Adaptive friction approximation

Lars

April 28, 2009

1 Introduction

We assume density measured or estimated on a slower time-scale. Therefore, this note considers estimation of friction only.

2 Function approximation

The approach to adaptive function approximation we will employ, is a 'local learning/neuro/fuzzy'-type approach [1]. We approximate (an unknown) function $h(x)$ with a normalized weighted average of N local approximators $\hat{h}_k(x)$. That is, the approximator $\hat{h}(x)$ is given as

$$\hat{h}(x) = \sum_{k=1}^{N} \phi_k(x)\hat{h}_k(x) = \Phi(x) \begin{pmatrix} \hat{h}_1(x) \\ \hat{h}_2(x) \\ \vdots \\ \hat{h}_N(x) \end{pmatrix},$$

where $\hat{h}_i(x)$ are local approximators, and we use 'basis functions' $\phi_i(x)$ that (for each x) forms a 'partition of unity',

$$\phi_i(x) = \frac{\omega_i(x)}{\sum_{k=1}^{n_\theta} \omega_k(x)}.$$

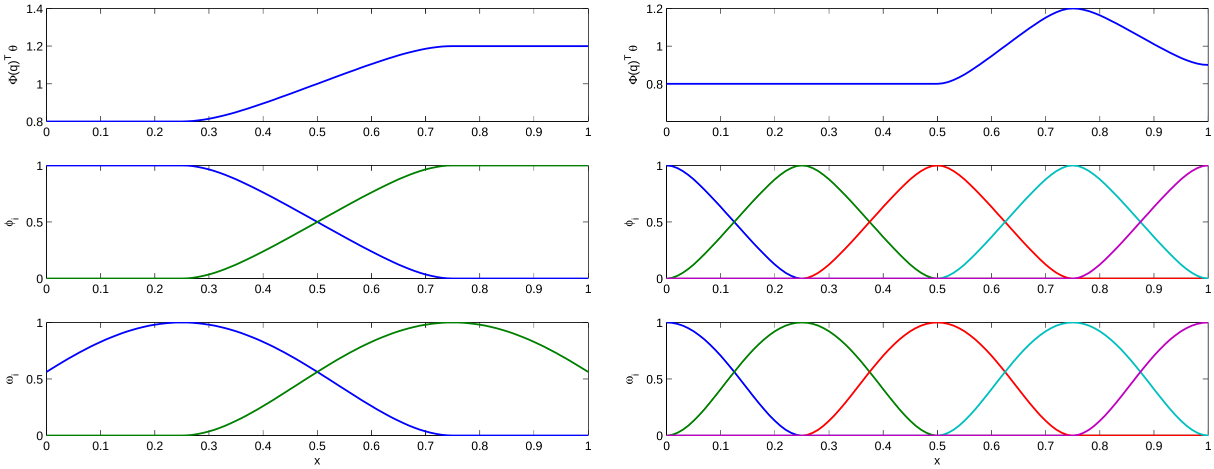
The function $\omega_k(x)$ is the local weighting function, or local support. Examples of such functions can be

$$\omega_i(x) = \begin{cases} \left(1 - \left(\frac{|x-c_i|}{\mu_i}\right)^2\right)^2, & \text{if } |x-c_i| < \mu_i, \\ 0, & \text{otherwise.} \end{cases}$$

Here, c_i is the center of the 'local support' of the local approximator $\hat{h}_i(x)$, while μ_i describes the area (radius).

Typically, the local approximators will be parameterized, $\hat{h}_k(x) = \hat{h}_k(x, \theta)$. The simplest example is that each local approximator is a constant, denoted θ_k, and that all of these are collected in a vector $\theta = (\theta_1, \theta_2, \ldots, \theta_N)$. In that case, the approximating function is $\Phi(x)^\mathsf{T}\theta$, and is plotted in Figure 1 for $N = 2$ and $N = 5$, for 'typical' choices of local weighting functions.

Conditions for when this class of approximators is a universal approximator for a class of functions, can be found [1].



(a) $N=2$, $c_i=\{0.25, 0.75\}$, $\mu_i=0.5$, $\theta_1=0.8$, $\theta_2=1.2$.

(b) $N=5$, $c_i=\{0, 0.25, 0.5, 0.75, 1\}$, $\mu_i=0.25$, $\theta_1=\theta_2=\theta_3=0.8$, $\theta_4=1.2$, $\theta_5=0.9$.

Figure 1: Approximation

3 Approximating the friction

Assume 'real friction' is a function of q (and of time, slowly-varying – disregard that for now), $\mathcal{F}(q)$.

Assume we have a model for $\mathcal{F}(q)$, denote this $F(q)$. Since $F(\cdot) \neq \mathcal{F}(\cdot)$, we parameterize model error with a parameter vector $\theta \in \mathbb{R}^{n_\theta}$ and accompanying basis functions, $\Phi(q) = (\phi_1(q), \phi_2(q), \ldots, \phi_{n_\theta}(q))^\mathsf{T}$. Corresponding to the above, we use $\hat{h}_k(q) = F(q)\theta_k$, and $N = n_\theta$. We assume n_θ large enough such that we can use this to construct a universal approximator, $F(q)\Phi(q)^\mathsf{T}\hat{\theta}$, where there exist an (unknown) θ such that

$$F(q)\Phi(q)^\mathsf{T}\theta = \mathcal{F}(q).$$

We will use $f(q)$ for the vector function $\Phi(q)F(q)^\mathsf{T}$. Specifically, we will assume annulus friction is $f_a^\mathsf{T}(q)\theta_a$ and drillstring friction is $f_d^\mathsf{T}(q)\theta_d$.

As an example, we plot a model of drillstring friction together with an approximation with $n_\theta = 2$, in Figure 2.

4 Simple adaptive scheme based on flow measurement

4.1 Case 1: Information from surface measurements only

A dynamic equation for bit flow ($q = Mq_{bit}$) is

$$\dot{q} = p_p - p_c - f_d^\mathsf{T}(q)\theta_d - f_a^\mathsf{T}(q)\theta_a + s(t), \tag{1}$$

where $s(t)$ is a possibly time-varying function describing difference in static head. We will assume that q is measured and bounded.

Even though we assume q measured, we suggest an observer for q,

$$\dot{\hat{q}} = p_p - p_c - f_d^\mathsf{T}(q)\hat{\theta}_d - f_a^\mathsf{T}(q)\hat{\theta}_a + s(t) + k\left(q - \hat{q}\right), \tag{2}$$

where we aim (with reservations regarding PE) to find update laws for $\hat{\theta} = \left(\hat{\theta}_a^\mathsf{T}, \hat{\theta}_d^\mathsf{T}\right)^\mathsf{T}$ such that $\hat{\theta} \to \theta$.

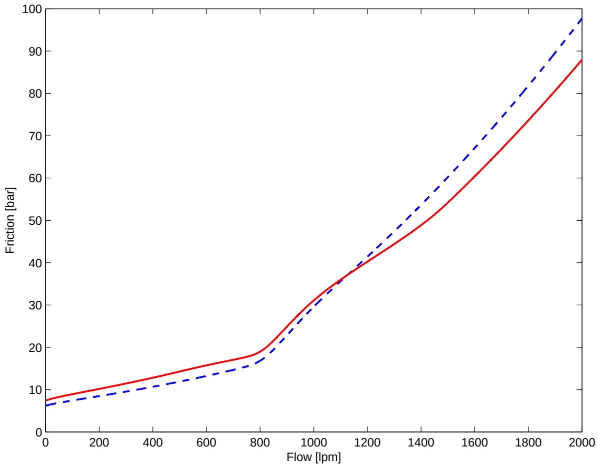


Figure 2: A model of friction (dashed line) with a two-parameter approximation (solid line). The approximation is for $\theta_1 = 1.2$ and $\theta_2 = 0.9$.

Define error variables $\tilde{q} = q - \hat{q}$, $\tilde{\theta} = \theta - \hat{\theta}$, and consider the Lyapunov-like function $V(\tilde{q}, \tilde{\theta}) = \frac{1}{2}\left(\tilde{q}^2 + \tilde{\theta}_a^\mathsf{T}\Gamma_a^{-1}\tilde{\theta}_a + \tilde{\theta}_d^\mathsf{T}\Gamma_d^{-1}\tilde{\theta}_d\right)$ where Γ_a and Γ_d are positive definite. The derivative of this function is

$$\begin{aligned}\dot{V} &= \tilde{q}\dot{\tilde{q}} + \tilde{\theta}_a^\mathsf{T}\Gamma_a^{-1}\dot{\tilde{\theta}}_a + \tilde{\theta}_d^\mathsf{T}\Gamma_d^{-1}\dot{\tilde{\theta}}_d \\ &= -k\tilde{q}^2 - \tilde{q}f_d^\mathsf{T}(q)\tilde{\theta}_d - \tilde{q}f_a^\mathsf{T}(q)\tilde{\theta}_a + \dot{\tilde{\theta}}_d^\mathsf{T}\Gamma_d^{-1}\tilde{\theta}_d + \dot{\tilde{\theta}}_a^\mathsf{T}\Gamma_a^{-1}\tilde{\theta}_a.\end{aligned}$$

Selecting parameter update laws

$$\begin{aligned}\dot{\hat{\theta}}_a &= -\Gamma_a f_a(q)\left(q - \hat{q}\right), \\ \dot{\hat{\theta}}_d &= -\Gamma_d f_d(q)\left(q - \hat{q}\right),\end{aligned}$$

we obtain

$$\dot{V} = -k\tilde{q}^2,$$

and hence we can conclude $\tilde{q} \to 0$ and boundedness of parameter estimates by standard arguments (Barbalat's lemma) under some technical assumptions on q and the friction functions.

However, we want more than boundedness; we ideally want parameter convergence for the parameters that have local support. For this, we need to examine the information content in the data; i.e., whether the data is persistently exciting (PE).

Note that the parameters that does not have local support (that is, $f(q) = 0$) will not be updated. Therefore, a PE-analysis must take this into account.

4.2 PE analysis of case 1

PE analysis of the above estimation problem is straightforward as the error system is essentially LTV and we can resort to standard theory (e.g. [3, 2]). The closed loop error system can be written as

$$\begin{pmatrix}\dot{\tilde{q}} \\ \dot{\tilde{\theta}}\end{pmatrix} = \begin{pmatrix}A & B\phi(t)^\mathsf{T} \\ -\phi(t)C^\mathsf{T} & 0\end{pmatrix}\begin{pmatrix}\tilde{q} \\ \tilde{\theta}\end{pmatrix}, \quad \text{where} \quad \tilde{\theta} = \begin{pmatrix}\sqrt{\Gamma_a^{-1}}\tilde{\theta}_a \\ \sqrt{\Gamma_d^{-1}}\tilde{\theta}_d\end{pmatrix}, \quad \phi(t) = \begin{pmatrix}-\sqrt{\Gamma_a}f_a(q) \\ -\sqrt{\Gamma_d}f_d(q)\end{pmatrix},$$

and $A = -k$ and $B = C = 1$. Since (A, B, C) is SPR, this system is (globally) exponentially stable if we have PE, that is, if there exists α such that at all time instants t, there exists a T such that

$$\int_t^{t+T} \phi(\tau)\phi^{\mathsf{T}}(\tau)d\tau \geq \alpha I.$$

If q varies within a small window, then only two parameters, one in each of θ_a and θ_d, is being estimated. The interesting question is: Is there enough variation to estimate both of these? Assume, as an approximation to this case, that $n_{\theta_a} = n_{\theta_d} = 1$, that is $f_a(q)$ and $f_d(q)$ are scalars. The question is then, does $f_a(q)$ and $f_d(q)$ vary "sufficiently differently", that is, will the matrix

$$\int_t^{t+T} \begin{pmatrix} f_a(q) \\ f_d(q) \end{pmatrix} \begin{pmatrix} f_a(q) \\ f_d(q) \end{pmatrix}^{\mathsf{T}} d\tau$$

become non-singular? Since both signals depend in a similar manner on q, this requires significant variations in q that really excite the nonlinearities in $f_a(q)$ and $f_d(q)$ differently. In practice, this is unlikely to happen apart from short time-periods (pipe connections), which probably are too short to obtain parameter convergence.

In practice, this means that there is only information to estimate one parameter, and this parameter will only converge to its true value provided the other parameter is set correctly. However, it is easy to see that the total friction is estimated correctly in steady state. Since $\tilde{q} \to 0$, when $\dot{\hat{q}} = 0$ (in steady state), then, from (2),

$$f_d^{\mathsf{T}}(q)\hat{\theta}_d + f_a^{\mathsf{T}}(q)\hat{\theta}_a = p_p - p_c + s(t).$$

Since the expression on the right hand side is the total pressure loss due to friction, the overall friction is estimated correctly.

In conclusion, it is only realistic to estimate one part of the friction, and assume the other known. That is either assume friction in the drillpipe known, and estimate the parameters related to annulus friction, or the other way around. The fact that annulus friction in general is the 'most unknown' friction, speaks in favor of estimating annulus friction. On the other hand, the following issues,

- drillpipe friction often is about an order of magnitude larger,
- annulus friction influences the main variable of interest (bottomhole pressure) directly, and hence we should be careful with updating annulus friction without using measurements of bottomhole friction,

points towards estimating only drillpipe friction as long as bottomhole pressure is not available.

However, an even better option is probably to adapt the total friction, and use a predefined distribution formula to update annulus and drillpipe friction parameters.

4.3 Case 2: Bottomhole pressure available

Consider the system (1), but assume now that we have available the additional measurement p_b,

$$p_b = p_c + f_a^{\mathsf{T}}(q)\theta_a + s_a(t). \tag{3}$$

We incorporate this information by extending the observer (2) for q in the following manner:

$$\dot{\hat{q}} = p_p - p_c - f_d^{\mathsf{T}}(q)\hat{\theta}_d - f_a^{\mathsf{T}}(q)\hat{\theta}_a + s(t) + k_q(q - \hat{q}) + k_p(p_b - \hat{p}_b), \tag{4}$$

where

$$\hat{p}_b = p_c + f_a^{\mathsf{T}}(q)\hat{\theta}_a + s_a(t).$$

This gives the following error dynamics:

$$\dot{\tilde{q}} = -f_d^\mathsf{T}(q)\tilde{\theta}_d - f_a^\mathsf{T}(q)\tilde{\theta}_a - k_q\tilde{q} - k_p\tilde{p}_b.$$

We first note that since $\tilde{p}_b = -f_a^\mathsf{T}(q)\tilde{\theta}_a$, it is straightforward to extend the design in Section 4.1 to obtain

$$\dot{\hat{\theta}}_a = \Gamma_a(1-k_p)f_a(q)\left(q-\hat{q}\right).$$

The information in the bottomhole pressure is now used indirectly (through $\hat{q}$) to affect the parameter estimates (in θ_a). It is of interest to see if we can obtain parameter update laws that use the information more explicitly.

We therefore propose

$$\begin{aligned}
\dot{\hat{\theta}}_a &= -\Gamma_a(1-k_p)f_a(q)\left(q-\hat{q}\right) + K_a\left(p_b - \hat{p}_b\right),\\
\dot{\hat{\theta}}_d &= -\Gamma_d f_d(q)\left(q-\hat{q}\right),
\end{aligned}$$

where matrix K_a has a specific form, $K_a = \Gamma_a f_a(q)k_a$. This gives the tuning parameters Γ_a, Γ_b, k_q, k_p, and k_a.

Considering the same Lyapunov function as before, we obtain

$$\dot{V} = -k_q\tilde{q}^2 - \tilde{q}f_d^\mathsf{T}(q)\tilde{\theta}_d - \tilde{q}f_a^\mathsf{T}(q)\tilde{\theta}_a - k_p\tilde{q}\tilde{p}_b + \tilde{\theta}_a^\mathsf{T}\Gamma_a^{-1}\dot{\tilde{\theta}}_a + \tilde{\theta}_d^\mathsf{T}\Gamma_d^{-1}\dot{\tilde{\theta}}_d$$

Inserting the parameter update laws and using $\tilde{p}_b = -f_a^\mathsf{T}(q)\tilde{\theta}_a$,

$$\begin{aligned}
\dot{V} &= -k_q\tilde{q}^2 - \tilde{q}f_d^\mathsf{T}(q)\tilde{\theta}_d - \tilde{q}f_a^\mathsf{T}(q)\tilde{\theta}_a + k_p\tilde{q}f_a^\mathsf{T}(q)\tilde{\theta}_a + \tilde{\theta}_a^\mathsf{T}\Gamma_a^{-1}\left(\Gamma_a(1-k_p)f_a(q)\tilde{q} - K_a f_a^\mathsf{T}(q)\tilde{\theta}_a\right) + \tilde{\theta}_d^\mathsf{T}\Gamma_d^{-1}\Gamma_d f_d(q)\tilde{q}\\
&= -k_q\tilde{q}^2 - \tilde{\theta}_a^\mathsf{T}\Gamma_a^{-1}K_a f_a^\mathsf{T}(q)\tilde{\theta}_a.
\end{aligned}$$

Now, since $K_a = \Gamma_a f_a(q)k_a$, we get

$$\dot{V} = -k_q\tilde{q}^2 - k_a\left(f_a^\mathsf{T}(q)\tilde{\theta}_a\right)^2,$$

allowing us to conclude that $\tilde{q} \to 0$ and $f_a^\mathsf{T}(q)\tilde{\theta}_a \to 0$ (using Barbalat's lemma, provided the derivatives of q and $f_a(q)$ are bounded).

4.4 PE analysis of case 2

Since we now have a measurement equation involving (some of) the unknown measurements, the analysis is not entirely standard anymore. Define

$$\mathcal{A}(t) = \begin{pmatrix} -k_q & -(1-k_p)f_a^\mathsf{T}(q) & -f_d^\mathsf{T}(q) \\ (1-k_p)\Gamma_a f_a^\mathsf{T}(q) & -k_a\Gamma_a f_a(q)f_a^\mathsf{T}(q) & 0 \\ \Gamma_d f_d^\mathsf{T}(q) & 0 & 0 \end{pmatrix}, \quad \mathcal{C}(t) = \begin{pmatrix} \sqrt{k_q} & 0 & 0 \\ 0 & \sqrt{k_a}f_a^\mathsf{T}(q) & 0 \end{pmatrix},$$

Then, the closed loop system is given by $\dot{x} = \mathcal{A}(t)x$, and the derivative of the Lyapunov function is $\dot{V}(x) = -x^\mathsf{T}\mathcal{C}^\mathsf{T}(t)\mathcal{C}(t)x$. We see that compared to Section 4.2, we have obtained a 'stabilizing' (2,2)-element, which should improve the stability properties, and an increased output dimension (larger $\mathcal{C}$-matrix), which should relieve some of the excitation requirements.

Since the system is LTV, we will have exponential stability of $x = 0$ if the pair $(\mathcal{A}(t), \mathcal{C}(t))$ is uniformly observable [3]. Uniform observability of $(\mathcal{A}(t), \mathcal{C}(t))$ is equivalent to uniform observability of $(\mathcal{A}(t) - K(t)\mathcal{C}(t), \mathcal{C}(t))$. Choose

$$K(t) = \begin{pmatrix} -\sqrt{k_q} & -\frac{(1-k_p)}{\sqrt{k_a}} \\ \frac{1-k_p}{\sqrt{k_q}}\Gamma_a f_a^\mathsf{T}(q) & 0 \\ \frac{1}{\sqrt{k_q}}\Gamma_d f_d^\mathsf{T}(q) & 0 \end{pmatrix},$$

then

$$\mathcal{A}(t) - K(t)\mathcal{C}(t) = \begin{pmatrix} 0 & 0 & -f_d^\mathsf{T}(q) \\ 0 & -k_a \Gamma_a f_a(q) f_a^\mathsf{T}(q) & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Due to the decoupled structure of the system $(\mathcal{A}(t) - K\mathcal{C}(t), \mathcal{C}(t))$, we can divide the problem in two. That is, the pair $(\mathcal{A}(t), \mathcal{C}(t))$ is uniformly observable if the systems

$$\dot{\xi}_1 = \begin{pmatrix} 0 & -f_d^\mathsf{T}(q) \\ 0 & 0 \end{pmatrix} \xi_1, \ \zeta_1 = \begin{pmatrix} \sqrt{k_q} & 0 \\ 0 & 0 \end{pmatrix} \xi_1, \tag{5a}$$

$$\dot{\xi}_2 = -k_a \Gamma_a f_a(q) f_a^\mathsf{T}(q) \xi_2, \ \zeta_2 = \sqrt{k_a} f_a^\mathsf{T}(q) \xi_2, \tag{5b}$$

are uniformly observable.

- Uniform observability of (5a) is implied by $f_d(q)^\mathsf{T}$ being PE, that is, by the existence of T and α such that for all t,

$$\int_t^{t+T} f_d(q) f_d(q)^\mathsf{T} d\tau \geq \alpha I.$$

- By looking at the measurement equation for (5b), it is clear that the corresponding condition,

$$\int_t^{t+T} f_a(q) f_a(q)^\mathsf{T} d\tau \geq \alpha I$$

  is a conservative[1] condition for uniform observability of (5b).

In the case of $n_\theta = 1$ (or, analysis within a local support), this reduces to that (the scalar) $f_a(q)$ and $f_d(q)$ must be positive.

That is, as long as the model predicts friction, we have enough information to obtain estimates of both annulus and drillpipe friction friction (within a local support). As we basically estimate two parameters from two equations ((1) and (3)), this is not surprising. We have less information when there is little friction (for small flows), which is not surprising either, since then the parameters does not influence the equations.

4.5 Time-delayed pressure measurement

Assume now that the pressure measurement is not a function of $q(t)$ as assumed above, but that it is T time-units delayed, that is,

$$p_b(t) = p_c(t-T) + f_a^\mathsf{T}(q(t-T))\theta_a + s_a(t-T).$$

Using the same setup as in Section 4.3 except for setting $k_p = 0$, that is, we use the observer (4) (with $k_p = 0$) and the parameter update laws

$$\dot{\hat{\theta}}_a = -\Gamma_a f_a(q)\left(q - \hat{q}\right) + K_a\left(p_b(t) - \hat{p}_b(t)\right),$$

$$\dot{\hat{\theta}}_d = -\Gamma_d f_d(q)\left(q - \hat{q}\right),$$

with $K_a = \Gamma_a f_a(q(t-T))k_a$ and

$$\hat{p}_b(t) = p_c(t-T) + f_a^\mathsf{T}(q(t-T))\hat{\theta}_a(t) + s_a(t-T).$$

[1]In this case, it is likely that less conservative conditions exist, in contrast to for the first system.

Noting that now, $\tilde{p}_b(t) = -f_a^\mathsf{T}(q(t-T))\tilde{\theta}_a$, we use the same Lyapunov function as above to arrive at

$$\dot{V} = -k_q\tilde{q}^2 - k_a\left(f_a^\mathsf{T}(q(t-T))\tilde{\theta}_a\right)^2.$$

Note that even though we have a time-delayed measurement, the system we analyze ($\tilde{q}$, $\tilde{\theta}_a$, $\tilde{\theta}_d$) is not a time-delay system, it merely involves a time-delayed signal, and therefore an analysis based on Barbalat's lemma can be readily invoked. That is, using Barbalat's lemma, we can conclude that $\dot{V} \to 0$ and hence $\tilde{q} \to 0$ and $f_a^\mathsf{T}(q(t-T))\tilde{\theta}_a \to 0$.

In principle, a PE analysis such as in Section 4.4 should also be possible. However, things become more complicated as it is not possible to arrive at the decoupled structure (5) in the same way. Nevertheless, from intuition one would expect that the same PE conditions should apply: Set the $(2,1)$-element of $K(t)$ to zero to get

$$\mathcal{A}(t) - K(t)\mathcal{C}(t) = \begin{pmatrix} 0 & -f_a^\mathsf{T}(q(t)) & -f_d^\mathsf{T}(q(t)) \\ 0 & -k_a\Gamma_a f_a(q(t-T))f_a^\mathsf{T}(q(t-T)) & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

If we try to divide this system as in Section 4.4, we see that (5b) is as before, while (5a) have an extra input. It is hard to see how this input would make things less observable than before.

It is notable that this way of including time-delay is considerably less complex than the standard way of doing it in the extended Kalman filter.

4.6 Robustifications

Several modifications for robustness are suggested in [2], the most important being normalization, leakage (localized σ-modification), projections and dead zone.

- **Normalization**: Since we have assumed q bounded (which also follows from the physics of the problem), normalization is not necessary to ensure boundedness. Nevertheless, normalization can be used to 'gain schedule' the adaptation laws to ensure similar loop gains for all conditions (varying qs).

  On the other hand, the local approach with scheduling in q allows us to achieve the same thing by setting the diagonal elements of Γ_a and Γ_d appropriately. Therefore, we do not pursue normalization any further.

- **Leakage**: Leakage can be implemented to ensure that the estimates converge to a nominal value (presumably 1) when there is no excitation. While this gives some robustness, it may compromise performance: If there is 'full circulation', should the parameter estimate corresponding to friction for small flows remain constant, or approach 1?

  If the answer is that the parameter should not be updated, then a localized σ-modification is probably smart [4], perhaps with an additional modification to avoid changing the parameter if the tracking performance is good (even if there is little excitation).

- **Projections**: Projections are easy to implement and do not affect performance (as long as the projections are sensible). Therefore, projections should be implemented. An example is $\theta_{\min} = 0$ and $\theta_{\max} = 2$, but probably tighter bounds must be used.

- **Dead zone**: "The principal idea behind the dead zone is to monitor the size of the estimation error and adapt only when the estimation error is large relative to the modeling error." Perhaps?

  Deadzone can be implemented localized [4].

5 Simulations

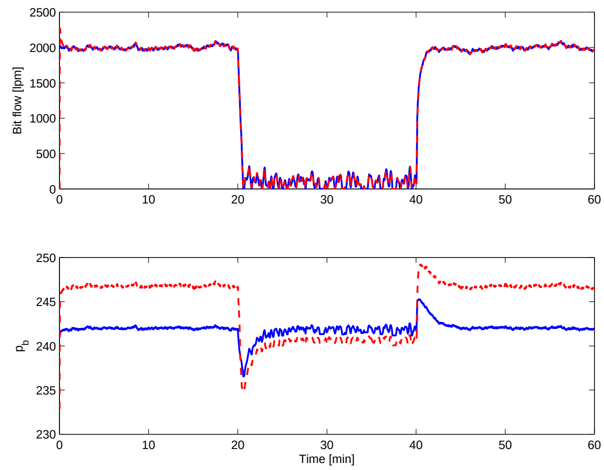
We test the approach on a simulator for the Kaasa model, with pressure loss (friction and hydrostatic) taken from a model of Grane. To make things slightly more realistic, we add noise to the pump flow input, and use estimated bit flow in friction calculation for the adaptive estimation (the latter improves adaption since it gives more excitation).

The scenario is initially full circulation, then rapid transition to low circulation (not zero flow, so not a pipe connection), then full circulation again.

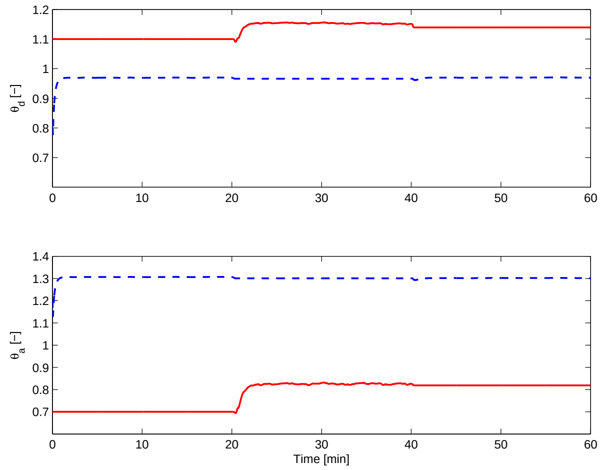
5.1 Simulation of case 1

A simulation of case 1 is shown in Figure 3. Estimated flow converges, and estimated parameters are bounded, but does not converge to their 'true' values (which in the simulation is 1 for all parameters). This gives large errors in predicted downhole pressure.

If we fix one of the parameter sets to its true value, then the other will converge to the true value for full circulation. For low circulation, low gain in the adaptation loop gives convergence problems.



(a) Top: simulated (whole) and estimated (dashed) bit flow; bottom: simulated (whole) and estimated (dashed) bottom-hole pressure.



(b) Top: drillstring parameters; bottom: annulus parameters.

Figure 3: Simulation of case 1 with $\Gamma_d = \text{diag}([0.2, 0.02])$, $\Gamma_a = \text{diag}([0.5, 0.2])$, $k_q = .001$.

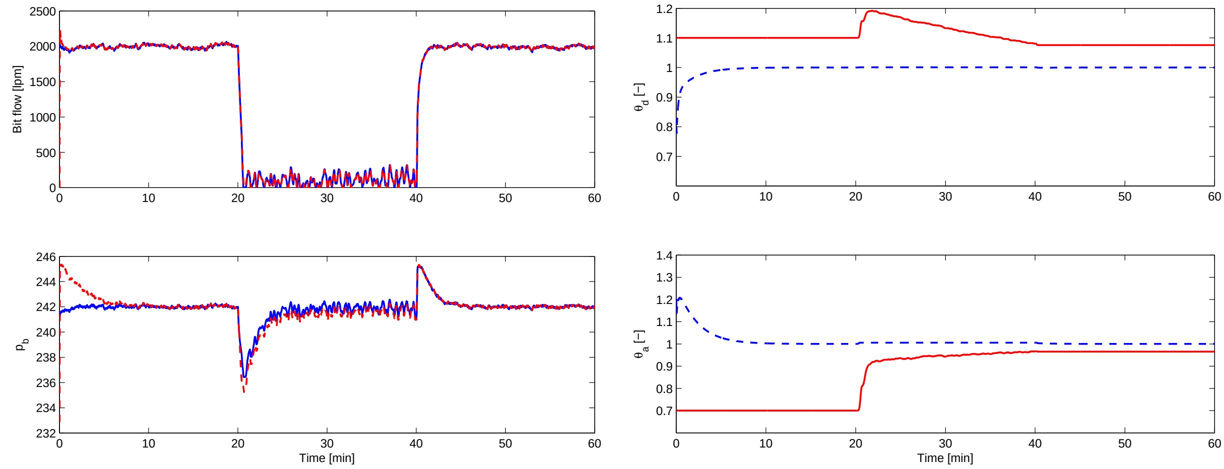
5.2 Simulation of case 2

A simulation of case 2 is shown in Figure 4. Now everything converges according to theory.

For small flows, convergence of θ_d is slow. This seems rather generic, increasing the gain gives faster convergence but larger 'overshoot' (and hence probably less robustness).

6 Testing data from Grane

We test the algorithm using (topside) data from Grane, recorded over the period 20:00-23:00 01/04/07. During this period it is drilled horizontally, with two pipe-connections. There are downlink events in the stand-pipe/bottomhole pressure measurements at ca. 65 min., 100 min., and 170 min.. Also, at ca. 90 min. the main choke become plugged, whereby the operators commands a large opening to the second choke, giving a dip in choke pressure which propagates to downhole and standpipe pressure.



(a) Top: simulated (whole) and estimated (dashed) bit flow; bottom: simulated (whole) and estimated (dashed) bottom-hole pressure.

(b) Top: drillstring parameters; bottom: annulus parameters.

Figure 4: Simulation of case 2 with $\Gamma_d = \text{diag}([0.2, 0.02])$, $\Gamma_a = \text{diag}([0.5, 0.2])$, $k_q = .001$, $k_p = 2 \cdot 10^{-4}$, $k_a = 2 \cdot 10^{-4}$.

The default adaptation algorithm is the one in Section 4.3. During pipe connections, the bottomhole pressure measurement is lost (topside), and the adaptation algorithm used is the one in Section 4.1.

We first try without correcting for time-delay, then with.

6.1 Assuming no time-delay

Figure 5 shows flow, pressure, and estimated parameters for a specific tuning of the adaptation algorithm. Figure 6 plots the calculated pressure (again) together with the topside measurement and measurements logged in the bit.

6.2 Including time-delay in measurement

We then run the estimation algorithm with the same parameters, but accounting for time-delay in the bottomhole pressure measurements. We assume the delay is 30s.

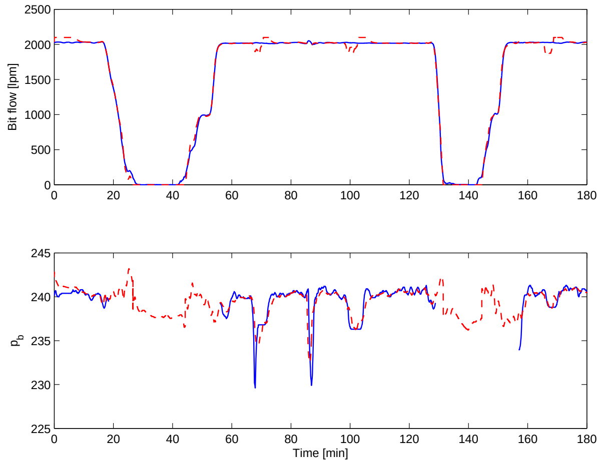
Some comments:

- The downlink events severely affect adaptation (best seen in the plot of θ_d). Probably, one should turn off estimation during these events. This would allow lower gains.
- The parameters corresponding to friction at small flows are seldom updated. This follows since the flow is zero most of the time with small flows, and the model used says that friction does not affect q or p_b at zero flow.

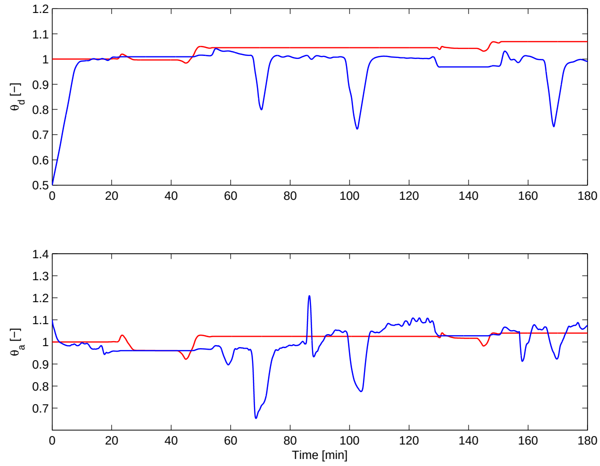
A Least squares approach

With (1) and (3) as a basis, define

$$z = \begin{pmatrix} \frac{s}{1+Ts} q - \frac{1}{1+Ts}\left(p_p - p_c + s(t)\right) \\ p_b \end{pmatrix}, \quad \phi^{\mathsf{T}} = \begin{pmatrix} -\frac{1}{1+Ts} f_d^{\mathsf{T}}(q) & -\frac{1}{1+Ts} f_a^{\mathsf{T}}(q) \\ 0 & f_a^{\mathsf{T}}(q) \end{pmatrix}, \quad \theta = \begin{pmatrix} \theta_d \\ \theta_a \end{pmatrix},$$



(a) Top: simulated/measured (whole) and estimated (dashed) bit flow; bottom: measured (whole) and estimated (dashed) bottomhole pressure.



(b) Top: drillstring parameters; bottom: annulus parameters.

Figure 5: Simulation of the algorithm in Section 4.3 with $\Gamma_d = 5 \cdot 10^{-3}\,\mathrm{diag}([2, 1])$, $\Gamma_a = 3 \cdot 10^{-2}\,\mathrm{diag}([2, 1])$, $k_q = .001$, $k_p = 0$, $k_a = 2 \cdot 10^{-3}$.

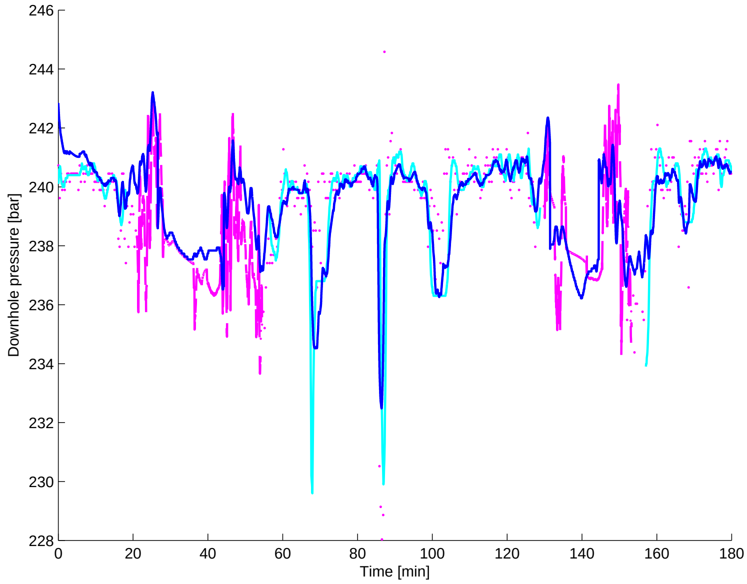
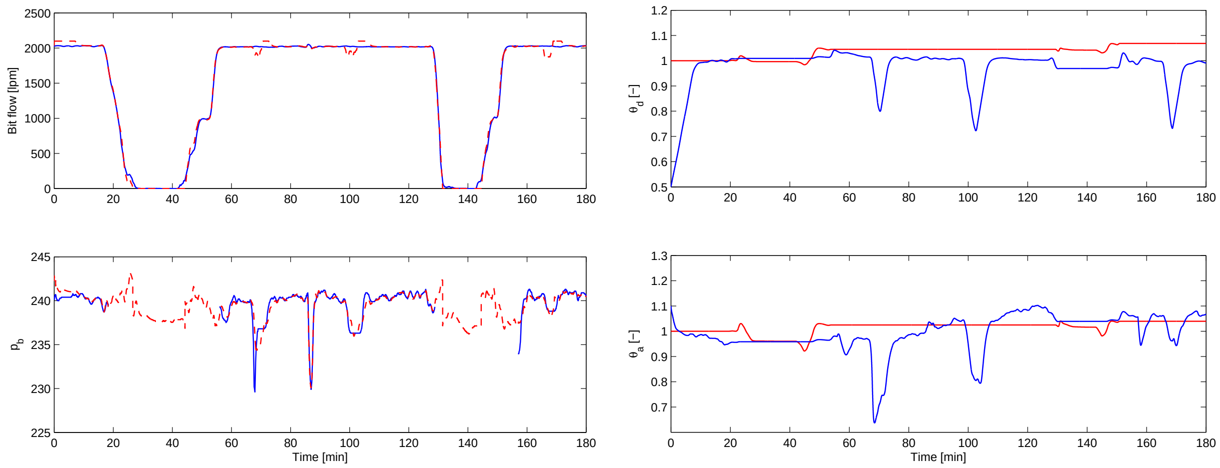


Figure 6: Adaptation based (blue, whole), topside measurement (cyan, dashed) and tool logged (magenta, whole/dotted) annulus bottomhole pressure for the algorithm in Section 4.3.



(a) Top: simulated/measured (whole) and estimated (dashed) bit flow; bottom: measured (whole) and estimated (dashed) bottomhole pressure. Note that estimated bottomhole pressure is delayed to fascilitate comparison with measurement.

(b) Top: drillstring parameters; bottom: annulus parameters.

Figure 7: Simulation of the algorithm in Section 4.5 with $\Gamma_d = 5 \cdot 10^{-3}\,\mathrm{diag}([2, 1])$, $\Gamma_a = 3 \cdot 10^{-2}\,\mathrm{diag}([2, 1])$, $k_q = .001$, $k_p = 0$, $k_a = 2 \cdot 10^{-3}$.

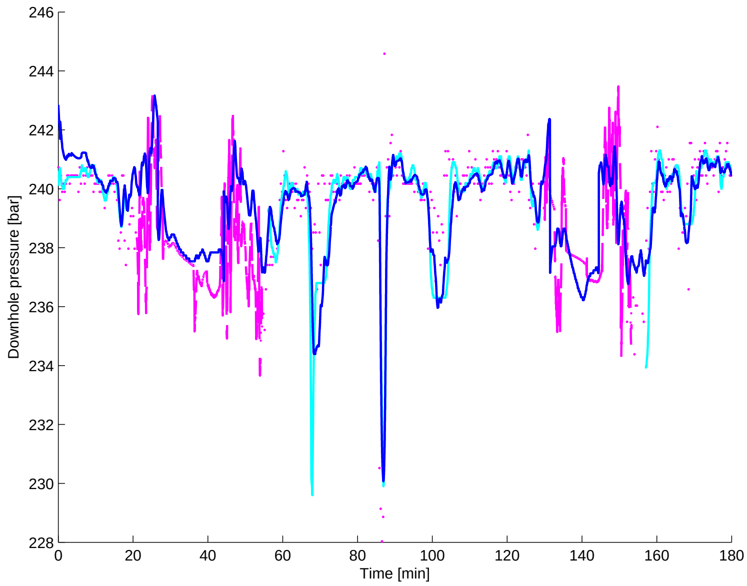
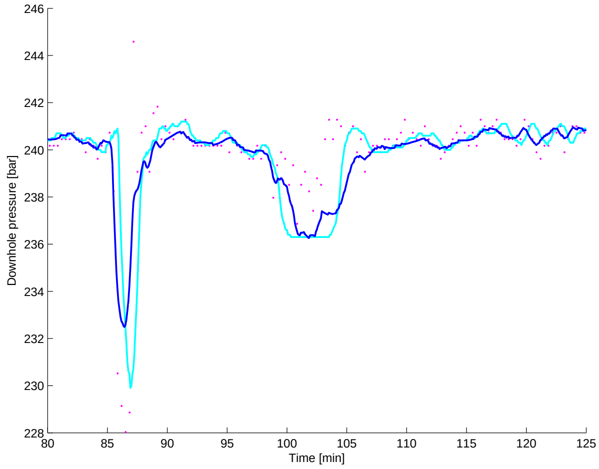
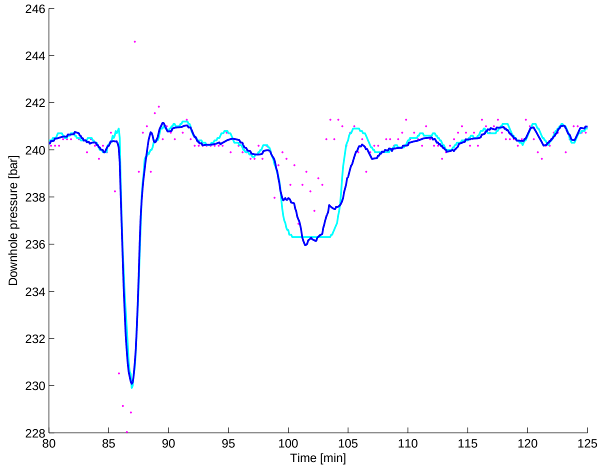


Figure 8: Adaptation based (blue, whole), topside measurement (cyan, dashed) and tool logged (magenta, whole/dotted) annulus bottomhole pressure for the algorithm in Section 4.5. Note that estimated bottomhole pressure is delayed to fascilitate comparison with topside measurement.



(a) A zoom on Figure 6.



(b) A zoom on Figure 8. Note that estimated bottomhole pressure is delayed to fascilitate comparison with measurement.

Figure 9: A comparison of Figure 6 and 8.

to obtain

$$z = \phi(t)^\mathsf{T}\theta.$$

With this in place, we are positioned to apply all adaptive approaches in [2, Chapter 4], and in particular the least squares-approach[2]. Since the "... use of the recursive least-squares algorithm with forgetting factor with $\phi \in \mathcal{L}_\infty$ and ϕ PE is appropriate in parameter estimation of stable plants where parameter convergence is the main objective", this might be worth trying.

The 'pure' least squares algorithm (without normalization) is

$$\dot{\hat{\theta}} = P(z - \hat{z})\phi$$
$$\dot{P} = -P\phi\phi^\mathsf{T}P, \quad P(0) = P_0.$$

Standard modifications to this algorithm are normalization, covariance resetting (P is the 'covariance matrix') and the use of a forgetting factor.

The main advantage with the least squares approach is that it might be easier to monitor convergence/PE-properties and taking corrective action using the covariance matrix.

The main disadvantage is perhaps the filtering that is necessary to obtain the linear parametric model. There is also less tuning parameters. Both these disadvantages can be rectified by taking 'the full step' towards dynamic least squares, and design an extended Kalman filter with (1) as dynamic model and q and (3) as measurements[3].

In some cases the increased computational complexity (of both least squares and extended Kalman filtering) might be a disadvantage.

References

[1] Jay A. Farrell and Marios M. Polycarpou. *Adaptive Approximation Based Control: Unifying Neural, Fuzzy and Traditional Adaptive Approximation Approaches*. Wiley-Interscience, 2006.

[2] A time-delay T in the p_b measurement can be incorporated by changing the $(2,2)$-element of $\phi(t)$ from $f_a(q)$ to $\mathrm{e}^{-Ts}f_a(q)$ (in other words, $f_a(q(t-T))$).

[3] For the EKF, a regressor is not explicitly used in the parameter update law, as it is in least squares adaptive control, but the same information is likely to be contained in the extended covariance matrix for EKF.

[2] Petros A. Ioannou and Jing Sun. *Robust Adaptive Control*. Prentice-Hall PTR, Upper Saddle River, NJ 07458, USA, 1995.

[3] Hassan K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2002.

[4] Yuanyuan Zhao and Jay A. Farrell. Localized adaptive bounds for approximation-based backstepping. *Automatica*, 44(10):2607–2613, 2008.