Belt-Pulley Friction Estimation for the Continuously Variable Transmission

Koos van Berkel, Toru Fujii, Theo Hofman, Maarten Steinbuch

Abstract—This study describes an approach to estimate the maximum friction between the pushbelt and pulley set in a mass-produced Continuously Variable Transmission. Up-to-date friction knowledge is useful to accurately determine the maximum transmittable torque and to monitor the variator condition. Besides standard sensors, torque signals are assumed to be available. An adaptive estimator is presented based on the Kalman filter, which estimates friction-related parameters to find the maximum friction. The filter requires no a priori knowledge of the maximum friction and needs no detailed friction model. Experiments show that this method can be used for both an undamaged and damaged pushbelt.

Index Terms—Automotive Control, Estimation, System Identification

I. INTRODUCTION

The market for Continuously Variable Transmissions (CVTs) is rapidly growing, especially in Asia [1]. The key component in the CVT is the variator, which transfers torque and varies the speed ratio. The variator consists of a metal V-belt, i.e., pushbelt which is clamped between two pulley sets; one on each side of the variator. Torque is transferred from the primary pulley set via the pushbelt to the secondary pulley set by means of friction. The amount of friction depends on the relative velocity between the contact surfaces, as for example characterized by a typical friction curve shown in Fig. 1. The belt-pulley friction is a key parameter in variator control designs to calculate the required clamping force for the pulley sets for the desired torque transfer. Clamping forces lower than the required minimum result in a relatively large (i.e., macroscopic or macro) belt slip that may cause severe damage [2], whereas higher clamping forces (over-clamping) result in higher internal losses — thereby decreasing the overall transmission efficiency [3]. The belt-pulley friction is not constant; it depends on its operating conditions such as the angular speed, torque, and speed ratio [4], but it may also vary in time as it is influenced by temperature, wear of the contact surfaces, and quality of the lubrication fluid, among others [5]. In current mass-produced CVTs, the actual friction curve is unknown, which contributes to a rather conservative variator control design, i.e., the pulley sets are over-clamped to avoid macro slip [6] — the transmission efficiency is not optimal. The following research question arises: is it possible to estimate the belt-pulley friction in a mass-produced CVT under driving conditions? Thereby, the purpose is twofold:

1) improving transmission efficiency: with up-to-date knowledge of the belt-pulley friction, the clamping forces (and internal losses) may be reduced;
2) condition monitoring: the change in belt-pulley friction can be used to indicate damage and wear of the CVT.

A small number of publications is found on friction estimation in CVTs. Friction estimators described in [7]–[9] are model-based and designed for macro slip detection, yet not to estimate the maximum friction. This study presents a novel method to estimate the maximum belt-pulley friction in a mass-produced CVT, based on the Kalman filter. The proposed method requires no a priori knowledge of the friction curve, but uses a random-walk-like friction slope model for the Kalman filter. The following is assumed:

1) measurements, or estimates are available for the speed, torque 1, and clamping pressure of both pulleys;
2) the belt-pulley friction estimation is not part of a closed-loop control system, i.e., there are no requirements on bandwidth, or causality.

The outline is given as follows: Section II describes the variator working principle and defines suitable measures for slip and friction. Section III presents an estimation algorithm, which distinguishes different regions in the

K. van Berkel, T. Hofman, and M. Steinbuch are with the Department of Mechanical Engineering, Eindhoven University of Technology, Den Dolech 2, 5600 MD Eindhoven, The Netherlands, Phone: +31404272811, Fax: +31402461418, k.v.berkel@tue.nl, t.hofman@tue.nl, m.steinbuch@tue.nl.
T. Fujii is with the Department of Mechanical Engineering, Doshisha University, Kyotofo Prefecture, Tataradani 1-3, Japan, tfujii@mail.doshisha.ac.jp.

Footnote—1Torque sensors are not common, mainly for reasons of cost and packaging [6]. Instead, estimates can be used, e.g., based on the throttle valve position by the Engine Control Unit, and based on vehicle acceleration by control systems such as Anti-lock Brake System, Electronic Stability Control, and Cooperative Adaptive Cruise Control system [10].

Fig. 1. Schematic friction curves for a new and damaged friction surfaces.
friction curve in order to estimate the maximum friction. Section IV presents experimental results for both an undamaged and damaged pushbelt. Finally, Section V gives conclusions and directions for future work.

II. SYSTEM DESCRIPTION

A. Variator working principle

The considered variator consists of a metal pushbelt, which is clamped between a pulley set on each side of the variator (cf. Fig. 3). Subscript \(i \in \{p, s\}\) denotes the primary (i.e., driving) pulley side \(p\) and the secondary (i.e., driven) pulley side \(s\). Each pulley pair consists of one axially fixed sheave and one axially moveable sheave. On each moveable sheave, an axially pulley force \(F_i\) is generated by an oil pressure \(p_i\), controlled by a hydraulic system. The level of the clamping forces determines the maximum transmittable torque, whereas the ratio of the clamping forces determines (non-linearly) the speed ratio.

![Fig. 2. Schematic view of the pushbelt variator, with angular speeds \(\omega_i\), torques \(T_i\), clamping forces \(F_i\), and corresponding pressures \(p_i, i \in \{p, s\}\).](image)

The pushbelt consists of around 400 V-shaped compression elements (i.e., segments, or blocks) that are held together by two sets of \(9 - 12\) thin tension bands (i.e., rings), cf. with Fig. 3. The amount of friction depends on the relative velocity between each contact surface within the variator. There exists several contact surfaces: (1) between the pulley and the elements, (2) between the element and the bands, and (3) between the bands [5]. For variator control purposes, the overall friction between the pushbelt and pulley is of interest, which can be described as a function of the overall relative velocity (i.e., slip) of the variator.

![Fig. 3. The metal V-shaped pushbelt [11].](image)

B. Belt-pulley slip

The overall slip of the variator is defined by

\[
\sigma := \frac{\omega_p R_p - \omega_s R_s}{\omega_p R_p} = 1 - \frac{r_s}{r_g}, \tag{1}
\]

where \(\omega_i\) denotes the angular speed of pulley \(i\) and \(R_i\) denotes the radius of the pushbelt to the pulley center. The speed ratio is defined by \(r_s = \omega_s/\omega_p\) and the geometrical ratio is defined by \(r_g = R_p/R_s\). The geometrical ratio is affected by the clamping force, caused by deformations within the variator [12], and may be formulated as

\[
r_g = f(r_{g0}, F_s). \tag{2}
\]

Here, the deformation function \(f(r_{g0}, F_s)\) monotonically increases with increasing \(F_s\) for fixed \(r_{g0}\), where \(r_{g0}\) denotes the geometrical ratio ideally defined at zero clamping force, i.e., \(r_{g0} = \{r_g | F_s = 0\}\). In our case, both the geometrical ratio \((r_g)\) and deformation function \((f(r_{g0}, F_s))\) are unknown, but \(r_{g0}\) is constant and known at the kinematical boundaries of the variator, i.e., at speed ratios \textit{low} and \textit{overdrive}. Therefore, at these boundaries, slip can be approximated by

\[
\sigma \approx \sigma^* := 1 - \frac{r_s}{r_{g0}}, \tag{3}
\]

where \(r_{g0} = 0.405\) for speed ratio \textit{low} and \(r_{g0} = 2.25\) for speed ratio \textit{overdrive}.

C. Belt-pulley friction

The overall belt-pulley friction \(\mu_i\) of pulley \(i\) is defined by

\[
\mu_i := \frac{T_i \cos \theta}{2R_i F_i}, \tag{4}
\]

where \(T_i\) denotes the torque acting on pulley \(i\) and \(\theta\) half the pulley wedge angle \((\theta = 11^\circ, [2])\). In the following, the secondary friction \((\mu_s)\) is considered, as this value is usually used to determine the minimum required clamping force \((F_s)\) for torque transfer [13]. In practice, often an approximate is used, i.e., \(\mu_s \approx \mu^*\), which relates the primary torque \(T_p\) with the secondary clamping force \(F_s\), such that the estimated engine torque can be used to calculate \(T_p\). Here, \(\mu^*\) is defined by

\[
\mu^* := \frac{T_p \cos \theta}{2R_p F_s}. \tag{5}
\]

For fixed \(T_p\) and \(R_p\), thus with varying \(F_s\), (3) and (5) describe a friction curve as schematically depicted in Fig. 4. This friction curve is dominated by variator deformation for low \(\mu^*\) (i.e., high \(F_s\)), whereas it is dominated by macro slip for high \(\sigma^*\). The shape of this curve differs with the curve shown in Fig. 1; this is explained with the difference in the definitions (1) and (3). Note that for fixed \(F_s\) (i.e., varying \(T_p\)), the variator deformation will be constant, so the shape of the friction curve will be similar to Fig. 1.
D. What is a suitable measure to quantify friction?

The friction will be estimated for two purposes, i.e., to determine the maximum transmittable torque (thus maximum friction) and to monitor the variator condition. The friction curve is, however, subject to change due to changes in temperature, material wear of the contact surfaces, and quality of the lubrication fluid. Hence, a suitable measure to quantify friction is: (i) related to the maximum friction, and (ii) independent of the value for \( \sigma^* \), as the shape of the friction curve may change. Suitable candidates are, e.g.,

1) the maximum friction \( \mu_{\text{max}} \); 
2) the friction at minimum \( \sigma^* \), denoted as \( \mu_{\text{esc}} \).

as shown in Fig. 4. The (dis-) advantages of each candidate will be explained next. The maximum friction is defined by

\[
\mu_{\text{max}} := \max \mu^*(\sigma^*). \tag{6}
\]

This candidate estimates directly the desired quantity. The main concern, however, is that the slip cannot be controlled at this point (without additional sensors as in [4]) and \( \mu_{\text{max}} \) locates on the edge of the instable slip region (cf. Fig. 4). The friction at minimum \( \sigma^* \) is defined by

\[
\mu_{\text{esc}} := \mu^*(\sigma_{\text{min}}^*), \tag{7}
\]
\[
\sigma_{\text{min}}^* := \min \sigma^*(\mu^*). \tag{8}
\]

This point coincides with the operation point of Extremum Seeking Control (ESC), as described in, e.g., [3], [6]. ESC uses the concave shape of the friction curve around the minimum in \( \sigma^* \) (i.e., maximum in \( r_s \)) to operate the variator at \( \mu^* \approx \mu_{\text{esc}} \). The main advantage of using \( \mu_{\text{esc}} \) is that ESC can be used to control the slip. However, the relation between \( \mu_{\text{esc}} \) and the maximum friction is not clear, especially when the friction curve changes.

In this study, the estimation of \( \mu_{\text{max}} \) is explored using experiments. Although the slip cannot be controlled, it is possible to cross this region with increasing slip (by decreasing the clamping pressure), then stop the experiment when the slip becomes too large and estimate \( \mu_{\text{max}} \) with the experimental data. It is sufficient to perform this experiment occasionally, as the friction estimation is not part of a closed-loop control system (see, assumption 2 in Section I-A). It might be difficult to avoid macro slip, however, in [2] and [4] the authors show that with either a relatively low applied torque, or a relatively low slip velocity, no damage occurs with macro slip. Then, a suitable situation to perform an experiment would be, e.g., during coasting down (with low torque) in front of a traffic light.

III. ESTIMATOR DESIGN

An example of a measured friction curve is shown in Fig. 5. It can be observed that the signal-to-noise ratio is relatively low, especially, for low values of \( \sigma^* \) (see, black line). Also, with this experiment, the amount of slip is more than necessary to estimate \( \mu_{\text{max}} \). For the estimator design, three questions arise: (1) what type of estimator is suitable?; (2) when can the experiment be stopped?; (3) how can \( \mu_{\text{max}} \) be estimated using the experimental data?

A. What type of estimator is suitable?

Many friction estimator designs are based on (detailed) friction models, such as the black-box Bakker-Pacejka model (magic formula) [14], [15], the dynamic LuGre model [16], and the static Coulomb model [11]. The main advantage of such estimators is that knowledge of the friction model is exploited for accurate estimation. The main concern is that in our case, the friction curve is subject to change and it is not clear how the friction curve changes, e.g., due to wear of the contact surfaces. Consequently, each model parameter needs to be estimated (e.g., 6 parameters for LuGre model), thereby increasing the complexity of the estimator. This motivates the use of a estimator that uses no detailed friction model, such as the Kalman Filter (KF) [17] as suggested in [18] and [19] for tire-road friction estimation. The KF can be used to estimate the friction slope \( k(t) \):

\[
k(t) = \frac{d\mu^*(t)}{d\sigma^*(t)}. \tag{9}
\]

which provides qualitative information of the friction curve, at time \( t \). The friction slope contains sufficient information to detect the maximum friction region (i.e., \( k(t) = 0 \)) and to detect macro slip (i.e., \( k(t) < 0 \)). Furthermore, \( k(t) \) can still

![Fig. 4. Schematic friction curve for varying clamping force \( F_s \).](image)

![Fig. 5. Measured friction curve for decreasing clamping force \( F_s \) in overdrive. Low-pass filters with cut-off frequencies of 5 Hz (gray line) and 1 Hz (black line) are applied.](image)
be used when the friction curve changes. This makes the KF very suitable for friction estimation.

**B. When to stop the experiment?**

The experiment is stopped when sufficient experimental data is obtained for accurate estimation. Estimators that contain (detailed) knowledge of the friction curve require less experimental data, so $\mu_{\text{max}}$ can be predicted before the experiment has reached $\mu_{\text{max}}$. Then, in order to stop the experiment, macro slip detectors can be used, e.g., as described in [7]–[9]. Another useful measure may be the so-called slip-state ID, based on the magnitude of the transfer function from $\omega_p$ to $\omega_s$ [20]. Estimators that use no (detailed) friction model require more experimental data for accurate estimation, so the experiment is preferably stopped after $\mu_{\text{max}}$ has reached. This requires a different detector to stop the experiment, e.g., $k(t) < c_k$ with $c_k < 0$ to be determined.

**C. How to estimate the maximum friction?**

In the following, the use of the KF is explored to estimate the friction slope $k(t)$. The KF design is inspired on the adaptive tire-road friction estimator as described in [19]. The estimated friction slope $k(t)$ is used to estimate $\mu_{\text{max}}$ and used in a detector to stop the experiment. Two problems are identified using $k(t)$: (i) in the microscopic (or, micro) slip region (i.e., before $\mu_{\text{esc}}$), the slip $\sigma^*$ is not sufficiently excited for accurate estimation; and (ii) the friction slope goes to infinity around $\mu_{\text{esc}}$, which results in numerical issues. One way to avoid these problems is suggested as follows: monitor a measure that indicates the slip excitation ($v(t)$), and activate the KF when the slip is excited, i.e., after point $\mu_{\text{esc}}$, using $v(t) > c_v$ with $c_v > 0$ to be determined. The next algorithm is proposed:

1) start experiment by reducing the $F_s$; estimate the slip excitation $v(t)$;
2) when $v(t) > c_v$, estimate friction slope $k(t)$;
3) when $k(t) < c_k$, stop experiment; estimate $\mu_{\text{max}}$ from experimental data for which $|\dot{k}(t)| < c_m$ holds;
with $c_m > 0$ to be determined. In the following, the estimation in each step is described in more detail.

1) **Estimate the slip variance**: The variance of $\sigma^*$ over a time interval of $T$ time samples is a suitable measure to indicate the slip excitation, given by

$$v(t) = \frac{1}{T} \sum_{\tau=t-T+1}^{t} (\sigma^*(\tau))^2 = \left( \frac{1}{T} \sum_{\tau=t-T+1}^{t} \sigma^*(\tau) \right)^2,$$

for $t > T$. This filter has one tuning parameter: $T$ determines the time constant of the estimation, i.e., a small value gives a quick, but noisy response to changes in the slip excitation, whereas a large value gives a slow, but smooth response. The variance $v(t)$ and threshold $c_v$ will be used to distinguish measurement noise from significant changes in $\sigma^*$.

2) **Estimate friction slope**: The friction slope $k(t)$ is estimated when $\sigma^*(t)$ is sufficiently excited. This is implemented as follows: if $v(t) > c_v$, the KF updates estimate $\dot{k}(t)$, if not, $\dot{k}(t)$ is held constant. For the KF design, the non-linear friction curve is modeled by a discrete-time linear regression model with time-varying parameters, given by

$$\mu^*(t) = \left[ \sigma^*(t) \right] \left[ \frac{k(t)}{\mu_0(t)} \right] + e(t),$$

$$z(t) = H(t)x(t) + e(t).$$

Here, $\mu_0(t)$ is a (time-varying) friction offset, $x(t) = \left[ k(t) \mu_0(t) \right]^T$ is a state vector, $z = \mu^*(t)$ is a measurement output term, $H(t) = [\sigma^*(t) 1]$ is a regression vector and $e(t)$ is an error term for the measurement and model errors. The states vary like a random walk, described by

$$x(t+1) = x(t) + w(t),$$

with variance $w(t)$. Both $e(t)$ and $w(t)$ are considered as independent white-noise processes, with zero mean and

$$Q(t) = E[w(t)w^T(t)],$$

$$R(t) = E[e(t)e^T(t)],$$

where $E[y]$ denotes the expectation value of variable $y$. With these assumptions, the KF gives the optimal (in the minimum-variance sense) state estimates $\hat{x}(t)$:

$$\dot{\hat{S}}(t) = \dot{\hat{P}}(t-1) + Q(t-1),$$

$$K(t) = \hat{S}(t)H^T(t)(H(t)\hat{S}(t)H^T(t) + R(t))^{-1},$$

$$\hat{x}(t) = \hat{x}(t-1) + K(t)(z(t) - H(t)\hat{x}(t-1)),$$

$$\dot{\hat{P}}(t) = (I - K(t)H(t))\hat{S}(t).$$

Here, the gain $K(t)$ can be interpreted as a weight for the estimations based on model and measurement knowledge. $\hat{P}(t)$ is interpreted as the covariance matrix of the state estimates. The following matrices are freely to choose: the tuning matrices $Q(t)$ and $R(t)$, and initializations $\hat{x}(0)$ and $\hat{P}(0)$. The ratio $||Q(t)||/||R(t)||$ may be considered as the time constant of the estimation: a large value gives a quick, but noisy response to changes in the states, whereas a small value gives a slow, but smooth response. It is assumed that no prior information is available when large state variations are expected, so the matrices can be chosen time-invariant. Note that the states vary dependently, but since there is no friction model, the following is chosen:

$$R(t) = 1,$$

$$Q(t) = \begin{bmatrix} q_k & 0 \\ 0 & q_{\mu_0} \end{bmatrix},$$

with constants $q_k$ and $q_{\mu_0}$. The magnitudes of $q_k$ and $q_{\mu_0}$ determine the time constants of each state estimate, i.e., $\hat{k}$ and $\hat{\mu}_0$. A smooth $\hat{k}$ is desired, so $q_k << q_{\mu_0}$. For
well-chosen time constants, the estimates converge quickly so that initializations $\hat{x}(0)$ and $\hat{P}(0)$ do not influence the experiment (if reasonably chosen). The estimated friction slope $\hat{k}$ and threshold $c_k$ will be used to stop the experiment.

3) Estimate maximum friction: $\mu_{\text{max}}$ is estimated after the experiment, hence a time-invariant estimator suffices. Here, sample mean estimation is used to find $\mu_{\text{max}}$ from a data set $\mu_{\text{set}}$ (with length $N$) for which the friction slope approximates zero, i.e.,

$$\hat{\mu}_{\text{max}} = \frac{1}{N} \sum_{n=1}^{N} \mu_{\text{set}}(n),$$

$$\mu_{\text{set}} = \{\mu^* | |\hat{k}| < c_m\}.$$ (22)

(23)

IV. EXPERIMENTS

A. Experimental setup

Experiments are performed on a test-rig as described in [21]. It consists of a variator with a speed-controlled electric machine on the input side, a torque-controlled electric machine on the output side, and two sensors that measure torque and speed on both sides of the variator. The clamping forces of the variator are controlled by two pressure control valves, supplied by an external hydraulic source. Data is sampled at 200 Hz, but a low-pass filter with a cut-off frequency of 5 Hz is applied to signals used in the estimator. The experiments are performed with a primary speed of $\omega_p \approx 105 \text{ rad/s}$ (1000 rpm), primary torque of $T_p \approx 50 \text{ Nm}$, decreasing clamping force $F_c$ at a rate of $-26 \text{ N/s}$, and the geometric ratio is held in overdrive. Under these conditions, excessive slip (i.e., $\sigma \gtrsim 10\%$) causes severe damage to the belt-pulley contact surfaces, as will be shown next. The experiments are stopped manually, when too much slip is observed.

B. Experimental results

The estimator discussed in Section III is applied to the data obtained with three experiments. For each experiment, the same parameters are used, as listed in Table 1.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>ESTIMATOR PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>parameter value</td>
<td>$T$</td>
</tr>
<tr>
<td>l</td>
<td>$1$</td>
</tr>
</tbody>
</table>

First, an experiment will be discussed that has resulted in damaged belt-pulley contact surfaces due to excessive slip. Fig. 6 shows from top to bottom respectively, the friction ($\mu^*$), slip ($\sigma^*$), slip variance ($v$), and estimated friction slope ($\hat{k}$) as a function of time. The thresholds $(c_i, i \in \{v, k, m\})$ are shown by the dotted gray lines, whereas the friction estimate ($\hat{\mu}_{\text{max}}$) is indicated by the solid gray line. The following is observed: during the first 85 s the friction steadily increases, whereas the slip remains constant — the slip variance remains low and the friction slope estimate remains its initial value. After 85 s, the slip starts to increase, also observed in its variance and the KF starts to estimate the friction slope. Around 89 s, the friction reaches its maximum value, also observed in $\hat{k}$ ($\hat{k} \approx 0$). After 90 s, the friction rapidly decreases, while the slip rapidly increases; this is where the macro slip starts, and the experiment should be stopped. This is also indicated by the negative $\hat{k}$. The corresponding friction curve is shown in Fig. 7, where the data sets for each estimation step are indicated with different gray tints. A low-pass filtered friction curve is shown in black (with cut-off frequency 1 Hz). It is seen that the estimator is able to identify the different data sets, i.e., from left to right: (1) data with a low variance in slip ($t < 85 \text{ s}$); (2) data suitable to estimate $k$ ($85 \text{ s} < t < 89 \text{ s}$); (3) data suitable to estimate $\mu_{\text{max}}$ ($89 \text{ s} < t < 90 \text{ s}$); and (4) macro slip ($t > 90 \text{ s}$). The resulting estimate $\hat{\mu}_{\text{max}} = 0.105$ seems a reasonable estimate by visual inspection.

Fig. 6. Experimental results with macro slip after 90 s.

Fig. 7. Measured friction curve with macro slip. Data sets for each estimation step are indicated with different gray tints. The estimated maximum friction equals $\hat{\mu}_{\text{max}} = 0.105$. 

6676
Two other experiments are shown in Figs. 8 and 9, which are obtained some experiments before and after this damaging experiment, respectively. Both experiments are stopped in time; no excessive slip occurred. Although these experiments show less slip, the estimator distinguishes the different data sets, necessary to estimate $\mu_{\text{max}}$. The estimator is able to detect a decreased maximum friction after the damaging experiment: from $\tilde{\mu}_{\text{max}} = 0.105$ to $\hat{\mu}_{\text{max}} = 0.100$. Finally, it can be seen that the shape of the friction curve indeed has changed after damaging the belt-pulley contact surfaces.

VI. ACKNOWLEDGMENTS

K. van Berkel thanks prof. T. Fujii for his support during a 2.5 months visit at the Department of Mechanical Engineering of Doshisha University, Japan, where most of this research has been performed. He also thanks K. Hirano and H. Yasui for their support with the experiments.

REFERENCES