘minimum intervention principle’ [24] follows this rule, where uncertain degrees of freedom (DOF) should have a lower stiffness [25], [26], which can also be interpreted as a risk-sensitive control. Nevertheless, such an approach is not reasonable for all interactive tasks that involve physical constraints [27].

The problem of designing robot response – including over a belief of human state – can be simplified by using Model Predictive Control (MPC). MPC can optimize a range of variables, such as motion trajectory [28]–[30], impedance parameters [31] and human joint pose, in order to improve a range of objectives that consider ergonomics, task performance, and uncertainty. Advances in optimization solvers, processing power, and toolboxes has increased the use of MPC in many control-engineering domains [32], including robotic locomotion [33], manipulation [34] and rehabilitation [35]. MPC can also accommodate constraints in state or control, allowing some safety or acceptance characteristics to be directly transcribed. However, MPC requires a model, and human models are challenging; a priori human models are limited and collecting human data is expensive. Furthermore, MPC on nonlinear systems has no intrinsic timing or stability guarantees [36, §7.3].

MPC has been applied in non-physical HRI for collision avoidance, treating human motion as a disturbance [28], or with a basic human model [37]. While MPC has also been applied in physical HRI on robots with one or two DOF [29], [30], these methods only planned a motion with safety constraints according to robot dynamics and did not model the human. MPC with human models in physical HRI has been explored for one DOF tasks in [31]. There, the impedance stiffness/damping were adapted with a neural-network based human model and the sampling-based Cross-Entropy Method optimization was employed for MPC.

In this work, we consider physical HRI tasks which are a discrete collection of modes, each with a distinct objective and human dynamics. We model the human force as a function of mode and robot position using Gaussian Processes (GP) regression and then employ them for:

- 1) a mode inference system, which uses the human models to estimate the current mode, possibly including prior probabilities or transitions for the modes.
- 2) a Model Predictive Controller to determine robot trajectory (and other control parameters) according to task- and human-related objectives, evaluated over the belief in mode.

This paper goes beyond the state of the art in physical HRI by allowing a general objective function with a belief over discrete modes, instead of weighting the action associated with each mode by the belief that the mode is currently true, as proposed in [12]–[14]. Compared to the state of the art MPC-based methods for physical HRI [31], this approach considers discrete modes, uses a deterministic optimization-based (not a stochastic sampling-based) MPC, provides 6-DOF assistance, and requires an order of magnitude fewer demonstrations.

While the proposed approach is general and flexible, a major challenge is computational efficiency. We devote special attention to computational techniques and approximations that improve the computational efficiency. The paper is structured in a following manner: we first introduce the models for the robot and human, the inference and MPC problem formulation, then we describe the proposed

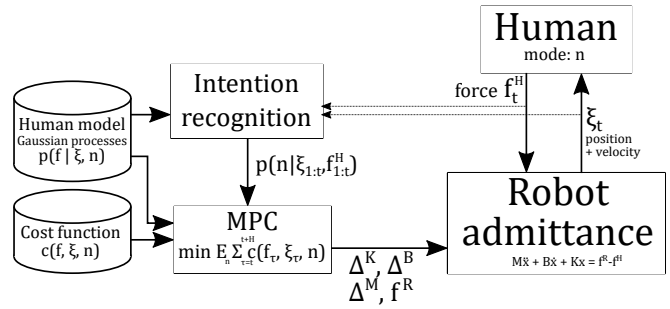


Fig. 2: Framework overview. The intention of the human and mode of operation is inferred by using human model encoded by Gaussian Processes and real-time measurements. The appropriate collaborative behaviour of the robot is optimised and governed by model predictive control based on the identified modes and corresponding cost functions.

approach and implementation. Experimental validation is done on a medium-payload industrial robot, performing a collaborative assembly task with a human (see Fig. 1).

II. HUMAN AND ROBOT MODELS

This section introduces the robot control architecture, robot dynamic model, and human modelling approach. The robot and human dynamics, when coupled, form the stochastic dynamics that is considered in the MPC problem.

A. Robot Admittance Control

As nonlinear MPC does not have intrinsic timing or stability guarantees, the proposed control architecture moves safety concerns, where possible, to a lower-level real-time controller. An admittance controller is used to allow the human or environmental forces to move the robot between MPC updates, which reduces the risk of excessive contact forces. The MPC controller then sets a virtual desired force, and (optionally) changes in the admittance parameters, as seen in Fig. 2.

A Cartesian admittance controller is realized in the robot's tool center point (TCP) frame to approximate the continuous-time dynamics of

$$f^R - f^H = M\ddot{x} + D\dot{x} + K(x - x_0), \quad (1)$$

with pose $x \in \mathbb{R}^6$, velocity $\dot{x} \in \mathbb{R}^6$, acceleration $\ddot{x} \in \mathbb{R}^6$, human force $f^H \in \mathbb{R}^6$ and virtual desired forces f^R . The impedance parameters are encoded by the inertia $M \in \mathbb{R}^{6 \times 6}$, damping $D \in \mathbb{R}^{6 \times 6}$, and stiffness $K \in \mathbb{R}^{6 \times 6}$ matrices, which are all diagonal. The rotational elements of x and f are the angles and torques, respectively, about the three axes of the TCP frame.

To formulate an MPC problem, the dynamics in (1) have to be discretized. Denote sample time T_s and let subscript t denote the value at time $t_0 + tT_s$, where t_0 is the start time at $t = 0$. Taking a first-order Euler discretization where $\xi = [x^T, \dot{x}^T]^T$, we derive

$$\xi_{t+1} = \begin{bmatrix} I & T_s I \\ -T_s M^{-1} K & I - T_s M^{-1} D \end{bmatrix} \xi_t + \begin{bmatrix} 0 \\ T_s M^{-1} \end{bmatrix} (f_t^H - f_t^R), \quad (2)$$

where $\mathbf{I} \in \mathbb{R}^{6 \times 6}$ and $\mathbf{0} \in \mathbb{R}^{6 \times 6}$ are the identity and zero matrix, respectively.

Using an explicit Euler integrator defined in (2) results in oscillation of the discrete state trajectory when $1 - T_s \frac{D_{i,i}}{M_{i,i}} < 0$, where $\cdot_{i,j}$ denotes the i, j -th element of a matrix. This oscillation can be eliminated if $T_s D_{i,i} < M_{i,i}$, which can only be achieved with a sufficiently small time step T_s for typical $M_{i,i}$ and $D_{i,i}$. To allow larger T_s , we rewrite the Euler damping term $\mathbf{I} - T_s \mathbf{M}^{-1} \mathbf{D}$ using the closed-form integral as $\exp(-T_s \mathbf{M}^{-1} \mathbf{D})$.

B. Human Modelling

1) *Human Forces*: The human is modelled stochastically as an impedance, where applied force depends on the pose and mode. For example, in an assembly task the modes represent different assembly goal locations where the part can be moved to (see Fig. 1). The models are independent for each mode n and are defined as

$$\mathbf{f}^H \sim p(\mathbf{f}^H | \mathbf{x}^*, n). \quad (3)$$

This relationship is realized with Gaussian processes (GP) [38], where the human force is regressed over the position/orientation of the robot, where each mode has its own GP model. Each element of the human force is regressed in parallel, where the force data is treated independently.

The pose \mathbf{x}^* reflects a change in orientation representation from \mathbf{x} . Rotation vectors are used to represent orientation for regression, where for a rotation vector $\mathbf{v} \in \mathbb{R}^3$, the direction describes the axis of rotation, and the magnitude the angle of rotation $\|\mathbf{v}\| = \theta$. This was found to have benefits over Euler angles and quaternions, which had discontinuities and were over-parameterized, respectively. The $*$ is dropped in the sequel, but all regression is done over rotation vectors, and admittance calculations over the TCP XYZ axis angles.

2) *Human Kinematics*: As some ergonomic costs are more naturally expressed in terms of human joint torques, we also consider a 4-DOF kinematic model of the human arm with three rotational DOF at the shoulder and one rotational joint for the elbow extension. This model has parameters of l_1 and l_2 , for the length of the upper and lower arm, and $\mathbf{x}^{sh} \in \mathbb{R}^3$ for the human shoulder position. Using these parameters and the human joint angles $\mathbf{q} \in \mathbb{R}^4$, we can calculate the human hand position $\mathbf{x}^H \in \mathbb{R}^3$ with forward kinematics as

$$\mathbf{x}^H = \text{FK}(\mathbf{q}, \mathbf{x}^{sh}), \quad (4)$$

Forces measured at the end-effector can be translated into joint torques $\boldsymbol{\tau}$ as

$$\boldsymbol{\tau} = \mathbf{J}_H(\mathbf{q})^T \mathbf{f}_l^H, \quad (5)$$

where \mathbf{f}_l^H are only the linear forces, and $\mathbf{J}_H(\mathbf{q}) \in \mathbb{R}^{3 \times 4}$ is the standard Jacobian matrix.

C. Stochastic dynamics

As the human model is stochastic and affects the robot trajectory, the future state trajectory will also be stochastic. Each mode induces a distribution on the future state, and

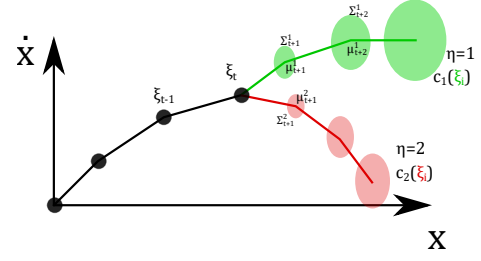


Fig. 3: Multi-modal future trajectory, where each possible human mode induces a different course of the trajectory, and each possible trajectory also has uncertainty.

when the mode is uncertain, the future trajectory will be a multi-modal distribution such as seen in Fig. 3.

Denoting the dynamics of (2) with state-space representation as $\xi_{t+1} = \mathbf{A}_t \xi_t + \mathbf{B}_t(\mathbf{f}_t^H - \mathbf{f}_t^R)$, the future trajectory in mode n is distributed as $\xi_t \sim \mathcal{N}(\mu_t^n, \Sigma_t^n)$, and the mode dynamics are defined by

$$\mu_t^{H,n}, \Sigma_t^{H,n} = \text{GP}_n(\mu_t^n, \Sigma_t^n), \quad (6)$$

$$\mu_{t+1}^n = \mathbf{A}_t \mu_t^n + \mathbf{B}_t(\mu_t^{H,n} - \mathbf{f}_t^R), \quad (7)$$

$$\Sigma_{t+1}^n = \mathbf{A}_t \Sigma_t^n \mathbf{A}_t^T + \mathbf{B}_t \Sigma_t^{H,n} \mathbf{B}_t^T, \quad (8)$$

where $\mu^{H,n}$ and $\Sigma^{H,n}$ are the mean and covariance of the n^{th} GP model.

When using (7) and (8) for prediction starting at time t , they are initialized with $\Sigma_t^n = \mathbf{0}$, and $\mu_t^n = \xi_t$, as it is assumed that the current state of the robot is measured.

III. INFERENCE AND MPC

The modelling method presented in Sec. II is then applied to the inference and planning problems. The proposed solutions to these two problems are described in the two subsections below.

A. Inference

Given a series of observations $\xi_{1:t} = [\xi_1, \dots, \xi_t]$ and $\mathbf{f}_{1:t}^H$, the human mode can be estimated using the assumed human dynamics from (3). We examine a case where the mode has a fixed prior distribution, and a case where the distribution of mode transitions are known.

1) *Mode without transition*: If the mode is fixed during the task (e.g. one of several goals is active), but has an initial distribution of $p(n)$, the posterior is defined as

$$\begin{aligned} p(n | \xi_{1:t}, \mathbf{f}_{1:t}^H) &= \frac{p(n) \prod_{i=1}^t p(\mathbf{f}_i^H | \xi_i, n)}{\sum_{m=1}^N p(m) \prod_{i=1}^t p(\mathbf{f}_i^H | \xi_i, m)} \\ &\propto p(\mathbf{f}_t^H | \xi_t, n) p(n | \xi_{1:t-1}, \mathbf{f}_{1:t-1}^H), \end{aligned} \quad (9)$$

which can be recursively calculated. We apply a floor to the belief before normalization, i.e. $p(n | \xi_{1:t}, \mathbf{f}_{1:t}^H) = \max(\underline{b}, p(n | \xi_{1:t}, \mathbf{f}_{1:t}^H))$, with typical values of $\underline{b} = 1e - 6$.

2) *Mode with transitions*: If the distribution of mode transitions are known, $p(n_{t+1} | n_t, \xi_t)$, the posterior can also

be calculated as

$$\begin{aligned}
p(n_{t+1}|\xi_{1:t}, \mathbf{f}_{1:t}^H) \\
\propto \sum_{n_{1:t}} p(n_1) \prod_{i=1}^t p(\mathbf{f}_i^H|\xi_i, n_t) p(n_{t+1}|n_t, \xi_t) \\
\propto \sum_{n_t} p(\mathbf{f}_t^H|\xi_t, n_t) p(n_{t+1}|n_t) p(n_t|\xi_{1:t-1}, \mathbf{f}_{1:t-1}^H).
\end{aligned}$$

This also allows an efficient recursive calculation of belief in a specific human mode.

3) *Transforming force to direction*: The GP model for force independently models each element in \mathbf{f} , $p(\mathbf{f}^h|\xi, n) = \prod_{i=1}^6 p(f_i^h|\xi, n)$, where f_i^h is the i^{th} element of the force vector. The direct approach described in (9) can be noisy, e.g., if one element has a very low probability this can render the total likelihood low. To account for that, we developed an ad-hoc similarity measure, where the pseudo-likelihood is calculated as

$$\begin{aligned}
s(\mathbf{f}^H, \boldsymbol{\mu}^H, \boldsymbol{\Sigma}^H) &= \beta \|\boldsymbol{\mu}^H\|_2 \log\left(\frac{1}{2}(\mathbf{f}^H)^T \boldsymbol{\mu}^H + \frac{1}{2}\right) \\
&\quad - \sum_{i=1}^6 \log(\boldsymbol{\Sigma}[i, i]), \quad (10)
\end{aligned}$$

where β weights the relative priority given to the direction matching (typical value, 0.05). This ad-hoc rule updates more strongly when forces are larger, and the value is higher when the model and measured force point in a similar direction, and smaller when the uncertainty in the GP model is larger. The developed ad-hoc similarity measure was found to result in a smoother mode inference for co-manipulation, when it is used to replace $p(\mathbf{f}^H|\xi, n)$ in the belief update (9).

B. MPC Problem

MPC iteratively solves an optimal control problem, choosing decision variables that (locally) minimize a cost function, while imposing consistency of the model dynamics (2). The MPC implemented algorithm uses a multiple-shooting transcription with a problem statement of

$$\begin{aligned}
\mathbf{u}_{t:t+H} &= \arg \min_{\mathbf{u}} J(b_t, \xi_t) \\
\text{s.t. } \forall n \in [1, \dots, N], \tau \in [t, \dots, t+H-1] : \\
&\quad \boldsymbol{\mu}_\tau^n = \xi_\tau, \boldsymbol{\Sigma}_\tau^n = \mathbf{0} \\
&\quad |\boldsymbol{\mu}_{\tau+1}^n - f(\boldsymbol{\mu}_\tau^n, \mathbf{u}_\tau)| \leq \varrho \\
&\quad |\boldsymbol{\Sigma}_{\tau+1}^n - g(\boldsymbol{\mu}_\tau^n, \boldsymbol{\Sigma}_\tau^n, \mathbf{u}_\tau)| \leq \varrho \\
&\quad \mathbf{u} \in U
\end{aligned} \quad (11)$$

where H is the planning horizon, U is the range of allowed inputs, ϱ the slack for the continuity constraints (the inequality is applied element-wise), $f(\mu_t, u_t)$ derives μ_{t+1} following (7), and g following (8).

The constraints in (11) are nonlinear, so an interior-point nonlinear optimization solver is used (details in §IV). While nonlinear, the GP models can be written to be automatically differentiated, allowing calculation of the gradient and Hessian of the objective and constraints, significantly improving the speed and stability of the optimization.

The MPC framework here allows for different choices of \mathbf{u} and \mathbf{J} . Recall \mathbf{x}^{sh} is the human shoulder position, \mathbf{q} are

the human joint angles, $\boldsymbol{\tau}$ are the human joint torques, \mathbf{f}^H is the human force, n is the mode, $\dot{\mathbf{x}}$ is the robot velocity, \mathbf{f}^R is the robot desired force.

1) *Decision variables \mathbf{u}* : The decision variable \mathbf{u} can include any of the following:

| | |
|---------------------------------------|--------------------------------|
| $\mathbf{f}_{t:t+H}^R$ | Robot trajectory |
| Δ_t^M, Δ_t^B | Change in robot impedance |
| $\mathbf{x}^{sh}, \mathbf{q}_{t:t+H}$ | Human shoulder and joint traj. |

If \mathbf{x}^{sh} and $\mathbf{q}_{t:t+H}$ are included as decision variables, an additional constraint of $\mathbf{x}^H = T(\boldsymbol{\mu}_t^n)$ is added to the MPC problem, where \mathbf{x}^H is from (4) and T represents the hand grasp location relative to robot TCP.

2) *Stage cost function c_n* : The general stage cost function is defined as:

$$\begin{aligned}
c_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\mu}_H, \boldsymbol{\Sigma}_H, \boldsymbol{\tau}, \mathbf{u}) &= \\
&\quad \boldsymbol{\mu}^T \mathbf{Q}_\mu \boldsymbol{\mu} + \text{tr}(\mathbf{Q}_\Sigma \boldsymbol{\Sigma}) + \\
&\quad \boldsymbol{\mu}_H^T \mathbf{Q}_H \boldsymbol{\mu}_H + \text{tr}(\mathbf{Q}_{\Sigma, H} \boldsymbol{\Sigma}_H) + \\
&\quad \boldsymbol{\tau}^T \mathbf{Q}_J \boldsymbol{\tau} + \mathbf{u}^T \mathbf{Q}_u \mathbf{u}, \quad (12)
\end{aligned}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are mean and covariance for the predicted state ξ in mode n , $\boldsymbol{\mu}_H$ and $\boldsymbol{\Sigma}_H$ are the mean and covariance of predicted human forces from (6), $\boldsymbol{\tau}$ are the human joint torques from (5), \mathbf{u} are control inputs, and \mathbf{Q}_\cdot are the related weight-matrices for each cost factor.

In some applications, replacing $\boldsymbol{\mu}_H^T \mathbf{Q}_H \boldsymbol{\mu}_H$ with $(\boldsymbol{\mu}_H + \mathbf{f}_t^R)^T \mathbf{Q}_H (\boldsymbol{\mu}_H + \mathbf{f}_t^R)$ can be more robust – the mean human force goes to zero outside the training data (in zero-mean GPs, as used here), and a large penalty on only $\boldsymbol{\mu}^H$ can lead to trajectories which seek to leave the training data.

3) *Total objective function $J(b, \xi)$* : The total objective sums the stage costs and considers the current belief b_t . Two varieties were used: the simple expectation over belief J_E , and risk-sensitive cost J_R , which is adapted from [39], where α adjusts the sensitivity to risk. The two cost functions are defined as

$$J_E = \sum_t \sum_{n=1}^N b_t[n] c_n(\xi_t^n, \mathbf{f}_t^R, \mathbf{f}_t^{H, n}), \quad (13)$$

$$J_R = -2\alpha^{-1} \ln \mathbb{E}_{n \sim b_t} e^{(\sum_t^{t+H} - \frac{1}{2} \alpha c_n(\xi_t^n, \mathbf{f}_t^R, \mathbf{f}_t^{H, n}))}. \quad (14)$$

IV. IMPLEMENTATION

This section presents the overall integration and details on the implementation of the proposed method. The code, parameter settings, training data, and experimental results are available at <https://gitlab.cc-asp.fraunhofer.de/hanikevi/gp-mpc-impedance-control>.

A. Data collection and commissioning

An overview of the commissioning and online execution is shown in Algorithm 1. To initialize the GPs, data must be collected for each mode. For co-manipulation tasks, this is done here by having the robot in a passive admittance control mode, where it acts as a mass-damper system. The default

damping gains used are $\mathbf{M} = \text{diag}([12, 12, 12, 1, 1, 1])$ where diag is the diagonalization operation, and $\mathbf{D} = \text{diag}([1100, 1100, 1100, 200, 200, 200])$. All units are in SI standard units, and radians are used.

Initialize GPs

```

for  $n = 1 \dots N$  do
    Collect data  $\mathcal{D}_n \leftarrow \{\{\xi_1, f_1\}, \dots\}$  with
    mass/damper admittance
    (optional)  $\phi = \text{fit\_hyperparams}(\mathcal{D}_n, \phi)$ 
    (optional)  $\mathcal{D}_n = \text{sparsify\_GP}(\mathcal{D}_n, \phi)$ 
     $\text{GP}_n \leftarrow \text{build\_GP}(\mathcal{D}_n, \phi)$ 
end

```

Online Execution

```

Belief Thread (50 Hz)
     $\xi, f \leftarrow \text{read\_state}()$ 
     $b \leftarrow \text{belief\_update}(b, \xi, f, \text{GP})$  by (9)
Control Thread(~5-20 Hz)
     $u \leftarrow \text{MPC\_solve}(\xi, b, \text{GP})$  by (11)

```

Algorithm 1: Data collection and execution, with GP hyperparameters ϕ

B. Software Implementation

To enable stable and smooth performance of the designed controller, an MPC loop speed > 5 Hz is targeted. The MPC problem is solved using CasADi [40] to interface with IPOPT [41] using the HSL MA57 linear solver [42]. Solver parameters used for IPOPT are detailed in the above-referenced cloud, and warm start is used for all subsequent solves. The horizon and shooting nodes were set to $H = 6$ with a sampling time $T_s = 0.10$ seconds. The continuity constraint slack was set at $\varrho = 10^{-5}$ meter, which is still within the limits of the robots precision. The maximum force applied by the robot \mathbf{f}^R was constrained to be less than 20 N and 6 Nm for linear and rotational force. The state space was limited with regard to the maximum velocity, but the position of the robot was unconstrained within the MPC calculations.

The GP regression we used builds on [36] and is implemented in CasADi. This allows the automatic calculation of the gradient and Hessian for the interior point solver. A squared exponential GP kernel was used with 35-50 training samples, sub-sampled from three the demonstration trajectories, which started at different positions, as seen in the left of Fig. 4. An example of the forces fitted over position for two modes can be seen in Fig. 4, right. Since GP model evaluations come with a high computational cost, the number of training samples necessary was reduced using Sparse GPs.

C. Sparse GPs

As every MPC loop involves at least HN GP evaluations for each optimizer step and the asymptotic complexity of a GP evaluation is $O(S^3)$, where S is the number of training samples, the calculation time is highly sensitive to the number of training points. Since the collected data set contains thousands of samples, a method of sub sampling is needed. To address these two issues, sparse GPs are used,

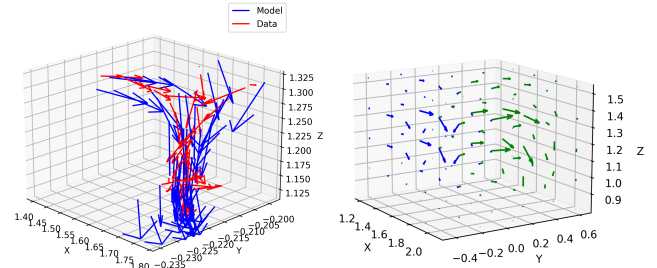


Fig. 4: Linear force over Cartesian position: (Left) Training data and mean of the GP model evaluated near the training data. The contact forces from the pins can be seen as red arrows to the left/right, and are smoothed out by the GP model. (Right) mean force of two GP models with blue and green vectors.

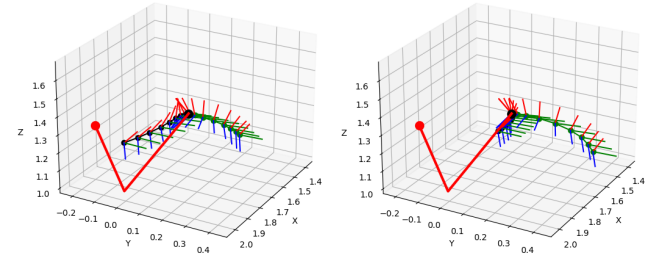


Fig. 5: View of MPC planned trajectory for each mode (blue and green), with human arm model in red and current TCP location/orientation as black point. On the left, when the belief is $b_t = [0.5, 0.5]$ for the left and right goal, and on the right is $b_t = [0.05, 0.95]$.

which can be calculated in $O(RS^2)$, where R is the number of inducing variables that replace the original training data. Within sparse GPs the inducing variables are generated by minimizing the Kullback-Leibler divergence between a GP trained on a larger subset of the original data and an GP using artificial training data as described in [43]. In our experiments this method effectively reduced the amount of training samples – 35 samples had equivalent log likelihood to 50 samples when subsampled by time.

V. EXPERIMENTAL VALIDATION

A. Hardware Validation

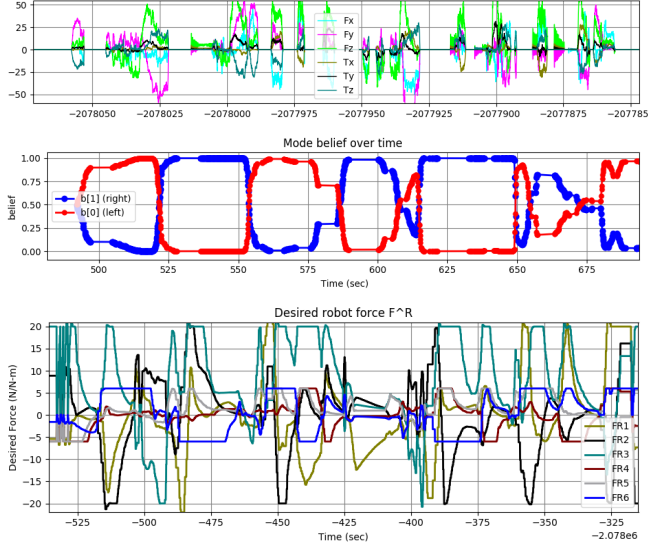
The approach is validated on a large industrial robot, where a 16 Kg steel plate had to be manipulated and/or have objects mounted to it, as seen in Fig. 1. This plate must be mounted on a set of two pins (insets of Fig. 1) which have a loose running clearance fit (hole is oversized by 0.5 mm), and there are two possible mounting positions for the plate. Here, we used $\mathbf{Q}_\mu = \text{diag}([0, \dots, 0, 0.1, \dots, 0.1])$ to penalize only velocity, $\mathbf{Q}_\Sigma = 0.1\mathbf{I}$, $\mathbf{Q}_H = 0.1\mathbf{I}$, $\mathbf{Q}_{\Sigma, H} = 270\mathbf{I} - \mathbf{Q}_u = 0.25\mathbf{I}$, and $\mathbf{Q}_{\Sigma, H}$ was found to be the most critical.

Two MPC planned trajectories can be seen in Fig. 5, where only the belief is changed. With an even belief $b_t = [0.5, 0.5]$ (left), the planned forces \mathbf{f}^R are smaller, whereas when the right belief is stronger (right), the MPC plans to move more in this direction.

The attached video shows the performance on physical hardware, and the resulting time plots can be seen in Fig. 6. The system is able to robustly detect changes in goal online, and provides appropriate active assistance. The system is also

TABLE I: Comparison of MPC problem statements and computational cost

| Full GP Cov. | State Cov. | # GP Points | J | Imp Params | $x^{sh}, q_{t:t+H}$ | Cold (sec) | Avg Warm | Worst Warm |
|--------------|------------|-------------|-----------|------------|---------------------|------------|----------|------------|
| Yes | Yes | 50 | Expected | No | No | 11.524 | 0.157 | 1.185 |
| No | Yes | 50 | Expected | No | No | 5.023 | 0.062 | 0.215 |
| No | No | 50 | Expected | No | No | 4.571 | 0.053 | 0.176 |
| No | Yes | 35 | Expected | No | No | 5.239 | 0.066 | 0.225 |
| No | Yes | 50 | Expected | Yes | No | 17.910 | 0.342 | 2.354 |
| No | Yes | 50 | Expected | No | Yes | 6.483 | 0.137 | 1.977 |
| No | Yes | 50 | Risk Sens | No | No | 44.223 | 0.330 | 2.399 |

Fig. 6: Time traces of applied forces (top), belief (middle) and desired force trajectory F^R (bottom).

robust to a variety of starting positions, and can recover when the human perturbs the robot away from the goal, providing the flexibility of online motion planning. While the final accuracy is limited (± 2 cm is typical), this fine positioning can be taken over by the human. Note that the goal position is not hard-coded, but is implicit from the demonstrations (human forces are smaller near the goal and model covariance is lower because more data is collected there).

B. Computational Cost

To compare the relative computational cost of several aspects of the approach, the MPC loop is solved over recorded experiment data, allowing repeatable comparison while covering the state space. A computer with Intel i3-5010 @ 1.70 Ghz was used for all validation, with $H = 5$ and $dt = 0.05$. If the ‘Full GP Cov.’ option is false, the GP covariance is simplified, replacing the $\text{Tr}(\mathbf{Q}_{H,\Sigma}\Sigma_H)$ term in (12) with $\text{Tr}(\mathbf{Q}_{H,\Sigma})\Sigma_{H,1,1}$, using just the first element (as the covariance kernel is shared between linear directions, and similar between linear and rotational). The ‘State Cov.’ option defines whether or not the state covariance in (8) is calculated. The cold start time (the first MPC solve, which includes building the MPC solver), average warm start time and worst-case warm start are compared.

It can be seen in Tab. I that simplifying the full GP covariance makes a large computational difference – this

simplification does not reduce performance when similar hyperparameters are used between the DOF, as the covariance will be similar between DOF. Reducing the GP size and not calculating state covariance make minor improvements in average calculation time.

Adding impedance parameters to the problem significantly increases both mean and worst-case calculation time. Similarly, adding the human joints/shoulder to the problem significantly increases the worst-case calculation. Using a risk-sensitive objective results in severe computational costs. These three problem statements are not currently feasible for online control in the current approach.

C. Co-optimization of human joints and impedance trajectory

The attached video shows the co-optimization of the human joint trajectory, where the co-optimized trajectory for human joints finds a solution that reduces the moment arm between point of force application and shoulder, thereby reducing the shoulder torques required. However, transcribing all of the ergonomic constraints and costs remains a topic for future work. Similar video results show the co-optimization of impedance parameters, although the current cost functions and test applications typically result in simply minimizing the impedance parameters.

VI. CONCLUSION

The feasibility of MPC for direct physical HRI in 6 DOF has been shown, with both realistic data-collection requirements (3 demonstrations) and online computation speed (15 Hz). We note that while the approach considers a stochastic human model and dynamics, the largest functional advantage of this approach is in the human model, in keeping the robot near the training data. Considering uncertainty in the trajectory has not yet been shown to offer new functionality, although new use-cases and objectives may establish benefits.

The proposed approach also has known limitations, many of which are shared with other approaches, but are listed for completeness. The approach does not distinguish between environmental forces and human forces, requiring care in contact tasks. GP covariance does not capture heteroscedastic properties (i.e. if there is larger variance in a region of the state space) – modelling changes in underlying covariance requires splitting the GP, which is then computationally expensive. Fitting GP hyperparameters with log likelihood on limited data can be problematic, to address this we limited the range of hyperparameters or used fixed ones.

REFERENCES

- [1] A. Takagi, G. Ganesh, T. Yoshioka, M. Kawato, and E. Burdet, "Physically interacting individuals estimate the partner's goal to enhance their movements," *Nature Human Behaviour*, vol. 1, no. 3, p. 0054, 2017.
- [2] L. Peternel, N. Tsagarakis, D. Caldwell, and A. Ajoudani, "Robot adaptation to human physical fatigue in human-robot co-manipulation," *Autonomous Robots*, vol. 42, no. 5, pp. 1011–1021, 2018.
- [3] B. Corteveille, E. Aertbeliën, H. Bruyninckx, J. De Schutter, and H. Van Brussel, "Human-inspired robot assistant for fast point-to-point movements," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*. IEEE, 2007, pp. 3639–3644.
- [4] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, "Progress and prospects of the human-robot collaboration," *Autonomous Robots*, vol. 42, no. 5, pp. 957–975, 2018.
- [5] S. Jain and B. Argall, "Recursive Bayesian human intent recognition in shared-control robotics," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 3905–3912.
- [6] Y. Demiris, "Prediction of intent in robotics and multi-agent systems," *Cognitive Processing*, vol. 8, no. 3, pp. 151–158, 2007.
- [7] Y. Li and S. S. Ge, "Human-Robot Collaboration Based on Motion Intention Estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007–1014, 2014.
- [8] C. Wang, Y. Li, S. S. Ge, and T. H. Lee, "Reference adaptation for robots in physical interactions with unknown environments," *IEEE transactions on cybernetics*, vol. 47, no. 11, pp. 3504–3515, 2017.
- [9] G. Kang, H. S. Oh, J. K. Seo, U. Kim, and H. R. Choi, "Variable Admittance Control of Robot Manipulators Based on Human Intention," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 3, pp. 1023–1032, 2019.
- [10] S. Gopinathan, S. Otting, and J. Steil, "A user study on personalized adaptive stiffness control modes for human-robot interaction," in *The 26th IEEE International Symposium on Robot and Human Interactive Communication*, 2017, pp. 831–837.
- [11] H. Ben Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," IEEE, 2014, pp. 2831–2837.
- [12] M. Khoramshahi and A. Billard, "A dynamical system approach to task-adaptation in physical human-robot interaction," *Autonomous Robots*, vol. 43, no. 4, pp. 927–946, 2019.
- [13] G. Raiola, S. S. Restrepo, P. Chevalier, P. Rodriguez-Ayerbe, X. Lamy, S. Tiba, and F. Stulp, "Co-manipulation with a library of virtual guiding fixtures," *Autonomous Robots*, vol. 42, no. 5, pp. 1037–1051, 2018.
- [14] L. Roza, S. Calinon, D. G. Caldwell, P. Jimenez, and C. Torras, "Learning physical collaborative robot behaviors from human demonstrations," *IEEE Transactions on Robotics*, vol. 32, no. 3, pp. 513–527, 2016.
- [15] Y. Maeda, T. Hara, and T. Arai, "Human-robot cooperative manipulation with motion estimation," in *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the Next Millennium (Cat. No. 01CH37180)*, vol. 4. IEEE, 2001, pp. 2240–2245.
- [16] F. Dimeas and N. Aspragathos, "Reinforcement learning of variable admittance control for human-robot co-manipulation," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference On*. IEEE, 2015, pp. 1011–1016.
- [17] C. Yang, G. Ganesh, S. Haddadin, S. Parusel, A. Albu-Schaeffer, and E. Burdet, "Human-like adaptation of force and impedance in stable and unstable interactions," *IEEE transactions on robotics*, vol. 27, no. 5, pp. 918–930, 2011.
- [18] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *Robotics and Automation (ICRA), 2011 IEEE International Conference On*. IEEE, pp. 4326–4332.
- [19] H. Lee, B. Lee, W. Kim, M. Gil, J. Han, and C. Han, "Human-robot cooperative control based on pHRI (Physical Human-Robot Interaction) of exoskeleton robot for a human upper extremity," *International Journal of Precision Engineering and Manufacturing*, vol. 13, no. 6, pp. 985–992, 2012.
- [20] T. Tsumugiwa, R. Yokogawa, and K. Hara, "Variable impedance control based on estimation of human arm stiffness for human-robot cooperative calligraphic task," in *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference On*, vol. 1. IEEE, 2002, pp. 644–650.
- [21] V. Duchaine, B. M. St-Onge, D. Gao, and C. Gosselin, "Stable and intuitive control of an intelligent assist device," *IEEE transactions on haptics*, vol. 5, no. 2, pp. 148–159, 2012.
- [22] X. Lamy, F. Colledani, F. Geffard, Y. Measson, and G. Morel, "Achieving efficient and stable comanipulation through adaptation to changes in human arm impedance," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference On*. IEEE, 2009, pp. 265–271.
- [23] J. R. Koller, D. H. Gates, D. P. Ferris, and C. D. Remy, "Body-in-the-Loop Optimization of Assistive Robotic Devices: A Validation Study," in *Robotics: Science and Systems*, 2016.
- [24] J. R. Medina, D. Lee, and S. Hirche, "Risk-sensitive optimal feedback control for haptic assistance," in *Robotics and Automation (ICRA), 2012 IEEE International Conference On*. IEEE, 2012, pp. 1025–1031.
- [25] S. Calinon, D. Bruno, M. S. Malekzadeh, T. Nanayakkara, and D. G. Caldwell, "Human-robot skills transfer interfaces for a flexible surgical robot," *Computer methods and programs in biomedicine*, vol. 116, no. 2, pp. 81–96, 2014.
- [26] E. Pignat and S. Calinon, "Learning adaptive dressing assistance from human demonstration," *Robotics and Autonomous Systems*, vol. 93, pp. 61–75, 2017.
- [27] L. Peternel, T. Petrič, and J. Babič, "Robotic assembly solution by human-in-the-loop teaching method based on real-time stiffness modulation," *Autonomous Robots*, vol. 42, no. 1, pp. 1–17, 2018.
- [28] P. Poignet and M. Gautier, "Nonlinear model predictive control of a robot manipulator," in *6th International workshop on advanced motion control. Proceedings (Cat. No. 00TH8494)*. IEEE, 2000, pp. 401–406.
- [29] S. Norouzzadeh, T. Lorenz, and S. Hirche, "Towards safe physical human-robot interaction: an online optimal control scheme," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2012, pp. 503–508.
- [30] S. M. Rahman and R. Ikeura, "Mpc to optimise performance in power-assisted manipulation of industrial objects," *IET Electric Power Applications*, vol. 11, no. 7, pp. 1235–1244, 2017.
- [31] L. Roveda, J. Maskani, P. Franceschi, A. Abdi, F. Braghin, L. M. Tosatti, and N. Pedrocchi, "Model-based reinforcement learning variable impedance control for human-robot collaboration," *Journal of Intelligent & Robotic Systems*, vol. 100, no. 2, pp. 417–433, 2020.
- [32] J. H. Lee, "Model predictive control: Review of the three decades of development," *International Journal of Control, Automation and Systems*, vol. 9, no. 3, pp. 415–424, 2011.
- [33] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, T. Koolen, P. Marion, and R. Tedrake, "Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," *Autonomous robots*, vol. 40, no. 3, pp. 429–455, 2016.
- [34] M. Rubagotti, T. Taunyazov, B. Omarali, and A. Shintemirov, "Semi-autonomous robot teleoperation with obstacle avoidance via model predictive control," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2746–2753, 2019.
- [35] T. Teramae, T. Noda, and J. Morimoto, "Emg-based model predictive control for physical human-robot interaction: Application for assist-as-needed control," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 210–217, 2017.
- [36] H.-A. Langäker, "Cautious MPC-based control with Machine Learning," *NTNU*, 2018.
- [37] A. Oleinikov, S. Kuskavlev, A. Shintemirov, and M. Rubagotti, "Safety-aware nonlinear model predictive control for physical human-robot interaction," *IEEE Robotics and Automation Letters*, 2021.
- [38] C. E. Rasmussen, "Gaussian processes for machine learning." MIT Press, 2006.
- [39] J. R. Medina, T. Lorenz, D. Lee, and S. Hirche, "Disagreement-aware physical assistance through risk-sensitive optimal feedback control," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 3639–3645.
- [40] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.
- [41] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.

- [42] “Hsl. a collection of fortran codes for large scale scientific computation.” 2021. [Online]. Available: <http://www.hsl.rl.ac.uk/>
- [43] M. Titsias, “Variational learning of inducing variables in sparse gaussian processes,” in *Artificial intelligence and statistics*. PMLR, 2009, pp. 567–574.