Control and Uncertainty Propagation in the Presence of Outliers by Utilizing Student-t Process Regression

Dimitris Papadimitriou Somayeh Sojoudi

***Abstract*— Gaussian Process Regression (GPR) has been extensively used to estimate unknown models and quantify model uncertainty in control tasks concerning safety-critical applications. However, one of the drawbacks of GPR is that it does not take into account the function evaluations of the observations in the uncertainty estimation, rendering it unsuitable for applications with observations prone to outliers or to an unassumed noise disturbance. In this work, we introduce the Student-t Process Regression (TPR) as a generalization of GPR for estimating dynamics models in the control literature. The key attribute of TPR is that the estimation variance explicitly depends on the function evaluations rendering it more robust to outliers. We prove uniform error bounds for the estimation based on TPR under certain continuity assumptions. Furthermore, we employ TPR to estimate unknown and nonlinear dynamical systems, which we incorporate in a Model Predictive Control (MPC) framework, and we show with simulations that the resulting estimation uncertainty compensates for the existence of outliers. Such informative variance estimates are of vital importance as they can lead to more informative uncertainty propagation and thus more robust and less constrained control policies.**

I. Introduction

In Control and Reinforcement Learning, the underlying dynamics of the systems are usually partially or totally unknown. Hence, estimation methods must be devised to learn these models that are often nonlinear. The arising estimation uncertainty however should be taken into consideration in the control design, especially in safety-critical applications, in order to provide safety guarantees during operation.

Subspace identification of unknown *linear* systems regarding control applications has been extensively studied for both non-sparse models [1] and sparse models [2]. Furthermore, the estimation uncertainty has been incorporated in the design of safe controllers using robust control formulations [3]. For *nonlinear* systems, various methods such as maximum likelihood estimation [4] and neural network regression [5] have been introduced in system identification. As before, estimates of the uncertainty can be incorporated in the control process [6].

Gaussian Processes (GPs) have been extensively utilized in the machine learning literature [7] and additionally in system identification [8]. One of the main advantages of GPs is the closed-form formulas for the prediction and the uncertainty that allow for system identification and robust control formulation [9], [10].

Email: dimitri@berkeley.edu and sojoudi@berkeley.edu.

Dimitris Papadimitriou is with the Department of Mechanical Engineering, University of California, Berkeley. Somayeh Sojoudi is with the Departments of Electrical Engineering and Computer Sciences and Mechanical Engineering, University of California, Berkeley.

Student-t Processes (TPs), which constitute a generalization of GPs, have been introduced as an alternative to the latter in stochastic optimization [11], [12]. The main benefit in comparison to GPs is that the estimation variance exclusively depends on the function evaluations of the observed data and not just the features as captured by the chosen kernel function. This attribute renders TPs *1)* more robust to outliers that are common in Control and Reinforcement Learning applications and *2)* suitable for applications where noise is heteroscedastic or more broadly some of its distributional characteristics are unaccounted for. Our main contributions can be summarized as follows. We introduce Student-t Processes in the control framework and provide uniform error bounds for the estimation of unknown nonlinear functions using TPR. Moreover, we incorporate the estimation uncertainty in a control framework. The resulting formulation is a convex optimization problem that can be solved efficiently.

Notation: We denote matrices and vectors with bold upper and lower case letters, respectively. For scalars, we use non-bold lower case letters. The identity matrix of size n is denoted with $\mathbb{I}_n$. We denote the cardinality of a set D with $|D|$. We use $\hat{f}$ for an estimate of the function f. D and D_τ represent a continuous domain and its discretization, respectively. $[z]$ is a point in the discrete domain D_τ closest according to the Euclidean metric to $z \in D$.

II. Problem Formulation

We study systems whose underlying dynamics can be modeled as

$$\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t) \coloneqq \mathbf{h}(\mathbf{x}_t, \mathbf{u}_t) + \mathbf{g}(\mathbf{x}_t, \mathbf{u}_t), \tag{1}$$

where $\mathbf{x}_t \in \mathbb{R}^{n_x}$ and $\mathbf{u}_t \in \mathbb{R}^{n_u}$ denote the state and the control input, respectively. We refer to the state and control input concatenation vector as $\mathbf{z}_t = [\mathbf{x}_t, \mathbf{u}_t] \in \mathbb{R}^{n_x+n_u}$ for compactness. The nonlinear function $\mathbf{h}$ is known, whereas $\mathbf{g}$ is unknown and presumed nonlinear. We assume that we have access to a dataset of size n comprised of tuples of noisy observations of the form $(\mathbf{z}_i, y_i)$ for $i = 1, \dots, n$. Our objectives are *1)* to obtain an estimate $\hat{\mathbf{g}}$ of $\mathbf{g}$, *2)* to quantify the uncertainty when predicting the system output at a new state-input pair using $\hat{\mathbf{g}}$ and *3)* to convert the dynamics functions in a form amenable for computationally efficient control.

III. Model Estimation

In this section, we present the details regarding TPR and we compare the closed form update formulas for the mean and variance with those of GPR.

A. Student-t Process Regression

A Student-t process can be seen as a generalization of Gaussian processes since the Student-t distribution converges to a Gaussian as its degrees of freedom approach infinity. The definition of a Student-t process is given below.

Definition 3.1: A function g forms a Student-t Process $\mathcal{TP}_\nu$ in D with mean $\mu : D \to \mathbb{R}$, kernel function $k : D \times D \to \mathbb{R}$ and ν degrees of freedom if any finite subset of function evaluations $g(\mathbf{z}_1), \dots, g(\mathbf{z}_n)$ with $\mathbf{z}_i \in \mathbb{D}$ follows a joint multi-variate Student-t distribution with mean $\mu := (\mu(\mathbf{z}_1), \dots, \mu(\mathbf{z}_n))$, kernel $K \in \mathbb{R}^{n \times n}$ with $K_{i,j} = k(\mathbf{z}_i, \mathbf{z}_j)$ and ν degrees of freedom.

Given the similarities between TPs and GPs, it is expected that the two share similar properties. In TPR, we assume that the prior distribution of the unknown function follows a zero mean Student-t distribution with ν degrees of freedom

$$g \sim \mathcal{T}_\nu(0, k), \tag{2}$$

and a kernel function k. Disregarding momentarily the measurement noise the joint multi-variate distribution of n measurements and a new query point $\mathbf{z}_*$ is given by

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{T}_\nu \left(\mathbf{0}, \begin{bmatrix} \mathbf{K}(\mathbf{Z}, \mathbf{Z}) & \mathbf{k}(\mathbf{Z}, \mathbf{z}_*) \\ \mathbf{k}(\mathbf{z}_*, \mathbf{Z}) & k(\mathbf{z}_*, \mathbf{z}_*) \end{bmatrix} \right), \tag{3}$$

where $\mathbf{y}$ is a vector of the y_is and $\mathbf{Z}$ is the data matrix of the $\mathbf{z_i}$s.

In contrast with GPR, if we assume that measurement noise follows a Student-t distribution with ν degrees of freedom, the sum of the unknown model and the measurement noise is no longer guaranteed to follow a Student-t distribution [13]. Hence, to incorporate measurement noise, we follow [11] and add the noise variance on the diagonal of the kernel function in (3)

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{T}_\nu \left(\mathbf{0}, \begin{bmatrix} \mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma^2 \mathbb{I}_n & \mathbf{k}(\mathbf{Z}, \mathbf{z}_*) \\ \mathbf{k}(\mathbf{z}_*, \mathbf{Z}) & k(\mathbf{z}_*, \mathbf{z}_*) \end{bmatrix} \right). \tag{4}$$

It should be noted that under this model, the noise term is uncorrelated but not independent from the function g [11]. The conditional distribution on a new point given past observations is given by (see [14])

$$y_* \mid \mathbf{y}, \mathbf{Z}, \mathbf{z}_* \sim \mathcal{T}_{\nu+n}(\mu_n(\mathbf{z}_*), \sigma_n^2(\mathbf{z}_*)), \tag{5}$$

where

$$\mu_n(\mathbf{z}_*) = \mathbf{k}(\mathbf{Z}, \mathbf{z}_*)(\mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma^2 \mathbb{I}_n)^{-1} \mathbf{y}, \tag{6a}$$

$$\sigma_n^2(\mathbf{z}_*) = \phi_n (k(\mathbf{z}_*, \mathbf{z}_*) - \mathbf{k}(\mathbf{Z}, \mathbf{z}_*)^T (\mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma^2 \mathbb{I}_n)^{-1} \mathbf{k}(\mathbf{Z}, \mathbf{z}_*)), \tag{6b}$$

$$\phi_n = \frac{\nu + \mathbf{y} \mathbf{K}(\mathbf{Z}, \mathbf{Z})^{-1} \mathbf{y} - 2}{\nu + n - 2}, \tag{6c}$$

and the subscript n denotes the number of data available to condition on. These update rules differ, from the corresponding GPR rules, only in the ϕ_n term that appears in the posterior variance. This is the key advantage of TPR compared to GPR since the posterior variance of the prediction explicitly depends on the evaluations of the observations $\mathbf{y}$. The quadratic term appearing in ϕ_n quantifies the variability of the measurements. For large values of this term the predictive variance is upscaled. This happens when the measurements vary significantly, increasing the variance and hence leading to more conservative confidence intervals, as will be shown in subsequent sections. On the other hand, if the measurements do not exhibit high variability, the predictive variance is down-scaled leading to "narrower" confidence intervals. Such dependence allows confidence intervals to be more informative in comparison to the ones obtained from GPR.

IV. Student-t Process Regression Motivation

Having access to informative confidence intervals is essential, especially for applications in safety-critical control systems. In this section, we will outline two scenaria under which TPR is more beneficial than GPR. Given the "adaptability" of TPR to varying levels of unaccounted noise, it is natural to expect TPR to perform better in the presence of *outliers* and *heteroscedastic* noise.

A statistical outlier is broadly defined as a sample point that lies abnormally many standard deviations away from the mean of the distribution from which it was most likely sampled [15]. If such outliers appear in the dataset, then we expect the posterior variance of TPR to adapt to their presence via the term (6c) appearing in the variance formula. Such anticipation for possible outliers will lead to a higher variance in the estimates rendering the control laws more conservative which would be beneficial in such a case.

In addition to our original theme about statistical outliers, we will also study the scenario with heteroscedastic noise. In the case of heteroscedastic noise, TPR can also have an advantage over GPR since it can "automatically" adjust for the different noise levels. For instance, if the distribution of noise varies throughout the state space, then using a local version of TPR as described in Algorithm 1 from Section V-A will adjust the variance levels throughout the state space depending on the observed noise levels in the data. GPR methods that adapt to heteroscedastic noise have indeed been proposed in the literature [16], but they generally lead to a higher computational complexity as opposed to TPR whose only additional substantial complexity is the matrix inversion appearing in (6c).

V. Theoretical Analysis

This section contains the uniform error bounds derived for the mean prediction using TPR. In what follows we assume that the unknown function g has a Lipschitz constant L_g. This assumption is needed to derive the following uniform error bound theorem.

Theorem 5.1: Assume that we have access to n noisy measurements. Let $D = [0, r]^d$, D_τ be a discretization of D of size τ^d, $N(D, \tau)$ be the covering number of D, and $\mathcal{K}_n := \|(\mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma^2 \mathbb{I}_n)^{-1}\|_2$. Let L_k be a Lipschitz constant for the kernel function k and let the gamma function

be denoted with $\Gamma(\cdot)$. Then the following inequality holds for the error when using the rule (6a) for prediction

$$\mathbb{P}\left(|g(\mathbf{z}) - \mu_n(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_n(\mathbf{z}) + \gamma_n(\tau),\ \forall\, \mathbf{z} \in D\right) \geq 1 - \delta$$

where

$$\beta_n = \frac{C_{\mathcal{T}} N(D,\tau)}{\delta}, \tag{7a}$$

$$C_{\mathcal{T}} = \frac{2\nu\sqrt{(\nu-2)}\Gamma(\frac{\nu+1}{2})}{(\nu-1)^2\sqrt{\pi}\Gamma(\frac{\nu}{2})}, \tag{7b}$$

$$\gamma_n(\tau) = (L_g + L_k\sqrt{n}\|\alpha\|_2)\tau + \sqrt[3]{\beta_n}\sqrt{2|\phi_n|nL_k\mathcal{K}_n \max_{\mathbf{z},\mathbf{z}'\in D} k(\mathbf{z},\mathbf{z}')}, \tag{7c}$$

$$\alpha = (\mathbf{K}(\mathbf{Z},\mathbf{Z}) + \sigma^2\mathbb{I}_n)^{-1}\mathbf{y}. \tag{7d}$$

To prove Theorem 5.1 we will first develop two useful Lemmas.

Lemma 5.2: A random variable T following a Student-t distribution with mean 0, variance $(\nu-2)/\nu$ and $\nu > 2$ degrees of freedom satisfies the upper tail bound

$$\mathbb{P}(T > x) \leq \frac{C_{\mathcal{T}}}{x^3},$$

with $C_{\mathcal{T}} \coloneqq \frac{2\nu\sqrt{(\nu-2)}\Gamma(\frac{\nu+1}{2})}{(\nu-1)^2\sqrt{\pi}\Gamma(\frac{\nu}{2})}$.

Proof: Let $c \coloneqq \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu-2)\pi}\Gamma(\frac{\nu}{2})}$. One can write

$$\begin{aligned}
\mathbb{P}(T > x) &= \int_x^\infty \frac{c}{(1+\frac{t^2}{\nu-2})^{(\nu+1)/2}}dt \\
&\leq \int_x^\infty \frac{t}{x}\frac{c}{(1+\frac{t^2}{\nu-2})^{(\nu+1)/2}}dt \\
&= \frac{\nu}{\nu-1}\frac{1}{x}\frac{c}{(1+\frac{x^2}{\nu-2})^{(\nu-1)/2}} \\
&\leq \frac{\nu}{\nu-1}\frac{1}{x}\frac{c}{(1+(\nu-1)x^2/(2(\nu-2)))} \\
&\leq \frac{2c\nu(\nu-2)}{(\nu-1)^2}\frac{1}{x^3},
\end{aligned}$$

where in the first inequality we used the fact that $t \geq x$ and in the second one we used Bernoulli's inequality. ∎

It is desirable to develop a bound for the deviation of g from the posterior mean (6a) in a discrete domain space D_τ. Afterwards, we will extend the result to a general compact space $D \subset \mathbb{R}^{n_x+n_u}$.

Lemma 5.3: For the observed data points $(\mathbf{z}_1,\ldots,\mathbf{z}_n)$ and $(\mathbf{y}_1,\ldots,\mathbf{y}_n)$, we select $\delta \in (0,1)$ and let $\beta_n = \frac{C_{\mathcal{T}}|D_\tau|}{\delta}$. For every $\mathbf{z} \in D_\tau$ and $n \geq 1$, where $D_\tau \subset \mathbb{R}^{n_x+n_u}$ is a discrete set, the inequality

$$|g(\mathbf{z}) - \mu_n(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_n(\mathbf{z}),$$

holds with probability at least $1-\delta$.

Proof: For every $\mathbf{z} \in D_\tau$, it follows from Lemma 5.2 that

$$\mathbb{P}(|g(\mathbf{z}) - \mu_n(\mathbf{z})| > \sqrt[3]{\beta_n}\sigma_n(\mathbf{z})) \leq \frac{C_{\mathcal{T}}}{\beta_n}.$$

From the union bound, we conlcude that $\forall \mathbf{z}$

$$\mathbb{P}(|g(\mathbf{z}) - \mu_n(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_n(\mathbf{z})) \geq 1 - \delta.$$

∎

Proof of Theorem 5.1: We generalize Lemma 5.3 to a continuous space $D \subset \mathbb{R}^{n_x+n_u}$. We proceed by bounding the difference in the unknown function g (2), the posterior mean (6a) and variance (6b) update rules between two points $\mathbf{z} \in D$ and $\mathbf{z}' \in D_\tau$ in the continuous and discrete domains, respectively. D_τ is a discretization of D such that

$$\max_{\mathbf{z}\in D}\|\mathbf{z} - [\mathbf{z}]\|_2 \leq \tau. \tag{8}$$

For the smallest discretization for which this bound holds, we obtain $|D_\tau| = N(D,\tau)$ where N denotes the covering number. Starting with g, and using its Lipschitz assumption along with (8) we obtain

$$|g(\mathbf{z}) - g(\mathbf{z}')| \leq L_g\|\mathbf{z} - \mathbf{z}'\|_2 \leq L_g\tau. \tag{9}$$

Assuming that the kernel function has a Lipschitz constant L_k, the mean difference between $\mathbf{z} \in D$ and its projection $[\mathbf{z}]$ on D_τ can be bounded as

$$\begin{aligned}
|\mu_n(\mathbf{z}) - \mu_n([\mathbf{z}])| &\leq |k(\mathbf{Z},\mathbf{z})\alpha - k(\mathbf{Z},[\mathbf{z}])\alpha| \\
&\leq L_k\sqrt{n}\|\alpha\|_2\|\mathbf{z} - [\mathbf{z}]\|_2 \\
&\leq L_k\sqrt{n}\|\alpha\|_2\tau,
\end{aligned} \tag{10}$$

by using Cauchy-Schwarz inequality and setting $\alpha = (\mathbf{K}(\mathbf{Z},\mathbf{Z}) + \sigma^2\mathbb{I}_n)^{-1}\mathbf{y}$. We can bound the variance term as follows

$$\begin{aligned}
&|\sigma_n^2(\mathbf{z}) - \sigma_n^2([\mathbf{z}])| \\
&= |\phi_n k(\mathbf{Z},[\mathbf{z}])^T(\mathbf{K}(\mathbf{Z},\mathbf{Z}) + \sigma^2\mathbb{I}_n)^{-1} \\
&\times k(\mathbf{Z},[\mathbf{z}]) - \phi_n k(\mathbf{Z},\mathbf{z})^T(\mathbf{K}(\mathbf{Z},\mathbf{Z}) + \sigma^2\mathbb{I}_n)^{-1}k(\mathbf{Z},\mathbf{z})| \\
&\leq |\phi_n|\|k(\mathbf{Z},[\mathbf{z}]) + k(\mathbf{Z},\mathbf{z})\|_2\|(\mathbf{K}(\mathbf{Z},\mathbf{Z}) + \sigma^2\mathbb{I}_n)^{-1}\|_2 \\
&\times \|k(\mathbf{Z},[\mathbf{z}]) - k(\mathbf{Z},\mathbf{z})\|_2.
\end{aligned}$$

To bound the individual terms, we use the inequalities $\|k(\mathbf{Z},[\mathbf{z}]) + k(\mathbf{Z},\mathbf{z})\|_2 \leq 2\sqrt{n}\max_{\mathbf{z},\mathbf{z}'\in D} k(\mathbf{z},\mathbf{z}')$ and $\|k(\mathbf{Z},[\mathbf{z}]) - k(\mathbf{Z},\mathbf{z})\|_2 \leq L_k\sqrt{n}\|[\mathbf{z}] - \mathbf{z}'\|_2 \leq L_k\sqrt{n}\tau$. Hence, by letting $\mathcal{K}_n \coloneqq \|(\mathbf{K}(\mathbf{Z},\mathbf{Z}) + \sigma^2\mathbb{I}_n)^{-1}\|_2$ we arrive at

$$|\sigma_n^2(\mathbf{z}) - \sigma_n^2([\mathbf{z}])| \leq 2|\phi_n|nL_k\mathcal{K}_n\tau \max_{\mathbf{z},\mathbf{z}'\in D} k(\mathbf{z},\mathbf{z}').$$

Using the fact that $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x-y|}$, we finally obtain that

$$|\sigma_n(\mathbf{z}) - \sigma_n([\mathbf{z}])| \leq \sqrt{2|\phi_n|nL_k\mathcal{K}_n\tau \max_{\mathbf{z},\mathbf{z}'\in D} k(\mathbf{z},\mathbf{z}')}. \tag{11}$$

Using Lemma 5.3 along with (9), (10) and (11) we obtain that

$$\mathbb{P}(|g(\mathbf{z}) - \mu_n(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_n(\mathbf{z}) + \gamma_n(\tau),\ \forall\, \mathbf{z} \in D) \geq 1 - \delta$$

where

$$\beta_n = \frac{C_{\mathcal{T}} N(D,\tau)}{\delta}$$

and

$$\gamma_n(\tau) = (L_g + L_k\sqrt{n}\|\alpha\|_2)\tau + \sqrt[3]{\beta_n}\sqrt{2|\phi_n|nL_k\mathcal{K}_n\tau \max_{\mathbf{z},\mathbf{z}'\in D} k(\mathbf{z},\mathbf{z}')}.$$

■

Theorem 5.1 allows us to quantify the uncertainty at a new query point $\mathbf{z}^*$. As someone would expect though from the heavier tails of the Student-t distribution, this bound is considerably more loose than similar bounds derived for GPR [10]. In the next section, we will present an extension of TPR that uses local information in the domain to reduce computational complexity. This variant of TPR can locally adapt to the presence of outliers or more generally, to varying levels of noise.

A. Local Student-t Regression

The calculation of the posterior of TP suffers form the same computational burden as that of GPR. The main bottleneck consists of the inversion of the Kernel function $\mathbf{K}(\mathbf{Z},\mathbf{Z})$, which requires $O(n^3)$ operations. For applications in which data collection happens once before the estimation, the kernel function can be inverted only one time alleviating the computational cost. If, on the other hand data collection takes place online, a new inversion of an increasingly higher dimensional kernel matrix $\mathbf{K}(\mathbf{Z},\mathbf{Z})$ is needed whenever the dataset grows rendering the entire process intractable. This is common especially for applications in dynamic decision making.

Implementation of local estimation methods using GPR has been studied in the literature [17]. We propose the use of local TPR in which only observations in the neighborhood of a query point are taken into consideration in the estimation. More specifically, for each prediction, only the m observations (usually $n \gg m$) closest to the query point are considered, reducing the computational complexity of inverting the kernel matrix to $O(m^3)$. Our proposed method is a simplification over the method in [17] that uses different models for different parts of the domain and constructs a weighted average prediction from those individual models.

Algorithm 1: Local Student-t Process Regression

Given query point $\mathbf{z}_* \in \mathbb{R}^{n_x+n_u}$, $\mathbf{Z}$ and m
Select m closest points from $\mathbf{Z}$ to $\mathbf{z}_*$ using ℓ_2 metric
Construct $\tilde{\mathbf{Z}} \in \mathbb{R}^{n_x+n_u\times m}$, $\tilde{\mathbf{y}} \in \mathbf{R}^m$ and compute:
$\phi_m = \frac{\nu+\tilde{\mathbf{y}}\mathbf{K}(\tilde{\mathbf{Z}},\tilde{\mathbf{Z}})^{-1}\tilde{\mathbf{y}}-2}{\nu+m-2}$
$\mu_m(\mathbf{z}_*) = \mathbf{k}(\tilde{\mathbf{Z}},\mathbf{z}_*)(\mathbf{K}(\tilde{\mathbf{Z}},\tilde{\mathbf{Z}}) + \sigma^2\mathbb{I}_m)^{-1}\mathbf{y}$
$\sigma_m(\mathbf{z}_*) = \phi_m(k(\mathbf{z}_*,\mathbf{z}_*) - \mathbf{k}(\tilde{\mathbf{Z}},\mathbf{z}_*)^T(\mathbf{K}(\tilde{\mathbf{Z}},\tilde{\mathbf{Z}}) + \sigma^2\mathbb{I}_m)^{-1}\mathbf{k}(\tilde{\mathbf{Z}},\mathbf{z}_*))$
Return $\mu_m(\mathbf{z}_*)$, $\sigma_m(\mathbf{z}_*)$

The local TPR variant described in Algorithm 1 has a number of advantages. If outliers appear only in a certain region of the state space, then local TPR would predict conservative confidence intervals only around that region and not throughout the state space. On the contrary, global TPR would return confidence intervals that are scaled up in the entire state space, which would lead to unnecessary conservatism in some parts of the state space.

Furthermore, as mentioned in the motivation section for TPR, homoscedastic noise can be an unrealistic assumption for many real-world applications. As such, a number of approaches that deal with heteroscedastic noise have been proposed in the literature, which for instance are based on maximum *a posteriori* estimation [16] or expectation propagation [18]. Local TPR can be used as another approach to deal with heteroscedastic noise with no significant additional computational complexity.

VI. Model Linearization and Uncertainty Propagation

In this work, we let both the known and unknown dynamics functions be nonlinear. To obtain a formulation conducive for control applications, a linearization of the dynamics functions is helpful. This model linearization leads to an additional error that should be accounted for in the control process.

This section quantifies the error of the linear approximation of the nonlinear system. Regarding model dynamics, the control algorithms implemented in this work require *1)* an estimate of the dynamics of the system over a horizon H, *2)* linearization of the nonlinear estimates and *3)* uncertainty regions in which the actual model is expected to be due to linearization and estimation errors. The linearization process outlined here is similar to the one in [9].

A. Estimation Uncertainty and Linearization Error

Before generalizing to multiple output dynamics we will focus on single output functions. Using TPR to estimate $g(\mathbf{z}_t)$ for a particular state-input pair $\mathbf{z}_t$, the dynamics function (1) now becomes

$$x_{t+1} = \hat{f}(\mathbf{z}_t) \coloneqq h(\mathbf{z}_t) + \hat{g}(\mathbf{z}_t), \tag{12}$$

where we substitute in the posterior mean $\hat{g}(\mathbf{z}) \coloneqq \mu_n(\mathbf{z})$ as the estimate for $g(\mathbf{z})$ at a point $\mathbf{z} \in \mathbb{R}^{(n_x+n_u)}$. This introduces uncertainty in the system due to the variance of the estimator (6b). Furthermore, (12) is a nonlinear function. In order to linearize it we assume that the gradient of the known function h has a Lipschitz constant L_{∇_h}, whereas the gradient of the posterior mean (6a) and the posterior variance (6b) have Lipschitz constants $L_{\nabla_{\mu_n}}$ and L_{σ_n}, respectively. In order to linearize the dynamics around a point $\mathbf{z}' \in \mathbb{R}^{(n_x+n_u)}$, we perform Taylor expansion of the known function h and the estimated function $\hat{g}$ around $\mathbf{z}'$. The linearized estimated function, denoted with $\tilde{f}$, is then given by

$$\tilde{f}(\mathbf{z}) \coloneqq h(\mathbf{z}') + \mathbf{J}_h(\mathbf{z}-\mathbf{z}') + \hat{g}(\mathbf{z}') + \mathbf{J}_{\hat{g}}(\mathbf{z}-\mathbf{z}'), \tag{13}$$

where $\mathbf{J}_h = [\mathbf{J}_{h,x}, \mathbf{J}_{h,u}] \in \mathbb{R}^{n_x\times(n_x+n_u)}$ and $\mathbf{J}_{\hat{g}} = [\mathbf{J}_{\hat{g},x}, \mathbf{J}_{\hat{g},u}] \in \mathbb{R}^{n_x\times(n_x+n_u)}$ are the Jacobians of h and $\hat{g}$, respectively, evaluated at $\mathbf{z}'_t$. We can bound the model

approximation error using the Lagrangian remainder theorem as

$$|\tilde{f}(\mathbf{z}) - \hat{f}(\mathbf{z})| \leq \frac{1}{2}(L_{\nabla_h} + L_{\nabla_{\mu_n}})\|\mathbf{z} - \mathbf{z}'\|_2^2. \quad (14)$$

Using the bound from Theorem 5.1, we can bound the output of $\hat{f}$ from the original dynamics function f by using a Lipschitz continuity argument for σ as follows

$$|f(\mathbf{z}) - \tilde{f}(\mathbf{z})| \leq \frac{1}{2}(L_{\nabla_h} + L_{\nabla_{\mu_n}})\|\mathbf{z} - \mathbf{z}'\|_2^2 + \sqrt[3]{\beta_n}(\sigma_n(\mathbf{z}) + L_{\sigma_n}\|\mathbf{z} - \mathbf{z}'\|_2) + \gamma_n(\tau). \quad (15)$$

For $\mathbf{z} \in \mathcal{Z}$, where $\mathcal{Z}$ denotes a generic state-action space, the error in (15) can be upper-bounded by

$$|f(\mathbf{z}) - \tilde{f}(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_n(\mathbf{z}) + \left(\frac{L_{\nabla_h} + L_{\nabla_{\mu_n}}}{2}\right) d_{max}(\mathbf{z}')^2 + \sqrt[3]{\beta_n}L_{\sigma_n} d_{max}(\mathbf{z}') + \gamma_n(\tau), \quad (16)$$

where $d_{max}(\mathbf{z}') = \sup_{\mathbf{z}\in\mathcal{Z}}\|\mathbf{z} - \mathbf{z}'\|_2$. Clearly, since $\mathbf{z} = [\mathbf{x}, \mathbf{u}]$, we can decompose $d_{max}(\mathbf{z}')$ as $d_{max}(\mathbf{x}') + d_{max}(\mathbf{u}')$. The Lipschitz constants appearing in (16) can be estimated since all the functions involved are known.

B. Uncertainty Propagation

The above subsection quantified the uncertainty due to model estimation and linearization. Given a state $\mathbf{z}_t$, by applying $\tilde{f}$ each coordinate of the system transitions inside the region specified by

$$f_i(\mathbf{z}) \in \tilde{f}_i(\mathbf{z}) \pm \sqrt[3]{\beta_n}\sigma_n(\mathbf{z}) + \left(\frac{L_{\nabla_h} + L_{\nabla_{\mu_n}}}{2}\right) d_{max}(\mathbf{z}')^2 + \sqrt[3]{\beta_n}L_{\sigma_n} d_{max}(\mathbf{z}') + \gamma_n(\tau), \quad (17)$$

where $i = 1, \ldots, n_x$. In the general case of a multi-output function f, we model each coordinate with an independent TP. The above uncertainty set for coordinate i is an interval in $\mathbb{R}$.

Lemma 6.1: Let $n_x > 1$ and each component $g_i(\mathbf{z}), i = 1, \ldots, n_x$ of $\boldsymbol{g}(\mathbf{z})$ be modeled with an independent Student-t process. The error of the multi-output function $\boldsymbol{g}(\mathbf{z})$ with posterior mean vector $\boldsymbol{\mu}_n \coloneqq [\mu_{n,1}(\mathbf{z}), \ldots, \mu_{n,n_x}(\mathbf{z})]^T$ and posterior variances $[\sigma_{n,1}(\mathbf{z}), \ldots, \sigma_{n,n_x}(\mathbf{z})]$ can be upper bounded by

$$\mathbb{P}\left(|\boldsymbol{g}(\mathbf{z}) - \boldsymbol{\mu}_n(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_n(\mathbf{z}) + \gamma_n(\tau),\ \forall\, \mathbf{z} \in D\right) \geq (1-\delta)^{n_x}. \quad (18)$$

Proof: The result follows from Theorem 5.1 and the independence of the Student-t processes for each coordinate by considering the intersection of the following events

$$\mathbb{P}(|\boldsymbol{g}(\mathbf{z}) - \boldsymbol{\mu}_n(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_n(\mathbf{z}) + \gamma_n(\tau),\ \forall\, \mathbf{z} \in D) = \prod_{i=1}^{n_x} \mathbb{P}(|\boldsymbol{g}_i(\mathbf{z}) - \boldsymbol{\mu}_{n,i}(\mathbf{z})| \leq \sqrt[3]{\beta_n}\sigma_{n,i}(\mathbf{z}) + \gamma_n(\tau),\ \forall\, \mathbf{z} \in D) \geq (1-\delta)^{n_x}.$$

■

Generalizing simultaneously to all coordinates, we obtain an uncertainty set that now belongs to $\mathbb{R}^{n_x}$. We can rewrite (17) as

$$\mathbf{f}(\mathbf{z}) \in \tilde{\mathbf{f}}(\mathbf{z}) \oplus \mathcal{E}(\mathbf{z}) \subset \mathbb{R}^{n_x}, \quad (19)$$

where now the set $\mathcal{E}(\mathbf{z})$ contains the deviation from the nominal system $\tilde{\mathbf{f}}(\mathbf{z})$ in each coordinate. For finite-horizon control tasks, we are interested in the propagation of the uncertainty sets throughout the horizon. In this subsection, we study the propagation of uncertainty when the estimated dynamics function is applied repeatedly to obtain a trajectory of horizon H. This is key for the MPC formulation as the uncertainty of the dynamics should be taken into account at every time step of the MPC horizon H. At each time step $t = 1, \ldots, H$, we recursively define the uncertainty set at time t as

$$\mathcal{E}_t = \tilde{\mathbf{f}}(\mathcal{E}_{t-1}) \oplus \mathcal{E}(\mathcal{E}_{t-1}), \quad (20)$$

with $\mathcal{E}_0$ containing only the current state of the system.

The expression in (20) allows us to propagate the system uncertainty to the end of the task horizon. Assume that the system starts at the known state x_0 and that we aim to linearize the system in (12) over a horizon H around a trajectory $\mathbf{x}'_0, \mathbf{u}'_0, \mathbf{x}'_1, \mathbf{u}'_1, \ldots, \mathbf{x}'_{H-1}, \mathbf{u}'_{H-1}$. Using (20), we can propagate the uncertainty sets until the end of the horizon H. Initially, $\mathcal{E}_0 = [\mathbf{x}_0] \coloneqq [\mathbf{x}'_0]$ as there is no uncertainty about the starting state of the system. The only unknown is the control input $\mathbf{u}_0$, which we assume is constrained within a set $\mathcal{U}_0$ that is centered around the linearization point $\mathbf{u}'_0$. The same assumption is made for all $\mathbf{u}_i$, $\mathcal{U}_i, i = 1, \ldots, H-1$. Hence, the confidence set $\mathcal{E}_1$ in which we can expect the system to be at the next time step with high probability is given by (17) for which $d_{max}(\mathbf{z}'_\mathbf{0}) = \sup_{\mathbf{u}\in\mathcal{U}_0}\|\mathbf{u} - \mathbf{u}'_0\|_2^2$ with $\mathbf{u}'_0 = \mathbf{0}$. At the next time step, the fact that the linearization state $\mathbf{x}'_1$ is fixed allows us to compute $d_{max}(\mathbf{z}'_1) = d_{max}(\mathbf{x}'_1) + d_{max}(\mathbf{u}'_1)$, which is needed to propagate the confidence set $\mathcal{E}_1$ to the next time step $\mathcal{E}_2$. For the maximum distance from the control linearization point, the same logic applies as before, $d_{max}(\mathbf{u}'_1) = \sup_{\mathbf{u}\in\mathcal{U}_1}\|\mathbf{u} - \mathbf{u}'_1\|_2^2$. To compute $d_{max}(\mathbf{x}'_1)$, we constrain the system to lie within a set $\mathcal{X}_1$ that is centered around $\mathbf{x}'_1$. The choice of that set is made by balancing the conservativeness and the feasibility of the controller. The same process is repeated throughout the horizon H using each time $\mathbf{u}'_i = \mathbf{0}$. The only thing left is to define the sequence of states $\mathbf{x}'_1, \ldots, \mathbf{x}'_{H-1}$ on which the system will be linearized. For that we propose using the time-shifted open-loop trajectory of the previous time step solution of a receding horizon control problem (21).

Proposition 6.2: Let H denote the horizon of a trajectory and suppose that the assumptions of Theorem 5.1 hold. By using (20) to propagate the uncertainty sets along a trajectory $x_1, \ldots, x_H$ obtained by applying the inputs $u_1, \ldots, u_{H-1}$, the realized state of the system $\mathbf{x}_t$ lies inside $\mathcal{E}_t$ for all $t = 1, \ldots, H$ with probability at least $(1-\delta)^{n_x}$.

Proof: This follows from the uniform error bound in Lemma 6.1 and the propagation rule (20). ■

C. Control Policy

We can now combine and incorporate the estimated dynamics from Section III and the linearization and uncertainty propagation from Section VI to design a control policy that completes a specific task with high probability. We will synthesize such a control policy π using an MPC formulation of horizon H_t for a task with duration N. More specifically, at time step t we solve the following convex optimization problem

$$
\begin{aligned}
&J^*(\mathbf{x}_t, H_t) = \\
&\min_{\mathbf{u}_{t|t},\dots,\mathbf{u}_{t+H_t-1|t}} \quad J(\bar{\mathbf{x}}_{t|t},\dots,\bar{\mathbf{x}}_{t+H_t|t},\mathbf{u}_{t|t},\dots,\mathbf{u}_{t+H_t-1|t}) \\
&\text{s.t.} \qquad \bar{\mathbf{x}}_{t|t} - \mathbf{x}_t \in \mathcal{E}_{t|t}, \mathbf{u}_{k|t} \in \mathcal{U}_{k|t} && \text{(21a)} \\
&\qquad\quad \bar{\mathbf{x}}_{k+1|t} = \tilde{\mathbf{f}}_{k|t}(\bar{\mathbf{x}}_{k|t}, \mathbf{u}_{k|t}) && \text{(21b)} \\
&\qquad\quad \bar{\mathbf{x}}_{k|t} \in \mathcal{X}_{k|t} \ominus \mathcal{E}_{k|t} && \text{(21c)} \\
&\qquad\quad \forall k = t,\dots,t+H_t-1 && \text{(21d)} \\
&\qquad\quad \bar{\mathbf{x}}_{t+H_t|t} \in \mathcal{O} \ominus \mathcal{E}_{t+H_t|t}, && \text{(21e)}
\end{aligned}
$$

where the subscript $t+k|t$ denotes the kth state when the input sequence $[u_{t|t},\dots,u_{t+k|t}]$ is applied on system (21b). The objective J is a convex cost function associated with a particular task, $\mathcal{E}_{t|t} = [\mathbf{x}_t]$, $\mathcal{X}_{k|t}$ and $\mathcal{U}_{k|t}$, $k = t+1,\dots,t+H_t-1$ denote the state and input constraint sets and the linearized estimation of the dynamics function $\tilde{\mathbf{f}}_{k|t}$ is given by (13). For each k this function is linearized around the time-shifted open-loop trajectory obtained at time $t-1$. We denote with $\bar{\mathbf{x}}$ the nominal state of the system that excludes the disturbance. For generality, we allow the state $\mathcal{X}_{k|t}$ and input $\mathcal{U}_{k|t}$ constraint sets to be time dependent. Given the solution to (21), the MPC policy at time step t is

$$\mathbf{u}_t = \pi(\mathbf{x}_t) = \mathbf{u}^*_{t|t}, \tag{22}$$

which is used to propagate the system to the next state. In the same manner, at the next time step $t+1$ we obtain a policy $\pi(\mathbf{x}_{t+1})$ by solving (21) again for the new measured state $\mathbf{x}_{t+1}$. The whole process is repeated until the goal set $\mathcal{O}$ is reached.

The optimization problem (21) is a convex optimization problem that can be solved efficiently. However, the problem may become infeasible at some time step, if for instance constraint (21c) results in an empty set. In this case, we employ a shrinking horizon strategy, outlined in Algorithm 2, that drives the system towards the goal set $\mathcal{O}$ with high probability.

To obtain an optimal policy for the system, we need to assume the existence of an optimal trajectory at the beginning of the task. This is required since the problem (21) can be infeasible at the first time step otherwise.

Assumption 6.3: We are initially given a feasible trajectory that can drive the system from a starting state $\mathbf{x}_s$ to the goal set $\mathcal{O}$.

Given this assumption, Algorithm 2 will drive the system to the terminal set $\mathcal{O}$ with high probability.

Theorem 6.4: Given β_n and γ_n as in (7a) and (7c), the control policy π in (22) obtained from Algorithm 2 drives the system to the terminal set $\mathcal{O}$ with high probability.

Algorithm 2: Shrinking Horizon Control

Given $\mathbf{x}_t$, H_t, optimal trajectory at the previous time step $\bar{\mathbf{x}}^*_{k|t-1},\ k = t,\dots,t+H_t$
if $J^*_t(\mathbf{x}_t, H_t)$ *feasible* **then**
  Set $H_{t+1} = H_t$
  Let $\mathbf{u}^*_{t|t},\dots,\mathbf{u}^*_{t+H_t-1} = \arg\min J^*_t(\mathbf{x}_t, H_t)$
else
  Set $\mathcal{E}_{k|t} = \mathcal{E}_{k|t-1}$, $\mathcal{X}_{k|t} = \mathcal{X}_{k|t-1}$, $\tilde{\mathbf{f}}_{k|t} = \tilde{\mathbf{f}}_{k|t-1} \forall k$
  Solve $J^*_t(\mathbf{x}_t, H_t - 1)$
  Set $H_{t+1} = H_t - 1$
  Let $\mathbf{u}^*_{t|t},\dots,\mathbf{u}^*_{t+H_t-2} = \arg\min J^*_t(\mathbf{x}_t, H_t - 1)$
end
Return H_{t+1}, $\mathbf{u}^*_{k|t},\ k = t,\dots,t+H_t-1$

Proof: The theorem can be proved by induction. Assume that at time t the finite-time optimal control problem solved by Algorithm 2 is feasible and let $[\mathbf{u}^*_{t|t},\dots,\mathbf{u}^*_{t+H_t-1|t}]$ and $[\bar{\mathbf{x}}^*_{t|t},\dots,\bar{\mathbf{x}}^*_{t+H_t|t}]$ be the optimal input and nominal state sequences. Notice that if

$$\mathbf{x}_{t+1} - \bar{\mathbf{x}}^*_{t+1|t} \in \mathcal{E}_{1|t}, \tag{23}$$

then at the next time step $t+1$, we have that the shifted input and nominal state sequences $[\mathbf{u}^*_{t+1|t},\dots,\mathbf{u}^*_{t+H_t-1|t}]$ and $[\bar{\mathbf{x}}^*_{t+1|t},\dots,\bar{\mathbf{x}}^*_{t+H_t|t}]$ are feasible for problem (21) with the prediction horizon H_t-1 and for $\mathcal{E}_{k|t+1} = \mathcal{E}_{k|t}$, $\mathcal{X}_{k|t+1} = \mathcal{X}_{k|t}$ and $\tilde{\mathbf{f}}_{k|t+1} = \tilde{\mathbf{f}}_{k|t}$, for $k = t+1,\dots,t+H_t-1$. From Proposition 6.2, we have that (23) holds with probability $(1-\delta)^{n_x}$, and therefore problem $J^*_t(\mathbf{x}_{t+1}, H_t - 1)$ from Algorithm 2 is feasible with probability $(1-\delta)^{n_x}$. Then at time $t+1$, Algorithm 2 returns a feasible sequence of input actions with probability $(1-\delta)^{n_x}$.

So far we have shown that if at time t Algorithm 2 returns a feasible sequence of inputs, then at time $t+1$ Algorithm 2 is feasible with high probability. By assumption, we have that at time $t = 0$ Problem (21) is feasible for $H_0 = H$. Therefore, we conclude by induction that for a control task of length N, Algorithm 2 returns a feasible control sequence for all $t \in \{0, N\}$ with high probability. ■

VII. Simulations

In this section, we provide simulations that showcase circumstances under which using TPR is preferred to GPR. Although this work focuses on measurements polluted with outliers, we will further provide results for the case of heteroscedastic measurement noise. To evaluate the quality of the confidence intervals between the two methods, we offer a simple control simulation for a one-dimensional problem. More specifically, the system dynamics function is

$$x_{t+1} = g(x_t, u_t) = \sqrt[3]{x_t} + u_t,$$

where $x_t, u_t \in \mathbb{R}$ and the known model h is zero. We collect a dataset of noisy transitions of size $n = 200$ by applying uniformly random control inputs at randomly chosen states. The goal of the task is to steer the system from the starting state $x_s = 5$ to goal state $x_f = -1$ by

solving (21) with a horizon $H_0 = 4$ and $N = 6$. We will focus in showcasing the accuracy of the confidence intervals and not the performance of Algorithm 2. Thus we will be removing constraint (21c) from the optimization problem. The control input and state constraint sets are $\mathcal{U}_{k|t} = [-1, 1]$ and $\mathcal{X}_{k|t} = [\bar{x}_{k|t-1} - 1, \bar{x}_{k|t-1} + 1], k = t, \dots, t + H_t - 1$, while $\mathcal{O} = [\bar{x}_f - 1, \bar{x}_f + 1]$. The cost function is a quadratic function with $Q = 1$ and $R = 5$. Theoretical estimation of the quantities appearing in the uniform error bound in Theorem 5.1 will lead to an excessive overestimation of the confidence sets. For that reason, and in order to have a more direct comparison to GPR, we follow [19] in which the authors set $\beta_n = 2$ and we set $\sqrt[3]{\beta_n} = 2$ and $\gamma_n(\tau) = 0$. The kernel function k is the squared exponential kernel. The Lipschitz constants are set to $L_{\nabla_\mu} = L_\sigma = 0.1$ and the degrees of freedom are $\nu = 5$. The comparison between TPR and GPR is done in two ways. First, by providing plots of the confidence sets $\mathcal{E}_{2|1}, \mathcal{E}_{3|2}, \dots$ to visually inspect the uncertainty sets. Second, for each of the cases studied below we carry out 10 simulations of the task, each time with a newly generated dataset, and measure the number of times the uncertainty sets failed to include the actual system $x_{t+1|t} \notin \mathcal{E}_{t+1|t}$ for $t = 1, \dots, N$.

A. Outliers

In the case of outliers, we assume that the system can potentially experience abnormal noise levels that lie in the excess of 3 standard deviations away from the unpolluted data. In this scenario we restrict the presence of outlier data points only in the region $[-1, 1]$ to further emphasize the importance of local models. The unpolluted observations are assumed to have an observation noise that follows a normal distribution $w_t \sim \mathcal{N}(0, 1)$. Although outliers are relatively infrequent in datasets, we pollute our datasets with up to 20% outlying data points for exposition purposes.

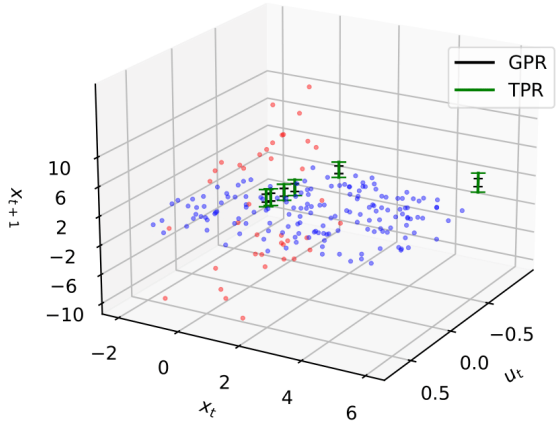

Fig. 1: Confidence intervals $\mathcal{E}_{t+1|t}$ using *global* models. Blue and red dots designate the ordinary data and the outliers, respectively.

In Figures 1 and 2, we plot the uncertainty sets $\mathcal{E}_{2|1}, \mathcal{E}_{3|2}, \dots$ that we obtain from the closed-loop solution for the cases of global and local estimations respectively. Note that the closed loop trajectories coincide, as the update rule (6a) is the same for both methods. The difference between the two methods lies in the confidence sets plotted with the error bars. These examples clearly demonstrate the benefits of TPR in comparison to GPR. In the global estimation case GPR returns uncertainty sets that are relatively uniform along the domain as the outliers do not affect the posterior variance. On the other hand, the uncertainty sets in TPR are scaled up due to the magnitude of the outliers. However, if outliers are present only in certain regions of the state space, then TPR will tend to predict scaled up uncertainty sets even for the regions that do not exhibit outliers.

To alleviate this issue we propose using the local variant of TPR with $m = 10$ so that local uncertainty cannot mitigate throughout the state space. Figure 2 shows that local GPR still fails to capture the higher variation between $[-1, 1]$ but local TPR exhibits scaled-up uncertainty sets in that region, in comparison to the rest of the state space. Table I demonstrates the fail Rates (FR), i.e. the times for which $x_{t+1|t} \notin \mathcal{E}_{t+1|t}$ for $t = 1, \dots, N$, averaged over 10 simulations.

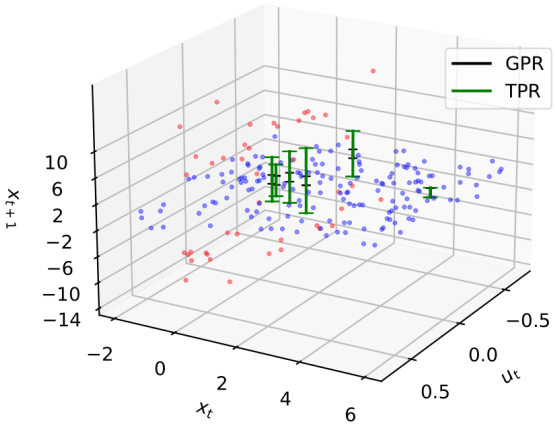

Fig. 2: Confidence intervals $\mathcal{E}_{t+1|t}$ using *local* models. Blue and red dots designate the ordinary data and the outliers, respectively.

B. Heteroscedasticity

This section studies the potential benefits of using TPR in the case of heteroscedastic noise in the state space. We argue that the local variant of TPR can be used in the case of heteroscedastic noise in a computationally efficient manner. The plots in Figures 3 and 4 show how the confidence sets $\mathcal{E}_{t+1|t}$ adapt to the local variability of the noise when using TPR. In these simulations, we used three different Gaussian noise levels with $\sigma = 4, 1$ and 0.1 in three distinct regions of the domain $[-1, 1], [1, 3]$ and $[3, 6]$, while the models assume that the actual noise of the data is white Gaussian. In the global estimation scheme, TPR uniformly scales up the confidence sets to compensate for the increased variability of the blue data points, whereas the local version scales up

or down the confidence sets depending on the high or low variability of the region.

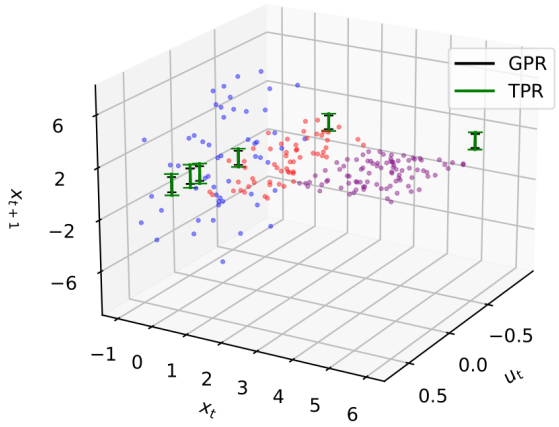

Fig. 3: Confidence intervals $\mathcal{E}_{t+1|t}$ using *global* models. Heteroscedastic noise with $\sigma = 0.01$ (purple), $\sigma = 1$ (red), and $\sigma = 4$ (blue).

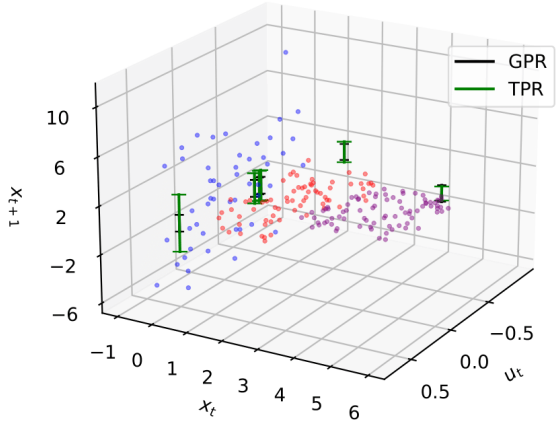

Fig. 4: Confidence intervals $\mathcal{E}_{t+1|t}$ using *local* models. Heteroscedastic noise with $\sigma = 0.01$ (purple), $\sigma = 1$ (red), and $\sigma = 4$ (blue).

		GPR	TPR	GPR (local)	TPR (local)
Outlier	FR (%):	55.7	**19.7**	60.7	**1.7**
Heteroscedastic	FR (%):	39.5	**8.1**	29.5	**9.6**

TABLE I: Percentage of times $x_{t+1|t} \notin \mathcal{E}_{t+1|t}$, $\forall t$ for GPR and TPR estimation averaged over 10 simulations.

VIII. Conclusion

We studied the benefits of using the Student-t Process Regression framework in control settings. The posterior variance of TPR can capture the additional variability due to statistical outliers, which are common in robotics applications, and thus it can be used to design control policies that are more robust to outliers. We proved uniform error bounds for the prediction error when using TPR, and then presented a method to linearize the estimated nonlinear dynamics while propagating the uncertainty throughout a trajectory. Finally, we presented simulations that showcase the benefits of TPR over GPR in the presence of outliers as well as in the case of heteroscedastic noise.

References

[1] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media, 2012.

[2] Y. Chen, Y. Gu, and A. O. Hero, "Sparse lms for system identification," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2009, pp. 3125–3128.

[3] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, pp. 1–47, 2019.

[4] T. B. Schön, A. Wills, and B. Ninness, "System identification of nonlinear state-space models," *Automatica*, vol. 47, no. 1, pp. 39–49, 2011.

[5] S. Chen, S. Billings, and P. Grant, "Non-linear system identification using neural networks," *International journal of control*, vol. 51, no. 6, pp. 1191–1214, 1990.

[6] D. Papadimitriou, U. Rosolia, and F. Borrelli, "Control of unknown nonlinear systems with linear time-varying mpc," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 2258–2263.

[7] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer School on Machine Learning*. Springer, 2003, pp. 63–71.

[8] J. Kocijan, R. Murray-Smith, C. E. Rasmussen, and A. Girard, "Gaussian process model based predictive control," in *Proceedings of the 2004 American control conference*, vol. 3. IEEE, 2004, pp. 2214–2219.

[9] T. Koller, F. Berkenkamp, M. Turchetta, and A. Krause, "Learning-based model predictive control for safe exploration," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 6059–6066.

[10] A. Lederer, J. Umlauft, and S. Hirche, "Uniform error bounds for gaussian process regression with application to safe control," in *Advances in Neural Information Processing Systems*, 2019, pp. 659–669.

[11] A. Shah, A. Wilson, and Z. Ghahramani, "Student-t processes as alternatives to gaussian processes," in *Artificial intelligence and statistics*, 2014, pp. 877–885.

[12] A. Shah, A. G. Wilson, and Z. Ghahramani, "Bayesian optimization using student-t processes," in *NIPS Workshop on Bayesian Optimisation*, 2013.

[13] C. Berg and C. Vignat, "On the density of the sum of two independent student t-random vectors," *Statistics & probability letters*, vol. 80, no. 13-14, pp. 1043–1055, 2010.

[14] P. Ding, "On the conditional distribution of the multivariate t distribution," *The American Statistician*, vol. 70, no. 3, pp. 293–295, 2016.

[15] S. K. Kwak and J. H. Kim, "Statistical data preparation: management of missing values and outliers," *Korean journal of anesthesiology*, vol. 70, no. 4, p. 407, 2017.

[16] Q. V. Le, A. J. Smola, and S. Canu, "Heteroscedastic gaussian process regression," in *Proceedings of the 22nd international conference on Machine learning*, 2005, pp. 489–496.

[17] D. Nguyen-Tuong and J. Peters, "Local gaussian process regression for real-time model-based robot control," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 380–385.

[18] L. Muñoz-González, M. Lázaro-Gredilla, and A. R. Figueiras-Vidal, "Heteroscedastic gaussian process regression using expectation propagation," in *2011 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE, 2011, pp. 1–6.

[19] F. Berkenkamp, M. Turchetta, A. P. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," *arXiv preprint arXiv:1705.08551*, 2017.