Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Nonparametric Identification of Linearizations and Uncertainty using Gaussian Process Models – Application to Robust Wheel Slip Control

Jostein Hansen*, Roderick Murray-Smith†‡ and Tor A. Johansen*§

*Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway.
E-mail: {josteih,torj}@itk.ntnu.no
†Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, Scotland, UK. E-mail: rod@dcs.gla.ac.uk
‡ Hamilton Institute, NUI Maynooth, Ireland
§ SINTEF Information and Communication Technology, Trondheim, Norway.

***Abstract*— Gaussian process prior models offer a nonparametric approach to modelling unknown nonlinear systems from experimental data. These are flexible models which automatically adapt their model complexity to the available data, and which give not only mean predictions but also the variance of these predictions. A further advantage is the analytical derivation of derivatives of the model with respect to inputs, with their variance, providing a direct estimate of the locally linearized model with its corresponding parameter variance. We show how this can be used to tune a controller based on the linearized models, taking into account their uncertainty. The approach is applied to a simulated wheel slip control task illustrating controller development based on a nonparametric model of the unknown friction nonlinearity. Local stability and robustness of the controllers are tuned based on the uncertainty of the nonlinear models' derivatives.**

I. INTRODUCTION

Robust control is a fairly mature field, in particular for the controller synthesis problem for linear systems, where numerous approaches exist, [1] and [2]. However, robust control synthesis relies on a description or model of plant uncertainty. Although system identification methods as in [3] and [4], may provide uncertainty information, this may be difficult to apply directly in robust control synthesis since this information may prove to be misleading in case of structural model mismatch. The availability of reliable uncertainty estimates is a major concern in applications of robust control, in particular for nonlinear systems.

Many model-based nonlinear control problems are still based on parametric models, where the functional form is fully described by a finite number of parameters, often a linear function of the parameters. Even in the cases where flexible parametric models are used, such as neural networks, spline-based models, multiple models etc, the uncertainty is usually expressed as uncertainty of parameters (even though the parameters often have no physical interpretation), and do not take into account uncertainty about model structure, or distance of current prediction point from training data used to estimate parameters.

In such cases, an alternative approach is that some parts or even the whole structure, may be given by nonparametric models, such as Gaussian Process prior (GP) models. In this paper, we study an approach which has the following properties. It is a non-parametric model which retains the available data and performs inference conditional on the current state and local data (called smoothing in some frameworks). The uncertainty of model predictions is dependent on local data density, noise on data, and model mismatch. The final model complexity is automatically related to the amount and distribution of available data (more complex models need more evidence to make them likely), as well as the complexity of the target system. These aspects are especially useful in transient regimes with sparse data, in system identification tasks.

II. GAUSSIAN PROCESS PRIOR

In a Bayesian framework the model must be based on a prior distribution over the infinite-dimensional space of functions. As illustrated in [5], such priors can be defined as Gaussian processes. These models have attracted a great deal of interest recently, in for example reviews such as [6]. Rasmussen [7] showed empirically that Gaussian processes were extremely competitive with leading nonlinear identification methods on a range of benchmark examples.

The further advantage that they provide analytic predictions of model uncertainty makes them very interesting for control applications. Early use of GPs in a control systems context is discussed in [8], [9]. A variation which can include ARMA noise models is described in [10]. k-step ahead prediction with GP's is described in [11], [12]. [13] provides a number of chapters on recent applications of Gaussian processes in control contexts.

A. Inference with Gaussian processes

Assume we are modelling a nonlinear target function $f(\mathbf{x})$ where the observed outputs y^i to inputs $\mathbf{x}^i$, subject to noise ϵ_i can be described by the equation

$$y^i = f(\mathbf{x}^i) + \epsilon_i \tag{1}$$

and that we can observe a set S of input/output pairs $X, \mathbf{y}$, or $\{(\mathbf{x}^i, y^i)\}$ are given, where $\mathbf{x}^i \in \mathbf{R}^D,\ y^i \in \mathbf{R},\ i = 1 \dots N$, hence:

$$X = \begin{bmatrix} \mathbf{x^1} \\ \vdots \\ \mathbf{x^i} \\ \vdots \\ \mathbf{x^N} \end{bmatrix}, y = \begin{bmatrix} y^1 \\ \vdots \\ y^i \\ \vdots \\ y^N \end{bmatrix} \tag{2}$$

Instead of parameterising the system as a parametric model, we are placing a prior directly on the space of functions where f is assumed to belong. A Gaussian process represents the simplest form of prior over functions, we assume that any N points have a N-dimensional multivariate Normal distribution. In the GP framework, the output values y^i are viewed as being drawn from a zero-mean multivariable Gaussian distribution whose covariance matrix is a function of the input vectors $\mathbf{x}^i$. Namely the output distribution is

$$(y^1, \ldots, y^N | \mathbf{x}^1, \ldots, \mathbf{x}^N) \sim \mathcal{N}(0, \Lambda(X, X)).$$

Where $\Lambda(\mathbf{x}^i, \mathbf{x}^j) = \mathrm{cov}(y^i, y^j)$ is the covariance matrix of the function observations. A general model, which reflects the higher correlation between spatially close (in some appropriate metric) points – a smoothness assumption in target function $f(\mathbf{x})$ – uses a covariance matrix with the following structure;

$$\Lambda(\mathbf{x}^i, \mathbf{x}^j) = \alpha \exp(-\frac{1}{2}\|\mathbf{x}^i - \mathbf{x}^j\|^2_\Gamma) + v_0 \delta_{i,j}, \quad (3)$$

where $\delta_{i,j} = 1$ for $i = j$, and zero otherwise. The norm $\|\cdot\|_\Gamma$ is defined as

$$\|\mathbf{u}\|_\Gamma = (u^T \Gamma u)^{\frac{1}{2}}, \quad \Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_D).$$

The $D + 2$ variables, $\alpha, \gamma_1, \ldots, \gamma_D, v_0$ are the hyperparameters of the GP model, which are constrained to be non-negative. In particular v_0 is included to capture the noise component of the covariance. The GP model can be used to calculate the distribution of an unknown output y^{N+1} corresponding to known input $\mathbf{x}^{N+1}$ as

$$(y^{N+1} | \mathbf{x}^1, \ldots, \mathbf{x}^N, \mathbf{x}^{N+1}, y^1, \ldots, y^N) \sim \mathcal{N}(\mu, \bar{\Lambda}),$$

where

$$\mu = \Lambda(\mathbf{x}^{N+1}, X)\Lambda^{-1}(X, X)\mathbf{y}, \quad (4)$$

$$\bar{\Lambda} = \Lambda(\mathbf{x}^{N+1}, \mathbf{x}^{N+1}) - \Lambda(\mathbf{x}^{N+1}, X)\Lambda^{-1}(X, X)\Lambda(X, \mathbf{x}^{N+1}) \quad (5)$$

so we can use μ as the expected model output, with a variance of $\bar{\Lambda}$. Note that the covariance matrix $\Lambda(X, X)$ will be $N \times N$ dimensional, so the computational cost of its inversion grows rapidly with the number of data points N.

1) Nonstationary covariance function: In this paper, instead of the stationary covariance function of (3) we use a nonstationary one,

$$\Lambda(\mathbf{x}^i, \mathbf{x}^j) = v_1 \sin^{-1}\left(\frac{\mathbf{x}^{i^T}\Gamma\mathbf{x}^j}{\sqrt{(1 + 2\mathbf{x}^{i^T}\Gamma\mathbf{x}^i)(1 + \mathbf{x}^{j^T}\Gamma\mathbf{x}^j)}}\right) + v_0\delta_{i,j} \quad (6)$$

as described in [14], using Rasmussen's MATLAB implementation.[1] $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_{D+1})$ where the positive γ_i, weight each input i (and an additional constant one acting as a bias term). The parameter vector $\Theta = \log[v_1, \gamma_{1,..,D+1}, v_0]^T$ (the log is applied elementwise and ensures positivity in the parameters) and D is the dimension of vector $\mathbf{x}$. The hyperparameters Θ, can be adapted as the model is fit to the identification data, using numerical methods such as standard gradient-based optimisation tools to optimise hyperparameters.

[1] Code available at http://www.kyb.tuebingen.mpg.de/bs/people/carl/code/gp/

B. Gaussian process derivatives

Differentiation is a linear operation, so the derivative of a Gaussian process remains a Gaussian process. We now use this fact to infer from and to a mixture of observations of values and derivatives. Suppose we are given new sets of pairs $S'_j = \{(\mathbf{x}^{j,i}, \omega^{j,i})\}, \ j = 1, \ldots, D, \ i = 1, \ldots K$, each S'_j corresponding to the K points of j^{th} partial derivative of the underlying function $y = f(\mathbf{x})$. In the noise-free setting this corresponds to the relation

$$\omega^{j,i} = \frac{\partial f(\mathbf{x})}{\partial x_j}|_{\mathbf{x}=\mathbf{x}^{j,i}}, \ i = 1, \ldots, K.$$

We now wish to find the joint probability of the vector of y's and ω's, which involves calculation of the covariance between the function and the derivative observations as well as the covariance among the derivative observations. Covariance functions are typically chosen to be differentiable, so the covariance between a derivative and function observation and the one between two derivative points satisfy

$$\mathrm{cov}(\omega^{j,m}, y^n) = \frac{\partial}{\partial x_j}\mathrm{cov}(y^m, y^n) \quad (7)$$

$$\mathrm{cov}(\omega^{j,m}, \omega^{i,n}) = \frac{\partial^2}{\partial x_j \partial x_i}\mathrm{cov}(y^m, y^n) \quad (8)$$

Including these in our covariance function allows us to identify and predict from a data set which includes a mixture of function and derivative observations. Predictions can be inferred function values or inferred derivative values, with standard deviations in both cases.

The use of derivatives of Gaussian processes is described in [15], [16], and in engineering applications in [8], [17], [18], [19]. This allows the integration of prior information in the form of state or control linearisations, as presented in [19], and importantly for this paper, GP models can provide local linearisation information with mean and uncertainty estimates. This is useful for controller development with robustness analysis in small regions around the point of operation, and make GP models interesting candidates for nonlinear robust control problems.

III. CASE STUDY;WHEEL SLIP CONTROL

An application to wheel slip control is studied to illustrate the controller development for a nonparametric nonlinear model.

A. Equations of motion of a quarter car

With reference to Figure 1, the quarter car model consists of a single wheel attached to a mass m. A tyre reaction force F_x is generated by the friction between the tyre surface and the road surface, while the wheel moves driven by inertia of the mass m in the direction of the velocity v. A rolling motion of the wheel ω will be initiated by a torque caused by the tyre reaction force. A brake torque applied to the wheel will act against the spinning of the wheel causing a negative angular acceleration. The equations of motion of the quarter car are

$$m\dot{v} = -F_x \quad (9)$$

$$J\dot{\omega} = rF_x - T_b sign(\omega) \quad (10)$$

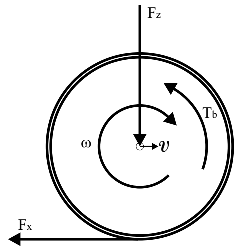


Fig. 1. Quarter car forces and torques.

where

v	longitudinal speed at which the car travels
ω	angular speed of the wheel
F_z	vertical force
F_x	tyre friction force
T_b	brake torque
r	wheel radius
J	wheel inertia

Further, the tyre friction force is given by $F_x = F_z \cdot \mu(\lambda, \mu_H, \kappa)$ where the friction coefficient μ is a nonlinear function of

λ	tyre slip
μ_H	maximum friction coefficient between tyre and road
κ	slip angle of the wheel

The slip $\lambda = (v - \omega r)/v$ describes the normalized difference between horizontal speed v and speed of the wheel perimeter ωr. The slip value of $\lambda = 0$ indicates that the wheel is in free motion and no friction force F_x is exerted. If the slip attains the value $\lambda = 1$ then the wheel is locked which means that it has come to a standstill. A change of variables is carried out where the angular speed of the wheel ω is replaced by the slip λ (assuming $\omega \geq 0$ and $v > 0$):

$$\dot{\lambda} = -\frac{1}{v}\left\{\frac{1}{m}(1-\lambda) + \frac{r^2}{J}\right\} F_z \mu(\lambda, \mu_H, \kappa) + \frac{1}{v} \cdot \frac{r}{J} T_b \tag{11}$$

$$\dot{v} = -\frac{1}{m} F_z \mu(\lambda, \mu_H, \kappa) \tag{12}$$

It can be seen that the time scale of the slip dynamics (11) scales with speed v. The qualitative dynamic behavior of slip is not affected by speed. Further assuming the slip angle of the wheel, κ, being zero, an example of such nonlinear tyre slip/friction curve, $\mu(\lambda, \mu_H)$, is shown in Figure 4. Several structural models of different complexities exist in the literature [20], [21], [22] and [23], but the detailed friction curve also depend on highly uncertain properties such as wear and tear of the tyres. A non-parametric GP model may therefore be a suitable alternative.

B. Control strategy

As in [24], assuming the velocity of the car varies much more slowly than the other variables involved and in addition that we have in general $\frac{J}{mr^2}(1-\lambda) \ll 1$, one obtains the dynamics of the tyre slip:

$$\dot{\lambda} v = -\frac{r^2 F_z}{J} \mu(\lambda, \mu_H) + \frac{r}{J} T_b \tag{13}$$

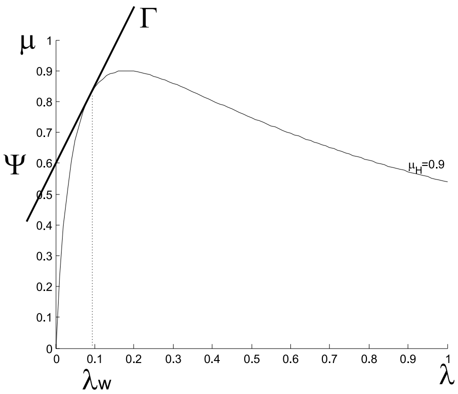


Fig. 2. Linearization $\mu_0(\lambda) = \Gamma\lambda + \Psi$ (solid line) of the nonlinear function $\mu(\lambda, \mu_H)$ (thin line) at $\mu_H = 0.9$ and $\lambda = \lambda_w$.

Further denoting $\beta = r^2 F_z / J$, $\alpha = r/J$ and adding hydraulic actuator dynamics, with time constant T_a:

$$\dot{T}_b = -\frac{1}{T_a} T_b + \frac{1}{T_a} u \tag{14}$$

$$\dot{\lambda} v = -\beta \mu(\lambda, \mu_H) + \alpha T_b \tag{15}$$

Further, a simplification is made by linearizing the nonlinear function $\mu(\cdot)$ at a point of operation λ_w, as seen in Figure 2. This new linearized function is given by $\mu_0(\lambda) = \Gamma\lambda + \Psi$, where Γ is the slope of the line and Ψ is the μ-intercept.

The total slip dynamics model may now be written as:

$$\dot{T}_b = -\frac{1}{T_a} T_b + \frac{1}{T_a} u \tag{16}$$

$$\dot{\lambda} v = -\beta(\Gamma\lambda + \Psi) + \alpha T_b \tag{17}$$

Assuming zero initial conditions the transfer function of (16) and (17) is:

$$h_p(s) = \frac{\frac{\alpha}{v}}{(s + \frac{\beta}{v}\Gamma)(1 + T_a s)} \tag{18}$$

We remark that Γ is to be considered a highly uncertain parameter, depending on both the operating point (reference slip value λ^*), and tyre/road properties. A GP model of the friction curve will be used to extract information about the uncertain parameters.

1) PID-controller: The literature presents a range of approaches to the use of PID-controllers in ABS problems, [25] and [26], to mention some. Here a different approach using a PID-controller to solve the ABS problem is presented.

An ideal PID-controller is given by:

$$h_c(s) = K_p \frac{(1 + T_i s)(1 + T_d s)}{T_i s} \tag{19}$$

Choosing $T_d = T_a$ leads to the open loop transfer function

$$h_0(s) = \frac{K_p \frac{\alpha}{v} s + \frac{K_p \alpha}{T_i v}}{s^2 + \frac{\beta}{v}\Gamma s} \tag{20}$$

which gives the following closed loop transfer function:

$$G(s) = \frac{Ts + 1}{1 + 2\zeta\frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} \tag{21}$$

where $2\zeta\omega_0 = \frac{\beta}{v}\Gamma + K_p\frac{\alpha}{v}$, $\omega_0^2 = \frac{K_p\alpha}{T_i v}$ and $T = T_i$. This gives the relative damping $\zeta = \left(\frac{\beta}{2v}\Gamma + K_p\frac{\alpha}{2v}\right)\sqrt{\frac{T_i v}{K_p\alpha}}$ and undamped resonance frequency $\omega_0 = \sqrt{\frac{K_p\alpha}{T_i v}}$.

2) Control performance requirement: For the closed loop system (21), the poles are given by the solutions of its characteristic equation. If the parameters of the characteristic equation of (21) is given as random variables with an expectation and variance, the poles are then also given as random variables.

Example 1: First assume for the system in (21) the parameter ζ is given as a random number with an expectation $E(\zeta)$ and a variance $\sigma^2(\zeta)$ and the parameter ω_0 is known exactly. The locations of the two complex conjugated poles are given by mean and standard deviation in a set of two arcs as seen in Figure 3 a). If instead the uncertainty is in the parameter ω_0, and ζ is known exactly, two lines will characterize locations of the mean and standard deviation of the poles, Figure 3 b). Finally, if ω_0 and ζ are independent random parameters, the mean and standard deviation of the poles are given as two areas, illustrated in Figure 3 c).

3) Controller Parameter tuning: Assume the controller is supposed to regulate the set-point $\lambda = \lambda^*$. The controller parameters may be tuned according to a linear model of the quarter car around this point. The GP model provides estimates of the expectation $E(\Gamma)$ and variance $\sigma^2(\Gamma)$.

In this case study, Γ is randomly given, hence the relative damping ζ will be a random parameter in (21), i.e $E(\zeta) = \left(\frac{\beta}{2v}E(\Gamma) + K_p\frac{\alpha}{2v}\right)\sqrt{\frac{T_i v}{K_p\alpha}}$. This is exactly the situation depicted in Figure 3 a).

Assume the desired bandwidth and the desired nominal relative damping are given by ω_{0d} and ζ_d respectively. The bandwidth ω_0 is known exactly, and given by $\omega_0 = \sqrt{\frac{K_p}{T_i}\frac{\alpha}{v}}$. If the relative damping ζ is assumed to be less than 1, the bandwidth may be chosen to be equal to ω_{0d}. Further, since the best we can do regarding tuning the relative damping, is given by $E(\zeta) = \zeta_d$, we get the following expressions for the controller parameters:

$$K_p = \zeta_d\omega_{0d}\frac{2v}{\alpha} - \frac{\beta}{\alpha}E(\Gamma) \tag{22}$$

$$T_i = \frac{K_p}{\omega_{0d}^2}\frac{\alpha}{v} \tag{23}$$

Given the above and the assumption of Γ being normally distributed with variance $\sigma^2(\Gamma)$, stability is ensured with at least 95% of confidence, i.e. behavior inside 2σ-contours, if $0 < \zeta^- < \zeta^+ < 1$ where:

$$\zeta^+ = \left(\frac{\beta}{2v}(E(\Gamma) + 2\sigma(\Gamma)) + K_p\frac{\alpha}{2v}\right)\sqrt{\frac{T_i v}{K_p\alpha}} \tag{24}$$

$$\zeta^- = \left(\frac{\beta}{2v}(E(\Gamma) - 2\sigma(\Gamma)) + K_p\frac{\alpha}{2v}\right)\sqrt{\frac{T_i v}{K_p\alpha}} \tag{25}$$

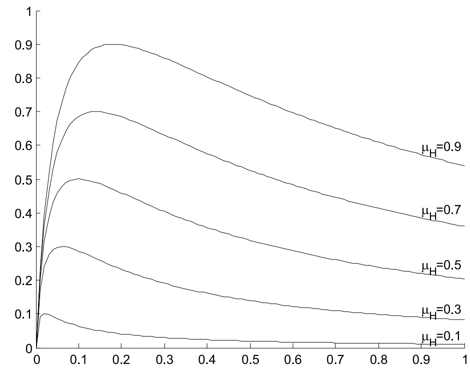


Fig. 4. Simplified tyre slip/friction curves $\mu(\lambda, \mu_H)$ by Daiss and Kiencke 1996.

C. Results

In the simulation that follows, a simplified model by [22] is used for the tyre friction curve $\mu(\lambda, \mu_H)$,

$$\mu(\lambda) = \frac{k\lambda}{a\lambda^2 + b\lambda + 1} \tag{26}$$

Note that this model is linear in the parameters a and b, where k is the slope at $\lambda = 0$.

First assuming the slope to be $k = 28$, and the slip at maximum friction to be given by $\lambda_0 = 0.2\mu_H$, where μ_H denotes the maximum friction coefficient at slip value λ_0. Further assuming the two parameters in (26) to be given as $a = 1/(\lambda_0^2)$ and $b = (k\lambda_0 - 2\mu_H)/(\lambda_0\mu_H)$, one obtains the following friction curve:

$$\mu(\lambda, \mu_H) = \frac{k\lambda}{\left(\frac{\lambda}{0.2\mu_H}\right)^2 + \frac{k0.2\mu_H - 2\mu_H}{0.2\mu_H}\lambda + 1} \tag{27}$$

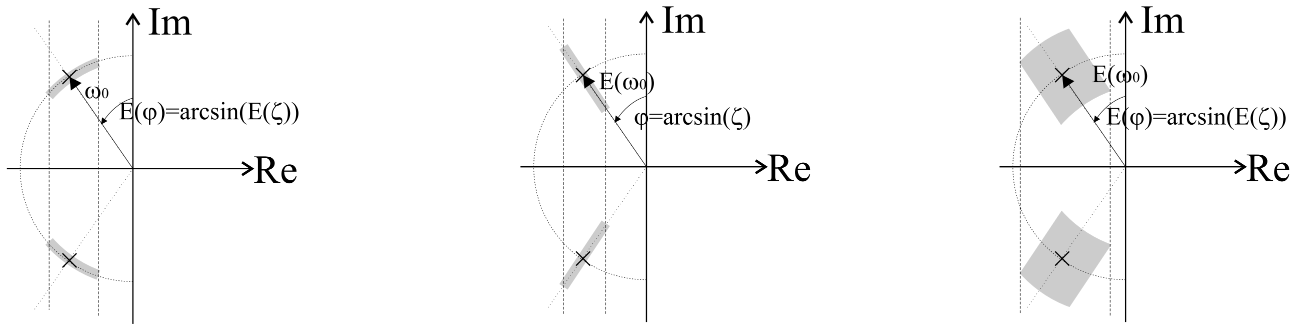
This simplified tyre friction curve is shown in Figure 4.

1) Modelling simulated data with a Gaussian Process: A set of only 10 noisy (Gaussian noise, zero mean with standard deviation, $\sigma = 0.05$) observations of the simulated curve is used to train a GP-model. Figure 4 shows curves which have significantly nonuniform curvatures, showing a rapid change for small values of λ, with more gradual change later. Because of this, we use a nonstationary covariance function, as defined in eq. (9). The parameters Θ were optimized for the given training data using a conjugate-gradient algorithm.

Figures 5 and 6 show the resulting GP model of the function value Ψ and the corresponding slope Γ. In addition to the noise property described above, also the *a priori* knowledge of the friction curve vanishing at the origin is included in the training data set.

The advantage of using a GP as a modelling tool is easily motivated by looking at Figure 5. Even with no structural knowledge of the nonlinear function and with fairly small, sparsely populated training set, the GP model is close to the correct one. The GP model in Figure 5 is based on a training set of 10 noisy data points, taken from a region where data typically are accessible. Note how the uncertainty increases in regions with no data, as would be expected.

Assume the controller is supposed to control a fixed slip



(a) (b) (c)

Fig. 3. Different pole locations. The vertical dotted lines indicates the real part of the mean with standard deviation of the pole placement. The uncertainties are shown as gray arcs, lines and areas. a) ζ uncertain. b) ω_0 uncertain. c) both ω_0 and ζ uncertain.

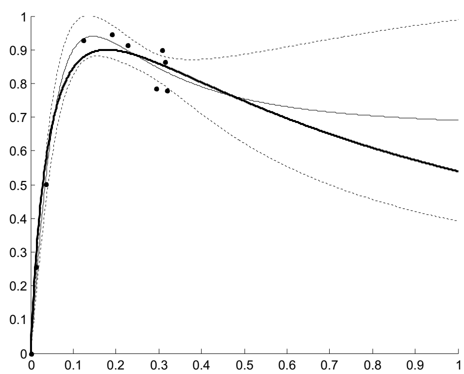


Fig. 5. GP model of friction coefficient. In addition to the set of accessible training data (sparsely populated dots), the model mean (thin line) and 2σ contours (dashed pair of lines) are shown. Also the real model curve (bold line) is shown for reference.

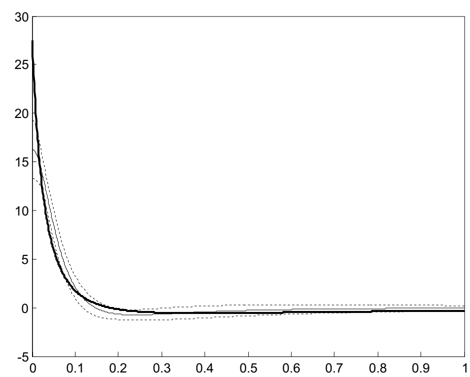


Fig. 6. Friction coefficient derivative $\frac{\partial \mu}{\partial \lambda}(\lambda)$. The model mean derivative (thin line) and 2σ contours (dashed pair of lines) are shown. Also the real model derivative curve (bold line) is shown for reference.

$\lambda^* = 0.15$. At the point of linearization (selected to be equal to the slip set-point, i.e. $\lambda_w = \lambda^*$) the nominal slope is $\Gamma = 0.3882$ with $\mu_H = 0.9$. But instead knowledge of an expectation and variance exist at this point, from our GP model. $E(\Gamma) = -0.0787$ and $\sigma^2(\Gamma) = 0.1610$. Further assume the parameters of the quarter car are given by $r = 0.35$, $F_z = 250 \cdot 9.81$ and $J = 0.68$. Hence $\beta = 441.8$ and $\alpha = 0.51$. The time constant of the actuator dynamics is $T_a = 0.014$ s. The desired closed-loop property is given by $\omega_{0d} = 20\pi$ and $\zeta_d = 0.5$. Since the relative damping is assumed to be less than 1, the bandwidth may be selected to be equal to the desired one, i.e. $\omega_0 = \omega_{0d}$. Plugging the mean into (22) and (23), a set of controller parameters, scheduled with speed v, are inferred:

$$K_p = \zeta_d \omega_{0d} \frac{2v}{\alpha} - \frac{\beta}{\alpha} E(\Gamma) = 122.07v + 67.55 \tag{28}$$

$$T_i = \frac{K_p}{\omega_{0d}^2} \frac{\alpha}{v} = 0.0159 + 0.0088 \frac{1}{v} \tag{29}$$

also with this given set of controller parameters, a 95% boundary of where the poles may be located inside is given by an upper and lower bound. For a given speed $v = 30$ m/s, these bounds are $\zeta^+ = 0.5940$ and $\zeta^- = 0.4060$, obtain from (24) and (25).

Figure 7 shows the pole placement of the controller parameter calculations with the 95% confidence interval, and the actual placement of the nominal poles. The are only 10 training points and we have set $v = 30$ m/s. The confidence interval is entirely inside the region of stability, hence stability is ensured with at least 95% of confidence.[2] Increasing the size of the training set, will tend to decrease the 95% confidence interval, Figure 8 shows the situation with 100 training points.

Only one road condition is considered in the case study, i.e. μ_H fixed. An alternative approach would be to include other road conditions as well. By training the model using data from different road conditions, a less certain model will lead to a more robust but hence a more conservative controller. Alternatively, an adaptive control approach with online training of the GP model could be implemented. In this case, tradeoffs between computational capacity and model update rate must be considered. Also, precautions due

[2]On a cautionary note, however, where we state e.g. '95% confidence intervals', these are conditioned on the GP with covariance parameters being an appropriate model. For simplicity, the examples in this paper used Maximum likelihood optimisation to find appropriate parameters for GP covariance function parameters. GPs are very flexible models, but optimisation will tend to make the model overconfident in its predictions. Taking the Bayesian approach of integrating over hyperparameter distributions, possibly implemented using MCMC algorithms, would be more robust, especially for small data sets [7].

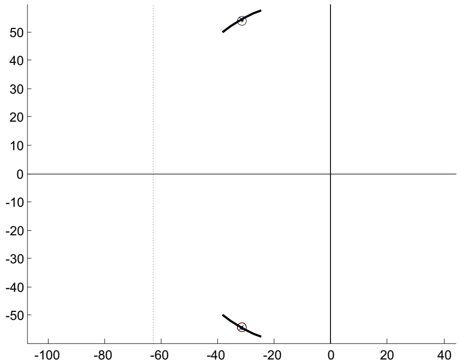


Fig. 7. Pole placement calculation. Mean placement of the poles are indicated with an x, the 95% confidence interval is indicated with the arc. The actual pole is indicated with the circle. The GP model is based on a training set of size 10. The dotted line is the line of ω_0.

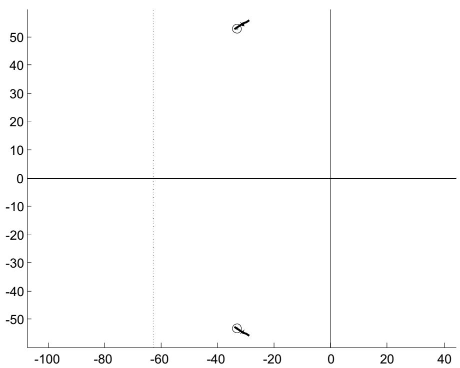


Fig. 8. Pole placement calculation. Mean placement of the poles are indicated with an x, the 95% confidence interval is indicated with the arc. The actual pole is indicated with the circle. The GP model is based on a training set of size 100. The dotted line is the line of ω_0.

to persistence of excitation must be made, e.g by turning off the estimator when information loss.

IV. Conclusions

We have shown that a nonparametric Gaussian process prior model can be used to model nonlinear, simulated tyre slip/friction curves, to a high degree of accuracy considering the sparseness of the training data. The inference based on the GP model provides not only mean predictions, with uncertainty estimates for the curves themselves, but also mean and uncertainty estimates of local linearisations of these curves, which is useful for robust control. We illustrate this with a pole-placement task for a PID controller.

Acknowledgements

This work was in part sponsored by the European Commission through the RTN project **MAC** (Multi Agent Control) HPRN-CT-1999-00107 , the STREP project **CEmACS** (Complex Embedded Automotive Control Systems), the Research Council of Norway and Science Foundation Ireland grant 00/PI.1/C067.

References

[1] K. Zhou, J.C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice-Hall, 1996.

[2] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*, John Wiley & Sons, 1996.

[3] A. Garulli, A. Tesi, and A. Vicino, Eds., *Robustness in Identification and Control*, Springer-Verlag, 1999, Number 245 in Lecture notes in Control and Information Sciences.

[4] W. Reinelt, A. Garulli, and L. Ljung, "Comparing different approaches to model error modelling in robust identification," *Automatica*, vol. (38:5), pp. 787–803, May 2002.

[5] A. O'Hagan, "On curve fitting and optimal design for regression (with discussion)," *Journal of the Royal Statistical Society B*, vol. 40, pp. 1–42, 1978.

[6] C. K. I. Williams, "Prediction with Gaussian processes: From linear regression to linear prediction and beyond," in *Learning and Inference in Graphical Models*, M. I. Jordan, Ed., pp. 599–621. Kluwer, 1998.

[7] C. E. Rasmussen, *Evaluation of Gaussian Processes and other Methods for Non-Linear Regression*, Ph.D. thesis, Graduate department of Computer Science, University of Toronto, 1996.

[8] R. Murray-Smith, T. A. Johansen, and R. Shorten, "On transient dynamics, off-equilibrium behaviour and identification in blended multiple model structures," in *European Control Conference, Karlsruhe, 1999*, 1999, pp. BA–14.

[9] D. J. Leith, R. Murray-Smith, and W. E. Leithead, "Nonlinear structure identification: A Gaussian Process prior/Velocity-based approach," in *Control 2000, Cambridge*, 2000.

[10] R. Murray-Smith and A. Girard, "Gaussian Process priors with ARMA noise models," in *Irish Signals and Systems Conference, Maynooth*, 2001, pp. 147–152.

[11] A. Girard, C. E. Rasmussen, J. Quiñonero Candela, and R. Murray-Smith, "Gaussian process priors with uncertain inputs – application to multiple-step ahead time series forecasting," in *Advances in Neural Information Processing Systems 15*, S. Thrun S. Becker and K. Obermayer, Eds., pp. 529–536. MIT Press, Cambridge, MA, 2003.

[12] J. Kocijan, R. Murray-Smith, C.Rasmussen, and A. Girard, "Gaussian process model based predictive control," in *American Control Conference, Boston*, 2004.

[13] R. Murray-Smith and R. Shorten, *Proceedings of the Hamilton Summer School on Switching and Learning in Feedback systems*, vol. 3355, Springer-Verlag, Lecture Notes in Computing Science, 2005.

[14] C. K. I. Williams, "Computation with infinite neural networks," *Neural Computation*, vol. 10, pp. 1203–1216, 1998.

[15] A. O'Hagan, "Some Bayesian numerical analysis," in *Bayesian Statistics 4*, J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Eds., pp. 345–363. Oxford University Press, 1992.

[16] C. E. Rasmussen, "Gaussian processes to speed up Hybrid Monte Carlo for expensive Bayesian integrals," Draft: available at http://www.gatsby.ucl.ac.uk/ edward/pub/gphmc.ps.gz, 2003.

[17] R. Murray-Smith and D. Sbarbaro, "Nonlinear adaptive control using non-parametric Gaussian process prior models," in *15th IFAC World Congress on Automatic Control, Barcelona*, 2002.

[18] D. J. Leith, W. E. Leithead, E. Solak, and R. Murray-Smith, "Divide & conquer identification: Using Gaussian process priors to combine derivative and non-derivative observations in a consistent manner," in *Conference on Decision and Control*, 2002.

[19] E. Solak, R. Murray-Smith, W. E. Leithead, D. J. Leith, and C. E. Rasmussen, "Derivative observations in Gaussian process models of dynamic systems," in *Advances in Neural Information Processing Systems 15*, S. Thrun S. Becker and K. Obermayer, Eds., pp. 1033–1040. MIT Press, Cambridge, MA, 2003.

[20] E. Bakker, L. Nyborg, and H. B. Pacejka, "Tyre modelling for use in vehicle dynamic studies," Tech. Rep. 870421, Society of Automotive Engineers, Warrendale PA, 1987.

[21] H. B. Pacejka and R. S. Sharp, "Shear force developments by pneumatic tyres in steady-state conditions: A review of modeling aspects," *Vehicle Systems Dynamics*, vol. 29, pp. 409–422, 1991.

[22] A. Daiss and U. Kiencke, "Estimation of tyre slip during combined cornering and braking observer supported fuzzy estimation," in *13th Triennal World Congress*, San Francisco, 1996, pp. 41–46.

[23] M. Burckhardt, *Fahrwerktechnik: Radschlupf-Regelsysteme*, Vogel Verlag, 1993.

[24] S. Solyom, *Control of Systems with Limited Capacity*, Ph.D. thesis, Lund Institute of Technology, August 2004.

[25] F. Jiang and Z. Gao, "An application of nonlinear PID control to a class of truck ABS problems," in *IEEE Conference on Decision and Control*, Orlando, Dec. 2001.

[26] Y. Wang, T. Schmitt-Hartmann, M. Schinkel, and K. J. Hunt, "A new approach to simultanious stabilisation and strong simultanious stabilisation with d stability and its application to abs control systems design," in *European Control Conference*, Porto, Sept. 2001.