2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Robust Stability of Gaussian Process Based Moving Horizon Estimation

Tobias M. Wolff, Victor G. Lopez, and Matthias A. Müller

Abstract— **In this paper, we introduce a Gaussian process based moving horizon estimation (MHE) framework. The scheme is based on offline collected data and offline hyperparameter optimization. In particular, compared to standard MHE schemes, we replace the mathematical model of the system by the posterior mean of the Gaussian process. To account for the uncertainty of the learned model, we exploit the posterior variance of the learned Gaussian process in the weighting matrices of the cost function of the proposed MHE scheme. We prove practical robust exponential stability of the resulting estimator using a recently proposed Lyapunov-based proof technique. Finally, the performance of the Gaussian process based MHE scheme is illustrated via a nonlinear system.**

I. INTRODUCTION

Moving horizon estimation (MHE) [1], [2], [3] is a nonlinear, optimization-based state estimation technique. Loosely speaking, at each time instant, we first measure the current output of the system. Then, we solve an optimization problem to determine an optimal estimated state sequence over some (finite) estimation horizon. Inherent physical constraints of the system, such as, e.g., nonnegativity constraints of chemical concentrations or hormone concentrations can be accounted for in the optimization problem. Finally, the state estimate is set to the last element of the optimal estimated state sequence. MHE is particularly suitable for nonlinear state estimation, as it can outperform other nonlinear state estimation techniques such as the extended Kalman filter [1].

However, MHE crucially relies on the knowledge of an accurate mathematical model of the dynamical system. The derivation of such a mathematical model from first principles can be difficult, time-consuming, and expensive. Alternatively, an MHE scheme can be set up by solely relying on data, or by learning the system dynamics using some machine learning technique. In this work, we focus on the latter approach, namely by learning a mathematical model of the system dynamics using Gaussian Processes (GPs) [4].

GPs are a Bayesian machine learning technique which are defined as a collection of random variables, any finite number of which follows a joint Gaussian distribution [4, Def. 2.1]. In recent years, GPs have been increasingly used in the area of learning-based control (compare, e.g., [5], [6]). Here, an advantage is that they inherently allow for a quantification of the model uncertainty, which is typically not the case when using other machine learning techniques such as, e.g., neural networks.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 948679).

Tobias M. Wolff, Victor G. Lopez, and Matthias A. Müller are with the Leibniz University Hannover, Institute of Automatic Control, Germany {wolff,lopez,mueller}@irt.uni-hannover.de

Concerning the design of learning-based estimators, there exist only few results in the literature, both in the context of MHE and regarding the usage of GP based techniques. In [7], a so-called state estimation function is learned by means of a feedforward neural network. The authors in [8] develop a learning-based MHE scheme, where the mapping of the input data (including the system matrices and the measured outputs) to the state estimates is learned offline. Moreover, GPs have been exploited in the design of an extended Kalman Filter [9] and to develop a joint dynamics and state estimation framework [10]. Interestingly, in the robotics community (which uses the term *sliding window* filtering instead of MHE [11, Chapter 4.3.4]), the combination of MHE and GPs has been suggested in various works, compare, e.g., [12], [13]. However, no stability analysis is offered in these works.

The contribution of this work is the introduction of a novel GP based nonlinear MHE framework. We exploit the posterior mean of the GP to approximate the system dynamics and the posterior variance in the design of the weighting matrices of the MHE. The advantages of this approach are two-fold. First, we do not require that any mathematical model of the system dynamics is available a priori. Second, we directly account for the uncertainty of the learned model by using weighting matrices that depend on the regression[1] inputs. Furthermore, we prove practical robust exponential stability of the resulting estimator. In a numerical example, we illustrate the performance of the proposed GP based MHE scheme. In contrast to [13], where the output map is assumed to be known and where the state estimates are directly approximated by a GP, we here learn the full state space model by means of GPs. Moreover, we rigorously prove practical robust stability of the proposed scheme, which has not been done in [13].

II. PRELIMINARIES AND SETTING

We denote the set of integers greater than or equal to $a \in \mathbb{R}$ by $\mathbb{I}_{\geq a}$. The set of non-negative real numbers is denoted by $\mathbb{R}_{\geq 0}$. The weighted vector norm for a vector $x = [x_1 \dots x_n]^\top \in \mathbb{R}^n$ and a symmetric positive definite matrix P is written as $\|x\|_P = \sqrt{x^\top P x}$. The identity matrix of dimension n is denoted by I_n. A diagonal matrix of dimension n with $q_1, \dots, q_n$ on the diagonal entries is written as $\mathrm{diag}(q_1, \dots, q_n)$. The standard maximal eigenvalue of a positive definite matrix P is denoted by $\lambda_{\max}(P)$. The maximum generalized eigenvalue of square matrices P_1 and P_2 is denoted by $\lambda_{\max}(P_1, P_2)$. For two symmetric

[1] In this paper, we clearly distinguish between *regression* inputs to the GP and *control* inputs to the physical system.

matrices $A,\ B,\ A \leq B$ means that $(B - A)$ is positive semidefinite.

GPs are an increasingly popular method to approximate a nonlinear function $f(d)$. GPs are fully defined by a mean function $m : \mathcal{Z} \to \mathbb{R}$ (where $\mathcal{Z} \subseteq \mathbb{R}^{n_d}$) and a covariance function (also referred to as kernel) $k : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}$ for some regression inputs $d,\ d' \in \mathcal{Z}$

$$f(d) \sim \mathcal{GP}(m(d), k(d, d')). \tag{1}$$

In this work, we consider (as commonly done in the context of GPs) a prior mean $m \equiv 0$ and the squared exponential automatic relevance determination (ARD) kernel, i.e.,

$$k(d_i, d_j) \coloneqq \sigma_f^2 \exp(-\frac{1}{2}(d_i - d_j)^\top \Lambda^{-1}(d_i - d_j)),$$

where $\sigma_f \in \mathbb{R}_{\geq 0}$, $\Lambda = \operatorname{diag}(\varphi_1^2, \varphi_2^2, \ldots, \varphi_{n_d}^2)$ with $\varphi_1,\ \varphi_2,\ \ldots,\ \varphi_{n_d} \in \mathbb{R}_{\geq 0}$.

The GP is trained by conditioning it on some given regression input data[2] $D^d = \begin{bmatrix} d_1^d & d_2^d & \ldots & d_N^d \end{bmatrix}^\top$ and some output data $Y^d = \begin{bmatrix} y_1^d & y_2^d & \ldots & y_N^d \end{bmatrix}^\top$, where each output data point y^d is given by $f(d^d) + \varepsilon^d$ with ε^d being normally distributed with distribution $\mathcal{N}(0, \sigma_\varepsilon^2)$. Then, the posterior mean at some regression test input d_* is given by

$$m_+(d_*|D^d, Y^d) = k(d_*, D^d)(K(D^d, D^d) + \sigma_\varepsilon^2 I)^{-1} Y^d$$

and the posterior variance (which corresponds to the inherent uncertainty quantification) by

$$\sigma_+^2(d_*|D^d, Y^d) = k(d_*, d_*) - k(d_*, D^d)(K(D^d, D^d) + \sigma_\varepsilon^2 I)^{-1} k(D^d, d_*),$$

where $k(d_*, D^d) = \big(k(d_*, d_i)\big)_{d_i \in D^d} = k(D^d, d_*)^\top$, with $k(d_*, D^d) \in \mathbb{R}^{1 \times N}$, and $K(D^d, D^d) = (k(d_i, d_j))_{d_i, d_j \in D^d}$ with $K(D^d, D^d) \in \mathbb{R}^{N \times N}$. The posterior mean depends on the hyperparameters $\sigma_f, \varphi_1, \ldots, \varphi_{n_d}, \sigma_\varepsilon$ that crucially influence the quality of the learned model, compare the discussion in [4, Sec. 2.3]. As commonly done in the literature, we determine the hyperparameters by maximizing the log marginal likelihood, see, e.g., [4, Eq. (2.30)].

After this general introduction to GPs, we now describe how they are exploited in this work. We consider discrete-time nonlinear systems with additive disturbances, i.e.,

$$x(t+1) = f(x(t), u(t)) + w(t) \tag{2a}$$
$$y(t) = h(x(t), u(t)) + v(t) \tag{2b}$$

with $x(t), w(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t), v(t) \in \mathbb{R}^p$, where w and v denote the process and the measurement noise, respectively.

Throughout this paper, we assume that the states and inputs evolve in compact sets, i.e., $x(t) \in \mathbb{X} \subset \mathbb{R}^n$ and $u(t) \in \mathbb{U} \subset \mathbb{R}^m\ \ \forall t \in \mathbb{I}_{\geq 0}$. In the following, we model f and h using GPs. Hence, for modeling purposes only, we consider w and v to be normally distributed. Note that such a setting (assuming bounded states and employing GP models) is common in the GP-based control/estimation literature, compare, e.g., [6], [10]. This corresponds to the realistic scenario in which the real process and measurement disturbances w and v in (2) are not unbounded in practical applications, despite being assumed to be normally distributed within the GP modeling. The hyperparameter σ_ε is determined such that the GP approximates the unknown function as well as possible, compare the discussion in [10, Sec. II C]. Furthermore, we assume that the state transition function f as well as the output mapping h are continuous.

The objective is to learn the state-space model (2) (i.e., the state transition function f and the output mapping h) by means of GPs. In this case, the regression input data is composed of the states and the control inputs at time t, i.e., $d(t) = \begin{bmatrix} x_1(t) & \ldots & x_n(t) & u_1(t) & \ldots & u_m(t) \end{bmatrix}^\top$. Since standard GPs only map on scalar regression outputs, we need to learn $n+p$ independent GPs to approximate the complete dynamics (i.e., n GPs for the components $f_1,\ f_2,\ \ldots,\ f_n$ of the function f, and p GPs for the components $h_1,\ h_2, \ldots,\ h_p$ of the function h). In the following, we denote the kernels associated to the GPs approximating the components of f and h by $k_{x_1},\ k_{x_2}, \ldots, k_{x_n}$ and $k_{y_1},\ k_{y_2}, \ldots, k_{y_p}$, respectively. We collect training data, condition the GPs on the training data, and tune the hyperparameters by maximizing the marginal log-likelihood. To simplify the notation, we write

$$m_{+,x}(d(t)|D^d, X^d) = \begin{pmatrix} m_{+,x_1}(d(t)|D^d, X_1^d) \\ m_{+,x_2}(d(t)|D^d, X_2^d) \\ \vdots \\ m_{+,x_n}(d(t)|D^d, X_n^d) \end{pmatrix} \tag{3}$$

to denote the stacked posterior means approximating the function f and analogously $m_{+,y}(d(t)|D^d, Y^d)$ to denote the stacked posterior means approximating the function h. Here, the regression output X_i^d (and analogously Y_j^d) is given by $X_i^d = \begin{bmatrix} x_i^d(1) & \ldots & x_i^d(N) \end{bmatrix}^\top$. Hence, the learned system dynamics can be expressed as

$$x(t+1) = m_{+,x}(d(t)|D^d, X^d) + \check{w}(t) \tag{4a}$$
$$y(t) = m_{+,y}(d(t)|D^d, Y^d) + \check{v}(t) \tag{4b}$$

with $\check{w} \in \mathbb{R}^n$ and $\check{v} \in \mathbb{R}^p$. Note that we recover the original system (2) for

$$\check{w}(t) \coloneqq f\big(x(t), u(t)\big) - m_{+,x}\big(d(t)|D^d, X^d\big) + w(t) \tag{5}$$
$$\check{v}(t) \coloneqq h\big(x(t), u(t)\big) - m_{+,y}\big(d(t)|D^d, Y^d\big) + v(t). \tag{6}$$

In this work, we consider two different phases. On the one hand, an offline phase, in which noise-free measurements of the control inputs and noisy measurements of the outputs *and* states are available. This assumption allows us to condition the GPs on the training data and to perform the hyperparameter optimization offline. On the other hand, an online phase, in which noise-free measurements of the control inputs, but only noisy measurements of the outputs (but *not* the states) are available, meaning that the states must be estimated. To

[2]Here, we use D^d to denote the regression input data (instead of the commonly used notation X) to avoid confusion with respect to the actual system states.

perform the state estimation, we apply the GP based MHE, which is explained in the following section. The assumption of having noisy state measurements available in an offline phase might be restrictive in general, but is certainly fulfilled in cases where one can measure all the states in a laboratory setting using sophisticated hardware that is not available online, compare also the discussion in [14], [15].

III. GP BASED MHE SCHEME

In this section, we explain in detail the GP based MHE scheme. As usual in MHE, at each time step t, an optimization problem is solved taking the past measurements over some horizon M into account. Namely,

$$\underset{\bar{x}(t-M_t|t),\bar{w}(\cdot|t)}{\text{minimize}} \quad J\big(\bar{x}(t-M_t|t),\bar{w}(\cdot|t),\bar{v}(\cdot|t),t\big) \tag{7a}$$

$$\text{s.t. } \bar{x}(j+1|t) = m_{+,x}(\bar{d}(j|t)|D^d,X^d) + \bar{w}(j|t), \tag{7b}$$

$$y(j) = m_{+,y}(\bar{d}(j|t)|D^d,Y^d) + \bar{v}(j|t), \quad \forall j \in \mathbb{I}_{[t-M_t,t-1]} \tag{7c}$$

$$\bar{x}(j|t) \in \mathbb{X} \quad \forall j \in \mathbb{I}_{[t-M_t,t]} \tag{7d}$$

with $M_t = \min\{t, M\}$ (M being the horizon length),

$$\bar{d}(j|t) \coloneqq \begin{bmatrix}\bar{x}_1(j|t) & \dots & \bar{x}_n(j|t) & u_1(j) & \dots & u_m(j)\end{bmatrix}^\top$$

and

$$\begin{aligned} J(\bar{x}(t-M_t|t),\bar{w}(\cdot|t),\bar{v}(\cdot|t),t) \\ \coloneqq & 2||\bar{x}(t-M_t|t)-\hat{x}(t-M_t)||^2_{P_2}\eta^{M_t} \\ & + \sum_{j=1}^{M_t} 2\eta^{j-1}\Big(||\bar{w}(t-j|t)||^2_{Q^{-1}_{\bar{d}(t-j|t)}} \\ & \qquad + ||\bar{v}(t-j|t)||^2_{R^{-1}_{\bar{d}(t-j|t)}}\Big). \end{aligned} \tag{7e}$$

In (7e), $\eta \in (0,1)$ is a discount factor. The notation $\bar{d}(j|t)$ denotes the estimated state (together with the measured control input) at time j, estimated at time t. The estimated process and measurement noise trajectories, estimated at time t, are denoted by $\bar{w}(\cdot|t)$ and $\bar{v}(\cdot|t)$, respectively. The cost function is composed of two terms: the prior weighting and the stage cost. Hence, the cost function trades off how much we believe the measurements within the current horizon and how much we believe the prior $\hat{x}(t-M_t)$. The optimal estimated state sequence is denoted by $\hat{x}(\cdot|t)$ (analogous for $\hat{d}(\cdot|t)$) and the estimated system state at time t is set to the last element of the estimated state sequence, i.e., $\hat{x}(t) \coloneqq \hat{x}(t|t)$.

Note the first main difference to standard model-based MHE schemes in (7b) and in (7c). We exploit the posterior mean functions of the GPs to approximate the state transition function f and the output mapping h.

The weighting matrices in the cost function are chosen as

$$\begin{aligned} Q_{\bar{d}(t-j|t)} = \operatorname{diag}\big(&\sigma^2_{+,x_1}(\bar{d}(t-j|t)|D^d,X^d_1),\dots, \\ &\sigma^2_{+,x_n}(\bar{d}(t-j|t)|D^d,X^d_n)\big) + Q_0 \end{aligned} \tag{8}$$

$$\begin{aligned} R_{\bar{d}(t-j|t)} = \operatorname{diag}\big(&\sigma^2_{+,y_1}(\bar{d}(t-j|t)|D^d,Y^d_1),\dots, \\ &\sigma^2_{+,y_p}(\bar{d}(t-j|t)|D^d,Y^d_p)\big) + R_0 \end{aligned} \tag{9}$$

with Q_0, R_0 positive definite (and P_2 positive definite) and $\sigma_{+,x_i}, \sigma_{+,y_j}$ denoting the posterior variances of the $n+p$ GPs modeling the components of the functions f and h, respectively. This choice of the weighting matrices constitutes the second main difference to standard MHE schemes. The weighting matrices $Q_{\bar{d}(t-j|t)}$ and $R_{\bar{d}(t-j|t)}$ are a sum of two matrices. The first one is a diagonal matrix, where the diagonal entries correspond to the posterior variances of the corresponding states/outputs, as in the work related to GP based extended Kalman filtering [9]. Loosely speaking, the beneficial effect of this choice is the following: in a region of low (high) training data availability, the posterior variance, representing the uncertainty of the learned model, is rather large (small). In turn, the inverse weighing matrices, on which the cost function is based, induce a low (high) weight on $\bar{w}$ and $\bar{v}$. Consequently, we allow for large (small) magnitudes of $\bar{w}$ and $\bar{v}$. This is meaningful, since in areas of low (high) training data availability, the mean functions will be poor (good) approximations of the true functions f and h. The second matrix corresponds to the standard MHE weighting matrix. The matrices Q_0 and R_0 are typically set according to the variance of the process/measurement noise affecting the online measurements [16]. The choice of the matrix P_2 is more difficult in the general nonlinear case [16, Sec. 3.1]. One model-based approach to design this matrix has recently been proposed in [17, Cor. 3].

As long as $t < M$, we use the so-called full information estimator, i.e., $M_t = t$, meaning that all available measurements are taken into account.

IV. ROBUST STABILITY ANALYSIS

In this section, we prove robust stability of the GP based MHE scheme based on the following definition, which is similar to [17, Def. 2], with the main difference that we consider a *practical* stability notion that can capture the mismatch between the posterior means and the true functions f and h.

Definition 1: A state estimator for system (2) is practically robustly exponentially stable (pRES) if there exist $C_1, C_2, C_3 > 0$, $\lambda_1, \lambda_2, \lambda_3 \in [0,1)$, and $\alpha > 0$ such that the resulting state estimates $\hat{x}(t)$ satisfy

$$\begin{aligned} \|x(t)-\hat{x}(t)\| \le \max\Big\{ & C_1\|x(0)-\hat{x}(0)\|\lambda_1^t, \\ & \max_{j\in\mathbb{I}_{[0,t-1]}} C_2\|w(j)\|\lambda_2^{t-j-1}, \max_{j\in\mathbb{I}_{[0,t-1]}} C_3\|v(j)\|\lambda_3^{t-j-1}, \alpha\Big\} \end{aligned} \tag{10}$$

for all $t \in \mathbb{I}_{\ge 0}$, all initial conditions $x(0), \hat{x}(0) \in \mathbb{X}$, and every trajectory $(x(t),u(t),w(t),v(t))_{t=0}^{\infty}$ satisfying the system dynamics (2).

Next, we introduce the matrices $Q^{-1}_{\min}$, $R^{-1}_{\min}$ and $Q^{-1}_{\max}$, $R^{-1}_{\max}$ such that

$$Q^{-1}_{\min} \le Q^{-1}_{\bar{d}(t)} \le Q^{-1}_{\max}, \tag{11}$$

$$R^{-1}_{\min} \le R^{-1}_{\bar{d}(t)} \le R^{-1}_{\max}. \tag{12}$$

The matrices $Q_{\min}^{-1}$ and $R_{\min}^{-1}$ represent the case when the regression test inputs are (infinitely) far away from the regression training inputs, meaning that the posterior variance is maximal. In case of the here considered squared exponential ARD kernel, an upper bound for the maximal posterior variance is given by σ_f^2, i.e., a lower bound for $Q_{\min}^{-1}$ is given by $Q_{\min}^{-1} \coloneqq \left[\mathrm{diag}(\sigma_{f,x_1}^2, \dots, \sigma_{f,x_n}^2) + Q_0\right]^{-1}$ and $R_{\min}^{-1}$ is defined analogously. In turn, the matrices $Q_{\max}^{-1},\ R_{\max}^{-1}$ correspond to the minimal possible posterior variances. A lower bound for the minimal posterior variance is 0, which occurs in the noise-free case, when a regression test input corresponds exactly to a regression training input. Consequently, an upper bound for $Q_{\max}^{-1}$ (and similarly $R_{\max}^{-1}$) is given by $Q_{\max}^{-1} \coloneqq Q_0^{-1}$. To prove robust stability of the GP based MHE scheme, we assume that the learned system (4) satisfies a detectability notion called incremental input/output-to-state stability (δ-IOSS), which has been frequently used to prove stability of various MHE schemes, compare [17], [18], [1]. The assumption applied in this work is proposed in a similar way in [17] with the main difference that we here consider the learned system dynamics and not the true system dynamics.

Assumption 1: The system (4) admits a δ-IOSS Lyapunov function $W_\delta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ with quadratic bounds and supply rate, i.e., there exist $\eta \in [0,1),\ P_1,\ P_2,\ Q_0,\ R_0 > 0$ such that

$$||x - \tilde{x}||_{P_1}^2 \leq W_\delta(x, \tilde{x}) \leq ||x - \tilde{x}||_{P_2}^2, \tag{13a}$$

$$\begin{aligned} W_\delta\Big(m_{+,x}(d|D^d, X^d) + \breve{w}, m_{+,x}(\tilde{d}|D^d, X^d) + \breve{w}'\Big) \\ \leq \eta W_\delta(x, \tilde{x}) + ||\breve{w} - \breve{w}'||_{Q_{\min}^{-1}}^2 \\ + ||m_{+,y}(d|D^d, Y^d) - m_{+,y}(\tilde{d}|D^d, Y^d)||_{R_{\min}^{-1}}^2 \end{aligned} \tag{13b}$$

for all $(x, u, \breve{w}), (\tilde{x}, u, \breve{w}')$ with $x, \tilde{x} \in \mathbb{X}$ and $u \in \mathbb{U}$, where $d = \begin{bmatrix} x_1 & \dots & x_n & u_1 & \dots & u_m \end{bmatrix}^\top$, $\tilde{d} = \begin{bmatrix} \tilde{x}_1 & \dots & \tilde{x}_n & u_1 & \dots & u_m \end{bmatrix}^\top$, and $Q_{\min}^{-1}$ and $R_{\min}^{-1}$ are from (11) and (12), respectively.

Note that existence of a δ-IOSS Lyapunov function is equivalent to the system being δ-IOSS [19]. This property is necessary and sufficient for the existence of state estimators for nonlinear systems and has widely been used in the recent MHE literature, compare [17] for a more detailed discussion.

Remark 1: After having determined the posterior mean and variance, one can verify Assumption 1 using the results of [17, Sec. IV]. An interesting property for future research is to study whether Assumption 1 is always satisfied (i.e., the learned GP model admits a δ-IOSS Lyapunov function) if the true unknown system (2) admits a δ-IOSS Lyapunov function (i.e., is detectable).

To simplify the notation in the following proof, we define

$$\alpha_1^{\max} \coloneqq \max_{x \in \mathbb{X}, u \in \mathbb{U}} \left\{ \|f(x,u) - m_{+,x}(d|D^d, X^d)\|_{Q_{\max}^{-1}} \right\} \tag{14}$$

$$\alpha_2^{\max} \coloneqq \max_{x \in \mathbb{X}, u \in \mathbb{U}} \left\{ \|h(x,u) - m_{+,y}(d|D^d, Y^d)\|_{R_{\max}^{-1}} \right\} \tag{15}$$

and $\alpha^{\max} \coloneqq \max\{\alpha_1^{\max}, \alpha_2^{\max}\}$. Notice that these constants exist, since we assume that (i) the states and the inputs evolve in compact sets, (ii) the functions f and h are continuous and since the here considered squared exponential ARD kernel leads to a continuous posterior mean.

Theorem 1: Let Assumption 1 hold. Then, there exist $\mu \in [0,1)$ and a minimal horizon length $\bar{M}$ such that for all $M \in \mathbb{I}_{\geq \bar{M}}$, the state estimation error of the GP based MHE (7) is bounded for all $t \in \mathbb{I}_{\geq 0}$ by

$$\begin{aligned} \|\hat{x}(t) - x(t)\|_{P_1} \leq \max \Bigg\{ & 6\sqrt{\mu}^t \|\hat{x}(0) - x(0)\|_{P_2}, \\ & \max_{q \in \mathbb{I}_{[0,t-1]}} \left\{ \frac{12}{1 - \sqrt[4]{\mu}} \sqrt[4]{\mu}^q \|w(t-q-1)\|_{Q_{\max}^{-1}} \right\}, \\ & \max_{q \in \mathbb{I}_{[0,t-1]}} \left\{ \frac{12}{1 - \sqrt[4]{\mu}} \sqrt[4]{\mu}^q \|v(t-q-1)\|_{R_{\max}^{-1}} \right\}, \\ & \frac{12}{1 - \sqrt[4]{\mu}} \alpha^{\max} \Bigg\}. \end{aligned} \tag{16}$$

Consequently, the GP based MHE (7) is pRES according to Definition 1.

The proof of Theorem 1 can be found in Appendix A. It mainly relies on the Lyapunov-based robust stability proof technique recently proposed in the context of model-based MHE in [17, Prop. 1, Thm. 1, Cor. 1]. Nevertheless, in our case we need to take into account two crucial differences, namely, (i) that the estimated system trajectory is based on the learned dynamics and (ii) that the weighting matrices in the cost function are not constant.

Theorem 1 shows that the state estimation error is upper bounded by means of (i) the initial state estimation error, (ii) the true process noise, (iii) the true measurement noise, and (iv) the mismatch between the learned system model and the true system dynamics.

Remark 2: In case of higher training data availability, $R_{\min}^{-1}$ and $Q_{\min}^{-1}$ increase, since the maximal posterior variances decrease. Larger values of $R_{\min}^{-1}$ and $Q_{\min}^{-1}$ in (13b) allow for a larger W_δ, which then allows for a larger P_1 in (13a). In turn, this results in a less conservative error bound (16). Hence, a higher training data availability results in less conservative state estimation error bounds.

Note that the estimation error bound (16) of Theorem 1 depends on $\alpha^{\max}$ that accounts for the mismatch between the learned system dynamics and the true system dynamics. A probabilistic upper bound for this mismatch can be obtained [20] by making the following additional assumption.

Assumption 2: Each component of the unknown functions f and h is Lipschitz continuous and a sample from a GP, i.e., f_i is a sample of $\mathcal{GP}(0, k_{x_i}(d, d'))$, $i = 1, \dots, n$ and h_j a sample of $\mathcal{GP}(0, k_{y_j}(d, d'))$, $j = 1, \dots, p$.

Furthermore, for a grid constant τ, we define

$$\beta(\tau) \coloneqq 2 \log\left(\frac{B(\tau, \mathbb{X})}{\delta}\right) \tag{17}$$

$$\gamma_{f_1}(\tau) \coloneqq (L_{m_{+,x_1}} + L_{f_1})\tau + \sqrt{\beta(\tau)}\omega_{\sigma_{+,x_1}}(\tau), \tag{18}$$

$$\gamma_{h_1}(\tau) \coloneqq (L_{m_{+,y_1}} + L_{h_1})\tau + \sqrt{\beta(\tau)}\omega_{\sigma_{+,y_1}}(\tau), \tag{19}$$

and similarly $\gamma_{f_i}, \forall i = 2,\dots,n$ and $\gamma_{h_j}, \forall j = 2,\dots,p$. The constant B denotes the covering number, which corresponds to the minimum number of points in a grid over $\mathbb{X}$ considering the grid constant τ. The constants $L_{m_{x_1}},\dots,L_{m_{x_n}},L_{m_{y_1}},\dots,L_{m_{y_p}}$ denote the Lipschitz constants of the mean functions and $\omega_{\sigma_{+,x_1}},\dots,\omega_{\sigma_{+,x_n}},\omega_{\sigma_{+,y_1}},\dots\omega_{\sigma_{+,y_p}}$, the moduli of continuity of the kernels. Finally, $\delta \in (0,1)$ and $L_{f_1},\dots,L_{f_n},L_{h_1},\dots,L_{h_p}$ are the Lipschitz constants of the components of the unknown functions of f and h. The definitions (17) - (19) were made in [20], and the reader is referred to this reference for additional details. Moreover, we introduce

$$\Delta_x^{\max}(\tau) \coloneqq \sqrt{\lambda_{\max}(Q_{\max}^{-1})}\times \sum_{i=1}^{n} \max_{x\in\mathbb{X},u\in\mathbb{U}} \left\{\|\sqrt{\beta(\tau)}\sigma_{+,x_i}(d|D^d,X_i^d)+\gamma_{f_i}(\tau)\|\right\}, \tag{20}$$

$$\Delta_y^{\max}(\tau) \coloneqq \sqrt{\lambda_{\max}(R_{\max}^{-1})}\times \sum_{i=1}^{p} \max_{x\in\mathbb{X},u\in\mathbb{U}} \left\{\|\sqrt{\beta(\tau)}\sigma_{+,y_i}(d|D^d,Y_i^d)+\gamma_{h_i}(\tau)\|\right\}, \tag{21}$$

which will be used in the following corollary to simplify the notation.

Corollary 1: Let Assumptions 1 - 2 hold. Then, there exist $\mu \in [0,1)$ and a minimal horizon length $\bar{M}$ such that for all $M \in \mathbb{I}_{\geq \bar{M}}$ the state estimation error of the GP based MHE (7) is (probabilistically) bounded for all $t \in \mathbb{I}_{\geq 0}$ by

$$\begin{aligned} P\Bigg(\|\hat{x}(t)-x(t)\|_{P_1} \leq \max\Bigg\{ & 6\sqrt{\mu}^t\|\hat{x}(0)-x(0)\|_{P_2}, \\ & \max_{q\in\mathbb{I}_{[0,t-1]}}\left\{\frac{12}{1-\sqrt[4]{\mu}}\sqrt[4]{\mu}^q\|w(t-q-1)\|_{Q_{\max}^{-1}}\right\}, \\ & \max_{q\in\mathbb{I}_{[0,t-1]}}\left\{\frac{12}{1-\sqrt[4]{\mu}}\sqrt[4]{\mu}^q\|v(t-q-1)\|_{R_{\max}^{-1}}\right\}, \\ & \frac{12}{1-\sqrt[4]{\mu}}\Delta_x^{\max}(\tau), \frac{12}{1-\sqrt[4]{\mu}}\Delta_y^{\max}(\tau), \Bigg\}\Bigg) \\ & \geq (1-\delta)^{n+p}. \end{aligned} \tag{22}$$

The proof of Corollary 1 is shown in Appendix B. The key idea of the proof is to bound each component of the difference between the functions f, h and the posterior means by applying the probabilistic bound developed in [20, Thm. 3.1].

Corollary 1 uses a probabilistic upper bound for the mismatch between the learned and the true unknown dynamics. As a result, also the obtained estimation error bound (22) is probabilistic in nature. Note that the final probability in (22) decreases for a higher state/output dimension. This is due to the current (conservative) proof technique of bounding component-wise the difference between the functions f, h and the posterior means. Developing a less conservative proof is an interesting subject for future research.

As can be seen from the definition of β in (17), a higher probability (i.e., a smaller δ) results in more conservative upper bounds and vice versa. Furthermore, as discussed in Remark 2, more training data will improve the posterior variance and thus the (probabilistic) estimation error bounds (22) that explicitly depend on $\sigma_{+,x_1},\dots,\ \sigma_{+,x_n}$ and $\sigma_{+,y_1},\dots,\ \sigma_{+,y_p}$.

V. APPLICATION TO BATCH REACTOR SYSTEM

In this section, we illustrate the performance of the GP based MHE. To this end, we consider the following Euler-discretized system

$$\begin{aligned} x_1(t+1) &= x_1(t)+T(-2k_1x_1^2(t)+2k_2x_2(t))+w_1(t) \\ x_2(t+1) &= x_2(t)+T(k_1x_1^2(t)-k_2x_2(t))+w_2(t) \\ y(t) &= x_1(t)+x_2(t)+v(t) \end{aligned}$$

with sampling time $T = 0.1$, constants $k_1 = 0.16$, $k_2 = 0.0064$ which corresponds to a batch reactor system [1, Ch. 4], [21]. This system is a benchmark example in the MHE literature, since other nonlinear state estimation techniques, such as the extended Kalman filter, can fail to converge, compare, [1].

As mentioned in Section II, we consider two different phases. In both phases, we consider $w \sim \mathcal{N}(0,\sigma_w^2 I_n)$ with $\sigma_w = 0.01$ and $v \sim \mathcal{N}(0,\sigma_v^2 I_p)$ with $\sigma_v = 0.1$. In the offline phase, we collect five different state/output trajectories (of length 31) with the following initial conditions $x_{01} = \begin{bmatrix}3 & 1\end{bmatrix}^\top$, $x_{02} = \begin{bmatrix}1.2 & 4.5\end{bmatrix}^\top$, $x_{03} = \begin{bmatrix}0.5 & 3.5\end{bmatrix}^\top$, $x_{04} = \begin{bmatrix}1 & 3\end{bmatrix}^\top$, $x_{05} = \begin{bmatrix}2 & 4\end{bmatrix}^\top$ and perform the hyperparameter optimization by maximizing the log marginal likelihood. In the online phase, we apply the MHE scheme (7) using $\eta = 0.91$, $M = 15$, initial condition $x_0 = \begin{bmatrix}3 & 1\end{bmatrix}^\top$, and $P_2 = I_n$, $R_0 = 100$ and $Q_0 = \mathrm{diag}(1000,1000)$. As in [17, Sec. V], we consider that the states evolve in a compact set $\mathbb{X} = \{x \in \mathbb{R}^2 : 0.1 \leq x_i \leq 4.5,\ i = \{1,2\}\}$. In addition, we illustrate the performance of the GP based MHE with the same parameters, when only three trajectories (for initial conditions $x_{01} = \begin{bmatrix}0.5 & 3.5\end{bmatrix}^\top$, $x_{04} = \begin{bmatrix}1 & 3\end{bmatrix}^\top$, $x_{05} = \begin{bmatrix}2 & 4\end{bmatrix}^\top$), i.e., less data, are collected in the offline phase. Finally, we implement a standard model-based MHE scheme (that is based on exact model knowledge) with the same characteristics, i.e., using the same P_2, Q_0, R_0, η, and M. The obtained results are illustrated in Figure 1. As guaranteed by Theorem 1, the GP based MHE scheme is robustly stable. The estimation performance improves when more training data is available. Furthermore, the GP based MHE (related to five collected trajectories) performs similarly well as the model-based MHE.

VI. CONCLUSION

In this paper, we introduced a GP based MHE framework for which we proved practical robust exponential stability. The framework leverages the posterior mean of the GP to replace the required mathematical model in the MHE scheme and the posterior variance to account for the uncertainty of the learned model within the cost function. This allows for

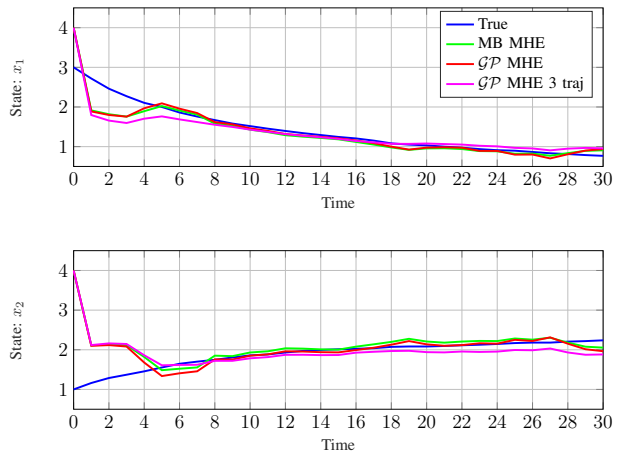

Fig. 1. Simulation results of the GP based MHE scheme (7).

an effective way to estimate the states of unknown nonlinear systems, as was also illustrated by a batch reactor example.

Future work includes investigating how detectability (δ-IOSS) can efficiently be verified for the learned system, in particular whether/when the learned model inherits this property if the true unknown system is detectable.

REFERENCES

[1] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model predictive control: theory, computation, and design*, 2nd ed. Nob Hill Publishing Madison, WI, 2020.

[2] A. Alessandri, M. Baglietto, and G. Battistelli, "Robust receding-horizon state estimation for uncertain discrete-time linear systems," *Systems & Control Letters*, vol. 54, no. 7, pp. 627–643, 2005.

[3] M. Gharbi and C. Ebenbauer, "Proximity moving horizon estimation for linear time-varying systems and a Bayesian filtering view," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 3208–3213.

[4] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[5] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using Gaussian process regression," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2736–2743, 2019.

[6] M. Maiworm, D. Limon, and R. Findeisen, "Online learning-based model predictive control with Gaussian process models and stability guarantees," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 18, pp. 8785–8812, 2021.

[7] A. Alessandri, M. Baglietto, G. Battistelli, and M. Gaggero, "Moving-horizon state estimation for nonlinear systems using neural networks," *IEEE Transactions on Neural Networks*, vol. 22, no. 5, pp. 768–780, 2011.

[8] W. Cao, J. Duan, S. E. Li, C. Chen, C. Liu, and Y. Wang, "Primal-dual estimator learning method with feasibility and near-optimality guarantees," in *IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 4104–4111.

[9] J. Ko and D. Fox, "GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models," *Autonomous Robots*, vol. 27, pp. 75–90, 2009.

[10] M. Buisson-Fenet, V. Morgenthaler, S. Trimpe, and F. Di Meglio, "Joint state and dynamics estimation with high-gain observers and Gaussian process models," in *2021 American Control Conference (ACC)*. IEEE, 2021, pp. 4027–4032.

[11] T. D. Barfoot, *State estimation for robotics*. Cambridge University Press, 2022.

[12] C. H. Tong, P. Furgale, and T. D. Barfoot, "Gaussian process Gauss-Newton: Non-parametric state estimation," in *2012 Ninth Conference on Computer and Robot Vision*. IEEE, 2012, pp. 206–213.

[13] C. H. Tong and T. D. Barfoot, "Gaussian process Gauss-Newton for 3D laser-based visual odometry," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 5204–5211.

[14] M. S. Turan and G. Ferrari-Trecate, "Data-driven unknown-input observers and state estimation," *IEEE Control Systems Letters*, vol. 6, pp. 1424–1429, 2022.

[15] T. M. Wolff, V. G. Lopez, and M. A. Müller, "Robust data-driven moving horizon estimation for linear discrete-time systems," *arXiv preprint arXiv:2210.09017*, 2022.

[16] C. V. Rao, "Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems," 2000.

[17] J. D. Schiller, S. Muntwiler, J. Köhler, M. N. Zeilinger, and M. A. Müller, "A Lyapunov function for robust stability of moving horizon estimation," *arXiv preprint arXiv:2202.12744*, 2022.

[18] D. A. Allan and J. B. Rawlings, "Moving horizon estimation," in *Handbook of Model Predictive Control*, S. V. Raković and W. S. Levine, Eds. Cham: Springer International Publishing, 2019, pp. 99–124.

[19] D. A. Allan, J. Rawlings, and A. R. Teel, "Nonlinear detectability and incremental input/output-to-state stability," *SIAM Journal on Control and Optimization*, vol. 59, no. 4, pp. 3017–3039, 2021.

[20] A. Lederer, J. Umlauft, and S. Hirche, "Uniform error bounds for Gaussian process regression with application to safe control," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[21] M. J. Tenny and J. B. Rawlings, "Efficient moving horizon estimation and nonlinear model predictive control," in *Proceedings of the 2002 American Control Conference*, vol. 6. IEEE, 2002, pp. 4475–4480.

APPENDIX

A. Proof of Theorem 1

The proof of Theorem 1 is based on the developments shown in [17]. Here, due to space restriction, we mainly comment on the steps that are conceptually different from the proof in [17], without describing the similar steps of the proof in all detail.

Proof: The constraints in the MHE problem guarantee that the (optimal) estimated system trajectory (denoted by $\hat{x}(j|t), u(t), \hat{w}(j|t), \hat{v}(j|t)$ for all $j \in \mathbb{I}_{[t-M_t,t-1]}$) fulfills the learned system dynamics. The unknown (true) system trajectory cannot necessarily be represented by the posterior means. Therefore, we use the introduced auxiliary variables $\check{w}$ (5) and $\check{v}$ (6) to represent the true system trajectory by the posterior means in the following. Exploiting that due to (7c) $m_{+,y}(d(t-j)|D^d, Y^d) - m_{+,y}(\hat{d}(t-j|t)|D^d, Y^d) = \hat{v}(t-j|t) - \check{v}(t-j)$ for all $j \in \mathbb{I}_{[t-M_t,t-1]}$ and applying inequality (13b) M_t times yields

$$\begin{aligned} W_\delta(\hat{x}(t), x(t)) \overset{(13b)}{\leq} \sum_{j=1}^{M_t} \eta^{j-1} \Big(&||\hat{w}(t-j|t) - \check{w}(t-j)||^2_{Q^{-1}_{\min}} \\ &+ ||\hat{v}(t-j|t) - \check{v}(t-j)||^2_{R^{-1}_{\min}} \Big) \\ &+ \eta^{M_t} W_\delta(\hat{x}(t-M_t|t), x(t-M_t)). \end{aligned}$$

Moreover, using (13a), the triangle inequality, the Cauchy-Schwarz inequality, and Young's inequality, it holds

$$\begin{aligned} &W_\delta(\hat{x}(t), x(t)) \\ &\leq 2\eta^{M_t} ||\hat{x}(t-M_t) - x(t-M_t)||^2_{P_2} \\ &\quad + \sum_{j=1}^{M_t} 2\eta^{j-1} \Big(||\check{w}(t-j)||^2_{Q^{-1}_{\min}} + ||\check{v}(t-j)||^2_{R^{-1}_{\min}} \Big) \\ &\quad + J_{\min}(\hat{x}(t-M_t|t), \hat{w}(\cdot|t), \hat{v}(\cdot|t), t). \end{aligned} \tag{23}$$

with

$$J_{\min}(\hat{x}(t-M_t|t),\hat{w}(\cdot|t),\hat{v}(\cdot|t),t) \coloneqq 2\eta^{M_t}||\hat{x}(t-M_t|t)-\hat{x}(t-M_t)||^2_{P_2} + \sum_{j=1}^{M_t} 2\eta^{j-1}\Big(||\hat{w}(t-j|t)||^2_{Q^{-1}_{\min}} + ||\hat{v}(t-j|t)||^2_{R^{-1}_{\min}}\Big).$$

Note that $J_{\min}(\hat{x}(t-M_t|t),\hat{w}(\cdot|t),\hat{v}(\cdot|t),t)$ does not correspond to the optimal cost of problem (7). In fact, $J_{\min}(\hat{x}(t-M_t|t),\hat{w}(\cdot|t),\hat{v}(\cdot|t),t)$ corresponds to the cost of the optimal trajectory, when $R^{-1}_{\min}$, $Q^{-1}_{\min}$ (and P_2) are considered in the cost function (7e) (but *not* the variable $R^{-1}_{\bar{d}(t-j|t)}$, $Q^{-1}_{\bar{d}(t-j|t)}$). Analogously, we define $J_{\max}$ by replacing $R^{-1}_{\bar{d}(t-j|t)}$ and $Q^{-1}_{\bar{d}(t-j|t)}$ in (7e) with $R^{-1}_{\max}$ and $Q^{-1}_{\max}$, respectively. Next, we upper bound $J_{\min}(\hat{x}(t-M_t|t),\hat{w}(\cdot|t),\hat{v}(\cdot|t),t)$ as follows

$$\begin{aligned} J_{\min}(\hat{x}(t-M_t|t),&\hat{w}(\cdot|t),\hat{v}(\cdot|t),t) \\ &\le J^*(\hat{x}(t-M_t|t),\hat{w}(\cdot|t),\hat{v}(\cdot|t),t) \\ &\le J(x(t-M_t),\breve{w}(\cdot),\breve{v}(\cdot),t) \\ &\le J_{\max}(x(t-M_t),\breve{w}(\cdot),\breve{v}(\cdot),t), \end{aligned}$$

where the first inequality holds by (11) - (12), the second is due to optimality (i.e., the true unknown system trajectory $x(\cdot),\breve{w}(\cdot),\breve{v}(\cdot)$ is a feasible but in general suboptimal solution to problem (7)), and the third again follows from (11) - (12). We consider these bounds in inequality (23) together with (11) and (12) and obtain

$$\begin{aligned} &W_\delta(\hat{x}(t),x(t)) \\ &\le 4\lambda_{\max}(P_2,P_1)\eta^{M_t}W_\delta(\hat{x}(t-M_t),x(t-M_t)) \\ &\quad + \sum_{j=1}^{M_t} 4\eta^{j-1}\Big(||\breve{w}(t-j)||^2_{Q^{-1}_{\max}} + ||\breve{v}(t-j)||^2_{R^{-1}_{\max}}\Big). \end{aligned}$$

We choose M large enough such that $\mu^M \coloneqq 4\lambda_{\max}(P_2,P_1)\eta^M < 1$ with $\mu \in [0,1)$, and get for all $t \in \mathbb{I}_{\ge M}$

$$\begin{aligned} W_\delta(\hat{x}(t),x(t)) &\le \mu^M W_\delta(\hat{x}(t-M),x(t-M)) \\ &\quad + \sum_{j=1}^{M} 4\eta^{j-1}\Big(||\breve{w}(t-j)||^2_{Q^{-1}_{\max}} + ||\breve{v}(t-j)||^2_{R^{-1}_{\max}}\Big). \end{aligned}$$

Performing similar steps as the ones in the proof of [17, Cor. 1] results in the following state estimation error bound

$$\begin{aligned} \|\hat{x}(t)-x(t)\|_{P_1} \le\ & \max\Bigg\{6\sqrt{\mu}^t\|\hat{x}(0)-x(0)\|_{P_2}, \\ & \max_{q\in\mathbb{I}_{[0,t-1]}}\left\{\frac{6}{1-\sqrt[4]{\mu}}\sqrt[4]{\mu}^q\|\breve{w}(t-q-1)\|_{Q^{-1}_{\max}}\right\}, \\ & \max_{q\in\mathbb{I}_{[0,t-1]}}\left\{\frac{6}{1-\sqrt[4]{\mu}}\sqrt[4]{\mu}^q\|\breve{v}(t-q-1)\|_{R^{-1}_{\max}}\right\}\Bigg\}. \end{aligned} \tag{24}$$

We replace $\breve{w}$ and $\breve{v}$ according to (5) and (6), respectively. Then, we apply the triangle inequality and bound the difference between the functions f, h and the posterior means $m_{+,x}$, $m_{+,y}$ by (14) and (15), respectively. Furthermore, using that $\max_{q\in\mathbb{I}_{[0,t-1]}}\sqrt[4]{\mu}^q = 1$ and $a+b \le \max\{2a,2b\}$ for any $a,b \ge 0$, we have

$$\begin{aligned} \|\hat{x}(t)-x(t)\|_{P_1} \le \max\Bigg\{& 6\sqrt{\mu}^t\|\hat{x}(0)-x(0)\|_{P_2}, \\ & \max_{q\in\mathbb{I}_{[0,t-1]}}\left\{\frac{12}{1-\sqrt[4]{\mu}}\sqrt[4]{\mu}^q\|w(t-q-1)\|_{Q^{-1}_{\max}}\right\}, \\ & \max_{q\in\mathbb{I}_{[0,t-1]}}\left\{\frac{12}{1-\sqrt[4]{\mu}}\sqrt[4]{\mu}^q\|v(t-q-1)\|_{R^{-1}_{\max}}\right\}, \\ & \frac{12}{1-\sqrt[4]{\mu}}\alpha_1^{\max}, \frac{12}{1-\sqrt[4]{\mu}}\alpha_2^{\max}\Bigg\}. \end{aligned} \tag{25}$$

Finally, we use $\alpha_{\max}$ to bound the last two terms of (25), which leads to the expression of Theorem 1. ■

B. Proof of Corollary 1

Proof: From the expressions in (14) - (15), we can bound $\alpha_1^{\max}$ and $\alpha_2^{\max}$ as follows

$$\alpha_1^{\max} \le \sqrt{\lambda_{\max}(Q^{-1}_{\max})}\times \sum_{i=1}^{n}\max_{x\in\mathbb{X},u\in\mathbb{U}}\big\{||f_i(x,u)-m_{+,x_i}(d|D^d,X_i^d)||\big\}. \tag{26}$$

$$\alpha_2^{\max} \le \sqrt{\lambda_{\max}(R^{-1}_{\max})}\times \sum_{i=1}^{p}\max_{x\in\mathbb{X},u\in\mathbb{U}}\big\{||h_i(x,u)-m_{+,y_i}(d|D^d,Y_i^d)||\big\}. \tag{27}$$

From here on, we apply [20, Thm 3.1] to probabilistically bound the difference between the true function components of f, h and the corresponding posterior means

$$\begin{aligned} P\Big(&||f_i(x,u)-m_{x_i}(d|D^d,X_i^d)|| \\ &\le \sqrt{\beta(\tau)}\sigma_{+,x_i}(d|D^d,X_i^d)+\gamma_{f_i}(\tau),\forall x\in\mathbb{X},u\in\mathbb{U}\Big) \\ &\ge 1-\delta \qquad\qquad i=1,\dots,n \\ P\Big(&||h_j(x,u)-m_{y_j}(d|D^d,Y_j^d)|| \\ &\le \sqrt{\beta(\tau)}\sigma_{+,y_j}(d|D^d,Y_j^d)+\gamma_{h_j}(\tau),\forall x\in\mathbb{X},u\in\mathbb{U}\Big) \\ &\ge 1-\delta \qquad\qquad j=1,\dots,p. \end{aligned}$$

Since the GPs are considered to be independent, the probability that all the components of f jointly fulfill their bounds is lower bounded by $(1-\delta)^n$ (the same holds for h with probability $(1-\delta)^p$). We replace the right-hand sides of (26) and (27) by these probabilistic bounds, i.e.,

$$P\big(\alpha_1^{\max} \le \Delta_x^{\max}(\tau)\big) \ge (1-\delta)^n \tag{28}$$

$$P\big(\alpha_2^{\max} \le \Delta_y^{\max}(\tau)\big) \ge (1-\delta)^p. \tag{29}$$

Finally, the probability that both $\alpha_1^{\max} \le \Delta_x^{\max}(\tau)$ and $\alpha_2^{\max} \le \Delta_y^{\max}(\tau)$ hold jointly is lower bounded by $(1-\delta)^{(n+p)}$. Using this bound in (25), we obtain the left-hand side of (22). ■