Linköping Studies in Science and Technology
Thesis No. 1547

Optimal Engine Operation in a Multi-Mode CVT Wheel Loader

Tomas Nilsson

Linköpings universitet
INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering
Linköpings universitet, SE–581 83 Linköping, Sweden

Linköping 2012

Optimal Engine Operation in a Multi-Mode CVT Wheel Loader

tnilsson@isy.liu.se
http://www.vehicular.isy.liu.se
Department of Electrical Engineering,
Linköpings universitet,
SE–581 83 Linköping,
Sweden.

ISBN 978-91-7519-829-3
ISSN 0280-7971
LIU-TEK-LIC-2012:32

Printed by LiU-Tryck, Linköping, Sweden 2012

Abstract

Throughout the vehicular industry there is a drive for increased fuel efficiency. This is the case also for heavy equipment like wheel loaders. The operation of such machines is characterized by its highly transient nature, the episodes of high tractive effort at low speed and that power is used by both the transmission and the working hydraulics. The present transmission is well suited for this operation, though the efficiency is low in some modes of operation. Both operational advantages and efficiency drawbacks are highly related to the use of a torque converter. Continuously variable transmissions (CVTs) may hold a potential for achieving similar operability with reduced fuel consumption, though such devices increase the demand for, and importance of, active control.

Common wheel loader operation is readily described in a framework of loading cycles. The general loading cycle is described, and a transmission oriented cycle description is introduced, in deterministic and stochastic forms, and a description is made on how such cycles are created from measurements. A loading cycle identifier is used for detecting cycles in a set of measured data, and a stochastic cycle is formed from statistics on the detected cycles.

CVTs increase the possibility for active control, since a degree of freedom is introduced in the engine operating point. Optimal operating point trajectories are derived, using dynamic programming (DP), for naturally aspirated (NA) and turbocharged (TC) engines. The NA-engine solution is analyzed with Pontryagin's maximum principle (PMP). This analysis is used for deriving PMP based methods for finding the optimal solutions for both engines. Experience show that these methods are $\sim$100 times faster than DP, but since the restrictions on the applicable load cases are severe, the main contribution from these is in the pedagogic visualization of optimization. Methods for deriving suboptimal operating point trajectories for both the NA and the TC engines are also developed, based on the optimization results. The methods are a factor $>$1000 faster than DP while producing feasible trajectories with less than 5% increase in fuel consumption compared to the optimal.

Multi-mode CVTs provide the possibility of even higher efficiency than single mode devices. At the same time, the added complexity makes control development increasingly time consuming and costly, while the quality of the control is decisive for the success of the system. It is therefore important to be able to evaluate the potential of transmission concepts before control development commence. Stochastic dynamic programming is used and evaluated as a tool for control independent comparing of the present transmission and a hydrostatic multi-mode CVT concept. The introduction of a stochastic load complicates the optimization, especially in the feasible choice of states for the optimization. The results show that the multi-mode CVT has at least 15% lower minimum fuel consumption than the present transmission, and that this difference is not sensitive to prediction uncertainties.

Acknowledgment

The work presented in this thesis has been carried out at the Division of Vehicular Systems at the Department of Electrical Engineering, Linköping University. I would like to start by thanking the funders of this research; Volvo Construction Equipment and the Swedish Energy Agency, for giving me the opportunity to work with projects in the interesting and challenging field of transmission control with a heavy equipment application.

I would like to express my gratitude to my supervisors Jan Åslund and Lars Nielsen at Linköping University, and Anders Fröberg at Volvo CE. Jan Åslund is acknowledged for all of his support and guidance during the work. I would like to acknowledge Lars Nielsen especially for his fresh ideas, ability to find weak points and making me take a step back every now and then. I would like to thank Anders Fröberg, who has my gratitude for his invaluable support, assistance and, not least, patience.

The research group of vehicular systems is acknowledged for the pleasant atmosphere and working environment. I would like to thank Martin Sivertsson for his input on the writing of this thesis. Erik Frisk, Mattias Krysander, Peter Nyberg and Christofer Sundström is acknowledged for their parts of the development of the loading cycle detector, a work which for me also became an introduction to wheel loader operation.

There are many at Volvo CE who deserve acknowledgment for their help and assistance. First and foremost is Anders Fröberg, who have not only helped in formulating the research questions, but also given much guidance in the methods used. Rickard Mäki is acknowledged for his support without which the project would not have run this smooth. I would like to thank Gianantonio Bortolin for assisting with models and data despite being fully occupied with his own high workload. I would also like to thank Per Mattsson and Mats Åkerblom for letting me use their MM-CVT concept in this work. Finally, I would like to thank Bobbie Frank, Jonas Larsson, Lars Arkeborn, Johan Carlsson, Reno Filla and all others who have provided material, assistance or inspiration which have helped in the work presented in this thesis.

Finally and most importantly, I would like to thank my wife Malin for her endless support, patience and encouragement.

Contents

1

Introduction

The common operation of wheel loaders differs from that of regular on-road vehicles in several important aspects. The present transmission is well suited to this operation, though the efficiency is low in some modes of operation. In the present transmission both operational advantages and efficiency drawbacks are highly related to the use of a torque converter. Continuously variable transmissions may have a potential for achieving similar operability with reduced fuel consumption. The work presented in this thesis focus on the optimal dynamic choice of engine operating point is such a transmission and using trajectory optimization for evaluating the fuel saving potential of complex transmissions.

1.1 Background

Wheel loader operation is in general highly repetitive, with cycle times below $30s$ and driving direction changes as often as every 5 seconds not being uncommon. The material handling also separate the wheel loader from most vehicles. Filling a bucket with gravel requires high tractive effort at low speed and it is not uncommon to have a peak working hydraulic power of the same magnitude as the peak transmission output power. A common wheel loader transmission layout is presented in Figure 1.1. The engine is connected to a working hydraulics pump and a hydrodynamic torque converter. The torque converter is connected to an automatic gearbox, which connects to the drive shaft. In this setup the torque converter is a crucial component, since it provides some disconnection between the engine and transmission speeds. This makes the system robust by, for example, preventing the engine from stalling if the vehicle gets

stuck. Unfortunately the torque converter also cause high losses at some modes of operation, such as during the high thrust and low speed combination when filling the bucket. This lack of efficiency is the reason for a desire to find other transmission concepts for wheel loaders. One alternative may be some type of continuously or infinitely variable transmission (CVT), such as the hydrostatic transmissions used in Zhang (2002) and in Lennevi (1995), or the diesel-electric transmission used in Filla (2008). The first major part of this thesis is focused on the choice of engine operating point in a CVT transmission. The second major part is focused on the evaluation of more complex CVT-based transmissions.

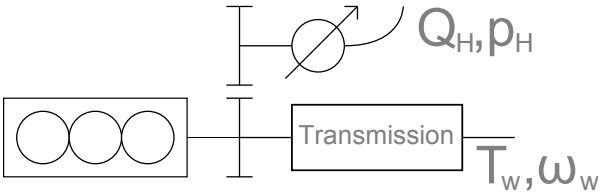


Figure 1.1: *A common wheel loader drivetrain setup.*

1.2 Motivation and previous work

Engine operating point selection

Common for CVTs are that these increase the freedom of choosing engine operating point and that the engine inertia will work as a small energy storage. If the output power demand is not constant the fuel optimal choice of operating point is an interesting and non-trivial problem. There have been a large amount of research regarding this choice. Liu and Paden (1997) presents a survey of causal heuristic control concepts, and corresponding operating point trajectories. These control concepts are referred to by Pfiffner (2001) and Srivastava and Haque (2009) among others. Since these are causal, the desired output power cannot be delivered for all transients, and the selection of strategy become highly affected by the penalty function for deviating from the desired output power. Pfiffner (2001) also derive non-causal optimal operating point trajectories, though the results are not thoroughly examined and explained. Rutquist et al. (2005) perform a theoretical investigation of optimal solutions, but only for fully stochastic future loads. There are several reasons for studying the optimization of engine operation, even though direct application of the solution requires prediction of the future load. Delprat et al. (2001) use optimization theory to derive a causal control concept which has become known as the ECMS (as described in the end of Section 3.3) and gained a lot of attention, as shown by Sciarretta and Guzzella (2007). A well designed online optimization algorithm may only require a short horizon prediction. Some proposals on how to obtain such predictions can be found in Asadi and Vahidi (2011), Mitrovic (2005) and Pentland and Andrew (1999). In case the vehicle

is made autonomous, as investigated by Ghabcheloo et al. (2009) and Koyachi and Sarata (2009) among others, the controller may also inform the optimizer about upcoming actions, thus providing a prediction. This motivates the work presented in Chapter 4.

Multi-mode CVT concept evaluation

The repeated power conversions in CVTs, in a hydrostatic CVT for example the conversions are from mechanical to hydraulic and back to mechanical, are negative for the overall efficiency. In power-split constructions, such as those used in Carl et al. (2006) and Gramattico et al. (2010), this is addressed by leading part of the power through a mechanical connection. Multi-mode CVTs are constructed so that several power-split layouts can be realized with the same device, thus enabling high efficiency at widely separated gear ratios. The increased complexity of such transmissions makes controller development time consuming and costly, as exemplified by the work in Zhang (2002), and at the same time the quality of the controller become decisive for the success of a transmission concept. It is therefore important to know what fuel consumption can be expected of a transmission concept before controller development commence. The potential of a concept, excluding control, can be determined by dynamic optimization methods, just as in Paganelli et al. (2000), Pfiffner (2001) and Sciarretta and Guzzella (2007). Since wheel loaders are off-road vehicles with highly transient operating patterns, access to accurate prediction is not to be expected. The evaluation of the potential of a transmission concept should therefore include an analysis of the sensitivity of the fuel saving potential to prediction uncertainties. This motivates the work presented in Chapter 5.

1.3 Contributions

Section 2.2.2 briefly describe an automatic loading cycle detector and identifier. The development of this detector is an ongoing project. In this work the detector is used for automatic extracting of a number of loading cycles from a measurement sequence.

Chapter 4 is based on the papers Nilsson et al. (2011) and Nilsson et al. (2012c) and provide a thorough investigation of dynamic optimization of the operation of naturally aspirated and turbocharged engines. Trajectory optimization methods, based on Pontryagin's maximum principle, for the two setups are developed and presented in Section 4.7. These are fast but highly restrictive on the applicable load cases and therefore of limited practical use. The method for the naturally aspirated engine though can be graphically interpreted by phase planes, which provide excellent visualization of optimization with Pontryagin's maximum principle. Section 4.8 present suboptimal trajectory derivation methods that does not depend on an analytic engine model and are extremely fast. These produce reference engine operating point trajectories

that give fuel consumptions with, in all cases studied, less than five percent fuel usage increase from the optimal.

Chapter 5 is based on the papers Nilsson et al. (2012a) and Nilsson et al. (2012b) and analyze the fuel saving potential of two drivetrain concepts and the use of stochastic dynamic programming for this analysis. The two concepts are the present transmission which is based on a torque converter and an automatic gearbox, and a hydrostatic multi-mode CVT concept. Section 5.4 shows that the stochastic load formulation highly affect the feasible choice of states in the optimization. Section 5.5 shows that the CVT concept has more than 15% better fuel saving potential than the present transmission, and that this number is not sensitive to prediction uncertainties, cycle smoothness or cycle length.

1.4 Thesis outline

Chapter 2 treats wheel loader operation and load case creation. Section 2.1 provide a description of typical wheel loader operation. Section 2.2 describe how deterministic and stochastic driving cycles can be derived from measurements with the aid of a driving cycle detector and identifier.

Chapter 3 describe the dynamic optimization methods and tools used. Section 3.1 formalize the optimization problem. Section 3.2 describe deterministic and stochastic dynamic programming and the algorithm used in this work for solving the recursion. Section 3.3 present Pontryagin's maximum principle and mention optimization methods based on this principle. Section 3.4 discuss applications and issues encountered when implementing trajectory optimization.

Chapter 4 examine optimization methods and results for a standalone engine with and without a turbocharger. Sections 4.1, 4.2, 4.3 and 4.4 sets up the models and optimization problem. Section 4.5 shows the engine map and static solutions. Section 4.6 presents the optimal solutions derived with dynamic programming. In Section 4.7 optimization methods based on Pontryagin's maximum principle are developed and evaluated. In Section 4.8 suboptimal trajectory derivation methods are developed and evaluated. Section 4.9 concludes the standalone engine analysis with a discussion of the methods developed.

Chapter 5 treats the comparative evaluation of the present transmission and a hydrostatic multi-mode CVT concept by deterministic and stochastic dynamic programming. Sections 5.1, 5.2 and 5.3 presents the models and load cases used. Section 5.4 discuss the important consequences on the choice of states from the optimization method and stochastic load formulation. Section 5.5 presents the results of the trajectory optimization. Section 5.6 concludes the chapter with a discussion on the optimization and the performance of the transmissions.

Chapter 6 presents conclusions drawn from the work presented in this thesis. Section 6.1 recapitulates the outcome of the methods developed in Sections 4.7 and 4.8 and summarizes the experiences and findings regarding the application of, especially stochastic, dynamic programming. Section 6.2 describe the conclusions drawn on the efficiency improvement capability of the CVT concept.

1.5 Nomenclature

General symbols

Symbol	Description	Unit
D	maximum displacement	m^3/rad
f	dynamic functions, discrete time	
F	dynamic functions, continuous time	
g	cost function, discrete time	
G	cost function, continuous time	
H	Hamiltonian	
I	moment of inertia	kgm^2
J	performance index	
m	mass	kg
M_{cvt}	CVT-mode	$-$
M_P	torque converter torque map	Nm
n	number of -	$-$
p	pressure	Pa
P	power	W
Q	flow	m^3/s
r_c	gear	$-$
s	optimization parameter	$-$
t	time	s
T	terminal time	s
T	torque	Nm
u	controls, continuous/singular points	
U	controls, discretized (vector)	
w	disturbance signals	
W	disturbance signals	
x	states, continuous/singular points	
X	states, discretized (vector)	
Y	load components	
z	number of cogs	$-$
ε	small value	$-$
η	efficiency, efficiency parameters	$-$
θ	angle	rad
λ	costate function	
μ	torque converter torque ratio	$-$
ν	torque converter speed ratio	$-$
ξ	miscellaneous constant	
Σ	”quasi static” peak efficiency curve	$-$
τ	time constant	s
ψ	relative displacement	$-$
ω	angular speed	rad/s

Subscripts

Symbol	Description
b	brake
c	torque converter
e	engine
f	fuel
H	working hydraulics
cp	torque converter pump
t	turbocharger
T	transmission
ct	torque converter turbine
v	variator
w	wheel/drive shaft
Σ	"quasi static" peak efficiency

Diacritics and Superscripts

Symbol	Description
$\hat{x}$	actual
$\tilde{x}$	interpolated
$\dot{x}$	time derivative
x^*	optimal

2

Wheel loader operation

2.1 Wheel loader operation

Wheel loaders are versatile heavy equipment machines, and as such they are used for a wide range of tasks. The most common usages are loading and short distance transportation of material, often bulk material such as gravel, shot rock or wood chips. The machine uses a bucket to scoop loose material from a source pile, transport it some distance and unload at a receiver. This type of operation is usually highly repetitive, since the source and receiver positions in general are stationary. Due to the repetitiveness, the operation is naturally described as repeating cycles, just like in Filla (2005) and Fengyuan et al. (2012), the most common being the short loading cycle, which is illustrated by Figure 2.1. This cycle can be partitioned into phases in different ways, one being the following:

1. **Forward.** From standstill, acceleration and driving toward the source pile. The phase ends at contact with the pile.

2. **Loading.** Penetration of the pile and filling the bucket, generally by a combination of bucket movement and forward driving with high tractive effort. The phase ends when the forward motion ends.

3. **Reversing.** Reversing away from the pile, including acceleration, driving and deceleration, often coordinated with raising of the bucket.

4. **Transport.** From driving direction change, forward acceleration, driving and deceleration at the load receiver. The bucket is in general raised during or at the end of this transport.

5. **Emptying.** The bucket is tilted down, emptying the load. This takes place around the direction change from transport to reversing.

6. **Reversing.** Reversing away from the load receiver, including acceleration, driving and deceleration, generally while lowering the bucket.

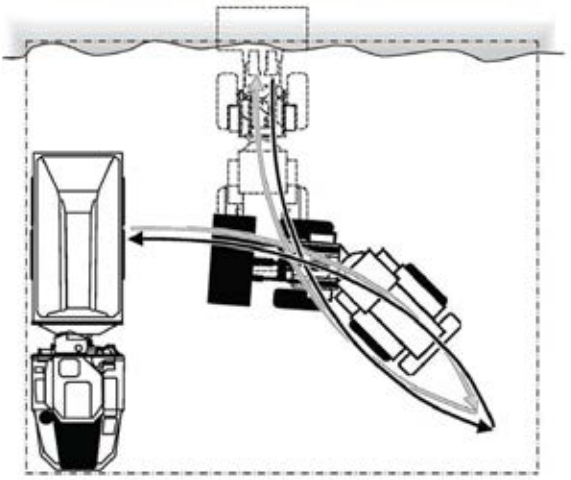


Figure 2.1: *Short loading cycle, based on a figure from Filla (2011).*

This general cycle is very common, even though the material handled and the distance driven vary. In the shortest cycles, each of the phases can be about 5 seconds, except for phase 5 which is close to instantaneous, which give a total cycle-time of about $25s$. Several interesting characteristics can be observed from, or should be added to, this cycle description:

- Each phase is short, to the extent that delays even in the order of tenths of a second would be significant.
- Driving direction changes are often made by changing gear direction, thus using engine torque to decelerate and change driving direction. This is convenient for the driver since it reduce jerk and stand-still time, though it is wasteful use of fuel.
- The loading phase is an extended (compared to the length of the cycle) low speed/high thrust driving episode. In this phase the transmission and hydraulics loads are mutually dependent, as described in Filla (2008).
- Driving and hydraulics usage is coordinated, and not separated in time.

The coupling between hydraulic and transmission load complicates controller evaluation. The coordination of traction and hydraulics use by the driver affects

the fuel consumption and makes it difficult to separate controller and driver effects in an even higher degree than for on-road vehicles. The evaluation would also have to weigh material moved to time and fuel used. In an optimization and simulation the evaluation would also require a model of the coupling, which means a model of the source pile. To avoid introduction of a complex pile model in the following optimization, no deviations from the desired position and force trajectories are allowed.

2.2 Derivation of load cycles from measurements

It is desirable that the optimization is made against realistic driving cycles, which consist of position and force trajectories. Such cycles can be derived from measurements on wheel loader usage. Section 2.2.1 describe a position and force trajectory formulation that suits the following optimization, and how these trajectories can be derived from the sensor signals available in an in production machine. The optimization also include a sensitivity analysis which require a load case with realistic disturbances. Section 2.2.2 describe how load cases with disturbances can be constructed from a measurement sequence.

2.2.1 Deterministic load

Measurements were made at the test-track of Volvo Construction Equipment in Eskilstuna, Sweden, with a machine without extra sensors. The operator performed a series of loading cycles, moving gravel from one pile to another. The cycles, as required for this project, can be condensed into the load components wheel speed ω_w, wheel torque T_w, hydraulic pump flow Q_H and hydraulic pump pressure p_H. The measurements available consist of the sensor signals displayed in Figure 2.2. In this figure continuous lines indicate mechanical connections and dashed lines indicate hydraulic connections.

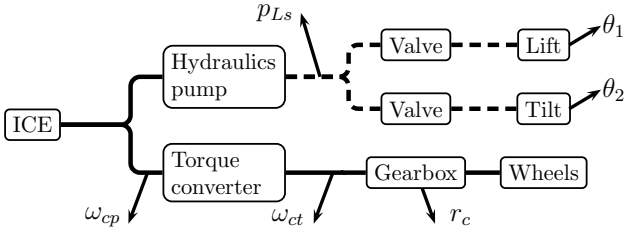


Figure 2.2: *Reference vehicle drivetrain and measurement setup.*

As seen, neither hydraulic flow nor wheel torque is directly measured. The hydraulic flow into the lift and tilt cylinders can instead be calculated from geometry and the derivative of the angles θ_1 and θ_2. Since the hydraulic fluid is near incompressible, this flow is approximately equal to the flow from the

working hydraulics pump Q_H. The pressure p_{LS} is used as hydraulics pump pressure p_H. The torque converter has two connections: the engine side which is denoted with index *cp* and the gearbox side which is denoted with index *ct*. The input and output torques of the torque converter depend only on the angular speeds at the input and output of the component. These torques can be calculated from the scalable maps M_P and μ, which have been measured at the reference speed $\omega_{c,ref}$, according to Equations (2.1).

$$\nu_c = \frac{\omega_{ct}}{\omega_{cp}} \tag{2.1a}$$

$$T_{cp} = M_P(\nu_c)\left(\frac{\omega_{cp}}{\omega_{c,ref}}\right)^2 \tag{2.1b}$$

$$T_{ct} = \mu(\nu_c)T_{cp} \tag{2.1c}$$

From ω_{ct}, T_{ct} and the engaged gear r_c the speed ω_w and torque T_w at the wheels are calculated. The torque T_{ct} includes the braking torque. Since no brake signal was available it was decided to include the brake torque in the load torque, though it would have been beneficial to instead have this as a control signal for the optimization. The deterministic load cases consist of one short loading cycle with a cycle time of $27s$; the 'DDP *sc*' cycle (Figure 2.3), and one long loading cycle; the 'DDP *lc*' with a cycle time of $137s$ (Figure 2.4).

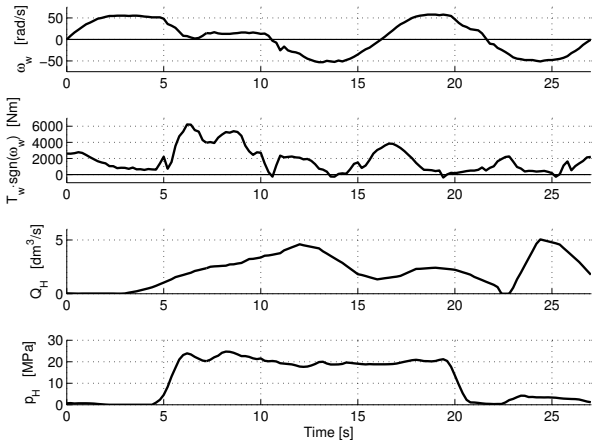


Figure 2.3: *The load case 'DDP sc'.*

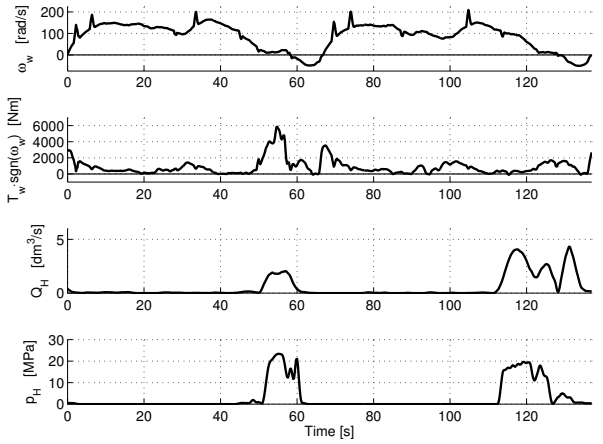


Figure 2.4: *The load case 'DDP lc'.*

2.2.2 Stochastic load

The sensitivity analysis of the proposed transmission concept requires a load that includes a disturbance model. This is created by introducing, for each time interval, several possible values for each load component W_Y and assigning probabilities $P(W_Y)$ to these. The probabilities are assumed to be independent

$$P(W_Y(t_k)|W_Y(t_{k-1})) = P(W_Y(t_k)) \tag{2.2}$$

This assumption affect which components of the load that may be described as non-deterministic. A stochastic vehicle speed ω_w, along with this assumption, would require non-physical accelerations. A more realistic disturbance would be in the acceleration, but this would require the vehicle speed to be a state. Therefore a deterministic ω_w is used. The same argument could be made for the hydraulic flow, since this roughly correspond to bucket raise speed. The hydraulic pump efficiency does on the other hand not depend on the bucket height, so the models do not require this as a state, and it is assumed that in general this height will not exceed its limits. Hydraulics flow is also supplied to the tilting of the bucket, and the lift angle dynamics is also affected by vehicle pitch dynamics. A fully stochastic hydraulic flow is therefore used. The deterministic ω_w also affect the wheel torque, since part of T_w depend on vehicle acceleration. This torque is divided into two parts; $T_w = T_A(\frac{d\omega_w}{dt}) + T_D$ where T_A depend on the acceleration and is deterministic, while T_D describe the rolling resistance, including force on the bucket, and is stochastic.

A loading cycle identifier has been developed for convenient measurement data analysis. This starts by detecting the discrete events '(change to) forward' f, '(change to) backward' b, 'bucket loading' l and 'bucket unloading' u. The identifier then search this event sequence for a pattern, described in automata language (see Kelley (1998)), which correspond to a loading cycle and which is

shown in Figure 2.5, and mark each occurrence. The identifier also mark the driving direction changes in each cycle. When all cycles in the dataset are found, the time scales are adjusted so that the direction changes in each cycle coincide. The measurement used consist of 34 short loading cycles with durations between $21.5s$ and $30.6s$. This time scale of each cycle has been changed to $10s$ forward toward the pile and filling the bucket, $5s$ reversing, $5s$ forward toward, and including, bucket emptying and finally $5s$ reversing; in total $25s$.

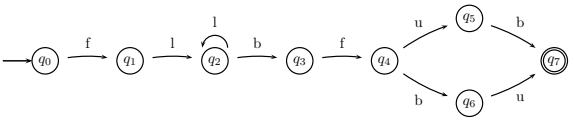


Figure 2.5: *Transition diagram of the automata which describe a loading cycle.*

At each time instant the mean E and standard deviation σ of each of the load components that are considered stochastic are calculated, among the 34 time-adjusted short loading cycles. The three load alternatives $W_Y = [E-\sigma, E, E+\sigma]$ of each of the three independent stochastic components $Y \in [T_D, Q_h, p_h]$ are created, while $W_Y = E$ for the components $Y \in [\omega_w, T_A]$. The probability vector $P(W_Y) = [0.25, 0.5, 0.25]$ is assigned to each of the components $Y \in [T_D, Q_h, p_h]$. The load $W(t) = [\omega_w, T_w, Q_h, p_h]$ along with the corresponding probability distributions $P(W_Y)$ makes the stochastic load cycle 'SDP *mc*'. Since this work examine the effect on the optimization of using a stochastic load, the deterministic reference case 'DDP *mc*' is also created, by only including the mean load $W_Y = E$, $Y \in [\omega_w, T_w, Q_h, p_h]$. In Figure 2.6 the load alternative of the 'DDP *mc*' cycle is indicated by the continuous lines, while the additional load alternatives of the 'SDP *mc*' cycle is indicated by the dashed lines.

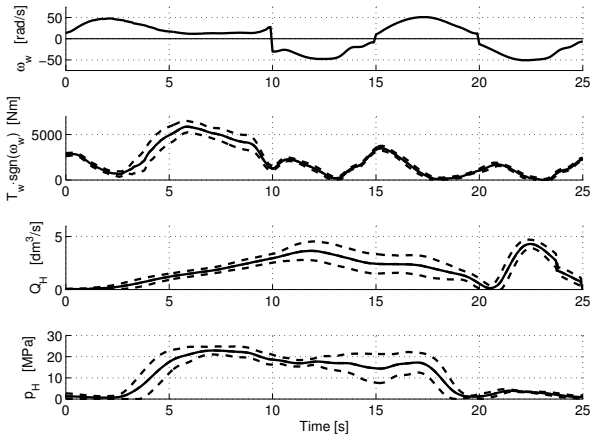


Figure 2.6: *The load cases 'DDP mc' and 'SDP mc'.*

3

Dynamic optimization

3.1 Problem statement

Dynamic optimization (DO) is the process of selecting some time dependent variable(s) in a dynamic system so that a cost function J is minimized (or maximized), according to Bryson (1999). The notation DO is most often used in the field of economics, such as in the book Kamien and Schwartz (1991), while in other fields the term optimal control is usually used, just as in Bryson (1975). The term 'optimal control' however is also commonly used for several other types of problems or specific subproblems. This often include optimal choice of parameters for causal controllers. For this reason, and for clarity, the phrase dynamic optimization will be used in this thesis. First the mathematical problem formulation will be sketched, then a few methods and principles for solving this type of problem will be presented.

Assume that the dynamic system can be described by a set of differential equations. Introduce the states $x(t)$ of the system, the decision variables, or control signals, $u(t)$ and the time dependent, non-controllable and possibly stochastic, signals $w(t)$. The signals $w(t)$ are often referred to as the disturbances, especially if they are stochastic. Here the disturbance signals are the load components, as introduced in Chapter 2.2. Let a dot represent time derivative, the dynamic system can then be written as

$$\dot{x} = F(x(t), u(t), w(t)), \quad x(0) = x_0 \tag{3.1}$$

A cost function, which is the target for the minimization, is formulated

$$J = \int_0^\infty G(x, u, w)\mathrm{d}t \tag{3.2}$$

The problem often has a finite terminal time, in which case the cost function is truncated and a cost J_N is assigned to the terminal state

$$J = J_N(x(T)) + \int_0^T G(x, u, w)\mathrm{d}t \tag{3.3}$$

In the problems of this thesis, the final time is also fixed, and not subject to optimization. If $w(t)$ is deterministic the problem can now be stated as

$$\begin{aligned} &\min_{u\in U} \left\{ J_N(x(T)) + \int_0^T G(x, u, w)\mathrm{d}t \right\} \\ &\dot{x} = F(x(t), u(t), w(t)) \\ &x(0) = x_0 \end{aligned} \tag{3.4}$$

along with possible state and control constraints. This problem is, regardless of the timespan, equivalent to an infinite dimension optimization problem. The problem is in general discretized for computerized numerical solving, transforming the problem into a large, but finite, dimensional optimization problem

$$\begin{aligned} &\min_{u\in U} \left\{ J_N(x(T)) + \sum_{k=0}^{N-1} g_k(u_k, x_k, w_k) \right\} \\ &x_{k+1} = f(x_k, u_k, t), \quad k = 0, \ldots, N-1 \end{aligned} \tag{3.5}$$

In case there are stochastic components in the time dependent variable $w(t)$, the minimization is instead made over the expected cost

$$\min_{u\in U} E\left\{ J_N(x(T)) + \int_0^T G(u, x, w)\mathrm{d}t \right\} \tag{3.6}$$

or

$$\min_{u_k\in U} E\left\{ J_N(x(T)) + \sum_{k=0}^{N-1} g_k(u_k, x_k, w_k) \right\} \tag{3.7}$$

There are several proposed approaches to solving problems of this type. The most intuitive may be to solve the problem (3.5) as a general large scale, but well structured, nonlinear minimization problem. This method works well for many problems, and may be expanded with more sophisticated discretization, due to the many good software packages that exist for nonlinear optimization. One problem is that this method requires a state and control signal initial guess, which may have to be rather good. Other methods are based on the optimality

principle, which states that if a trajectory is optimal for the problem (3.7) then all sections of this are optimal for the corresponding time intervals. Dynamic programming (DP) use this principle to recursively construct the optimal trajectory for the discrete optimization problem by solving a series of one-stage optimization problems, starting at the last time-step. This method is used both in Chapter 4 and in Chapter 5, and will therefore be described in more detail. A range of other methods are based on optimality criteria related to the calculus of variations. The basic fact of this field is that if x^*, u^* is an optimal solution, any variation of x or u will produce a higher cost J. Pontryagin's maximum principle (PMP) can be used for deriving a number of necessary conditions for a solution to a DO problem to be optimal. Since the conditions are necessary, but not sufficient, fulfilling these criteria does not guarantee global optimality for a general problem, but for some problems the criteria can be used for finding an analytic solution. There is a wide range of DO algorithms that are based on PMP. In the first two papers included in this thesis, PMP is directly applied to the problem. Therefore this principle is described further, along with short descriptions of some of the general PMP-based algorithms.

3.2 Dynamic programming (DP)

Dynamic programming is a recursive method for solving optimization problems which develop in stages, such as a discrete time. The method is based on the optimality principle, which states that if one trajectory u_k^*, x_k^*, $k = 0, \ldots, N-1$, is optimal for the problem (3.7) then each truncated trajectory u_k^*, x_k^*, $k = n, \ldots, N-1$, $0 < n$ is optimal for the corresponding time interval. This suggest that the problem can be solved as a series of one stage optimization problems, starting with the last stage and proceeding backward in time. According to Bellman (1957) and Bertsekas (2005) the recursion can, for the stochastic case, be stated as

$$J_k(x_k) = \min_{u \in U} E\big\{ g(x_k, u_k, w_k) + J_{k+1}(x_{k+1}(x_k, u_k, w_k)) \big\} \tag{3.8}$$

$$J_N(x_N) = g_N(x_N) \tag{3.9}$$

Details about stochastic dynamic programming (SDP) can be found in Ross (1983). If the external load w_k is deterministic the expectation E vanishes

$$J_k(x_k) = \min_{u \in U} \big\{ g(x_k, u_k, w_k) + J_{k+1}(x_{k+1}(x_k, u_k, w_k)) \big\} \tag{3.10}$$

in which case the method is labeled deterministic dynamic programming (DDP). The implementation of the recursion as an algorithm includes a strategic choice. Denote the discretized states $x \in X$. The 'cost-to-go', J_{k+1}, is then only calculated and stored at the grid points $x_{k+1} \in X$, and is not explicitly known for other $x_{k+1} \notin X$. The method selected for handling this highly affects the calculatory effort. Three possible choices are presented here.

If the function $x_{k+1}(x_k, u_k, w_k)$ is invertible, that is if $u_k(x_k, w_k, x_{k+1})$ is well defined, then $g + J_{k+1}$ can be evaluated for each $\{x_k, x_{k+1}\} \in X$ combination. With this choice the calculatory effort increase with the square of the size of X but is independent of the controls. If inverting $x_{k+1}(x_k, u_k, w_k)$ is not possible or desirable (for example if X is large) $x_{k+1}(x_k, u_k, w_k)$ can be calculated for the discretized $u \in U$, not requiring that $x_{k+1} \in X$. Then $\tilde{u}_k(x_k, w_k, x_{k+1} \in X)$ can be found by interpolation among these u_k, followed by the calculation of $g(x_k, \tilde{u}_k, w_k)$. Another option is to make the same calculation of $x_{k+1}(x_k, u_k, w_k)$, but to determine $\tilde{J}_{k+1}(x_{k+1}(x_k, u_k w_k))$ by interpolation among the $J_{k+1}(x_{k+1} \in X)$. In this case the calculatory effort increase linearly with the number of possible state and control combinations. In this thesis the third option is used, producing the following algorithm

1: For $x_N \in X_N$, declare $J_N(x) = J_N$
2: **for** $k = N-1, \ldots, 1$ **do**
3: For each $x_k \in X_k$, simulate $\frac{dx}{dt}$ for t_k to t_{k+1} for all $u \in U$ to find $x_{k+1}(x_k, u, w_k)$
4: For each $x_k \in X_k$

$$J_k(x_k) = \min_{u \in U} \big(g(x_k, u, w_k) + \tilde{J}_{k+1}(x_{k+1}(x_k, u, w_k)) \big) \tag{3.11}$$

in which $\tilde{J}_{k+1}(x_{k+1})$ is interpolated from $J_{k+1}(x_{k+1} \in X)$
5: **end for**

If the load is stochastic, step 3 is performed for each possible load combination $w_l \in W_k$, and Equation (3.11) is altered to

$$J_k(x_k) = \min_{u \in U} \sum_{w_l \in W_k} \big(P(w_l) g(x_k, u, w_l) + \tilde{J}_{k+1}(P(w_l) x_{k+1}(x_k, u, w_l)) \big) \tag{3.12}$$

This first part establishes a cost-to-go map $J(x \in X, t)$. In the following part the optimal trajectory $x^*(t), u^*(t)$ is calculated

1: Select an initial state $x_0^* = x_0$
2: **for** $m = 1, \ldots, N$ **do**
3: For x_{m-1}^*, simulate $\frac{dx}{dt}$ for t_{m-1} to t_m for all $u \in U$ to find $x_m(x_{m-1}^*, u)$
4: Select

$$u_{m-1}^* = \underset{u \in U}{\operatorname{argmin}} \big(g(x_{m-1}^*, u, w_{m-1}) \mathrm{d}t + \tilde{J}_m(x_m(x_{m-1}^*, u, w_{m-1})) \big) \tag{3.13}$$

in which $\tilde{J}_m(x_m)$ is interpolated from $J_m(x_m \in X)$
5: $x_m^* = x_m(x_{m-1}^*, u_{m-1}^*, w_{m-1})$
6: **end for**

In the second part the load w_k, $k = 0, \ldots, N-1$ is deterministic. This second part also show how DP can be used to implement an optimal state feedback scheme. In each repetition of the for-loop the optimal control action u_{m-1}^* is calculated, depending on the state x_{m-1}^*. Here the state x_{m-1}^* is found

by simulation, but in a feedback application the actual state of the system at $t = m - 1$ would be used instead. If there is then an unexpected state disturbance so that $\hat{x}_{m-1} \neq x^*_{m-1}$, in which $\hat{x}$ is the actual state of the system, the algorithm will find the control that minimizes the cost-to-go from this state $\hat{x}_{m-1}$. Apart from this attractive property, the method also guarantees that if a solution is found, this is the global optimum. This does however require that the state grid is sufficiently dense. Consider for example the situation in which $J_{k+1}(x^i_{k+1}) = \infty$ for $i \leq j$ in which x^i should be interpreted as the i:th component of X. Then there will have to exist $x_{k+1}(x^h_k, u) > x^{j+1}_{k+1}$ or else $J_k(x^h_k)$ will also be infinite. This is a situation that will more easily emerge if the state grid is made more sparse or the time grid is made more dense. An increased time resolution may therefore cause erroneous infinite-cost spread if the state resolution is not also increased. Guzzella and Sciarretta (2007) contains an introduction to dynamic programming, which also mention this pitfall.

3.3 Pontryagin's maximum principle (PMP)

Pontryagin's maximum (or minimum) principle is a condition necessary for optimality. Before the condition is stated, a function called the Hamiltonian is introduced

$$H = G(x(t), u(t), w(t)) + \lambda^T(t) F(x(t), u(t), w(t)) \tag{3.14}$$

in which G and F is the cost and dynamics functions from (3.4) and λ is a set of continuous functions with one component corresponding to each of the components of x. Then Pontryagin's maximum principle, which was presented in Pontryagin et al. (1962) and is described and used in Bryson (1975), state that for x^*, u^* to be optimal, λ^* must exist and

$$H(x^*, u^*, w, \lambda^*) \leq H(x^*, u, w, \lambda^*) \quad \forall u, t \in [t_0, T] \tag{3.15}$$

along with boundary conditions for λ^*, which depend on whether the final time T is fixed or subject of optimization, must be fulfilled. By differentiating H this condition can be rewritten as a set of necessary conditions. For the unconstrained problem (3.4)

$$\frac{\partial H}{\partial u} = 0 \tag{3.16a}$$

$$\frac{\partial H}{\partial x} = -\dot{\lambda} \tag{3.16b}$$

$$\frac{\partial H}{\partial \lambda} = \dot{x} \tag{3.16c}$$

$$x(0) = x_0, \quad \lambda(T) = \frac{\partial J_N}{\partial x}(x^*(T)) \tag{3.16d}$$

must be fulfilled for x^*, u^* to be optimal. Condition (3.16c) is trivially fulfilled, as can be seen by differentiating (3.14). If the problem includes state or control

constraints the Hamiltonian must be expanded, but the conditions (3.16) are sufficient for the analysis in Chapter 4. The general problem cannot be solved analytically, but there is a wide range of methods that are based on PMP, for solving DO problems. They generally relax some of the conditions, use an initial guess for the trajectories, and iteratively tighten the relaxed condition(s). A few types of methods are:

Gradient methods: The condition (3.16a) is relaxed, and an arbitrary control signal is applied. This control signal is iteratively updated until

$$\int_0^T \left| \frac{\partial H}{\partial u} \right|^2 \mathrm{d}t$$

become sufficiently small.

Shooting: The condition (3.16d) is relaxed, and an arbitrary $\lambda(0)$ is applied. The system is simulated for $t = [0, T]$ and $\lambda(0)$ is updated until

$$\left| \lambda(T) - \frac{\partial J_N}{\partial x}(x^*(T)) \right|$$

become sufficiently small. In multiple shooting methods the problem is divided into shorter time intervals to reduce instability problems.

Collocation: The condition (3.16b) and/or (3.16c) is relaxed. The problem is discretized and formulated as a large non-linear optimization problem, in which the differential equations are piecewise approximated by simple functions. These approximations are updated until the discontinuities at the discretization, or collocation, points become sufficiently small.

In some other special cases the application of the necessary conditions of PMP has also helped in finding some simple structure of the optimal control trajectory. One example can be found in Delprat et al. (2001), which show that in the case of hybrid electric vehicles, a simplification of the battery model can make the adjoint function $\lambda(t)$ become a cycle-specific constant. This constant tells the value of stored electricity and can be used in a causal controller that select whether to charge or discharge the battery. Though the value of the constant is only known if the cycle is known, this method, which is known as the 'Equivalent Consumption Minimization Strategy' or ECMS, has drawn much attention. The theory and several varieties and references are collected and presented in Sciarretta and Guzzella (2007).

3.4 Implementation

The results from dynamic optimization show how the vehicle should have been controlled if the disturbance, or load, $w(t)$ had been known beforehand. Predicting the future load is not a trivial problem, especially for off-road heavy equipment such as that treated in this thesis. The following sections provide and discuss examples of how to utilize and benefit from optimization results.

3.4.1 Benchmarking

The probably most straight-forward and obvious use of the dynamic optimization results is for evaluation of controllers or mechanical concepts. A controller can always be made to seem good by comparing it to other controllers that perform worse. As an alternative, the minimized cost can be used as an objective reference. This cost also show the magnitude of the unutilized potential of the physical system; how much better the controller may become. In Nilsson et al. (2011) and Nilsson et al. (2012c) the quality of the extremely simple but suboptimal and non-causal, state and control trajectory derivation is evaluated by comparing the cost when using suboptimal and optimal controls. In a similar way, the analysis in Nilsson et al. (2012a) and Nilsson et al. (2012b) is motivated by the need for a fair evaluation and comparison of the potential of two different mechanical systems, which strictly eliminate the influence of any controller.

3.4.2 Utilizing principles

The results of the optimization, or the optimization procedure itself, may in the best case reveal some underlaying structure of the optimal solutions. For example; Hellström et al. (2010) use the Euler-Lagrange equation from the calculus of variations to show that the optimal speed for a fixed time, fixed distance driving mission is constant and Delprat et al. (2002) use PMP to derive the causal ECMS control strategy for hybrid vehicles. In Nilsson et al. (2011) and Nilsson et al. (2012c) the general appearance of the optimal trajectories inspire the derivation of the presented suboptimal methods.

3.4.3 Predictive control

The basic assumption of the optimization is that the future load is known. If this assumption had been fulfilled the control trajectory from the optimization could be applied directly. Applying the optimal control signal trajectory, and assuming that this give the corresponding optimal state trajectory would be feed-forward control, so errors would get integrated and the state errors would eventually become unacceptably large. This could be augmented with a feedback controller that correct state deviations, but though this would correct state trajectory deviations, the control signals may instead deviate from the optimal with an adverse effect on the optimality. Dynamic programming has an advantage in this aspect; $J(x)$ is calculated not only near the optimal trajectory, but for the entire state space. This means that if there is an unexpected disturbance v_k that affect the state $\hat{x}_{k+1}(x_k^*, u_k, w_k + v_k)$ so that $\hat{x}_{k+1} \neq x_{k+1}^*$ a new control signal that is optimal when starting in $\hat{x}_{k+1}$ is found simply by solving

$$\min_{u \in U} \Big(g(\hat{x}_{k+1}, u_{k+1}, w_{k+1}) + \tilde{J}_{k+2}(x_{k+2}(\hat{x}_{k+1}, u_{k+1}, w_{k+1}) \Big) \tag{3.17}$$

in which $\tilde{J}_{k+2}(x_{k+2})$ is interpolated among the known $J_{k+2}(x_{k+2})$. This is identical to a step in the simulation for finding the optimal path, after the

calculation of $J(X,t)$, as described in Section 3.2. Still, $J(X,t)$ is only valid for the load, deterministic or stochastic, for which it has been calculated. Therefore this solution is optimal from the disturbed state $\hat{x}_{k+1}$ if and only if there are no more unexpected disturbances. The key to being able to directly use the optimization results is therefore to obtain a load prediction, in deterministic or stochastic form.

Predicting the future load

There are several possibilities for predicting the future load, that has been proposed or is used in different applications.

If the vehicle is made autonomous, new possibilities for prediction and optimization arise. The complete operation of the vehicle, including positioning, bucket movement and engine and transmission control may be co-optimized. This would allow a truly global optimum to be found, though the optimization would require a large number of states and controls, making the problem difficult to solve. If there are widely separated time constants, it may be a good approximation to partition the optimization into one part for the slow states and one for the fast. For wheel loaders it could be possible to separate vehicle and bucket movement from engine and transmission control. Regardless of whether the vehicle trajectories are optimized or not, they may for an autonomous vehicle be determined some time in advance, and therefore also be used as a load prediction for the engine and transmission. Research on autonomous wheel loaders can be found, for example, in Ghabcheloo et al. (2009) and Koyachi and Sarata (2009).

Careful selection of states and possibly changing the stage variable from time to some other may also open new possibilities for obtaining a prediction that is sufficient for optimization. In Hellström (2010) the speed of a truck is subject to optimization, and road inclination is the future load. This is solved by using GPS altitude data and changing the stage variable from time t to position or distance s, so that the inclination prediction become the known $W(s)$ instead of the state and control dependent $W(t)$.

There has also been some research on early detection of events based on driver behavior. In Mitrovic (2005) and Pentland and Andrew (1999) this includes detecting imminent turning in a crossing or overtaking of other vehicles, based on the first signs of positioning and braking or accelerating. In these detectors there is a finite number of events that can be detected. If these are not highly parametrized (parameters such as speed, power level etc) the optimal control solution for each event can be calculated in advance and applied when the event is detected. If there is a high number of parameters it might still be possible to use the prediction, but in that case the optimization may have to be performed online when the start of the event has been detected.

4

Engine

For an alternative transmission to be interesting for commercialization, it has to enable increased efficiency without reducing productivity in the typical operating conditions of the machine. Since the losses discussed in Chapter 1 are related to the torque converter, the solution should include elimination of this component. Elimination of the torque converter, combined with the low speeds often encountered, makes it impractical to use a stepped gearbox. The immediate alternative is to consider continuously or infinitely variable transmissions (CVTs or IVTs), such as the diesel-electric transmission used in Filla (2008) or the hydrostatic transmission used in Rydberg (1998). Such transmissions make it possible to separate power demand at the wheels, from power production at the engine, thereby enabling a free choice of engine operating point. This chapter examine this choice of operating point during transients for minimization of the amount of fuel used, for naturally aspirated and turbocharged engines.

4.1 System setup

In a series hybrid without energy storage or a CVT vehicle the transmission can approximately be divided into one power consuming and one power producing part. In an electric series hybrid the partitioning could be made at the electric connection by using electric power instead of voltage and current, in a hydraulic hybrid it could be made by using hydraulic power instead of pressure and flow, and in a belt type CVT it could be made by using belt power instead of belt force and speed. In this chapter it is assumed that the device has no maximum or minimum gear ratio. If such a partitioning can be made, any driving cycle

can be translated, including efficiencies on the power consuming side, to a power demand trajectory $P_{load}(t)$. The efficiencies in the power producing side of the transmission, see Figure 4.1, can be included in the engine efficiency.

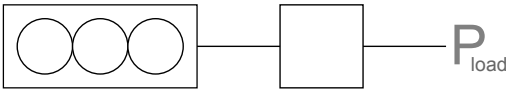


Figure 4.1: *The system consist of an engine, the engine side of an infinitely variable transmission (e.g. an electric generator) and an output power.*

The engines used in the papers Nilsson et al. (2011), Nilsson et al. (2012c), Nilsson et al. (2012a) has different maximum output powers. In this chapter the engine from Nilsson et al. (2012a) is used, since this is also the one used in Chapter 5 and is matched to the transmission of that chapter. Chapter 4 treats both a naturally aspirated and a turbocharged engine, while Chapter 5 only include the turbocharged engine. The differences between the setups are mentioned as they appear. The naturally aspirated engine is referred to as the NA-engine, while the turbocharged is referred to as the TC-engine.

Engine model

The engine speed ω_e dynamics is modeled as an inertia I_e which is affected by the engine torque T_e and a load power P_{load}.

$$\frac{d\omega_e(t)}{dt} \cdot I_e = T_e(t) - \frac{P_{load}(t)}{\omega_e(t)} \tag{4.1}$$

The engine torque T_e depend on fuel mass per injection m_f and engine speed ω_e according to a quadratic Willan's model, as described in Rizzoni et al. (1999). Introduce the lower heating value q_{lhv}, the number of cylinders n_{cyl}, the number of strokes per injection n_r and the parameters $\eta_{e00}, \eta_{e01}, \eta_{e02}, \eta_{e10}, \eta_{e11}, \eta_{eL0}, \eta_{eL2}$ and define

$$A = \frac{q_{lhv} n_{cyl}}{2\pi n_r} \tag{4.2}$$

$$\eta_e = \eta_{e0} - \eta_{e1} m_f \tag{4.3a}$$

$$\eta_{e0} = \eta_{e00} + \eta_{e01}\omega_e + \eta_{e02}\omega_e^2 \tag{4.3b}$$

$$\eta_{e1} = \eta_{e10} + \eta_{e11}\omega_e \tag{4.3c}$$

$$\eta_{eL} = \eta_{eL0} + \eta_{eL2}\omega_e^2 \tag{4.3d}$$

The Willan's model, expanded with an additional torque loss T_t caused by lack of air intake pressure, can then be described by Equation (4.4). The torque loss T_t is introduced for the modeling of the turbocharged engine, and for the naturally aspirated engine this loss is zero $T_t = 0$.

$$T_e = A \cdot \eta_e \cdot m_f - \eta_{eL} - T_t \tag{4.4}$$

The engine is also subject to the state and control restrictions

$$\begin{aligned} &\omega_{e,min} \leq \omega_e \\ &0 \leq m_f \\ &T_e \leq T_{e,max}(\omega_e) \end{aligned} \tag{4.5}$$

Turbocharger model

The torque loss T_t is caused by low air intake pressure, a pressure which depend on the rotational speed of the turbocharger. The turbocharger speed is assumed to be a first order dynamic system with the the time constant $\tau_t(\omega_e)$ and an asymptotic speed that is a function of ω_e, m_f. The dynamic relations are expressed in the corresponding asymptotic and dynamic air intake pressures. Denote the asymptotic intake pressure by $p_{t,set}$ and the time dependent pressure by p_t. Introduce the model and effeiciency parameters $\xi_{\tau 0}$, $\xi_{\tau 1}$, ξ_{t1}, ξ_{t2}, ξ_{t3}, η_{t10}, η_{t11}, η_{t20} and η_{t21} and define

$$\tau_t = \xi_{\tau 0} + \xi_{\tau 1}\omega_e \tag{4.6a}$$
$$p_{t,set} = \xi_{t1}\omega_e + \xi_{t2}m_f + \xi_{t3} \tag{4.6b}$$
$$\eta_{t1} = \eta_{t10} + \eta_{t11}\omega_e \tag{4.6c}$$
$$\eta_{t2} = \eta_{t20} + \eta_{t21}\omega_e \tag{4.6d}$$

The pressure dynamics can then be described by

$$\frac{dp_t(t)}{dt} \cdot \tau_t(\omega_e) = p_{t,set}(\omega_e, m_f) - p_t(t) \tag{4.7}$$

By defining $p_{t,off} = p_{t,set}(\omega_e, m_f) - p_t$ the torque loss can then be described by

$$T_t = \begin{cases} \eta_{t1}(\omega_e) \cdot p_{t,off}^2 + \eta_{t2}(\omega_e) \cdot p_{t,off} & \text{if } p_{t,off} > 0 \\ 0 & \text{if } p_{t,off} \leq 0 \end{cases} \tag{4.8}$$

Efficiency definitions

The quasi-static peak efficiency line Σ is defined as the ω_e, T_e that maximize (4.9a) as a function of P_{load} under the restrictions (4.5) and $\frac{d\omega_e}{dt} = \frac{dp_t}{dt} = 0$ as described by the Equations (4.9).

$$\eta_{e,static} = \frac{P_{load}}{P_{m_f}} = \frac{T_e\omega_e}{\omega_e A m_f} \tag{4.9a}$$
$$\omega_{e,\Sigma}(P_{load}) = \underset{\omega_e}{\text{argmax}}\eta_{e,static}(P_{load}) \tag{4.9b}$$
$$m_{f,\Sigma}(P_{load}) = \underset{m_f}{\text{argmax}}\eta_{e,static}(P_{load}, \omega_{e,\Sigma}) \tag{4.9c}$$

The Equations (4.9) also define $T_{e,\Sigma} = T_e(\omega_{e,\Sigma}, m_{f,\Sigma})$. Individual points along the line Σ is referred to as (quasi) static optimal operating points or SOOPs.

4.2 Problem statement

The problem studied in this chapter is the minimization of the total amount of fuel used, according to Equation (4.10)

$$\min \int_0^T A\omega_e m_f \mathrm{d}t \tag{4.10}$$

while fulfilling the engine dynamics Equation (4.1) the constraints (4.5) and, in case the engine is turbocharged, the turbo dynamics (4.7). This also means that no deviations from the output load trajectory $P_{load}(t)$ will be allowed.

4.3 Load cases

In Equation (4.1) the time dependent load $P_{load}(t)$ is introduced. In this chapter two different types of loads are used. The first type is the total output power from the recorded 'DDP *sc*' and 'DDP *lc*' load cases, which is presented in Chapter 2. These output trajectories are calculated as $P_{load} = \omega_w T_w + Q_h p_h$, and are presented in Figure 4.2.

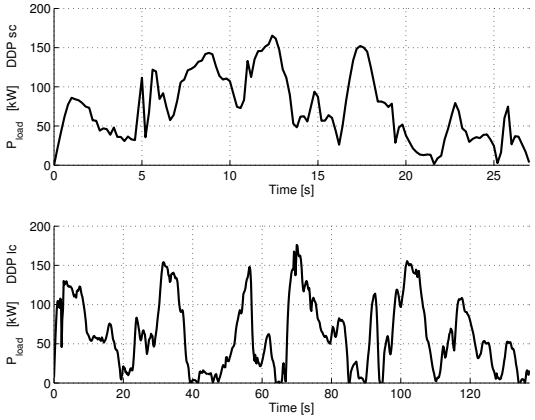


Figure 4.2: *The output power in the load cases 'DDP sc' and 'DDP lc'.*

The other type is artificial load cases, and consist of the four pulse and step cases presented in Table 4.1. The 'DDP *sc*' and 'DDP *lc*' load cases are applied to both engine setups while the pulse load cases are primarily intended for the NA-engine and the steps load cases are primarily intended for the TC-engine. In all four artificial load cases the time before the first and after the last steps are selected so that an increase in any of the times would not affect the result. The time scales in the pulse load cases are selected so that in the slow pulse the engine has time to settle at the static optimal operating point (SOOP) of the intermediate output power, while in the quick step it does not. Due to the

increased complexity of the TC-engine, this is only subjected to the steps load cases. The power levels in these load cases are selected so that the low step is between two SOOPs on the minimum engine speed limit, while both of the SOOPs of the high step are above this limit ($\sim 85kW$).

Table 4.1: *Stylized load cases for the standalone engine.*

Name	Load case: Power(Duration)
Slow pulse	$100kW(5s)$-$180kW(5s)$-$100kW(5s)$
Quick pulse	$100kW(5s)$-$180kW(0.8s)$-$100kW(5s)$
Low step	$50kW(5s)$-$80kW(5s)$
High step	$100kW(5s)$-$180kW(5s)$

4.4 Application of optimization

The application of dynamic programming to this problem is straightforward. The cost to be minimized is the total amount of fuel used. In general this cost formulation will cause all energy stored in the system to be drained at the end of the cycle. Here this would be seen as the engine speed approaching $\omega_{e,min}$, regardless of the terminal output power. Especially for output power steps and pulses, it is instead desired that the engine settle at the SOOP corresponding to the terminal output power. This can be stated as $x(T) = x_\Sigma(T)$. Due to the state discretization for DP, this cannot be exactly achieved. Since the energy in the system increase with increasing $\omega_e(T), p_t(T)$, introducing a J_N with a sufficient penalty for $\omega_e(T) < \omega_{e,\Sigma}(T)$, $p_t(T) < p_{t,set}(\omega_{e,\Sigma}(T), m_{f,\Sigma}(T))$ is sufficient for bringing the end state toward the static optimal operating point. In this work the terminal cost

$$J_N = \begin{cases} 0 & \text{for } x_N \geq \Omega \\ \infty & \text{else} \end{cases} \tag{4.11}$$

is used, with Ω being equal to $x_\Sigma(P_{load}(T))$ except when stated otherwise. The states and controls for the two engine setups are collected in Table 4.2.

Table 4.2: *Standalone engine states and controls.*

	NA-engine	TC-engine
States X	ω_e	ω_e, p_t
Controls U	m_f	m_f

Also recapitulate the PMP conditions for these two setups. For the unconstrained TC-engine the Hamiltonian become

$$H = A\omega_e m_f + \frac{\lambda_1}{I_e}(T_e - \frac{P_{load}}{\omega_e}) + \frac{\lambda_2}{\tau_t}(p_{t,set} - p_t) \tag{4.12}$$

in which λ_1 is the adjoint variable related to the engine speed dynamics (4.1) and λ_2 is the adjoint variable related to the turbo pressure dynamics (4.7). This gives the following conditions necessary for optimality

$$\frac{\partial H}{\partial m_f} = A\omega_e + \lambda_1 \frac{\partial}{\partial m_f}\frac{d\omega_e}{dt} + \lambda_2 \frac{\partial}{\partial m_f}\frac{dp_t}{dt} = 0 \tag{4.13a}$$

$$\frac{\partial H}{\partial \omega_e} = Am_f + \lambda_1 \frac{\partial}{\partial \omega_e}\frac{d\omega_e}{dt} + \lambda_2 \frac{\partial}{\partial \omega_e}\frac{dp_t}{dt} = -\frac{d\lambda_1}{dt} \tag{4.13b}$$

$$\frac{\partial H}{\partial p_t} = \lambda_1 \frac{\partial}{\partial p_t}\frac{d\omega_e}{dt} + \lambda_2 \frac{\partial}{\partial p_t}\frac{dp_t}{dt} = -\frac{d\lambda_2}{dt} \tag{4.13c}$$

The optimality conditions for the unconstrained NA-engine can be retrieved by using $\lambda_2 = 0$ and disregarding equation (4.13c).

4.5 Engine map and static optimal solution

The quasi-static optimal line Σ is defined in (4.9). The Σ for the turbo engine is identical to that of the naturally aspirated engine, since $\dot{p}_t = 0 \Rightarrow T_t = 0$. This is a simple problem which can be solved either direct as the problem (4.9) or by solving the PMP problem with $\frac{d}{dt}[\omega_e, \lambda_1, p_t, \lambda_2] = 0$. The later is valid only when the solution fulfills $\omega_{e,min} \leq \omega_e$ though, since the state and control constraints is not included in the presented PMP formulation. The engine efficiency map given by equations (4.2) to (4.5) is presented in Figure 4.3 along with $\omega_{e,min}$, $T_{e,max}$, output power ($T_e\omega_e$) lines and Σ.

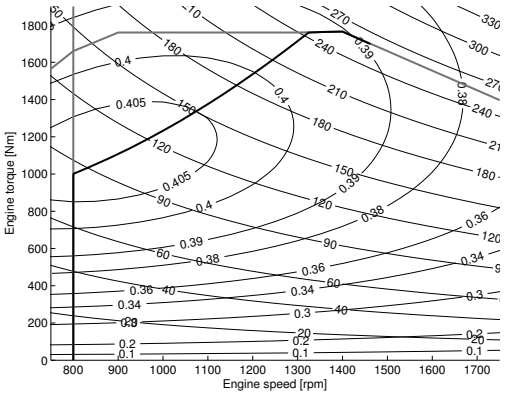
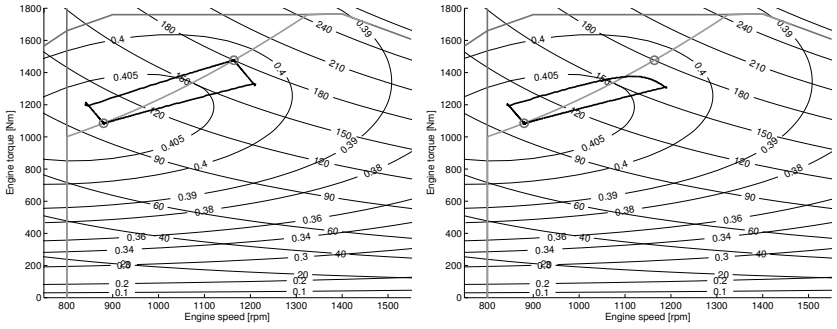


Figure 4.3: *Engine map showing efficiency curves, output power lines with kW markings, state and control restrictions according to* (4.5) *and the quasi-static optimal line which for output powers below* $\sim 85kW$ *coincide with* $\omega_{e,min}$ *and above* $\sim 240kW$ *with* $T_{e,max}$.

4.6 DP derived optimal trajectories

The optimal engine map trajectories for the pulse load cases for the NA-engine are presented in Figure 4.4. In both these cases the operating point move in a counter clockwise direction; before the output power increase the operating point diverges toward high speed. When the step occur, the operating point motion changes direction toward the new static optimum by reducing the speed and increasing the torque. Before the power reduction the engine speed decreases, and at the step the motion changes direction and the speed increases while the torque falls and the operating point converges to the new static optimum.

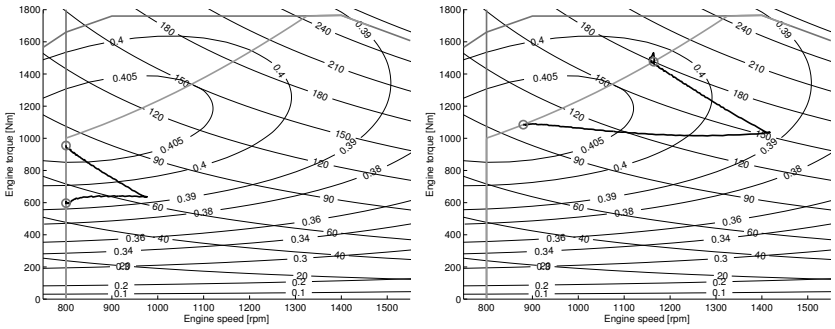


(a) (b)

Figure 4.4: *Engine map trajectories for the naturally aspirated engine in the slow (4.4(a)) and quick (4.4(b)) pulse load cases.*

The optimal engine map trajectories for the steps load cases for the TC-engine are presented in Figure 4.5. Just as for the NA-engine, the engine speed increase before the step, and when the step occur the direction of movement of the operating point changes. After the step the engine speed drops while the torque increases, converging toward the new static optimum. Both the trajectories displayed in Figure 4.5 are less smooth than those for the NA-engine. This is caused by a somewhat sparse discretization, which is motivated by the increase in calculation time caused by the extra state.

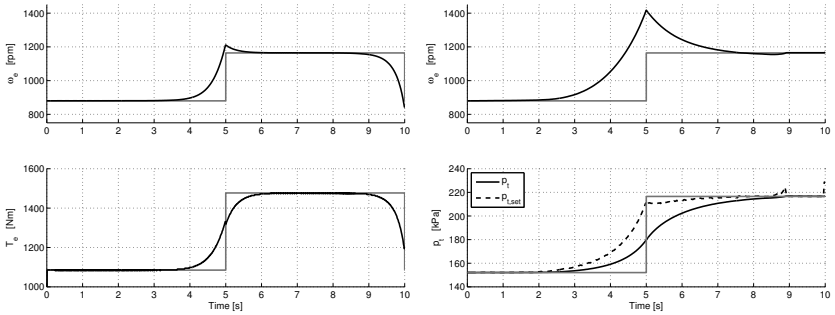
In Figure 4.6 the engine operation trajectories of the NA- and TC-engines are compared. Figure 4.6(a) shows the engine speed and torque during the first $10s$ of the slow pulse load case for the NA-engine and Figure 4.6(b) shows the engine speed and turbo-pressure during the high step load case for the TC-engine. The load case parts are identical, apart from that the NA-engine does not need to remain at the higher SOOP at $10s$. The NA-engine starts changing its state about one second before the step, while the TC-engine starts about three seconds before the step. Note that while both setups cause a speed overshoot, this is substantially larger for the TC-engine. Figure 4.6(b) shows that before the step, the increasing engine speed alters the turbo set-pressure



(a) (b)

Figure 4.5: *Engine map trajectories for the turbocharged engine in the low (4.5(a)) and high (4.5(b)) step load cases.*

so that it is roughly at the new static optimal level when the step occur. The actual pressure start to increase as soon as the set pressure starts to change, but at the time of the step it still is far from the new static level. After the step, the pressure keep increasing while the set pressure remain fairly constant and the engine speed falls back toward the new static optimum.



(a) Engine speed and torque in the slow pulse load case for the NA-engine

(b) Engine speed and turbo pressure in the high step load case for the TC-engine

Figure 4.6: *Engine operation during steps for the NA- and TC-engines.*

Figure 4.7 shows the engine map trajectories for the two engine setups in the short loading cycle. These trajectories should be compared to those in Figures 4.4 and 4.5. The movement is still counter clockwise, and the patterns of the movement remain, though the direction changes are less pronounced than in the solutions for the steps and pulses load cases since the output power changes are more ramped. The engine speed is generally higher for the TC-engine (972rpm mean) than for the NA-engine (861rpm mean), which is caused

by the need for keeping the turbo pressure up. It should be noted that this is despite having access to perfect prediction of future load. Note that the initial operating point for the TC-engine is at a much higher engine speed than for the NA-engine. The initial conditions $x(t_0)$ are selected so that the results could be readily used for evaluation of the suboptimal methods described in Section 4.8.

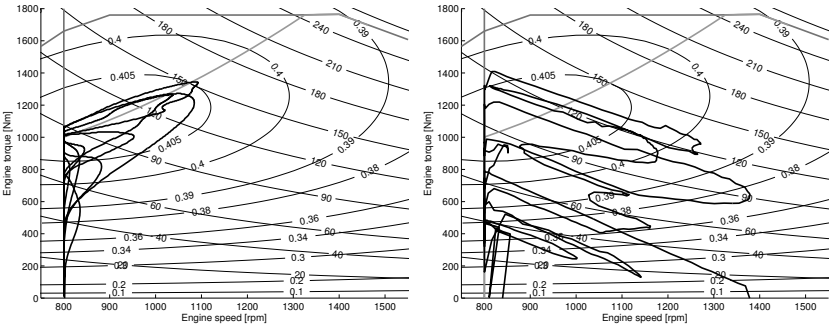


Figure 4.7: *Engine map trajectories for the naturally aspirated (4.7(a)) and the turbocharged (4.7(b)) engine in the 'DDP sc' cycle.*

4.7 PMP trajectory derivation

The solution to a DO problem must fulfill the conditions stated by Pontryagin's maximum principle (PMP). Section 4.7.1 analyze the NA-engine step/pulse results presented in Section 4.6 using these conditions. In Section 4.7.2 this analysis is utilized for developing a method for deriving the same optimization results. Section 4.7.3 expands this method for application on the TC-engine.

The PMP formulation in Section 4.4 does not include the constraints (4.5). A solution to the unconstrained problem (4.10) for a specific load $P_{load}(t)$ is optimal also for the constrained problem if and only if it does not violate the constraints (4.5). It is obvious that solutions for the unconstrained problem for steps to or from loads with $\omega_{e,\Sigma}(P_{load}) = \omega_{e,min}$ will violate these constraints. Therefore this section only treat load cases with $\omega_{e,\Sigma}(P_{load}) > \omega_{e,min}$.

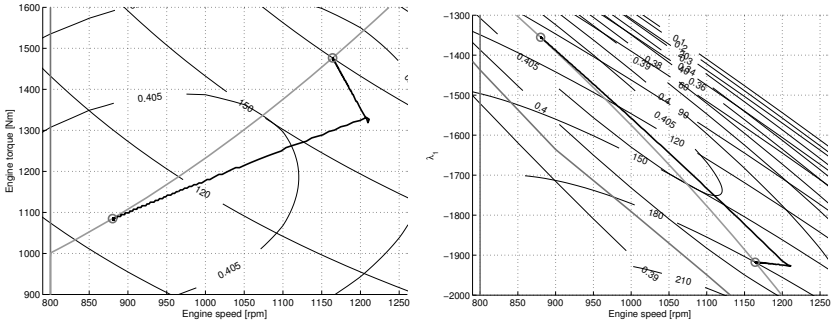
4.7.1 Analysis of optimization results

This analysis treats the high step load case, which is identical to the first part of the slow pulse load case, applied to the NA-engine. The DP result for the slow pulse load case is presented in Figure 4.4(a), and the part used is presented again in Figure 4.8(a). Equation (4.13a) can be used for transformation of positions in an ω_e-T_e engine map into an ω_e-λ_1 engine map. For the NA-engine

this relation can be rewritten as

$$\lambda_1 = \frac{\omega_e I_e}{2\eta_{e1} m_f - \eta_{e0}} \tag{4.14}$$

Figure 4.8(b) shows such a transformation of the map of Figure 4.8(a), including efficiency curves, output power lines with kW markings, the static optimal line Σ, the constraints (4.5) and the DP derived optimal operating point trajectory. In Figure 4.8(a) the trajectory starts at the lower left, moving toward the upper right, and when the step occur the direction of motion changes so that the maximum engine speed occur at the instant of the step. In Figure 4.8(b) this translates to initial movement toward the lower right and a change of direction of motion at the instant of the step.



(a) (b)

Figure 4.8: *DP derived Optimal solution for the high step load case in ω_e-T_e (4.8(a)) and ω_e-λ_1 (4.8(b)) engine maps.*

The dynamics of the adjoint variable $\lambda_1(t)$ is described by Equation (4.13b) (with $\lambda_2 = 0$). This equation can for the NA-engine be rewritten as

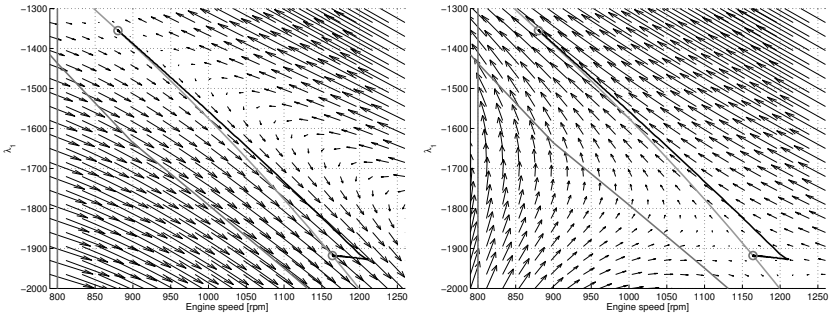
$$\dot{\lambda}_1 = -Am_f - \frac{\lambda_1}{I_e}\left(\frac{\partial T_e}{\partial \omega_e} + \frac{P_{load}}{\omega_e^2}\right) \tag{4.15}$$

in which

$$\frac{\partial T_e}{\partial \omega_e} = (\eta_{e01} + 2\eta_{e02}\omega_e - \eta_{e11} m_f) A m_f - 2\eta_{eL2}\omega_e \tag{4.16}$$

Since Equation (4.14) eliminate the only degree of freedom, all dynamics of the optimal solution is governed by Equations (4.1) (the engine speed) and (4.15) (the adjoint variable). The properties of a two dimensional autonomous dynamic system can be visualized by phase planes. The time dependent load means this system is not autonomous, though for piecewise constant loads, such as steps or pulses, the system can be regarded as piecewise autonomous. The phase planes for the system (4.1),(4.15) at the two output power levels

of the high step load case are presented in Figure 4.9. The figure also shows the constraints (4.5), the static optimal line Σ and the DP-derived optimal trajectory, as shown in Figure 4.8(b).



(a) (b)

Figure 4.9: *DP derived Optimal solution for the high step load case along with the* $100kW$ *(4.9(a)) and* $180kW$ *(4.9(b))* ω_e-λ_1 *phase planes.*

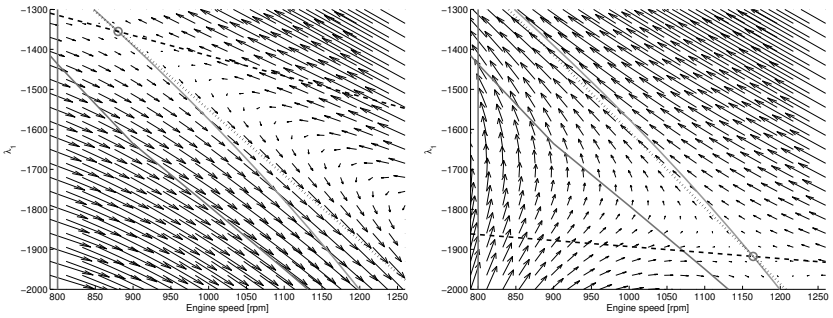
Figure 4.9 shows the dynamics behind the optimal solution for the high step load case. The first segment, the movement toward the lower right, occur when $P_{load} = 100kW$ and is therefore governed by the $100kW$ phase plane (Figure 4.9(a)), while the second segment, the approach of the second SOOP, is governed by the $180kW$ phase plane (Figure 4.9(b)). Section 4.7.2 starts with these phase planes and presents a method not only for visualizing but also for deriving the optimal solutions for similar load cases.

4.7.2 Optimal trajectory derivation for the NA-engine

This section shows how the reasoning in the previous section can be reversed and optimal trajectories be derived from the PMP conditions. The phase planes shown in Figure 4.9 indicate that, for each constant P_{load}, the SOOP is a saddle point of the corresponding autonomous system (4.17). This is confirmed by the eigenvalues of the Jacobian of this system, evaluated at the corresponding SOOP, since one is positive and the other is negative.

$$\frac{d}{dt}[\omega_e, \lambda_1]^T(P_{load}) \tag{4.17}$$

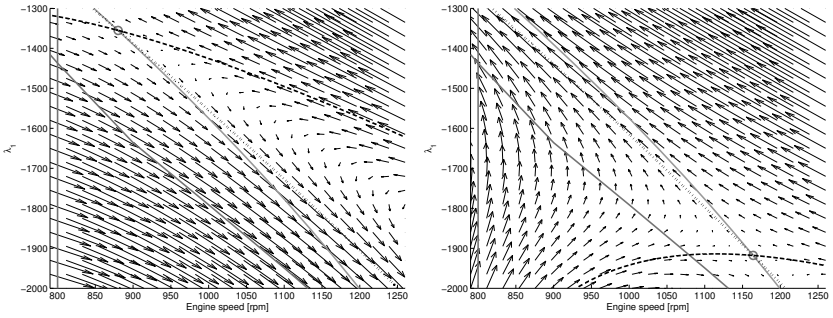
The unstable and stable manifolds of the autonomous system can, in a small region near the SOOP, be approximated by the eigenvectors of the Jacobian. The stable (dashed) and unstable (dotted) eigenvectors and the previously presented phase-planes corresponding to $P_{load} = 100kW$ and $P_{load} = 180kW$ are shown in Figure 4.10. More accurate approximations of the manifolds, valid outside the vicinity of the SOOP, can be obtained by simulations backward in time for the stable manifolds and forward in time for the unstable manifolds



(a) (b)

Figure 4.10: *Phase planes along with stable (dashed) and unstable (dotted) eigenvectors of the Jacobian of the dynamic system* (4.17) *with* $P_{load} = 100kW$ *(4.10(a)) and* $180kW$ *(4.10(b)).*

initiated from the SOOP with small, ε, disturbances in the directions of the eigenvectors. The result of such simulations, corresponding to the situations of Figure 4.10, are displayed in Figure 4.11.



(a) (b)

Figure 4.11: *Simulation derived stable (dashed) and unstable (dotted) manifolds of the system* (4.17) *with* $P_{load} = 100kW$ *(4.11(a)) and* $180kW$ *(4.11(b)).*

The optimal operating point trajectory for an output power step (in this example $100kW - 180kW$) which start and end at the SOOPs of the initial and terminal output powers, must start by leaving the first SOOP along a path in the unstable manifold of the earlier autonomous system. At the instant of the step the operating point must switch to a path in the stable manifold of the later autonomous system. Since the trajectory must be continuous the operating point must be at an intersection of these manifolds at the instant of the step. In general there is only one such intersection, which is easily found from the

simulated paths. When the point of intersection is found the excess parts of the simulated paths are cropped of and the time-scales of the simulations behind Figure 4.11 are adjusted so that a single, continuous, $\omega_e(t)$,$\lambda_1(t)$ trajectory is obtained. This trajectory is then the optimal solution. Graphically, this solution can be found by simply superposing Figure 4.11(a) with Figure 4.11(b) and cropping of excessive parts of the paths. Figure 4.12 shows the results as derived with this method (continuous) and with dynamic programming (dashed) for the upward and downward steps of the slow pulse load case. This solution can then be translated into an $\omega_e(t)$,$T_e(t)$ trajectory by Equation (4.14).

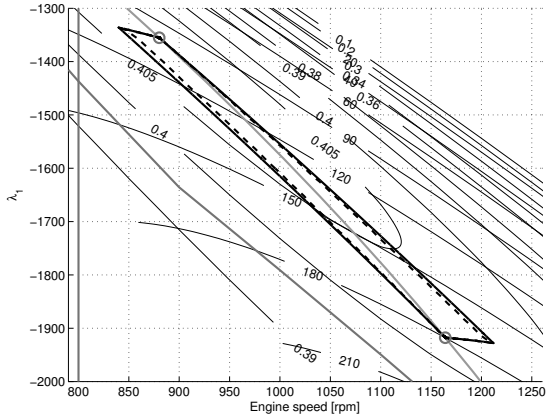
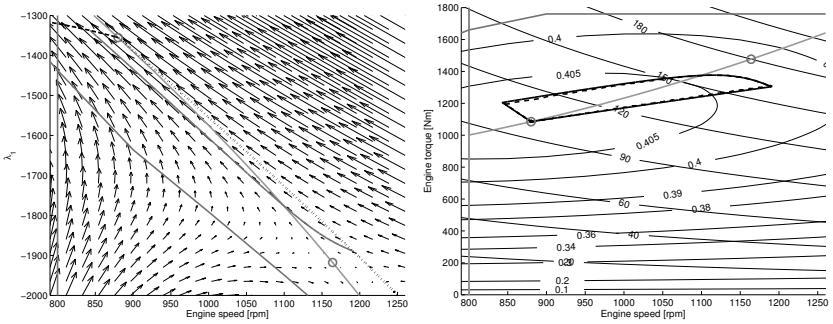


Figure 4.12: *PMP (continuous) and DP (dashed) derived optimal solutions for the slow pulse load case.*

This method can be expanded to somewhat more complicated load cases. If the case starts and ends with episodes of constant power, the optimal $\omega_e(t)$,$\lambda_1(t)$ trajectory must start with a leaving of the SOOP of the initial output power along the corresponding unstable manifold, and end with an approach of the SOOP of the terminal output power along the stable manifold. This is illustrated in Figure 4.13 by the solving of the quick pulse load case. This case consist of $5s$ at $100kW$, $0.8s$ at $180kW$ and finally $5s$ at $100kW$. The optimal trajectory must therefore start with a leaving of the $100kW$ SOOP along a path in the corresponding unstable manifold (dotted) and end by approaching the same SOOP along the stable manifold (dashed). Solving the quick pulse optimization problem therefore translates to finding a path in the $180kW$ phase plane, as shown in the figure, that starts on the dotted line, ends on the dashed line and has a transition time $t_T = 0.8s$. If the starting point of the transition is at t_i from the initial SOOP along the unstable manifold, the problem can be formulated as $\min_{t_i} |t_T - 0.8|$, which is locally convex, making the problem easily solved. The resulting transition trajectory is indicated in Figure 4.13(a) by the gray line. In Figure 4.13(b) this solution (continuous) is translated to an ω_e,T_e map and compared to the solution derived with DP (dashed).



(a) (b)

Figure 4.13: *Illustration of the PMP-method for solving the quick pulse load case. Figure 4.13(a) shows the stable and unstable* $100kW$ *manifolds along with the* $180kW$ *phase plane and the* $0.8s$*,* $180kW$ *transition path. In Figure 4.13(b) the PMP (continuous) and DP (dashed) derived optimal solutions are compared.*

4.7.3 Optimal trajectory derivation for the TC-engine

This section expands the method derived in the previous section for use with the TC-engine. The optimal solutions for the TC-engine is governed by the four dynamics Equations (4.18) and the static control relation (4.13a). The four dimensions of this problem means that phase planes can no longer be drawn and the problem can therefore not be solved graphically.

$$\frac{d}{dt}[\omega_e, \lambda_1, p_t, \lambda_2](P_{load}) \tag{4.18}$$

The formulation of the torque loss T_t in Equation (4.8) may cause discontinuities in the optimality conditions (4.13) due to the differentiation, which severely complicates simulation. One solution may be to approximate the discontinuities with a tangent function. In a step however it can instead be assumed that the intake pressure will not cross the discontinuity; p_t will fulfill $p_t < p_{t,set}$ in an upward step and $p_{t,set} < p_t$ in a downward step, so that for steps the discontinuity can be disregarded. In this section, just as in the previous, the upward high step load case is studied.

In the same way as for the NA-engine, the Jacobian of the system (4.18) is evaluated at the SOOPs of, in this example, $P_{load} = 100kW$ and $P_{load} = 180kW$ and the eigenvalues are calculated. These show that the SOOPs are saddle points, since two of the four eigenvalues are positive while the other two are negative. For the NA-engine, the optimization problem is easily solved since the trajectories simulated and presented in Figure 4.11 covers the entire stable and unstable manifolds within the reasonable engine operating region, and the point of intersection is easily found. For the TC-engine however, each of the manifolds are two dimensional. Calculation of the complete unstable manifold

would therefore require infinitely many simulations, initiated from the SOOP with small disturbances in all directions that are combinations of the eigenvectors corresponding to the positive eigenvalues, and vice versa for the stable manifold. Recall however that the objective is not to find the manifolds, but only the trajectories within these manifolds that connect the SOOPs of the initial and terminal P_{load}. Since the manifolds are two dimensional and the state space is four dimensional, there is in general a single point at which the manifolds intersect, and therefore only one combination of eigenvectors that produce trajectories that intersect. Since the location of the intersection is unknown, the problem is reformulated as a problem of finding the combination of eigenvectors that minimizes the minimum distance between the simulated trajectories. Similar problems are treated for example in Dellnitz et al. (2001). Denoting the initial and terminal output powers P_1 and P_2 and using the notation $v_{1,1}, v_{1,2}$ for the unstable eigenvectors corresponding to P_1 and $v_{2,1}, v_{2,2}$ for the stable eigenvectors corresponding to P_2 the problem is formulated as

$$\min_{s_1,t_1,s_2,t_2} \|X_1(P_1,t_1) - X_2(P_2,t_2)\|_2 \tag{4.19}$$

$$0 < [t_1, -t_2]^T, \quad 0 \leq [s_1, s_2]^T \leq 2\pi \tag{4.20}$$

in which

$$X_n = [\omega_e, \lambda_1, p_t, \lambda_2]^T(P_n, t_n), \quad n = 1, 2 \tag{4.21}$$

are simulated from $t_n = 0$ forward and backward in time with initial conditions that are small, ε, perturbations from the SOOPs according to

$$X_n(t_n = 0) = X_\Sigma(P_n) + \varepsilon\big(\sin(s_n)v_{n,1} + \cos(s_n)v_{n,2}\big), \quad n = 1, 2 \tag{4.22}$$

and the components of X_n in (4.19) being scaled with the average of the values of the component at the two SOOPs. Numerically this is solved as one external and one internal minimization problem. The external minimizes $\|X_1 - X_2\|_2$ over the disturbance direction combination s_1, s_2. Inside this, with s_1, s_2 given, $X_1(0<t_1)$,$X_2(t_2<0)$ is simulated and the minimum distance between the trajectories is determined by minimizing $\|X_1 - X_2\|_2$ over t_1, t_2. Each of the two internal simulations start at $t_1 = t_2 = 0$ and proceed until some state leave a predefined reasonable operating range. If a solution to the problem is found, the result of (4.19) should approach 0. The resulting point $X_1(t_1) \approx X_2(t_2)$ is then the intersection of the manifolds. This is the point at which the output power step occur and the operating point movement switch from one manifold to the other. Finally the times are shifted so that t_1 and t_2 coincide with the instant of the step. The result is a continuous operating point trajectory that start at $X_\Sigma(P_1)$, ends at $X_\Sigma(P_2)$ and has the step correctly placed in time.

The method is illustrated by the high step load case. Figure 4.14 shows the static optimal line (gray), the SOOPs (markers), the unstable (dotted) and stable (dashed) trajectories and a dark gray line which indicate the position of

the minimum distance between the trajectories. Figure 4.15 shows the ω_e,T_e translated trajectories in an engine map. Figure 4.16 shows the time-adjusted unstable and stable engine speed and turbo pressure trajectories along with the DP-derived solution (gray). Typical calculation times experienced for finding this solution have been around $30s$, which is considerably faster than the more than $2500s$ needed for finding the solution with dynamic programming. On the other hand, this method works only for load steps and, since the engine speed overshoots are larger for the TC-engine than for the NA-engine, at a narrow output power range.

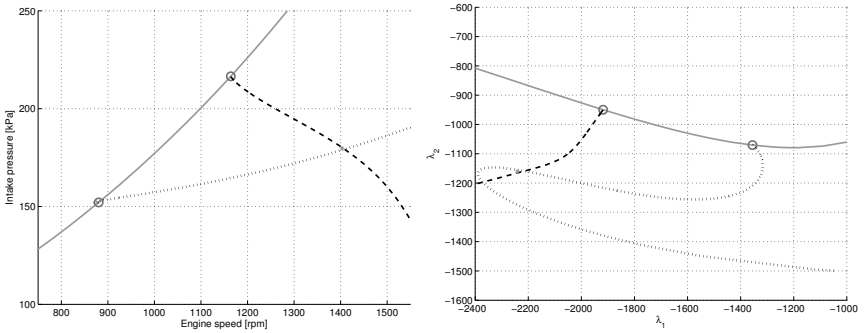


Figure 4.14: *Intersecting stable (dashed) and unstable (dotted) trajectories for the high step load case in ω_e,p_t (Figure 4.14(a)) and λ_1,λ_2 (Figure 4.14(b)) maps. The minimum distance between the trajectories is marked with gray.*

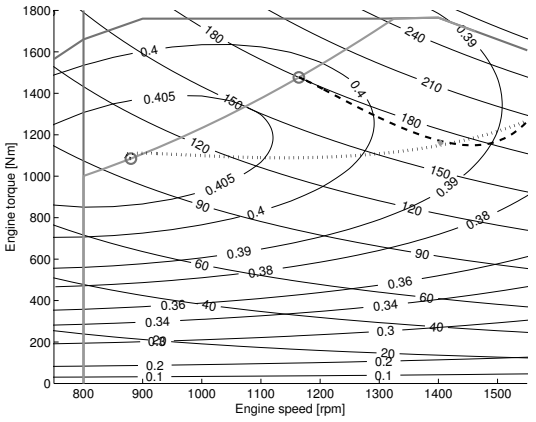


Figure 4.15: *Intersecting stable (dashed) and unstable (dotted) trajectories for the high step case in an ω_e,T_e map. Note the minimum distance marker (gray).*

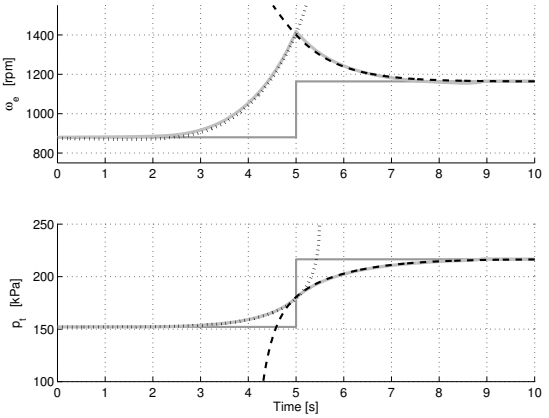


Figure 4.16: *PMP-derived (dotted & dashed) compared to DP-derived (gray) solution for the high step load case. The dark gray lines indicate* $X_\Sigma(P_{load}(t))$.

4.8 Suboptimal method development

Method for the NA-engine:

As mentioned, DP has several advantages but is slow while the PMP methods presented above are fast but very restrictive in which load cases can be treated. Another method which is fast and works for all load cases is desired, even if the resulting trajectories become suboptimal. Using $\omega_e(t) = \omega_{e,\Sigma}(P_{load}(t))$ is not possible, since output power steps would then imply engine speed steps. Inspiration for a method can instead be found in the optimal trajectories, for example in Figure 4.7. The operating point of the NA-engine seldom move far from the static optimal line Σ. A natural suboptimal strategy is to keep the operating point exactly on the line Σ at all times. Such a trajectory can be found by adding a large cost for deviation from this line to the DP algorithm, but solving this problem would be as computationally costly as solving the original problem. Instead start by redefining the static optimal line by introducing a small inclination in the minimum engine speed, so that at high torque the minimum speed is somewhat higher, to make $T_{e,\Sigma}(\omega_e)$ well defined. The rule

$$T_e(t) = T_{e,\Sigma}(\omega_e(t)) \tag{4.23}$$

then define the control signal, and thereby eliminate the only degree of freedom. The problem is therefore reduced from an optimization problem to finding the state and control trajectories that correspond to a set of admissible boundary conditions. Observe that as long as $T_{e,\Sigma}(\omega_e) \cdot \omega_e$ increase with increasing ω_e applying (4.23) will make the system unstable. This means that at the instant of an output power step the engine must already have exactly reached the terminal stationary operating point by a preceding divergence from the initial stationary operating point, initiated by a small disturbance. Since the system is always

unstable it can easily be simulated backward in time from an arbitrary terminal engine speed, for example using the Euler method according to Equation (4.24).

$$\omega_{e,k-1} = \omega_{e,k} - \left(\frac{T_{e,\Sigma}(\omega_{e,k})\omega_{e,k} - P_{load}}{\omega_{e,k} I_e} \right) \mathrm{d}t \tag{4.24}$$

This method works well, as illustrated by Table 4.3, for all cases tested. The table shows fuel usage in the solutions derived with DP and the suboptimal method, along with typical calculation times experienced. The same $x(T)$ is used in both methods and the $x(0)$ from the suboptimal method is used as initial condition for the DP solving. The last row shows the relative increase in fuel consumption and reduction of calculation time for the suboptimal method compared to DP. Figure 4.17 shows the suboptimal and optimal engine speed and torque trajectories. The $\omega_{e,\Sigma}(P_{load}(t))$, $T_{e,\Sigma}(P_{load}(t))$ trajectories that would have been applicable and indeed optimal for an engine with zero inertia I_e are included as a reference. The figure shows that the engine speed reacts somewhat later to upcoming load changes in the suboptimal solution than in the optimal. The example is a cutout from the 'DDP *sc*' load case.

Table 4.3: *Calculation effort and fuel usage with the suboptimal method.*

	Fuel usage $[ml]$		Calculation time $[s]$	
	DDP '*sc*'	DDP '*lc*'	DDP '*sc*'	DDP '*lc*'
Dynamic Programming	152.8	675.9	1270	6480
Suboptimal method	152.9	676.5	0.38	1.89
Relation	+.086%	+.099%	1 : 3340	1 : 3430

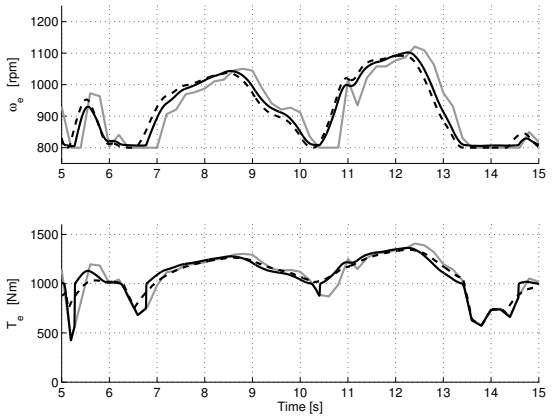


Figure 4.17: *Engine speed and torque. Gray is static optimum ($\omega_{e,\Sigma}(P_{load}(t))$,$T_{e,\Sigma}(P_{load}(t))$), continuous is suboptimal and dashed is optimal.*

Method for the TC-engine:

The expansion to the TC-engine is not trivial. The turbocharger stable in the forward direction, so it appears unstable in the backward direction and cannot be included in the simulation (4.24). It is tempting to derive an $\omega_e(t), T_e(t)$ trajectory while disregarding $p_t(t)$, and then simulate (4.7) forward in time while compensating for T_t with increased m_f. Unfortunately this is not possible for a general load case for this engine. This is most obvious for an upward step between two SOOPs with $\omega_{e,\Sigma} = \omega_{e,min}$. With this method, and with a neglectable minimum speed inclination, a step in P_{load} requires a step in T_e, and thereby in m_f. Equations (4.6)-(4.8) indicate that the p_t dynamics prevents making arbitrarily big steps in T_e simply by steps in m_f. It is therefore necessary to increase p_t in preparation for upcoming output power steps and/or to use power from the engine inertia I_e. Preparatory increasing of p_t has to be done by altering the engine speed and torque trajectories, possibly deviating from the static optimal line. The following algorithm is therefore proposed:

1) Find $\omega_e(t)$,$m_f(t)$ either by backward simulation of (4.1) assuming $p_{t,off} = 0$ or by assuming $I_e = \tau_t = 0 \Rightarrow \omega_e T_e = P_{load}, p_{t,off} = 0$ with $T_e = T_{e,\Sigma}(\omega_e)$.

2) Using $\omega_e(t)$,$m_f(t)$ from 1), simulate (4.7) forward in time to find a first estimate of $p_t(t)$, and thereby also of $T_t(t)$.

3) Update $\omega_e(t)$,$T_e(t)$ by simulating (4.1) backward in time while adding the result form 2) to the load; $T_e(t) = T_{e,\Sigma}(\omega_e) - T_t(t) = \frac{P_{load}}{\omega_e} - \frac{d\omega_e}{dt} I_e$.

4) Update $m_f(t)$,$p_t(t)$,$T_t(t)$ by simulating p_t forward in time, in each step solving Equations (4.1)-(4.6) for m_f so that $T_e = \frac{P_{load}}{\omega_e} - \frac{d\omega_e}{dt} I_e$.

If I_e=τ_t= 0 is assumed in step 1), this step can be performed inside step 2). After step 4) a feasible $\omega_e(t)$,$p_t(t)$,$m_f(t)$ trajectory has been found. This method works well for all cases tested, as illustrated by Table 4.4. The table shows the fuel usage in the trajectories derived with DP and the suboptimal method, along with typical calculation times experienced. The same $x(T)$ is used in both methods and the $x(0)$ from the suboptimal method is used as initial condition for the DP solving. This is also the cause of the high initial engine speed in Figure 4.7(b). The last row shows the relative increase in fuel consumption and reduction of calculation time for the suboptimal method compared to DP.

Table 4.4: *Calculation effort and fuel usage with the suboptimal method.*

	Fuel usage $[ml]$		Calculation time $[s]$	
	DDP 'sc'	DDP 'lc'	DDP 'sc'	DDP 'lc'
Dynamic Programming	154.8	701.0	6800	38500
Suboptimal method	157.2	725.2	2.10	10.2
Relation	+1.54%	+3.46%	1 : 3240	1 : 3800

An example of resulting engine speed and turbo pressure trajectories are compared to the optimal in Figure 4.18. The example is a cutout from the 'DDP *sc*' load case. The figure shows that while the suboptimal engine speed differs significantly from the optimal, the suboptimal turbo pressure trajectory is close to the optimal. Since the operating point is forced to leave the static optimal line, the engine map trajectories for the low and high steps load cases are also presented in Figure 4.19. In the high step load case the suboptimal and optimal trajectories are close. In the low step load case, just as in the 'DDP *sc*' case, the engine speed reacts later in preparation for upcoming loads in the suboptimal solution.

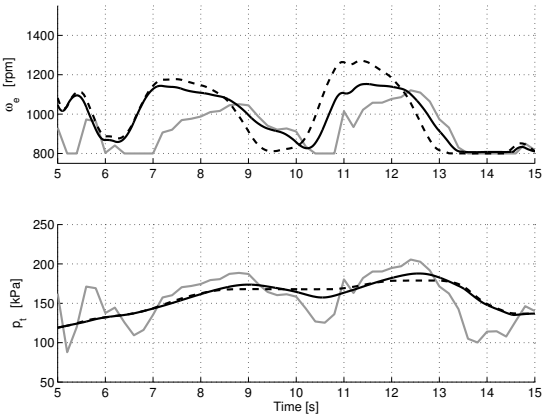
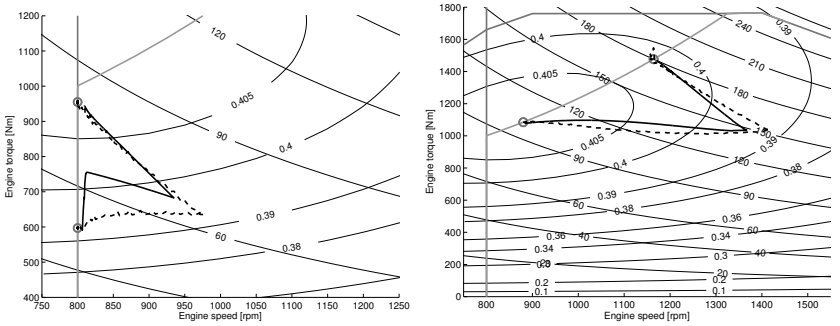


Figure 4.18: *Engine speed and turbo pressure. Gray is static optimum ($\omega_{e,\Sigma}(P_{load}(t))$,$p_{t,\Sigma}(P_{load}(t))$), continuous is suboptimal and dashed is optimal.*



(a) (b)

Figure 4.19: *Suboptimal (continuous) and optimal (dashed) trajectories for the TC-engine in the low (4.19(a)) and high (4.19(b)) step load cases.*

4.9 Discussions and comments

4.9.1 Dynamic programming

The dynamic programming optimization in this chapter is fairly straight-forward. The result for the naturally aspirated engine is a bit unexpected though; before output power steps it is optimal to accelerate or decelerate past the upcoming static optimal engine speed, and approach the new static optimum from the 'wrong' direction after the step. The motion of the engine operating point is counter clockwise in all cases studied, so that it travels toward higher engine speeds below the static optimal line and toward lower speeds above this line. This differs from the result presented in Pfiffner (2001), in which the operating point move in a clockwise direction. The clockwise motion however seems to be caused by a bad choice of initial and terminal states. In this chapter the engine is forced to start and finish at the static optimal points corresponding to the initial and terminal output powers, and given sufficient time to move between these so that the trajectories would not change if more time were added to the beginning or the end of the load cases. The primary problem with DP, which is encountered in both engine setups but especially for the turbocharged engine, is the high calculatory effort. The most obvious way of countering this is to reduce the discretization grid densities, though care has to be taken to avoid large simulation errors and faulty infinite-cost spread (as mentioned in Section 3.2).

4.9.2 PMP based methods

The phase planes in Section 4.7.1 is used to validate the results derived with dynamic programming and to provide insight into the the mechanisms behind the trajectories. This insight is enhanced by the actual derivation of optimal trajectories in Section 4.7.2, and the expansion in Section 4.7.3 which show that the reasoning is valid also for the TC-engine. The actual solving of the dynamic optimization problems in this section is also fast, compared to dynamic programming. The treatment therefore provide an excellent pedagogic example of optimization with Pontryagin's maximum principle. The methods are however highly restrictive in the load cases which can be treated. The PMP formulation used does not include the state and control constraints (4.5) and the methods are only practically usable for output power steps or, for the NA-engine, slightly more complicated cases.

4.9.3 Suboptimal methods

The developed methods for finding suboptimal solutions works well for both of the engine setups. In both cases the time for finding a solution is reduced by a factor > 3000, while the amount of fuel required only increase by $< 0.1\%$ for the NA-engine and $< 5\%$ for the TC-engine. It should be noted that in both cases, and in particular for the TC-engine, finding even a feasible solution

is not a trivial problem. The developed methods does not require analytic expressions neither for the engine efficiency nor for the static optimal line. The only requirements for the NA-engine are that $T_{e,\Sigma}(\omega_e)$ is well defined for all ω_e and that $T_{e,\Sigma}(\omega_e) \cdot \omega_e$ is strictly increasing with increasing ω_e, so that the rule (4.23) makes the system unstable.

5

Multi-Mode CVT Drivetrain

Continuously variable transmissions (CVTs) give a large freedom for control of the engine, but involve two power conversions, each of which produce losses. In power-split constructions, such as those in Carl et al. (2006) and Gramattico et al. (2010), this is adressed by transmitting part of the power mechanically. The most famous power-split device is probably the Toyota Hybrid System of the Toyota Prius, which is described in Sasaki (1998). The efficiency of such devices do in general vary with the gear ratio. If high efficiency at widely separated gear ratios is desired, the power-split CVT can be expanded. Multi-mode CVTs are constructed so that several power-split layouts can be realized with the same device, in general by applying and releasing clutches. The continuously variable component in the device can be of any type, for example belt, electric or hydraulic. In this chapter a hydraulic multi-mode CVT is evaluated by comparing the fuel saving potential of this concept with that of the present setup under deterministic and stochastic loads. The chapter presents the transmission models, continue with a discussion on the application of dynamic programming to the two setups and finally present and discuss the results.

5.1 System setup

The system consist of an engine which on one side is connected to the working hydraulics pump and on the other side to the transmission. Two different types of transmissions are studied; the reference torque converter/automatic gearbox transmission and a multi-mode CVT transmission. The engine, the transmission concepts and the hydraulic pump are described in the following sections.

Engine

The engine used is a turbocharged diesel engine, identical to that presented in Sections 4.1 and 4.5, though in Equation (4.1) the load torque P_{load}/ω_e is replaced by the transmission load T_T and the hydraulic load T_H, according to Equation (5.1). The reader is referred to Sections 4.1 and 4.5 for details on the engine model. The relation between the loads T_T and T_H, and the load cycle depend both on the transmission setup and on ω_e.

$$\frac{d\omega_e(t)}{dt} \cdot I_e = T_e(t) - T_T(t) - T_H(t) \tag{5.1}$$

Reference transmission

The present transmission consist of a torque converter, connected to a four speed $r_c \in \{1, 2, 3, 4\}$ forward/reverse automatic gearbox, which connects to the drive shaft. The main source of losses in this transmission is the torque converter, which is modeled according to Equation (5.2). Denote the engine side connection with index *cp* and the gearbox side connection with index *ct*. The input and output torques of the torque converter depend only on the angular speeds at the input and output of the device. The torques are calculated from the maps M_P and μ, which have been measured at the reference speed $\omega_{c,ref}$.

$$\nu_c = \frac{\omega_{ct}}{\omega_{cp}} \tag{5.2a}$$

$$T_{cp} = M_P(\nu_c) \left(\frac{\omega_{cp}}{\omega_{c,ref}} \right)^2 \tag{5.2b}$$

$$T_{ct} = \mu(\nu_c) T_{cp} \tag{5.2c}$$

The torque converter is connected to the engine so that $\omega_{cp} = \omega_e$ and $T_{cp} = T_T$, while the relation between ω_{ct} and T_{ct} and the load cycle is a fix ratio which depend on the engaged gear. This layout and model is identical to that used in Chapter 2 to derive the load cases. The transmission layout is displayed in Figure 5.1, in which the gray box represent the automatic gearbox.

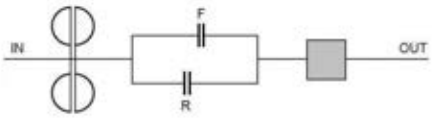


Figure 5.1: *Reference transmission scheme.*

MM-CVT transmission

Multi-mode constructions combine several power split setups, thus allowing for high efficiency at several widely separated gear ratio regions. This thesis studies

the three mode ($M_{cvt} \in \{1, 2, 3\}$) hydrostatic CVT, which is described in the patent Mattson and Åkerblom (2012) and has a structure similar to that of devices used in Savaresi et al. (2004), Tanelli et al. (2007) and Lauinger et al. (2007). The CVT functionality is provided by two hydraulic machines which together is referred to as the variator. Changing gear ratio within a mode is done by altering the displacement of the hydraulic machines and switching between the modes is done by applying and releasing clutches. The proposed layout is presented in Figure 5.2, in which the dark gray box represents a Ravigneaux planetary gearset and the light gray box represents a regular planetary gearset. Regular gears are omitted in the figure. The clutches 1, 2 & 3 are used to select the mode by applying the one that corresponds to the desired mode M_{cvt}. The mechanical structures of the three possible modes are presented in Figure 5.3.

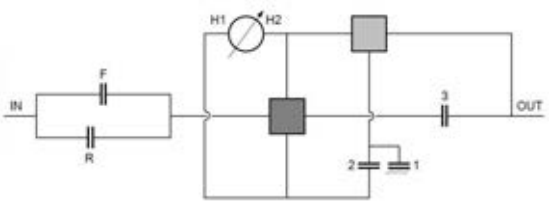


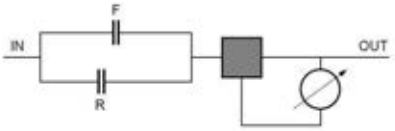
Figure 5.2: *Multi-mode CVT scheme.*

It is assumed that all axles and cogwheels in the CVT are stiff and massless. The planetary gearset can then be described by a kinematic condition, a power balance and a torque balance. Denote the number of cogs by z and use the indices r for the ring wheel, c for the planetary carrier and s for the sun wheel. The gearset can the be described by Equation (5.3).

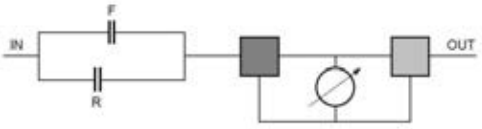
$$z_s\omega_s + z_r\omega_r - (z_s + z_r)\omega_c = 0 \tag{5.3a}$$
$$T_s\omega_s + T_r\omega_r + T_c\omega_c = 0 \tag{5.3b}$$
$$T_s + T_r + T_c = 0 \tag{5.3c}$$

A Ravigneaux gearset is a planetary gearset which has been expanded with an additional sun wheel. The expressions that describe this device is therefore similar to those for the regular planetary gearset, but include one more kinematic condition. By denoting the number of cogs by z and use the indices r for the ring wheel, c for the planetary carrier, sl for the larger sun wheel and ss for the smaller sun wheel, this can be expressed according to Equation (5.4).

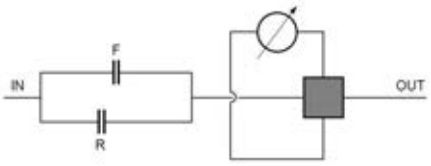
$$z_{sl}\omega_{sl} + z_r\omega_r - (z_{sl} + z_r)\omega_c = 0 \tag{5.4a}$$
$$z_{ss}\omega_{ss} - z_r\omega_r - (z_{ss} - z_r)\omega_c = 0 \tag{5.4b}$$
$$T_{sl}\omega_{sl} + T_{ss}\omega_{ss} + T_r\omega_r + T_c\omega_c = 0 \tag{5.4c}$$
$$T_{sl} + T_{ss} + T_r + T_c = 0 \tag{5.4d}$$



(a) 1st mode



(b) 2nd mode



(c) 3rd mode

Figure 5.3: *Layouts of each of the three modes of the MM-CVT.*

The ideal (no-load) relation between variator displacement and overall gear ratio for the three modes is displayed in Figure 5.4. The overall gear ratio has been normalized with the maximum gear ratio. Mode shifts are performed at

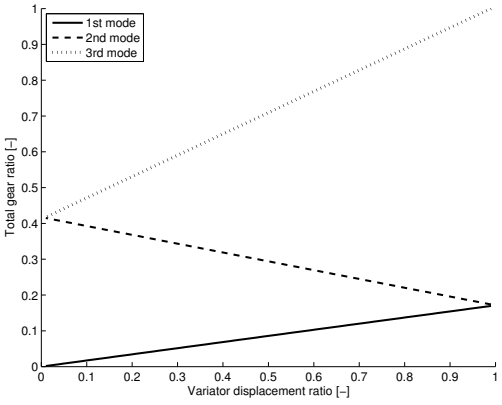


Figure 5.4: *No-load relation between variator displacement ratio and normalized overall gear ratio for the three modes of the MM-CVT. Continuous is first mode, dashed is second mode and dotted is third mode.*

the extremals of the variator displacement ratio, and mode shifts at these points does not change the overall gear ratio for a lossless transmission. At mode shifts the speed differences over the involved clutches are ideally zero, and therefore there are no losses in the clutches, which simplifies the models. The main source of losses in this concept is the variator, which therefore is described in more detail. Let the index $n = 1, 2$ denote the two hydraulic machines and use the notation D_v for the maximum displacement, $\psi_{v,n} \in (0, 1)$ for the relative displacement, ω_n for the axle speed and T_n for the axle torque of the machines. Also introduce p_v for the variator hydraulic pressure and $\eta_{v,1}$,$\eta_{v,2}$,$\eta_{v,3}$,$\eta_{v,4}$ as efficiency parameters. The model for this device, which is similar to and based on models used in Lennevi (1995), can then be expressed as

$$\psi_{v,1} D_v \omega_1 \pm p_v(2\eta_{v1} + (|\omega_1| + |\omega_2|)\eta_{v,2}) - \psi_{v,2} D_v \omega_2 = \tau_{p_v} \dot{p}_v \tag{5.5}$$

$$\psi_{v,n} D_v p_v - T_n \pm (\eta_{v3}|\omega_n| + \eta_{v4} p_v) = 0 \tag{5.6}$$

The signs in the equations depend on the power flow direction. Equation (5.5) describes hydraulic fluid flow and Equation (5.6) describes torque at each machine. The variator is constructed so that $\psi_1 + \psi_2 = 1$. The variator pressure dynamics is assumed to be fast compared to the engine dynamics, that is; it is assumed that the time constant $\tau_{p_v} \approx 0$. The pump which powers the CVT actuators are modeled as a constant torque loss at the transmission input.

The combination of the equations above into relations between the transmission inputs (ω_e,T_T) and outputs (ω_w,T_w) depend on driving direction, power flow direction and the engaged mode, as presented in Figure 5.3.

Working hydraulics

The bucket and boom are hydraulically driven. Pressure p_H and flow Q_H of hydraulic fluid is supplied by a pump which is connected to the engine axle, so that the pump speed is the same as the engine speed. The pump has a variable displacement, so that the pressure and flow are not dependent on the engine speed, apart from the efficiency and that the maximum displacement give a time dependent lower limit for the speed. It is assumed that the pressure dynamics are quick compared to other dynamics of the system, and that there are no flow-losses. Introduce the maximum displacement D_H, the relative displacement $\psi_H \in [0, 1]$ and the efficiency $\eta_H(p_H, \psi_H)$ of the hydraulic pump. The torque T_H for driving this pump can then be calculated from Equations (5.7) and (5.8).

$$Q_H = \psi_H D_H \omega_e \tag{5.7}$$

$$Q_H p_H = \eta_H T_H \omega_e \tag{5.8}$$

The pump efficiency $\eta_H(p_H, \psi_H)$ is modeled according to

$$\eta_H = \eta_{H1}\psi_H + \eta_{H2} p_H \psi_H + \eta_{H3}\sqrt{p_H \psi_H} \tag{5.9}$$

5.2 Problem statement

The problem studied in this chapter is the minimization of the total amount of fuel used, according to Equation (4.10), while fulfilling the engine dynamics Equation (5.1), the turbo dynamics Equation (4.7), the constraints (4.5) and the relevant relations between load cycle and engine load (T_T & T_H). This also means that the load trajectories $\omega_w(t)$, $T_w(t)$, $Q_H(t)$ and $p_H(t)$ will be exactly followed. Optimization is performed with both deterministic and stochastic loads. In the stochastic cycles a deterministic load is applied, but optimization is performed against an uncertain prediction of future loads.

5.3 Load cases

The load cases are specified by the wheel speed ω_w, wheel torque T_w, hydraulic flow Q_H and hydraulic pressure p_H. The derivation of the load cases from measurement data is described in Chapter 2. The evaluation is based on the stochastic load case 'SDP *mc*', but since one aim is to evaluate the method itself, the corresponding deterministic case 'DDP *mc*' is also used. The averaging in the creation of these cycles essentially low pass filter the signals, so one individual cycle from the set is also used; the 'DDP *sc*'. Along with these the longer cycle 'DDP *lc*' is included in the set of applied load cycles.

5.4 Application of dynamic programming

With the driving cycle description presented in Section 2.2, the dynamics of the vehicle is in the engine speed and the turbo pressure, regardless of transmission. These dynamics are described by Equations (4.1) and (4.7). The controls available, regarless of transmission, is the injected fuel and a possible brake torque. Additional control signals are choice of gear for the reference vehicle and variator displacement ratio and choice of mode for the MM-CVT. When formulating these models in the DP framework a few non-trivial choices emerge, particularly on which states and controls to actually use when applying a stochastic load. The choices made in the two setups are discussed here. An analogue analysis is made for a naturally aspirated engine in Nilsson et al. (2012b).

Application to the reference vehicle

The following argument is written for a single gear but is still valid for this vehicle since the number of gears is finite and frequent changes are not desired.

Equation (5.2) shows that the torque from the torque converter only depend on the input and output speeds. A rapid change of output torque therefore require an equally rapid change of engine speed. The rate of engine speed change is limited by the engine inertia, as described by Equation (5.1). Therefore the inertia also limit the rate of output torque change $\frac{dT_{ct}}{dt}$. Faster output torque

changes requires preparatory increase of the engine speed, while the produced excess torque must be balanced by the brakes. Fast torque reductions must in the same way be smoothened by the brakes. In the deterministic load cases this is seldom a problem, especially since the load cases has been derived from measurements performed with a torque converter vehicle. In the stochastic load case the probabilities are independent, as stated in Equation (2.2). This can be interpreted as a perpetual possibility of steps in the stochastic load components, while failure to fulfill any of the applied load combinations yield an infinite cost. The engine speed must therefore always be high enough to manage the most demanding possible future load, thus in general be higher than made necessary by the actual load. The transmission output torque will therefore in general also be higher than the load torque, and the brakes will have to be applied almost continuously, producing unrealistic losses in both torque converter and brakes.

An alternative is to assume that the engine can change speed instantly by assuming zero inertia, $I_e = 0$, in Equation (5.1). The engine speed will then be implicitly given by the load and the selected gear, and will not be a state of the system. This choice can be expected to cause an underestimation of the fuel requirement for this vehicle, and thereby also of the increase in potential when changing to an MM-CVT transmission. This underestimation is considered less severe than a possible overestimation. The $I_e = 0$ assumption will inevitably lead to non-physical engine speed changes. Since this model is primarily intended as a fuel consumption reference for the MM-CVT vehicle, it was decided that this is to be preferred over constant use of the brakes.

Since there is no engine inertia and this vehicle may not loose speed nor lack thrust for any time, instant gear changes are assumed and the clutch losses are disregarded by not associating any fuel use to these, which will cause an underestimation of the fuel needed for this vehicle. A small (1% of $\max m_f$) cost for changing gear is added to J though, to prevent excessive gear changes.

The states for the reference vehicle are therefore gear r_c and turbo pressure p_t and the controls are injected fuel m_f, brake torque T_b and gear change Δr_c.

Table 5.1: *Reference vehicle states and controls.*

	DDP	SDP
States X	r_c,p_t	r_c,p_t
Controls U	m_f,T_b,Δr_c	m_f,T_b,Δr_c

Application to the MM-CVT vehicle

Since the driving cycle specify the vehicle speed, an engine speed trajectory is equivalent to a gear ratio trajectory. The relation presented in Figure 5.4 can therefore be interpreted as a relation between CVT mode, variator displacement ratio and engine speed; $\omega_e(\psi_{v,1}, M_{cvt})$. Therefore all of these cannot be states, or the system will be overdetermined. The function $\omega_e(\psi_{v,1})$ is invertible within the allowed operating region, so that within a CVT-mode the choice of ω_e or

$\psi_{v,1}$ as state is equivalent. The possibility of restrictions on $\frac{d\psi_{v,1}}{dt}$, especially during mode-shifts, points to using $\psi_{v,1}$ as state. Since one of the hydraulic machines speeds up when $\psi_{v,1}$ gets close to 0 or 1, the losses increase in these regions. Because of this it is desirable to have higher state grid density near the extremes of $\psi_{v,1}$, which also points to using $\psi_{v,1}$ as state. The dynamics however are described in terms of ω_e, so using $\psi_{v,1}$ as state imply the following scheme for the simulations in the calculation of the cost-to-go map $J(X,t)$:

$$\psi_{v,1}(t_k) \xrightarrow{W_k} \omega_e(t_k) \xrightarrow{\frac{d\omega_e}{dt}} \omega_e(t_{k+1}) \xrightarrow{W_\kappa} \psi_{v,1}(t_{k+1})$$

In the first and last steps the load is required, since $\omega_e(\psi_{v,1})$ depend on the load. At the last step a choice has to be made whether to use $\kappa = k$ or $\kappa = k+1$. Using $\kappa = k$ does not guarantee continuity in ω_e and makes it possible for the optimizer to draw a net power from the engine inertia. This is illustrated in Figure 5.5 by the solution to the 'SDP *sc*' load case, as calculated with $\psi_{v,1}, M_{cvt}$ and ω_e as states. The figure shows $\omega_e(t_k)$ and $\omega_e(t_{k+1})$ trajectories, which should coincide for the engine speed to be continuous. $\kappa = k+1$ on the other hand guarantees continuous ω_e, but causes a quadratic increase in the number of possible load combinations in the stochastic case. Even though a small load variation space (3^3 combinations) is used, the quadratic increase in combinations would cause an unacceptable increase in calculation effort. This means that for SDP it is not practical to use $\psi_{v,1}$ as a state.

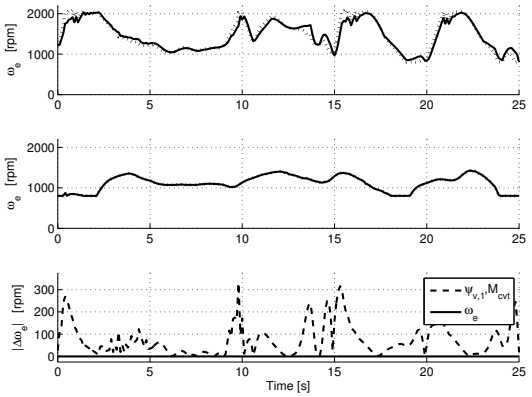


Figure 5.5: *The top two figures show the engine speeds $\omega_{e,k}$ (continuous) and $\omega_{e,k+1}$ (dotted) for $\psi_{v,1}, M_{cvt}$ (top) and ω_e (middle) as state. The bottom figure shows the speed errors for the two alternatives.*

The choice of $\psi_{v,1}$ as a state would require that M_{cvt} is also used as at state. As Figure 5.4 indicate, the function $\omega_e(\psi_{v,1}, M_{cvt})$ is invertible except exactly at the mode switching points. When load is applied, the regions at which $M_{cvt}(\omega_e)$ is not well defined may no longer be infinitesimal. M_{cvt} should therefore be a

state also if ω_e is used as a state. Still, since the regions at which $M_{cvt}(\omega_e)$ may not be well defined are small, the state M_{cvt} is skipped and in ambiguous cases the M_{cvt} which give highest efficiency is used. In Section 5.1 it is assumed that the speed differences over the clutches are zero at mode changes. Mode changes are therefore only allowed at the extremes of $\psi_{v,1}$, but since mode changes are not explicitly controlled, this is has to be verified in the results.

In DDP the choice is made to use $\psi_{v,1}, M_{cvt}$ as states, while in SDP ω_e is used as state. The turbo pressure p_t is a state in both DDP and SDP. The fuel injection m_f is a control signal in both DDP and SDP, and in DDP the change of mode ΔM_{cvt} is also used as a control signal.

Table 5.2: *MM-CVT vehicle states and controls.*

	DDP	SDP
States X	$\psi_{v,1}$,M_{cvt},p_t	ω_e,p_t
Controls U	m_f,ΔM_{cvt}	m_f

5.5 Optimization results

5.5.1 General results

The reason for introducing the multi-mode CVT transmission is to increase the potential for fuel consumption reduction. The minimum fuel requirement for the two transmission concepts in the four driving cycles 'DDP *mc*', 'SDP *mc*', 'DDP *sc*' and 'DDP *lc*' is presented in Table 5.3. The fuel saving potential is consistently around 17% higher for the MM-CVT transmission. The saving in the 'DDP *mc*' loadcase is roughly the same as that in the 'SDP *mc*' load case. This indicates that the MM-CVT transmission is not more sensitive than the present transmission to roughnesses nor to prediction uncertainties.

Table 5.3: *Reference and MM-CVT vehicle fuel usage.*

	Reference [ml]	MM-CVT [ml]	Saving [%]
DDP *mc*	203	170	16.3
SDP *mc*	210	174	17.1
DDP *sc*	225	187	16.9
DDP *lc*	932	772	17.2

5.5.2 Reference vehicle 'SDP *mc*' results

Figures 5.6 and 5.7 presents the state and control trajectories for the reference vehicle. The states are gear and turbo pressure, with the set pressure indicated by the dotted line. Since the speed is low in this cycle, the vehicle is at higher than second gear only briefly at $17s$ into the cycle. Despite the cost associated to gear changes, these are still frequent. These could have been suppressed

further by increasing this cost, but this artificial cost would then risk obscuring the true loss mechanisms. The controls are injected fuel and brake torque. The brake is applied primarily at around $14 - 15s$ but also briefly at $20s$. At both these instances there is an infinite cost associated to low turbo pressure. The brake is applied since the additional torque translates to increased engine speed, which also result in higher intake set pressure, thus avoiding the infinite cost.

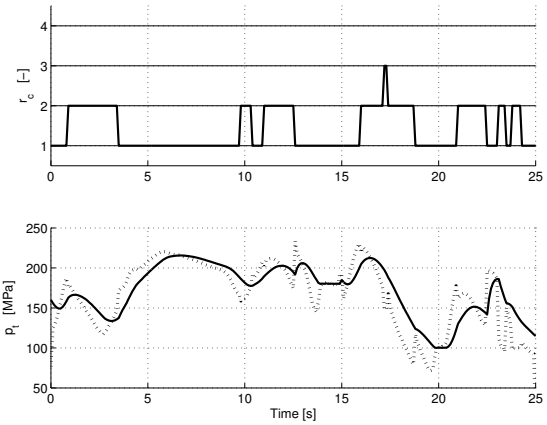


Figure 5.6: *Reference vehicle state trajectories for the 'SDP mc' load case.*

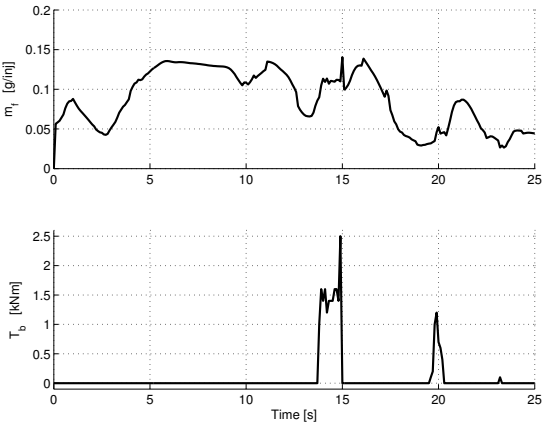


Figure 5.7: *Reference vehicle control trajectories for the 'SDP mc' load case.*

Figure 5.8 presents the engine speed, torque and turbo pressure for the reference vehicle. The dotted lines in the engine speed figure indicate the minimum speeds which are given by the maximum pump displacement and the three alternative hydraulics flows in the cost-to-go map calculation. The alternative in the middle, which is dashed, is that which is actually applied in the subsequent

optimal path simulation. At general the engine speed is above the highest alternative, but at around 23s it briefly goes below this limit. In the calculation of the cost-to-go map, this would have produced an infinite cost, but since the engine speed is not a state this is not explicitly linked to an infinite cost in the later forward simulation. Comparing Figure 5.8 to Figure 5.6 shows that at around 13s and 23s the hydraulics flow causes shifts to a lower gear. At both of these instances the engine speed peaks at just above 2000rpm.

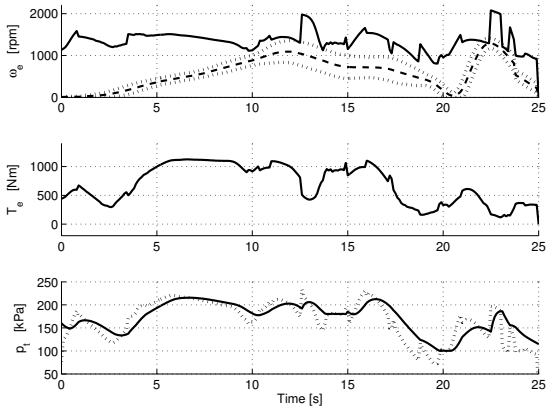


Figure 5.8: *Reference vehicle engine trajectories for the 'SDP mc' load case.*

Figure 5.9 presents the speed, torque and power at the input and output of the torque converter of the reference vehicle. The continuous lines show input (engine) side and the dotted show output (gearbox) side. As expected, the losses are the highest at the bucket filling between 4s and 9s, during which time the vehicle speed is low and the required tractive effort is high.

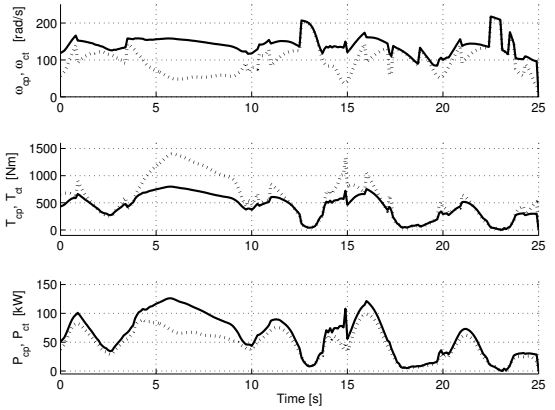


Figure 5.9: *Reference vehicle transmission trajectories for the 'SDP mc case.*

5.5.3 MM-CVT vehicle 'SDP *mc*' results

Figures 5.10 and 5.11 shows the state and control trajectories for the MM-CVT vehicle. The states are engine speed and turbo pressure. The turbo pressure set point is indicated with a dotted line. This set pressure is smoother for the MM-CVT than for the reference vehicle, as shown in Figure 5.6. The only control signal is the injected fuel, though Figure 5.11 also shows the engine torque.

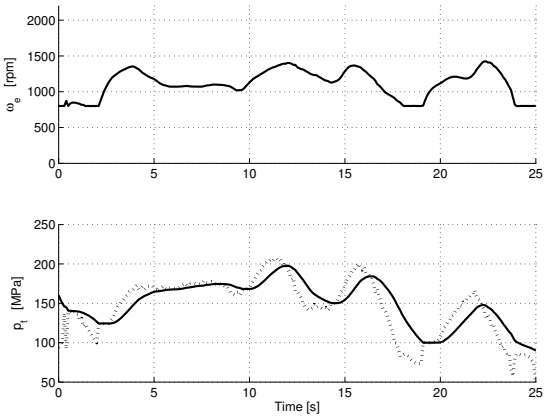


Figure 5.10: *MM-CVT vehicle state trajectories for the 'SDP mc' load case.*

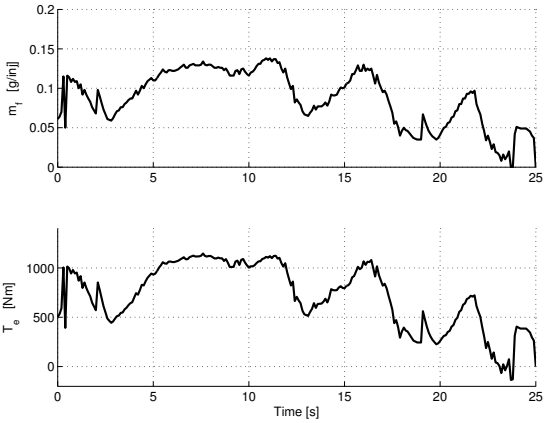


Figure 5.11: *MM-CVT vehicle control trajectories for the 'SDP mc' load case.*

Figure 5.12 shows engine speed, torque and turbo pressure for the MM-CVT vehicle. Along with the engine speed, the minimum speeds as given by the hydraulic pump displacement and alternative flows, are indicated by the dotted lines. The dashed alternative is that which is applied in the optimal path simulation. Unlike for the reference vehicle, the engine speed is never lower than

the highest alternative since the engine speed is a state, and therefore there can be an infinite cost directly associated to speeds below this limit. The engine speed is below $1500rpm$ for the duration of the cycle.

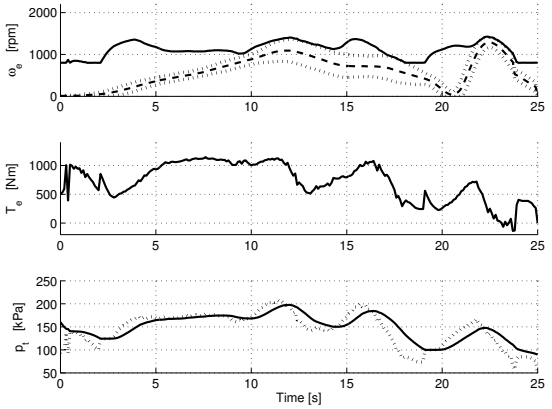


Figure 5.12: *MM-CVT vehicle engine trajectories for the 'SDP mc' load case.*

Figure 5.13 presents the variator displacement ratio, CVT-mode and speed differences in the mode-clutches. Most of the time the vehicle operate at the lowest mode, with a few excursions to the second mode. The variator displacement trajectory is continuous and at its extreme at mode changes, as required for using the engine speed as state. At the mode changes the speed differences in the involved clutches are close to zero, as assumed in the loss model.

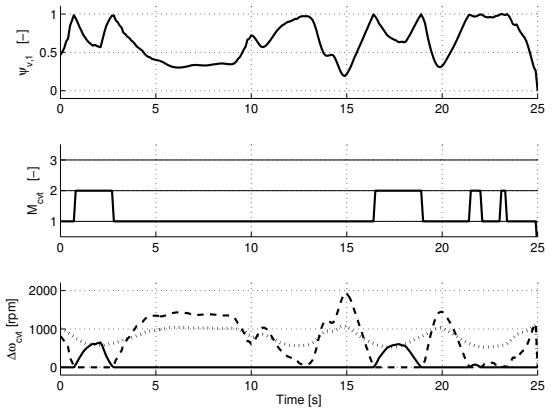


Figure 5.13: *MM-CVT vehicle CVT trajectories for the 'SDP mc' load case.*

Figure 5.14 presents the speed, torque and power at the input and output of the CVT for the MM-CVT vehicle. The continuous lines show input (engine) side and the dotted show output (driveshaft) side. Comparing this figure to

Figure 5.9 shows that the transmission efficiency is higher for the MM-CVT vehicle, especially during the bucket filling at $4s$ to $9s$.

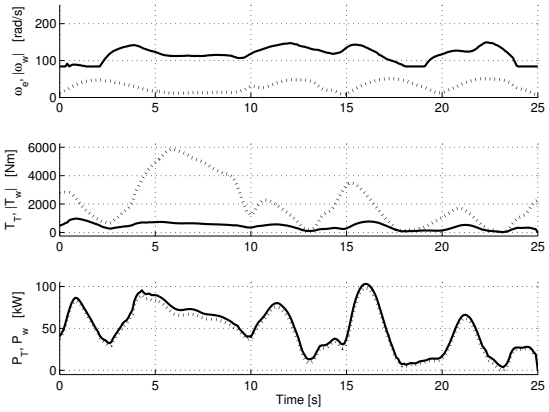


Figure 5.14: *MM-CVT vehicle CVT trajectories for the 'SDP mc' load case.*

5.5.4 Reference/MM-CVT comparison

The engine speed, torque and turbo pressure for the reference and MM-CVT vehicles in the 'SDP *mc*' load case are compared in Figure 5.15. Though the torques are similar for most of the time, the engine speed and turbo pressure are in general higher for the reference vehicle. Figure 5.16 presents the efficiencies and losses for the two concepts, divided into engine, transmission and hydraulics pump. The efficiencies can be increased and the losses reduced in all three presented components by changing from the present transmission to the proposed MM-CVT alternative. First, the CVT provide an increased freedom in the choice of engine operating point. This means that the mean engine efficiency can be increased. Since the engine used has a quite flat efficiency map, the mean efficiency increase is quite small (from 36.4% to 37.7%), but due to the high average output power this still correspond to more than $30kW$ loss reduction (from $173kW$ to $140kW$). Second, the removal of the torque converter reduces the average transmission losses by more than three quarters (from $16.0kW$ to $3.95kW$), and increase the efficiency from 72.1% to 92.1%. About one fifteenth of the reference vehicle transmission losses are caused by braking and the rest is losses in the torque converter. It should be noted that more than one third of the MM-CVT transmission losses is caused by the transmission actuator pump, and stressed once again that the reference vehicle do not have any clutch losses. The third effect is that more efficient operating points, which is at high displacement and low engine speed, can be selected for the working hydraulics pump. The average pump efficiency is increased from 70.2% to 75.0%, reducing the average loss from $11.1kW$ to $8.65kW$. Similar values are experienced for all four load cases, except for the average hydraulics

pump loss which is significantly smaller in the 'DDP *lc*' load case. In this case the longer transports reduce the proportion of the cycle in which the hydraulics are used, and thereby also the average hydraulic output and loss powers.

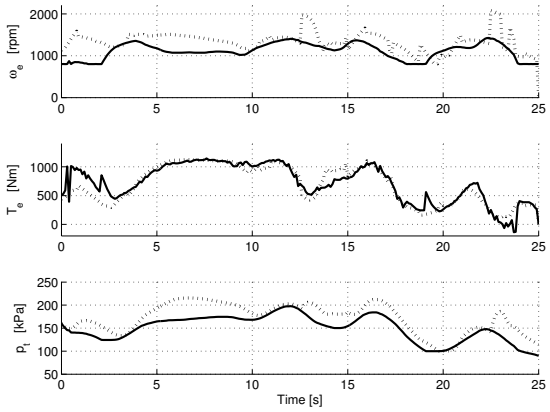


Figure 5.15: *Reference/MM-CVT vehicle engine operation comparison for the 'SDP mc' load case. Dotted is Reference vehicle and continuous is MM-CVT.*

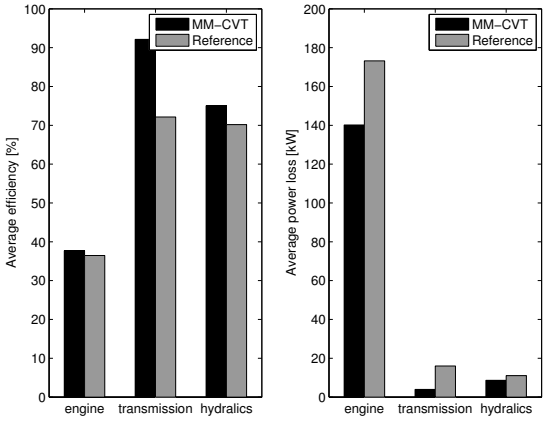


Figure 5.16: *Reference/MM-CVT vehicle efficiency and loss comparison*

5.5.5 DDP/SDP comparison

Figure 5.17 shows the engine speed, torque and turbo pressure, and Figure 5.18 shows the engaged gear and torque converter power loss, for the reference vehilce in the stochastic 'SDP *mc*' (continuous) and deterministic 'DDP *mc*' (dotted) load cases. The gear used in the DDP cycle is always equal to or higher than in the SDP cycle. The result is lower engine speed, which is most clearly visible

in that the peaks at 13s and 23s do not occur for the deterministic solution. Because of this, the turbo pressure immediately after these instances is lower in the DDP solution. The higher engine speed at the regions at 15s and 20s in the SDP solution also cause higher transmission losses at these instances. Much of the additional SDP losses are caused by braking, since in the DDP case the vehicle only make a few minor brake pulses in the region 16s to 24s.

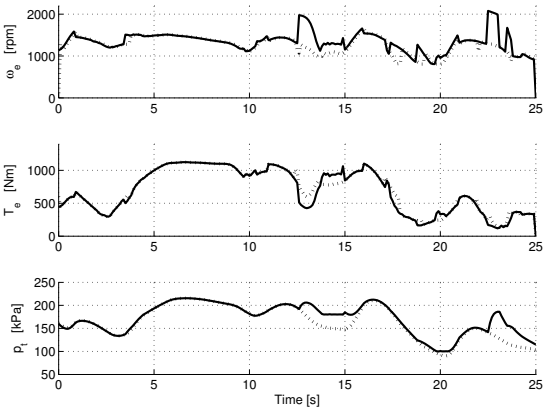


Figure 5.17: *Reference vehicle engine operation comparison for the 'DDP mc' and 'SDP mc' load cases. Continuous is SDP and dotted is DDP.*

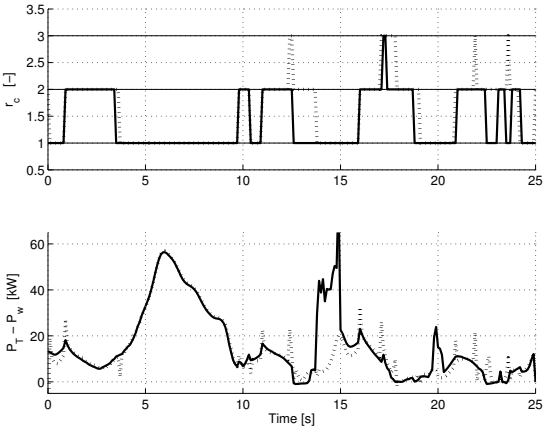


Figure 5.18: *Reference vehicle transmission operation comparison for the 'DDP mc' and 'SDP mc' load cases. Continuous is SDP and dotted is DDP.*

The engine speed, torque and turbo pressure for the MM-CVT vehicle in the stochastic 'SDP *mc*' and deterministic 'DDP *mc*' load cases are compared in Figure 5.19. In Figure 5.20 the variator displacement ratio and CVT-mode

for the same vehicle and cycles are compared. The dotted lines are the deterministic and the continuous are the stochastic solution. In the deterministic case the engine speed is in general somewhat lower. The engine torque varies more rapidly in the deterministic case. This is due to the several possible load alternatives in the SDP load case making the cost-to-go map for this case less rugged. The most visible difference is however that in the deterministic case the vehicle change to second CVT-mode later, as seen around $2s$ and $22s$, but spend more time at this mode, at around $23s$, and make one additional second mode excursion at around $13s$.

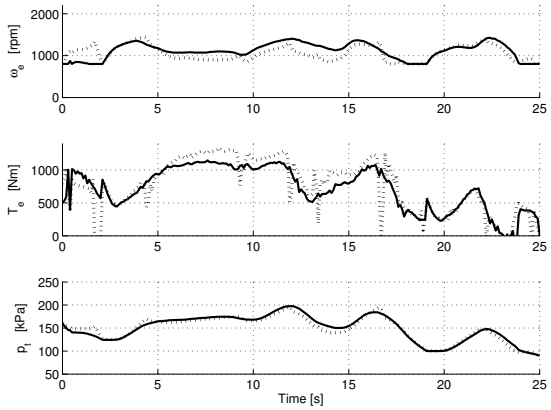


Figure 5.19: *Comparison of the MM-CVT vehicle engine operation in the 'DDP mc' and 'SDP mc' load cases. Continuous is SDP and dotted is DDP.*

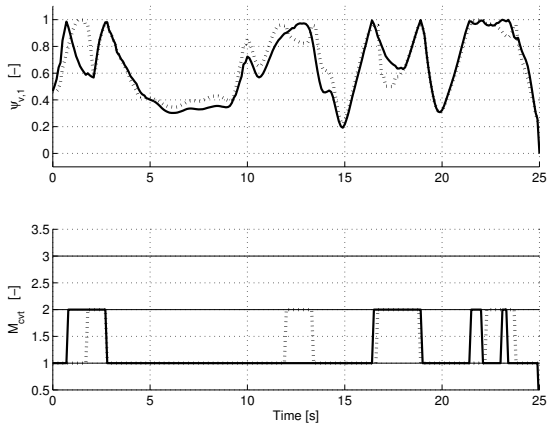


Figure 5.20: *Comparison of the MM-CVT vehicle CVT operation in the 'DDP mc' and 'SDP mc' load cases. Continuous is SDP and dotted is DDP.*

5.6 Discussion and comments

5.6.1 Optimization

The optimization in Chapter 5 is more intricate than in Chapter 4, especially in the choice of states. Intuitively the engine speed and turbo pressure should be used as a states for both vehicles. For the reference vehicle the engine speed can also be calculated from the output torque, making the system overdetermined. Since an underestimaton of the reference vehicle fuel consumption, and thereby of the potential increase of the MM-CVT vehicle, is considered less severe than an overestimation, this is solved by assuming zero engine inertia and not using the engine speed as state. For the MM-CVT vehicle the variator displacement ratio is for several reasons preferred as state instead of the engine speed. In the stochastic case this could only be realized to an extreme increase in calculatory effort, and therefore this ratio is used as state only in the deterministic case.

The results of the deterministic and stochastic loads differ somewhat for both transmissions. The most pronounced differences are that the engine speed and turbo pressure are in general higher for the stochastic load. The differences follow expectations, which supports the applied stochastic load formulation. It was decided that SDP is a method that suits the transmission evaluation of Chapter 5 since it combines the benefits of DDP; simplicity and guaranteed global optimality, with easily controlled disturbances.

5.6.2 MM-CVT potential increase

Since dynamic optimization has been used on both transmission concepts and effort has been made not to overestimate the fuel consumption of the reference vehicle, it can with a high degree of certainty be stated that the increase in fuel saving potential by changing the transmission is no less than 15% (Table 5.3).

The concept provide an increase in engine efficiency from 36.4% to 37.7%. The mean value of the maximum engine efficiency of the instantaneous output power is 38.5% while the maximum efficiency of the engine at the mean ouput power is just above 40.5%. This shows that there is still an unexploited potential for fuel consumption reduction. The concept does however reduce the transmission losses by more than three quarters, and allows for conciderable engine and hydraulics pump efficiency improvements. The efficiency of the system increase from 27% to 32% and the fuel requirement is reduced by more than 15%. The performed analysis also indicate that the potential of the MM-CVT concept is not more sensitive to prediction uncertainties, cycle smoothness or cycle length than the present transmission. Introducing the proposed MM-CVT may lead to savings far greater than 15%, due to the prioritization of driveability over fuel efficiency in the present transmission, this has not been exploited though since that would make an unfair comparison. On the other hand, the MM-CVT concept requires active control, contrary to the present transmission. Any further analysis should therefore include an evaluation of different controller concepts.

6

Conclusions

6.1 Methods used and developed

Dynamic programming

Dynamic programming is used in both Chapter 4 and Chapter 5. This method and the resulting solutions has several important benefits, but the calculatory effort is high, and increase fast with the number of states and control signals. In Chapter 4 dynamic programming is primarily used as a baseline method for evaluation of the other trajectory derivation methods developed. In Chapter 5 both stochastic and deterministic dynamic programming is used in the evaluation of a transmission concept. The introduction of a stochastic load is motivated by a need for evaluating the sensitivity to prediction uncertainties, but the stochastic load formulation also highly affect the solving of the dynamic optimization problem in that restrictions is imposed on which states of the system that can be used as states in the optimization.

PMP based methods

Section 4.7 presents methods based on Pontryagin's maximum principle (PMP) for finding optimal state and control trajectories. These methods are fast compared to dynamic programming, but difficult to apply in practice due to the severe restrictions on the load cases that can be treated. The methods do on the other hand provide an excellent pedagogic example of dynamic optimization with Pontryagin's maximum principle, which for the naturally aspirated engine includes visualization by phase planes.

Suboptimal methods

Due to the high calculatory effort of dynamic programming and the load case restrictions of the developed PMP based methods, alternative methods are desired. In Section 4.8 such suboptimal methods for both the naturally aspirated and the turbocharged engines are presented. These are close to optimality while being extremely fast and not restricting on the load case applied. The methods are still non-causal and require a load prediction, but follows this load exactly.

6.2 MM-CVT transmission

The evaluation presented in Chapter 5 shows that the MM-CVT concept has at least 15% lower minimum fuel consumption than the torque converter/automatic gearbox reference drivetrain. That is, a vehicle with the MM-CVT transmission requires 15% less fuel than a vehicle with the reference transmission for performing exactly the same driving mission, as specified in Section 2.2. The evaluation also show that this potential is not sensitive to prediction uncertainties, cycle smoothness or cycle length. The efficiency improvements are made by better choice of operating points for the engine and hydraulics pump and higher transmission efficiency due to the ellimination of the torque converter. The transmission losses are reduced by more than three quarters, though the largest power-loss reduction is caused by more efficient engine operation.

The fuel saving potential of the MM-CVT is greater than the presented 15% since the reference transmission is not controlled according to the optimal as presented in Chapter 5, due to lack of prediction and prioritization of driveability over fuel consumption. The flexibility of the torque converter, on the other hand, makes it possible to run the present drivetrain with little active control, while the MM-CVT concept will not be operable without active control. Due to the decisive importance of the controller, any further analysis of the MM-CVT concept should include investigation of performance, requirements and applicability of different controller concepts. The methods and results of Chapter 5 can in that case be used for benchmarking of such controllers.

References

Asadi, B. and Vahidi, A. (2011). Predictive cruise control: utilizing upcoming traffic signal information for improved fuel economy and reduced trip time. *IEEE Transactions on Control Systems Technology*, 19:707–714.

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.

Bertsekas, D. (2005). *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, 3 edition.

Bryson, A. (1975). *Applied Optimal Control; Optimization, Estimation and Control*. Taylor and Francis.

Bryson, A. (1999). *Dynamic Optimization*. Addison Wesley Longman.

Carl, B., IIvantysynova, M., and Williams, K. (2006). Comparison of operational characteristics in power split continuously variable transmissions. In *Comercial Vehicle Engineering Congress and Exhibition*. SAE.

Dellnitz, M., Junge, O., and Thiere, B. (2001). The numerical detection of connecting orbits. *Discrete and Continuous Dynamical Systems - Series B*, 1:125–135.

Delprat, S., Guerra, T., and Rimaux, J. (2001). Control strategy optimization for a parallel powertrain. In *Proceedings of the 2001 American Control Conference*, pages 1315–1320. IEEE.

Delprat, S., Guerra, T., and Rimaux, J. (2002). Optimal control of a parallel powertrain: from global optimization to real time control strategy. In *IEEE Vehiclular Technology Conference*, pages 2082–2088. IEEE.

Fengyuan, W., Jian, Z., Ruifeng, S., and Fei, Y. (2012). Analysis on the performance of wheel loades in typical work cycle. *Applied Mechanics and Materials*, 148:526–529.

Filla, R. (2005). An event driven operator model for dynamic simulation of contruction machinery. In *Proceedings from the Ninth Scandinavian International Conference on Fluid Power,*. Linköping University.

Filla, R. (2008). Alternative systems solutions for wheel loaders and other construction equipment. In *1st CTI Forum Alternative, Electric and Hybrid Drive Trains*. CTI.

Filla, R. (2011). *Quantifying Operability of Working Machines*. PhD thesis, Linköping University.

Ghabcheloo, R., Hyvönen, M., Uusisalo, J., Karhu, O., Jara, J., and Huhtala, K. (2009). Autonomous motion control of a wheel loader. In *Proceedings of the ASME 2009 Dynamic Systems and Control Conference*, pages 1339–1346. ASME.

Gramattico, S., Balluchi, A., and Cosoli, E. (2010). A series-parallel hybrid electric powertrain for industrial vehicles. In *2010 IEEE Vehicle Power and Propulsion Conference*, pages 1–6. IEEE.

Guzzella, L. and Sciarretta, A. (2007). *Vehicle Propulsion Systems*. Springer Verlag, 2 edition.

Hellström, E. (2010). *Look-ahead Control of Heavy Vehicles*. PhD thesis, Linköping University.

Hellström, E., Åslund, J., and Nielsen, L. (2010). Design of an efficient algorithm for fuel-optimal look-ahead control. *Control Engineering Practice*, 18(11):1318–1327.

Kamien, M. and Schwartz, N. (1991). *Dynamic Optimization; the Calculus of Variations and Optimal Control in Economics and Management*. North-Holland, 2 edition.

Kelley, D. (1998). *Automata and Formal Languages*. Prentice Hall.

Koyachi, N. and Sarata, S. (2009). Unmanned loading operation by autonomous wheel loader. In *ICCAS-SICE 2009*, pages 2221–2225. IEEE.

Lauinger, C., Englisch, A., Gotz, A., Teubert, A., Muller, E., and Baumgartner, A. (2007). Cvt components for powersplit commercial vehicle transmissions. In *Proceedings of the 6th International CTI Symposium*. CTI.

Lennevi, J. (1995). *Hydrostatic Transmission Control, Design Methodology for Vehicular Drivetrain Applications*. PhD thesis, Linköping University.

Liu, S. and Paden, B. (1997). A survey of today's cvt controls. In *Proceedings of the 36th Conference on Decision and Control*, pages 4738–4743. IEEE.

Mattson, P. and Åkerblom, M. (2012). Continuously variable transmission and a working maching including a continuously variable transmission. Patent. WO 2012/008884 A1.

Mitrovic, D. (2005). Reliable method for driving events recognition. *IEEE Transactions on Intelligent Transportation Systems*, 6:198–205.

Nilsson, T., Fröberg, A., and Åslund, J. (2011). Optimized engine transients. In *7th IEEE Vehicle Power and Propulsion Conference*. IEEE.

Nilsson, T., Fröberg, A., and Åslund, J. (2012a). Fuel potential and prediction sensitivity of a power-split cvt in a wheel loader. In *IFAC Workshop on Engine and Powertrain Control, Simulation and Modeling*. IFAC.

Nilsson, T., Fröberg, A., and Åslund, J. (2012b). On the use of stochastic dynamic programming for evaluating a power-split cvt in a wheel loader. In *8th IEEE Vehicle Power and Propulsion Conference*. IEEE.

Nilsson, T., Fröberg, A., and Åslund, J. (2012c). Optimal operation of a turbocharged diesel engine during transients. In *SAE world congress*. SAE.

Paganelli, G., Guerra, T., Delprat, S., Santin, J., Delhom, M., and Combes, E. (2000). Simulation and assessment of power control strategies for a parallel hybrid car. *Proceedings of the Institution of Mechanical Engineers, part D: Journal of Automobile Engineering*, 214:705–717.

Pentland, A. and Andrew, L. (1999). Modeling and prediction of human behavior. *Neural Computation*, 11:229–242.

Pfiffner, R. (2001). *Optimal Operation of CVT-Based Powertrains*. PhD thesis, ETH, Zurich.

Pontryagin, L., Boltyanskii, V., Gamkrelidze, R., and Mishchenko, E. (1962). *The Mathematical Theory of Optimal Processes*. Interscience Publishers.

Rizzoni, G., Guzzella, L., and Baumann, B. (1999). Unified modeling of hybrid electric vehicle drivetrains. *IEEE/ASME Transactions on Mechatronics*, 4:246–257.

Ross, S. (1983). *Introduction to Stochastic Dynamic Programming*. Academic Press.

Rutquist, P., Brietholtz, C., and Wik, T. (2005). An eigenvalue approach to infinite-horizon optimal control. In *Proceedings of the 16th IFAC World Congress*. IFAC.

Rydberg, K.-E. (1998). Hydrostatic drives in heavy mobile machinery - new concepts and development trends. In *International Off-Highway & Powerplant Congress & Exposition*. SAE.

Sasaki, S. (1998). Toyota's newly developed hybrid powertrain. In *Proceedings of the 10th International Symposium on Power Semiconductor Devices and IC's*, pages 17–22. IEEE.

Savaresi, S., Taroni, F., Previdi, F., and Bittanti, S. (2004). Control system design on a power-split cvt for high-power agricultural tractors. *IEEE/ASME Transactions on Mechatronics*, 9(3):569–579.

Sciarretta, A. and Guzzella, L. (2007). Control of hybrid electric vehicles. *Control Systems, IEEE*, 27:60–70.

Srivastava, N. and Haque, I. (2009). A review on belt and chain continuously variable transmissions (cvt): dynamics and control. *Mechanism and Machine Theory*, 44:19–41.

Tanelli, M., Codeca, F., Savaresi, S., Taroni, F., and Previdi, F. (2007). On transmission-ratio computation for the control of a continuously variable transmission in agricultural tractors. In *Proceedings of the 2007 American Control Conference*, pages 5730–5735. IEEE.

Zhang, R. (2002). *Multivariable Robust Control of Nonlinear Systems with Application to an Electro-Hydraulic Powertrain*. PhD thesis, University of Illinois.