Control of delay dominant systems with costs related to switching

Honglian Deng and Lars Larsen and Jakob Stoustrup and Henrik Rasmussen

Abstract—The objective of this paper is to extend a novel low complexity method for optimizing switch control developed by the authors earlier to work with delay dominant systems and demonstrate that the method works in practice with a refrigeration test system. The extended method solves switching problems for high order systems which can be approximated as a first order with a delay as well. The extension of the method is realized with an observer to retrieve the delay-free information. Experimental validation of the extended method is carried out with a test system. A comparison to a baseline relay controller with fixed bounds shows that the optimizing switch control outperforms the baseline.

I. INTRODUCTION

Within the recent years optimal control of hybrid switching systems has attained a lot of focus. Among the recent results can be mentioned [1], [2], [3]. In this paper we will focus on systems with discrete inputs, which can be categorized as a subclass of hybrid switching systems. Examples of such systems in the industry are numerous. In this paper, we will extend the method proposed by the authors in [4] and [5] to include delay dominant systems. The method is analyzed using a generic delay dominant model and the applicability is tested using an air conditioning system (AC-system) with a ON/OFF compressor capacity as case study.

A method for synthesizing optimal control for hybrid systems is hybrid model prediction control (MPC) using the mixed logical dynamical framework (MLD)[1], but also other MPC methods based on solving finite horizon optimization problems exist such as [6] and [7]. Common for these methods is that they typically produce rather complex controllers requiring large computational power. Furthermore, if special care is not taken using finite horizon methods on systems with discrete inputs, it can lead to rather poor performance,[2],[4]. There exist also other methods based on steady state optimization of limit cycles, for instance [3] on page 253. Some of these methods give good results for particular applications, but are complex and difficult to apply especially in the lower levels of the control hierarchy where sufficient computational power is not available. This motivates the authors to develop a much less complex optimizing switch control method [4] and [5] overcoming the potentially poor performance that finite horizon methods leads to. The focus of this paper is on extending the developed method to systems with delay and multi-order systems, and demonstrating that the method is applicable in practice.

The paper is organized in the following way. The development of the optimizing switch control method is briefly introduced in Section II. Observer design to retrieve delay-free information is presented in Section III. Experimental validation of the optimizing switch control method and a comparison to a baseline relay controller is presented in Section IV. Finally conclusions are drawn in Section V.

II. BRIEF REVIEW OF THE METHOD

The process of deriving optimizing switch control method will be briefly presented in this section. More details can be found in [4] and [5]. The process can be divided into three steps: defining a cost function, cost function convexity study and deriving the method.

A. Cost function definition

For an open loop stable system with a binary input, we assume that the optimal steady state output trajectory follows a stable first order limit cycle (which is defined as a period including two switches and a constant duty cycle) instead of a fixed point. The system performance is therefore evaluated as the average squared control error over one switch period. For a first order limit cycle two switches are done in each period. The total switch cost for one period therefore accounts to 2R, where R is the cost of making a switch. The total average cost over one cycle period (T) can then defined as follows

\[ J = \frac{Q}{T} \int_{t_0}^{t_0+T} (y(t) - y_{ref})^2 dt + 2R, \]

where \( t \) is the time, \( y(t) \) is the measured output, \( y_{ref} \) is the output reference, \( Q \) is the cost associated with having a certain control error, and \( t_0 \) is the initial time for the period in question. In the following we assume that the cost \( R \) accounts for as well the costs of component wear as the efficiency loss at start/stops. This cost function represents the performance of the system over one cycle period. This cost function will be applied in the remainder of this paper.

B. Convexity

Having defined the cost function we will now study convexity based on first order (SISO) system, i.e.

\[
\begin{align*}
\dot{x} &= ax + bu + d \\
y &= cx,
\end{align*}
\]

where \( u \in \{0,1\} \) is the binary valued input and \( d \) is a disturbance. Further more (2) is stable, \( a \leq 0 \), and \( c = 1 \).

A first order limit cycle consists of an ON-period and an OFF-period as indicated in Fig.1. The output error accumulation can be divided into two parts: error from the
ON-period \( t \in [t_0, t_0 + \alpha T] \), and error from the OFF-period \( t \in [t_0 + \alpha T, t_0 + T] \), where \( \alpha \) denotes the duty cycle. The cost function (1) can therefore also be written as

\[
J(T) = \frac{Q}{T} \int_{t_0}^{t_0 + \alpha T} (y_1(t) - r_{ref})^2 dt + \frac{Q}{T} \int_{t_0 + \alpha T}^{t_0 + T} (y_2(t) - r_{ref})^2 dt + 2R
\]

(3)

where \( y_1 \) and \( y_2 \) refer to the output in respectively the ON and OFF period as indicated in Fig.1.

![Diagram](image.png)

Fig. 1. One cycle period

Operating at steady state limit cycles, the system output is periodical, meaning that the output at the end of one period equals the output at the beginning of the next, \( y_{1\text{,init}} = y_{2\text{,end}} \) and \( y_{1\text{,end}} = y_{2\text{,init}} \). Therefore \( y_1 \) and \( y_2 \) can be solved explicitly with the above conditions and become only functions of \( T \) and \( \alpha \). Hence the cost according to (3) can be computed as a function of \( \alpha \) and \( T \). Even so it is not easy to prove that the cost function (3) is mathematically convex. The alternative is simply to calculate the cost (3) within a reasonable range of \( \alpha \in [0, 1] \) and \( T > 0 \) and investigate the existence of a unique optimum.

The cost function calculation with first order system (2) is carried out the following parameters,

\[
\begin{align*}
\alpha &= -2.00 \times 10^{-4} \\
b &= -4.40 \times 10^{-6} \\
c &= 1 \\
\beta &= 0 \\
\bar{p} &= 300
\end{align*}
\]

which are from a first order air conditioned room model (presented in [4] and [5]) where the air conditioning unit is not modeled in detail but as energy input. The output set-point is \( r_{ref} = 23 \). \( d \) is corresponding to different outdoor temperature disturbance.

In Fig.2 the computed cost with \( d = 5.19 \times 10^{-3} \) is plotted as function of \( \alpha \) and \( T \). It clearly appears from the figure the cost function for this system indeed is convex.

Following this way, it can be demonstrated that the cost function (3) is convex for SISO high order systems with real poles.

![Graph](image.png)

Fig. 2. The computed cost, for a 1st order system following a steady 1st order limit cycle, as function of the period time \( T \) and the duty cycle \( \alpha \).

C. The optimizing switch control method

From the study of the cost function (3) above, we know that (3) is convex, therefore

\[
\frac{\partial J}{\partial T} = 0
\]

(4)

is true for the optimal solution.

Insert (1) into (4), we derive (5), which can be written as (6) by applying \( T = T_{\text{now}} + T_{\text{last}} \), \( T_{\text{now}} \) is the current half cycle period (either ON or OFF half cycle period) and \( T_{\text{last}} \) is the last half cycle period (either OFF or ON). It can be understood in this way: the left side of (6) is the current output deviation from the reference and the right side is the accumulated output error from the beginning of last switch and switch related cost.

So an optimizing switch method using (6) is derived, and the simulation results show that it drives the system towards the optimal solution both in duty cycle and period. More details regarding derivation of the optimizing switch control method and simulation results is referred to [4] and [5]

\[
(y_{0+T} - y_{\text{ref}})^2 = \int_{0}^{0+T} (y_t - y_{\text{ref}})^2 dt + \frac{2R}{Q}
\]

(5)

\[
(y_{0+T} - y_{\text{ref}})^2 (T_{\text{now}} + T_{\text{last}}) = \int_{0}^{T_{\text{now}}} (y_t - y_{\text{ref}})^2 dt + \frac{2R}{Q}
\]

(6)

The new method is developed based on a first order single input single output model. However, the method works as well for integration systems, because the cost function (1) for integration system is also convex. The proof of the convexity of the cost function for integration switching system is straightforward therefore it is not included in the paper.

Delay dominant systems pose challenges for the optimizing switch method. The reason is that, when the controller gets the delayed output information, it is already too late to react. To extend the method to delay dominant systems, the method can solve a set of multi-order system problems as well. The reason is that multi-order systems in many applications can be approximated by a first order system and
a delay. So the next section will focusing on delay dominant systems.

III. SYSTEM WITH DELAYS

The reason why it is important to extend the optimizing switch control method to work with delay dominant systems has been explained. Applying the optimizing method directly to a first order system with a delay leads to suboptimal solution due to that the feedback to the controller is ‘old’. The optimal solution for a first order switching system is independent of whether the system has a delay or not, meaning that if the delay-free output is fed back to the controller, the system should perform exactly the same as for the system without delay under same disturbances. The question is how the delay-free output information can be achieved.

Extensive research has been carried out on controlling systems with delays [8]. One traditional method is based on the Smith Predictor (SP) [9]. Due to the problem that SP does not work with unstable systems for example systems with an integrator, [10], [11], have proposed modified SP methods. A common prerequisite for successful employment of SP (both the original as well as the modified ones) is the accurate knowledge of the model and delay. However, for most industrial cases this knowledge is not easily available.

Another approach is based on the idea of representing the delay with Pade approximation, e.g. [12]. A systematical approach of designing observer based on Pade approximation is introduced in [13], which will be explained briefly here.

A. Observer based on Pade approximation

The proposed observer design approach in [13] is to first partition the delay \( \theta \) into \( p \) non-overlapping delays as (7), then introduce Pade approximation (8) to each partition of the delay and formulate a state space model of the Pade approximated delay together with the system model (2). This approach facilitates using well-established design techniques to construct an observer that provides the delay-free output.

\[
e^{-\theta s} = (e^{-\frac{\theta}{m}})^p
\]

\[
e^{-\theta s} \approx \frac{1 - \frac{\theta}{m}}{1 + \frac{\theta}{m}}
\]

B. A new approach of designing observer for systems with delay

Our proposal of designing observer for delay system can be explained in the following steps:

1. Discretize the system including delay with a large sampling time \( \frac{\theta}{m} \), so that the delay can be partitioned into \( m \) partitions, and each of the delay partition can be modeled as one sampling time delay, and therefore the whole system becomes discrete.
2. Design an observer for this discrete time system - find the gain \( L_d \).
3. Check whether the observer is stable if the discrete observer gain is directly applied into the continuous system (or a system with a much faster sampling time). If it is stable, then the observer is found, otherwise, increase \( m \), and repeat the above procedure. It is preferred to have a minimal \( m \) because it ensures that the observer has a minimal number of states.

In continuous time, the delay in the observer is divided into \( m \) number of delays \( e^{\frac{\theta}{m}} \), and the corresponding terms in gain \( L_d \) to compensate each term of \( e^{\theta s} \). The observer stability can be checked in this way: find the observer transfer function from the error (between the measured and estimated output) to the estimated output, and check the stability based on e.g. a Nyquist diagram. Such a observer stability Nyquist diagram is shown in Fig.3, where the observer gains are designed based on different sampling times. From the figure, it can be seen that the observer stability increases with decreasing sampling time.

The proposed observer design approach is systematic. Comparing with the observer based on Pade approximation, the advantage is, that the observer stability is ensured. However, in the experiments the observer based on Pade observer \( p = 1 \) turns out to be stable. Thus, the more general method described above is given for completeness, but in the experimental testing, the simpler design with Pade observer \( p = 1 \) has been applied, since that the focus of the paper is on the optimizing switch control method.

IV. TEST SYSTEM DESCRIPTION

To verify the optimizing switch control method, experiments are carried out with a refrigeration plant in Aalborg University in Denmark [14]. The schematic view of the refrigeration plant is shown in Fig.4. The system emulates an air conditioning system, where the tank simulates a room, and the refrigeration system works as an air conditioning unit installed in the room. The overall control goal is to maintain the tank water temperature close to the reference by switching the refrigeration system ON/OFF.

The heat load on the system is maintained by an electrical heater with a adjustable power supply for the heat element in the water tank. Since the tank is well insulated, the heat loss to the ambient is so little that it can be neglected, which
means that the tank behaves like a integration system. The compressor is with a variable speed drive from 35Hz–60Hz. Temperature sensors and mass flow sensor are installed on the water pipe passing the heat exchanger.

A. Modeling of the test system

The refrigeration system is considered as a whole unit with discrete input control signal 1 or 0 to the compressor. When 1 is sent to the compressor, the compressor starts and therefore the refrigeration system switches on. In the experiments, when the compressor is on, it runs with 35Hz, which in steady state can remove approximately 3200W from the water through the heat exchanger. When 0 is sent to the compressor, the refrigeration system switches off, which means no power is removed. The power removed by the refrigeration system is expressed as $\dot{Q}_e = 3200,0W$. More details of the refrigeration system can be found in [14].

The water tank is roughly modeled as

$$T_{\text{water}} = \frac{\dot{Q}_{\text{load}} - \dot{Q}_e}{c_{p,\text{water}} \cdot m_{\text{water}}} \quad (9)$$

where $T_{\text{water}}$ is the delay-free average tank water temperature, $T_{\text{w,out}}$ is the measured outlet water temperature, $\dot{Q}_e$ is the estimated power removed from the water tank, which can be modeled from the water side,

$$\dot{Q}_e = c_{p,\text{water}} \cdot m_{\text{water}} (T_{\text{w,out}} - T_{\text{w,in}}) \quad (10)$$

$T_{\text{water}}$ is the average tank water temperature. $c_{p,\text{water}}$ and $m_{\text{water}}$ is the specific heat capacity and mass of water in the tank. $\dot{Q}_e$ is the cooling capacity from the refrigeration system. $\dot{Q}_{\text{load}}$ is the heating load from the electrical heater. The disturbance from the ambient to the water tank is ignored due to the very good insulation of the tank.

The water temperature coming into and out of the tank is $T_{\text{w,in}}$, $T_{\text{w,out}}$ which is corresponding to the water temperature after and before passing the evaporator. A mixer is installed to keep the water in the tank well mixed so that the water temperature is even and therefore the temperature of the water coming out of the tank $T_{\text{w,out}}$ can be approximated as the water temperature inside the tank. However, there is a delay from when the refrigeration system removes the power until $T_{\text{w,out}}$ reacts. This could be caused by the flow transportation in the pipe etc. In this case the delay is approximately $\theta = 90s$, and it is modeled as

$$T_{\text{w,out}} = T_{\text{water}}(t - \theta) \quad (11)$$

The parameter for the models are $c_{p,\text{water}} = 4180J/(kg \cdot ^{\circ}C)$, $m_{\text{water}} = 65kg$, $\theta = 90s$, $\dot{Q}_e = 0$ or 3200W.

In all of our experiments, the reference for the tank ($T_{\text{w,out}}$) is $22^{\circ}C$, and the ambient temperature is controlled around $22^{\circ}C$ so that there is no significant heating disturbance from the ambient. The disturbance from the electrical heating element is $1000W$.

It should be noted that the optimizing switch control method was developed based on a first order model, but it works on an integration system as well, which can be seen as a special first order system with infinite time constant. The proof of convexity of the cost function (3) for an integration system is not hard, therefore is not included here.

V. Experimental results

In this part, two groups of experiments are carried out. The first group of experiments are designed to validate the developed optimizing switch control method. The second group is to compare the optimizing switch control method with a baseline relay control. An observer is needed to retrieve the delay-free output.

The observer applied in the experiments is based on a first order Pade approximation as described in III-A. The model of the refrigeration plant used for the observer is (10). Due to the difference between the input energy to the refrigeration plant and the estimation of the input to the observer, an offset appears on the estimated delay-free output. To remove the offset, an extra integration term is introduced, which at last becomes a proportional and integral (PI) observer. Design of such a PI observer can be found in [15], and the procedure will not be repeated here. The resulting observer is shown as in Fig.5, which is referred to as Pade PI observer in this paper. The following observer gains $L = [1,23 \times 10^1, 6.75 \times 10^{-2}, 1.25 \times 10^{-4}]$ which are obtained from off-line simulations with some recorded data from one experiment are applied in the experiments in this section.
Fig. 6. Cost vs. periods. \(Q=20, R=500\). Heating load 1000W. ‘pulse’ refers to the cost results from the pulse input experiments. ‘polyfit’ is the polynomial fit of the discrete test points of the pulse results.

Fig. 7. Comparison of optimizing controller and real optimal solution at heating load 1000W. \(R\) is kept at 500. ‘sim’ is the optimal solution for the model. ‘optimizing’ is the solutions from applying the optimizing switch control on the test system. ‘real’ is the optimal solution from the pulse experiments with the test system.

### A. Validation of the optimizing switch control method

Validation of the optimizing switch control method is carried out by comparing the control result with the optimal solutions.

The first question is what the optimal solution for the ON/OFF switching refrigeration plant with respect to the water outlet temperature is. The optimal solution for the refrigeration plant is different from the calculated optimal solution based on the model (10), because the model (10) (integration and delay) is a crude approximation of a high order system. The only way to achieve the optimal solution for the test system is to run pulse experiments with different periods until it reaches steady states, then calculate the cost using (3) to find the period corresponding to the smallest cost. Due to that the refrigeration plant is very close to an integration system, therefore duty cycle and the load percentage are very close to each other, which can be calculated directly by the heating to cooling ratio, while for a first order system, the optimal duty cycle will be significantly different from the load percentage when the load is different from 50%.

The above procedure of finding the real optimal solutions is illustrated by Fig.6, where the penalty is \(Q=20, R=500\) and heating load is 1000W. A polynomial fit of the test points has minimal cost at around 350s. The benefit of the polynomial fit is that with discrete test points, a continuous expression of the cost as a function of the period can be achieved.

To investigate the ability of the optimizing switch control method to drive a system towards the optimal solution at different penalty ratios, another two optimal solutions for the real penalty \(Q = 2, R = 500\) and \(Q = 80, R = 500\) are also demonstrated with pulse experiments. The resulting three optimal solutions at three different Q (R is kept at 500 for all the experiments) are shown in Fig.7. For comparison, the optimizing switch control resulting costs at different Q/R ratio and heating disturbance 1000W are also plotted Fig.7.

Fig.7 demonstrates that the solutions from the optimizing switch control (legend ‘optimizing’) are very close to the optimal (legend ‘real’). The error defined by (12) at the three \(Q = 2, 20, 80\) are 2.30%, 4.85%, 3.00%. These results prove that, the optimizing controller achieves results very close to the optimal solutions.

### B. Optimizing switch method vs. relay control

In [5], simulation results from the optimizing and relay controller for a first order system has been compared, and they show that the optimizing switch control method outperforms relay control when the systems are not at the nominal working condition for the relay.

Several experiments have been carried out to investigate the performances of the optimizing switch control method and relay controller with and without observer at different heating load.

1. Optimizing switch control method. Run the optimizing switch control method including the Padé PI observer at different heating loads with parameter \(Q = 2, R = 50\). The observer gains are \(L = [1.23 \times 10^1, 6.75 \times 10^{-2}, 1.25 \times 10^{-4}]\).
2. Standard relay controller. Tune the relay controller at about 50% corresponding to about 1500W heating disturbance, which reaches results very close to the optimizing switch control. The resulting relay band is 0.36°C, which means that the refrigeration system switches on when \(T_{\text{w, out}}\) is higher than 22 + 0.36°C and switches off when \(T_{\text{w, out}}\) is lower than 22 − 0.36°C. Then run the experiments with this controller at different heating loads.
3. Relay controller with observer. An improvement for the relay controller could be to include the PI observer. Apply the Padé PI observer and feed the non-delayed information to the relay controller. Repeat the procedure of tuning the standard relay controller. The resulting relay band is 1°C, which means that the refrigeration system switches on when \(T_{\text{w, out}}\) is higher than 22 + 1°C and switches off when \(T_{\text{w, out}}\) is lower than 22 − 1°C. Then run the experiments with this controller at different heating loads.

The results from the above three experiments are shown in Fig.8, where it can be seen that close to the nominal condition 50% load, the relay controller with and without produces fine results, but when the heating load moves away from 50%, for example at heating load 500W, the standard controller cost deviation from the optimizing controller is very large, while the relay controller including observer improves, but still the cost is much larger than optimizing switch control result.
The temperature outputs from these three controllers at heating load 500W are plotted in Fig.9. Comparing Fig.9a,b,c, it can be easily seen that the standard relay control output has an offset, which results in a large cost. The reason for the offset is that during the delay time, the heating power \( \dot{Q}_{\text{load}} \) to raise the temperature (when refrigeration system is switched ON) and the heating power minus the cooling power \( \dot{Q}_{\text{load}} - \dot{Q}_c \) (when refrigeration system is switched OFF) to remove heat is different. The offset in the standard relay has been removed and it performs more close to the optimizing switch control result. The optimizing switch control output is offset free.

VI. CONCLUSIONS

A newly developed optimizing switch control method for optimal control of system with binary inputs that have costs related to switching has been tested with a refrigeration test systems. An observer design has been made to attack the delay problems for the optimizing switch control method. The method including the observer drives the system to the optimal limit cycle solution and it performs better than relay controllers with/without observer at non-nominal condition. The optimizing switch control method has only been tested with SISO first order systems and integration, but it has been proven to work also with SIMO first order and integration systems as long as the cost function has no local optima. For multi-order single input systems, the phase shift problem can also be solved with this method by approximating the system with a first order system with a delay in a proper frequency range. The method has low complexity which at the same time avoids the fundamental problem of finite horizon prediction methods.

REFERENCES