Adaptive Pulse Amplitude Pulse Width Control of Systems subject to Coulomb and Viscous Friction

Jeroen J.M. van de Wijdeven
Dept. of Mechanical Engineering
Technische Universiteit Eindhoven
Eindhoven, 5600 MB,
The Netherlands

Tarunraj Singh
Dept. of Mech. & Aero. Eng.,
University of Buffalo,
Buffalo, NY 14260
http://code.eng.buffalo.edu/tdf/

April 18, 2002

Abstract

The focus of this paper is on adaptive control of maneuvering rigid bodies in the presence of friction. The paper describes a simple technique which include Pulse Amplitude and Pulse Width modulation to progressively move the system to the desired final position. To account for uncertainty in estimated friction coefficients and approximated system model, an adaptation algorithm is necessary to accurately track the desired position. The proposed technique is suited for discrete time implementation and is illustrated on a rest-to-rest maneuver. The proposed technique is shown to considerably reduce the steady state error which exists in previously proposed Pulse Width controllers.

1 Introduction

Friction is a phenomenon which is ubiquitous. It is desirable in applications such as tires, clutches, brakes etc. and is a challenging problem when precise position control is desired. In control systems, the presence of friction can result in undesirable behavior such as limit cycling and steady state errors. It is therefore necessary that the phenomenon of friction has to be well understood, and compensated for. There exist numerous models for friction spanning the range from the simple Coulomb model [1] to the comprehensive Lund-Grenoble model [2] which accounts for effects such as stiction, Stribeck effect and hysteresis. The model which is selected for the design of controllers is application dependent.

Yoshida and Tanaka [3], proposed applying a dither signal to linearize the nonlinear friction, which in conjunction with a feedback controller is used to position a flexible arm system. Friedland and Park [4] proposed to adaptively estimate the coefficient of Coulomb friction and used the estimate in a feedforward control to cancel the effects of friction. Recently, Liao and Chien [5] modified the adaptive estimation algorithm for tracking control systems which ensures that the tracking error and the parameter errors converge to zero exponentially, in the presence of persistent excitation.

One of the more novel approaches for precise positioning using an adaptive Pulse Width Control (PWC) has been proposed by Yang and Tomizuka [6]. They consider a laboratory positioning table and study the effect of applying a pulse input on the displacement of the system subject to friction. They arrive at a closed form solution for the total displacement in the presence of Coulomb friction and show it to be a quadratic function of the pulse width. Next, they derive a closed form expression for the displacement as a function of the pulse width, including viscous friction and approximate it to make it compatible with their adaptation algorithm. When their algorithm is implemented with a discrete time controller, the discrepancy between the required pulse width and the pulse width permitted by sampled data system results in limit cycling of the closed loop response.

In this paper the technique proposed by Yang and Tomizuka [6] is modified to require the pulse width to be coincident with an integer multiple of the sampling interval. Simultaneously, the pulse amplitude is calculated to achieve the desired final displacement. An adaptation algorithm is an integral part of the proposed work. Unlike the approach proposed by Yang and Tomizuka where one parameter, which is a function of the friction coefficient, system mass and permitted pulse amplitude is estimated, the new algorithm also estimates the coefficient of Coulomb friction.

Section 2 reviews the adaptive Pulse Width Control technique proposed by Yang and Tomizuka [6]. Section 3.1 describes the variation of the parameter estimated in the work by Yang and Tomizuka as a function of varying displacement. This is the motivation for the use of an adaptation algorithm for precisely positioning the system at the desired value. This is followed by the description of the adaptive Pulse Amplitude Pulse Width Control (PAPWC) in Section 3.2. Section 4 illustrates the proposed technique on a simple rest-to-rest maneuver and compares its performance with adaptive PWC. The paper concludes with some remarks in Section 5.

2 Adaptive Pulse Width Control

2.1 System Model

The basic idea and motivation for the use of pulse width control (PWC) was introduced by Yang and Tomizuka [6]. With the assumption that the sampling frequency is much lower than the first resonance peak, their model of a X-Y table can be reduced to a single mass subject to friction. The friction that is acting on the mass m is assumed to consist of Coulomb friction f_c, stiction f_s and viscous damping c. The differential equations for this model are:

$$\ddot{x} = \begin{cases} \frac{1}{m}(u - f_c - c\dot{x}) & \text{if } \dot{x} \neq 0 \\ 0 & \text{if } \dot{x} = 0 \text{ and } |u| \leq f_s \\ \frac{1}{m}(u - sgn(u) f_s) & \text{if } \dot{x} = 0 \text{ and } |u| > f_s. \end{cases} \tag{1}$$

If the system described in Eq.(1), is driven by a single pulse with pulse height f_p and pulse width t_p, a closed form expression for the displacement can be derived. In [6] this is done for the model with and without viscous damping. The results are given by:

$$d = \frac{f_p(f_p - f_c)}{2mf_c} t_p^2 \quad \text{for } f_p > 0 \tag{2}$$

$$d = \frac{f_p t_p}{c} - \frac{mf_c}{c^2} ln\left[\frac{f_p}{f_c}\left(e^{ct_p/m} - 1\right) + 1\right]. \tag{3}$$

Assuming that the displacement is small, it was shown [6] that viscous damping can be neglected compared to the Coulomb friction. Thus, Eq.(2) is a reasonable approximation for small displacement. The displacement now is linearly proportional to the square of the pulse width. The coefficient of the term t_p^2 is represented by one parameter b, with $b \geq 0$. The final expression for the displacement is given as:

$$d(t_p) = bt_p^2 sgn(f_p), \quad \text{with } b = \frac{f_p(f_p - f_c)}{2mf_c}. \tag{4}$$

Since the direction of the displacement must be equal to the sign of f_p, Eq.(4) includes $sign(f_p)$.

2.2 Adaptation Algorithm

The control scheme that is presented for the system has two components. The first one is a simple feedback controller used in conjunction with a feedforward controller to compensate for the Coulomb friction force.

This controller is used to move the system from its initial position to the vicinity of the desired position. Once the system sticks within an error tolerance area around the reference position, the controller switches from feedback control to the second component, the PWC. The feedback controller must be designed in such a way that the maximum steady state error is smaller than a predefined error tolerance.

The input to the PWC is the error e between the desired position and the current position and is used to calculate the pulse width. In [6] two equations are derived to accomplish the above, which are:

$$\begin{aligned} d(k+1) &= b u_p(k) \qquad (5) \\ \text{with } u_p(k) &= t_p^2(k)\ sgn\left(f_p(k)\right) \\ u_p(k) &= \frac{K_c}{b} e(k). \qquad (6) \end{aligned}$$

In Eq.(6), K_c is a control parameter with $0 < K_c < 2$ for stability reasons. k stands for the k^{th} pulse and should not be mistaken with the sampling time.

Since b is not known exactly, the pulse width cannot be derived from Eq.(6). Instead t_p is calculated with an estimate $\hat{b}$ of b. With the use of an adaptation algorithm $\hat{b}$ is updated after each pulse. Although [6] presents multiple ways to estimate $\hat{b}$ and $1/\hat{b}$, this paper only presents the self tuning regulator approach for estimating $\hat{b}$. The adaptation algorithm is given by the equations:

$$\begin{aligned} \epsilon_0^0(k) &= d(k) - \hat{b}(k-1) u_p(k-1) \qquad (7) \\ F^{-1}(k) &= \lambda_1 F^{-1}(k-1) + \lambda_2 u_p^2(k-1) \qquad (8) \\ \hat{b}(k) &= \hat{b}(k-1) + F(k) u_p(k-1) \epsilon_0^0(k) \qquad (9) \\ & F(0) > 0,\ 0 < \lambda_1 \leq 1,\ 0 \leq \lambda_2 < 2. \end{aligned}$$

ϵ_0^0 represents the error between the real displacement d and the estimated displacement $\hat{b}u_p$. F is referred to as the time-varying gain matrix and λ_1 and λ_2 are parameters related to forgetting previous data. For λ_1 and λ_2 equal to one, all data is equally weighted.

The adaptation algorithm results in pulse widths that can take any positive value. However, in a discrete time system, the pulse width will automatically be rounded to the smallest integer number of the sampling time larger than the calculated pulse width. This is caused by the D/A converter that is assumed to be Zero Order Hold (ZOH). In this way, the calculated pulse isn't the real input pulse on the system.

In [6] it is assumed that the pulse widths are relatively small, so that the pulse width calculated using Eq.(4) results in the desired displacement. However, when the required displacement increases, the approximation for the displacement described in Eq.(4) isn't accurate anymore. The impact of b on the final displacement of the PWC, when the required motion is large, needs to be studied.

In section 3 a technique to address the issue of the finite sampling time on the control performance is proposed.

3 Adaptive Pulse Amplitude Pulse Width Control

In this section, the influence of relatively large displacements on the estimation of $\hat{b}$ is studied. Next, the original PWC algorithm is modified, such that only integer number of the sampling time are used to describe the pulse width. The consequences of this on the algorithm will be analyzed.

3.1 Motivation for Adaptation

Figure 1 illustrates the displacement as a function of the pulse width t_p. The solid line corresponds to the displacement, described by Eq.(3), the dotted line corresponds to the Eq.(4) which is derived by ignoring viscous friction. In this plot it is assumed that all the parameters are known. Their values are presented in Table 1.

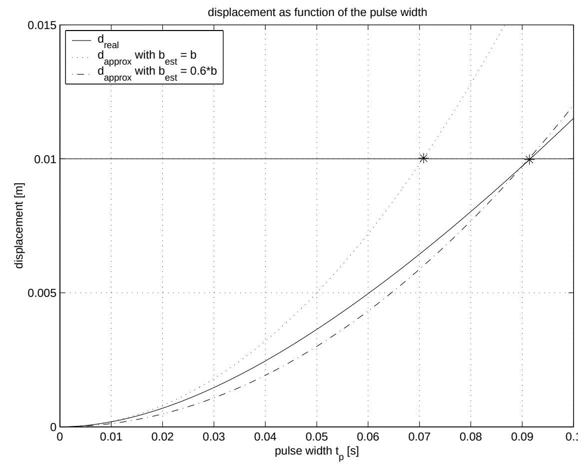

Figure 1: Displacement d as function of the pulse width t_p

Table 1: Parameter values

m	150 kg
f_c	100 N
c	700 kg/s
f_s	130 N
f_p	300 N
T	0.01 s

For a displacement of 0.01 m, Eq.(3) results in a pulse width of 0.092 s, while the approximate Eq.(4), results in a t_p of 0.070 s.

In order to find the appropriate t_p, the adaptation algorithm modifies $\hat{b}$ such that the dash-dotted line in figure 1, which corresponds to Eq.(4), coincides with the exact curve (solid line), for the specific desired displacement. For example, for a displacement of 0.01 m, the $\hat{b}$ which forces the curve to correspond to Eq.(4) is $0.6b$.

To prevent the large fluctuations in the estimate of $\hat{b}$, due to large difference in displacements, t_p is constrained to a maximum. If a maximum for t_p=$t_{p,max}$ of 0.03 s is chosen, the estimate $\hat{b}$ is between 81% and 100% of that of the theoretical b. The lower percentage is a function of the parameters of the system.

A point to be noted is that, if the feedback controller results in the same error, every time this controller is used, there's no need for a $t_{p,max}$. The estimated $\hat{b}$ will not converge to the theoretical b, rather, the adaptation algorithm will estimate the $\hat{b}$ in such a way, that only one pulse is needed to move the system from its initial position to the desired position.

If for some reason the system isn't at the reference position after the first pulse, more pulses have to be used. The pulse widths of these subsequent pulses result in the overshoot of the system response. If for example, the residual error is 2.0 mm, the dash-dotted line will correspond to a pulse width of 41 ms, while the real pulse width should be 36 ms. With the t_p calculated from the dash-dotted line, the true displacement is 2.5 mm instead of the desired 2.0 mm and thus overshoot occurs.

As the residual error decreases with subsequent pulses, the difference between the estimated and real displacement (ϵ_0^0) is small resulting in minimal adaptation of $\hat{b}$.

For a finite $t_{p,max}$, the difference between $\hat{b}$ and b is smaller than without limits on the pulse width, resulting in a better estimation of the displacement. Thus, the adaptation algorithm estimates $\hat{b}$ more precisely, and overshoot is less likely to occur. A drawback of introducing a maximum pulse width is that the system takes longer to reach the final position.

3.2 Adaptive PAPWC

The pulse width resulting from Eq.(4) should be transformed to be equal to an integer times the sampling time (nT), so that it can be implemented in a discrete time system. The pulse width needs to be rounded to the higher integer multiple of T, i.e. $T_x = nT$, so that the corresponding f_p^* is smaller than or equal to the maximum permitted pulse amplitude. This is the starting point for the transformation, proposed in this section.

The displacement in Eq.(4) is dependent on the sign of f_p. Note that $sign(f_p)$ can be replaced by $sign(e)$ without changing the outcome of Eq.(4). The desired displacement must be the same for both t_p and T_x. This results in:

$$\begin{aligned} d\,(t_p) &= d\,(T_x) \\ b\; t_p^2 sgn(e) &= b^*\; T_x^2 sgn(e) \\ \frac{f_p(f_p - f_c)}{2mf_c} t_p^2 sgn(e) &= \frac{f_p^*(f_p^* - f_c)}{2mf_c} T_x^2 sgn(e). \end{aligned} \tag{10}$$

The constant b must change to b^*, in order to satisfy Eq.(10). This can only be done by varying the pulse height f_p^*, since this is the only free parameter in b^*. Solving Eq.(10) for f_p^* results in:

$$f_p^* = 0.5 f_c \pm 0.5 \sqrt{f_c^2 + 4 f_p (f_p - f_c) \frac{t_p^2}{T_x^2}}. \tag{11}$$

Instead of using t_p and f_p, the control pulse is specified by T_x and f_p^*, where $\pm$ should be replaced by $+$ for a positive pulse height. The above algorithm works, if the Coulomb friction coefficient f_c is known. But unfortunately that isn't the case.

In order to solve this problem, we propose to divide b^* into two parameters:

$$\begin{aligned} b^* &= A_1 f_p^2 - A_2 f_p \\ \text{with } A_1 &= \frac{1}{2mf_c} \text{ and } A_2 = \frac{1}{2m} \\ a &= \begin{bmatrix} A_1 & A_2 \end{bmatrix}^T \end{aligned} \tag{12}$$

From the estimates $\hat{A}_1$ and $\hat{A}_2$, an estimate of f_c can be derived by dividing $\hat{A}_2$ by $\hat{A}_1$. Using $\hat{f}_c$ in Eq.(11), the pulse height is calculated. The next pulse width, is calculated from Eq.(4) using the $\hat{b}$ obtained from $\hat{A}_1$ and $\hat{A}_2$.

In [7] an adaptation algorithm is proposed, that doesn't need to invert F^{-1}. Now that there are two parameters to be estimated, this algorithm can save computation time. The algorithm is given by the equations:

$$\begin{aligned}
\pi(k) &= P(k-1)u(k) & (13)\\
P(k) &= \lambda^{-1}P(k-1) - \lambda^{-1}K(k)u^H(k)P(k-1) & (14)\\
K(k) &= \frac{\pi(k)}{\lambda + u^H(k)\pi(k)} & (15)\\
\xi &= d(k) - \hat{a}^H(k-1)u(k) & (16)\\
\hat{a}(k) &= \hat{a}(k-1) + K(k)\xi^H(k) & (17)
\end{aligned}$$

$$u(k) = \left[f_p^2 t_p^2 \quad f_p t_p^2\right]^T \tag{18}$$

The superscript H stands for the Hermitian transpose. In this case, all values are real and thus H can be replaced by T. u, d and ξ in Eq.(13)-(17) have the same interpretation as u_p, d and ϵ_0^0 respectively, in Eq.(7)-(9). In [7] K is described as the time-varying gain vector and P as the inverse correlation matrix. The λ in the algorithm represents the forgetting factor. Generally the forgetting factor is selected to be smaller than and close to 1 and is set to 1 if all data is to be equally weighted.

When applied λ should have a value close to, but just less than one. If the forgetting factor is not wanted, λ should be set to one.

For the initiation of the adaptation algorithm, initial conditions for $P(0)$, $\hat{A}_1(0)$and $\hat{A}_2(0)$ are required. $\hat{A}_1(0)$ and $\hat{A}_2(0)$ should be chosen as accurately as possible and the inverse covariance matrix $P(0)$ is set to $P(0) = \delta^{-1}I$. δ should be large if the sensors are noisy and is small otherwise.

4 Numerical Simulations

For the system shown in Figure 2, with parameters listed in Table 1, the proposed algorithm is simulated and compared to the adaptive PWC. In Figure 2 the Coulomb friction and stiction are represented by f.

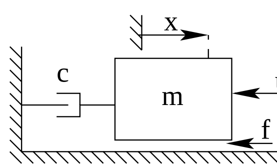

Figure 2: Model of the system

The forgetting factor is set to 1, so there is no forgetting of data. The initial value for $F^{-1}(0)$ and δ^{-1} are both $1e^{-5}$, since there's no noise present in the system. In order to compare the two algorithms, the initial estimations for $\hat{b}(0)$ and $\hat{b}(0) = f\left(\hat{A}_1(0), \hat{A}_2(0)\right)$ are the same. This is done by calculating both $\hat{b}(0)$'s assuming that m is exactly known and $f_{c,\ est}$ is some percentage of the the real f_c. This provides a fair means of comparing $\hat{b}(0)$, $\hat{A}_1(0)$ and $\hat{A}_2(0)$. The initial value for $f_{c,\ est}$ is set to $0.8 f_c$, so that $\hat{b}(0)$ is 1.375 b.

Figure 3 and 4 illustrate the evolution of the position errors for the algorithm proposed by [6] and from the adaptive PAPWC algorithm respectively. Both the figures present the error profiles for the first iteration and the 10th iteration respectively.

From Figure 3, it can be seen that for the case when $t_{p,max}$ is not bounded, overshoot occurs, as predicted. However, for the case when $t_{p,max}$ is bounded, overshoot continues to exist. This is caused by a minimum possible displacement due to a minimum pulse width of one sample time. The theoretical minimum displacement can be found using

$$d_{min}(T) = \frac{f_p T}{c} - \frac{m f_c}{c^2} ln\left[\frac{f_p}{f_c}\left(e^{cT/m} - 1\right) + 1\right]. \tag{19}$$

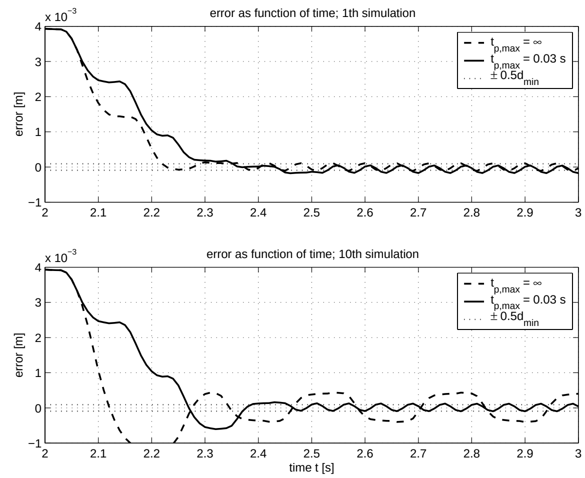

Figure 3: Error as a function of time using the original PWC approach

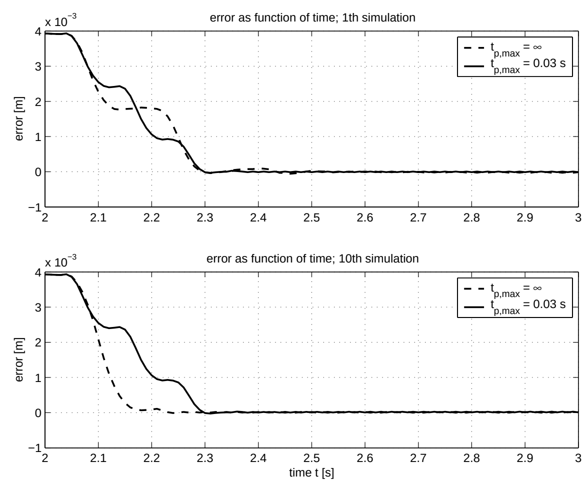

Figure 4: Error as a function of time using adaptive PAPWC

Since d_{min} is equal to the maximum final error, d_{min} causes a limit cycle around the reference point.

In the second plot of Figure 3, the overshoot for $t_{p,max} = \infty$ is larger than the displacement given by Eq.(19). This is caused by the fact that the pulse width now is $2T$.

In Figure 4, one can see that the the final error is around zero. Due to Coulomb friction, fluctuations with small velocities around zero result in numerical problems in Simulink. To avoid this situation an error tolerance of $2e^{-5}m$ is introduced. If the error is smaller than this value, the control input is set to zero and the system will stay at rest. The final results thus have a maximum allowed value of $2e^{-5}m$.

Although much smaller, adaptive PAPWC also results in a limit cycle. This occurs when $f_p^* = f_s$. The minimum displacement is given by:

$$d_{min}(T) = \frac{f_s T}{c} - \frac{m f_c}{c^2} ln \left[\frac{f_s}{f_c} \left(e^{cT/m} - 1 \right) + 1 \right]. \tag{20}$$

To compare the two presented minimum displacements, their values have been calculated using Table 1. PWC results in a d_{min} of $1.9e^{-4}m$, while PAPWC has a d_{min} of $1.3e^{-5}m$. There is thus a reduction in the final error of a factor 15.

In the second plot of Figure 4 for $t_{p,max} = \infty$, one can note that theoretically only one pulse is needed to get the system very close to zero error. The final error from the first pulse will practically never be zero, since a second pulse which needs to move the system by a small amount will marginally modify $\hat{b}$ such that in the next iteration, the first pulse cannot satisfy the desired motion.

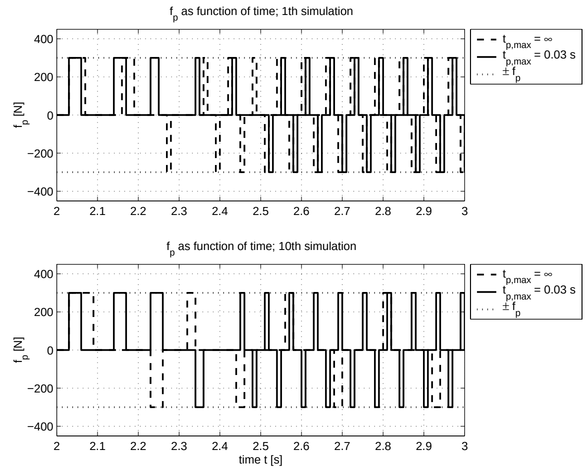

Figure 5: f_p As a function of time using the original PWC approach

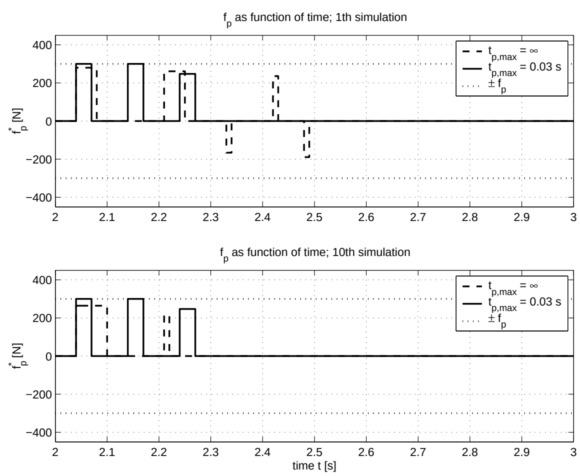

Figure 6: f_p As a function of time using adaptive PAPWC

The pulse sequences for the adaptive PWC and PAPWC are presented in Figures 5 and 6 respectively. Figure 5 illustrates that the input f_p into the system contains only maximum positive and negative pulses. The switching between the positive and negative pulses is caused by the limit cycle.

Figure 6 reveals that the pulses die out after a short period. If a $t_{p,max}$ is used, only positive pulses drive the system. Without a bound on $t_{p,max}$, overshoot occurs, which results in subsequent negative pulses. After ten iterations, the unbounded pulse width control sequence, needs two pulses to reach the reference position, while the input with the bounded pulse width hasn't changed much.

The estimates for $\hat{b}$ are plotted in Figures 7 and 8.

Although both algorithms have used the same initial value for the inverse covariance matrix, it can be seen that the adaptation speed of the adaptive PAPWC is much faster than that of the adaptive PWC.

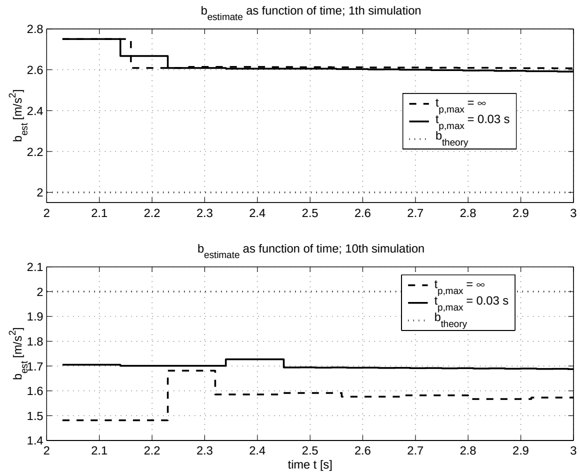

Figure 7: $\hat{b}$ As a function of time using the original PWC approach

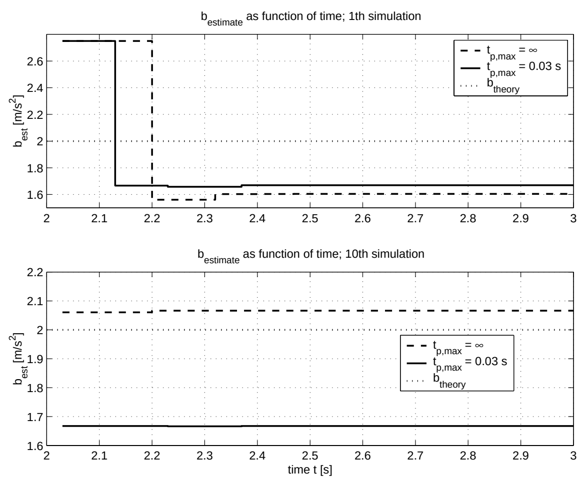

Figure 8: $\hat{b}$ As a function of time using PAPWC

In the second plot of Figure 7, it can be seen that the estimate for $\hat{b}$ with bounded $t_{p,max}$ is closer to the theoretical b than for the unbounded $t_{p,max}$. This confirms earlier findings.

The estimate for $\hat{b}$ for the case that $t_{p,max} = \infty$ in the second plot of Figure 8 shows some unexpected results. The estimated value is larger than the theoretical one. Looking at the simulation results, only one large pulse is driving the system. After this first pulse, only a small error in position is left. Small position errors result in a relatively small ξ. Looking at Eq.(17), a small ξ results in almost no update of the parameters.

Since there are two parameters to be estimated from one equation, only the ratio between the two parameters can be derived. Apparently the adaptation algorithm found values for $\hat{A}_1$ and $\hat{A}_2$ that result in an unexpected value for $\hat{b}$ that still result in the desired displacement. If the sampling time is decreased, the system returns to the situation where $\hat{b}$ is smaller than the theoretical b.

The final results are presented in Figure 9. This plot presents the estimated Coulomb friction.

From Eq.(12), one can conclude that $\hat{b}$ increases when $\hat{f}_c$ decreases and vice versa. The results that are shown are thus to be expected, after seeing Figure 7 and 8.

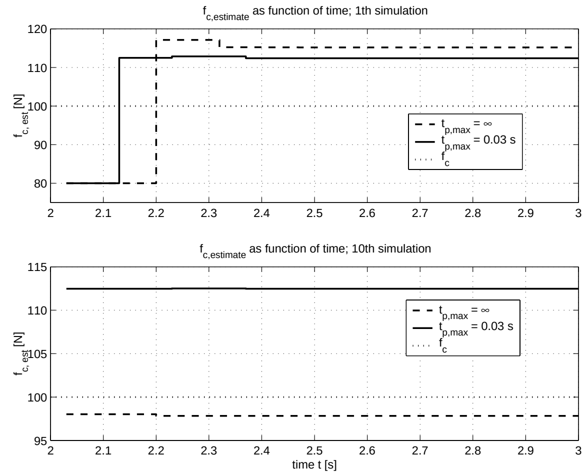

Figure 9: $\hat{f}_c$ As a function of time using adaptive PAPWC

When simulations are done *without* updating the inverse covariance matrix, the values for $\hat{b}$ and $\hat{f}_c$ at the tenth simulation are fluctuating more. This results in more overshoot in the error signal and more sign switching of the input pulses. This however doesn't influence the final error significantly.

5 Conclusions

This paper proposed a modification to the Pulse Width Control technique by forcing the input pulse width to be coincident with an integral multiple of the sampling period. This, in conjunction with adaptation of the pulse amplitude results in the Pulse Amplitude Pulse Width Controller. To account for uncertainties in the estimated system parametes, an adaptation algorithm is used to estimate the coefficient of friction which is subsequently used by the PAPWC.

Numerical simulations illustrate the significant reduction in the steady state error which corresponds to a reduction in amplitude of the limit cycle. The adaptive PAPWC controller will be implemeted on an experimental testbed in the future and extended to systems with low frequency resonance.

References

[1] C. A Coulomb. Theorie des machines implies, en ayant egard au frottement de leurs parties, et a la roideur dews cordages. *Mem. Math Phy.*, pages 161–342, 1785.

[2] C. Canudas de Wit, H. Olsson, K. J. Åström, and P. Lischinsky. A new model for control of systems with friction. *IEEE Transactions on Automatic Control*, 40(3):419–425, 1995.

[3] Y. Yoshida and M Tanaka. Position control of a flexible arm using a dither signal. *JSME International Journal*, Series C, 36(1):93–99, 1993.

[4] B. Friedland and Y-J. Park. On adaptive friction compensation. *IEEE Transactions on Automatic Control*, 37(10):1609–1612, October 1992.

[5] T-L. Liao and T-I. Chien. An exponentially stable adaptive friction compensator. *IEEE, Transactions on Automatic Control*, 45(5):977–980, May 2000.

[6] S. Yang and M. Tomizuka. Adaptive pulse width control for precise positioning under the influence of stiction and coulomb friction. *Journal of Dynamic Systems, Measurements, and Control*, Vol. 110:221–227, September 1988.

[7] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, Inc., fourth edition, 2002.